

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

DEPARTAMENTO DE ANÁLISIS ECONÓMICO APLICADO



TESIS DOCTORAL

ANÁLISIS Y PREDICCIÓN DE LA DEMANDA DE TRANSPORTE DE PASAJEROS : UNA APLICACIÓN AL ESTUDIO DE DOS CORREDORES DE TRANSPORTE EN GRAN CANARIA

RAQUEL ESPINO ESPINO

Las Palmas de Gran Canaria, 2003

Análisis y predicción de la demanda de transporte de pasajeros

Una aplicación al estudio de dos corredores
de transporte en Gran Canaria

Trabajo realizado por Raquel Espino Espino y dirigido por
Juan de Dios Ortúzar Salas y Concepción Román García

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

Agradecimientos

La tesis doctoral es el fruto del trabajo de horas, días, meses y años y en este camino han participado muchas personas a las que quiero agradecer su colaboración.

Mi más sincera gratitud a mis directores de tesis el Dr. Don Juan de Dios Ortúzar Salas de la Pontificia Universidad Católica de Chile y la Dra. Doña Concepción Román García de la Universidad de Las Palmas de Gran Canaria por haber confiado en mí para realizar esta investigación y guiado en el camino del conocimiento científico. A Concepción Román porque siempre ha estado dispuesta a trabajar codo con codo en los distintos aspectos de este trabajo. A Juan de Dios Ortúzar por haber hecho que los kilómetros que nos separan no fueran un impedimento, atendiendo siempre a mis consultas, animándome a salvar los distintos obstáculos con los que me encontré y aprender de cada uno de ellos. Quiero agradecer también el trato cálido y el cariño mostrado durante este tiempo.

A Elisabetta Cherchi de la Universidad de Cagliari por haber estado siempre dispuesta a discutir conmigo algunas cuestiones importantes. A Jorge Pérez y Matías González, de la Universidad de Las Palmas de Gran Canaria; a Mauricio Sillano, de Steer Davies Gleave; a Kjartan Saelensminde, del Instituto de Economía de Transporte de Oslo; y a Sergio Jara-Díaz de la Universidad de Chile; cada uno de ellos han contribuido a la mejora de esta investigación.

Mi agradecimiento a Nérida Ruíz de Transred, S.A. por haberme facilitado toda la información necesaria para el estudio del Sistema integrado de tarifas, así como a los jefes de explotación de las distintas empresas de transporte: Ramón Hernández de Global, S.A. (Salcai y Utinsa) y Luis Cabrera de Guaguas Municipales, S.A. que me han atendido siempre que lo he requerido. A Oscar Bergasa de la

Universidad de Las Palmas de Gran Canaria y Pablo Suárez Rocha de Grafcan, S.A. por facilitarme una herramienta de medición que ha permitido obtener datos de distancias muy precisos que ha contribuido a mejorar la base de datos. A Moisés Martín de la Universidad de Las Palmas de Gran Canaria que me ayudó a resolver algunas cuestiones del manejo de esta herramienta.

A mis compañeros de departamento por ayudarme en los momentos de flaqueza y animarme a continuar; y al departamento de Ingeniería de Transporte de la Pontificia Universidad Católica de Chile que siempre me ha recibido con los brazos abiertos en mis dos estancias; y especialmente a Ana M^a Ivelic, Claudio Valeze y Luis Rizzi, a los cuales he recurrido en ocasiones para resolver algunas de las cuestiones de este trabajo.

También deseo agradecer el financiamiento recibido gracias al Programa de Mecenazgo de Fundación Universitaria de Las Palmas, siendo el patrocinador Domingo Alonso, S.A, sin cuyo apoyo económico no hubiera sido posible realizar este trabajo.

Finalmente, agradezco a mi familia, amigos y amigas por su comprensión, este proyecto nos ha obligado a estar menos tiempo juntos, y a Matías con el que he compartido los momentos más difíciles.

Índice

Capítulo 1. Introducción	1
Capítulo 2. La modelización de la demanda de transporte	8
2.1 Introducción	8
2.2 Los modelos de demanda desagregados	11
2.2.1 Fundamentos microeconómicos	11
2.2.2 La teoría microeconómica del valor del tiempo	14
2.2.3 La teoría de la utilidad aleatoria	24
2.3 La teoría de la elección discreta	30
2.3.1 Principales modelos de elección discreta	30
2.3.2 Procedimientos de estimación	40
2.4 La predicción de demanda con modelos desagregados	53
2.5 Las fuentes de datos	57
2.5.1 Preferencias reveladas	57
2.5.2 Preferencias declaradas	58
2.5.3 Datos mixtos	63
Capítulo 3. La oferta y la demanda de servicios de transporte en los corredores	72
3.1 Introducción	72
3.2 La oferta de servicios de transporte	74
3.2.1 El transporte regular de pasajeros	74
3.2.2 Sistema tarifario	75
3.2.3 Sistema integrado de transporte	79
3.3 La demanda de servicios de transporte	86
3.4 La encuesta de preferencias reveladas	88
3.4.1 Personas entrevistadas	88
3.4.2 Diseño del cuestionario	89

3.4.3	Análisis descriptivo de los resultados de la encuesta.....	93
3.5	La encuesta de preferencias declaradas.....	96
3.5.1	Las personas entrevistadas.....	96
3.5.2	Diseño del experimento.....	97
3.5.3	Modelización con los resultados de la encuesta piloto	109
3.5.4	Análisis de los resultados de la encuesta.....	113
Capítulo 4.	Modelización de la demanda de transporte	
	de pasajeros y aplicaciones	120
4.1	Introducción	120
4.2	Los datos y las variables.....	121
4.2.1	Los datos	121
4.2.2	Las variables	125
4.3	Estimación de los modelos de elección modal.....	128
4.4	Cálculo de disposiciones a pagar.....	147
4.5	Análisis de la respuesta de la demanda	157
Capítulo 5.	Conclusiones.....	172
	Referencias bibliográficas	178
	Anexos.....	189

Índice de Tablas

Tabla 3.1 Títulos por empresas.....	79
Tabla 3.2 Títulos de viaje del Sistema integrado de tarifas.....	83
Tabla 3.3 Tarifas según títulos	84
Tabla 3.4 Ahorros por nuevas tarifas.....	85
Tabla 3.5 Clasificación por motivo de viaje	94
Tabla 3.6 Clasificación por frecuencia de viaje	95
Tabla 3.7 Definición de niveles	99
Tabla 3.8 Primer diseño	100
Tabla 3.9 Segundo diseño.....	102
Tabla 3.10 Tercer diseño	105
Tabla 3.11 Funciones de utilidad estimadas- encuesta piloto	108
Tabla 3.12 Resultados de las estimaciones de la encuesta piloto.....	110
Tabla 4.1 Modelos con parámetros genéricos.....	141
Tabla 4.2 Modelos con parámetros específicos	146
Tabla 4.3 Valores subjetivos del tiempo de viaje	152
Tabla 4.4 Valores subjetivos del tiempo de caminata.....	153
Tabla 4.5 Valores subjetivos de la frecuencia.....	154
Tabla 4.6 Disposiciones a pagar por el cambio de comodidad baja a estándar.....	155
Tabla 4.7 Disposiciones a pagar por el cambio de comodidad estándar a alta	155
Tabla 4.8 Porcentajes de variación de las políticas individuales	159
Tabla 4.9 Porcentajes de variación de combinaciones de políticas (1).....	169
Tabla 4.10 Porcentajes de variación de combinaciones de políticas (4).....	170

Índice de Figuras

Figura 2.1 Estructura del Logit Jerárquico	33
Figura 2.2 Estructura de árbol para estimación con datos mixtos.	67
Figura 2.3 Representación de parámetros de PR y PD.....	69
Figura 2.4 Representación de parámetros de PR y PD no iguales.....	70
Figura 3.1 Concesiones de las empresas	76
Figura 3.2 Corredores analizados de transporte de viajeros	87
Figura 3.3 Estructura jerárquica para estimación mixta	108
Figura 4.1 Estructura zonal para la medición de tiempos	122
Figura 4.2 Estructura jerárquica para la estimación con datos mixtos	130
Figura 4.3 Estimación con datos mixtos con nido real entre alternativas de PR.....	131
Figura 4.4 Variación de la probabilidad de elegir bus ante un aumento de la frecuencia (50%).....	160
Figura 4.5 Variación de la probabilidad de elegir bus ante un aumento de la frecuencia (100%).....	161
Figura 4.6 Variación de la probabilidad de elegir bus ante una reducción de la tarifa (Tarjeta Valor).....	162
Figura 4.7 Variación de la probabilidad de elegir bus ante una reducción de la tarifa (Tarjeta Cliente)	163
Figura 4.8 Variación de la probabilidad de elegir bus ante una reducción del tiempo de viaje (10%).....	165
Figura 4.9 Variación de la probabilidad de elegir bus ante un incremento del coste de aparcamiento	167

1. Introducción

El transporte de pasajeros en áreas urbanas constituye un problema cada vez más importante en nuestra sociedad. En los últimos años se ha demostrado empíricamente que la construcción de infraestructura o la ampliación de las vías existentes no es una solución adecuada a este problema, sino que se debe planificar el transporte de pasajeros otorgando incentivos al uso del transporte público, para intentar paliar los efectos concomitantes que produce la operación del sistema de transporte: congestión, accidentes, contaminación, etc.

Los desarrollos en el análisis de la demanda de transporte son importantes, incorporando en ocasiones metodologías que han sido desarrolladas en otras áreas del conocimiento científico. Sin embargo, no siempre se da la importancia necesaria al estudio riguroso de los aspectos que influyen de alguna manera en el comportamiento de viajes. Nos encontramos con un sistema compuesto de múltiples y variados elementos que requieren de un ejercicio de planificación importante para no malgastar recursos públicos, así como evitar los problemas derivados de una oferta de transporte que realmente no provea el servicio demandado. Estamos hablando de intentar considerar todos los elementos que interactúan en el desarrollo de las ciudades y la necesidad de, por ejemplo, planificar paradas y nuevas líneas de buses ante una ampliación territorial de la ciudad. También se requiere desarrollar la capacidad de

valorar todos los costes ambientales que cada modo de transporte genera para que la toma de decisiones de política sea lo más correcta posible.

Por otra parte, en áreas densamente pobladas como es el caso de Gran Canaria con una densidad¹ de población de 494 habitantes por metro cuadrado, la escasez del territorio es un problema añadido a los del sistema de transporte. Además, el crecimiento² del parque automovilístico ha sido aproximadamente del 6% desde el año 1997 al 2001. Este importante crecimiento ha derivado en un aumento de los problemas de tráfico en la isla y concretamente en los accesos a su capital. Por esto, el objetivo principal de esta tesis se centra en estudiar la demanda de transporte de pasajeros en los corredores Norte (Arucas-Las Palmas de Gran Canaria) y Sur (Telde-Las Palmas de Gran Canaria) de acceso a la capital, con el fin de intentar analizar la respuesta de la demanda de transporte de pasajeros ante una reducción en los precios del transporte público debida a una integración de tarifas entre las empresas que prestan servicio.

Las últimas décadas han sido testigo del importante desarrollo experimentado por las técnicas de análisis y predicción de demanda en el área de transporte. Desde principios de los años 70 (veáse Domencich y McFadden, 1975; Daganzo, 1979) donde comenzaron a sentarse las bases metodológicas de los modelos de elección discreta hasta la actualidad, son muchos los avances que se han producido, especialmente en los últimos cinco años gracias al desarrollo de nuevas técnicas de estimación basados en la simulación (Train, 2002).

¹ Fuente: Instituto Canario de Estadística, ISTAC, año 2002. No incluye a la población no residente.

² Dato obtenido a partir de los datos de matriculación a 31 de diciembre publicados por el ISTAC.

A los modelos de demanda desagregados, a pesar de constituir una metodología totalmente aceptada por la comunidad científica y de ser considerados un instrumento clave para la política de transporte en muchos países, en España no se les ha concedido, a nuestro juicio, la importancia y el apoyo que merecen. Es por esta razón que no existen muchos trabajos ni muchos especialistas en esta disciplina. Esta tesis trata de contribuir en este campo del conocimiento siguiendo la línea planteada en trabajos previos (veáse González, 1995), e incorporando algunos de los avances producidos desde entonces. Desde un punto de vista metodológico, abordamos esta investigación tratando de separarnos de las hipótesis de los modelos clásicos como el logit multinomial y de las especificaciones de la función de utilidad basadas en aproximaciones polinómicas de primer orden. Por otra parte, también tratamos de analizar el efecto que tienen sobre las decisiones de los individuos otras variables diferentes de los atributos típicos de nivel de servicio como el tiempo de viaje, el coste y la frecuencia. Gran parte de estos objetivos se logran gracias al empleo de técnicas de preferencias declaradas (PD) y a la combinación de estos datos con las preferencias reveladas (PR).

Los modelos de demanda desagregados encuentran su fundamento teórico en la microeconomía de las elecciones discretas (McFadden, 1981) y en la teoría de la utilidad aleatoria (McFadden, 1974; Domencich y McFadden, 1975). Esta última representa la herramienta estadística que permite abordar de forma empírica el problema de modelización de la demanda en un contexto de elecciones discretas. De esta forma, la utilidad del consumidor está representada por una variable aleatoria que se incorpora aditivamente en una parte observable y otra no observable de naturaleza estocástica. Las distintas hipótesis acerca de la distribución del término estocástico darán lugar a los distintos modelos de elección discreta.

La información relativa a la parte observable se obtiene a partir de datos PR y PD. Los datos de PR reflejan el comportamiento actual de los individuos en sus decisiones de viaje, esto es, se trata de preferencias efectivas y han sido los más utilizados en la modelización de la demanda de transporte hasta mitad de los años 80. Por su parte, los datos de PD representan preferencias eventuales y tratan de reflejar lo que los individuos harían ante determinadas situaciones hipotéticas. Los datos de PR presentan algunas limitaciones que pueden resolverse, parcialmente, mediante el diseño de experimentos de PD (Ortúzar y Willumsen, 2001; Louviere *et al*, 2000).

Desde un punto de vista econométrico, la diferencia entre ambos tipos de datos está en que presentan errores de distinta naturaleza. Los datos de PR presentan errores de medición de las variables independientes mientras que los datos de PD en la variable dependiente, esto es, en la elección. La estimación conjunta con datos de PR y PD se denomina estimación con datos mixtos. Este procedimiento fue desarrollado por Ben Akiva y Morikawa (1990) y permite aprovechar las ventajas de ambos tipos de datos ofreciendo mejores resultados.

En este trabajo se plantea el estudio de la demanda de transporte de pasajeros en dos corredores partiendo, por un lado, de una integración tarifaria que permitiría reducir la tarifa del bus y por otro, de la hipótesis que hay más elementos, además del coste del viaje, que son tenidos en cuenta en la elección del modo de transporte. Bajo este planteamiento realizamos dos tipos de encuestas: una de PR con el objeto de conocer las elecciones reales de los individuos y otra de PD para plantear un transporte público (bus) más atractivo, de manera que en el diseño de PD se incorporaran, además de la reducción de tarifas, otros elementos que se consideraron relevantes en la elección modal basándonos en el grupo focal realizado para la mejor comprensión del problema

que estábamos estudiando. Así pues, parte de la investigación se dedicó al diseño de un experimento de PD que nos permitiera medir estos elementos, incluyendo una variable latente como la comodidad, que no es fácilmente medible o no presenta suficiente variabilidad. Para la obtención de la encuesta definitiva se realizaron tres pretest, que exigieron realizar fuertes modificaciones en el diseño hasta conseguir un experimento que se adaptara al objeto del estudio y que permitiera definir de forma adecuada la percepción de los distintos niveles de la variable comodidad. Además, el diseño creado permitió analizar el efecto de interacciones entre el tiempo, el coste y la frecuencia, esto es, el efecto que produce sobre la utilidad la variación conjunta de pares de estas variables.

Respecto a la modelización de la demanda, se estudiaron distintas especificaciones para la utilidad. Se investigó una especificación no lineal, ya que existen numerosos estudios (Hensher, 1998; Jara-Díaz y Videla, 1989; Ortúzar *et al*, 2000) que demuestran que una especificación lineal da lugar a estimaciones incorrectas para explicar el comportamiento del individuo y por tanto del valor subjetivo del tiempo y otras disposiciones a pagar. En concreto, la utilidad definida en nuestra modelización incorpora la renta dividiendo los costes de viaje y de aparcamiento por la tasa de gasto; también incluye la interacción entre el coste y la frecuencia y la especificación de variables socioeconómicas definidas en las variables modales, a través de un parámetro base para la variable modal más otros para las variables socioeconómicas que interactúan con ésta. La especificación que se obtuvo finalmente nos permite explicar el comportamiento del individuo de una manera bastante rica, obteniéndose valores de las disposiciones a pagar que varían entre individuos y que son agregadas posteriormente de acuerdo a las categorías definidas por las características socioeconómicas.

Este documento se divide en cinco capítulos. El capítulo 2 se centra en el análisis de los aspectos más relevantes relativos a la modelización de la demanda de transporte basada en el empleo de modelos de elección discreta. En su segunda sección se estudian los principales modelos existentes, analizando tanto los fundamentos microeconómicos como los econométricos. En la tercera sección se estudian los procedimientos de estimación que se usan habitualmente para estimar este tipo de modelos. En la sección cuarta se aborda el problema de predicción agregada, esto es, cómo predecir el comportamiento de grupos o clases a partir de estimaciones procedentes de modelos individuales. Por último, en la sección quinta se analizan las diferentes fuentes de datos empleadas por estos modelos, haciendo especial referencia a la modelización con PD.

El capítulo 3 está dedicado a la caracterización del mercado objetivo, distinguiendo entre la oferta y la demanda de transporte. En la segunda sección se describe de manera general la oferta de servicios de transporte planteada desde los aspectos del servicio ofrecido por las distintas empresas, el sistema tarifario y el sistema integrado de tarifas que tiene implicaciones sobre el coste del transporte público, y que será tenido en cuenta en nuestro diseño de PD. En la tercera sección, se caracteriza la demanda de servicios de transporte antes de la realización de nuestra investigación, mientras que en las secciones cuarta y quinta se hace referencia a la información utilizada en el estudio; concretamente se explican los dos tipos de encuestas realizadas, la encuesta de PR y la de PD, incidiendo en el caso de esta última en el diseño del experimento.

En el capítulo 4 se presenta la modelización realizada para los corredores estudiados. En la primera sección se describen las bases de datos utilizadas, así como la definición de las distintas variables consideradas en la modelización. En la segunda sección se presenta la modelización realizada utilizando la metodología de estimación con datos

mixtos. Se probaron diferentes especificaciones de la función de utilidad con el objeto de encontrar un modelo que explicase adecuadamente el comportamiento de los individuos entrevistados. Así, se estudió la posible existencia de correlación entre las alternativas de transporte privado, esto es, coche conductor y coche acompañante; la existencia o no de efecto renta siguiendo el planteamiento desarrollado por Jara-Díaz y Videla (1989), y la hipótesis de linealidad de la función de utilidad. Una vez modelizado el comportamiento de viaje, se calculan las disposiciones a pagar, en la tercera sección. Finalmente, en la última se estudia el efecto de la respuesta de la demanda de transporte de pasajeros ante distintas medidas de política.

Para finalizar, en el capítulo 5 se presentan las principales conclusiones de este trabajo y aquellos aspectos que consideramos deben ser objeto de futuras investigaciones.

2. La modelización de la demanda de transporte

2.1 Introducción

El objetivo de un sistema de transporte es satisfacer la demanda mediante la provisión de una oferta adecuada que responda a sus exigencias. Los servicios de transporte surgen como consecuencia de la necesidad que tienen los individuos de realizar actividades que implican desplazamientos; por ejemplo, ir al trabajo, llevar los niños al colegio, ir al teatro, al médico, etc. Por esta razón se dice que la demanda de transporte es derivada. En general, no se demanda viajar *per se* sino con el objeto de realizar alguna actividad localizada en el espacio y en el tiempo. Estamos, por tanto, ante una demanda cualitativa y diferenciada en el sentido de que existen viajes por múltiples motivos, a distintas horas del día y que pueden realizarse en distintos modos de transporte. En lo que respecta a la oferta, nos encontramos con un servicio que no puede ser almacenado para ser ofertado, por ejemplo, en períodos donde existe una mayor demanda. Se produce, por tanto, un desequilibrio entre la cantidad ofertada y la demandada debido a la existencia de períodos con mayor nivel de demanda (hora punta) y períodos con menores niveles de demanda (hora valle), siendo imposible el “trasvase” de la capacidad ociosa de un período a otro. Además, la demanda se encuentra

localizada en un espacio, que se va ampliando con cierta celeridad, lo que suele producir problemas de coordinación que afectan al equilibrio del sistema.

Para satisfacer de forma eficiente la demanda de transporte es necesaria la planificación en el corto y largo plazo con un conocimiento bueno y preciso de los distintos atributos que la determinan. Existen varios argumentos que justifican esta planificación. La provisión de infraestructura de transporte es costosa dado su alto grado de indivisibilidad, y requiere, en muchos casos, de largos periodos desde que se planifica su construcción hasta que ésta se ve totalmente finalizada. En segundo lugar, existe una segregación entre el proveedor, generalmente público, de la infraestructura y el operador de la misma, lo que implica la necesidad de una regulación que defina los estándares del servicio de transporte que se desea ofrecer. Por último, la operación del servicio de transporte está asociada a determinadas externalidades (accidentes, contaminación, etc) que distorsionan las decisiones a tomar, ya que los usuarios del sistema no internalizan el coste que su elección de viaje genera.

Los argumentos expuestos justifican la importancia de realizar estudios de demanda que permitan evaluar los proyectos de transporte de la manera más completa posible, en el sentido de que sea factible analizar de forma conjunta los aspectos de la demanda y de la oferta con el fin de conocer el efecto de futuras políticas de transporte que puedan introducirse (por ejemplo, variación de algunos atributos de alternativas ya existentes o la introducción de nuevos modos de transporte).

La modelización de la demanda de transporte se ha basado tradicionalmente en el empleo de dos tipos de modelos: los modelos agregados o de *primera generación* y los modelos desagregados o de *segunda generación*. Los primeros fueron utilizados mayoritariamente en los estudios de transporte hasta finales de los años 70. Estos

modelos emplean datos agregados, por ejemplo, promedio de las variables a nivel zonal para las distintas zonas en las que se divide el área de estudio. La agregación de los datos presenta algunos inconvenientes. En primer lugar, es necesario disponer de una cantidad de datos importante ya que cada observación resulta de la obtención de un promedio de un cierto número de datos individuales. En segundo lugar, la existencia de sesgo entre las unidades agregadas es más probable que en los modelos desagregados. La conducta individual puede ser ocultada por características no identificadas asociadas a la zona, lo que se conoce como *falacia ecológica*³.

Por su parte, los modelos desagregados comienzan a ser populares en los años 80 por el hecho de que presentan ciertas ventajas sobre los modelos utilizados hasta el momento. Este tipo de modelos, se basa en el uso de los datos a nivel individual, lo que permite una mejor comprensión de los comportamientos de viaje, puesto que se basan en las teorías de la elección individual⁴. Las ventajas que presenta este tipo de modelos respecto a los de *primera generación* son varias. En primer lugar, el hecho de utilizar cada dato individual como una observación hace que sea más eficiente en el uso de la información en que los modelos agregados. En segundo lugar, ocupan toda la variabilidad de la que se dispone con los datos individuales. En tercer lugar, los resultados de los modelos a nivel individual pueden ser utilizados a cualquier nivel de agregación⁵. En cuarto lugar, la estimación de los parámetros para cada una de las variables explicativas consideradas en el modelo es explícita, lo que deriva en cierta

³ Un ejemplo para explicar la *falacia ecológica* es el caso de un modelo de frecuencia de viajes por zonas a través del cual se obtiene que el número de viajes por hogar decrece con el ingreso. Al observar los datos individuales se demuestra lo contrario que es lo que cabría esperar. Para un mayor detalle ver Ortúzar y Willumsen, 2001, pp. 221.

⁴ Concretamente, vamos a estudiar la Teoría de la Utilidad Aleatoria en el epígrafe 2.2.3 y a partir de la misma se derivan distintos modelos desagregados o de elección discreta.

⁵ Para la predicción con este tipo de modelos es necesaria la agregación que no es trivial en muchos casos y que será estudiada en el epígrafe 2.4.

flexibilidad para representar las variables de política; además, los parámetros estimados tienen una interpretación directa en términos de la importancia relativa de cada variable explicativa considerada en la elección. Y finalmente, se trata de modelos más estables en el tiempo y en el espacio.

Este capítulo se centra en el análisis de los aspectos más relevantes relativos a la modelización de la demanda de transporte basada en el empleo de modelos desagregados. En la segunda sección se estudian los principales modelos analizando tanto los fundamentos microeconómicos como los econométricos. En la tercera sección se estudian los procedimientos de estimación que se usan habitualmente para estimar este tipo de modelos. En la sección cuarta se aborda el problema de la predicción agregada a partir de estimaciones procedentes de modelos desagregados. Por último, en la sección quinta se analizan las diferentes fuentes de datos empleadas por estos modelos, haciendo especial referencia a la modelización con preferencias declaradas.

2.2 Los modelos de demanda desagregados

2.2.1 Fundamentos microeconómicos

La teoría económica clásica plantea un modelo de comportamiento del consumidor donde las preferencias que definen dicho comportamiento se expresan a través de una función de utilidad que depende de la cantidad consumida de distintos bienes de naturaleza divisible. Este modelo está caracterizado por la maximización de su nivel de satisfacción sujeto a una restricción presupuestaria, que indica que el gasto total en bienes no debe exceder a la renta disponible por el individuo. Postulados posteriores (Lancaster, 1966) que modifican esta teoría, establecen que una fuente de utilidad proviene de las características de los bienes (sabor, color, etc) y no de los bienes *per se*, de modo que unas características determinadas se obtendrían a partir de la compra de un

conjunto de bienes de mercado. La incorporación del consumo de bienes de naturaleza discreta (ver la formalización en McFadden, 1981) permite reformular el problema clásico de comportamiento del consumidor. Siguiendo a Jara-Díaz (1998), la función de utilidad a maximizar va a depender de las cantidades consumidas de bienes de naturaleza continua, representadas en el vector (X) y por la elección entre un conjunto de alternativas discretas representadas por un vector Q_j que recoge sus principales características o atributos. El problema de decisión del consumidor será:

$$\begin{aligned}
 & \underset{x,j}{\text{Max}} U(X, Q_j) \\
 & \text{s.a.} \\
 & \sum P_i X_i + c_j \leq I \\
 & X_i \geq 0 \quad j \in M
 \end{aligned} \tag{2.1}$$

siendo:

P_i , el precio del bien i

X_i , cantidad consumida del bien i

c_j , el coste de la alternativa discreta j

I , la renta del individuo

M , el conjunto de alternativas disponibles

La resolución de las condiciones de primer orden del problema (2.1) para cada j permite obtener las demandas condicionadas $X_j(P, I - c_j, Q_j)$ a la elección discreta j . La sustitución de estas funciones en la utilidad da lugar a la función de utilidad indirecta

condicional (FUIC) $V_j = V_j(P, I - c_j, Q_j)$ ⁶. Esta función representa la máxima utilidad que puede obtener el individuo una vez elegida la alternativa j . Al maximizar en j , el individuo elegirá aquella alternativa que le proporcione máxima utilidad. Si se define $V^* = \text{Max}_{j \in J} V_j(P, I - c_j, Q_j)$, la aplicación de la identidad de Roy permite obtener la demanda de las alternativas discretas:

$$\frac{\frac{\partial V^*}{\partial c_i}}{\frac{\partial V^*}{\partial I}} = \mathbf{d}_i = \begin{cases} 1 & \text{si } V_i \geq V_j \quad \forall j \neq i \\ 0 & \text{en otro caso} \end{cases} \quad (2.2)$$

que en este caso es una variable discreta.

A partir de V_j podemos obtener la utilidad marginal de la renta y el valor subjetivo de la característica k :

$$\text{Utilidad marginal de la renta: } UMR = \frac{\partial V_j}{\partial I} = - \frac{\partial V_j}{\partial c_j} \quad (2.3)$$

$$\text{Valor subjetivo de la característica } k: SV_k = - \frac{\partial V_j / \partial q_{kj}}{\partial V_j / \partial I} \quad (2.4)$$

Para obtener estas magnitudes es preciso especificar la forma funcional de V_j y estimarla empíricamente. En la modelización basta considerar la parte relevante de V_j en la toma de decisiones del individuo. Ésta se denomina función de utilidad indirecta condicional truncada⁷. En general, cualquier función de utilidad indirecta que cumpla las

⁶ Esta función es continua en $(I - c_j, p, Q_j)$, dos veces diferenciable y homogénea de grado cero en $(I - c_j, p)$, estrictamente cuasiconvexa en p , y además se cumple que $\frac{\partial V_j}{\partial (I - c_j)} > 0$, McFadden (1981).

⁷ Se obtiene al comparar V_i con V_j .

propiedades matemáticas adecuadas⁸ puede ser aproximada a través de su desarrollo de Taylor por una especificación lineal en los parámetros (McFadden, 1981). Muchos de los modelos aplicados en la literatura consideran la aproximación local de primer orden que da lugar a una especificación lineal en los parámetros y en las variables. Bajo esta especificación, la UMR coincide con menos el parámetro del coste de la alternativa y el valor subjetivo de la característica k se obtiene a partir del cociente entre el parámetro de dicha característica y el del coste. Un inconveniente de definir funciones de utilidad lineales es que la renta no juega ningún papel en las decisiones de los individuos, es decir, las elecciones son independientes de la renta. Jara-Díaz y Videla (1989) demuestran que basta considerar una expansión de Taylor de orden igual o superior a dos (términos de orden dos para el coste) para poder considerar la renta en la elección de la alternativa.

2.2.2 La teoría microeconómica del valor del tiempo

Entre los modelos alternativos al modelo clásico de comportamiento del consumidor están aquellos que reconocen que el tiempo influye en las decisiones y en las restricciones a las que se enfrentan los consumidores. La valoración económica de los ahorros de tiempo dedicado al consumo de ciertos bienes o servicios constituye una pieza clave en la evaluación de políticas que tienen como objetivo la reducción de esos tiempos de consumo. En el caso del transporte la obtención de una cuantificación económica de estos ahorros de tiempo cobra una especial relevancia.

⁸ U es continua en (Y, Q_j) para cada j , dos veces diferenciable con continuidad en Y con $\frac{\partial U}{\partial Y} > 0$ y estrictamente cuasiconcava para cada Q_j y cada j .

En este epígrafe se presenta una revisión de los principales modelos que incorporan el tiempo al modelo de comportamiento del consumidor⁹. Podemos distinguir fundamentalmente dos tipos de modelos. En primer lugar, los modelos clásicos que se sitúan en un contexto general donde se asume que los bienes de consumo son de tipo continuo o divisibles (Becker, 1965; DeSerpa, 1971; 1973). En segundo lugar, los modelos desarrollados en el ámbito del transporte que se sitúan en el contexto microeconómico de las elecciones discretas (elección modal, generalmente) y por tanto consideran la posibilidad de consumir bienes de tipo discreto: (Train y McFadden, 1978; Jara-Díaz y Farah, 1987; Bates y Roberts, 1986).

La primera consideración del tiempo en el modelo de comportamiento del consumidor fue desarrollada por Gary Becker en 1965. La idea fundamental es que el consumo de los bienes requiere tiempo (tiempo para comprar, para prepararlos, para consumirlos, etc) y bajo esta hipótesis es necesario incluir una restricción de tiempo, además de la ya conocida restricción de renta.

El tiempo es considerado como un recurso económico no almacenable del cual todos los individuos están dotados en la misma cantidad. Cada individuo dispone del mismo número de unidades de tiempo por período (24 horas al día, 168 horas a la semana, etc). Los consumidores distribuyen su tiempo en trabajar y realizar otras actividades¹⁰, entendiéndose que la realización de una actividad determinada lleva implícito el consumo de varios bienes de mercado y de tiempo. Tenemos entonces, que el tiempo dedicado a trabajar (T_w) más el tiempo que dedicamos a realizar otras actividades (t_i), debe ser igual

⁹ Una revisión más detallada puede verse en González (1997).

¹⁰ Becker se refiere a estas actividades con el término mercancías básicas (*basic commodities*).

al tiempo total disponible (T). Por tanto, la restricción de tiempo según el modelo de Becker es:

$$\sum_{i=1}^{i=n} t_i q_i + T_w = T \quad (2.5)$$

siendo q_i la cantidad consumida de la mercancía i .

Por su parte, la restricción de renta está compuesta por un lado, por la renta salarial y la renta no salarial (H). La renta no salarial (pensiones, dividendos, pagos de la Seguridad Social, etc) no requiere del consumo de tiempo para su obtención. La renta salarial dependerá del número de horas trabajadas. En este modelo el tiempo de trabajo es endógeno, esto es, el individuo decide libremente el número de horas que dedica a trabajar, lo que significa que puede influir sobre su renta salarial. Por otro lado, tenemos que el individuo dedica su renta a la obtención de aquellos bienes de mercado que le permitan realizar aquellas actividades que desee. Por tanto, la restricción presupuestaria tiene la expresión:

$$\sum_{i=1}^{i=n} P_i q_i = w T_w + H \quad (2.6)$$

siendo

P_i , el precio de la mercancía i ¹¹

w , el salario por hora de trabajo

¹¹ En realidad no existe un mercado donde realizar transacciones con estas mercancías, pero si suponemos que para adquirir la mercancía q_i necesitamos $b_i q_i$ unidades del bien de mercado x_i al precio p_i (es decir $x_i = b_i q_i$) entonces el precio de la mercancía P_i es equivalente a $b_i p_i$.

Al considerar el tiempo que se dedica al consumo, el precio total de la mercancía i no va a ser el precio de mercado, sino éste más el coste de oportunidad del tiempo dedicado a dicho consumo, esto es, el ingreso al que se renuncia debido a que este tiempo se podría dedicar a trabajar. Este precio total del bien se obtiene de despejar las horas de trabajo de la ecuación (2.6) y reemplazar en la ecuación (2.5), de lo que resulta la siguiente restricción que incluye tanto la restricción de tiempo como de renta:

$$\sum_{i=1}^{i=n} (P_i + wt_i) q_i = wT + H \quad (2.7)$$

siendo el precio total de la mercancía i : $P_i + wt_i = P_i^T$ (2.8)

Existe, por tanto, un *compromiso* entre precio y tiempo; precios más altos requieren un tiempo de consumo menor y viceversa. Se pueden describir las elecciones de precio-tiempo a través de una función decreciente que expresa el tiempo de consumo en función del precio de mercado, de manera que el precio total va depender, en definitiva, del precio de mercado. El consumidor minimiza el precio total, decide de forma equilibrada entre el precio de mercado y el coste de oportunidad del tiempo que dedica al consumo de cada bien. La elección de los consumidores está sujeta a su renta salarial; esto es, consumidores de mayores rentas (mayor tasa salarial) estarán dispuestos a pagar mayores precios de mercado para ahorrar en el tiempo que dedican al consumo de los bienes que los consumidores con tasas salariales más bajas. De la condición de primer orden del problema de minimización del precio total se obtiene que el valor del tiempo en el modelo de Becker coincide con la tasa salarial.

El modelo de DeSerpa (1971; 1973) incluye el tiempo dedicado a la realización de actividades y el tiempo de trabajo en la función de utilidad. Además de las restricciones

de renta y tiempo, ya definidas en otros modelos, se incluyen las *restricciones tecnológicas* que expresan la necesidad de un tiempo mínimo para realizar cada actividad, distinguiendo la actividad laboral del resto de actividades. Se define un número mínimo de horas de trabajo con el fin de relajar la hipótesis poco realista acerca de la absoluta flexibilidad que dispone el consumidor para decidir el número de horas que trabaja. De este modo, la formulación del modelo es la siguiente:

$$\begin{aligned}
 & \text{Max}U(X, t, t_w) \\
 & \text{s.a.} \\
 & pX \leq wt_w + H \quad (\mathbf{l}) \\
 & \sum_{j=1}^{j=n} t_j + t_w \leq T \quad (\mathbf{m}) \\
 & t_w \geq t_w^* \quad (\mathbf{f}) \\
 & t_j \geq t_j^* \quad (\mathbf{y}_j)
 \end{aligned} \tag{2.9}$$

siendo:

w , salario por hora

t_w , horas de trabajo

H , renta no salarial

T , tiempo total disponible

t_j , el tiempo que dedica a la actividad j

t_w^* , tiempo mínimo de trabajo

t_j^* , tiempo mínimo para la actividad j

A cada restricción se le asocia un multiplicador de *Lagrange*, de manera que I es la utilidad marginal de la renta, m la utilidad marginal del tiempo como recurso, f la utilidad marginal por disminuir el número mínimo de horas de trabajo y Y_j la utilidad marginal por disminuir el número mínimo de horas dedicadas a la actividad j .

Lo relevante en este modelo es la distinción entre actividades de ocio (*pure leisure activities*) y actividades intermedias (*intermediate activities*). Las actividades de ocio son aquellas a las que el individuo dedica más tiempo del mínimo exigido (la restricción no se cumple con igualdad, por tanto $Y_i=0$). Las actividades intermedias son aquellas en las cuales se dedica el tiempo mínimo necesario. En estas actividades nuestra utilidad se ve incrementada cuando se reduce el mínimo tiempo que hay que dedicarles, es decir, $Y_j>0$; por tanto la condición de holgura complementaria asegura que la restricción se cumple con igualdad $t_j = t_j^*$.

La valoración marginal del tiempo dedicado a la actividad j es el cociente entre la utilidad marginal del tiempo dedicado a la actividad j y la utilidad marginal de la renta, que es igual a la suma de cuatro elementos: la tasa salarial, la desutilidad/utilidad del trabajo, la disposición a pagar por disminuir las horas de trabajo y la disposición a pagar por ahorrar tiempo en la actividad j que puede ser transferido a actividades de ocio, esto es¹²:

$$\frac{1}{I} \frac{\partial U}{\partial t_j} = w + \frac{1}{I} \frac{\partial U}{\partial t_w} + \frac{f}{I} - \frac{Y_j}{I} \quad (2.10)$$

¹² Esta expresión se puede obtener a partir de las condiciones de primer orden del problema (2.9)

En el modelo de Becker, los tres últimos términos de la expresión (2.10) son cero porque no hay desutilidad asociada a dedicar tiempo al trabajo, no hay un mínimo de horas de trabajo y no hay restricciones de tiempo mínimo para las distintas actividades que puede realizar el individuo. En este caso, el valor marginal del tiempo dedicado a la actividad j coincide con la tasa salarial, como hemos comentado anteriormente.

De las condiciones de primer orden, también se puede obtener que la valoración marginal del tiempo en la actividad j es igual al valor del tiempo como recurso (m/I) menos el valor de ahorrar tiempo en la actividad j (y_j/I), es decir:

$$\frac{1}{I} \frac{\partial U}{\partial t_j} = \frac{m}{I} - \frac{y_j}{I} \quad (2.11)$$

En las actividades de ocio ($y_j = 0$), la valoración marginal del tiempo dedicado a la actividad j coincide con el valor del tiempo como recurso, mientras que para las actividades que son intermedias ($y_j \neq 0$), ésta va a ser menor que el valor del tiempo como recurso en la cuantía $\frac{y_j}{I}$. Esta magnitud representa el valor del ahorro de una unidad de tiempo en la actividad j , es decir:

$$\frac{y_j}{I} = \frac{-\partial U^* / \partial t_j^*}{\partial U^* / \partial I} = \frac{dI}{dt_j^*} \quad (2.12)$$

donde I representa la renta del individuo. En otras palabras, (2.12) representa lo que el individuo está dispuesto a sacrificar de su renta por ahorrar una unidad de tiempo en la actividad intermedia j y poder disfrutar de una unidad adicional de tiempo de ocio. Éste

es el concepto al que convencionalmente se hace referencia como *valor subjetivo del tiempo* en la terminología de evaluación de proyectos de inversión en transporte.

Train y McFadden (1978) plantean un modelo de intercambio entre bienes y ocio donde tratan de analizar cuál es el papel del salario en las decisiones de los consumidores cuando se enfrentan a un problema de elección de modo de transporte y cómo éste varía al considerar distintas formas funcionales en la función utilidad que representa las preferencias individuales.

La función objetivo a la que se enfrenta el consumidor depende del consumo de bienes (G) y del tiempo de ocio disponible (L) sujeto a dos restricciones; la primera establece que el gasto total en bienes consumidos (incluido el gasto en transporte) debe ser igual al total de renta disponible y la segunda, que el tiempo de ocio debe ser igual al total de tiempo disponible una vez descontado el tiempo de dedicado al viaje y al trabajo. El problema de optimización del consumidor es el siguiente:

$$\begin{aligned}
 & \text{Max}U(G_j, L_j) \\
 & \text{s.a.} \\
 & G_j = H + wW - c_j \\
 & L_j = T - W - t_j
 \end{aligned} \tag{2.13}$$

donde t_j y c_j representan el tiempo y coste de viaje, respectivamente, del modo de transporte j . Al sustituir G_j y L_j se obtiene que la función de utilidad depende de las horas de trabajo (W) y el problema se transforma en:

$$\text{Max}_j \left\{ \text{Max}_w U [G_j(W, c_j), L_j(W, t_j)] \right\} \tag{2.14}$$

Al maximizar en W y sustituir en la función objetivo se obtiene la FUIC:

$$U_j = U [G_j(W^*(c_j, t_j), c_j), L_j(W^*(c_j, t_j), t_j)] \quad (2.15)$$

En este modelo, la renta del individuo se determina de forma endógena ya que él elige el número de horas que dedica a trabajar a un salario dado. La forma en que se introduce la variable renta en la función de utilidad va a estar definida por el mapa de curvas de indiferencia entre bienes y ocio del consumidor. Considerando una función de utilidad tipo Cobb-Douglas se obtiene que la tasa salarial puede especificarse ya sea multiplicando al tiempo o bien dividiendo al coste. En el primer caso, se supone implícitamente que las personas con mayor nivel de renta (tasas salariales más altas) perciben más las pérdidas de tiempo en su viaje al trabajo que las personas con menor nivel de renta (tasas salariales más bajas); mientras que en el segundo caso, implícitamente se supone que el trabajador con alto nivel de renta es menos consciente del coste del viaje al trabajo que las personas con bajo nivel de renta. Una consecuencia que se deriva de esta formulación es que tanto la utilidad marginal de la renta como el valor subjetivo del tiempo de viaje son valores constantes, coincidiendo este último con la tasa salarial.

Jara-Díaz y Farah (1987) reformulan el modelo planteado por Train y McFadden (1978) argumentando que, por un lado, la renta no debe determinarse de forma endógena sino exógena debido a que muchos individuos no pueden decidir fácilmente el número de horas que trabajan ni obtener un salario adicional por trabajar más horas de las establecidas en su jornada laboral; y por otro lado, que la proporción gastada en transporte, especialmente para viajes recurrentes, no es despreciable.

Partiendo del modelo de Train y McFadden (1978) obtienen un valor del tiempo que depende directamente de lo que denominan *tasa de gasto* (g), que no es más que la renta

por unidad de tiempo disponible. Si el número de horas de trabajo es fijo, lo importante es el tiempo disponible para gastar nuestra renta y en este caso, el efecto renta es relevante.

Bates y Roberts (1986) proponen una adaptación del modelo de DeSerpa al caso de la elección de alternativa de transporte, diferenciando explícitamente el tiempo dedicado al transporte del dedicado a realizar otras actividades en su formulación del problema del consumidor.

Así definen una función de utilidad que depende de la cantidad consumida de un bien generalizado (X), del tiempo dedicado a realizar una actividad generalizada (q) y del tiempo empleado en cada una de las alternativas de transporte disponibles (t_j " $j=1,2,\dots,n$), siendo las elecciones entre alternativas mutuamente excluyentes, como por ejemplo la elección de ruta o de modo de transporte para un viaje específico. Para este modelo, el problema de optimización será:

$$\begin{aligned}
 & \text{Max}U(X, q, t_1, \dots, t_n) \\
 & \text{s.a.} \\
 & pX + \sum d_j c_j \leq y \quad (\mathbf{l}) \\
 & q + \sum d_j t_j \leq T \quad (\mathbf{m}) \\
 & t_j \geq t_j^* \quad (\mathbf{y}_j)
 \end{aligned} \tag{2.16}$$

Las restricciones son las ya conocidas, de renta y tiempo, añadiendo respectivamente a cada una de ellas el coste (c_j) y el tiempo (t_j) de las distintas alternativas. Se define una variable ficticia (d_j) que toma valor 1 si la alternativa j es elegida y 0, en otro caso. Y finalmente, una tercera restricción que define un tiempo mínimo para cada alternativa.

Sustituyendo las condiciones de primer orden del problema (2.16) en la aproximación local de primer orden de la función de utilidad, obtenemos una especificación lineal de la FUIC de la alternativa j , de manera que si solo tenemos en cuenta las variables que dependen de la alternativa, la FUIC¹³ toma la siguiente expresión:

$$V_j = -Ic_j - \mathbf{y}_j t_j \quad (2.17)$$

La disposición a pagar por ahorrar tiempo en la alternativa j vendrá dada por el cociente de las utilidades marginales del tiempo y el coste $\left(\frac{\mathbf{y}_j}{I}\right)^{14}$, que son los parámetros del tiempo y coste respectivamente en la función de utilidad V_j . Este es el caso concreto de una función lineal en los parámetros y en las variables; no obstante, si se consideran aproximaciones locales de la función de utilidad de orden superior se pueden obtener especificaciones no lineales que incluyan interacciones entre variables como las señaladas en Train y McFadden (1978) y Jara-Díaz y Farah (1987). En estos casos, la obtención de la disposición a pagar no corresponde con el cociente de los parámetros del tiempo y coste, sino que debe obtenerse a partir de la expresión (2.4).

2.2.3 La Teoría de la utilidad aleatoria

Cuando abordamos el problema de modelización de la demanda en un contexto de alternativas discretas, es preciso tener en cuenta los aspectos relevantes que rodean el proceso de toma de decisiones de los individuos. Entre los más importantes cabe señalar los siguientes: definición de la unidad que toma decisiones, determinación del conjunto

¹³ Es la parte representativa de la función de utilidad, que comentaremos más adelante.

¹⁴ Obsérvese que este valor coincide con el valor del ahorro de tiempo en la actividad j en el modelo de DeSerpa.

de elección, selección de los atributos que explican la utilidad de cada alternativa j y determinación de la regla de decisión que sigue la unidad tomadora de decisiones.

El individuo es la unidad típica de toma de decisiones, pero alternativamente también puede considerarse a un grupo, como la familia o el hogar, a una empresa o institución pública. Cada uno de ellos puede representar una única unidad de decisión. Es importante señalar que aunque estudiemos la predicción de la demanda agregada, hay que tener en cuenta las diferencias entre los individuos en el proceso de decisión, debido principalmente a la existencia de distintos patrones de gustos.

Dado un conjunto universal de alternativas (C), el conjunto de elección de un individuo n (C_n) está determinado por aquellas que éste tiene disponibles; a su vez, este es un subconjunto del conjunto universal ($C_n \subset C$) y representa la región factible en el problema de optimización del consumidor, definida a partir de las distintas restricciones del individuo.

En un contexto de elección modal, el conjunto universal está determinado por todos los modos de transporte existentes para realizar un viaje entre el par origen-destino considerado. Por ejemplo, si los modos existentes son coche, bus y metro, un individuo particular podría no tener disponible alguna de las alternativas (por ejemplo, si no tiene permiso de conducir no tendría disponible la alternativa coche como conductor; si la parada de metro más cercana está a una distancia superior a un kilómetro probablemente no sería factible para él ir en metro, etc) y esto debe ser tomado en cuenta a la hora de determinar su conjunto de elección.

Identificadas las distintas alternativas, cada una de ellas es evaluada en términos de un vector de atributos que representan sus características. Así, por ejemplo, los atributos

para el coche pueden ser tiempo, coste y confort; para el bus, tiempo, coste, frecuencia y confort, y para el metro, tiempo, coste, frecuencia y confort. Para cada alternativa, los distintos atributos toman diferentes valores. El individuo elige, dados los valores de los atributos, aquella alternativa que considera mejor, estableciendo una regla de decisión que describe el mecanismo interno utilizado para procesar la información disponible y realizar la elección. Los procesos de decisión pueden ser compensatorios, esto es, cambios en uno o más atributos pueden ser compensados por cambios en otros; o no compensatorios, donde los buenos atributos de una alternativa no permiten compensar los malos. Los procesos de naturaleza compensatoria definen una función objetivo (normalmente, una función de utilidad) que exprese el atractivo de la alternativa en términos de sus características. Está basado en el *compromiso* entre atributos, pues al elegir la alternativa con la utilidad más alta, el individuo selecciona aquella que ofrece una combinación óptima de los mismos. Este tipo de modelos no son apropiados para los procesos de decisión caracterizados por discontinuidades, las cuales son mucho mejor expresadas a través de los procesos no compensatorios. Por su parte, estos procesos definen normas o umbrales que restringen el campo de alternativas factibles, como por ejemplo que el coste del viaje no exceda de una determinada cantidad, así se define el coste del viaje máximo. Estas restricciones son consideradas por los individuos en sus procesos de decisión. Se distinguen la elección por eliminación y la satisfacción.

En la elección por eliminación el individuo jerarquiza los atributos (por ejemplo, el tiempo de viaje es más importante que el coste, el coste que el tiempo de espera, etc.) y define un valor mínimo aceptable o umbral para cada uno de ellos. Posteriormente, se jerarquizan las alternativas de mejor a peor en relación al atributo, por ejemplo que el tiempo de viaje no sea mayor de veinte minutos, y todas aquellas alternativas que no satisfagan esta restricción son eliminadas y así sucesivamente hasta quedarnos con una

sola alternativa. Es posible que repetido el proceso tengamos varias alternativas que cumplan todas las restricciones establecidas, en este caso se elige una de manera compensatoria (ver Tverski, 1972).

Por su parte, la satisfacción presenta varias maneras de búsqueda de la estrategia de elección. Por ejemplo, podría tratarse de un proceso cíclico donde los umbrales se modifican secuencialmente hasta encontrar una alternativa. Asimismo, se podría predefinir una regla de manera que la búsqueda se alcance en el punto en el que se satisface la regla. En este caso, no todos los atributos o alternativas tienen que ser considerados.

De los dos procesos de decisión comentados, el que ha sido más utilizado para explicar el comportamiento de elección del individuo es el compensatorio. No obstante, en algunos casos el individuo podrá no considerar este proceso como criterio de elección, y en estas situaciones se analizará si las preferencias del individuo descritas por sus elecciones cumplen con los axiomas que establece la teoría del consumidor. Este tema será tratado más adelante.

La teoría en que se sustenta el comportamiento del consumidor parte del supuesto que los individuos son racionales (*homo economicus*) y su toma de decisiones se basa en alcanzar unos objetivos dadas unas restricciones. En el contexto de las elecciones discretas, estos objetivos se alcanzan maximizando la FUIIC sobre el conjunto de elección del individuo.

La teoría de la utilidad aleatoria¹⁵ constituye el cuerpo teórico que permite tratar empíricamente el problema de las elecciones discretas. A la hora de modelizar las decisiones de los individuos y obtener empíricamente la FUIIC, el investigador no dispone de información perfecta; es decir, no puede observar todos los factores que participan en el proceso de elección. Puede darse la situación en que dos individuos con el mismo conjunto de alternativas disponibles, sujetos a las mismas condiciones para realizar un viaje concreto e idénticas características socioeconómicas, elijan distintas alternativas. Para resolver este problema, se acude a la teoría probabilística con la idea de tener en cuenta el efecto de factores de naturaleza aleatoria y de este modo corregir las percepciones incorrectas del analista. Así pues, se define una función de utilidad que presenta dos componentes: una parte observable, determinada por el vector de características o atributos medibles de la alternativa y el vector de características socioeconómicas del individuo; y una componente no observable o aleatoria, definida para incorporar las distintas fuentes de aleatoriedad que han sido identificadas; en primer lugar, cualquier atributo no observado. El vector de atributos de las alternativas es normalmente incompleto, no es posible medir el conjunto total de elementos que el individuo considera en sus decisiones. Otra fuente de aparente aleatoriedad es la variación en los gustos no observada; la función de utilidad puede tener elementos no observados que varían entre individuos. En tercer lugar, los errores en la medición de los distintos atributos y en cuarto lugar, elementos de la función de utilidad no observables que se reemplazan por otros elementos que incluyen variables *proxy*.

¹⁵ Formalizada por McFadden (1974), Domencich y McFadden (1975) y Manski (1977). Desarrollos más recientes se pueden encontrar en Ortúzar y Willumsen (2001).

De este modo, la utilidad de una alternativa está representada por una variable aleatoria que se descompone en la suma de dos componentes: la componente observable (V_{jn}) y la no observable (e_{jn}), según se establece en la siguiente expresión:

$$U_{jn} = V(Q_{jn}, S_n) + e(Q_{jn}, S_n) = V_{jn} + e_{jn} \quad (2.18)$$

Tanto la componente observable como la no observable dependen de los atributos de las alternativas y de las características socioeconómicas del individuo.

Al ser la utilidad una variable aleatoria, el individuo n elegirá la alternativa j en su conjunto de opciones disponibles, C_n , con probabilidad igual a:

$$P(j|C_n) = \Pr[V_{jn} + e_{jn} \geq V_{in} + e_{in}, \forall i \in C_n] \quad (2.19)$$

$$P(j|C_n) = \Pr[e_{in} - e_{jn} \leq V_{jn} - V_{in}, \forall i \in C_n] \quad (2.20)$$

Para obtener un modelo de utilidad aleatoria específico, es necesario suponer una distribución de probabilidad conjunta para el término de perturbación aleatoria. En función de las distintas hipótesis formuladas acerca de la distribución de los errores se podrán derivar distintos modelos de elección discreta.

En la especificación de la parte observable¹⁶ se distinguen dos tipos de variables: en primer lugar, los atributos que caracterizan a la alternativa j para el individuo n , esto es, tiempo de viaje, coste del viaje, confort, seguridad, etc. Y en segundo lugar, el vector de variables socioeconómicas que caracterizan al individuo n como la renta, la posesión o no de vehículo, el tamaño del hogar, edad, sexo, ocupación, nivel de estudios, etc. Dos

¹⁶ También se suele denominar utilidad sistemática, determinística o medible.

cuestiones se derivan una vez determinadas las variables a considerar. La primera, que constituye el principal problema en la especificación de la parte determinística de la función de utilidad, está relacionada con la determinación de una combinación de atributos y variables socioeconómicas que reflejen una hipótesis razonable sobre el efecto de tales variables en la utilidad. Mientras que la segunda consiste en determinar la forma funcional. En este sentido, resulta interesante trabajar con especificaciones de funciones que, por un lado, definan una relación entre las variables consideradas y que a su vez estén basadas en supuestos teóricos sobre el comportamiento de los individuos, establecidos por el investigador con base en el conocimiento del mercado objetivo que se analiza; por otro lado, que la especificación presente buenas propiedades en el proceso de estimación. Normalmente, se utilizan funciones que sean lineales en los parámetros, donde las variables pueden entrar en forma polinómica, o empleando cualquier otra transformación que permita relacionar los atributos con la función de utilidad. Por otro lado, las variables pueden ser genéricas o específicas; son genéricas cuando el efecto marginal es el mismo para todas las alternativas y específicas cuando la misma variable presenta efectos diferentes en función de la alternativa que se trate. El problema radica en estudiar cuál debe ser la mejor especificación de la parte observable de la función de utilidad dadas las variables y las características del mercado objetivo.

2.3 La teoría de la elección discreta

2.3.1 Principales modelos de elección discreta

Partiendo de la teoría de la utilidad aleatoria se pueden obtener distintos modelos de elección discreta suponiendo una distribución de probabilidad conjunta para el término de la perturbación aleatoria.

Si consideramos la hipótesis de que los errores distribuyen idéntica e independientemente Gumbel con media cero y varianza s^2 , obtenemos el modelo Logit Multinomial o Logit Simple (MNL). De acuerdo a este modelo, la expresión de la probabilidad de elección de la alternativa j para el individuo n (McFadden, 1974) es:

$$P_{jn} = \frac{e^{bV_{jn}}}{\sum_{A_i \in C_n} e^{bV_{in}}} \quad (2.21)$$

siendo:

$$b = \frac{p}{s\sqrt{6}} \quad (2.22)$$

El parámetro b no se puede estimar por separado del resto de los parámetros definidos en la FUIC (V_{in}) y para muchos propósitos se puede suponer, sin pérdida de generalidad, que b toma valor uno.

El modelo MNL descansa sobre el supuesto de independencia de alternativas irrelevantes. Este se expresa prácticamente en que el cociente entre las probabilidades de elección de dos alternativas i y j es constante y sólo depende de las utilidades de ambas alternativas; así, es independiente de las utilidades del resto de las alternativas.

La expresión de esta propiedad es fácil de obtener a partir de (2.21):

$$\frac{P_{in}}{P_{jn}} = e^{b(V_i - V_j)} \quad (2.23)$$

En un principio, esta propiedad tuvo su importancia ya que permitía analizar la demanda de nuevas alternativas sin necesidad de reestimar el modelo, bastaba simplemente con conocer sus atributos. El problema se presenta cuando existen alternativas

correlacionadas (por ejemplo, coche como conductor y coche como acompañante, o distintos modos de transporte público como metro y tren), porque el modelo conduce a predicciones sesgadas. Una ilustración interesante de este problema puede verse en Mayberry (1973).

Este modelo es homocedástico, esto es, presenta la misma varianza para todas las alternativas y no permite medir correlación entre ellas, de manera que la matriz de varianza y covarianza presenta forma diagonal:

$$\Sigma = s^2 \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix} \quad (2.24)$$

Otros modelos de elección discreta permiten considerar correlación entre las variables e incluso la existencia de heterocedasticidad¹⁷. Conforme más general es la matriz de covarianza, más compleja es la implementación del modelo y su estimación. El modelo Logit Jerárquico (HL) o Anidado (*hierarchical, nested o tree logit*) resuelve en parte el problema de la independencia de alternativas irrelevantes, de manera que es posible considerar cierto patrón de correlación entre alternativas agrupándolas en jerarquías o nidos (Williams, 1977; Daly y Zachary, 1978). Cada nido o jerarquía es representado por una alternativa compuesta frente al resto de las alternativas disponibles. Un ejemplo de una estructura jerárquica o anidada puede ser la de la Figura 2.1.

La obtención del modelo HL se basa en los siguientes supuestos:

¹⁷ Distinta varianza entre alternativas.

Se agrupa en jerarquías o nidos a todos los subconjuntos de opciones correlacionadas entre sí, es decir, aquellas que presentan ciertas semejanzas.

Cada jerarquía o nido es representado por una alternativa compuesta frente a las demás que están disponibles a la población.

La estimación sigue un proceso secuencial donde se estima primero un MNL para aquellas opciones más similares que se han agrupado en un nido, por ejemplo para las alternativas A_3 y A_4 del nido inferior de la figura 2.1, omitiendo todas aquellas variables¹⁸ que tengan el mismo valor para este subconjunto de opciones, $A^I(q)$. La necesidad de omitir estas variables se debe al hecho de que el MNL trabaja en diferencias. Éstas son introducidas posteriormente en el nido superior, puesto que afectan a la elección entre la jerarquía inferior y el resto de opciones en $A^S(q)$.

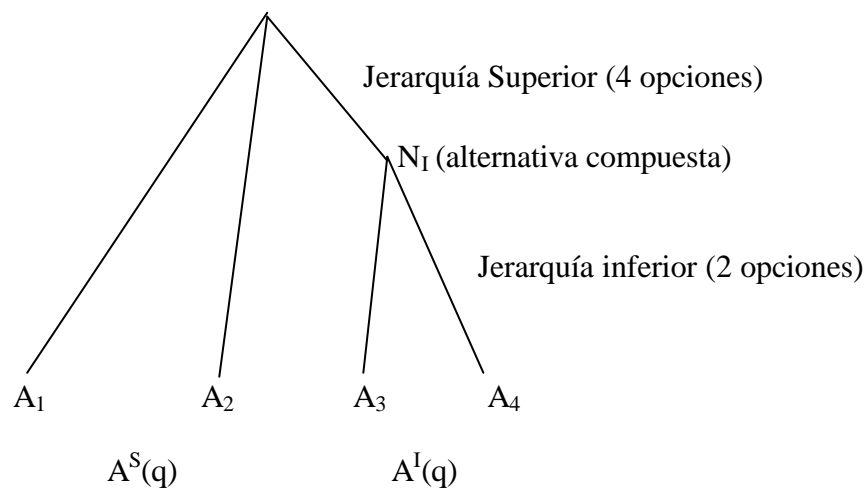


Figura 2.1: Estructura del Logit Jerárquico

La forma en que se introduce el nido inferior en la jerarquía superior es a través de la alternativa compuesta N_I , a la cual se le asocia una utilidad representativa de todo el

¹⁸ Denominadas como vector W .

nido, que va a tener dos componentes (ver Sobel, 1980; Ortúzar, 1983); la primera considera como variable la utilidad máxima esperada de las alternativas del nido (EMU, *expected maximum utility*):

$$EMU = \ln \sum_{A_j \in A^I(q)} \exp(V_j) \quad (2.25)$$

siendo V_j la utilidad de la alternativa j del nido, en la cual no se tiene en cuenta el vector de variables W , es decir, aquellas que toman el mismo valor para las alternativas del subconjunto de alternativas $A^I(q)$.

Y la segunda que considera el vector de los atributos comunes a todos los componentes del nido, W . De este modo, la utilidad compuesta del nido viene dada por:

$$\tilde{U}_j = \mathbf{f}EMU + \mathbf{a}W \quad (2.26)$$

donde \mathbf{f} y \mathbf{a} son parámetros a estimar.

Realizado todo el proceso de estimación para el nido inferior, se estima un MNL para la jerarquía superior donde se incluye la alternativa compuesta y el resto de alternativas contenidas en el nido superior.

De acuerdo con este modelo, la probabilidad de que un individuo n elija la alternativa del nido $A_j \in A^I(q)$ es igual a la probabilidad de que elija la alternativa compuesta (en el nido superior) por la probabilidad condicional de que opte por A_j , en la jerarquía inferior, dado que escogió la alternativa compuesta.

En el caso del ejemplo mostrado en la Figura 2.1 donde hay dos niveles jerárquicos, esta probabilidad sería:

$$P_{jn} = \frac{\exp(\mathbf{f}EMU + \mathbf{a}W + \mathbf{q}^S Z^S)}{\sum_{A_j \in A^S(q)} \exp(\mathbf{f}EMU + \mathbf{a}W + \mathbf{q}^S Z^S)} \cdot \frac{\exp(\mathbf{q}^I Z_j^I)}{\sum_{A_j \in A^I(q)} \exp(\mathbf{q}^I Z_j^I)} \quad (2.27)$$

denotando los superíndices I y S (como en la Figura 2.1) los niveles jerárquicos inferior y superior, respectivamente; Z^I y Z^S los vectores de los atributos de las alternativas disponibles en ambos niveles y \mathbf{q} , \mathbf{a} y \mathbf{f} , los parámetros a estimar.

La consistencia interna requerida por el modelo establece, para el caso de dos niveles jerárquicos¹⁹, que:

$$0 < \mathbf{f} \leq 1 \quad (2.28)$$

Si \mathbf{f} es igual a 1, el modelo HL es matemáticamente equivalente al MNL. Si \mathbf{f} es menor que 0 significa que el aumento de la utilidad de una alternativa del nido disminuirá la probabilidad de elegir dicha opción. Si \mathbf{f} es igual a 0 significa que el aumento de la utilidad de una alternativa del nido no alteraría la probabilidad de elección de la alternativa. Si \mathbf{f} es mayor que 1, el aumento de la utilidad de una alternativa del nido aumentaría tanto su probabilidad de elección como la del resto de las alternativas del nido.

Si el número de jerarquías o nidos definidos es superior a 2, la consistencia interna del modelo vendría definida de la siguiente manera:

$$0 < \mathbf{f}_1 \leq \mathbf{f}_2 \leq \dots \leq \mathbf{f}_s \leq 1 \quad (2.29)$$

siendo \mathbf{f}_I el parámetro correspondiente al nido del nivel inferior y \mathbf{f}_S al nido de nivel superior de cada rama del árbol (Ortúzar y Willumsen, 2001).

¹⁹ El considerado en la Figura 2.1 y en la definición de la probabilidad de elección de la alternativa j para el individuo n .

La matriz de covarianza para este modelo para el caso que hemos supuesto en nuestro planteamiento será:

$$\Sigma = \begin{pmatrix} \mathbf{s}^2 & \cdots & \mathbf{0} \\ \vdots & \ddots & \mathbf{s}_{34}^2 \\ \mathbf{0} & \mathbf{s}_{43}^2 & \mathbf{s}^2 \end{pmatrix} \quad (2.30)$$

Al igual que en el modelo MNL, se trata de un modelo homocedástico y presenta correlación entre las alternativas 3 y 4 que están agrupadas en un mismo nido (la covarianza entre la alternativa 3 y 4 es distinta de cero y $\mathbf{s}_{34} = \mathbf{s}_{43}$). Para una explicación más clara de este modelo ver Carrasco y Ortúzar (2002).

A pesar de que el modelo HL permite resolver, parcialmente, el problema de la correlación entre alternativas y se presenta como un modelo más completo que el MNL, éste posee aún algunas limitaciones (Ortúzar y Willumsem, 2001):

No permite considerar variaciones en los gustos, al igual que el MNL, ni funciones de utilidad heterocedásticas.

No permite tratar los problemas de correlación entre alternativas pertenecientes a distintos nidos; sólo considera las interdependencias entre opciones de los nidos que se haya especificado en la estructura jerárquica a modelizar.

Las posibles estructuras jerárquicas aumentan conforme aumenta el número de alternativas en proporción no lineal. Sobel (1980) encuentra 26 estructuras posibles para un conjunto de cuatro alternativas. No obstante, el problema combinatorio se reduce al estudiar aquellas estructuras que teórica o intuitivamente son más razonables.

Uno de los problemas potenciales de la estimación del modelo por el método secuencial es la transmisión de errores y distinto valor de los parámetros en los distintos nidos (ver Ortúzar *et al.*, 1987). Sin embargo, éstos se han podido solucionar gracias a la aparición de distintos programas²⁰ que permiten la estimación simultánea del modelo.

El modelo Probit Multinomial (MNP) se obtiene de suponer una distribución de probabilidad conjunta para el término de perturbación aleatoria Normal Multivariada con media cero y matriz de covarianza arbitraria (Daganzo, 1979). En este sentido, el MNP permite considerar heterocedasticidad y correlación entre los distintos términos del error. La posibilidad de poder definir cualquier matriz de covarianzas otorga gran flexibilidad y le confiere al modelo MNP la capacidad para estudiar una gama más amplia de comportamientos.

Una cuestión importante de este modelo es la dificultad a la hora de establecer la función de verosimilitud debido a que no existe una expresión cerrada para la probabilidad de elección de las alternativas. Para más de tres opciones, no es posible, analíticamente, estimar el modelo y hay que solventar este problema empleando procedimientos de simulación.

Los problemas observados en la estimación de modelos MNP derivaron en su no utilización por muchos años, y también en la búsqueda de modelos flexibles que fueran más fácilmente tratables desde el punto de vista computacional. Así surgió el modelo Logit Mixto (ML) que presenta una flexibilidad igual a la del MNP y una implementación computacional marginalmente menos compleja. No obstante, avances en las técnicas de simulación y algoritmos de búsqueda, así como en poder de los

²⁰ Software: ALOGIT (Daly, 1987; 1992); LIMDEP (Economic Software, Inc. 1995); HieLoW (Bierlaire y Vandevyvere 1995); GAUSS (Aptech Systems 1994).

ordenadores, hacen que ambos modelos sean factibles de implementar en la práctica; de hecho, actualmente el mayor problema es la mayor dificultad asociada a la interpretación de los resultados.

McFadden y Train (2000) demuestran que el modelo ML puede aproximar cualquier modelo de utilidad aleatoria. Entre sus principales ventajas, tenemos que, con respecto al MNL, permite estudiar la variación en los gustos, patrones de sustitución no restrictivos y correlación entre elementos no observados en el tiempo. Y con respecto al MNP, que no está restringido a la distribución Normal; de hecho es posible definir otras distribuciones²¹ de probabilidad y es algo más sencilla la simulación de sus probabilidades de elección.

La definición del ML está basada en la forma funcional de sus probabilidades de elección. Las probabilidades de elección son las integrales de una probabilidad logit estándar, donde los parámetros no tienen por qué ser fijos y se comportan según una determinada distribución de probabilidad caracterizada por su función de densidad. De este modo:

$$P_{jn} = \int S_{jn}(\mathbf{b}) f(\mathbf{b}) d\mathbf{b} \quad (2.31)$$

siendo

$$S_{jn}(\mathbf{b}) = \frac{e^{V_{jn}(\mathbf{b})}}{\sum_{i=1}^I e^{V_{in}(\mathbf{b})}} \quad (2.32)$$

²¹ Revelt y Train (1998) y Ben-Akiva y Bolduc (1996) han especificado distribuciones Normal o Lognormal, mientras que Revelt y Train (2000), Hensher y Greene (2001), Sillano y Ortúzar (2002) y Train (2001) han utilizado además distribuciones Uniformes y Triangulares.

la probabilidad logit evaluada para el parámetro β , $f(\mathbf{b})$ la función de densidad y V_{jn} la parte determinística de la función de utilidad.

Partiendo de un modelo de comportamiento maximizador de utilidad podemos derivar un modelo ML que permita estudiar variación en los gustos, definiendo un ML de parámetros aleatorios. Para ello se especifica que los parámetros de las variables explicativas de la función de utilidad para cada una de las alternativas varíen según una distribución de probabilidad. Si queremos definir un modelo que considere la existencia de correlación entre las funciones de utilidad de las distintas alternativas, formularemos un modelo de componentes de error. En este modelo, alguno de los parámetros de las variables explicativas de la función de utilidad son fijos y otros aleatorios. Esto da lugar a que el error aleatorio tenga dos componentes: un término que distribuye iid Gumbel y uno adicional que distribuye con media cero.

La utilidad de la alternativa j para el individuo n en el modelo de parámetros aleatorios está dada por:

$$U_{jn} = \mathbf{b}_n x_{jn} + \mathbf{e}_{jn} \quad (2.33)$$

siendo

\mathbf{e}_{jn} , el término de perturbación aleatoria que distribuye iid Gumbel.

x_{jn} , el conjunto de las variables explicativas de la alternativa j .

\mathbf{b}_n , un vector de coeficientes, el cual sigue una determinada distribución, $\beta \sim f(\mathbf{b}, \Omega)$ que depende de los parámetros \mathbf{b} y Ω (generalmente, vector de medias y matriz de varianza y covarianza).

Por su parte, la función de utilidad de la alternativa j para el modelo de componentes de error tiene la expresión:

$$U_{jn} = \mathbf{a} x_{jn} + \mathbf{m}_h z_{jn} + \mathbf{e}_{jn} \quad (2.34)$$

en que

\mathbf{a} , es el vector de parámetros de las variables explicativas de la alternativa j , que es un vector de coeficientes fijos.

x_{jn} y z_{jn} , son vectores de variables observadas.

\mathbf{m}_h , vector de errores aleatorios con media cero.

\mathbf{e}_{jn} , vector de errores que distribuye iid Gumbel.

El modelo de parámetros aleatorios es equivalente al modelo de componentes de error cuando el vector de parámetros \mathbf{b} se descompone en su media \mathbf{a} más las desviaciones \mathbf{m}_h ($\mathbf{b}_n = \mathbf{a} + \mathbf{m}_h$). De forma recíproca, si $z_{jn} = x_{jn}$ el modelo de componentes de error es equivalente a un modelo de parámetros aleatorios; y si $z_{jn} \neq x_{jn}$ se obtendría un modelo de parámetros aleatorios con coeficientes fijos para x_{jn} y coeficientes aleatorios con media cero para z_{jn} .

2.3.2 Procedimientos de estimación

Existen diversos procedimientos para estimar modelos de elección discreta. Los que se emplean con más frecuencia son el método de máxima verosimilitud y el método de

máxima verosimilitud simulada. Los modelos MNL y HL se estiman por el primer método, mientras que los modelos MNP²² y ML se estiman por el segundo. Además de estos métodos, es posible aplicar técnicas Bayesianas para obtener parámetros individuales en modelos de tipo ML.

El método de estimación de máxima verosimilitud está basado en que una determinada muestra puede ser obtenida de varias poblaciones distintas; existe una probabilidad mayor de que sea generada por una cierta población que por otras. La idea del método de estimación se puede ilustrar considerando una muestra de n observaciones de una variable $Z=\{Z_1, Z_2, \dots, Z_n\}$ representativa de una población caracterizada por un vector de parámetros \mathbf{q} , por ejemplo, su media y varianza. La variable Z es una variable aleatoria que tiene asociada una función de densidad, $f(Z, \theta)$, que depende de los valores de los parámetros \mathbf{q} . Siendo independientes en la muestra los valores de la variable Z , podemos escribir la función de densidad conjunta:

$$f(Z_1, Z_2, \dots, Z_n | \mathbf{q}) = f(Z_1 | \mathbf{q}) \cdot f(Z_2 | \mathbf{q}) \cdot \dots \cdot f(Z_n | \mathbf{q}) \quad (2.35)$$

donde estadísticamente se interpreta que f es una función con Z variables y un vector de parámetros \mathbf{q} fijo. Considerando f como función de \mathbf{q} en (2.35) se obtiene la función de verosimilitud $L(\mathbf{q})$. Al maximizar con respecto a \mathbf{q} , se obtienen los estimadores por máxima verosimilitud. Estos representan los parámetros que reproducen con mayor probabilidad la muestra observada (Ortúzar, 2001).

²² Existen otros métodos de estimación como la integración numérica o la aproximación de Clark, véase Daganzo (1979).

En el caso de una muestra de N elecciones discretas independientes, la función de verosimilitud está representada por el producto de las probabilidades de que cada individuo elija la alternativa efectivamente seleccionada:

$$L(\mathbf{q}) = \prod_{n=1}^N \prod_{A_j \in A(n)} (P_{jn})^{g_{jn}} \quad (2.36)$$

siendo

$$g_{jn} = \begin{cases} 1, & \text{si } A_j \text{ es elegida por } n \\ 0, & \text{en otro caso} \end{cases} \quad (2.37)$$

Es habitual maximizar el logaritmo natural de la función de verosimilitud $L(\mathbf{q})$, $l(\theta)$, porque se obtiene el mismo resultado y es más manejable. La nueva función es:

$$l(\mathbf{q}) = \log L(\mathbf{q}) = \sum_{n=1}^N \sum_{A_j \in A(n)} g_{jn} \log P_{jn} \quad (2.38)$$

El conjunto de parámetros obtenidos al maximizar la log-verosimilitud sigue una distribución asintóticamente Normal, $N(\mathbf{q}, \mathbf{S}^2)$, siendo la expresión de la varianza:

$$\Sigma^2 = \frac{-1}{E\left(\frac{\partial^2 l(\mathbf{q})}{\partial \mathbf{q}^2}\right)} \quad (2.39)$$

La obtención de los parámetros implica la aplicación de procedimientos numéricos iterativos. Cuando la función de utilidad es lineal en los parámetros, el método converge rápidamente y siempre existe un máximo único. Esto es una ventaja para la estimación del modelo pues es fácil disponer de software apropiado.

Una cuestión a destacar es que no se puede utilizar como medida de bondad del ajuste el R^2 del método de mínimos cuadrados ordinarios, que está basado en los residuos estimados, porque si el MNL especifica constantes específicas para cada alternativa, se puede asegurar que siempre reproduciría las cuotas de mercado de cada una de ellas. La constante específica recoge el efecto de variables que no han sido consideradas en la especificación del modelo a estimar. Así, no es apropiado comparar la suma de probabilidades de elegir una alternativa con el número total de observaciones de las alternativas que fueron elegidas como medida de bondad del ajuste porque esta condición siempre se cumple si el MNL especifica constantes para todas las alternativas.

El método de estimación por máxima verosimilitud simulada fue propuesto por Lerman y Manski (1981) para la estimación del modelo MNP. Se basa en evaluar la probabilidad de elección generando un número de simulaciones de una distribución Nomal Multivariada. Se considera un éxito cuando U_i es el valor más alto. La proporción de éxitos se aproxima a la probabilidad de elección para un número de simulaciones suficientemente grande. Este método presentaba algunos problemas en la práctica que se han podido resolver en los últimos años gracias a los avances en los procesos de simulación en la modelización de las elecciones discretas. El enfoque propuesto por McFadden (1989) no precisa evaluar la integral múltiple sustituyendo la probabilidad de elección en la ecuación de momentos por un simulador insesgado. El método de momentos simulados puede considerarse el precursor del modelo ML.

Los dos problemas de este método en la práctica eran:

Si el número de éxitos fuera igual a cero, la log verosimilitud tendería a infinito y el método colapsa. Para resolver este problema Lerman y Manski (1981) recomiendan

sustituir el ratio del número de éxitos (N_i) sobre el número total de simulaciones (N) por el cociente $(N_i+1)/(N+J)$ siendo J el número de alternativas; sin embargo, esta modificación introduce sesgo.

El error relativo asociado con el método de simulación es inversamente proporcional a la raíz cuadrada del número de éxitos. Esto implica hacer más simulaciones lo que se traduce en un método más caro.

Para el modelo ML, tenemos que la probabilidad de elección de la alternativa j para el individuo n tiene la expresión general (2.31). Si consideramos que existe variación en los gustos, los parámetros \mathbf{b}_n distribuyen con cierta media y covarianza de manera que la probabilidad de elección (2.31) depende de dos tipos de parámetros; por un lado, del vector de coeficientes que representan los gustos en la población, \mathbf{b}_n y por otro, de los parámetros (\mathbf{t}^*) asociados a la distribución de estos gustos en la población ($f(\mathbf{b}_n | \mathbf{t}^*)$).

La probabilidad de la secuencia de elecciones de un individuo n es:

$$P_n(\mathbf{t}^*) = \int S_n(\mathbf{b}_n) f(\mathbf{b}_n | \mathbf{t}^*) d\mathbf{b}_n \quad (2.40)$$

donde S_n representa la probabilidad conjunta de elección del individuo n en distintas situaciones de elección t dado un conjunto de parámetros \mathbf{b}_n fijo.

$$S_n(\mathbf{b}_n) = \prod_{it} [P_{jnt}]^{g_{jnt}} = \prod_{it} \left[\frac{e^{V_{it}(\mathbf{b}_n)}}{\sum_{it} e^{V_{jnt}(\mathbf{b}_n)}} \right]^{g_{jnt}} \quad (2.41)$$

siendo $g_{jnt} = \begin{cases} 1, & \text{si en la secuencia } t \text{ elige } i \\ 0, & \text{en otro caso} \end{cases}$

y $f(\mathbf{b}_n | \mathbf{t}^*)$ la función de densidad de \mathbf{b}_n .

y la función de verosimilitud tiene la expresión:

$$L(\mathbf{t}^*) = \prod_n P_n(\mathbf{t}^*) \quad (2.42)$$

La función de log-verosimilitud:

$$l(\mathbf{t}^*) = \sum_n \ln P_n(\mathbf{t}^*) \quad (2.43)$$

La integral (2.40) no puede resolverse analíticamente y, por tanto, no es posible maximizar la función de verosimilitud. Esta probabilidad se aproxima por simulación y se maximiza la función de log-verosimilitud simulada. El procedimiento de simulación es el siguiente:

Dar valores iniciales a la media y a la covarianza, \mathbf{t}^* ; dados estos valores se simulan los valores de \mathbf{b}_n a partir de su distribución, $f(\beta_n/\tau^*)$.

Conocidos los valores de \mathbf{b}_n , que han sido obtenidos en la primera etapa, se calcula la probabilidad de la secuencia de elección para el individuo n , $S_n(\mathbf{b}_n)$.

Este proceso se repite para varias simulaciones de \mathbf{b}_n y el promedio se toma como aproximación de la probabilidad de elección:

$$PS_n(\mathbf{t}^*) = \frac{1}{R} \sum_{r=1}^R S_n^r(\mathbf{b}_n^r) \quad (2.44)$$

siendo R el número de repeticiones y \mathbf{b}_n^r la simulación r -ésima de la distribución f .

El estimador que se obtiene por este procedimiento de simulación es consistente y asintóticamente Normal bajo condiciones de regularidad.

También se puede utilizar el método de estimación Bayesiana (HB); éste requiere dos distribuciones: una *a priori* y otra *a posteriori*. La distribución *a priori* añade información sobre los valores de los parámetros (\mathbf{b}) del modelo y está caracterizada por su función de densidad $P(\mathbf{b})$; y representa el conocimiento previo que se tiene sobre los parámetros antes de la obtención de los datos, esto es, las elecciones observadas. Además se obtiene información adicional, concretamente una muestra de elecciones independientes para un conjunto de individuos. Así tendremos un conjunto de T elecciones para cada individuo n , $y_n = \{y_{n1}, \dots, y_{nT}\}$ y el conjunto de elecciones de la muestra total será $Y = \{y_1, \dots, y_N\}$. Con este conjunto Y de elecciones se actualiza la información *a priori* obteniéndose, una nueva función de densidad $P(\mathbf{b}/Y)$ que caracteriza a la distribución *a posteriori*. Esta es una distribución condicional de los parámetros del modelo dada la información *a priori* y el conjunto de elecciones observadas. El teorema de la probabilidad condicionada de Bayes establece la relación definida entre la distribución *a priori* y la distribución *a posteriori*. Tenemos que:

$$\begin{aligned} P(\mathbf{b}, Y) &= P(\mathbf{b}|Y) \cdot P(Y) \\ P(\mathbf{b}, Y) &= P(Y|\mathbf{b}) \cdot P(\mathbf{b}) \end{aligned} \quad (2.45)$$

y por tanto,

$$P(\mathbf{b}|Y) = \frac{P(Y|\mathbf{b}) \cdot P(\mathbf{b})}{P(Y)} \quad (2.46)$$

donde $P(Y)$ es la probabilidad marginal de Y que es constante (no depende de los valores de los parámetros del modelo). Esta probabilidad es la constante de normalización, que asegura que la suma de las probabilidades sea igual a uno (propiedad requerida para las funciones de densidad); por otro lado, $P(Y|\mathbf{b})$ es la probabilidad de observar la elección dados los valores de los parámetros del modelo. Este valor es conocido como la verosimilitud de los datos.

La distribución *a posteriori* se puede expresar como una proporción de la verosimilitud de los datos y de la distribución *a priori*, tal que:

$$P(\mathbf{b}|Y) \propto P(Y|\mathbf{b}) \cdot P(\mathbf{b}) \quad (2.47)$$

Es posible dar una interpretación clásica para los estimadores bayesianos lo que permite su obtención de éstos para los mismos objetivos de análisis que los estimadores máximo verosímiles. Así el Teorema de Bernstein-von Mises²³ (Bernstein, 1917; von Mises, 1931) establece que:

Dado un modelo con parámetros \mathbf{b} , siendo el verdadero parámetro \mathbf{b}^* , el estimador máximo verosímil $\hat{\mathbf{b}}$ y el estimador bayesiano $\bar{\mathbf{b}}$ se tiene que:

La distribución *a posteriori* de \mathbf{b} converge a una distribución Normal con matriz de covarianza $-H^{-1}/N$ al aumentar el tamaño de la muestra²⁴. Es decir,

²³ El planteamiento inicial del teorema es de Laplace (1820) y los trabajos posteriores de Bernstein (1917) y de von Mises (1931) dieron lugar a lo que hoy se conoce como el Teorema de Bernstein-von Mises. En realidad, la explicación y notación corresponden a la utilizada por Train (2002).

²⁴ $-H$ es la matriz de información y se define como $-H = -E \left[\frac{\partial^2 \log P(y_n | \mathbf{b}^*)}{\partial \mathbf{b} \partial \mathbf{b}'} \right]$ para que sea definida positiva, Train (2002).

$\sqrt{N}(\mathbf{b} - \bar{\mathbf{b}}) \xrightarrow{d} N(0, -H^{-1})$. Esto implica que la desviación estándar de la distribución *a posteriori* corresponde a los errores estándar del enfoque clásico.

La media de la distribución *a posteriori* converge al máximo de la función de verosimilitud. Esto es, $\sqrt{N}(\bar{\mathbf{b}} - \hat{\mathbf{b}}) \xrightarrow{p} 0$. Este resultado es una implicación del primer punto. La distribución *a posteriori* es asintóticamente Normal y su media y su máximo coinciden. Asimismo, la distribución *a posteriori* es asintóticamente proporcional a la función de verosimilitud y desaparece la diferencia entre los estimadores máximo verosímil y bayesiano. Esto se debe al hecho de que la distribución *a posteriori* es proporcional a la distribución *a priori* y para muestras grandes desaparece el efecto de la segunda sobre la primera.

La distribución *a posteriori* se distribuye asintóticamente igual que el máximo de la función de verosimilitud. Es decir, $\sqrt{N}(\bar{\mathbf{b}} - \mathbf{b}^*) \xrightarrow{d} N(0, -H^{-1})$. Esto implica que el estimador bayesiano es equivalente al estimador máximo verosímil desde el punto de vista clásico.

El procedimiento para la estimación de la distribución *a posteriori* de los parámetros en un modelo de tipo ML es el denominado Bayes Jerárquico (HB) que fue desarrollado por Allenby (1997), SawtoothSoftware (1999) y Train (2001). La distribución *a posteriori* de los parámetros a estimar es proporcional a la verosimilitud de los datos y a la distribución *a priori*. Se supone, por tanto, que los parámetros distribuyen según una determinada función de densidad conocidos los valores de la media y varianza poblacionales, esto es, $\mathbf{b}_n \sim \mathbf{f}(\mathbf{b}, \Sigma)$. Asimismo, la media poblacional distribuye

Normal mientras que la varianza poblacional distribuye Wishart invertida, quedando definida la distribución *a priori* como $k(b, \Sigma)^{25}$. De esta manera, tenemos:

$$K(b, \Sigma, \mathbf{b}_n \forall n | Y) \propto \prod_n L(y_n | \mathbf{b}_n) \mathbf{f}(\mathbf{b}_n | b, \Sigma) k(b, \Sigma) \quad (2.48)$$

siendo $L(y_n | \mathbf{b}_n)$ la probabilidad de que la persona n realice la secuencia de elecciones observada condicionada al vector de parámetros \mathbf{b}_n .

La estimación del modelo implica un proceso iterativo que requiere de valores²⁶ iniciales de los parámetros. Este proceso aplica el muestreo de Gibbs para obtener simulaciones de la distribución *a posteriori* en la ecuación (2.48). El procedimiento tiene los siguientes pasos:

Se genera un nuevo valor de b condicionado a los valores de Σ y de $\mathbf{b}_n \forall n$. El nuevo valor se muestrea de una distribución Normal con media igual al valor medio de los parámetros \mathbf{b}_n en la población y matriz de covarianza conocida Σ .

Se simula un nuevo valor para la varianza Σ a partir de una distribución Wishart invertida condicionado a los valores de b y \mathbf{b}_n .

Una vez obtenidos los nuevos valores para la media y la varianza (b y Σ , respectivamente), se generan nuevos valores para los parámetros \mathbf{b}_n mediante el algoritmo de Metrópolis-Hastings (SawtoothSoftware, 1999). Este algoritmo se basa en

²⁵ Se puede asumir independencia entre b y Σ y entonces $k(b, \Sigma) = \mathbf{g}(b) \mathbf{y}(\Sigma)$.

²⁶ Estos valores iniciales pueden tomarse de estudios previos y/o de suposiciones del investigador. Suelen considerarse distribuciones poco informativas (por ejemplo media cero y varianza grande).

la obtención sucesiva de valores de \mathbf{b}_n hasta encontrar aquel valor que consigue el mejor ajuste a los datos, de manera que no sea posible alcanzar un ajuste superior.

Cada uno de estos pasos es repetido en cada iteración. El conjunto de parámetros b , Σ y \mathbf{b}_n se re-estiman en cada una de ellas. Este procedimiento de simulación basado en el muestreo de Gibbs converge a la distribución *a posteriori* de todos los parámetros. La convergencia del procedimiento se asegura con un número elevado de simulaciones, entre 100.000 y 200.000. El muestreo de la distribución *a posteriori* requiere de nuevas simulaciones, una vez que el procedimiento converge. Dado que cada iteración se construye a partir de la iteración anterior, existe correlación entre las simulaciones. Para evitarlo se extraen una de cada diez o más. El estimador bayesiano del parámetro es el valor medio de la distribución *a posteriori*.

Como en el caso de método clásico de estimación, es posible realizar distintos test de ajuste general (Ortúzar y Willumsem, 2001). En cada una de las iteraciones se calcula el valor de la función de verosimilitud de los datos con los parámetros estimados hasta ese punto de la simulación. Con este dato se puede calcular la log-verosimilitud como una variable adicional que es reestimada en cada iteración. Alcanzada la convergencia, con los datos de esta variable adicional se obtiene su distribución *a posteriori*. Tenemos así un valor de la log-verosimilitud²⁷ que sirve como indicador del ajuste del modelo y permite realizar los tests de selección de modelos.

Para la estimación de parámetros individuales, además del método de Bayes Jerárquico descrito anteriormente, existe un segundo procedimiento desarrollado por Revelt y Train (2000). Su planteamiento se concreta en obtener la distribución de gustos

²⁷ Este valor presenta las mismas propiedades de la log-verosimilitud estimada por el método clásico.

individuales condicionado a las elecciones observadas de los individuos y la distribución de gustos poblacionales caracterizados por su media y su varianza partiendo de la discusión de Allenby y Rossi (1999).

Dada una secuencia de T_n elecciones y_n y dada la distribución de los gustos en la población, esto es, conocidos los parámetros poblacionales, la densidad condicional de cualquier parámetro individual se expresa a través de la regla de Bayes como:

$$h(\mathbf{b}_n | y_n, b, \Sigma) = \frac{P(y_n | \mathbf{b}_n) \cdot g(\mathbf{b}_n | b, \Sigma)}{P(y_n | b, \Sigma)} \quad (2.49)$$

La esperanza condicional de \mathbf{b}_n se calcula integrando sobre todo el dominio de su distribución, cuya expresión formal es:

$$\begin{aligned} E(\mathbf{b}_n | y_n, b, \Sigma) &= \int \mathbf{b}_n \cdot h(\mathbf{b}_n | y_n, b, \Sigma) \cdot d\mathbf{b}_n = \\ &= \frac{\int \mathbf{b}_n \cdot P(y_n | \mathbf{b}_n) \cdot g(\mathbf{b}_n | b, \Sigma) d\mathbf{b}_n}{P(y_n | b, \Sigma)} = \\ &= \frac{\int \mathbf{b}_n \cdot P(y_n | \mathbf{b}_n) \cdot g(\mathbf{b}_n | b, \Sigma) \cdot d\mathbf{b}_n}{\int P(y_n | \mathbf{b}_n) \cdot g(\mathbf{b}_n | b, \Sigma) \cdot d\mathbf{b}_n} \end{aligned} \quad (2.50)$$

La aproximación simulada de esta expresión se obtiene promediando R observaciones de \mathbf{b}_n^r . Estas observaciones son extraídas de la función de densidad poblacional y cada una de ellas es ponderada por la probabilidad logit condicionada a cada observación. La expresión es, por tanto:

$$ES(\mathbf{b}_n | y_n, b, \Sigma) = \frac{\sum_{r=1}^R \mathbf{b}_n^r \cdot P(y_n | \mathbf{b}_n^r)}{\sum_{r=1}^R P(y_n | \mathbf{b}_n^r)} \quad (2.51)$$

Esta expresión permite obtener los estimadores de los parámetros individuales.

Por otra parte, Revelt y Train (2000) proponen un método alternativo donde los coeficientes del modelo están condicionados por las elecciones individuales. Así la relación de proporcionalidad se puede establecer partiendo de la ecuación (2.49), puesto que el denominador es un valor constante que no depende de los parámetros individuales. Tenemos, entonces:

$$h(\mathbf{b}_n | y_n, b, \Sigma) \propto P_n(y_n | \mathbf{b}_n) \cdot g(\mathbf{b}_n | b, \Sigma) \quad (2.52)$$

Las observaciones de la distribución de los parámetros individuales (2.49) se obtienen utilizando el algoritmo de Metrópolis-Hastings (Chib y Greenberg, 1995) que consiste en extraer simulaciones de manera iterativa de la distribución de los parámetros individuales condicionada a los parámetros poblaciones, $g(\mathbf{b}_n | b, \Sigma)$, que serán aceptadas como valores de los parámetros individuales si incrementan la verosimilitud de los datos. Cada valor muestreado se almacena después de un cierto (elevado²⁸) número de iteraciones. Finalizado el proceso iterativo se construye la distribución de los parámetros calculando las frecuencias observadas para los parámetros individuales durante el proceso de simulación. Este método es bastante sencillo de implementar.

Cabe señalar dos importantes diferencias entre este método y el método HB. La primera está en la forma en que la incertidumbre sobre los parámetros de la población, b y Σ , es definida y calculada. En el HB para calcular la matriz de covarianza se especifica una distribución *a priori* para los parámetros poblaciones que refleja el conocimiento y/o creencia previa sobre éstos antes de examinar los datos, de manera que la matriz de

²⁸ Algunas miles de iteraciones.

covarianza de la distribución *a posteriori* se calcula a partir de la regla de Bayes combinando esta información *a priori* con información exógena (es decir, los datos). Mientras que en este método la matriz de covarianza se obtiene por la maximización de la log-verosimilitud y es calculada en base a una formulación asintótica. La segunda diferencia está en que el HB simula las distribuciones *a posteriori* de los parámetros poblacionales y de cada uno de los parámetros individuales; mientras que el otro requiere calcular la función de verosimilitud para los parámetros de la población.

La información *a priori* es más sencilla de introducir en el análisis bayesiano. Sin embargo, para trabajar con distribuciones distintas a la Normal es más sencillo implementar el método de Revelt y Train (2000).

2.4 La predicción de demanda con modelos desagregados

En la mayoría de las ocasiones, la estimación de modelos de elección discreta que permiten obtener las probabilidades individuales tiene como objetivo final su empleo para la predicción de demanda agregada y su variación ante cambios de política. Pasamos del ámbito individual o desagregado al agregado, fase que requiere de la agregación de las estimaciones obtenidas a nivel individual. Esta agregación puede introducir sesgos y por otra parte, el proceso mismo de agregación no es trivial en la mayoría de los casos²⁹.

Partiendo de un modelo desagregado en el que la probabilidad de que el individuo n elija la alternativa j es:

²⁹ Sólo cuando el modelo desagregado es lineal, entonces el proceso de agregación es trivial; se trata solamente de sustituir los promedios de las variables explicativas en la especificación del modelo.

$$P_{jn} = f_j(\vec{X}_n) \quad (2.53)$$

donde \vec{X}_n , es el vector de variables explicativas, esto es, de atributos de la alternativa j y f_j , es la función de elección que puede ser logit, probit, etc., para la alternativa j . La proporción de mercado estimada de la alternativa j es el valor esperado de las probabilidades individuales sobre el grupo, es decir:

$$P_{jN} = \frac{1}{N} \sum_{j=1}^N f_j(\vec{X}_n) \quad (2.54)$$

siendo N el número de individuos en el grupo. Este procedimiento es conocido como enumeración.

Asimismo, conocida la distribución del vector de atributos para toda la población, el problema de determinar la probabilidad de elección de la alternativa j en el ámbito agregado se reduce al cómputo de la siguiente integral:

$$P_{jN} = \int_{\vec{X}_N} f_j(\vec{X}) g(\vec{X}_N) d\vec{X}_j \quad (2.55)$$

donde $g(\vec{X}_N)$ es la distribución de probabilidad conjunta de las variables explicativas, \vec{X}_j , para el grupo de individuos N .

Ambas expresiones, (2.54) y (2.55), presentan cierta complejidad, y no se utilizan en la práctica ya sea porque requieren una importante cantidad de datos, o bien porque sólo es posible de resolver mediante procedimientos numéricos complejos. Existen diversas técnicas que permiten aproximar estas expresiones. Las más utilizadas en la práctica son el enfoque inocente, el método de clasificación y la enumeración muestral.

El *enfoque inocente* consiste en reemplazar los valores promedio de los atributos en la función individual. Así la probabilidad (2.53) se aproxima por:

$$P_j \approx f_j(\bar{X}_{jn}; \vec{q}) \quad (2.56)$$

Este método sólo es apropiado cuando la función de elección, f_j , es lineal. En el caso que sea no lineal, se pueden cometer sesgos importantes en la predicción. Una manera de reducir este sesgo es dividiendo la población en un número finito de clases homogéneas y aplicar en cada una de ellas el enfoque inocente. Este método se denomina *método de clasificación* y en este caso la predicción agregada está dada por la siguiente expresión:

$$P_{jN} = \sum_{c \in N} \frac{n_c}{N} f_j(\bar{X}_c) \quad (2.57)$$

donde \bar{X}_c es el vector que representa el promedio de los atributos de la clase c y n_c/N es la proporción de individuos en dicha clase. El problema en este método surge a la hora de determinar el número de clases y cómo van a estar definidas. McFadden y Reid (1975) definen un criterio para determinar las clases en función del tamaño de la varianza de las diferencias entre los distintos pares de utilidades estimadas y demuestran que con ocho clases definidas de esta forma, el sesgo de agregación es despreciable.

El *método de enumeración muestral* consiste en aplicar el método de enumeración³⁰ a una muestra representativa de la población. Este enfoque implica una serie de etapas; en primer lugar, predecir el comportamiento de cada individuo en la muestra y obtener el promedio de tales predicciones. En segundo lugar, predecir el comportamiento de cada

³⁰ Ecuación (2.54)

individuo para las condiciones de cambio y obtener el promedio de las mismas; y en tercer lugar, calcular las diferencias entre las predicciones de los pasos 1 y 2 con el objetivo de determinar el efecto agregado que tuvo la aplicación de la política.

Los tres pasos descritos no generan ninguna complicación adicional siempre y cuando la muestra sea aleatoria. Si la muestra es estratificada, es necesario ponderar³¹ cada observación conforme a su representatividad en la población. Este método resulta eficiente cuando el conjunto de elección no es muy grande y las predicciones se realizan a corto plazo. En el largo plazo, no se puede garantizar que la distribución de alternativas y atributos en la muestra se mantenga constante a lo largo del tiempo. Para estos casos, Daly y Gunn (1986) han propuesto aplicar este método para una muestra artificial.

El método de clasificación coincide con el enfoque inocente cuando sólo existe una clase. Y si en el método de enumeración muestral cada individuo es una clase, este enfoque es idéntico al método de clasificación.

Existen otros métodos de agregación basados en la obtención de la distribución del vector de atributos a lo largo de la población empleando diferentes procedimientos. En Ben-Akiva y Lerman (1985) se encuentra una revisión detallada de estos métodos.

Finalmente, se debe señalar que en la etapa predictiva puede aparecer un problema serio cuando se producen cambios en la varianza de los errores de los datos usados en cada una de las distintas fases: de estimación y de predicción. Esto puede ser debido a un cambio en el método de medición de los datos o a un cambio real en la componente

³¹ El factor de ponderación se calcula como el cociente entre la incidencia de la observación en la población y en la muestra.

aleatoria. Ortúzar e Ivelic (1987) muestran que si los errores en la fase de estimación son menores que los errores en la fase predictiva, la demanda será sobrestimada y viceversa.

2.5 Las fuentes de datos

2.5.1 Preferencias Reveladas

Las Preferencias Reveladas (PR) son datos que reflejan el comportamiento actual de los individuos en sus decisiones de viaje. Estos se obtienen a partir de encuestas que permiten recoger información de las variables que explican la utilidad de las distintas alternativas y de las elecciones realizadas. Hasta la mitad de los años 80 éste fue el tipo de datos más utilizado en la modelización de la demanda de transporte. Sin embargo, presentan una serie de limitaciones en términos de comprensión del comportamiento de viajes (ver Ortúzar y Willumsem, 2001):

Las observaciones de las elecciones actuales pueden no presentar suficiente variabilidad para la construcción de buenos modelos que permitan evaluar proyectos y realizar predicciones.

Las variables más interesantes suelen estar correlacionadas, por ejemplo, el tiempo de viaje y el coste. En este caso, es difícil, separar dicho efecto en la modelización y por lo tanto también en la fase predictiva.

No es posible el estudio de variables latentes³². Pueden existir factores que dominen el comportamiento actual, lo que dificulta detectar la importancia relativa de otras

³² Las variables latentes son variables difíciles de medir o que no presentan suficiente variabilidad.

variables igualmente importantes como el confort, la seguridad, la puntualidad del servicio, etc.

No permiten estudiar los efectos de nuevas políticas, como por ejemplo la introducción de un nuevo modo de transporte.

No existe información completa sobre las condiciones del mercado, lo que dificulta determinar el conjunto real de elecciones disponibles.

Están sujetas a potencialmente importantes errores de medida, especialmente de nivel de servicio. Esto se intenta resolver realizando el investigador las mediciones para el conjunto de alternativas disponibles.

2.5.2 Preferencias Declaradas

Las Preferencias Declaradas (PD) son datos que tratan de reflejar lo que los individuos harían ante determinadas situaciones hipotéticas construídas por el investigador. Las PD se desarrollaron inicialmente en el ámbito de la investigación de mercado y comenzaron a ser utilizadas en la modelización de transporte a fines de los años 70. A diferencia de los datos de PR, que entregan información sobre los viajes que realiza un individuo habitualmente, los datos de PD informan sobre los viajes que el individuo realizaría si, por ejemplo, se introdujera un nuevo modo de transporte, se mejorase la calidad del servicio, se ofreciese una ruta alternativa más rápida, etc.

La posibilidad de diseñar experimentos de PD permite, en principio, resolver los problemas que presentan las PR (ver Ortúzar y Willumsen, 2001):

- ✓ Se puede ampliar el rango de variación hasta el nivel en el que existe un *compromiso* entre las distintas alternativas consideradas en el diseño.

- ✓ En la construcción de los escenarios se puede evitar la existencia de correlación entre variables.
- ✓ Permiten incorporar tanto atributos como alternativas no disponibles en el momento del análisis.
- ✓ Se puede aislar el efecto de un determinado atributo así como considerar variables latentes.
- ✓ El conjunto de elección se puede pre-especificar.
- ✓ En el diseño, se pueden evitar los errores de medida.

No obstante, no podemos estar seguros de que el individuo se comporte como dice que haría cuando contesta a una encuesta de PD. Es importante, por tanto, diseñar ejercicios que sean plausibles y realistas para que el entrevistado se implique en el juego correctamente. Los tipos de errores clasificados para este tipo de datos son cuatro (ver por ejemplo Bradley y Kroes, 1990):

Sesgo de afirmación: El entrevistado contesta, consciente o inconscientemente, lo que cree que el entrevistador quiere.

- *Sesgo de racionalización:* El entrevistado intenta ser racional en sus respuestas con el objetivo de justificar su comportamiento en el momento de la entrevista.
- *Sesgo de política:* El entrevistado contesta con el objetivo de influir en las decisiones de política en función de su creencia de cómo pueden afectar los resultados de la encuesta.
- *Sesgo de no restricción:* A la hora de responder no se toma en cuenta todas las restricciones que afectan a su comportamiento, de manera que las respuestas no son reales.

En un ejercicio de PD se pueden distinguir tres elementos principalmente. En primer lugar, tenemos la situación en la que el individuo se encuentra para declarar sus preferencias; ésta puede ser una situación real (un viaje que realice en este momento como ir al trabajo) o hipotética (un viaje que realizaría en el futuro dadas una serie de condiciones), y constituye el contexto de decisión. En segundo lugar, se deben seleccionar las alternativas, normalmente hipotéticas aunque algunas de ellas pueden existir en la actualidad, que se presentan en el ejercicio como función de un conjunto de atributos. En tercer lugar, está la forma en que los individuos pueden declarar sus preferencias (ver Ortúzar y Garrido, 2000); las más frecuentes son: Jerarquización (*Ranking*), Escalamiento o Elección Generalizada (*Rating*) y Elección (*Choice*).

Un punto importante del experimento de PD es la selección de los atributos a considerar en cada alternativa del ejercicio. Una forma de identificar los atributos más relevantes es realizando un grupo de discusión (*focus group*) con una muestra representativa de individuos. Un grupo de discusión no es más que una reunión de individuos dirigida por un moderador de manera que los participantes hablen sobre el tema considerado y durante el desarrollo de la misma se obtengan tanto los elementos clave en el análisis como una mejor comprensión del problema tratado.

El número de atributos a considerar para cada juego es determinado por el investigador; no obstante, la literatura recomienda que no debe ser muy elevado (no más de cuatro) por cada juego para evitar el efecto fatiga (Carson *et al.* 1994) o que contesten de manera lexicográfica (Saelensminde, 1999), tema que trataremos más adelante.

Las unidades de medida en que se expresan los atributos pueden ser triviales en algunos casos, como el tiempo o el coste, pero no lo son en otros como el confort, la seguridad, etc. Estos atributos requieren un estudio más detallado en grupos de discusión, unido a

la realización encuestas piloto que ayuden a encontrar la métrica más adecuada³³. La definición de niveles es también decisión del investigador; sólo se recomienda definir mayores niveles de variación para las variables consideradas más importantes en la elección a estudiar (Wittink *et al*, 1982), pero se debe cuidar que esto no sesgue los resultados del experimento.

El diseño factorial completo determina el número total de opciones dados los atributos y los niveles de variación para cada uno de ellos. Si a es el número de atributos y n el número de niveles de cada atributo, el número total de opciones es n^a . Si tenemos un diseño con dos atributos a dos niveles y tres atributos a tres niveles, el número de opciones resultantes es 108 ($2^2 \times 3^3$). En la práctica no es factible presentar a un individuo 108 opciones para que declare sus preferencias. Este número se puede reducir si consideramos un diseño factorial fraccional, que no es otra cosa que un diseño formado por un subconjunto de opciones del diseño factorial completo. La diferencia entre un diseño factorial completo y un diseño factorial fraccionado está en que mientras el primero permite medir todos los efectos principales e interacciones entre las variables, en el segundo se sacrifica la medición de algunas (o todas) las interacciones con el fin de reducir el número de opciones. Los efectos principales se definen como la respuesta de pasar al siguiente nivel de la variable en cuestión manteniendo constante el resto de variables, mientras que las interacciones consideran los efectos debidos a la variación conjunta de más de una variable; por ejemplo, en la elección modal el efecto de la interacción entre el tiempo y el coste. Normalmente, los cambios en la elección están determinados por los efectos principales y según Louviere (1988):

³³ En la siguiente sección se presenta el proceso de generación del diseño experimental, donde la definición de la variable latente considerada requirió de un trabajo minucioso.

- El 80% o más de la varianza de los datos lo explican los efectos principales.
- Las interacciones de dos términos explican más de un 2% o 3% de la varianza.
- Las interacciones de tres términos explican una proporción muy pequeña de la varianza, del orden del 0.5% al 1% y rara vez sobre el 2% o el 3%.
- Las interacciones de mayor orden explican una proporción minúscula de la varianza de los datos.

Si sólo se desea analizar los efectos principales, el diseño de 108 opciones anterior se reduciría a 16, y a 27 opciones si se considera interacciones de dos términos (Kocur *et al*, 1982), que son números más manejables. No obstante, la literatura (Bradley y Daly, 1994) recomienda no presentar más de 10 (entre 8 y 10) situaciones de elección a cada individuo (ver la discusión de Cherchi y Ortúzar, 2002). En casos en que se requiera estudiar interacciones y el número de opciones sea elevado, se puede utilizar el diseño en bloques; para esto se subdivide la muestra y a cada submuestra se le presenta uno de los bloques (Louviere *et al*, 2000). El tamaño de cada submuestra recomendado (Kocur *et al*, 1982) es de un mínimo de 30 personas por grupo, aunque en la práctica este tamaño parece bajo (ver también la discusión de Ortúzar y Willumsen, 2001).

Las formas de obtener la respuesta en un experimento de PD son, como comentamos anteriormente, Jerarquización, Escalamiento y Elección:

Jerarquización: Se presentan todas las opciones simultáneamente al individuo y se le pide que las ordene en función de sus preferencias, de más a menos preferida. Al ordenar las opciones, el individuo está jerarquizando los valores de utilidad de forma que la opción más preferida le reportará un mayor nivel de utilidad.

Escalamiento: Se le pide al individuo que exprese su grado de preferencia para una opción utilizando una escala arbitraria que puede ser numérica (de 1 a 5 o de 1 a 10) o semántica; por ejemplo: 1 = siempre elijo A, 2 = probablemente elijo A, 3 = ninguna; 4 = probablemente elijo B, 5 = siempre elijo B.

Elección: El individuo selecciona una de las distintas opciones que se le presentan que pueden ser dos (elección binaria) o más de dos (elección múltiple). Se considera que ésta es la forma más sencilla de responder a una encuesta de PD para un individuo porque es la forma habitual en la que toma decisiones. En estos casos, puede incluirse la alternativa “ninguna de ellas” para no forzar al entrevistado a elegir cuando ninguna le parece conveniente.

Un aspecto importante en este tipo de diseños es la necesidad de hacer creíble las distintas opciones presentadas al individuo. Para ello podemos definir el contexto en el que el experimento se realiza y en ese contexto hacer factibles las opciones de elección. Por ejemplo, para que sea considerado viable que exista un bus que realice un recorrido en un tiempo menor que un vehículo privado, puede ser necesario que existan una serie de condiciones tales como: carriles exclusivos para el bus, preferencia en los semáforos, etc. Este tipo de elementos se pueden considerar en el contexto en el que se presenta el experimento haciendo más realistas las opciones.

2.5.3 Datos Mixtos

Con el objetivo de aprovechar las ventajas de cada tipo de datos, PR y PD, e intentar evitar sus desventajas se plantea la *estimación con datos mixtos*³⁴ cuya formalización

³⁴ También conocida con el nombre de *estimación mixta*.

econométrica se debe a Ben-Akiva y Morikawa (1990). Esta estimación se basa en la utilización de ambos conjuntos de datos, PR más PD, en la estimación econométrica de modelos de demanda de transporte de forma que sea posible aprovechar las complementariedades de estas dos fuentes de información.

Desde el punto de vista econométrico, la diferencia fundamental entre los datos de PR y de PD es que presentan distintos tipos de error. Como ya hemos comentado, los datos de PR tienen errores en la medición de las variables independientes³⁵; mientras que los datos de PD presentan errores en la variable dependiente³⁶, debido a que no existe compromiso entre b que el individuo declara que haría y lo que realmente hará cuando se dé la misma situación.

De acuerdo a su naturaleza, las distintas fuentes de datos se clasifican en: *datos primarios* y *datos secundarios*. La fuente de *datos primarios* proporciona información directa acerca de los parámetros del modelo mientras que la fuente de *datos secundarios* proporciona información indirecta. En el caso de modelos de elección modal la principal fuente de información sería una encuesta de PR, *fuentes de datos primarios*, y la información adicional proporcionada por la encuesta de PD se consideraría como *fuentes de datos secundarios*.

La estructura econométrica de estimación conjunta de ambas fuentes de datos se basa en la idea que la diferencia entre los errores de los datos de PR y PD se puede especificar considerando términos de error con diferente varianza. Si ε es el error estocástico de los datos de PR y η el de los datos de PD, podemos expresar la diferencia entre las varianzas a través de la expresión:

³⁵ En nuestro caso, las variables independientes son los atributos de las distintas alternativas.

³⁶ En nuestro caso, la variable independiente es la alternativa elegida.

$$\mathbf{s}_e^2 = \mathbf{m}^2 \cdot \mathbf{s}_h^2 \quad (2.58)$$

siendo \mathbf{m} un parámetro desconocido.

Esta consideración determina que las funciones de utilidad³⁷ de la alternativa j para cada una de las fuentes de datos sean:

$$\begin{aligned} U_j^{PR} &= V_j^{PR} + \mathbf{e}_j = \mathbf{q} \cdot X_j^{PR} + \mathbf{a} \cdot Y_j^{PR} + \mathbf{e}_j \\ \mathbf{m}U_j^{PD} &= \mathbf{m}(V_j^{PD} + \mathbf{h}_j) = \mathbf{m}(\mathbf{q} \cdot X_j^{PD} + \mathbf{w} \cdot Z_j^{PD} + \mathbf{h}_j) \end{aligned} \quad (2.59)$$

donde \mathbf{q} , \mathbf{a} y \mathbf{w} son los parámetros a estimar. X_j^{PR} y X_j^{PD} son atributos comunes de la alternativa j para los datos de PR y de PD respectivamente; mientras que Y_j^{PR} y Z_j^{PD} son atributos no comunes de la alternativa j para cada conjunto de datos.

Al multiplicar la función de utilidad de los datos de PD por el parámetro desconocido \mathbf{m} ecuación (2.59), se consigue que el error estocástico de este tipo de datos tenga la misma varianza³⁸ que los datos de PR.

Normalmente, se supone que los datos de PD deberían tener más ruido que los datos de PR, si éste es el caso, el valor de \mathbf{m} conocido como el *coeficiente de escala* del modelo, debiera estar entre 0 y 1. Si el valor resultara mayor que 1 estaría indicando que los datos con mayor nivel de ruido son los de PR.

Cuando los errores distribuyen Gumbel con media cero y distinta varianza, las probabilidades de elección están dadas por las siguientes expresiones (Ben-Akiva y Morikawa, 1990):

³⁷ Las funciones de utilidad están basadas en el enfoque de la *utilidad aleatoria* expuesto en el epígrafe 2.2.3.

³⁸ La varianza viene definida por la ecuación (2.58).

$$P_j^{PR} = \frac{\exp(\mathbf{q} X_j^{PR} + \mathbf{a} Y_j^{PR})}{\sum_i \exp(\mathbf{q} X_i^{PR} + \mathbf{a} Y_i^{PR})} \quad (2.60)$$

$$P_j^{PD} = \frac{\exp\{\mathbf{m}(\mathbf{q} X_j^{PD} + \mathbf{w} Z_j^{PD})\}}{\sum_i \exp\{\mathbf{m}(\mathbf{q} X_i^{PD} + \mathbf{w} Z_i^{PD})\}}$$

siendo la función de verosimilitud conjunta:

$$L(\mathbf{q}, \mathbf{m}, \mathbf{a}, \mathbf{w}) = \left(\prod_{h=1}^{H^{PR}} \prod_{A_j \in A(n)} (P_{jn}^{PR})^{g_{jn}} \right) \cdot \left(\prod_{h=1}^{H^{PD}} \prod_{A_j \in A(n)} (P_{jn}^{PD})^{g_{jn}} \right) \quad (2.61)$$

donde g_{jn} se define como en la ecuación (2.37)

Los parámetros del modelo se obtienen maximizando la función de verosimilitud que es no-lineal debido a que \mathbf{m} multiplica al resto de los parámetros en la función de utilidad PD. Para la estimación de este tipo de modelos se han desarrollado dos procedimientos: estimación simultánea y estimación secuencial³⁹.

La estimación simultánea fue propuesta por Bradley y Daly (1997). Esta se basa en proponer una estructura jerárquica como la de la Figura 2.2 para resolver la no linealidad (en los parámetros) en la ecuación (2.59). Según esta estructura las alternativas de PR caen directamente del nido raíz y las alternativas PD estarían, cada una de ellas, incluidas en un nido de alternativa única, siendo \mathbf{m} el parámetro que acompaña a la utilidad representativa de cada nido (EMU). El ejemplo mostrado en la Figura 2.2 corresponde a un caso de tres alternativas de PR y tres alternativas de PD.

³⁹ Swait *et al* (1994) han propuesto otro método de estimación secuencial similar al desarrollado por Ben-Akiva y Morikawa (1990).

Para cada observación de PR se considera que no están disponibles las alternativas de PD y para cada observación de PD se considera que no están disponibles las alternativas de PR. La modelización de la elección en este caso se realiza mediante un HL. Este procedimiento de estimación simultánea se puede generalizar a más de dos conjuntos de datos, basta con añadir los nidos artificiales para las alternativas de los conjuntos de datos adicionales teniendo en cuenta que el coeficiente de escala sea el mismo para cada conjunto de datos. Dado que se trata de un modelo tipo HL, la estimación se puede realizar con software estándar como por ejemplo ALOGIT (Daly, 1992).

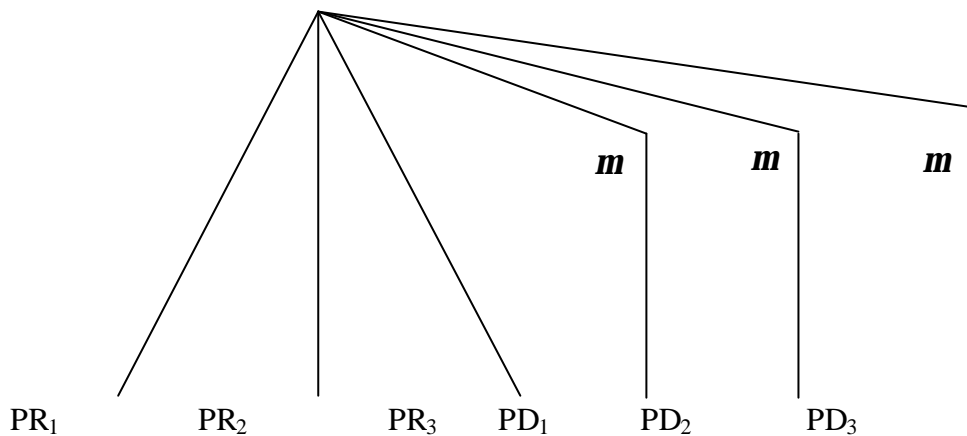


Figura 2.2: Estructura de árbol para estimación con datos mixtos.

El método secuencial fue propuesto por Ben-Akiva y Morikawa (1990). Se basa en un algoritmo con los siguientes tres pasos:

Se estima un modelo con los datos de PD según la expresión (2.59) con el objetivo de estimar los parámetros m_q y m_w . Obtenidos estos valores se define una nueva variable, tal que:

$$\hat{V}_j^{PR} = m_q \cdot X_j^{PR} \quad (2.62)$$

Se estima un modelo con los datos de PR incluyendo la nueva variable definida anteriormente, (2.62), para obtener los parámetros \mathbf{I} y \mathbf{a} según la expresión

$$U_j^{RP} = \mathbf{I}\widehat{V}_j^{PR} + \mathbf{a}Y_j^{PR} + \mathbf{e}_j, \text{ siendo } \mathbf{I} = 1/\mathbf{m}.$$

Finalmente, se ponderan X_j^{PD} y Z_j^{PD} por el *coeficiente de escala* del modelo, \mathbf{m} . Se obtiene así un conjunto de *datos secundarios* que tienen el mismo error aleatorio⁴⁰ que el conjunto de *datos primarios*, lo que posibilita que se estimen los parámetros de forma simultánea trabajando conjuntamente con ambos datos.

Ortúzar y Willumsen (2001) recomiendan utilizar el método simultáneo, que es más fácil de implementar, pues aplicaron tanto éste como el método secuencial al mismo conjunto de datos y los resultados fueron muy parecidos.

Una cuestión a resolver es si es posible combinar dos (o varias) fuentes de datos. Para combinar varias fuentes de datos, se supone que son iguales los parámetros de las variables que son comunes a las distintas fuentes de datos. Para el caso de combinar PR y PD sería \mathbf{q} en las funciones de utilidad de PR y PD, expresión (2.59). Louviere *et al* (2000) proponen estimar modelos MNL para los datos de PR y PD definiendo la misma especificación para las variables comunes a ambos datos. Estimados los parámetros, representar gráficamente los parámetros comunes a PR y PD (θ^{PR} y θ^{PD}), ver Figura 2.3:

⁴⁰ Ambos errores tienen media cero e igual varianza.

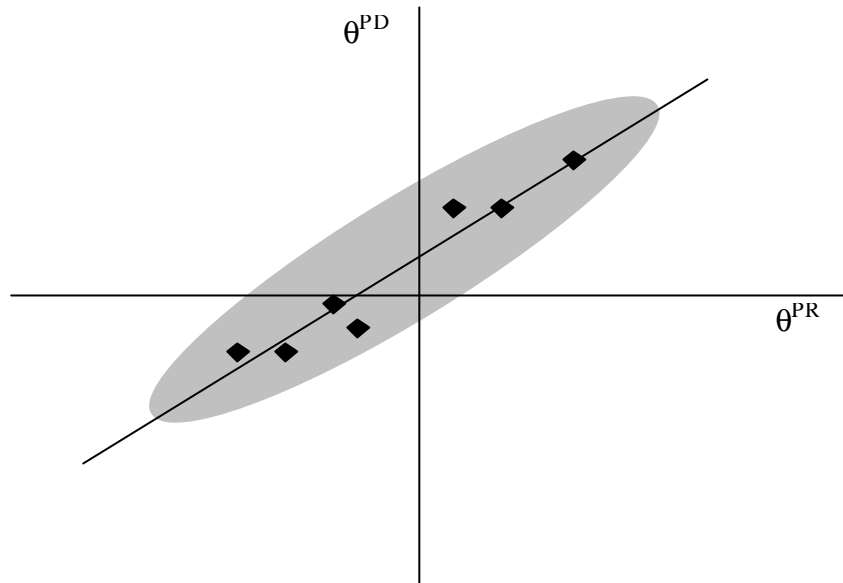


Figura 2.3: Representación de parámetros de PR y PD

Si los parámetros son iguales o parecidos debe resultar una nube de puntos que se aproxima a una línea recta, en este caso se podría estimar el modelo con datos mixtos. La situación contraria, nos estaría indicando que no se debe estimar el modelo combinando las fuentes de datos. Puede darse el caso de que algunos de los parámetros no se concentren en la nube de puntos que da lugar a una línea recta, en este caso se debería realizar una combinación parcial de datos, definiendo iguales parámetros para aquellas variables que resultan iguales entre PR y PD y para los que no, parámetros específicos para cada base de datos. Por ejemplo, según la Figura 2.4, los parámetros específicos serían para las variables tiempo de aparcamiento y frecuencia.

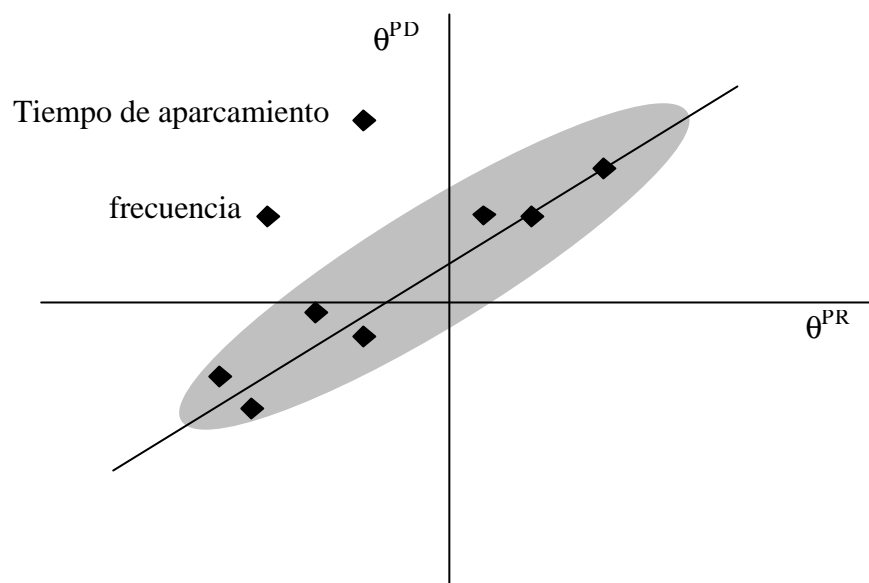


Figura 2.4: Representación de parámetros de PR y PD no iguales

Para la predicción Louviere *et al* (2000) proponen que el modelo contenga las constantes específicas de las alternativas de PR así como todos los parámetros estimados conjuntamente. Si los parámetros no son estimados conjuntamente, surge el problema de cuáles considerar, sugieren que si el parámetro de PR⁴¹ no estimado conjuntamente tiene el signo incorrecto, deberíamos utilizar el correspondiente al parámetro de PD y si existe una diferencia importante entre el parámetro de PR y de PD, no está claro cuál debe ser la solución; se podría estudiar cuán de sensible es la elección al parámetro con los datos de PR o una muestra de validación (holdout data set).

Alternativamente, existe un test para verificar la hipótesis de homogeneidad de los parámetros propuesto por Swait y Louviere (1993) cuyo procedimiento es el siguiente:

⁴¹ Según explican estos autores, es más frecuente que los datos de PR den lugar a estimaciones de parámetros con signo intuitivamente incorrectos (pp. 246).

Estimar un MNL para cada conjunto de datos de manera independiente, según la expresión (2.21).

Estimar el modelo conjuntamente según el método manual propuesto por los autores en dicho artículo.

Calcular el estadístico siguiente:

$$\mathbf{c}^2 = -2 \left[(L^{PR} + L^{PD}) - L^{PR+PD} \right] \sim \mathbf{c}_{q-1}^2 \quad (2.63)$$

Este test es generalizable a cualquier número de bases de datos.

Si se acepta la hipótesis nula, existe homogeneidad de parámetros entre ambas bases de datos, y por tanto, se pueden combinar ambos datos. Si no se aceptara, probaríamos con la representación gráfica, pues el test puede rechazarse porque algunos parámetros no sean iguales, situación de la Figura 2.4.

3. La oferta y la demanda de servicios de transporte de pasajeros en los corredores

3.1 Introducción

La ampliación de la capacidad viaria no siempre es la mejor solución para resolver los problemas derivados del crecimiento de la demanda de transporte de pasajeros; mayor capacidad incentiva mayores desplazamientos, lo que tiene un impacto medioambiental y monetario que no compensa la mejora en la fluidez del tráfico. Asimismo, el transporte privado es menos eficiente en el uso del espacio y energía que el transporte colectivo; las emisiones de gases contaminantes son de cuatro a ocho veces menores, el espacio necesario para un autobús es sólo el 5% del necesario para transportar el mismo número de personas en automóviles y el coste estimado en transporte privado es cuatro veces superior que en transporte colectivo (Boletín de la Comisión Europea, suplemento 4/95).

Por tanto, es de vital importancia definir estrategias para el sistema de transporte basadas en las ventajas del transporte colectivo y en la internalización de los efectos externos que generan los distintos modos, principalmente el vehículo privado. Las líneas de acción deben estar dirigidas al fomento del uso del transporte colectivo siendo

necesaria una red para los ciudadanos y por tanto, una integración de los distintos modos de transporte.

La tasa de motorización⁴² en la Unión Europea es de 469 vehículos por cada mil habitantes; en España, es de 442; y en Gran Canaria, es aún mayor, del orden de 624 vehículos por cada mil habitantes. Dada la escasez de espacio, es cada vez más importante una planificación de los sistemas de transporte dentro de las líneas directrices de la ordenación del territorio. En este punto, el análisis de la demanda de transporte juega un papel vital para la adecuación de la oferta de transporte a la misma.

Este capítulo está dedicado a la caracterización del mercado objetivo, distinguiendo entre la oferta y la demanda de transporte. En la segunda sección se describe de manera general la oferta de servicios de transporte planteada desde los aspectos del servicio ofrecido por las distintas empresas, el sistema tarifario y el sistema integrado de tarifas que tiene implicaciones sobre el coste del transporte público, y que será tenido en cuenta en nuestro diseño de PD. En la tercera sección, se caracteriza la demanda de servicios de transporte antes de la realización de nuestra investigación, mientras que en las secciones cuarta y quinta se hace referencia a la información utilizada en el estudio; concretamente se explican los dos tipos de encuestas realizadas, encuesta de PR y PD, incidiendo en el caso de esta última en el diseño del experimento.

⁴² Fuente: Instituto Nacional de Estadística e Instituto Canario de Estadística, datos del año 2001.

3.2 La oferta de servicios de transporte

3.2.1 El transporte regular de pasajeros

El transporte regular de viajeros es ofrecido en la isla de Gran Canaria por tres empresas⁴³, *Guaguas Municipales*, *Salcai* y *Utinsa*⁴⁴. La empresa *Guaguas Municipales* ofrece servicios urbanos⁴⁵ en la ciudad de Las Palmas de Gran Canaria, capital de la isla; mientras que las empresas *Salcai* y *Utinsa*, enlazan la capital de la isla con el resto de municipios, ofreciendo un servicio interurbano.

La empresa *Salcai* es titular de la concesión del transporte regular de viajeros en la zona Este-Sur-Suroeste⁴⁶ de Gran Canaria. Ofrece un servicio interurbano entre la ciudad de Las Palmas de Gran Canaria y ocho municipios de la isla (Telde, Ingenio, Agüimes, Santa Lucía de Tirajana, San Bartolomé de Tirajana, Mogán, Valsequillo y San Nicolás de Tolentino) además de los principales núcleos turísticos y playas. La oferta de servicios es de 61 líneas⁴⁷, de las cuales 15 operan en el corredor objeto de estudio: Telde-Las Palmas de Gran Canaria.

La empresa *Utinsa* es concesionaria del transporte regular de viajeros entre la ciudad de Las Palmas de Gran Canaria y los municipios del Norte-Noroeste⁴⁸ de la isla, un total de

⁴³ Existen otras empresas del ámbito urbano en varios municipios de la isla que dan servicio a determinados núcleos de población del mismo municipio. Normalmente, se trata de núcleos retirados que es necesario conectar con el resto del municipio.

⁴⁴ Las empresas de *Salcai* y *Utinsa* se fusionaron en enero de 2000 creando la empresa *Global*. Esta fusión no ha afectado al servicio de transporte ofertado en los corredores objeto de estudio.

⁴⁵ El trayecto urbano es el que considera un trayecto entre dos puntos o zonas urbanas de un mismo municipio; mientras que el trayecto es interurbano cuando cada zona pertenece a distintos municipios.

⁴⁶ Desde este punto, cuando hagamos referencia a la zona en la que opera *Salcai* abreviaremos diciendo zona Sur, aunque obviamente se trate de la zona Sur-Este-Suroeste.

⁴⁷ Actualmente, existen 60 líneas. En el año 2003 se crearon cinco líneas nuevas y dejaron de operar seis líneas respecto a la situación del 2000. No obstante, esto no afecta a la investigación, pues se trata de líneas que no operan en el corredor Telde-Las Palmas.

⁴⁸ A partir de este momento, para simplificar nos referiremos a zona Norte cuando hablemos del ámbito territorial en el que trabaja *Utinsa*.

doce municipios (Artenara, Agaete, Firgas, Teror, Santa Brígida, Valleseco, San Mateo, Tejeda, Gáldar, Santa María de Guía, Moya y Arucas). La oferta de servicios está constituida por 76 líneas, de las cuales 16 operan en el corredor Arucas-Las Palmas de Gran Canaria, objeto de nuestra investigación.

La empresa *Guaguas Municipales* es titular de la concesión del transporte regular de viajeros en la ciudad de Las Palmas de Gran Canaria. La oferta de servicios es de 38 líneas⁴⁹, seis de las cuales tienen servicio permanente, con una frecuencia que varía en las distintas franjas horarias del día.

Cada empresa funciona como un monopolio dentro de su servicio concesionado. El contrato de concesión define la tarifa y el nivel de servicio.

En la Figura 3.1 se puede ver en color naranja la concesión de *Salcai* y en azul, la concesión de *Utinsa*.

3.2.2 Sistema tarifario

La distinción de trayectos en interurbano y urbano define un sistema tarifario en función del número de kilómetros recorridos. En el caso de trayectos urbanos se define una tarifa única que es aplicada por *Guaguas Municipales*; y en el caso de trayectos interurbanos se define una tarifa kilométrica con un mínimo de percepción para aquellos trayectos inferiores a un determinado número de kilómetros, esta tarifa es la que aplican las empresas *Salcai* y *Utinsa*.

⁴⁹ En el año 2003 han ampliado su oferta a 40 líneas.

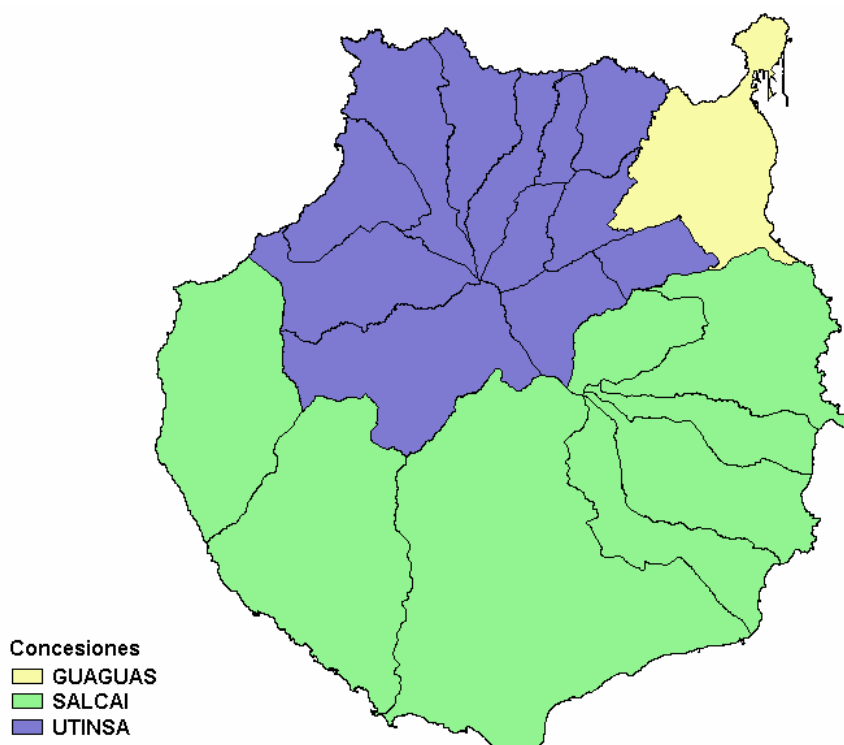


Figura 3.1: Concesiones de las empresas

El mínimo de percepción en la tarifa kilométrica se define para los trayectos inferiores o iguales a doce kilómetros y a partir de trece kilómetros no se aplica el mínimo de percepción y la tarifa es igual al número de kilómetros por pts/kilómetro. Asimismo, en cada empresa se distinguen dos tipos de títulos de viaje: un primer título se denomina el billete sencillo, corresponde al pago directo en el momento de subirse al bus; el segundo es un bono. Este último implica un pago anticipado mediante la compra de una tarjeta que permite realizar más de un viaje. Normalmente, cada empresa tiene sus propios tipos de bonos, pero todos tienen en común que su compra permite realizar diez viajes y que al usuario el precio de cada uno de esos diez viajes le supone un descuento sobre el precio del billete sencillo. Este descuento varía dependiendo del tipo de bono y de la empresa que se trate entre un 30 y un 70% sobre la tarifa normal. Además, el 20% del descuento aplicado es subvencionado por la Autoridad Única del Transporte del Cabildo de Gran Canaria, con fondos de la administración central.

Las tarifas sobre las que se aplican los descuentos son diferentes para cada empresa. *Salcai* establece un mínimo de percepción de 130 pts (0,78 €) y 10,9578 pts (0,07 €) por viajero/km. *Utinsa*, establece un mínimo de percepción de 135 pts (0,81 €) y 11,118 pts/km (0,07 €). Y finalmente, *Guaguas Municipales* define una tarifa única de 130 pts (0,78 €).

La empresa *Salcai* ofrece un mayor número de títulos de viaje que el resto de empresas. Por ejemplo, la tarjeta de Abono es de 10 viajes para un origen- destino definido ex-ante; se diferencian distintos tipos de descuento en función de ciertas características del individuo. En el abono general el descuento que obtiene el usuario es del 35% sobre la tarifa del trayecto elegido. En el abono de estudiantes es del 40% y en el de familia numerosa del 40, 60 y 70% según se trate de familia numerosa de primera, segunda o tercera, respectivamente⁵⁰. La tarjeta dinero no necesita la definición de ningún trayecto concreto, permite realizar cualquier recorrido de los que la empresa realiza con un descuento del 25% sobre la tarifa normal con un coste o saldo de 2.000 pts (12 €). La tarjeta playa es para viajar a las playas (San Agustín, Maspalomas, Playa del Inglés, Puerto Rico, Puerto de Mogán, etc) en los días azules (sábados, domingos y festivos) con un descuento del 50%.

La empresa *Utinsa*, por su parte, ofrece dos tipos de bonos. El bono de trayecto definido, similar al abono general de *Salcai*, es una tarjeta de diez viajes para un trayecto concreto, siendo el descuento a aplicar a la tarifa normal del 33%. Y la tarjeta monedero, similar a la tarjeta dinero de *Salcai*, es una tarjeta por valor de 2.000 pts (12

⁵⁰ Una familia numerosa de primera es aquella en la que entre tres y cinco hijos menores de 21 años (o 25 si son estudiantes universitarios), dependen económicamente de su padre o madre. La familia numerosa es de segunda si el número de hijos dependientes está entre seis y ocho, y de tercera si los hijos son nueve o más.

€) con la que se puede realizar cualquier tipo de viaje, dentro del servicio que ofrece la empresa, con un descuento del 30% sobre el coste del trayecto.

Finalmente la empresa *Guaguas Municipales*, ofrece varios tipos de bonos; al tener definida una tarifa única los bonos pueden ser utilizados para cualquier línea y recorrido. El bono “guagua⁵¹”, el más demandado, es un bono de diez viajes donde el usuario se beneficia de un descuento del 38,5% sobre la tarifa única. El bono de estudiantes es una tarjeta que permite realizar 80 viajes. El precio de este bono es de 3.070 pts, lo que supone un ahorro del 51,25% sobre el coste del viaje (cada viaje tiene un coste de 38,875 pts). El bono fácil es una tarjeta mensual que permite realizar todos los viajes que el usuario desee durante el mes sin ningún tipo de restricción por un precio fijo de 4.590 pts. Los individuos que optan por este título de viaje son los que se mueven con una frecuencia más elevada para que esta opción les sea más rentable. Haciendo unos pequeños cálculos podemos observar que para que este título sea más rentable que el bono guagua, el viajero debe realizar más de 60 viajes al mes, número que no es tan difícil de conseguir.

Un trabajador normal que realiza todos los días laborables el mismo trayecto desde su hogar al lugar de trabajo dos veces, correspondientes a la ida y a la vuelta. Suponiendo que este individuo se desplaza en transporte público (la guagua) y cada trayecto consta de una sola etapa motorizada, en un mes realizaría entre 40 y 44 viajes para desplazarse al trabajo. Esta cantidad se duplicaría en el caso en que cada trayecto contase de dos etapas motorizadas en transporte público. De esta forma podemos definir unos intervalos de coste por cada viaje para este individuo en función del número de viajes

⁵¹ Guagua es como se conoce comúnmente al bus en las Islas Canarias.

que realice al mes. Estos intervalos serían [114,75 pts; 104,31 pts] en el caso de una sola etapa motorizada y [57,375 pts; 52,16 pts] en el caso de dos etapas motorizadas donde no hay despenalización por transbordo. De este análisis, se deduce que la oferta de *Guaguas Municipales* resulta atractiva para aquellas personas que realizan un elevado número de trayectos. Finalmente, el bono para jubilados depende de las características socioeconómicas y exige ciertos requisitos, pero supone la gratuidad del viaje. En la Tabla 3.1 se presenta un resumen de los títulos de viaje ofrecidos por cada una de las empresas.

Tabla 3.1: Títulos por empresas

Salcai		Utinsa		Guaguas Municipales	
Billete sencillo		Billete sencillo		Billete sencillo	
Abono general	35%	Bono trayecto definido	33%	Bono guagua	38,5%
A. estudiantes	40%			Bono estudiantes	51,25%
A. Familia numerosa	40%			Bono fácil	variable
	60%				
Tarjeta dinero	25%	Tarjeta monedero	30%	Bono jubilados	100%
Tarjeta playa	50%				

3.2.3 Sistema integrado de transporte

El *Sistema Integrado de Transporte*⁵² en la isla de Gran Canaria pretende una integración insular y territorial del transporte, siguiendo las directrices de planificación definidas por la Comisión Europea⁵³. El objetivo principal de la integración es que se permita que cualquier individuo pueda desplazarse entre dos puntos cualesquiera de la

⁵² A partir de este punto abreviaremos diciendo *S.I.T.*

⁵³ Véase el Boletín de la Unión Europea (1995), suplemento 4/95: “Una red para los ciudadanos”

geografía del territorio insular con la posibilidad de realizar transbordo de una línea a otra.

El *S.I.T.* plantea una reorganización de los servicios de transporte público ofrecidos en la isla de Gran Canaria. Esta reorganización va a implicar en un primer lugar, una *integración tarifaria*, que incluye la definición de títulos de viaje y tarifas sin distinción de empresas. En segundo lugar, una *coordinación de los servicios*, para facilitar los transbordos entre las distintas empresas, siendo el tiempo disponible para realizar el transbordo de 45 minutos a partir de que finaliza el primer viaje. Y en tercer lugar, una *cámara de compensación* que está definida como la institución encargada de llevar a cabo las compensaciones monetarias entre las distintas empresas, por un lado, debido a los transbordos, y por otro, debido a la venta de los títulos que se realizan en los puntos de venta de las empresas.

Los títulos definidos para la *integración tarifaria* son:

- i. **Billete sencillo:** Es la tarifa definida para la realización de viajes de una etapa en alguna de las líneas de las diferentes empresas, con un mínimo de percepción homogéneo, es decir, el mismo para todas y cada una de las empresas, pero manteniendo tarifas kilométricas para *Salcai* y *Utinsa*⁵⁴ y tarifa plana para *Guaguas Municipales*⁵⁵. En el billete sencillo no existen variaciones respecto a la situación anterior.
- ii. **Tarjeta Valor:** Es una tarjeta monedero que permite acceso a toda la red insular. Está dirigida a los usuarios frecuentes del transporte público en general, sin trayecto

⁵⁴ Estas empresas realizan trayectos interurbanos, y siempre que hablemos de tales trayectos, a partir de este punto, estarán referidos a los que realizan estas empresas.

⁵⁵ Cuando hablemos de trayectos urbanos nos referiremos a los que realiza esta empresa.

definido. A la hora de la cancelación del importe del trayecto, cualquiera que sea éste, urbano o interurbano, el individuo percibe un descuento del 30% sobre el importe del trayecto si éste es la primera etapa del viaje. Si fuera una segunda etapa, es decir, el individuo ha realizado transbordo de una línea a otra, el descuento es del 70% si alguna de las dos etapas es urbana y del 30% si no lo es. Ahora bien, el descuento sobre la tarifa urbana no supone problema, pues es una tarifa plana; sin embargo, la tarifa interurbana exige un mínimo de percepción; por lo tanto, en este caso el descuento en la segunda etapa se aplica sobre el coste por kilómetros sin el mínimo de percepción, pues el individuo ya soportó un mínimo de percepción en la primera etapa.

iii. **Tarjeta Cliente:** Es una tarjeta monedero que permite acceso a toda la red insular al igual que la *tarjeta valor*. Está dirigida a los usuarios *recurrentes*, y por tanto, opera en un trayecto definido. El descuento aplicable a la hora de cancelar el importe del trayecto es progresivo hasta un 50%, mientras que para la segunda etapa es del 100%; es decir, el transbordo es gratuito siempre. El descuento progresivo se establece en función del volumen de gasto del individuo partiendo de un 30% de descuento como mínimo y hasta un 50% como máximo. En la definición de este descuento se tienen en cuentas las cuestiones siguientes:

- a. Los descuentos son acumulables, pues se trata de descuentos personales y cuando el usuario compra otra tarjeta cliente se le define en la misma el descuento que hasta ese momento tiene acumulado.
- b. La reducción progresiva se establece en función del volumen de gasto para beneficiar al usuario de la manera más equitativa posible.
- c. A pesar de existir un trayecto definido para la *tarjeta cliente*, el usuario tiene la posibilidad de realizar otros trayectos; en este caso el título

trabaja como una *tarjeta valor* y no contabiliza para la reducción progresiva del importe del trayecto definido.

Los distintos títulos de viaje así como las posibles combinaciones de trayectos urbanos e interurbanos se presentan en la Tabla 3.2. En esta tabla se presentan las combinaciones para viajes con dos etapas. Si el viaje tuviera más de dos etapas, sería como un nuevo viaje donde, por ejemplo, si el viaje requiere de cuatro etapas, la tercera etapa tendría el mismo tratamiento que la primera etapa y la cuarta el mismo que la segunda etapa.

Es importante señalar que el proceso de integración tarifaria se ha realizado de manera parcial. Hasta el momento, solo se ha introducido la tarjeta valor, llamada tarjeta insular, que convive con los títulos de viaje de las distintas empresas y no hay fecha para la integración total.

En nuestra investigación, consideramos la integración que implicaba la introducción de la tarjeta valor o tarjeta insular para la definición de las tarifas en el bus mejorado presentado en la encuesta de preferencias declaradas (PD), que veremos más adelante.

Tabla 3.2: Títulos de viaje del S.I.T.

Billete sencillo		
Trayecto urbano	Tarifa plana	130 pts
Trayecto interurbano	Tarifa kilométrica	Km*pts/km con un mínimo de percepción
Tarjeta Valor		
Combinaciones	1ª etapa	2ª etapa
Urbano-urbano	30% s/ tarifa plana	70% s/ tarifa plana
Urbano-interurbano	30% s/ tarifa plana	40% s/ tarifa plana
		30% s/ tarifa plana
Interurbano-urbano	30% s/ tarifa kilométrica	70% s/ tarifa plana
Interurbano-interurbano	30% s/ tarifa kilométrica	30% s/ tarifa kilométrica ¹
Tarjeta Cliente		
Urbano-urbano	[30-50%] s/ tarifa plana	100% s/ tarifa plana
Urbano-interurbano	[30-50%] s/ tarifa plana	[70-50%] s/ tarifa plana
		[30-50%] s/ tarifa kilométrica ¹
Interurbano-urbano	[30-50%] s/ tarifa kilométrica	100% s/ tarifa plana
Interurbano-interurbano	[30-50%] s/ tarifa kilométrica	[30-50%] s/ tarifa kilométrica ¹

¹sin mínimo de percepción.

Para analizar los efectos de la introducción de los nuevos títulos sobre el coste del viaje, vamos a considerar tres tipos de desplazamiento. El primero tiene solamente un trayecto interurbano; el segundo, un trayecto interurbano y otro urbano; y el tercero dos trayectos interurbanos. Estas son las tres opciones con que nos podríamos encontrar en los corredores analizados. Consideremos ahora tres casos particulares de viajes incluyendo estas opciones en cada corredor. Para Arucas tenemos que el primer viaje es a la zona de Mesa y López (trayecto interurbano), el segundo a la zona de Siete Palmas (trayecto interurbano + urbano) y el tercer viaje al campus universitario de Tafira (dos trayectos interurbanos). Para Telde, el primero es hasta Triana (trayecto interurbano), el segundo hasta el centro comercial Las Arenas (trayecto interurbano + urbano) y el tercero a Tamaraceite (dos trayectos interurbanos).

Para cada uno de estos viajes se analizará las distintas tarifas (normal, bono, tarjeta valor y cliente), así como los porcentajes de ahorro, respecto al nuevo sistema de tarificación (Tabla 3.3).

Tabla 3.3: Tarifas según títulos

Origen	Viaje	Tarifa Normal	Tarifa Bono	Tarjeta Valor	Tarjeta Cliente
Arucas	I ¹	175 ³	117	123	123-88
	I+U ²	460	234	200	161-115
	I+I	460	241	238	238-170
Telde	I	175	114	123	123-88
	I+U	305	194	162	123-88
	I+I	305	198	172	172-123

¹Trayecto interurbano.

²Trayecto urbano.

³En pesetas.

En la tabla se presentan las tarifas en pesetas para las distintas opciones: *normal* para la tarifa correspondiente al pago directo en el bus, esto es, el billete sencillo; *bono* cuando la tarifa que se aplica es la derivada de la utilización de los bonos generales de las distintas empresas; *valor* cuando se aplican los descuentos definidos en la tarjeta valor y *cliente* cuando se aplican los descuentos de la tarjeta cliente. En este último caso, se puede aplicar un descuento mínimo o máximo, por lo que se llega a un intervalo de precios definido según el criterio del descuento; por ejemplo, 123-88, indica que la tarifa de 123 pts corresponde al descuento mínimo y la tarifa de 88 pts al descuento máximo. En la Tabla 3.4 se presentan los ahorros (en porcentaje) que implican las nuevas tarifas (tarjeta valor y cliente) con respecto a las ya existentes (normal y bono) y el ahorro sobre la tarifa normal por la utilización del bono, para poder comparar las mejoras en coste debido los nuevos descuentos que definen las tarjetas valor y cliente.

Para los viajes de un solo trayecto interurbano (I), la tarjeta valor no supone mejoras en términos de menor coste en el bus porque el descuento aplicado es menor que el

aplicado por las empresas en la actualidad, y comparado con la tarifa bono el coste es mayor con la tarjeta valor. Con la tarjeta cliente⁵⁶, el menor coste se da cuando el descuento es el máximo, lo que supone un ahorro del 50% sobre la tarifa normal y entre el 23 y el 25% sobre la tarifa bono.

Tabla 3.4: Ahorros por nuevas tarifas

Origen	Viaje	Ahorro Bono	Ahorro Tarjeta Valor		Ahorro Tarjeta Cliente	
			Tarifa Normal	Tarifa Bono	Tarifa Normal	Tarifa Bono
Arucas	I ¹	33 ³	30	(-5)	30-50	(-5)-25
	I+U ²	49	57	15	65-75	31-51
	I+I	48	48	1	48-63	1-30
Telde	I	35	30	(-8)	30-50	(-8)-23
	I+U	37	47	17	60-71	37-55
	I+I	35	34	13	34-60	13-39

¹Trayecto interurbano.

²Trayecto urbano.

³En porcentaje.

En el caso de los viajes con dos trayectos (I+U), uno interurbano y otro urbano, es donde los ahorros para ambas tarjetas son mayores. Para la tarjeta valor, el ahorro está entre un 47-57% con respecto a la tarifa normal y entre un 15-17% para la tarifa bono. Para la tarjeta cliente, los ahorros van desde un 60-75% sobre la tarifa normal y entre un 30-55% sobre la tarifa bono.

Y para los viajes con los dos trayectos interurbanos (I+I), los ahorros no son tan importantes como en el caso anterior. Con la tarjeta valor, los porcentajes de ahorros son los mismos que con la utilización de los bonos de las empresas y con la tarjeta cliente, tenemos un ahorro entre el 34-63% sobre la tarifa normal y entre 1-39% sobre la tarifa bono.

⁵⁶ La tarjeta cliente exige un trayecto definido ex-ante.

Como conclusión, podemos decir que la tarjeta valor supone una mejora en términos de coste para los usuarios de bus que realizan trayecto interurbano y urbano, con lo que se estaría fomentando el uso del bus dentro de la ciudad de Las Palmas para aquellas personas que residen en otros municipios. Con respecto a la tarjeta cliente, supone importantes ahorros para aquellos usuarios que realicen un trayecto concreto con cierta frecuencia, puesto que a mayor número de viajes, mayor es el ahorro. Igualmente, los mayores ahorros se dan para la combinación de trayectos interurbano y urbano.

3.3 La demanda de servicios de transporte

La investigación se ha centrado en la isla de Gran Canaria, concretamente en los municipios de Arucas, Telde y Las Palmas de Gran Canaria, esta última capital de la isla. La superficie de la isla es de 1.560 kilómetros cuadrados, con una población⁵⁷ de 771.333 habitantes y una densidad de aproximadamente 494 habitantes por kilómetro cuadrado para el total de la isla; de 997 para Arucas; 890 para Telde y 3.686 para Las Palmas de Gran Canaria. Se analiza la demanda de transporte en los corredores Norte y Sur (ver Figura 3.2). El primero está definido entre Arucas y Las Palmas de Gran Canaria, mientras que el corredor Sur es entre Telde y Las Palmas de Gran Canaria. En estos tres municipios colindantes reside el 64% de la población residente de la isla. Con respecto al parque de vehículos, éste ha experimentado un crecimiento anual importante, con una media anual del 6% en el periodo⁵⁸ 1997-2001, siendo la tasa de motorización de 624 vehículos por cada mil habitantes, como ya hemos comentado.

⁵⁷ Fuente: Instituto Canario de Estadística, dato referido al año 2002.

⁵⁸ Fuente: Instituto Canario de Estadística, datos anuales a fecha de 31 de diciembre.



Figura 3.2: Gran Canaria-corredores

Según la encuesta de movilidad realizada en 1994 por IPD Española, en un día laboral medio, el número de viajeros de Telde a Las Palmas es de 26.462, siendo el reparto modal de un 70,25% en transporte privado, 22,22% en transporte público y 7,53% en otros modos; desde Arucas a Las Palmas el número de viajeros es de 9.937; un 70,52% viaja en transporte privado, 26,6% en transporte público y 2,88% en otros modos de transporte. Como las cifras muestran, existe una alta preferencia por el transporte privado y el porcentaje de usuarios del transporte público no distingue entre usuarios cautivos y no cautivos.

3.4 La encuesta de preferencias reveladas

3.4.1 Personas a entrevistar

Con objeto de conocer las preferencias actuales de los viajeros se realizó una encuesta de preferencias reveladas (PR) dirigida a aquellas personas que viajaran a Las Palmas de Gran Canaria, en alguno de los dos corredores objeto de estudio, al menos una vez

por semana por cualquier motivo de viaje. Las personas debían ser mayores de 18 años, porque es a partir de esta edad cuando se puede disponer de carné de conducir y por tanto, de coche como conductor como alternativa disponible.

Para determinar cuántas encuestas realizar en cada corredor se consideró el número total de viajeros en ambos corredores de acuerdo con el Plan de Coordinación de la Oferta en el área metropolitana de Las Palmas de Gran Canaria de 1998 (Edei Consultores, 1998). El total de viajeros es de 36.399; 9.937 en el corredor Norte y 26.462 en el corredor Sur. En base a estos datos, se decidió que el número de encuestas a realizar fuera de 250 en el corredor Norte (Aruca) y 655 en el corredor Sur (Telde), lo que supone un total de 905 encuestas.

3.4.2 Diseño del cuestionario

Para la estimación de modelos de demanda desagregados es preciso obtener información de los atributos de todas las alternativas que tiene disponibles el individuo así como de sus principales características socioeconómicas. Este hecho condiciona muchas de las preguntas en la encuesta a realizar. En este trabajo se tomaron como referencia otros estudios con un objetivo similar (Cherchi, 2000).

Dentro del cuestionario se pueden distinguir tres bloques. El primero corresponde a la identificación del hogar. Esta información puede ser facilitada por cualquier persona perteneciente al mismo y, concretamente, se resume en las siguientes variables:

- ✓ Dirección del hogar
- ✓ Personas que viven en el hogar
- ✓ Personas que poseen carnet de conducir
- ✓ Personas que estudian

- ✓ Personas que trabajan
- ✓ Ingreso familiar neto mensual
- ✓ Sexo
- ✓ Edad (preguntando por año de nacimiento)
- ✓ Relación familiar
- ✓ Ocupación

El segundo bloque del cuestionario hace referencia a algún viaje a Las Palmas realizado por el entrevistado. Igualmente, se distinguen dos partes. Una primera en la que se realiza una descripción del viaje y otra en la que se especifican las distintas etapas así como sus atributos. Los elementos relevantes son:

- ✓ Motivo del viaje
- ✓ Frecuencia del viaje
- ✓ Hora de salida y de llegada
- ✓ Origen del viaje
- ✓ Destino del viaje
- ✓ Modo de transporte utilizado
- ✓ Modos de transporte disponibles
- ✓ Distancia de caminata recorrida en origen y destino
- ✓ Tiempo de viaje
- ✓ Tiempo de aparcamiento (si viaja en coche)
- ✓ Lugar de estacionamiento (si viaja en coche)
- ✓ Coste del aparcamiento (si viaja en coche)
- ✓ Tiempo de espera (si viaja en bus)
- ✓ Línea utilizada (si viaja en bus)

- ✓ Tipo de billete utilizado (si viaja en bus)

Y como tercer bloque, la caracterización socioeconómica del entrevistado:

- ✓ Nivel de estudios
- ✓ Nivel de estudios del cabeza de familia (si el entrevistado no es el cabeza de familia)
- ✓ Posesión de carné de conducir
- ✓ Posesión de vehículo
- ✓ Tipo de vehículo
- ✓ Jornada laboral
- ✓ Ingreso neto mensual

En cuanto al tipo de formato del cuestionario, existen varias opciones como puede verse en Richardson *et al*, (1995). En un principio, planteamos realizar las encuestas por motivo de trabajo en el lugar de trabajo, pues esto permite reducir bastante los costes de recolección y en el hogar, las encuestas por cualquier otro motivo. Las encuestas en el trabajo deberían ser autocumplimentadas, pues a pesar de disponer de la autorización de la dirección de las empresas contactadas y el apoyo de los departamentos de recursos humanos, no se nos permitía entrevistar personalmente durante la jornada laboral. Se realizaron varias pruebas piloto para diseñar un cuestionario fácil de cumplimentar y detectar problemas con el mismo. Finalmente se logró diseñar un cuestionario manejable por el entrevistado, pero la tasa de respuesta fue bastante baja (apenas un 10%), por lo que se optó por realizar las encuestas por motivo de trabajo también en el hogar.

Así, todas las encuestas se realizaron mediante entrevista personal en el hogar. Para la selección de los hogares se siguieron rutas aleatorias definiendo cuotas por sexo (50% hombres y 50% mujeres). En la primera visita se identificaba si en el hogar había personas que realizaran viajes con una frecuencia de al menos una vez a la semana por cualquier motivo. Sólo se consideraban los viajes de las personas mayores de 18 años porque a partir de esta edad tiene la capacidad legal para tener carné de conducir y por tanto, tener disponible como alternativa de transporte el coche como conductor. Excepcionalmente, se podían considerar aquellos viajes que realizasen menores de edad (pero mayores de 14 años) que viajaran en coche como acompañante y pudieran realizarlo en bus o viceversa. Si el individuo no se encontraba en el hogar en ese momento, se concertaba una segunda visita. En el Anexo I se presenta el cuestionario empleado en la encuesta de PR.

Se realizó un pequeño pretest para analizar si la forma de preguntar era clara para los entrevistados y determinar aquellos aspectos que no hubieran sido considerados o fuera necesario modificar. Así se observó que era mejor preguntar por los motivos de no disponibilidad de diversos modos de transporte; esto permite al investigador analizar si realmente existe disponibilidad o no de cada uno de los modos. Además, se crearon tarjetas que el entrevistador presentaba al individuo en determinadas preguntas para facilitar su respuesta y reducir el tiempo total de la encuesta. Las tarjetas se realizaron para el motivo del viaje, la frecuencia del viaje, los medios de transporte, la ocupación y el nivel de estudios (ver Anexo I).

En esta encuesta sólo se preguntaba por el viaje realizado en uno de los dos corredores con una frecuencia determinada⁵⁹. Podría haberse solicitado información del diario de viajes de los individuos del hogar a entrevistar y a partir de ahí obtener la información correspondiente del corredor. El problema que presenta este formato de solicitud de información es que obtenemos información de viajes que no son relevantes para nuestro estudio y puede suceder también que tras recolectar una alta cantidad de información, con el coste que esto supone, finalmente no dispongamos de una muestra de un tamaño razonable (ver por ejemplo, Cherchi, 2000). Sin embargo, hemos de decir que tal vez hubiera sido interesante recopilar el diario de viajes de aquellos individuos que si viajaran al menos una vez por semana a la ciudad de Las Palmas de Gran Canaria, con el objetivo de poder analizar el peso que tiene el viaje en el corredor sobre el total de viajes que realiza el individuo. Esto nos hubiera permitido considerar en la modelización de la elección modal aspectos como la proporción gastada en transporte, la frecuencia del viaje, etc. Este es un tema que se debe estudiar en profundidad en futuras investigaciones. Para poder ponderar el gastado en transporte del viaje a modelizar necesitamos conocer el total de viajes que el individuo realiza en la semana, ya que el viaje relevante en nuestro trabajo es aquel que se realiza al menos una vez por semana. No obstante, plantear un diario semanal para que sea cumplimentado bien por el entrevistador o el individuo encuestado no resulta una tarea fácil.

3.4.3 Análisis descriptivo de los resultados de la encuesta

Se realizó un total de 950 encuestas durante el último trimestre de 1999. Durante el periodo de recolección de datos se supervisó el 30% de las encuestas realizadas para

⁵⁹ Al menos una vez por semana.

verificar si éstas eran correctas y el viaje descrito real. En esta supervisión se detectaron algunas encuestas que presentaban datos incoherentes, por lo que fue necesario realizar una segunda supervisión en la que fueron anuladas 28 encuestas de las 950 realizadas porque no se habían efectuado realmente (93% sobre el total de encuestas anuladas) o no cumplían las condiciones para ser consideradas en el estudio (7% sobre el total de encuestas anuladas).

Considerando las 922 encuestas válidas (266 del corredor Norte y 656 del corredor Sur), se obtuvo que por modos de transporte un 51,95% de los viajes se realizaba en coche como conductor⁶⁰, un 10,95% en coche como acompañante⁶¹ y un 37,1% en bus⁶². Por tanto, la partición modal entre transporte privado y público se sitúa, aproximadamente, en un 63% y un 37%, respectivamente. Estos datos reflejan la elevada preferencia por el vehículo privado existente en ambos corredores.

Estudiando la disponibilidad de los individuos entrevistados, esto es, qué modos de transporte tienen disponibles para realizar su viaje del total de modos existentes, se encontraron 212 individuos cautivos, esto es, un 23% del total; 15 en coche conductor, seis en coche acompañante y 191 en bus; esto es, un 3,13% de los usuarios de coche conductor son cautivos de este modo de transporte; un 6% del coche como acompañante y un 55,84% del bus. Como puede observarse, el porcentaje de cautivos en bus es bastante elevado, ya que más de un 50% de los usuarios entrevistados no dispone de otro modo de transporte para realizar su viaje. Estas cifras evidencian el hecho de que el

⁶⁰ Aquí se incluyó a los que viajan en moto porque es un modo no disponible para el resto de entrevistados y antes de eliminarlos preferimos probar si era factible considerarlos como usuarios de coche conductor.

⁶¹ Incluye a los individuos que viajan en coche acompañante y bus, porque la mayor parte de su viaje lo realizaban en coche acompañante.

⁶² Incluye a los individuos que viajan en bus y taxi, considerando la parte del viaje en taxi como una etapa más en bus.

transporte público no siempre constituye una alternativa al transporte privado y este último se consolida como la primera opción de transporte para la gran mayoría de los entrevistados siempre y cuando lo tengan disponible.

El número de individuos con posibilidad de elección entre distintos modos para realizar el viaje, es de 710 (205 del corredor Norte y 505 del corredor Sur). De éstos, un 65,35% elige coche como conductor; un 13,38% coche como acompañante y un 21,27%, bus. Al eliminar a los individuos cautivos, el transporte privado incrementa su cuota de mercado a un 78,73% del total de viajes.

En cuanto a la clasificación por motivo de viaje, tenemos, para los individuos con elección, que un 55,07% realiza viajes recurrentes (47,89% por motivo de trabajo y 7,18% por estudios). Los viajes por motivos compras y ocio suponen aproximadamente un 17% y un 15% del total de viajes, respectivamente (ver Tabla 3.5).

Tabla 3.5: Clasificación por motivo de viaje

Motivo del Viaje	Porcentaje	
	Sobre el total	Sobre no cautivos
Trabajo	44,00	47,89
Estudios	9,51	7,18
Compras	16,97	16,76
Dejar/recoger a alguien	1,41	1,70
Visita a un familiar-amigo/a	5,95	6,05
Deporte	0,65	0,70
Visita médica	4,43	4,37
Ocio	16,97	15,35

Un 45,77% de los viajes se realizan cinco veces por semana y un 30%, aproximadamente, una vez a la semana. Los viajes que se realizan con una frecuencia menor a cinco a la semana suponen un total del 45% (aproximadamente) del total de

viajes; mientras que los viajes de más de cinco veces a la semana, un 9% (ver Tabla 3.6).

Por sexo, un 46,2% de los individuos con posibilidad de elección son mujeres y un 53,8%, hombres. En cambio, considerando el total de individuos entrevistados (incluye a los individuos cautivos), tendríamos un 50% de hombres y de mujeres.

Tabla 3.6: Clasificación por frecuencia de viaje

Frecuencia (n° de veces por semana)	Porcentaje	
	Sobre el total	Sobre no cautivos
1	29,98	29,44
2	10,39	9,72
3	4,65	4,37
4	1,41	1,41
5	41,56	45,77
6	4,65	5,35
7	5,09	2,39
8	0,11	0,14
10	0,76	0,99
11	0,11	0,14
15	0,22	0,28

3.5 La encuesta de preferencias declaradas

3.5.1 Personas a entrevistar

Con el fin de analizar el efecto producido por los cambios introducidos en el sistema de transporte se realizó una encuesta de preferencias declaradas (PD) dirigida a aquellas personas que pudieran ser afectadas por la introducción de las nuevas tarifas. El mismo diseño experimental se aprovechó, además, para estudiar el efecto de variables latentes (la comodidad, en concreto) sobre las decisiones de los viajeros. Se planteó un diseño adaptado a la situación de cada individuo tomando como base la información de PR. Del total de encuestas de PR (922 individuos), se seleccionó para la entrevista de PD a

aquellas personas que declaraban viajar en coche como conductor y tuvieran disponible la alternativa bus (456⁶³ individuos).

3.5.2 Diseño del experimento

De las tres formas usuales de presentar un experimento de PD (Ortúzar y Willumsen, 2001) se planteó un juego de elección entre coche como conductor y transporte público (bus); esto es, se presentaron distintas opciones de elección hipotéticas para que, en cada caso, los individuos contestaran qué alternativa elegirían. Se optó por un experimento de este tipo porque es más sencillo para el entrevistado, siendo la calidad de los datos que se obtienen similar a la de los otros experimentos (Ortúzar y Garrido, 2000).

Para el conjunto de atributos de cada alternativa, se planteó un grupo de discusión compuesto por ocho personas (cuatro hombres y cuatro mujeres) que viajaban en los corredores objeto de estudio y en alguno de los modos de transporte considerados (coche conductor y bus). La reunión duró unas dos horas y durante el desarrollo de la misma se trataron diferentes aspectos del sistema de transporte; en un primer momento estos aspectos fueron introducidos por la psicóloga que dirigía el grupo focal, pero posteriormente surgían de la propia iniciativa de los individuos. De este grupo de discusión se obtuvieron como variables relevantes el tiempo de viaje, el coste del viaje, el problema del aparcamiento para la alternativa coche (bien sea en tiempo que dedica a buscar o bien en la necesidad de pagar estacionamiento), la frecuencia para la alternativa bus, la comodidad y la puntualidad del servicio público. Para cada alternativa se definieron cuatro variables; en el caso del coche como conductor se incluyó el tiempo de viaje, el coste del viaje, el tiempo o coste de aparcamiento y la comodidad, mientras que

⁶³ Este número se reduce a 407 personas porque se eliminan algunos individuos que por su motivo de viaje o trabajo eran, de hecho, cautivos del coche.

para el bus se consideró el tiempo de viaje, el coste del viaje, la frecuencia y la comodidad. En cada alternativa, se consideraron tres variables genéricas expresadas en diferencias para reducir el número de niveles, y dos específicas, el tiempo o coste de aparcamiento para el coche, y la frecuencia para el bus.

Se realizaron distintos diseños, que fueron modificados a la vista de los resultados de las distintas encuestas piloto. En todos los casos hemos trabajado con un diseño factorial fraccional, que reduce el número de escenarios a sólo 27. Este diseño permite medir, además de los efectos principales, las interacciones de segundo orden entre tres variables: tiempo, coste y frecuencia (Kocur *et al*, 1982). Como no es recomendable presentar las 27 opciones de elección a un mismo individuo, se realizó un diseño formado por tres bloques de nueve opciones cada uno; a cada subgrupo de la muestra se le presentó uno de los tres bloques; la definición de los bloques se hizo de manera aleatoria, así como la presentación de las distintas opciones de elección a cada uno de los individuos.

Los distintos diseños utilizados se adaptan a la situación de cada entrevistado tomando como nivel base el tiempo de viaje, el tiempo y el coste de aparcamiento, el recorrido realizado (para definir la tarifa correspondiente en bus), el gasto en combustible y el tipo de vehículo que cada individuo declaró en la encuesta previa de PR. La transformación del diseño general al caso particular de cada entrevistado se llevó a cabo a partir de un programa en MATLAB (Hunt *et al.*, 2001).

Para poder definir el compromiso entre modos de transporte es necesario establecer una diferencia mínima absoluta tanto para el caso de variables genéricas (por ejemplo, el tiempo de viaje) como para los distintos niveles de una misma variable (por ejemplo, la frecuencia). Si la diferencia mínima absoluta para el tiempo de viaje es de 10 minutos,

esto quiere decir que la diferencia mínima entre el tiempo de viaje del coche y del bus debe ser al menos de 10 minutos. En el caso de la frecuencia, una diferencia mínima absoluta de cuatro minutos, significa que el tiempo mínimo entre dos buses debe ser de cuatro minutos.

En la Tabla 3.7 se muestra la definición general de niveles utilizada en el diseño definitivo, donde T es el tiempo de viaje declarado por el individuo en la encuesta de PR; C es el gasto en combustible en función de la distancia recorrida por el individuo multiplicada por dos para añadir parte del coste de mantenimiento del vehículo; F es la frecuencia del bus y CA es el coste de aparcamiento del individuo. La tarifa actual es el precio del billete del bus sin descuento de ningún tipo y el bono es el coste del bus cuando se aplican los descuentos definidos en la tarjeta insular, explicada anteriormente.

La definición de los bloques se realizó de manera aleatoria y teniendo en cuenta que en cada bloque hubiera escenarios con los tres niveles de tiempo de viaje y con los tres niveles de comodidad. De esta forma se evitaba que existieran bloques con un elevado número⁶⁴ de opciones en que la diferencia entre el tiempo de viaje entre coche conductor y bus fuera, por ejemplo, siempre positiva. La restricción definida a la asignación aleatoria daba lugar a que cada bloque tuviera, por ejemplo, opciones en que la diferencia entre el tiempo de viaje en coche conductor y bus fuera positiva, igual o negativa, esto es, nivel cero, uno y dos, respectivamente, como puede verse en la Tabla 3.7 y Tabla 3.10.

⁶⁴ El máximo era nueve.

Tabla 3.7: Definición de niveles

VARIABLES	Niveles	Coche	Bus	% de Variación	Diferencia Mínima
Tiempo	0	T	1,25*T	-25%	10
	1	T	T	0%	-
	2	1,25*T	T	+25%	10
Coste	0	1,15*C	Tarifa actual	-	200 pts
	1	1,15*C	Bono	-	200 pts
	2	C	Bono	-	200 pts
Frecuencia	0	-	F	0%	-
	1	-	0,75*F	-25%	4
	2	-	0,50*F	-50%	4
C. de Ap.	0	CA	-	-	-
	1	1,5*CA	-	-	-
Comodidad	0	Buena	Mala	-	-
	1	Buena	Estándar	-	-
	2	Buena	Buena	-	-

En el primero de los diseños analizados se consideró el tiempo, el coste y la frecuencia a tres niveles, y el tiempo de aparcamiento y la comodidad a dos, resultando un diseño factorial fraccionado de 27 opciones. La definición cualitativa y cuantitativa de los niveles en este primer diseño se presenta en la Tabla 3.8; cuando se trata de variables genéricas (por ejemplo, tiempo y coste), se presenta el valor de la variable para el coche menos la del bus, la comodidad (también genérica) se presenta siendo la del coche fija y el nivel más alto y la del bus es definida en relación a ésta, es decir, cómo es la comodidad del bus en relación a la del coche. Cuando se trata de variables específicas (tiempo de aparcamiento para el coche y frecuencia para el bus), se toma como nivel base el valor que el individuo reveló en la encuesta de PR. El signo “<” significa que el valor de la variable del coche es menor que el del bus, “<<” significa que es mucho menor, “>” que es mayor y “=” que es igual. Así mismo, se especifica en porcentaje la variación en términos cuantitativos. Por ejemplo, un valor de “-50%” significa que el valor del atributo del coche es un 50% menor que el del bus.

En la Tabla 3.8 se puede observar que el tiempo de viaje es 50% menor, 25% menor o igual en el coche que en el bus; por otro lado, el coste en coche siempre es mayor que en bus (50%, 25% o 10%). El intervalo de tiempo entre dos buses consecutivos, es distinto para cada individuo y varía entre 50% mayor, igual o 50% menor que el actual. El tiempo de aparcamiento es el que el entrevistado declaró y no varía, o aumenta un 50%.

Tabla 3.8: Primer diseño

Diferencia de Coche y Bus		Bus	Coche	Bus	
Niveles	Tiempo	Coste	Frecuencia	Tiempo de Aparcamiento	Comodidad
0	<<	>>>	>	=	<
	-50%	+50%	+50%	0%	
1	<	>>	=	>	=
	-25%	+25%	0%	+50%	
2	=	>	<	-	
	0%	+10%	-50%		

Finalmente, la comodidad está referida a la del coche que se puede considerar siempre mejor que la del bus y está en función de las características del habitáculo del vehículo. Así, se definió una comodidad en el bus que es peor que la del coche y una segunda, que permite considerar la comodidad del bus “igual” a la del coche. De esta forma, para el coche se tiene:

Comodidad estándar: Sería la que experimenta cuando va en su coche, es decir, viajar de forma independiente, sin necesidad de esperar para iniciar el viaje o de caminar hasta una parada de bus, pero con la posibilidad de soportar aglomeraciones de tráfico que le pueden producir cierta molestia o tensión (en las tarjetas este nivel fue presentado como: “Estándar”).

Para el bus se definió:

Comodidad Baja: Cuando viaja en bus que va casi lleno, donde se puede encontrar con situaciones no gratas como por ejemplo contactos físicos, empujones, voces elevadas, olores desagradables, etc (en las tarjetas fue presentado como: “*Baja*”).

Comodidad Alta: Cuando viaja en bus sentado cómodamente con música de fondo agradable, y puede ir realizando alguna actividad, como leer, sin que las posibles aglomeraciones de tráfico le produzcan molestia o tensión (en las tarjetas fue presentado como: “*Alta*”).

Con este primer diseño se realizaron treinta y dos encuestas, de las cuales diecisiete fueron contestadas eligiendo la alternativa coche en las nueve opciones de elección presentadas, en tres casos se eligió la alternativa bus en todas las situaciones planteadas, y solo doce individuos eligieron alternadamente entre las dos alternativas posibles en cada juego de elección. Dado que un 53% de entrevistados eligieron siempre la alternativa que estaban usando actualmente, esto es, el coche, se dedujo que posiblemente el experimento planteado no permitía captar la variación en la elección modal, es decir, la determinación de aquellos valores de los atributos de las dos alternativas que harían al individuo variar sus preferencias por los dos modos de transporte. Con los datos obtenidos se estimaron modelos logit simple con el fin de obtener una primera aproximación al valor de los distintos parámetros, resultando poco significativa la variable comodidad, y la constante modal con signo intuitivamente incorrecto.

Ante estos dos problemas, alto porcentaje de cautivos del coche y parámetros poco significativos, decidimos modificar el diseño presentando una alternativa de bus

mejorada. Para ello, se cambiaron los niveles del tiempo de viaje de manera que el del bus pudiera ser menor que el del coche; se incluyó el tiempo de aparcamiento en el tiempo del viaje y se consideró el coste de aparcamiento como nueva variable para el coche. También se modificaron los niveles de la frecuencia, definiendo valores iguales o menores que la frecuencia actual y se matizó la definición de la variable comodidad en las respectivas tarjetas (ver Tabla 3.9).

Tabla 3.9: Segundo diseño

Diferencia de Coche y Bus			Bus	Coche	Bus
Niveles	Tiempo	Coste	Frecuencia	C. Aparcamiento	Comodidad
0	<	>>	=	=	<
	-25%	Mínimo	0%	0%	
1	=	>>>	<	>	=
	0%	Mínimo	-25%	+50%	
2	>	>	<<	-	-
	+25%	Mínimo	-50%		

En este segundo diseño se tiene un tiempo de viaje en coche un 25% menor, igual o un 25% mayor que el tiempo de viaje en bus. Se ampliaron las diferencias entre los costes de ambas alternativas, definiendo un umbral mínimo entre ambos costes. El intervalo entre buses es el actual, 25% menor o 50% menor, fijando la diferencia mínima entre opciones en cinco minutos.

El coste de aparcamiento tiene un tratamiento especial, pues se dan dos situaciones. En primer lugar, los individuos que declaran no pagarlo, ya sea porque estacionan en la vía pública o porque aparcan en centros comerciales que no tienen definido un precio por estacionar. En el primer caso, para definir el valor se examinó el precio que se pagaba en otras zonas por aparcar en la calle (“zona azul”); en el segundo caso, se consideró el precio por hora de un estacionamiento privado. Por otro lado, existen individuos con

estacionamiento propio, que pagan mensualmente una cierta cantidad; para ellos se definió un coste de aparcamiento diario en función del número de veces que viajaban a la semana. Este valor fue utilizado para calcular el dato del nivel 1, que implica un 50% más del coste. Para facilitar la comprensión, se elaboró una tabla de equivalencias expresando en pago mensual lo que suponía cada pago diario.

La definición de comodidad se mantuvo, pero se consideraron los siguientes cambios:

En la tarjeta: “estándar” por “la que usted experimenta cuando viaja en su coche”.

En la tarjeta: “baja” por “guagua casi llena”.

En la tarjeta: “alta” por “sentado cómodo y tranquilamente en la guagua”.

También se hizo necesario definir un contexto en el que se dieran las condiciones necesarias para que las nuevas situaciones hipotéticas parecieran realistas, como es el caso de un tiempo de viaje menor en bus que en coche. Así, en relación a las mejoras en el servicio de transporte se consideraron las siguientes especificaciones (ver Anexo I):

- ✓ *Habrá un carril sólo bus desde La Laja hasta la estación de Guaguas de Las Palmas. Por este carril sólo podrán circular los vehículos de transporte regular de pasajeros y se velará por el estricto cumplimiento de esta norma.*
- ✓ *Una vez que la guagua haya entrado en la ciudad, tendrá preferencia en los semáforos (los semáforos se pondrán en verde cuando la guagua se aproxime).*
- ✓ *Habrá servicio de guaguas desde las 5.00 horas hasta las 24.00 horas.*
- ✓ *Los horarios de las tres empresas de guaguas se coordinarán para facilitar el trasbordo en el menor tiempo posible.*
- ✓ *Con un mismo bono se podrá viajar en las tres empresas de guaguas, con descuentos que pueden ir desde el 30 al 70%.*

En relación a los aparcamientos, se consideró:

- ✓ *La zona azul se irá ampliando hasta prácticamente alcanzar la mayoría del aparcamiento en la calle.*
- ✓ *Los estacionamientos privados que ahora son gratuitos acabarán cobrando por el aparcamiento.*

Otras consideraciones:

- ✓ *El precio del combustible podría seguir aumentando, lo que supone un mayor coste del coche.*
- ✓ *La existencia de carriles sólo bus probablemente supondrá un aumento de los tiempos de viaje en coche.*

Con estos cambios, se hizo una nueva encuesta piloto pero desgraciadamente la tasa de respuesta fue muy baja, sólo once encuestas fueron contestadas; de éstas, dos no fueron consideradas buenas porque eligieron la alternativa bus en todas las opciones. Esto puede haber sucedido ya sea porque el entrevistado contestó lo que creía que el entrevistador esperaba (sesgo de afirmación) o porque eligió la alternativa de menor coste (es decir, eligió de manera lexicográfica, cuestión que trataremos más adelante).

A pesar de las mejoras sustanciales en el diseño, los resultados de las estimaciones no resultaron muy esperanzadores: el signo del parámetro de la comodidad no era correcto y los parámetros presentaron bajos niveles de significación (aunque esto puede deberse al pequeño tamaño muestral). El problema del alto porcentaje de cautivos del coche pareció resolverse pero, aparentemente, la variable comodidad no fue percibida por los entrevistados. Por este motivo, nos centramos en la definición de esta variable y procedimos a contactar telefónicamente con los entrevistados de esta segunda encuesta

para que nos explicaran que era para ellos la comodidad en el bus. De esta manera detectamos que, efectivamente, nuestra definición de comodidad *peor* en el bus que en el coche, no era lo realmente percibido por los entrevistados. Contrastamos con las empresas de transporte si era posible que se dieran las situaciones que los entrevistados comentaban como por ejemplo, la posibilidad de viajar de pie, pues se trata de viajes interurbanos, siendo posible en un número no superior a doce personas.

Por esto se decidió aumentar los niveles de la variable comodidad de dos a tres; con ello el número total de opciones del nuevo diseño (el tercero) fue de 162, que se redujeron a 27 (como en los anteriores diseños) al utilizar el diseño factorial fraccional. En este tercer diseño, el resto de variables no se modificaron (ver Tabla 3.10).

Tabla 3.10: Tercer diseño

Diferencia de Coche y Bus			Bus	Coche	Bus
Niveles	Tiempo	Coste	Frecuencia	C. Aparcamiento	Comodidad
0	<	>>	=	=	<<
	-25%	Mínimo	0%	0%	
1	=	>>>	<	>	<
	0%	Mínimo	-25%	+50%	
2	>	>	<<	-	=
	+25%	Mínimo	-50%		

Así, la comodidad presenta tres niveles, un nivel bajo o malo, que implicaría viajar de pie, otro medio o estándar, que considera viajar sentado pero sin elegir el asiento y un valor alto o bueno, que implica viajar sentado eligiendo donde sentarse. Estas nuevas definiciones fueron presentadas en la encuesta de la siguiente manera:

Para el bus:

Comodidad Baja: Cuando viaja en guagua que va muy llena y debe viajar de pie; además, a veces se encuentra con situaciones no gratas como por ejemplo contactos físicos, empujones, voces elevadas, olores desagradables, etc. (en las tarjetas este nivel se presentó como: *“Guagua llena, viajando de pie”*).

Comodidad Media o Estándar: Cuando viaja en guagua que va casi llena y puede ir sentado, pero se puede encontrar con situaciones no gratas como contactos físicos, empujones, voces elevadas, olores desagradables, etc (en las tarjetas fue presentado como: *“Guagua casi llena, con espacio para viajar sentado, pero sin posibilidad de elegir donde sentarse”*).

Comodidad Alta: Cuando viaja en guagua sentado cómodamente y con música de fondo agradable; además, puede ir realizando alguna actividad, como leer, sin que las posibles aglomeraciones de tráfico le produzcan molestia o tensión (en las tarjetas fue presentado como: *“guagua con espacio para viajar sentado cómoda y tranquilamente, pudiendo elegir donde sentarse”*).

Para el coche:

Comodidad Base-Estándar: Sería *“la que experimenta cuando viaja en su coche”*, es decir, viajar en coche de forma independiente, sin necesidad de esperar para iniciar el viaje o de caminar hasta una parada de guaguas, pero con la posibilidad de soportar aglomeraciones de tráfico que pueden producir molestia o tensión (en las tarjetas fue presentado como: *“la que experimenta cuando viaje en su coche”*).

Con este tercer diseño se llevó a cabo un nuevo pretest, realizando un total de diecinueve encuestas de las cuales cinco eligieron siempre coche (cautivos en coche) y

siete eligieron siempre bus (cautivos en bus o lexicográficos en coste). Al igual que con las encuestas piloto anteriores, estimamos modelos logit para ver qué resultados se obtenían. Todos los parámetros presentaban los signos esperados; destacando la poca significación de la variable frecuencia, pero hay que señalar que la muestra es pequeña. Comparando resultados, finalmente consideramos que se podía realizar la encuesta definitiva utilizando este tercer diseño.

Para la realización de la encuesta de PD se instruyó debidamente a los entrevistadores, ya que este tipo de encuestas son más complejas tanto para el individuo como para el entrevistador. El primer paso a seguir era contactar al entrevistado para realizar una nueva entrevista y tras establecer una cita realizar la encuesta. En primer lugar, se le recordaba el viaje que había declarado en la encuesta previa de PR para que situara y recordara el viaje y sus características (tiempo, frecuencia, etc.). Una vez constatado que el entrevistado recordaba el viaje se le presentaba la hoja de contexto de la encuesta (ver Anexo I) que era leída por el entrevistador haciendo hincapié en aquellos aspectos fuera necesario. Posteriormente, se explicaba el significado de las distintas variables consideradas en el experimento y finalmente, se planteaban las distintas situaciones de elección, con el objeto de recoger las preferencias del individuo (ver Anexo I).

3.5.3 Modelización con los resultados de las encuestas piloto

Con los datos obtenidos en las distintas encuestas piloto se estimaron modelos logit binomial considerando funciones de utilidad lineales en los parámetros y en las variables. También se estimaron modelos combinando datos procedentes de las distintas encuestas piloto siguiendo el procedimiento de estimación con datos mixtos (Bradley y Daly, 1997).

Todas las variables de los modelos son genéricas, salvo tiempo y coste de aparcamiento y frecuencia que sólo aparecen especificadas en el modo correspondiente. Para la estimación de los modelos sólo se consideraron las respuestas de individuos no cautivos⁶⁵.

Las funciones de utilidad especificadas para cada modo y modelo se presentan en la Tabla 3.11.

Tabla 3.11: Funciones de utilidad estimadas- encuestas piloto

Modelos	Utilidad del coche	Utilidad del bus
Modelo 1	$V_{coche} = q_{cmc} + q_{tv}tv + q_c c + q_{tap}tap + q_{com}com$	$V_{bus} = q_{tv}tv + q_c c + q_f f + q_{com}com$
Modelo 2	$V_{coche} = q_{cmc} + q_{tv}tv + q_c c + q_{cap}cap + q_{com}com$	$V_{bus} = q_{tv}tv + q_c c + q_f f + q_{com}com$
Modelo 3	$V_{coche} = q_{cmc} + q_{tv}tv + q_c c + q_{com}com$	$V_{bus} = q_{tv}tv + q_c c + q_f f + q_{com}com$
Modelo 4		
Modelo 5	$V_{coche} = q_{cmc} + q_{tv}tv + q_c c + q_{cap}cap$	$V_{bus} = q_{tv}tv + q_c c + q_f f + q_{cB}cB + q_{cE}cE$
Modelo 6	$V_{coche} = q_{cmc} + q_{tv}tv + q_c c$	$V_{bus} = q_{tv}tv + q_c c + q_f f + q_{cB}cB + q_{cE}cE$

Las estructuras especificadas para la estimación según la metodología de datos mixtos se muestran en la Figura 3.3.

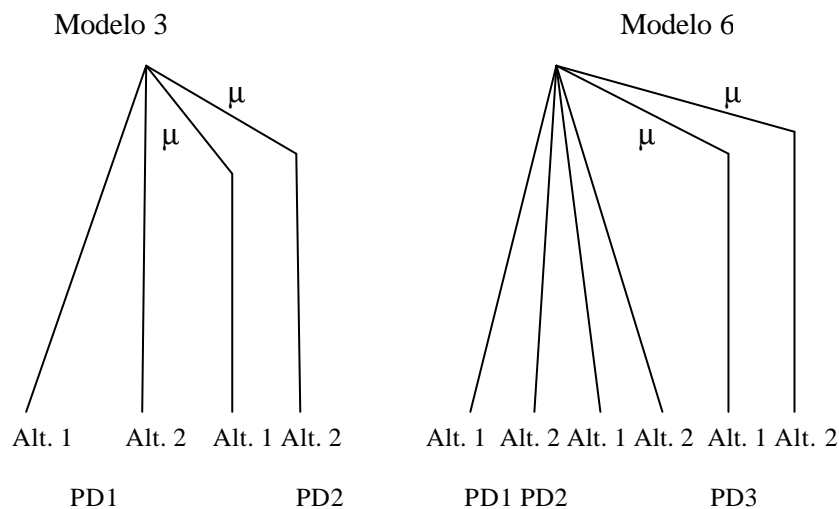


Figura 3.3: Estructura jerárquica para estimación mixta

⁶⁵ Análisis realizado en la sección posterior.

Los modelos estimados que se presentan en la Tabla 3.12 son:

- Modelo 1: Estimación con los datos la primera encuesta piloto.
- Modelo 2: Estimación con los datos la segunda encuesta piloto.
- Modelo 3: Estimación con los datos de ambas encuestas piloto según la metodología de estimación con datos mixtos (ver Bradley y Daly, 1997), en este caso las alternativas de la segunda encuesta caen del nido artificial (ver Figura 3.3).
- Modelo 4: Estimación con los datos de la primera y segunda encuesta piloto realizando estimación conjunta, mezclando los datos directamente.
- Modelo 5: Estimación con los datos de la tercera encuesta piloto.
- Modelo 6: Estimación con los datos las tres encuestas piloto según la metodología de estimación con datos mixtos (ver Figura 3.3).

Es importante señalar que en los cuatro primeros modelos la variable comodidad se especifica con referencia a la comodidad mala, siendo igual a 1 cuando es buena y 0, en otro caso, por tanto debe resultar con signo positivo. Sin embargo, para los modelos restantes se especificó con relación a la comodidad buena, esperándose signos negativos para ambos parámetros. La variable cB toma valor 1 cuando la comodidad es mala (viajar de pie, sólo datos de la tercera encuesta) y 0, en otro caso; mientras que la variable cE toma valor 1 cuando es estándar (sentado sin posibilidad de elegir en la tercera encuesta, y comodidad mala en las dos primeras) y 0, en otro caso.

Los resultados correspondientes a la estimación de la tercera encuesta piloto son alentadores pues todos los parámetros presentan los signos esperados. Destaca la poca significación de la variable frecuencia (definida como intervalo de tiempo entre buses consecutivos), pero hay que señalar que la muestra es pequeña. Respecto a la combinación de los distintas fuentes de datos, tenemos que según la metodología de estimación con datos mixtos, cuando se combina los datos de la primera y segunda

encuestas resulta un valor del factor de escala PD cercano a uno al 95% de confianza, lo que implica que se puede suponer que ambos datos tienen “igual” varianza y se pueden mezclar directamente (ver Ortúzar y Willumsen, 2001) como en el modelo 4; en este caso las estimaciones mejoran en nivel de significación. Por último, la combinación de las tres fuentes de datos dio lugar a resultados aceptables, todos los signos están de acuerdo a lo esperado (quizás, el de la constante modal) y casi todas las variables son significativas al 95%, excepto el coste y la comodidad estándar.

Tabla 3.12: Resultados de las estimaciones de las encuestas piloto

Parámetros		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
<i>Constante</i>	q_{cmc}	-2,962 ¹ (-3,1) ²	-0,8961 (-0,6)	-0,8246 (-2,2)	-1,391 (-2,7)	0,8465 (0,6)	-1,728 (-3,4)
<i>Tiempo</i>	q_v	-0,1244 (-2,9)	-0,01814 (-0,7)	-0,01492 (-1,8)	-0,03679 (-2,2)	-0,08335 (-2,2)	-0,01944 (-2,0)
<i>Coste</i>	q_c	-0,009148 (-2,0)	-0,004703 (-0,8)	-0,002205 (-2,8)	-0,006047 (-4,1)	-0,006 (-1,6)	-0,0006194 (-0,6)
<i>T. Aparc.</i>	q_{iap}	-0,02962 (-3,1)	-	-	-	-	-
<i>C. Aparc.</i>	q_{cap}	-	-0,003002 (-0,9)	-	-	-0,02119 (-0,5)	-
<i>Frecuencia</i>	q_f	-0,1385 (-4,2)	-0,07466 (-2,9)	-0,05581 (-3,9)	-0,1051 (-5,4)	-0,007923 (-1,7)	-0,07635 (-4,6)
<i>Comod</i>	q_{com}	0,7541 (1,3)	-0,2637 (-0,5)	0,04002 (0,2)	0,1502 (0,4)	-	-
<i>c_Baja</i>	q_{cB}	-	-	-	-	-2,129 (-2,5)	-1,095 (-2,1)
<i>c_Estándar</i>	q_{cE}	-	-	-	-	-0,5692 (-0,7)	-0,3298 (-1,0)
<i>PD</i>	m	-	-	1,008 (14,7)	-	-	0,5516 (7,8)
$r^2(C)$		0,2964	0,1023	0,0389	0,1899	0,2263	0,0339
$l(\hat{q})$		-46,9886	-49,1482	-235,8492	-101,0421	-28,9597	-278,7887
<i>Muestra</i>		99	81	180	180	63	234

¹Valor estimado del parámetro.

²t estadístico.

Se debe destacar que en principio el signo de la constante modal se esperaría positivo (está en la alternativa coche), ya que normalmente debiera existir una preferencia natural por el coche (aunque la comodidad ya está considerada) cuando todas las

variables toman valor cero. No obstante, en varios de los modelos estimados resultó negativa, lo que, a priori, no es un resultado incorrecto⁶⁶.

Las estimaciones con las distintas encuestas piloto se presentan tal y como se realizaron en cada momento del diseño. En base a estas estimaciones se fueron tomando decisiones respecto a las distintas modificaciones del diseño.

3.5.4 Análisis de los resultados de la encuesta

El número final de encuestas PD a realizar se redujo de un total potencial de 456 a sólo 345 individuos por dos razones: la primera es que se revisó detalladamente la disponibilidad y cautividad de los individuos; así se eliminaron aquellos que realmente no disponían del bus como modo de transporte (45 individuos) y a los que habían declarado que su motivo de viaje era recoger/dejar a alguien por tratarse de cautivos del coche conductor (cuatro individuos). La segunda es que tampoco se consideraron las personas entrevistadas en las distintas encuestas piloto (62 individuos). Por otro lado, la tasa de respuesta de la encuesta fue relativamente baja, 97 encuestas contestadas de 345 (28,12%). Esto se debe a que no fue posible localizar a un 21,14% de los individuos a entrevistar a pesar de realizar cuatro intentos de localización, a distintas horas del día por teléfono y mediante visita al hogar; un 21,16% se había trasladado a otra residencia y no fue posible conocer su nueva dirección, y un 29% se detectó que se negó a contestar; finalmente, una encuesta realizada debió ser anulada porque en realidad no había sido realizada por el entrevistado.

⁶⁶ De hecho, estrictamente hablando, las constantes específicas sólo permiten que el modelo reproduzca en forma exacta la proporción de mercado de cada opción (Ortúzar y Willumsen, 2001).

Según la teoría del comportamiento del consumidor, las preferencias deben cumplir una serie de supuestos que nos permiten expresarlas a través de una función de utilidad. Si estos supuestos no se cumplen, no sería factible obtener la función de utilidad a partir de las preferencias. Por esta razón es necesario realizar tests que permitan detectar a aquellos individuos que no utilizan la regla compensatoria al realizar su elección en el juego de PD y por tanto, no declaran una correcta estructura de sus preferencias. Los tests realizados fueron tres: análisis de individuos cautivos (individuos que no consideran el compromiso definido entre las alternativas), análisis de individuos que eligen de manera lexicográfica (aquellos que violan el axioma de continuidad) y análisis de individuos inconsistentes (aquellos que violan el axioma de transitividad).

Los individuos cautivos son aquellos que siempre eligen la misma alternativa. En nuestro caso, serían aquellos que siempre eligen coche como conductor (cautivos del coche) o aquellos que siempre eligen transporte público (cautivos del bus). Esto puede deberse a que el diseño no es capaz de plantear un compromiso adecuado entre los atributos considerados, lo que lleva a seleccionar siempre el mismo modo de transporte, o bien a que el individuo desee influir en las decisiones de política que puedan tomarse como resultado de la encuesta (*sesgo de política*).

Del total de encuestas realizadas (97), resultaron veintitrés individuos cautivos (catorce cautivos del coche como conductor y nueve cautivos del bus), esto es, un 21,65 %. En la práctica, el porcentaje de individuos cautivos en este tipo de encuestas está entre un 15% y un 20% del total de individuos entrevistados (ver por ejemplo, Ortúzar e Iacobelli, 1998; Cherchi y Ortúzar, 2002); nuestro porcentaje es algo mayor. Ahora bien, se debe señalar que los entrevistados son usuarios del coche en el momento de realizar la encuesta, por lo que podemos considerar solamente como cautivos a los que

siempre eligen coche (a pesar de las mejoras del bus seguirían eligiendo el coche como modo de transporte). También se debe añadir que los que siempre eligen bus, son además individuos que eligen de manera lexicográfica⁶⁷ en coste, pues el bus mejorado que se presentaba en el experimento siempre tiene un coste menor que el coche como conductor. Así, sólo consideramos a los cautivos del coche como conductor, el porcentaje es del 14,43%. En cualquier caso, este alto porcentaje de cautivos (21,65%) está reflejando un hecho importante: existe un sector de la población reacio a cambiar de modo de transporte (de coche a bus).

En cuanto al segundo tipo de análisis, se examina a los individuos que siempre eligen la alternativa superior en alguno de los atributos (eligen de manera lexicográfica, ver Saelensminde, 1998a). La detección de individuos lexicográficos, en cualquiera de las variables genéricas en ambos modos de transporte es sencilla; por ejemplo, individuos lexicográficos en tiempo, elegirán siempre la alternativa que presente un menor tiempo de viaje. No obstante, cuando las variables son específicas del modo (por ejemplo, tiempo o coste de aparcamiento para el coche y frecuencia para el bus), el problema de detección no es sencillo. Aún cuando parecería posible proceder de igual forma que en el caso de variables genéricas, no existe claridad al respecto (Saelensminde, 2001).

La elección de manera lexicográfica puede darse cuando los entrevistados a pesar de no estar interesados en participar en el experimento, tampoco se niegan a contestar optando por considerar un solo atributo para simplificar el proceso de elección. También puede aparecer este efecto, aunque de forma espúrea, cuando los valores mostrados en el

⁶⁷ En base a un sólo atributo.

diseño son tales que no representan un verdadero compromiso para ellos (ver la discusión en Saelensminde, 1998b).

El análisis se realizó sobre el total de encuestas realizadas (97 encuestas), llegándose a reconocer un total de dieciocho individuos lexicográficos, siete en tiempo, nueve en coste y dos en comodidad baja. Estos porcentajes son consistentes con los encontrados en otros estudios de esta naturaleza (ver por ejemplo Rizzi y Ortúzar, 2003).

En cuanto al tercer análisis, consiste en examinar las características de aquellos individuos que no contestan de manera consistente. Dadas las distintas opciones de elección se pueden construir reglas lógicas que definan la consistencia en las elecciones realizadas por las personas. Si se presentan más de dos respuestas inconsistentes, se elimina al individuo y si son dos o menos, solo se eliminan las respuestas inconsistentes.

Se detectaron ocho personas inconsistentes (lo que suponen un total de setenta y dos respuestas inconsistentes) más dieciséis respuestas inconsistentes correspondientes a individuos que presentan dos o menos respuestas inconsistentes. Esto hace un total de 88 (10%) respuestas inconsistentes, que se eliminaron de la base de datos para analizar si su inclusión afectaba a las estimaciones del modelo.

Una vez depurados los datos, se realizaron estimaciones con varias bases de datos para comprobar si la consideración o no de individuos cautivos, inconsistentes y lexicográficos afectaba los resultados de las estimaciones. Se cuenta con un total de seis bases de datos distintas: la primera, *todos*, considera las respuestas de todos los

individuos entrevistados, lo que hace un total de 871 observaciones⁶⁸; la segunda base de datos, *consistentes*, elimina a los individuos que son inconsistentes en sus respuestas (se eliminaron ocho personas inconsistentes más dieciséis respuestas inconsistentes), resultando una base con 783 observaciones; la tercera base de datos, *no cautivos*, elimina de la primera a los individuos cautivos (23 individuos), dando como resultado 664 observaciones. En cuarto lugar, se eliminó de la primera base de datos a los individuos lexicográficos; así la cuarta base de datos, *sin lexicográficos*, tiene un total de 621 observaciones. Finalmente, en la quinta base de datos (*sin cautivos e inconsistentes*), se eliminaron tanto a los individuos cautivos como a los inconsistentes (576 observaciones); mientras que la sexta base, *buenos*, se eliminaron además a los individuos lexicográficos (495 observaciones).

Se estudiaron diferentes especificaciones para la función de utilidad del coche⁶⁹ y del bus, tales como análisis de interacciones entre las variables tiempo, coste y frecuencia; inclusión de la renta en la elección, y definición de interacciones entre los atributos y las características socioeconómicas de los individuos. Éstas fueron estimadas con las distintas bases de datos con el objetivo de detectar el efecto sobre los resultados de la inclusión o no de individuos que presentaran preferencias diferentes a las que se pretende modelizar. No se está diciendo que los individuos cautivos, inconsistentes o lexicográficos no tengan definidas sus preferencias respecto a los dos modos de transporte; si no que para poder representarlas a través de una función de utilidad deben verificar los supuestos que establece la teoría del consumidor. Esta función es

⁶⁸ El total de observaciones es 873, 97 entrevistados por nueve respuestas cada uno, pero hay dos observaciones menos porque un entrevistado dejó sin contestar las dos últimas preguntas. Se intentó localizarlo nuevamente sin éxito. No se detectó ninguna inconsistencia en las siete preguntas contestadas, ni se trataba de un individuo cautivo o que eligió de manera lexicográfica, por lo que se consideraron válidas.

⁶⁹ En la encuesta de PD sólo se tiene la alternativa de coche como conductor.

posteriormente estimada en base a la teoría de la utilidad aleatoria (Domencich y McFadden, 1975).

De los distintos modelos estimados, aquí sólo presentamos los más sencillos ya que permiten observar las diferencias en los resultados cuando se estima con las distintas bases de datos. La especificación presentada no define una constante específica para la alternativa coche conductor ya que ésta no resultó significativamente distinta de cero. Por otra parte, desde un punto de vista microeconómico, se podría argumentar que no se justifica especificar una constante de la alternativa ya que esta permite recoger todo aquello que no es percibido por el resto de variables, y como en este caso los entrevistados realizaban su elección basándose en los valores de los atributos del experimento, no sería necesario especificarla.

Concretamente, las funciones de utilidad especificadas fueron las siguientes:

$$\begin{aligned} V_{coche} &= \mathbf{q}_{tv} \cdot tv + \mathbf{q}_c \cdot c + \mathbf{q}_{cap} \cdot cap \\ V_{bus} &= \mathbf{q}_{tv} \cdot tv + \mathbf{q}_c \cdot c + \mathbf{q}_f \cdot f + \mathbf{q}_{cB} \cdot cB + \mathbf{q}_{cE} \cdot cE \end{aligned} \quad (3.1)$$

donde,

tv es el tiempo de viaje expresado en minutos;

c es el coste del viaje expresado en pesetas;

cap es el coste de aparcamiento expresada en pesetas por hora;

f es la frecuencia expresada en buses por hora;

cB toma el valor uno cuando se trata de la comodidad baja;

cE toma el valor uno cuando se trata de comodidad estándar; y

\mathbf{q}_i son los parámetros a estimar.

Se esperan signos negativos para los parámetros del tiempo de viaje, coste de viaje, aparcamiento, y para las distintas comodidades, ya que un aumento de cualquiera de estas variables produce una menor utilidad. Para la frecuencia se espera un signo positivo, ya que está especificada como buses por hora; esto es, un aumento de un bus a la hora genera más utilidad para la alternativa bus.

En la Tabla 3.13 se presentan los resultados para la encuesta definitiva y las seis bases de datos comentadas anteriormente. Se obtiene que los resultados mejoran cuando se van eliminando de la base de datos a los individuos que no verifican los supuestos definidos por la teoría del comportamiento del consumidor. Concretamente, se destacan dos efectos interesantes. En primer lugar, obtenemos que cuando se eliminan los individuos que eligieron de manera inconsistente, el signo de la variable frecuencia es positivo (y por tanto correcto en función de la definición de la variable frecuencia), mientras que cuando se considera todos los entrevistados es negativo y, por tanto, incorrecto. Esto nos indica que es importante identificar a estos individuos, siempre que sea posible, y analizar el efecto sobre las estimaciones. Un comportamiento inconsistente puede significar que el peso que se le da a los distintos atributos sea erróneo y se obtenga, por ejemplo, un parámetro negativo para la variable frecuencia.

El segundo efecto que vale la pena destacar es que cuando se elimina de la base de datos a los individuos que eligen de manera lexicográfica, todos los signos de los parámetros pasan a ser correctos excepto el del coste del viaje que resulta con signo positivo. Aquí se elimina a individuos que, por ejemplo, elegían la alternativa con menor tiempo de viaje. Cuando esta era el coche conductor, a pesar de que su coste era mucho mayor que el del bus, elegían esta alternativa.

Tabla 3.13: Resultados de las estimaciones con la encuesta definitiva

Comodidad como variable ficticia							
Parámetros (t-estadístico)		Todos	Consistentes	No cautivos	Sin lexicográficos	Sin cautivos e inconsistentes	Buenos
<i>Tiempo</i>	q_{iv}	-0,04045 (-4,9)	-0,0484 (-5,4)	-0,05535 (-5,6)	-0,04052 (-3,9)	-0,07374 (-6,5)	-0,05313 (-4,4)
<i>Coste</i>	q_c	-0,003379 (-2,0)	-0,001506 (-2,2)	-0,001021 (-1,3)	0,0006017 (0,8)	-0,001941 (-2,2)	-0,002252 (-2,4)
<i>C. Aparc.</i>	q_{cap}	-0,002753 (-4,0)	-0,003003 (-4,1)	-0,003069 (-3,8)	-0,004158 (-4,8)	-0,003644 (-4,1)	-0,003838 (-4,1)
<i>Frecuencia</i>	q_f	-0,9065 (-3,6)	0,03483 (1,5)	0,5321 (2,2)	0,0797 (3,0)	0,08293 (2,8)	0,07958 (2,7)
<i>c_Baja</i>	q_{cB}	-0,4594 (-2,0)	-1,29 (-6,8)	-1,173 (-5,8)	-1,508 (-6,8)	-1,773 (-7,5)	-1,809 (-7,1)
<i>c_Estándar</i>	q_{cE}	-0,2704 (-1,4)	-0,4694 (-2,6)	-0,2113 (-1,1)	-0,526 (-2,6)	-0,5088 (-2,3)	-0,6394 (-2,7)
$r^2(C)$		0,0619	0,0904	0,0983	0,1104	0,1600	0,1424
$l(\hat{q})$		-566,09	-493,18	-411,86	-380,43	-331,36	-290,34
<i>Observaciones</i>		871	783	664	621	576	495

Cuando se elimina a los individuos cautivos e inconsistentes; y también a los lexicográficos; se obtienen los mejores modelos. Los signos de los distintos parámetros estimados son todos correctos, mejora la log-verosimilitud de los modelos y aumenta su poder explicativo (comparar los valores del índice $r^2(C)$). En cuanto a la significatividad de los parámetros, los modelos con ambas bases de datos son muy similares.

En cuanto a la inclusión o no de individuos lexicográficos (quinta y sexta base de datos) se obtiene modelos muy similares, teniendo menor valor de log-verosimilitud el modelo con datos buenos, mientras que el modelo con individuos lexicográficos tiene mejor $r^2(C)$. En nuestra opinión, en base a éstos y otros resultados obtenidos (véase Rizzi y Ortúzar, 2003) se concluye que la inclusión de los individuos lexicográficos es opcional, especialmente cuando se trabaja con datos mixtos, esto es, Preferencias Reveladas y Declaradas, ya que en el caso de las PR estos individuos no son identificables.

En la Tabla 3.14 se presentan las medidas de disposición a pagar obtenidas para los cuatro mejores modelos. En general se observa que estos valores se ven afectados por el tipo de individuos incluidos en la base de datos. Este hecho viene a confirmar la necesidad de identificar a los individuos inconsistentes, cautivos y lexicográficos según se ha expuesto con anterioridad.

Tabla 3.14: Disposiciones a pagar

Disposiciones a Pagar (DAP)	Consistentes	No cautivos	Sin Cautivos e inconsistentes	Buenos
Valor subjetivo del tiempo de viaje (ptas/hora)	1928,29	3252,69	2279,44	1415,54
DAP por mejoras en la frecuencia (ptas/bus-hora)	23,13	521,16	42,73	35,34
DAP por incrementar la comodidad de baja a estándar (ptas)	544,89	941,92	651,31	519,36
DAP por incrementar la comodidad de estándar a alta (ptas)	311,69	206,95	262,13	283,93

4. Modelización de la demanda de transporte de pasajeros y aplicaciones

4.1 Introducción

En este capítulo se presenta la modelización realizada para los corredores estudiados. En la primera sección se describen las bases de datos utilizadas, así como la definición de las distintas variables consideradas en la modelización. En la segunda sección se presenta la modelización realizada utilizando la metodología de estimación con datos mixtos. Se han probado diferentes especificaciones de la función de utilidad a estimar con el objeto de determinar la mejor especificación para el patrón de viajes de los individuos entrevistados y la disponibilidad de los datos.

Hemos estudiado la posible existencia de correlación entre las alternativas de transporte privado, esto es, coche conductor y coche acompañante; la existencia o no de efecto renta siguiendo el planteamiento desarrollado por Jara-Díaz y Videla (1989), y el estudio de especificaciones no lineales de la función de utilidad. Una vez modelizado el comportamiento de viaje en la tercera sección, se calculan las disposiciones a pagar por cambios en las variables de nivel de servicio, y en la cuarta y última sección, se estudia

el efecto de la respuesta de la demanda de transporte de pasajeros ante distintas medidas de política.

Se han estimado modelos con datos mixtos según el procedimiento de estimación simultánea desarrollado por Bradley y Daly (1997). Con este procedimiento hay que definir una estructura jerárquica artificial para las alternativas de PD, de manera que los datos sean homogenizados por el factor de escala. Para la estimación de estos modelos se ha utilizado el software ALOGIT (Daly, 1992).

4.2 Los datos y las variables

4.2.1 Los datos

Los datos de PR requieren de un proceso de medición ex-post y la transformación del dato obtenido en la encuesta en variable, por ejemplo el origen y el destino del viaje son datos importantes para la obtención de varias variables como la distancia recorrida en vehículo privado, la distancia de caminata en origen y destino a la parada de bus, para la medición de los tiempos de viajes, qué líneas de bus debe tomar, etc.

Para la preparación de la base de datos de PR hemos realizado diversas mediciones. El objetivo de realizar las mediciones es doble: por un lado, es necesario conocer los niveles de servicio que hay que asignar a cada individuo de acuerdo a su conjunto de alternativas disponibles; y por otro lado, para intentar reducir los errores de medición en los datos de PR que como hemos comentado en el capítulo 2 vienen recogidos en las variables explicativas.

En primer lugar, hemos realizado las mediciones de los tiempos de viaje. Distinguimos entre tiempos de viaje en vehículo privado y bus. Para la medición de los tiempos de viaje en vehículo privado se definieron zonas en origen y en destino para simplificar las

mediciones, estableciendo puntos comunes de llegada para los orígenes (O en la Figura 4.1) y de partida para los destinos (D en la Figura 4.1). Se realizaron mediciones entre O_i y O, O-D y D- D_j . Estas mediciones se realizaron a distintas horas del día para intentar captar la diferencia entre hora punta y hora valle, distintos días⁶⁷ de la semana, obteniéndose la media de las cuatro mediciones realizadas para cada trayecto.

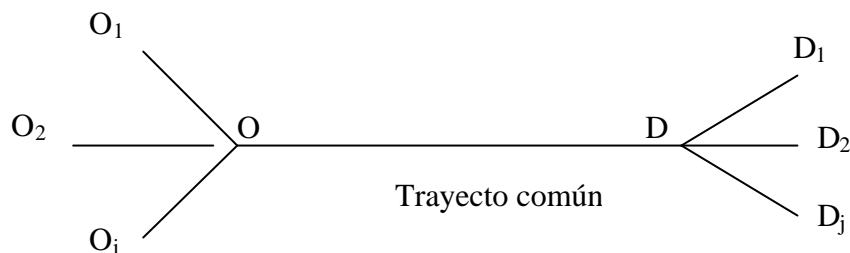


Figura 4.1: Estructura zonal para la medición de tiempos

Idealmente estas mediciones deberían realizarse con un vehículo instrumentado que dispone de un ordenador a bordo que recoge el tiempo de parada, consumo de combustible, etc. Al no disponer de este vehículo se realizaron con un vehículo normal.

En cuanto a las mediciones de tiempos en bus, se hizo uso del servicio de ayuda a la explotación (SAE) que tienen instalado las empresas *Salcai* y *Utinsa* y que permite, entre otras cosas, disponer de los tiempos de viaje entre las paradas de cada una de las líneas. La empresa *Guaguas Municipales* no dispone de SAE y por tanto, para todos aquellos trayectos realizados con esta empresa fue necesario realizar las mediciones correspondientes. Éstas se realizaron entre paradas de una misma línea, obteniéndose la media de los tiempos entre paradas y del tiempo total del viaje.

⁶⁷ Días laborables.

También se realizó la medición de las distancias recorridas por los individuos en los distintos modos de transporte así como las caminadas. Éstas han sido medidas con un sistema de información geográfica básico desarrollado por Grafcan, S.A. llamado *Canarymap* (Grafcan, 1998). Integra la información vectorial, las ortofotos y el modelo digital sobre un territorio continuo.

Este programa ofrece varias herramientas que nos permiten realizar las distintas mediciones. Destacan la posibilidad de dibujar una línea que es medida automáticamente, la identificación de topónimos (nombres y números de calles, etc) y la de dibujar recorridos mediante la creación de una capa. Para la medición de las distancias de caminata y las recorridas se utilizaron dos procedimientos ligeramente diferentes.

En primer lugar, se realizaron las mediciones de las distancias de caminata. En el caso del transporte privado, éstas no se pudieron obtener con esta herramienta porque no se disponía de los recorridos realizados de manera precisa. En origen, normalmente el individuo tiene su coche aparcado en el garaje de su casa por lo que no es posible realizar la medición y por otra parte, esta distancia puede considerarse despreciable. En destino, el problema está en que no siempre se tenía el dato donde el individuo había dejado estacionado su coche, especialmente para aquellos que estacionan en la calle ya que no siempre aparcan en el mismo sitio. Estos son más del cincuenta por ciento de los usuarios de transporte privado. Para aquellos que estacionan en aparcamiento privado también es difícil realizar la medición porque más del 75% de ellos lo hacen en el mismo destino, es el caso de centros comerciales, hospitales y centros de trabajo que disponen de aparcamiento privado gratuito.

Por otra parte, es importante medir la distancia de caminata a la parada de bus con precisión porque un individuo puede no tener disponible el transporte público si la distancia a caminar supera un determinado número de metros. Para esta medición fue necesario solicitar a las empresas de transporte la ubicación exacta de cada una de las paradas. Las mediciones se realizaron con la herramienta facilitada por el programa y a una escala de 1:5000, lo que implica un error de medición de más menos veinte metros.

La disponibilidad en el bus se define en función de la distancia a la que se encuentre la parada. Para establecer la distancia que determine la disponibilidad se estudiaron a aquellos individuos que declaraban viajar en bus y tenían disponibles otros modos de transporte. En nuestra muestra la distancia se establece en los mil metros, es decir, el bus es un modo de transporte disponible si el individuo tiene la parada a una distancia inferior o igual a mil metros.

En segundo lugar, se realizó la medición de las distancias recorridas por los vehículos. Para estas mediciones no se podía utilizar la herramienta de medida automática del programa porque ésta solo permite medir lo que se visualiza en la pantalla. En este caso, realizamos el dibujo del recorrido y éste era guardado como una capa. El dibujo se realizó a escala 1:5000, por lo que se dibujó varias líneas que conformaban el recorrido completo. Las mediciones en metros de estas líneas se obtuvieron con el programa Arc-View (ESRI Inc, 1996; Lantada y Nuñez, 2002). Al realizarse los dibujos a escala 1:5000 el error de medición sigue siendo de más menos veinte metros.

Con esta medición de distancia recorrida se obtiene el gasto en combustible de los usuarios de vehículo privado y es una variable que permite estratificar la muestra por distancia si se considerara relevante.

4.2.2 Las variables

En cuanto a las variables, vamos a comentar como se han definido algunas de las consideradas en la elección modal:

El tiempo de viaje (t_v) es una variable común a todas las alternativas de transporte, está medida en minutos. Hemos considerado el tiempo de viaje medido para el modo de transporte elegido por el individuo, porque, por ejemplo, en el caso del vehículo privado encontramos varios casos en los que el tiempo revelado era relativamente alto y poco creíble para el viaje en cuestión. En las pruebas preliminares de estimación de los modelos se comprobó que los mejores resultados se obtuvieron con los tiempos de viaje medidos, así que decidimos trabajar con estos valores. El tiempo de viaje incluye el tiempo de aparcamiento. Esta variable va acompañada de un parámetro cuyo signo se espera negativo en la función de utilidad.

El coste es otra variable común a todas las alternativas de transporte. Está medida en pesetas⁶⁸. Para el vehículo privado este coste viene dado por el gasto en combustible que se obtiene a partir de la distancia recorrida y tipo de vehículo utilizado. Este gasto en combustible se multiplica por 2 ó 2,5 para incorporar otros costes variables como el gasto de neumáticos, aceite, etc. Para el transporte colectivo el coste va a venir definido por la tarifa pagada por el individuo, en el caso de los usuarios. Para los no usuarios, pero con disponibilidad, podemos definir dos tarifas: la tarifa normal o la tarifa bono. Para los que viajan por motivo obligado y/o con una frecuencia igual o superior a las

⁶⁸ Hemos trabajado siempre en pesetas (del año 2000) porque es la moneda que estaba vigente en el momento que se realizaron las dos encuestas.

cinco veces por semana hemos definido la tarifa bono, mientras que para el resto la tarifa normal. El signo del parámetro del coste se espera negativo.

El coste de aparcamiento es una variable específica de las alternativas de transporte privado, coche conductor y coche acompañante en nuestro caso; y está en pesetas. Para los usuarios de coche conductor que paguen un aparcamiento mensual se ha definido un coste de aparcamiento al día dividiendo el coste mensual por la frecuencia mensual del viaje como procedió en la encuesta de PD para definir el nivel 1 del coste de aparcamiento. Se espera un signo negativo para el parámetro.

El tiempo de caminata, una vez medida la distancia de caminata, se obtiene a partir de la definición de velocidades de caminata por sexo y edad, que fueron obtenidas de un estudio realizado en la Universidad de Chile (Bojórquez, 2002). Se espera un signo negativo para el parámetro.

La frecuencia (f) es una variable específica del bus. Está medida como número de buses por hora porque los usuarios de bus tienen información del horario del servicio y por tanto, programan su salida en función de esta información para reducir el tiempo de espera. Partiendo de este supuesto, tiene menos sentido captar el efecto de esta variable a partir del tiempo medio que transcurre entre dos buses consecutivos. El parámetro de la frecuencia se espera con signo positivo.

La comodidad es una variable específica de las alternativas de PD. Como hemos comentado en el anterior capítulo, esta variable presenta tres niveles para el bus: una comodidad baja, una estándar y una alta, esta última intenta igualar la comodidad en el bus con la del coche. Así pues, vamos a definir la comodidad tomando como referencia la del coche, esto es, comodidad alta. Tenemos, por tanto, dos variables ficticias: una

que toma valor 1 cuando la comodidad es baja en el bus y 0 en otro caso; y una segunda que toma valor 1 cuando la comodidad es estándar en el bus y 0 en otro caso. Al definir las en referencia a la comodidad alta, sus parámetros deben resultar con signo negativo.

Para analizar el papel de la renta en las decisiones de los individuos se incorpora la tasa de gasto en la función de utilidad dividiendo al coste de viaje y al coste de aparcamiento. Esta especificación sigue el planteamiento propuesto por Jara-Díaz y Farah (1987) cuando el salario viene dado de forma exógena y se explicará en el epígrafe siguiente cuando abordemos el problema de la especificación de la función de utilidad.

En cuanto a las variables socioeconómicas, hemos considerado si el individuo trabaja o no, el sexo, el motivo del viaje, el origen del viaje y la edad. La variable socioeconómica *trabaja* toma valor 1 si trabaja y 0 en otro caso. *Sexo* se define como una variable ficticia que toma valor 1 si es hombre y 0 si es mujer. El *motivo del viaje* está definido en función de si es un viaje obligatorio (trabajo y estudios) o no; se trata de una variable ficticia que toma valor 1 cuando el motivo del viaje es obligado y 0 cuando no lo es. Con respecto al *origen*, esta variable identifica a los individuos en relación a su lugar de residencia (Arucas o Telde) y por tanto, origen del viaje descrito en la encuesta. Esta variable toma valor 1 si el individuo reside en Arucas y 0 si reside en Telde. La variable socioeconómica *edad* toma valor 1 cuando el individuo es mayor de treinta y cinco años y 0 en otro caso. En cuanto al signo de los parámetros de las variables socioeconómicas, éste dependerá de la variable modal con la que se haya especificado. Por ejemplo, la variable *trabaja* especificada interactuando con el tiempo de viaje se esperaría con signo negativo, pues partimos del supuesto de que las personas

que trabajan disponen de menos tiempo libre que las que no lo hacen y por tanto, el tiempo de viaje le generará más desutilidad. Si esta misma variable especificada interactuando con el coste el signo se esperaría positivo partiendo de la hipótesis de que el coste del viaje le genera menos desutilidad a los individuos que trabajan, pues disponen de mayor renta que los que no y además daría lugar a mayores disposiciones al pago para los que trabajan.

4.3 Estimación de los modelos de elección modal

En este apartado presentamos la parte de la modelización realizada más interesante o que mejores resultados ofrece según las hipótesis planteadas en esta investigación. Hemos partido de especificaciones de funciones de utilidad lineales para pasar a analizar cuestiones más complejas como la existencia de efecto-renta, correlación entre alternativas, la consideración de interacciones y el efecto de variables latentes y su interacción con las variables de política sobre la elección modal. En relación a este último punto hemos considerado una especificación que define la comodidad como interacción entre esta variable y el tiempo de viaje (según sugerencia de Jara-Díaz, 2001), ya que los niveles de la comodidad definidos están relacionados con la posibilidad de ir sentado o no (es el caso de la comodidad baja), la percepción de la misma debería estar fuertemente relacionada con el tiempo de viaje; ir de pie en un viaje de quince minutos es menos “incómodo” que ir de pie en un viaje de veinticinco minutos.

Se ha estimado modelos con datos mixtos siguiendo el procedimiento de estimación simultánea desarrollado por Bradley y Daly (1997). Para la estimación de los modelos se ha utilizado el software Alogit (Daly, 1992).

Todos los modelos estudiados han sido estimados para las dos especificaciones de la comodidad; la primera que supone definirla como una variable ficticia tomando como referencia la comodidad buena, donde tenemos la variable cB que toma valor 1 si la comodidad es baja en el bus y 0 en otro caso, y la variable cE que toma valor 1 si la comodidad es estándar en el bus y 0 en otro caso; la segunda especificación define las mismas variables que la primera, pero ambas aparecen multiplicadas por el tiempo de viaje.

En cuanto a la información disponible se utilizaron seis bases de datos distintas y la diferencia entre ellas está en el número de observaciones de PD. La primera considera a todos los individuos de PD (1581 observaciones); la segunda elimina a los individuos que eligieron siempre la misma alternativa en el juego de PD, esto es, individuos cautivos (1374 observaciones); la tercera elimina a los individuos que contestaron inconsistentemente (1493 observaciones), la cuarta a los que eligen la alternativa que es mejor en alguno de los atributos, es decir, realizan su elección de manera lexicográfica (1331 observaciones); la quinta elimina conjuntamente a los cautivos e inconsistentes (1286 observaciones) y la sexta, además de los cautivos e inconsistentes, elimina a los individuos lexicográficos (1205 observaciones). Para la gran mayoría de los modelos estimados, los mejores resultados se obtuvieron con la quinta⁶⁹ y sexta base de datos, es decir, aquella que elimina a los individuos cautivos e inconsistentes y la que elimina además a los lexicográficos.

Los modos de transporte para los datos de PR son tres: coche conductor, coche acompañante y bus. Para los datos de PD son sólo dos: coche conductor y bus. La

⁶⁹ Los modelos presentados están estimados con esta base de datos y por tanto se recoge al comportamiento de los individuos que eligen de forma lexicográfica.

estructura⁷⁰ jerárquica a definir para la estimación con datos mixtos se representa en la Figura 4.2; siendo μ el factor de escala que permite escalar los datos de PD de manera que ambos conjuntos de datos tengan igual varianza y pueda usarse conjuntamente. Como explicamos en el capítulo 2, Louviere et al (2000) sugieren especificar los mismos parámetros para aquellas variables que sean iguales en PR y PD. Para saber qué variables debemos especificar con igual parámetro recomiendan hacer un gráfico con los parámetros de dichas variables estimados individualmente con datos de PR y PD, de manera que para aquellas variables que sean iguales sus coeficientes deban situarse en una nube de puntos que se aproxima a una línea recta. En nuestro caso, todas las variables comunes a PR y PD (tiempo, coste, frecuencia y coste de aparcamiento) pueden especificarse con el mismo parámetro, es decir, $q_i^{PR} = q_i^{PD} = q_i$.

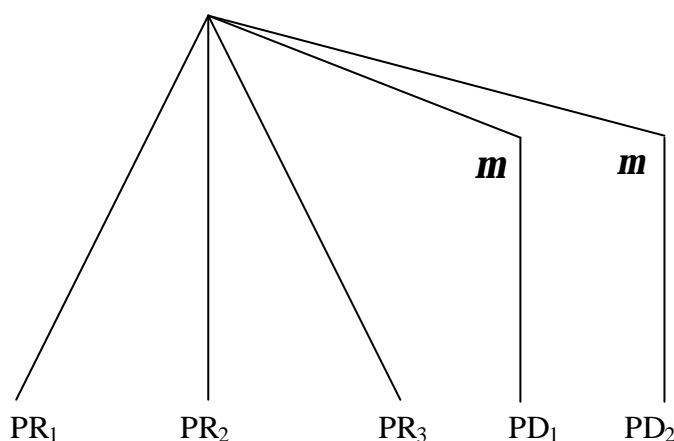


Figura 4.2: Estructura jerárquica para la estimación con datos mixtos

Para la modelización hemos partido de funciones que sólo consideraban variables modales (tiempo, coste, frecuencia, etc) donde hemos estudiado la existencia o no de correlación entre las alternativas de PR así como la importancia de las interacciones que

⁷⁰ No se considera correlación entre las alternativas de transporte privado.

permite medir el diseño de PD, su signo y significación. Posteriormente se pasó al análisis del efecto de las variables socioeconómicas sobre la percepción de los atributos de las alternativas en la elección modal.

En primer lugar, se definió un modelo básico especificando como variables el tiempo de viaje (que incluye el tiempo de aparcamiento), el coste del viaje, el coste de aparcamiento para las alternativas de transporte privado, la frecuencia para el transporte colectivo (bus), la comodidad para las alternativas de PD y las constantes específicas para cada una de las alternativas. Partiendo de este modelo, estudiamos la existencia de correlación entre las alternativas de PR; que exige definir un nuevo nido para las alternativas correlacionadas, siendo la nueva estructura⁷¹ similar a la de la Figura 4.3.

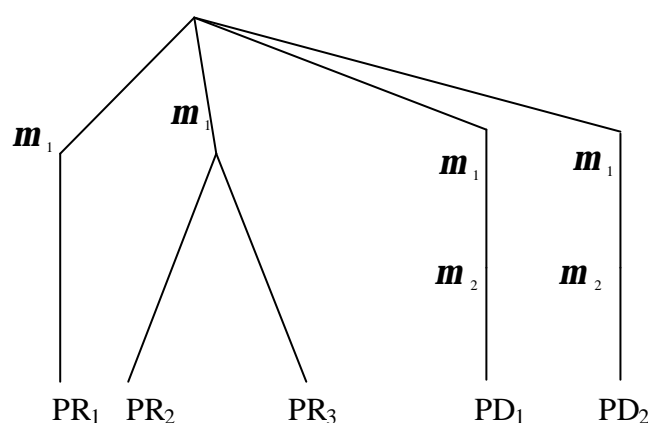


Figura 4.3: Estimación con datos mixtos con nido real entre alternativas de PR

Se analizó la correlación entre las dos alternativas de transporte privado, esto es, coche conductor y coche acompañante y en segundo lugar, entre coche acompañante y bus. En ambos casos se obtuvo que no existe correlación entre las alternativas, esto es que valor de m_i era igual a 1 o no es significativamente distinto de 1 (ver Figura 4.3) y por lo tanto el HL colapsa a un MNL.

⁷¹ Para la especificación de un modelo jerárquico consistente usando el software ALOGIT, es preciso definir los nidos artificiales de parámetros μ_2 en PD₁ y PD₂ y μ_1 en PR₁.

En relación a la forma funcional de la función de utilidad se probaron especificaciones que consideraban todas las interacciones que el diseño de PD nos permitía incluir, esto es, las interacciones dos a dos entre las variables tiempo, coste y frecuencia, resultando significativa únicamente la interacción entre el coste y la frecuencia⁷². El signo de esta interacción va a estar definido a partir de la condición de que la utilidad marginal del coste sea negativa y la de la frecuencia positiva (al estar definida ésta como buses/hora) y el signo que pueden tomar los parámetros de los efectos individuales del coste y la frecuencia. Una vez definida una especificación concreta analizaremos las condiciones que establecen el signo de la interacción⁷³.

Otra cuestión determinante en la especificación de las funciones de utilidad es la existencia de efecto-renta, esto es, la presencia de la renta en la elección modal. Del análisis microeconómico de la demanda desarrollado anteriormente, tenemos dos modelos que permiten incorporar la renta a la toma de decisión en la elección de las alternativas. Train y McFadden (1978) consideran que el individuo puede elegir el número de horas que trabaja siendo su renta endógena; ésta dependerá de las horas que trabaje. En este caso, el coste en la función de utilidad está dividido por la tasa salarial (w). Jara-Díaz y Farah (1987) sugieren que si la renta es exógena debido a que el individuo no puede decidir el número de horas que trabaja el coste debe ir dividido por la tasa de gasto, esto es, la renta que el individuo dispone para gastar por unidad de tiempo disponible. En nuestro estudio hemos considerado más adecuadas las especificaciones que incluyen la tasa de gasto debido a que el mercado laboral en España está bastante regulado, por lo que el trabajador no puede decidir de forma flexible cuántas horas trabajar. En la práctica, una recomendación habitual consiste en

⁷² Para otras especificaciones de interacciones ver Ortúzar *et al* (2000).

⁷³ Este análisis se presentará más adelante, pp. 137-138.

consultar este dato en la encuesta de PR; existen ya ejemplos de encuestas que incorporan un diario de actividades para determinar si realmente el número de horas trabajadas es endógeno o exógeno para el individuo (ver Jara-Díaz *et al* 2002) y en función de esto justificar la especificación uno u otro modelo.

Siguiendo a Jara-Díaz y Videla (1989) estudiamos la existencia del efecto-renta. Estos autores proponen probar para los distintos estratos de la muestra una especificación para la FUIC⁷⁴ en la que aparece, además de las variables características de la alternativa, el coeficiente del coste al cuadrado. El signo de este parámetro debe ser positivo y disminuir a medida que aumente la renta, al igual que la utilidad marginal de la renta (I). Si esto es así se verifica la existencia de efecto renta y ésta debería incluirse en la especificación de las FUIC. La FUIC a estimar tiene la siguiente expresión:

$$V_j = \mathbf{q}_j + \mathbf{q}_c \cdot c_j + \frac{1}{2} \cdot \mathbf{q}_{c^2} \cdot c_j^2 + V'(\mathcal{Q}_j) \quad (4.1)$$

En nuestro estudio se definieron tres estratos de ingreso: el primero formado por individuos con renta mensual inferior a 200.000 pts; el segundo por individuos con renta comprendida entre 200.000 pts y 300.000 pts y el tercero por individuos con renta mayor o igual que 300.000 pts. Para esta definición de estratos, el parámetro del coste al cuadrado (\mathbf{q}_{c^2}) resulta con signo positivo y significativamente distinto de cero. La utilidad marginal de la renta no es menor en el tercer estrato que en el segundo. Redefinimos los estratos, pasando de tres a dos estratos: el primero para rentas menores o iguales que 200.000 pts y el segundo para rentas mayores. En este caso, se verificó que, en primer lugar, el parámetro del coste al cuadrado (\mathbf{q}_{c^2}) resultaba positivo y

⁷⁴ Función de utilidad indirecta condicional, ver capítulo 2-pp. 13-14.

significativamente distinto de cero; y en segundo lugar, que tanto éste como la utilidad marginal del renta decrecían con la renta (ver anexo II). Por tanto, la estratificación considerada nos da un indicio de la existencia de efecto renta. De acuerdo con esto, se justifica la inclusión de la renta a través de la tasa de gasto en la especificación de la utilidad.

La siguiente cuestión es cómo definir la tasa de gasto, ya que se trata de la renta por unidad de tiempo disponible podemos bien considerar la renta individual o la renta familiar *per cápita*. Con respecto al tiempo total disponible bien las 24 horas del día o éstas menos las horas que se dedican a necesidades fisiológicas como son dormir, comer, etc. Para este caso, establecimos que serían unas ocho horas para dormir y dos horas para otras necesidades como higiene, alimentación, etc; siendo entonces el tiempo total disponible para el individuo de catorce horas al día. Esto establece una restricción fuerte en el sentido de que a todos los individuos se les asigna el mismo número de horas de sueño, cuando este valor puede ser bastante variable entre ellos. Con respecto a la renta, si definimos la renta individual no se puede considerar aquellos que no han declarado ingreso alguno como es el caso de las amas de casa y los estudiantes, que a pesar de no percibir una renta, disponen de una cantidad para sus gastos. Tras probar los cuatro posibles casos para la definición de la tasa de gasto, los mejores resultados se obtuvieron cuando se consideraba el ingreso familiar *per cápita* y como tiempo total disponible el total de horas del día.

Definida la parte de la FUIC correspondiente a las variables modales, investigamos el efecto de las variables socioeconómicas en la elección modal. En primer lugar, estudiamos estratificaciones por distancia recorrida y motivo de viaje, distinguiendo entre motivo obligado y no obligado. Ninguna de las estratificaciones que se probaron,

resultaron significativas. En segundo lugar, especificamos las variables socioeconómicas interactuando con las variables modales. Así especificamos un parámetro para el sexo en el tiempo de viaje, otro para el coste, etc, es decir, un parámetro del sexo para cada variable modal y así con todas las variables socioeconómicas que queríamos estudiar. De esta forma pudimos establecer la hipótesis acerca de la percepción de las variables modales en función de las características socioeconómicas del individuo.

Para cualquier atributo de la alternativa, X_i , se define un parámetro básico (q_{X_i}) y un parámetro para cada variable socioeconómica ($q_{X_i - VSE_h}$), expresión (4.2). El parámetro básico mide la percepción del atributo cuando todas las variables socioeconómicas toman valor cero y el parámetro de éstas mide la percepción dicho atributo en función de la característica socioeconómica correspondiente (Ortúzar y Willumsen, 2001):

$$\left(q_{X_i} + \sum_{h=1}^n q_{X_i - VSE_h} \cdot VSE_h \right) \cdot X_i \quad (4.2)$$

De todas las variables socioeconómicas probadas resultaron significativas cinco, cada una de ellas para una variable modal diferente. Así tenemos la variable *trabaja* especificada en el tiempo de viaje, el *sexo* en el coste, el *motivo del viaje* en el coste de aparcamiento, el *origen* en el tiempo de caminata y la *edad* en la frecuencia del bus.

Con respecto al parámetro del tiempo de viaje, hemos estudiado la especificación de parámetros específicos para el tiempo de viaje en transporte privado (alternativas de coche conductor en PR y PD y coche acompañante sólo en PR) y en transporte colectivo (alternativa de bus en PR y PD) basándonos en la hipótesis de que a las personas que viajan en transporte privado les genera más desutilidad el tiempo de viaje que a las que

viajan en transporte colectivo. Por ésto se esperaría un parámetro del tiempo de viaje en transporte privado mayor en términos absolutos que el del transporte colectivo. Esta definición de parámetros específicos implica una mayor disposición al pago por ahorros de tiempo de viaje en transporte privado que en transporte colectivo. Para poder determinar si es mejor el modelo con parámetros específicos que el modelo genérico se utiliza el test de razón de verosimilitud (4.3), que establece como hipótesis nula que el modelo restringido (en nuestro caso, el modelo con parámetros genéricos para el tiempo de viaje) es una especificación correcta.

$$LR = -2 \left\{ l^* (\hat{\mathbf{q}}_r) - l^* (\hat{\mathbf{q}}) \right\} \quad (4.3)$$

donde $l^* (\hat{\mathbf{q}}_r)$ es la log-verosimilitud en convergencia del modelo restringido y $l^* (\hat{\mathbf{q}})$ es la del modelo más general. Este test distribuye asintóticamente χ^2 con r grados de libertad, siendo r el número de restricciones lineales. La hipótesis nula se rechaza si $LR > c_{r,95\%}^2$, lo que quiere decir que el modelo restringido es erróneo. Sólo es posible aplicar este test cuando un modelo es una versión restringida de otro.

Para la comparación de modelos distintos se puede utilizar el índice rho cuadrado corregido (Tardiff, 1976):

$$r^2 = 1 - \frac{l^* (\hat{\mathbf{q}})}{l^* (C)} \quad (4.4)$$

siendo $l^* (\hat{\mathbf{q}})$ la log-verosimilitud en convergencia del modelo estimado y $l^* (C)$ la log-verosimilitud del modelo especificando sólo constantes. Este valor está comprendido entre 0 y 1 y permite comparar distintos modelos y distintas muestras.

En las expresiones (4.5) y (4.6) se presentan las especificaciones de las funciones de utilidad de los modelos 1 y 2, respectivamente. En el modelo 1 se especifica la comodidad como variable ficticia y en el modelo 2 interactuando con el tiempo de viaje.

Para el modelo 1 son:

$$\begin{aligned}
V_{conductor}^{PR} &= \mathbf{q}_{cmc} + (\mathbf{q}_{tv} + \mathbf{q}_{tv_T} \cdot T) \cdot tv + (\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S) \cdot c / g \\
&+ (\mathbf{q}_{cap/g} + \mathbf{q}_{cap/g_mvo} \cdot mvo) \cdot cap / g \\
V_{acompañante}^{PR} &= \mathbf{q}_{cma} + (\mathbf{q}_{tv} + \mathbf{q}_{tv_T} \cdot T) \cdot tv + (\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S) \cdot c / g \\
&+ (\mathbf{q}_{cap/g} + \mathbf{q}_{cap/g_mvo} \cdot mvo) \cdot cap / g \\
V_{bus}^{PR} &= (\mathbf{q}_{tv} + \mathbf{q}_{tv_T} \cdot T) \cdot tv + (\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S) \cdot c / g \\
&+ (\mathbf{q}_{tcam} + \mathbf{q}_{tcam_O} \cdot O) \cdot tcam + (\mathbf{q}_f + \mathbf{q}_{f_E} \cdot E) \cdot f \\
V_{conductor}^{PD} &= \mathbf{q}_{cmcc} + (\mathbf{q}_{tv} + \mathbf{q}_{tv_T} \cdot T) \cdot tv + (\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S) \cdot c / g \\
&+ (\mathbf{q}_{cap/g} + \mathbf{q}_{cap/g_mvo} \cdot mvo) \cdot cap / g \\
V_{bus}^{PD} &= (\mathbf{q}_{tv} + \mathbf{q}_{tv_T} \cdot T) \cdot tv + (\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S + \mathbf{q}_{f \cdot c/g} \cdot f) \cdot c / g \\
&+ (\mathbf{q}_f + \mathbf{q}_{f_E} \cdot E) \cdot f + \mathbf{q}_{cB} \cdot cB + \mathbf{q}_{cE} \cdot cE
\end{aligned} \tag{4.5}$$

Para el modelo 2 son:

$$\begin{aligned}
V_{conductor}^{PR} &= \mathbf{q}_{cmc} + (\mathbf{q}_{tv} + \mathbf{q}_{tv_T} \cdot T) \cdot tv + (\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S) \cdot c / g \\
&+ (\mathbf{q}_{cap/g} + \mathbf{q}_{cap/g_mvo} \cdot mvo) \cdot cap / g \\
V_{acompañante}^{PR} &= \mathbf{q}_{cma} + (\mathbf{q}_{tv} + \mathbf{q}_{tv_T} \cdot T) \cdot tv + (\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S) \cdot c / g \\
&+ (\mathbf{q}_{cap/g} + \mathbf{q}_{cap/g_mvo} \cdot mvo) \cdot cap / g \\
V_{bus}^{PR} &= (\mathbf{q}_{tv} + \mathbf{q}_{tv_T} \cdot T) \cdot tv + (\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S) \cdot c / g \\
&+ (\mathbf{q}_{tcam} + \mathbf{q}_{tcam_O} \cdot O) \cdot tcam + (\mathbf{q}_f + \mathbf{q}_{f_E} \cdot E) \cdot f \\
V_{conductor}^{PD} &= \mathbf{q}_{cmcc} + (\mathbf{q}_{tv} + \mathbf{q}_{tv_T} \cdot T) \cdot tv + (\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S) \cdot c / g \\
&+ (\mathbf{q}_{cap/g} + \mathbf{q}_{cap/g_mvo} \cdot mvo) \cdot cap / g \\
V_{bus}^{PD} &= (\mathbf{q}_{tv} + \mathbf{q}_{tv_T} \cdot T + \mathbf{q}_{tv_cB} \cdot cB + \mathbf{q}_{tv_cE} \cdot cE) \cdot tv \\
&+ (\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S + \mathbf{q}_{f \cdot c/g} \cdot f) \cdot c / g + (\mathbf{q}_f + \mathbf{q}_{f_E} \cdot E) \cdot f
\end{aligned} \tag{4.6}$$

donde

tv es el tiempo de viaje.

c es el coste del viaje.

g es la tasa de gasto.

cap es el coste de aparcamiento.

$tcam$ es el tiempo de caminata.

f es la frecuencia.

cB es comodidad baja.

cE es comodidad estándar.

T es trabaja.

S es el sexo.

mvo es el motivo de viaje.

O es el origen.

E es la edad.

qs los parámetros desconocidos.

De acuerdo a estas especificaciones podemos analizar el signo de la interacción entre el coste y la frecuencia. Las utilidades marginales del coste y de la frecuencia se obtienen derivando la FUIC con respecto al coste y la frecuencia respectivamente, expresión (4.7). De estas condiciones, siendo el signo del parámetro del coste negativo y de la frecuencia positiva se obtiene que el parámetro de la interacción debe pertenecer al intervalo definido en la expresión (4.8).

$$\begin{aligned} \frac{\partial V}{\partial c} &= \frac{1}{g} (\mathbf{q}_{c/g} + \mathbf{q}_{c/g-S} \cdot S + \mathbf{q}_{f \cdot c/g} \cdot f) < 0 \\ \frac{\partial V}{\partial f} &= (\mathbf{q}_f + \mathbf{q}_{f-E} \cdot E + \mathbf{q}_{f \cdot c/g} \cdot c/g) > 0 \end{aligned} \quad (4.7)$$

A partir de de (4.7) cuando $\mathbf{q}_{c/g} < 0$ y $\mathbf{q}_f > 0$, entonces

$$\begin{aligned}
q_{f \cdot c/g} &< -\frac{q_{c/g} + q_{c/g-S} \cdot S}{f} \text{ y} \\
q_{f \cdot c/g} &> -\frac{q_f + q_{f-E} \cdot E}{c} \cdot g \\
\therefore q_{f \cdot c/g} &\in \left[-\frac{q_f + q_{f-E} \cdot E}{c} \cdot g, -\frac{q_{c/g} + q_{c/g-S} \cdot S}{f} \right]
\end{aligned} \tag{4.8}$$

El parámetro del sexo en el coste del viaje ($q_{c/g-S}$) y el de la edad en la frecuencia (q_{f-E}) tienen signo positivo en nuestros modelos, así el límite inferior del intervalo es un número negativo mientras que el superior es un número positivo.

El parámetro estimado para la interacción debe pertenecer al intervalo que definen estas restricciones. El signo de la interacción puede ser positivo o negativo siempre y cuando pertenezca al intervalo. Este intervalo varía en función de las utilidades especificadas para las alternativas, por tanto, para cada modelo es necesario determinar el intervalo en el que se encuentra el signo de las interacciones. Además la definición de alguna otra interacción puede afectar a la definición del intervalo.

En nuestro caso se constató que el parámetro de la interacción siempre resultaba negativo y esta condición se verificó para prácticamente todo el rango de la muestra utilizada en la investigación. En concreto, para modelo 1 no se verifica un 12,25% de las observaciones y para el modelo 4 en un 20%. Dado que no se trata de porcentajes despreciables, en la aplicación de los modelos para el cálculo de disposiciones a pagar y análisis de política estudiaremos la sensibilidad de los resultados a la consideración, o no, de estos individuos.

En la Tabla 4.1 se presentan los modelos estimados para la especificación de parámetros genéricos para el tiempo de viaje y las dos formas de especificar la comodidad; en el modelo 1 se especifica la comodidad como variable ficticia y en el modelo 2

interactuando con el tiempo de viaje. Los resultados fueron estimados con la base de datos que elimina sólo a los individuos cautivos e inconsistentes en PD.

En cuanto a los signos de los parámetros, todos resultan con el signo esperado, a excepción de la constante de la alternativa coche acompañante que *a priori* se esperaría positiva. El signo negativo de esta constante se podría interpretar como que existe una mayor preferencia por el bus que por el coche acompañante, *ceteris paribus*, pero esta es una interpretación un poco forzada ya que la constante básicamente busca asegurar que la proporción de mercado modelizada de la alternativa sea igual a la proporción de mercado observada (Carrasco y Ortúzar, 2002).

También se especificó constante específica para la alternativa de coche conductor en PD. En este caso, resultó ser casi cero y poco significativa. Se estimó nuevamente el modelo eliminando esta constante de la especificación; estas últimas son las estimaciones que se presentan en la tabla.

Respecto al signo de las variables socioeconómicas, el parámetro de la variable *trabaja* en el tiempo de viaje resulta negativo, lo que quiere decir que para los individuos que trabajan (cuando la variable trabaja toma valor 1) la desutilidad del tiempo de viaje es mayor en la cuantía de este parámetro que para los individuos que no trabajan (cuando toma valor 0). Los individuos que trabajan tienen menos tiempo disponible que los que no trabajan, por lo tanto, es lógico que la desutilidad que genera el tiempo de viaje sea mayor que para los que no trabajan. Por otra parte, esta mayor desutilidad va a implicar una mayor disposición al pago por ahorros de tiempo de viaje (como veremos en el siguiente epígrafe), lo cual confirma la hipótesis de que las personas que trabajan tienen, en términos generales, una mayor disposición al pago, como ya apuntamos.

Tabla 4.1: Modelos con parámetros genéricos

Parámetros		Modelo 1	Modelo 2
C_conductor	q_{cmc}	3,463 (7,7) ¹	3,582 (7,9)
C_acompañante	q_{cma}	-0,8219 (-2,7)	-0,773 (-2,5)
Tiempo de viaje	q_{iv}	-0,01914 (-1,9)	-0,01494 (-1,4)
Tv_Trabaja	q_{iv_T}	-0,04859 (-2,7)	-0,05189 (-2,9)
Coste del viaje	$q_{c/g}$	-0,3309 (-3,7)	-0,3747 (-4,2)
C/g_Sexo	q_{c/g_S}	0,1695 (2,5)	0,2041 (2,8)
C/g-Frecuencia	q_{c/g_f}	-0,04388 (-2,5)	-0,04498 (-2,3)
Coste de aparcamiento	$q_{cap/g}$	-0,0483 (-0,5)	-0,03251 (-0,3)
Cap/g_Motivo viaje	q_{cap/g_mvo}	-0,4746 (-3,1)	-0,5002 (-2,9)
Tiempo de caminata	q_{tcam}	-0,07938 (-2,1)	-0,07846 (2,0)
Tcam_Origen	q_{tcam_O}	-0,1354 (-2,4)	-0,1341 (-2,4)
Frecuencia	q_f	0,1294 (2,9)	0,1174 (-2,5)
Frecuencia_Edad	q_{f_E}	0,1137 (2,3)	0,1483 (2,7)
Comodidad Baja	q_{cB}	-1,967 (-4,2)	-
Comodidad Estándar	q_{cE}	-0,5641 (-2,4)	-
Tv_ Comodidad Baja	q_{iv_cB}	-	-0,0584 (-4,2)
Tv_ Comodidad Estándar	q_{iv_cE}	-	-0,01341 (-1,9)
Factor de escala	m	0,8884	0,7213
		(4,0) [0,50] ²	(4,2) [1,63]
ρ^2	$\rho^2(C)$	0,1304	0,1237
Log-verosimilitud	$l(\theta)$	-583,4861	-588,0076
Nº de observaciones		1286	1286

¹estadístico t²test del factor de escala con respecto a uno.

La segunda variable socioeconómica significativa es el *sexo* especificada en el coste del viaje. Su parámetro toma signo positivo, lo que significa que para los hombres (cuando

la variable toma valor 1) la percepción del coste es igual a la suma de ambos parámetros, $q_{c/g} + q_{c/g_S}$; mientras que para las mujeres (cuando la variable toma valor 0) sería $q_{c/g}$. Esta especificación del sexo en el coste va a determinar que la disposición a pagar de los hombres sea mayor que la de las mujeres, pues el denominador es menor en la cuantía del parámetro q_{c/g_S} . Esto quiere decir la percepción entre hombres y mujeres es diferente, lo que define patrones de comportamiento distintos en función del género.

Como tercera variable socioeconómica tenemos el *motivo de viaje*, que fue especificado en el coste de aparcamiento. Sin embargo, al incluirla el parámetro base del coste de aparcamiento, $q_{cap/g}$, perdió significación, siendo ésta bastante baja; no obstante, nos pareció más interesante dejar la variable socioeconómica que no incluirla para captar el efecto del motivo y obtener distintas valoraciones del atributo en función de si el motivo del viaje es obligado (cuando la variable toma valor 1) o no. Además, el parámetro del motivo del viaje en el coste de aparcamiento resulta bastante significativo, y su signo negativo nos está indicando que a las personas que viajan por motivo obligado les genera más desutilidad pagar el aparcamiento que a las personas que viajan por otro motivo.

Normalmente, los viajes obligados (por trabajo y estudios) son viajes con una mayor frecuencia que los viajes por otro motivo y esta mayor frecuencia, en términos de coste de aparcamiento, implica una mayor proporción de renta gastada en transporte.

La variable *origen* fue especificada en el tiempo de caminata. Esta variable toma valor 1 para los entrevistados que residen en Arucas y 0 para los que residen en Telde. Como su parámetro resultó con signo negativo, esto implica que a las personas que residen en Arucas caminar hasta la parada de bus les genera más desutilidad que a los residentes en Telde, y por tanto, tienen una disposición a pagar mayor por reducir este tiempo de

caminata. Este parámetro está recogiendo un hecho importante y es que en Arucas la distancia a la parada de bus es, normalmente, mayor que en el caso del municipio de Telde, la señalización es escasa y en muchas ocasiones no hay marquesina o alguna estructura física que permita a la persona esperar sentada o resguardarse cuando las condiciones climatológicas no son buenas. En este sentido, en general, la zona Norte tiene un déficit en cuanto al acceso al transporte colectivo de viajeros.

Y finalmente, la *edad* que está especificada con la frecuencia; el parámetro resulta con signo positivo e implica que para las personas mayores de treinta y cinco años (cuando la variable edad toma valor 1) las mejoras en la frecuencia son más valoradas, en parte porque a estas personas les incomoda más esperar en la parada de bus, los aumentos en frecuencia reducen el tiempo medio de espera.

El parámetro del tiempo es poco significativo en el modelo 2 (ver Tabla 4.1); esto se debe a que se han definido tres variables interactuando con el tiempo de viaje: si trabaja, la comodidad baja y la comodidad estándar, lo que puede disminuir la significación del parámetro básico. Sin embargo, todas estas son significativas⁷⁵ al 95%.

Para ambos modelos el factor de escala⁷⁶ (m) es menor que uno, lo significa que la varianza de los datos de PD es mayor que la de los datos de PR, esto es, tienen más ruido los datos de PD. Y si comparamos ambos modelos en función del r^2 calculado según la expresión (4.4), el modelo 1 presenta un mejor ajuste que el modelo 2.

⁷⁵ El parámetro del tiempo de viaje por la comodidad estándar es significativo al 90%, pero está muy próximo al 95%, $t=-1,9$.

⁷⁶ El factor de escala es el cociente entre las desviaciones típicas de los datos de PR y PD (ver expresión 2.58).

En la Tabla 4.2 se presentan los modelos que consideran parámetros específicos para el tiempo de viaje, distinguiendo entre transporte privado (coche conductor en PR y PD; y acompañante en PR) y colectivo (bus en PR y PD). El modelo 3 se corresponde al modelo 1, mientras que el modelo 4 al modelo 2, esto es, el modelo 1 es el modelo restringido del modelo 3 y el modelo 2 del 4.

Los resultados de la estimación son similares a los de modelos restringidos, signos de acuerdo a lo esperado y significación al 95% para la gran mayoría de los parámetros. El parámetro del coste de aparcamiento sigue resultando poco significativo, y también pasa a serlo el del tiempo de viaje en transporte colectivo, especialmente para el modelo 4. Como la comodidad está definida en la alternativa del bus en PD, el parámetro del tiempo en transporte colectivo resulta muy poco significativo en este modelo. Igual que en los anteriores modelos, el factor de escala resulta menor que uno y el ajuste del modelo 3 es un poco mejor que el del modelo 4.

Para comparar el modelo de parámetros genéricos con el de parámetros específicos debemos contrastar la hipótesis de que los parámetros del tiempo de viaje en transporte privado sean iguales a los del transporte colectivo, esto implica la definición de dos restricciones lineales:

$$\begin{aligned} \mathbf{q}_{tp} - \mathbf{q}_{tc} &= 0 \\ \mathbf{q}_{tp_T} - \mathbf{q}_{tc_T} &= 0 \end{aligned} \tag{4.9}$$

Según el test de razón de verosimilitud, expresión (4.3), aceptaríamos que el modelo 1 es una especificación correcta pues $LR_{\text{modelo 1-3}} = 2,609 < \mathbf{c}_{2,95\%}^2 = 5,99$; mientras que para los modelos donde la comodidad se especifica interactuando con el tiempo de

viaje, se rechazaría la hipótesis nula, aceptándose el modelo de parámetros específicos, modelo 4, $LR_{\text{modelo 2-4}} = 6,2616 > c_{2,95\%}^2 = 5,99$.

La siguiente cuestión sería decidir entre el modelo 1 y el modelo 4. Como ambos modelos son bastantes similares, es difícil determinar cuál es estadísticamente superior, por lo que hemos optado por presentar ambos y comparar los resultados. Las disposiciones a pagar que se derivan de cada uno son muy interesantes, ya que permiten obtener valores según diferentes características socioeconómicas de la muestra, y éstas difieren entre los modelos 1 y 4; concretamente, para el tiempo de viaje con el modelo 4 se obtienen dieciséis valores subjetivos del tiempo de viaje en función de si viaja en transporte privado o colectivo, del sexo, si trabaja y nivel de comodidad, frente a los ocho que se obtienen con el modelo 1. Nos pareció interesante presentar ambos modelos y analizar la sensibilidad de los resultados (las disposiciones a pagar y análisis de predicción de distintas medidas de política) a la especificación.

Tabla 4.2: Modelos con parámetros específicos.

Parámetros		Modelo 3	Modelo 4
C_conductor	q_{cmc}	3,617 (7,0) ¹	4,138 (6,9)
C_acompañante	q_{cma}	-0,6559 (-1,7)	-0,1882 (-0,4)
Tv_Tprivado	q_{ttp}	-0,02176 (-1,9)	-0,02743 (-2,0)
Tv_Tprivado_Trabaja	q_{ttp_T}	-0,05407 (-2,8)	-0,05323 (-2,7)
Tv_Tcolectivo	q_{ttc}	-0,01650 (-1,6)	-0,007941 (-0,7)
Tv_Tcolectivo_Trabaja	q_{ttc_T}	-0,04946 (-2,7)	-0,05124 (-2,9)
Coste del viaje	$q_{c/g}$	-0,3287 (-3,5)	-0,370 (-4,0)
C/g_Sexo	q_{c/g_S}	0,1856 (2,6)	0,2267 (2,9)
C/g_Frecuencia	q_{c/g_f}	-0,0428 (-2,2)	-0,05021 (-2,2)
Coste de aparcamiento	$q_{cap/g}$	-0,04808 (-0,5)	-0,0242 (-0,2)
Cap/g_Motivo viaje	q_{cap/g_mvo}	-0,4337 (-2,7)	-0,4447 (-2,4)
Tiempo de caminata	q_{tcam}	-0,07714 (-2,0)	-0,07302 (-1,9)
Tcam_Origen	q_{tcam_O}	-0,1422 (-2,2)	-0,1442 (-2,6)
Frecuencia	q_f	0,1214 (2,6)	0,1032 (2,0)
Frecuencia_Edad	q_{f_E}	0,1057 (2,0)	0,1593 (2,6)
Comodidad Baja	q_{cB}	-2,198 (-3,7)	-
Comodidad Estándar	q_{cE}	-0,748 (-2,5)	-
Tv_ Comodidad Baja	q_{ttc_cB}	-	-0,07894 (-3,9)
Tv_ Comodidad Estándar	q_{ttc_cE}	-	-0,02806 (-2,5)
Factor de escala	m	0,8495	0,6211
		(3,7) [0,65] ²	(4,0) [2,43]
ρ^2	$\rho^2(C)$	0,1323	0,1283
Log-verosimilitud	$l(\theta)$	-582,1816	-584,8768
Nº de observaciones		1286	1286

¹estadístico t²test del factor de escala con respecto a uno.

4.4 Cálculo de disposiciones a pagar

En este epígrafe se van a obtener las disposiciones a pagar por mejoras en el atributo de una alternativa, lo que se denomina normalmente valor subjetivo del atributo o característica k , y que se obtiene, según la teoría microeconómica, como el cociente negativo entre la utilidad marginal de dicha característica y la utilidad marginal de la renta. Este índice mide lo que el individuo está dispuesto a pagar por la mejora del atributo en cuestión. La utilidad marginal de la renta es igual a menos la utilidad marginal del coste, así tenemos:

$$SV_k = -\frac{\partial V / \partial k}{\partial V / \partial I} = \frac{\partial V / \partial k}{\partial V / \partial c} \quad (4.10)$$

Utilizando la expresión (4.10), derivaremos los valores subjetivos de las distintas características de las alternativas para el modelo 1 y 4. Las disposiciones a pagar que difieren entre ambos modelos son las del tiempo de viaje y la comodidad; las del tiempo de caminata y de la frecuencia permanecen invariables en cuanto a la expresión a utilizar para la obtención del valor.

Las utilidades marginales se han obtenido sobre una utilidad híbrida definida a partir de los atributos comunes a PR y PD y a aquellos que sólo están especificadas en PR o en PD.

Para el modelo 1, el valor subjetivo del tiempo de viaje depende de si es hombre o mujer y de si trabaja o no, la expresión es:

$$\begin{aligned}
 VStv_{t_privado} &= \frac{\mathbf{q}_{iv} + \mathbf{q}_{iv_T} \cdot T}{\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S} \cdot g \\
 VStv_{t_colectivo} &= \frac{\mathbf{q}_{iv} + \mathbf{q}_{iv_T} \cdot T}{\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S + \mathbf{q}_{c/g_f} \cdot f} \cdot g
 \end{aligned} \tag{4.11}$$

Mientras que para el modelo 4, el valor subjetivo del tiempo de viaje va a depender además del nivel de comodidad:

$$\begin{aligned}
 VStv_{t_privado} &= \frac{\mathbf{q}_{tp} + \mathbf{q}_{tp_T} \cdot T}{\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S} \cdot g \\
 VStv_{t_colectivo} &= \frac{\mathbf{q}_{tc} + \mathbf{q}_{tc_T} \cdot T + \mathbf{q}_{tc_cB} \cdot cB + \mathbf{q}_{tc_cE} \cdot cE}{\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S + \mathbf{q}_{c/g_f} \cdot f} \cdot g
 \end{aligned} \tag{4.12}$$

Para el tiempo de caminata, el valor subjetivo depende de si es hombre o mujer y de si reside en Arucas o Telde:

$$VStcam = \frac{\mathbf{q}_{tcam} + \mathbf{q}_{tcam_O} \cdot O}{\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S + \mathbf{q}_{c/g_f} \cdot f} \cdot g \tag{4.13}$$

El valor subjetivo de la frecuencia está en función de la edad y del sexo:

$$VSf = \frac{\mathbf{q}_f + \mathbf{q}_{f_E} \cdot E + \mathbf{q}_{c/g_f} \cdot c/g}{\mathbf{q}_{c/g} + \mathbf{q}_{c/g_S} \cdot S + \mathbf{q}_{c/g_f} \cdot f} \cdot g \tag{4.14}$$

Con respecto a la comodidad, vamos a tener una disposición a pagar por cambiar de comodidad baja a comodidad estándar, esto es, disposición a pagar por no ir de pie y otra por cambiar de comodidad estándar a alta, disposición a pagar por elegir asiento. Para el modelo 4 estas disposiciones van a estar en función del tiempo de viaje.

Para el modelo 1, la disposición a pagar por mejoras en la comodidad en función del sexo son:

$$\begin{aligned}
 DAP_{cB_{-}cE} &= -\frac{\mathbf{q}_{cE} - \mathbf{q}_{cB}}{\mathbf{q}_{c/g} + \mathbf{q}_{c/g_{-}S} \cdot S + \mathbf{q}_{c/g \cdot f} \cdot f} \cdot g \\
 DAP_{cE_{-}cA} &= \frac{\mathbf{q}_{cE}}{\mathbf{q}_{c/g} + \mathbf{q}_{c/g_{-}S} \cdot S + \mathbf{q}_{c/g \cdot f} \cdot f} \cdot g
 \end{aligned}
 \tag{4.15}$$

Y para el modelo 4, tenemos:

$$\begin{aligned}
 DAP_{cB_{-}cE} &= -\frac{(\mathbf{q}_{tc_{-}cE} - \mathbf{q}_{tc_{-}cB}) \cdot tv}{\mathbf{q}_{c/g} + \mathbf{q}_{c/g_{-}S} \cdot S + \mathbf{q}_{c/g \cdot f} \cdot f} \cdot g \\
 DAP_{cE_{-}cA} &= \frac{\mathbf{q}_{tc_{-}cE} \cdot tv}{\mathbf{q}_{c/g} + \mathbf{q}_{c/g_{-}S} \cdot S + \mathbf{q}_{c/g \cdot f} \cdot f} \cdot g
 \end{aligned}
 \tag{4.16}$$

Como puede observarse, las expresiones anteriores dependen de variables que son individuales, como es la tasa de gasto o la frecuencia de la alternativa bus. Por tanto, debemos calcular tales valores para cada individuo y posteriormente obtener un valor medio del total de individuos por clases, éstas estarán definidas en función de las características socioeconómicas que afecten a cada disposición. La agregación se realiza mediante la enumeración muestral (ver epígrafe 2.4).

Presentamos los valores subjetivos calculados con los modelos 1 y 4. Estos cálculos se realizan solo para las observaciones de PR, ya que se trata de elecciones efectivas y no de preferencias eventuales como es el caso de los datos de PD. La FUIC se construye utilizando los parámetros comunes estimados así como los no comunes, según recomiendan Louviere *et al* (2000). Por otra parte, los datos de PD han sido escalados para la estimación conjunta. Si utilizamos parámetros de variables que solo están definidas en los datos de PD (la comodidad en nuestro caso), debemos multiplicar dichos parámetros por el factor de escala para el cálculo de disposiciones al pago y predicción de demanda. Sin embargo, cuando las variables están definidas para los datos

de PR (caso de la interacción del coste y la frecuencia) y se utiliza un parámetro estimado en la utilidad de PD no es preciso escalarlo (Cherchi, 2003).

Para cada modelo hay un número no despreciable de individuos que no verifican las condiciones microeconómicas definidas en relación al signo de la interacción (12,5% para el modelo 1 y 20% para el modelo 4). Para estos individuos, resultan valores subjetivos de la frecuencia con el signo contrario debido a que violan la condición de que la utilidad marginal de la frecuencia sea positiva; para el resto de las disposiciones al pago los signos son correctos para todos los individuos, incluidos estos últimos.

Hemos realizado los cálculos de dos formas: para todos los individuos y eliminando a aquellos que no verifican las condiciones microeconómicas según el modelo. En el primer caso, a estos individuos se les asigna valor cero para calcular el promedio de la disposición a pagar de cada categoría. En el segundo caso, simplemente los eliminamos.

Existen diferencias entre los valores subjetivos calculados considerando o no a los individuos que violan los supuestos microeconómicos, las cuales se pueden considerar despreciables en el modelo 1 que están en torno al 10% (salvo algunas excepciones) pero son importantes para los cálculos realizados para el modelo 4; donde el número de individuos que no verifican la condición de la interacción es mayor que para el modelo 1, (ver tablas 4.3 a 4.7).

Podemos concluir que el cálculo de las disposiciones a pagar es sensible a la consideración o no de aquellas personas para las que no se verifican las condiciones microeconómicas, puesto que el modelo no está modelizando correctamente su comportamiento y sería recomendable comprobar si efectivamente las variaciones en los cálculos son relevantes o no.

En la Tabla 4.3 se presentan los valores subjetivos del tiempo de viaje calculados para el modelo 1 y el 4 en pts⁷⁷/hora, así como la diferencia en porcentaje entre estos valores considerando como base los cálculos realizados con todos los individuos. Para el modelo 1 tenemos valores subjetivos para el transporte privado o colectivo, si el que viaja es hombre o mujer⁷⁸ y si trabaja o no, esto es, un total de ocho valores. Para el modelo 4, tenemos los valores además en función de la comodidad en el bus, obteniéndose dieciséis disposiciones a pagar.

Para las mujeres, los valores son menores que para los hombres, al igual que para los no trabajadores y para el transporte colectivo. Por ejemplo, en el modelo 1 un hombre que trabaja y viaja en transporte privado, tendríamos que estaría dispuesto a pagar 3767 pts por ahorrar una hora de su viaje frente a uno que no trabaja que solo estaría dispuesto a pagar 850 pts⁷⁹. Mientras que este mismo individuo solo estaría dispuesto a pagar 2281 pts por ahorrar una hora en transporte colectivo si trabajara y 550 pts si no trabajara. Por otro lado, una mujer que viaja en transporte privado estaría dispuesta a pagar 2282 pts por ahorrar una hora de viaje si trabaja y 401 pts si no trabaja; mientras que si viajara en transporte colectivo, 1827 pts y 315 pts, respectivamente. Como se ha comentado anteriormente, las disposiciones a pagar de los hombres son mayores que las de las mujeres, así como las de los trabajadores mayores que las de personas que no trabajan.

⁷⁷ Todos los cálculos se han realizado en pesetas por ser la unidad monetaria vigente durante la de toma de datos.

⁷⁸ Todos los valores van a estar en función del sexo porque esta variable interactúa con el coste.

⁷⁹ Estamos considerando los valores calculados sin los individuos que no verifican los supuestos que establece la teoría microeconómica.

Tabla 4.3: Valores subjetivos del tiempo de viaje

Valor subjetivo del tiempo de viaje					
Modelo 1			Todos ¹	Sin ¹	Diferencia ²
T. Privado	Mujer	No trabaja	359	401	11,88
		Trabaja	2206	2282	3,46
	Hombre	No trabaja	724	850	17,39
		Trabaja	3747	3767	0,54
	Promedio Total		1857	2061	10,97
T. Colectivo	Mujer	No trabaja	282	315	11,43
		Trabaja	1768	1827	3,3
	Hombre	No trabaja	474	551	16,09
		Trabaja	2273	2281	0,38
	Promedio total		1228	1357	10,55
Modelo 4			Todos ¹	Sin ¹	Diferencia ²
T. Privado	Mujer	No trabaja	460	534	16,17
		Trabaja	2349	2518	7,20
	Hombre	No trabaja	1169	1457	24,59
		Trabaja	5026	5162	27,08
	Promedio total		2421	2809	16,45
T. Colectivo C. Alta	Mujer	No trabaja	104	120	15,20
		Trabaja	1376	1468	6,67
	Hombre	No trabaja	204	252	23,36
		Trabaja	2048	2092	2,13
	Promedio total		977	1141	16,75
T. Colectivo C. Estándar	Mujer	No trabaja	333	384	15,20
		Trabaja	1781	1900	6,67
	Hombre	No Trabaja	652	804	23,36
		Trabaja	2652	2708	2,13
	Promedio total		1405	1618	15,14
T. Colectivo C. Baja	Mujer	No trabaja	748	862	15,20
		Trabaja	2516	2683	6,67
	Hombre	No trabaja	1464	1806	2,,36
		Trabaja	3745	3825	2,13
	Promedio total		2181	2483	13,83

¹Valores en pts/hora.²Diferencia en porcentaje sobre el valor considerando a todos los individuos.

Cuando incluimos todos los individuos obtenemos valores de la disposición a pagar por ahorros de tiempo inferiores al caso en que se eliminan a aquellos para los cuales el modelo no se ajusta bien a su comportamiento. De hecho, observando los valores de la última columna apreciamos diferencias cercanas al 25% en el modelo 4 para los hombres que no trabajan.

En la Tabla 4.4 se presentan los valores subjetivos del tiempo de caminata. Estos valores están en función del sexo y del origen residencial del entrevistado, teniendo un total de cuatro valores. Las personas residentes en Arucas tienen una mayor disposición a pagar por reducir el tiempo de caminata en un minuto que las residentes en Telde. Los valores entre los modelos no son muy distintos, siendo los valores puntuales sin distinguir clases por sexo y origen exactamente iguales.

Tabla 4.4: Valores subjetivos del tiempo de caminata

Valor subjetivo del tiempo de caminata				
Modelo 1		Todos ¹	Sin ¹	Diferencia ²
Mujer	Telde	23	25	9,82
	Arucas	82	87	6,55
Hombre	Telde	36	39	6,65
	Arucas	132	140	6,28
Promedio Total		53	57	8,48
Modelo 4		Todos ¹	Sin ¹	Diferencia ²
Mujer	Telde	27	21	-21,22
	Arucas	74	82	11,11
Hombre	Telde	34	38	9,99
	Arucas	139	151	9,01
Promedio Total		53	57	7,59

¹Valores en pts/minuto.

²Diferencia en porcentaje sobre el valor considerando a todos los individuos.

El valor subjetivo de la frecuencia resulta con signo negativo porque se trata de una variable cuya utilidad marginal es positiva. Por tanto, la disposición a pagar por aumentar la frecuencia en una unidad debe interpretarse como el valor absoluto de los valores presentados en la Tabla 4.5. Se obtuvieron valores positivos para aquellos

individuos que no verificaban las condiciones microeconómicas de la interacción. Cuando consideramos a todos los individuos, a los anteriores les asignamos valor cero para calcular el promedio de cada clase. En este caso, las clases están definidas por el sexo y la edad (menores o mayores de treinta y cinco años). Las personas mayores de treinta y cinco años estarían dispuestas a pagar más por un aumento de la frecuencia en un bus a la hora. Las diferencias entre los modelos se producen para los menores de treinta y cinco años, mientras que para los mayores las disposiciones entre modelos no varían significativamente. Las mayores diferencias comparando entre todos los individuos o sólo los que cumplen con la teoría microeconómica son importantes justamente para los menores de treinta y cinco años, especialmente en el modelo 4.

Tabla 4.5: Valores subjetivos de la frecuencia

Valor subjetivo de la frecuencia				
Modelo 1		Todos ¹	Sin ¹	Diferencia ²
Mujer	Menor de 35 años ³	-23	-29	26,52
	Mayor de 35 años	-58	-61	5,23
Hombre	Menor de 35 años	-39	-48	22,88
	Mayor de 35 años	-81	-85	4,86
Promedio Total		-51	-58	14,76
Modelo 4		Todos ¹	Sin ¹	Diferencia ²
Mujer	Menor de 35 años	-12	-19	53,21
	Mayor de 35 años	-55	-59	7,33
Hombre	Menor de 35 años	-24	-36	49,21
	Mayor de 35 años	-88	-93	6,01
Promedio Total		-45	-57	26,49

¹Valores en pts/bus a la hora.

²Diferencia en porcentaje sobre el valor considerando a todos los individuos.

³Incluye a los que tienen 35 años.

Finalmente, presentamos las disposiciones a pagar por mejoras en la comodidad. En la Tabla 4.6 tenemos las disposiciones por mejorar de comodidad baja a estándar y en la Tabla 4.7 de comodidad estándar a alta, esta última debe ser menor. Para el modelo 4, estas disposiciones se obtienen para cada individuo, pues dependen del tiempo de viaje,

ver expresión (4.16). En este caso, los hemos agrupado por sexo para calcular un valor puntual.

Tabla 4.6: Disposiciones a pagar por el cambio comodidad baja a estándar

Disposición a pagar de comodidad baja a estándar			
Modelo 1	Todos ¹	Sin ¹	Diferencia ²
Mujeres	396	431	9,01
Hombres	625	666	6,67
Promedio Total	519	559	7,70
Modelo 4	Todos ¹	Sin ¹	Diferencia ²
Mujeres	322	516	60,25
Hombres	604	652	7,85
Promedio Total	474	590	24,48

¹Valores en pts.

²Diferencia en porcentaje sobre el valor considerando a todos los individuos.

Tenemos que las mujeres en el modelo 1 están dispuestas a pagar 431 pts por no viajar de pie, frente a las 666 pts que estarían dispuestos a pagar los hombres. Mientras que en el modelo 4, las disposiciones serían de 516 pts para las mujeres y de 652 pts para los hombres. Se puede observar que la disposición a pagar de las mujeres en este modelo varía más de un sesenta por ciento con respecto al caso con todos los individuos.

Tabla 4.7: Disposiciones a pagar de comodidad estándar a alta

Disposición a pagar de comodidad estándar a alta			
Modelo 1	Todos ¹	Sin ¹	Diferencia ²
Mujeres	159	173	9,01
Hombres	251	268	6,67
Promedio Total	209	225	7,70
Modelo 4	Todos ¹	Sin ¹	Diferencia ²
Mujeres	178	196	10,11
Hombres	333	359	7,85
Promedio Total	261	285	8,94

¹Valores en pts.

²Diferencia en porcentaje sobre el valor considerando a todos los individuos.

Las disposiciones a pagar por elegir asiento (cambiar de comodidad estándar a alta) son mucho menores. En el modelo 4, una mujer estaría dispuesta a pagar sólo 196 pts por

elegir asiento frente a las 359 pts de un hombre, esto es, aproximadamente el doble que una mujer.

Se debe señalar que para especificaciones lineales de la FUIIC es posible calcular intervalos de confianza para estas disposiciones a pagar (ver Armstrong *et al*, 2001). Sin embargo, para especificaciones no lineales como es nuestro caso, tendríamos que calcular un intervalo por individuo y no está claro cómo determinarlo para el valor promedio de una clase compuesta por un número determinado de individuos.

Podemos concluir por un lado, que la inclusión de los individuos que violan las condiciones microeconómicas que establece la especificación de la interacción sesga a la baja los valores de las disposiciones a pagar, recomendándose por tanto la no consideración de los mismos; y por otro, destacar que en relación a las disposiciones a pagar obtenidas, para todos los hombres están dispuestos a pagar más que las mujeres. Además, cada una de ellas varía en función de alguna otra característica socioeconómica. El valor subjetivo del tiempo de viaje es mayor para los trabajadores que para los no trabajadores y disminuye cuando la comodidad aumenta en el caso del modelo 4. La disposición a pagar por mejoras en la frecuencia del bus tiene un mayor valor para las personas mayores de treinta y cinco años. Para el tiempo de caminata, la disposición al pago es mayor para los que residen en Arucas (origen de uno de los corredores); y finalmente, las disposiciones a pagar por mejoras en la comodidad aumentan con el tiempo de viaje en el caso del modelo 4.

4.5 Análisis de la respuesta de la demanda

En esta sección estudiamos determinadas medidas de política con el objetivo de predecir su efecto sobre la elección modal. Este análisis se realiza sólo con datos de PR como en

el epígrafe anterior; e igualmente considerando a todos los individuos de PR y eliminando a los que no verifican las condiciones microeconómicas para cada modelo estimado.

La respuesta a los cambios estudiados se calcula como el porcentaje de variación de la cuota⁸⁰ de mercado agregada de la alternativa j respecto de la situación inicial, esto es:

$$\Delta P_j = \frac{P_j^1 - P_j^0}{P_j^0} \quad (4.17)$$

donde P_j^0 es la probabilidad agregada de la alternativa j antes de la política y P_j^1 , la probabilidad agregada de la alternativa j una vez realizada la política; ambas son calculadas por enumeración muestral.

Para definir las distintas políticas se ha tomado como referencia el diseño del experimento de PD realizado en esta investigación. Así hemos estudiado mejoras en la frecuencia de los buses, incrementos en el coste de aparcamiento, y reducción de la tarifa y del tiempo de viaje del bus. Además hemos estudiado la combinación de varias de estas medidas de política.

Comparando los resultados obtenidos para las distintas políticas analizadas, con todos los individuos o eliminando los que no verifican la condición que define la interacción, encontramos que las diferencias son significativas para el modelo 1 cuando estudiamos el aumento del coste de aparcamiento y la reducción de las tarifas; mientras que la combinación de políticas presenta diferencias en todas las analizadas. Para el modelo 4, en el caso de combinación de políticas las diferencias no son significativas y si lo son

⁸⁰ En realidad no se trata de cuotas de mercado porque no conocemos el número de individuos que en la población viaja en cada modo de transporte, sino de las probabilidades de elección para nuestra muestra.

cuando analizamos las medidas de mejora de la frecuencia y de reducción del tiempo de viaje en bus individualmente. En general, los valores de las variaciones porcentuales de la elección modal son menores cuando eliminamos a los individuos que no verifican los supuestos de la teoría microeconómica. Presentaremos los resultados de las medidas de política sin considerar a estos individuos porque creemos que distorsionan el efecto de las políticas sobre la elección. En el Anexo II se presenta los resultados con todos los individuos.

La primera medida analizada es la mejora de la frecuencia en el bus. En el diseño de PD se establecieron dos niveles de mejora de la frecuencia, definida ésta en la encuesta como el intervalo de tiempo entre dos buses consecutivos. En el nivel 1 se reducía en un veinticinco y en el nivel 2 un cincuenta por ciento. Estas reducciones del tiempo entre buses implican un incremento del número de buses/hora del cincuenta y cien por ciento, respectivamente; y han sido éstas las variaciones de frecuencia estudiadas.

Las variaciones porcentuales de la probabilidad del bus son similares para ambos modelos, tanto al incrementar la frecuencia en un cincuenta como en un cien por cien. En la Tabla 4.8 se presentan, para los dos modelos, los porcentajes de variación del bus para los tres niveles de comodidad. Las diferencias entre ambos modelos no son significativas como se puede observar en la Figura 4.4. Así un incremento de la frecuencia en un cincuenta por ciento implicaría un aumento en la probabilidad de ir en bus del 6,18% si la comodidad es alta, de un 7,52% si la comodidad es estándar y 11,60% si es baja para el modelo 1. Para el modelo 4, estos porcentajes de variación serían del 6,24%, del 7,59% y del 10,15%, respectivamente.

Tabla 4.8: Porcentajes de variación de las políticas individuales

Modelo 1	Bus		
Políticas	C. Alta ¹	C. Estándar ¹	C. Baja ¹
50% de Frecuencia	6,18	7,52	11,60
100% de Frecuencia	12,37	15,17	24,54
Tarjeta Valor	5,21	6,17	8,81
Tarjeta Cliente	11,32	13,80	21,18
-10% Tiempo de viaje en bus	6,66	7,67	10,24
50% de Coste de aparcamiento	41,75	47,71	71,02
Modelo 4	Bus		
Políticas	C. Alta	C. Estándar	C. Baja
50% de Frecuencia	6,24 ¹	7,59	10,15
100% de Frecuencia	12,80	15,55	21,29
Tarjeta Valor	6,27	8,15	11,34
Tarjeta Cliente	14,26	18,52	26,11
-10% Tiempo de viaje en bus	3,87	7,36	15,38
50% de Coste de aparcamiento	38,70	40,65	46,85

¹Variaciones expresadas en porcentaje.

Respecto al incremento de la frecuencia en un cien por cien, los porcentajes de variación de la probabilidad de elegir bus son muy similares entre los modelos 1 y 4, la mayor diferencia se encuentra cuando el nivel de comodidad en el bus es baja (ver Tabla 4.8 y Figura 4.5). Así tenemos para el modelo 1 un aumento de la probabilidad de ir en bus del 12,37% si la comodidad es alta, del 15,17% si la comodidad es estándar y del 24,54% si es baja. Para el modelo 4, estos porcentajes de variación serían del 12,80%, del 15,55% y del 21,29%, respectivamente.

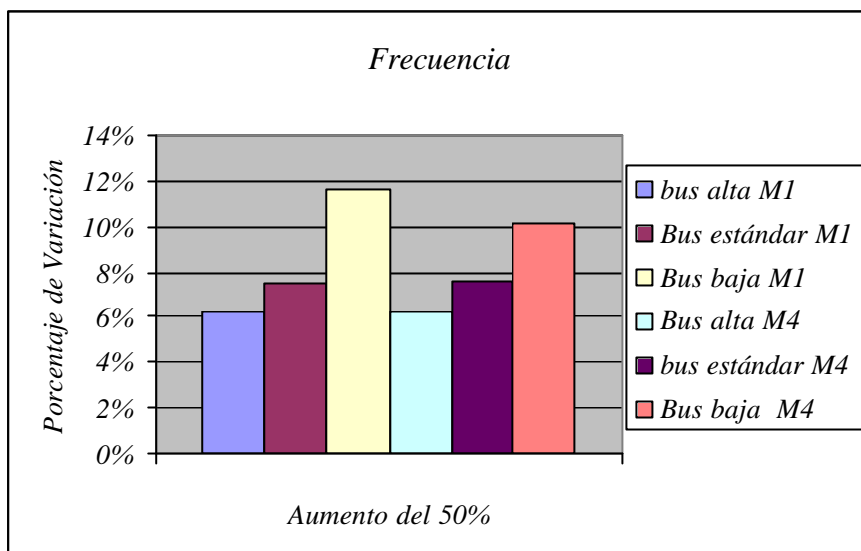


Figura 4.4: Variación de la probabilidad de elegir bus ante un aumento de la frecuencia (50%)

La segunda medida es la reducción de la tarifa del bus basándonos en la nueva definición de tarifas explicadas en el capítulo 3 (ver epígrafe 3.2.3 y Tabla 3.2). Hemos estudiado el caso de la tarjeta valor y la tarjeta cliente. La primera implica una reducción del treinta por ciento sobre la tarifa de la primera etapa y el trasbordo gratuito si la segunda etapa es urbana y un descuento del treinta por ciento si es interurbana. En cambio para la tarjeta valor, el descuento es progresivo en función del número de viajes que se realice y va del treinta al cincuenta por ciento como máximo. En este caso, hemos estudiado el supuesto del descuento máximo, es decir, una reducción en la tarifa del bus del cincuenta por ciento.

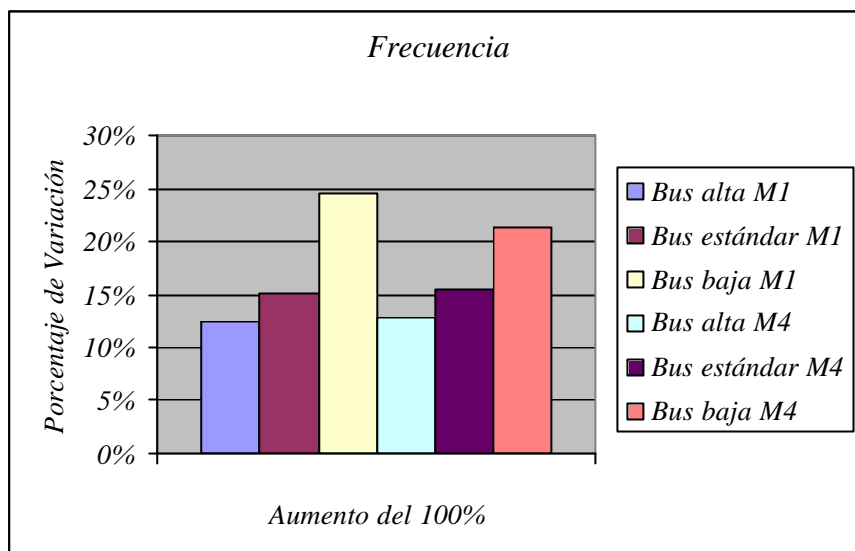


Figura 4.5: Variación de la probabilidad de elegir bus ante un aumento de la frecuencia (100%)

Las variaciones se presentan en la Tabla 4.8. Éstas son mayores para el modelo 4 que para el modelo 1 como puede observarse en la Figura 4.6 y Figura 4.7. Si comparamos estos resultados con las variaciones que se experimentan al variar la frecuencia tenemos que para el modelo 1 las reducciones en la tarifa determinan una menor variación para el bus que las derivadas del incremento de la frecuencia; mientras que estas variaciones son similares, para el modelo 4, si comparamos tarjeta valor con incrementar la frecuencia en un cincuenta por ciento. Por otro lado si se compara tarjeta cliente con un incremento de la frecuencia del cien por cien, las variaciones son un poco mayores para la reducción de la tarifa.

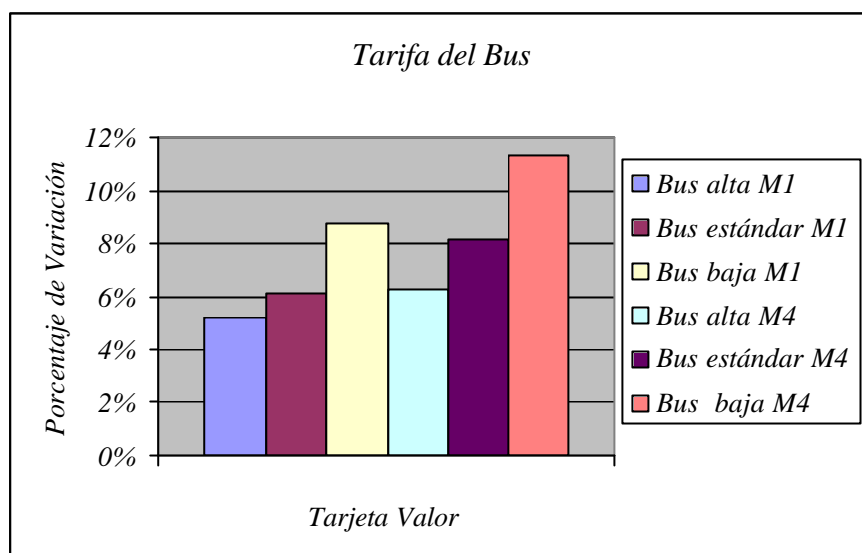


Figura 4.6: Variación de la probabilidad de elegir bus ante reducción de la tarifa (Tarjeta Valor)

La tercera política implica una reducción del tiempo de viaje en bus en un diez por ciento. Para analizar esta política consultamos a las empresas de transporte si era viable reducir los tiempos de viajes en bus si se establecían medidas como la creación de carriles bus y preferencia en los semáforos que permitieran al transporte público realizar los recorridos en menor tiempo; además se consultó si tenían algún estudio que permitiera determinar en cuánto se podrían reducir estos tiempos y si esta reducción variaba entre líneas, horas del día, etc. Las empresas no disponen de estudios sobre la reducción de los tiempos de viaje. Sin embargo, aseguran que es posible reducir los tiempos de viaje si se establecieran medidas de incentivo al transporte público; hoy sólo existen en la ciudad de Las Palmas seis kilómetros de carril⁸¹ bus y un único semáforo en el que tienen preferencia. Opinan, además, que estas medidas son viables en la práctica, aunque tal vez su implantación pueda ser más costosa que otras medidas.

⁸¹ Los kilómetros de carril bus existentes son utilizados sólo por Guaguas Municipales. Global sólo utiliza el carril bus de la Avda. de Rafael Cabrera, que no supera el kilómetro.

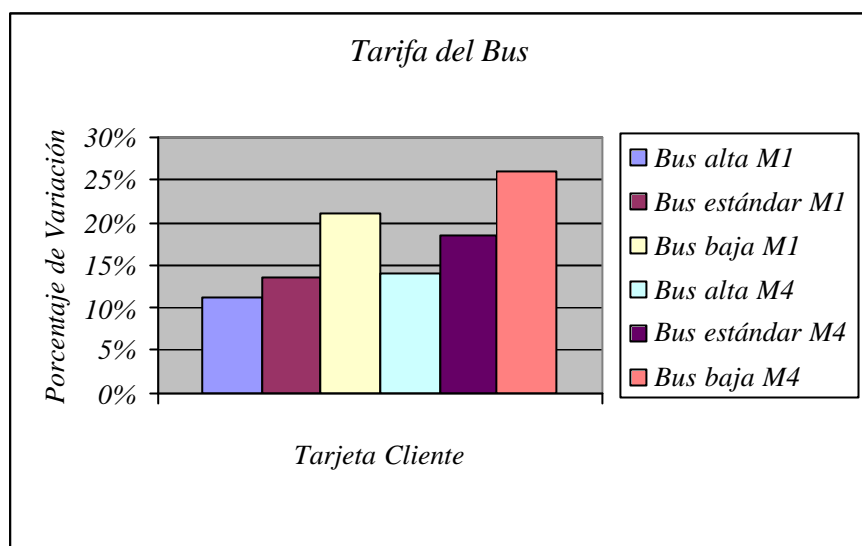


Figura 4.7: Variación de la probabilidad de elegir bus ante una reducción de la tarifa (Tarjeta Cliente)

Sin embargo, no disponían de mediciones de cuánto se reducirían estos tiempos en este momento. Nuestra intención era poder disponer de reducciones de tiempo por líneas y hora para asignar estas reducciones a aquellos individuos que se vean afectados y en la cuantía correspondiente porque estas medidas pueden afectar de distinta manera a cada línea y depender esta reducción de la hora del día.

Es por este motivo, que hemos estudiado sólo una reducción del 10% para analizar si el tiempo de viaje es tan relevante como se supone *a priori*. Analizando las diferencias de los tiempos de viaje en transporte privado y bus, si en el tiempo de viaje del transporte privado incluimos el tiempo de aparcamiento; un 16,62% de los entrevistados tiene un tiempo de viaje en transporte privado mayor que en bus y tan solo un 0,007% de éstos, elige la alternativa bus. De este análisis se puede concluir que es importante por un lado, definir políticas dirigidas a reducir los tiempos de viaje en bus para que éste se consolide como una alternativa al transporte privado y por otro, que no siempre el

tiempo de viaje es lo relevante⁸², pues este 16,62% prefiere adelantar su hora de salida con el objetivo de buscar aparcamiento aunque esto implique un tiempo de viaje total mayor que en transporte público⁸³.

Los resultados de esta política se pueden ver en la Tabla 4.8 y Figura 4.8, donde se puede observar que cuando la comodidad es estándar, ambos modelos predicen una variación similar, del 7,67% en el modelo 1 y del 7,36% en el 4; mientras que cuando la comodidad es alta o baja, las diferencias son significativas, un 6,66% y un 3,87% cuando la comodidad es alta para los modelos 1 y 4, respectivamente y un 10,24% y un 15,38% cuando la comodidad es baja.

La cuarta política analizada es el incremento del coste de aparcamiento en un cincuenta por ciento para los usuarios de transporte privado. Esta política define un aumento del coste de aparcamiento del cincuenta por ciento para los que declaran pagar coste de aparcamiento en la encuesta de PR; mientras que para los que declaran un coste cero realizamos un tratamiento similar al de la encuesta de PD para definir el coste de aparcamiento en el nivel 1 de dicha encuesta.

⁸² Una alternativa es valorada en función de un conjunto de atributos, siendo el tiempo de viaje uno de ellos.

⁸³ Este porcentaje de individuos presentan tiempos de aparcamientos elevados.

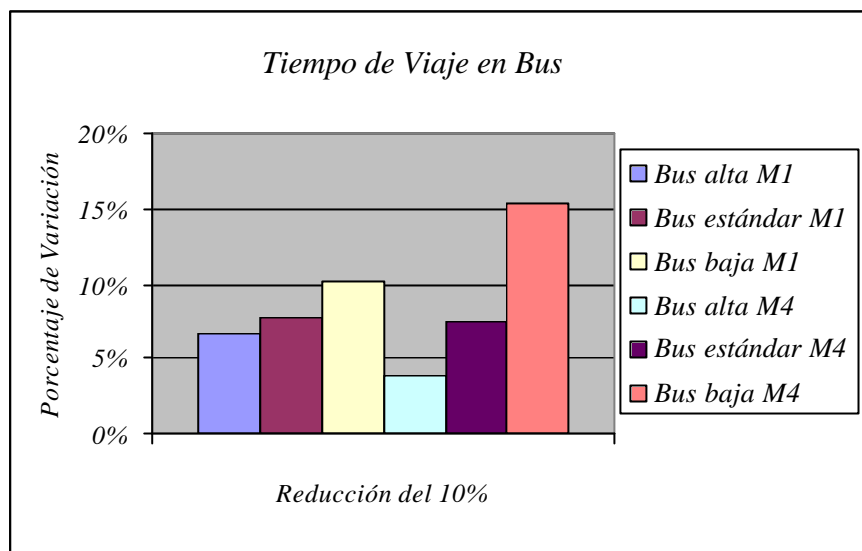


Figura 4.8: Variación de la probabilidad de elegir bus ante una reducción del tiempo de viaje (10%)

Para los que declaran aparcar en la calle se les estableció el precio del aparcamiento azul y para los que aparcan en aparcamiento privado el precio medio por hora de un aparcamiento privado. Este precio es muy similar entre los distintos aparcamientos de la ciudad. Asimismo, se distinguió por motivos de viaje de manera que a los que viajan por motivo de trabajo un mínimo de cinco veces a la semana se les asignó un coste de aparcamiento mensual igual a la media de la muestra. El coste de aparcamiento por día se obtiene como el cociente entre este valor y la frecuencia mensual⁸⁴. Las personas que viajan con una frecuencia menor son tan solo un 6,18% del total de personas que viajan por motivo de trabajo. A estas se les trató como a los que viajan por otros motivos, interpretando que en estos casos las personas no trabajan en la ciudad de Las Palmas de Gran Canaria, sino que viajan para realizar algún trámite por motivos de trabajo que no les implica estar una jornada de ocho horas. No obstante, lo correcto hubiera sido haber

⁸⁴ La frecuencia del viaje es semanal, la mensual la calculamos multiplicando ésta por las cuatro semanas del mes.

consultado esta cuestión en la encuesta de PR; sin embargo, no creemos que el error sea muy grave porque se trata de un porcentaje pequeño.

Para el resto de los motivos⁸⁵ de viaje se estableció un tiempo de aparcamiento de dos horas que es el tiempo estimado para la realización de gestiones, compras, etc. Este tiempo es el máximo en el aparcamiento azul, ya que el cobro del aparcamiento en la calle pretende regular un espacio que es escaso en la ciudad y la definición de un tiempo máximo de estacionamiento permite la rotación de usuarios de este servicio.

Esta política es en parte la que lleva a cabo la institución local, ya que en los últimos años ha aumentado el número de plazas de aparcamiento privado a la vez que se ha reducido el espacio de aparcamiento en la calle mediante la ampliación de las aceras. Con esta política indirectamente se está obligando al que habitualmente aparcaba en la calle que asuma un coste de aparcamiento si quiere seguir desplazándose en vehículo privado. El que no está dispuesto a asumir este coste adicional deberá elegir otra alternativa de transporte para sus desplazamientos.

Los resultados de esta política predicen que la probabilidad del bus aumentaría considerablemente, ver Tabla 4.8 y Figura 4.9. El modelo 1 predice que el aumento del bus sería de un 41,75%, de un 47,71% y de un 71,02% para comodidad alta, estándar y baja, respectivamente. El modelo 4, por su parte, predice que estas variaciones serán del 38,70%, del 40,65% y del 46,85%, respectivamente. En este caso, no existen importantes diferencias en función del nivel de comodidad como ocurre en el modelo 1, concretamente para el nivel de comodidad baja.

⁸⁵ Se incluye el viaje por estudios porque un 72% de éstos declaran no trabajar y en este caso, suponemos que no estarían dispuestos a pagar un aparcamiento mensual.

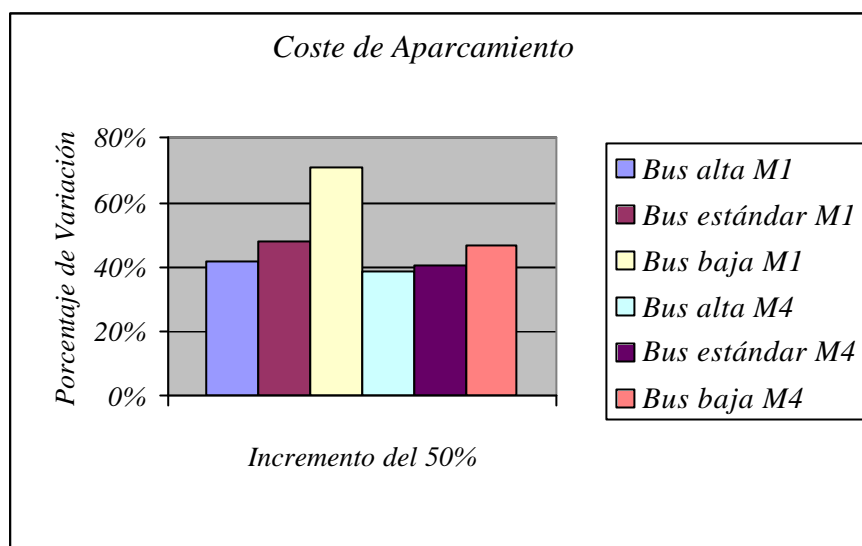


Figura 4.9: Variación de la probabilidad de elegir bus ante un incremento del coste de aparcamiento

Estos resultados revelan que una medida que encarece el transporte privado tiene un importante efecto sobre la respuesta de la demanda de transporte de pasajeros, en el sentido de que el porcentaje de variación para el bus es mayor en este caso que en el caso de las medidas de mejora del transporte público. De esto se deriva que tan importante es ofrecer un buen servicio de transporte público como encarecer el transporte privado con el objetivo de que cada modo de transporte internalice los costes reales que genera. De hecho en esta línea camina el desarrollo de las políticas de planificación del transporte como es el caso reciente de Londres, donde se cobra un peaje a los vehículos que accedan al centro; con esta medida se ha reducido los niveles de congestión en cerca de un 20%.

En cuanto a la combinación de distintas políticas, se estudiaron varias opciones; así se analizó el efecto de los aumentos de frecuencia combinadas con las reducciones de tarifa y con el incremento del coste del aparcamiento. También se estudiaron estas dos últimas conjuntamente y la combinación de las tres. Y finalmente, la combinación de estas tres con la reducción del tiempo de viaje.

En las Tabla 4.9 y Tabla 4.10 se presentan los resultados de la respuesta de la demanda cuando combinamos las distintas políticas para el modelo 1 y 4, respectivamente. Cuando se combinan más de dos políticas, ambos modelos predicen una variación positiva para el bus de más de un 50% y si la comodidad es baja, estos porcentajes superan el 100% de variación en determinadas políticas. Ahora bien, las políticas que implican la combinación de más de dos medidas distintas de manera conjunta son costosas y más difíciles de implantar, por lo que nos centraremos en el análisis de las políticas que combinan solamente dos medidas distintas.

En este caso, se observa que ambos modelos predicen un mayor aumento en la elección del bus cuando se combina el incremento del coste de aparcamiento con alguna política de mejora del transporte público (incremento de la frecuencia o reducción de la tarifa) que cuando se combinan sólo éstas. Y además, estas políticas son equivalentes en el sentido de que los porcentajes de variación son muy similares. Por ejemplo, para un aumento de la frecuencia del cincuenta por ciento combinada con un incremento en el coste de aparcamiento en la misma cuantía el modelo 1 predice que el bus aumentaría en un 49,05%, un 56,55% y un 85,76% según el nivel de comodidad sea alto, estándar o bajo (ver Tabla 4.9). Si el incremento del coste de aparcamiento se combina con una reducción de la tarifa de un treinta por ciento (tarjeta valor) las variaciones son similares, de un 47,36%, un 54,51% y un 81,99%. Para el modelo 4 las diferencias son aún menores (ver Tabla 4.10).

Tabla 4.9: Porcentajes de variación de combinaciones de políticas (1)

Modelo 1 Políticas	Bus		
	C. Alta ¹	C. Estándar ¹	C. Baja ¹
50% F. + T. Valor	12,19 ¹	14,74	22,81
50% F. + T. Cliente	19,22	23,58	38,03
100% F. + T. Valor	19,21	23,39	38,15
100% F. + T. Cliente	27,20	33,44	56,24
50% F. + 50% Cap.	49,05	56,55	85,76
100% F. + 50% Cap.	56,41	65,56	101,69
T. Valor + 50% Cap	47,36	54,51	81,99
T. Cliente + 50% Cap	54,20	63,09	97,45
50% F. + T. Valor + 50% Cap.	55,51	64,41	99,14
50% F. + T. Cliente + 50% Cap.	63,41	74,35	117,66
100% F. + T. Valor + 50% Cap.	63,75	74,46	117,36
100% F. + T. Cliente + 50% Cap.	72,72	85,77	138,97
50% F. + T. Valor + 50% Cap+ Tv.	64,31	74,78	115,64
50% F. + T. Cliente + 50% Cap. + Tv.	72,29	84,87	134,83
100% F. + T. Valor + 50% Cap.+ Tv.	72,68	85,03	134,58
100% F. + T. Cliente + 50% Cap.+ Tv.	81,72	96,49	156,89

¹Variaciones expresadas en porcentaje.

Por tanto, si el objetivo de la medida de política es que aumente el porcentaje de usuarios de bus, la política a definir sería un incremento del coste de aparcamiento combinado con una mejora del servicio de transporte público de pasajeros, bien mediante una intensificación de la frecuencia o bien mediante una reducción de las tarifas mediante la introducción de las tarjeta valor y/o cliente.

El incremento de la frecuencia tiene un coste considerable para las empresas, ya que implica un aumento de la flota y del número de trabajadores. Esto requiere que la cuota del bus aumente tras la medida y se mantenga en el tiempo para que sea viable desde un punto de vista económico para la empresa.

Tabla 4.10: Porcentajes de variación de combinaciones de políticas (4)

Modelo 4	Bus		
Políticas	C. Alta	C. Estándar	C. Baja
50% F. + T. Valor	13,40 ¹	17,03	24,55
50% F. + T. Cliente	22,57	29,00	42,49
100% F. + T. Valor	20,93	26,18	38,63
100% F. + T. Cliente	31,36	39,76	59,67
50% F. + 50% Cap.	47,27	50,82	60,29
100% F. + 50% Cap.	56,23	61,58	75,07
T. Valor + 50% Cap	46,21	50,25	60,31
T. Cliente + 50% Cap	55,72	62,51	78,09
50% F. + T. Valor + 50% Cap.	55,84	61,96	77,28
50% F. + T. Cliente + 50% Cap.	66,80	76,21	98,78
100% F. + T. Valor + 50% Cap.	65,85	74,11	95,39
100% F. + T. Cliente + 50% Cap.	78,21	90,31	120,56
50% F. + T. Valor + 50% Cap+ Tv.	62,89	74,06	102,49
50% F. + T. Cliente + 50% Cap. + Tv.	74,01	88,66	125,57
100% F. + T. Valor + 50% Cap.+ Tv.	73,27	86,82	122,01
100% F. + T. Cliente + 50% Cap.+ Tv.	85,81	103,39	148,85

¹Variaciones expresadas en porcentaje.

Si las medidas de mejora del transporte público van a ser financiadas por la administración pública, lo correcto sería determinar cuál de las dos medidas es más eficiente desde el punto de vista de la asignación de subvenciones. López (2001) investiga este tema desde una perspectiva del bienestar y concluye que las subvenciones asignadas a intensificación de frecuencia incrementan el bienestar en una cuantía mayor que la subvención vía precios.

De cualquier manera, los resultados obtenidos confirman la importancia de definir medidas que encarezcan el transporte privado, de forma que éste internalice los costes reales que genera.

Finalmente, podemos concluir que la consideración de los individuos que no verifican los supuestos microeconómicos sesga al alza el efecto de las distintas medidas de política analizada sobre la demanda, lo que implica que éstos deben ser eliminados. Señalar que el análisis realizado supone que permanecen constantes todas aquellas

variables sobre las que no se define una medida de política. No obstante, si podemos concluir que la demanda de transporte es sensible en mayor o menor medida a las políticas analizadas.

5. Conclusiones

En la planificación del transporte de pasajeros en las ciudades es necesario realizar estudios de demanda que permitan adecuar la oferta de servicios de transporte, por un lado, y realizar análisis que contribuyan a resolver los problemas derivados del crecimiento secular del parque automovilístico en el largo plazo, por otro.

El objetivo de nuestra investigación estuvo centrado en el análisis y predicción de la demanda de transporte de pasajeros en dos corredores de transporte. Entre los aspectos que nos interesaba estudiar, se encuentra el efecto producido por la integración de tarifas de transporte urbano e interurbano así como el análisis de diversas medidas de política que establecen variaciones sobre el sistema actual y que permiten ofrecer una alternativa de transporte público más atractiva.

La elección por parte del consumidor de una alternativa de transporte para realizar su viaje, se basa en el enfoque teórico de la microeconómica de las elecciones discretas desarrollado por McFadden (1981) y en la teoría de la utilidad aleatoria cuya formalización se debe a McFadden (1974).

La modelización de la demanda se basó en la metodología de estimación con datos mixtos, que combina observaciones de PR y PD. Para ello se realizaron dos encuestas; la encuesta de PR permitió obtener información sobre el comportamiento actual de los viajeros y sirvió para adaptar el diseño de PD, y la encuesta posterior, a la experiencia

particular de cada individuo. De esta manera, conseguimos que el ejercicio fuera creíble para los entrevistados. También se decidió estudiar el efecto de la comodidad, ya que esta variable latente es difícil de medir cuando se trabaja sólo con observaciones de PR. Por otra parte, decidimos construir un experimento que fuese capaz de medir las interacciones entre las principales variables de nivel de servicio para, de esta forma, decidir si éstas podían explicar de forma significativa el comportamiento del individuo frente a la elección modal. Para medir estas interacciones fue necesario utilizar un diseño factorial fraccionado de 27 escenarios que fue dividido en tres bloques con el objeto de reducir el efecto fatiga en los entrevistados. Hasta conseguir el diseño definitivo se realizaron tres pretest que permitieron disponer de información relevante en lo referente a la percepción de algunas de las variables, prestando especial atención a la comodidad.

Una conclusión correspondiente a la fase del diseño del experimento es que tanto las reuniones de grupo de focal como la realización de encuestas piloto, redundan de forma positiva en la capacidad predictiva del modelo. En nuestro caso, el especial cuidado tenido en la identificación de la comodidad nos permitió, por un lado, obtener estimaciones consistentes de las disposiciones a pagar en función de los distintos niveles de esta variable y por otro, obtener la disposición a pagar por mejoras en la comodidad en el servicio de transporte público.

Otra cuestión importante en este tipo de diseños es estudiar los posibles sesgos en las respuestas de las personas entrevistadas. Basándonos en los supuestos que establece la teoría microeconómica sobre el comportamiento del consumidor, detectamos a aquellos individuos que no verificaban estas condiciones, esto es, individuos cautivos, inconsistentes y lexicográficos, con el fin de analizar si la inclusión de los mismos

afectaba a los resultados de la modelización (Saelensminde, 1998a; 1998b). Se concluyó que los mejores resultados, utilizando sólo los datos de PD, se obtenían cuando no se consideran estos individuos, por lo que se recomienda realizar este análisis siempre que sea posible. Sin embargo, cuando se combinan estos datos con información de PR, se plantea la cuestión de si se debe incluir o no a los individuos que eligen de manera lexicográfica en la encuesta de PD. En nuestro caso, se obtuvieron mejores resultados para la estimación con datos mixtos (PR+PD) incluyendo a éstos últimos, esto es interesante ya que en general éstos individuos no son identificables en PR y su inclusión normalmente se plantea como opcional (ver por ejemplo, Rizzi y Ortúzar, 2003).

En cuanto a la modelización, encontramos que las especificaciones no lineales de la utilidad daban lugar a mejores estimaciones de la elección modal, coincidiendo con los resultados obtenidos en otros estudios previos (Hensher, 1998; Jara-Díaz y Videla, 1989; Ortúzar *et al*, 2000). En primer lugar, se estudió la existencia de efecto renta según el planteamiento desarrollado por Jara-Díaz y Videla (1989) y se verificó la existencia del mismo al dividir la población en dos estratos de renta. La renta en la elección modal se incorporó dividiendo los costes de viaje y de aparcamiento por la tasa de gasto, esto es, renta por unidad de tiempo disponible; esto considera que las horas de trabajo son una variable exógena, ya el individuo no puede decidir con flexibilidad el número de horas que trabaja (Jara-Díaz y Farah, 1987).

En segundo lugar, se obtuvieron estimaciones significativas de la interacción entre el coste del viaje y la frecuencia, concluyendo, por tanto, que el efecto de ambas variables no era aditivo. De esta forma, la inclusión del término correspondiente a la interacción dio lugar a modelos estadísticamente superiores. De acuerdo a los supuestos microeconómicos, este parámetro debe satisfacer ciertas condiciones para cada uno de

los individuos. El no cumplimiento de estas condiciones implica la violación de dichos supuestos. Esto nos llevó a cuestionarnos si la no consideración de estos individuos en las aplicaciones posteriores de los modelos (cálculos de disposiciones a pagar y predicción) afectaba de forma sustancial a los resultados. Así, realizamos el cálculo de la disposición a pagar y la predicción del efecto de las distintas políticas estudiadas, bajo dos modalidades: (i) con todos los individuos de la muestra y (ii) eliminando a los que no verificaban la condición de la interacción. Para el caso de las disposiciones a pagar, la inclusión de éstos sesga a la baja estos valores; mientras que el análisis de medidas de política, sesga al alza las variaciones en la probabilidad de elegir bus. De este análisis, concluimos que, en primer lugar, es importante comprobar que se verifican las condiciones microeconómicas para el rango de nuestra muestra y, en segundo lugar, que si no se verifica para un porcentaje mayor del diez por ciento, deben eliminarse a estos individuos de la muestra, pues su inclusión sesga los resultados.

En tercer lugar, se especificaron las variables socioeconómicas en las variables modales. Se incluyó un parámetro base para la variable modal más otro para la variable socioeconómica que interactuaba con ésta. La hipótesis sobre la que descansa esta forma de especificar las variables socioeconómicas radica en que la percepción de los atributos modales de las alternativas varía en función de las características socioeconómicas y en que éstas no tienen por qué afectar de igual manera a todos los atributos modales. En nuestro caso, encontramos una variable socioeconómica distinta para cada atributo. Esta especificación nos permitió obtener disposiciones a pagar diferentes para cada individuo. El valor promedio de la disposición a pagar se obtuvo aplicando enumeración muestral de acuerdo a las clases que determinaba cada variable socioeconómica.

Es interesante señalar que en todos los casos los hombres están dispuestos a pagar más que las mujeres, pero se debe recordar que cada disposición a pagar varía en función de alguna otra característica socioeconómica. Por otro lado, el valor subjetivo del tiempo de viaje es mayor para los trabajadores que para los no trabajadores. La disposición a pagar por mejoras en la frecuencia del bus tiene un mayor valor para las personas mayores de treinta y cinco años. Y Para el tiempo de caminata, la disposición al pago es mayor para los que residen en Arucas (origen de uno de los corredores).

En cuarto lugar, se estudiaron dos formas alternativas de especificar la comodidad: una en la que la variable se especificaba de forma aditiva y la otra interactuando con el tiempo de viaje. Esta última relaciona la percepción del tiempo de viaje en función de la comodidad del modo de transporte y la percepción de la comodidad en función del tiempo de viaje. De esta forma, el valor subjetivo del tiempo de viaje aumenta al disminuir la comodidad y las disposiciones a pagar por mejoras en la comodidad aumentan con la duración del viaje.

Desde un punto de vista metodológico, pensamos que esta investigación aporta mayor evidencia empírica sobre la utilidad de las encuestas de PD y los distintos aspectos tratados en la modelización de la demanda: estimación con datos mixtos, estudio de especificaciones no lineales, existencia de efecto renta, análisis del efecto de variables latentes, análisis de la existencia de interacciones y la inclusión de variables socioeconómicas interactuando con los atributos modales.

Como líneas abiertas de futura investigación, se pretende extender la modelización a la aplicación de modelos más flexibles como es el caso del *logit mixto* y obtener, mediante estimación bayesiana jerárquica, parámetros diferentes para cada individuo con el

objeto de comparar los resultados obtenidos de acuerdo a las distintas estrategias de modelización.

Por último, con el objeto de contrastar la robustez de las estimaciones de las disposiciones a pagar presentadas en esta memoria, se pretende construir intervalos de confianza a partir de simulaciones de la distribución de los estimadores de los parámetros.

Referencias bibliográficas

Allenby, G. (1997). An introduction to hierarchical Bayesian modeling, *Tutorial notes, Advanced Research Techniques Forum*, America Marketing Association.

Allenby, G. y Rossi, P. (1999). Marketing models of consumer heterogeneity. *Journal of Econometrics* 8, 13-21.

Aptech Systems, (1994). *GAUSS User's Manual* (Mapple Valley, USA).

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

Bates, J. y Roberts, M. (1986). Value of time research: summary of methodology and findings. *14th PTRC Summer Annual Meeting*, University of Sussex, Inglaterra, 14-18 de Julio.

Becker, G. (1965). A theory of the allocation of time. *The Economic Journal* 75, 493-517.

Ben-Akiva, M. y Bolduc, D. (1996). Multinomial probit with a logit kernel and general parametric specification of the covariance structure. *Working Paper*, Department of Civil Engineering, MIT.

Ben-Akiva, M. y Lerman, S.R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. The MIT Press, Cambridge. Mass.

Ben-Akiva, M.E. y Morikawa, T. (1990). Estimation of travel demand models from multiple data sources. *Proceedings 11th International Symposium on Transportation and Traffic Theory*, Yokohama, Japón.

Bernstein, S. (1917). *Calcul des Propabilités*.

Bierlaire, M. y Vandevyvere, Y. (1995). *HieLoW: The Interactive User's Guide* (Namur: Transportation Research Group, FUNDP).

Bojórquez, E. (2002). Generación de variables de nivel de servicio para un modelo conjunto de viajes y asignación de tiempo a actividades. Memoria para optar al título de Ingeniero Civil. Departamento de Ingeniería Civil, Universidad de Chile.

Boletín de la Unión Europea (1995)., suplemento 4/95: Una red para los ciudadanos.

Bradley, M.A. y Daly, A. J. (1994). Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data. *Transportation* 21, 167-184

Bradley, M.A. y Daly, A. J. (1997). Estimation of logit choice models using mixed stated preference and revealed preference information. En P.R. Stopher y M. Lee-Gosselin (eds.), *Understanding Travel Behaviour in an Era of Change*, Pergamon, Oxford.

Bradley, M.A. y Kroes, E. (1990). Forecasting issues in stated preference survey research. *69th TBR Annual Meeting*, Washington DC, January 1990, USA.

Carrasco, J.A. y Ortúzar, J. de D. (2002). Review and assessment of the nested logit model. *Transport Reviews* 22, 197-218.

Carson, R.T., Louviere, J. J., Anderson, D.A., Arabie, P., Bunch, D.S., Hensher, D.A., Jonson, R.M., Kuhfeld, W.F., Steinberg, D., Swait, J., Timmermans, H y Wiley, J.B. (1994). Experimental analysis of choice. *Marketing Letters* 5, 351-368.

Cherchi, E. (2000). Derivazione del Valore del Tempo Utilizando Dati Misti RP/SP con Funzioni di Utilità Non Lineari. Tesi di Dottorato, Facoltà di Ingegneria, Università di Cagliari, Italia.

Cherchi, E. (2003). Comunicación privada.

Cherchi, E. y Ortúzar, J. de D. (2002). Mixed RP/SP models incorporating interaction effects: modelling new suburban train services in Cagliari. *Transportation* 29, 371-395.

Chib, S. y Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm, *American Statistician* 49, 327-335.

Daganzo, C.F. (1979). Multinomial Probit: The Theory and its Applications to Demand Forecasting. Academic Press, New York.

Daly, A. (1987). Estimating “tree” logit models. *Transportation Research* 21B, 251-268.

Daly, A. (1992). *ALOGIT 3.2 User's Guide*. Hague Consulting Group, La Haya.

Daly, A. y Zachary, S. (1978). Improved múltiple choice models. En D.A. Hensher y M. Q. Dalvi (eds), *Determinants of Travel Choice*, Saxon House, Westmead.

Daly, A.J. y Gunn, H.F. (1986). Cost effective methods for national level demand forecasting. En A. Ruhl (ed); *Behavioural Research for Transport Policy*. VNU Science Press, Utrecht.

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

DeSerpa, A. (1973). Microeconomic theory and the valuation of travel time: some clarifications. *Regional Urban Economics* 2, 401-410.

Domencich, T. A. y McFadden, D. (1975). *Urban Travel Demand. A Behavioral Analysis*. North Holland, Amsterdam.

Economic Software, Inc. (1995). *LIMDEP, Version 7.0* (Bellport, USA).

Edei Consultores, S.A (1998). Plan de Coordinación de la Oferta en el Área Metropolitana de Las Palmas de Gran Canaria. *Informe Final*, Las Palmas de Gran Canaria.

ESRI Inc (1996). ArcView GIS. *The Geographic Information System for Everyone*.

González, R.M. (1995). La Demanda de Transporte de Pasajeros en el Trayecto Gran Canaria-Tenerife. Una Aplicación de los Modelos de Elección Discreta. Tesis doctoral. Departamento de Economía y Dirección de Empresas. Universidad de La Laguna, España.

González, R.M. (1997). The value of time: a theoretical review. *Transport Reviews* 17, 245-266.

GRAFSCAN, S.A. (1998). *Canarymap 4.1, Manual de usuario*.

Hensher, D. A. (1998). Extending valuation to controlled value functions and non uniform scaling with generalised unobserved variances. En T. Garling, T. Laitila K. and Westin (eds), *Theoretical Foundations of Travel Choice Modeling*. Pergamon Press.

Hensher, D. A. y Greene, W. (2001). The mixed logit model: the state of practice and warnings for the unwary. *Working Paper*, School of Business, The University of Sidney.

Hunt, B.R., Lipsman, R.L. y Rosenberg, J.M. (2001) *A guide to Matlab: For Beginners and Experienced Users*. Cambridge University Press, Cambridge.

IPD Española (1994). Encuesta domiciliaria Las Palmas de Gran Canaria, *Memoria del trabajo*, Ministerio de Obras Públicas, Transportes y Medio Ambiente.

Jara-Díaz, S y Farah, M. (1987). Transport demand and user's benefits with fixed income: the goods/leisure trade-off revisited. *Transportation Research 21B*, 165-170.

Jara-Díaz, S. (2001). Comunicación privada.

Jara-Díaz, S. y Videla, J. (1989). Detection of income effect in mode choice: theory and application. *Transportation Research*, 23B, nº 6, 393-400.

Jara-Díaz, S., Munizaga, M.A. y Palma, C. (2002). Generación de una base de datos para modelación conjunta de asignación de tiempo a actividades y viajes. *Memorias del XII Congreso Panamericano de Ingeniería de Tránsito y Transporte*, Quito-Ecuador.

Jara-Díaz, S.R. (1998). "Time and income in travel choice: towards a microeconomic activity-based theoretical framework" En T. Garling, T. Laitila y K. Westin (eds): *Theoretical Foundations of Travel Choice Modelling*. Elsevier Science, New York.

Kocur, G., Adler, T., Hyman, W. y Aunet B. (1982). Guide to forecasting travel demand with direct utility assessment. *Report No UMTA-NH-11-0001-82*, Urban Mass Transportation Administration, US Department of Transportation, Washington, DC.

Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy* 74, 132-157.

Lantada, N. y Nuñez, M.A. (2002). *Sistemas de Información Geográfica. Prácticas con ArcView*. Ediciones Universitat Politècnica de Catalunya.

Laplace, P. (1820). *Théorie Analytique des Probabilités*, 3rd ed., Paris.

Lerman, S. y Manski, C. (1981). On the use of simulated frequencies to approximate choice probabilities. In C. Manski y D. McFadden (eds.) *Structural Analysis of discrete Data with Econometric Applications*. MIT Press, Cambridge, Mass.

López, F. (2001). Regulación y Subvenciones en el Transporte Público: Una Aplicación al caso de Canarias. Tesis doctoral, Departamento de Análisis Económico Aplicado, Universidad de Las Palmas de Gran Canaria, España.

Louviere, J.J. (1988). *Analysing Decision Making: Metric Conjoint Analysis*. Sage Publications, Newbury Park.

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

Manski, C. (1977). The structure of random utility models. *Theory and Decision* 8, 229-254.

Mayberry, J. P. (1973). Structural requirements for abstract-mode models of passenger transportation. En R.E. Quandt (ed.), *The Demand for Travel: Theory and Measurement*. D.C. Health and Co., Lexington.

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

McFadden, D. (1981). Econometric models of probabilistic choice. En C. Manski y D. McFadden (eds), *Structural Analysis of Discrete Choice Data*. MIT Press, Cambridge, Mass.

McFadden, D. (1989). A method of simulated moments for estimation of discrete choice response models without numerical integration. *Econometrica* 57, 995-1026.

McFadden, D. y Reid, F.A. (1975). Aggregate travel demand forecasting from disaggregate behavioural models. *Transportation Research Record* 534, 24-37.

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

Ortúzar, J. de D. (1983). Nested logit models for mixed-mode travel in urban corridors. *Transportation Research* 17A, nº 4, 283-299.

Ortúzar, J. de D. (2001). On the development of the nested logit model. *Transportation Research* 35B, 213-216.

Ortúzar, J. de D. y Garrido, R.A. (2000). Rank, rate or choice?: an evaluation of SP methods in Santiago. En J. de D. Ortúzar (ed), *Stated Preference Modelling Techniques*, Perspectives 4, PTRC, Londres.

Ortúzar, J. de D. e Iacobelli, A. (1998). Mixed modelling of interurban trips by coach and train. *Transportation Research* 32A, 345-357.

Ortúzar, J. de D. e Ivelic, A.M. (1987). Effects of using more accurately measured level of service variables in the specification and stability of mode choice models. *Proceeding 15th PTRC Summer Annual Meeting*, P290,117-130. PTRC, Londres.

Ortúzar, J. de D. y Willumsen, L. G. (2001). *Modelling Transport*. 3^a edición, John Wiley & Sons, Chichester.

Ortúzar, J. de D., Achondo, F.J. y Ivelic, A.M. (1987). Sequential and full information estimation of hierarchical logit models: some new evidence. *11th Triennial Conference on Operations Research*, Buenos Aires, Argentina.

Ortúzar, J. de D., Roncagliolo, D.A. y Velarde, U.C. (2000). Interactions and independence in stated preference modelling. En J. de D. Ortúzar (eds), *Stated Preference Modelling Techniques*, Perspectives 4, PTRC, Londres.

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

Revelt, D. y Train, K. (2000). Specific taste parameters and mixed logit. *Working Paper, N° E00-274*, Department of Economics, University of California, Berkeley.

Richardson, A.J., Ampt, E.H. y Meyburg, A.M. (1995). *Survey Methods for Transport Planning*. Eucaliptus Press, University of Melbourne, Australia.

Rizzi, L.I. y Ortúzar, J. de D. (2003). Stated preference in the valuation of interurban road safety. *Accident Analysis and Prevention* 35, 9-22.

Saelensminde, K. (1998a). Causes and consequences of lexicographic choices in stated choice studies. *Working Paper*, Institute of Transport Economics: Oslo.

Saelensminde, K. (1998b). The impact of choice inconsistencies on the valuation travel time in stated choice studies. *Working Paper*, Institute of Transport Economics: Oslo.

Saelensminde, K. (1999). *Valuation of Nonmarket Goods for Use in Cost-Benefits Analyses: Methodological Issues*. PhD Thesis, Department of Economics and Social Sciences, Agricultural University of Norway.

Saelensminde, K. (2001). Comunicación privada.

SawtoothSoftware (1999). The CBC/HB module for hierarchical Bayes at www.sawtoothsoftware.com

Sillano, M. y Ortúzar, J. de D. (2002). WTP estimation with mixed logit models: some new evidence. *Mimeo*, Departamento de Ingeniería de Transporte, Pontificia Universidad Católica de Chile.

Sobel, K. (1980). Travel demand forecasting by using the nested multinomial logit model. *Transportation Research Record* 775, 45-55.

Swait, J. y Louviere, J. (1993). The role of the scale parameter in the estimation and use of multinomial logit models, *Journal Marketing Research* 30, 305-314.

Swait, J., Louviere, J.J. y Williams, M. (1994). A sequential approach to exploiting the combined strengths of SP and RP data: application to freight shipper choice. *Transportation* 21, 135-152.

Tardiff (1976). A note on goodness-of-fit statistics for probit and logit models. *Transportation* 5, 377-388.

Train, K. (2001). A comparison of hierarchical Bayes and maximum simulated likelihood for mixed logit. *Working paper*, Department of Economics, University of California, Berkeley.

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

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

Tverski, A. (1972). Elimination by aspects: a theory of choice. *Psychological Review* 79, 281-299.

von Mises, R. (1931). *Wahrscheinlichkeitsrechnung*, Springer-Verlag, Berlín.

Williams, H. C. W. L. (1977). On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning* 9A, 167-219.

Wittink, D., Krishnamurthi, L. y Nutter, J. (1982). Comparing derived importance weights across attributes. *Journal of Consumer Research* 8, 471-474.

Anexos

Anexo I

Encuesta de Preferencias Reveladas

Encuesta de Preferencias Declaradas

Anexo II

Otros resultados:

Análisis de efecto renta

Análisis de políticas con todos los individuos

Tabla II.1: Estratificación por ingresos

Parámetros		Modelo 5		Modelo 6	
		1 ^{er} estrato	2 ^o estrato	1 ^{er} estrato	2 ^o estrato
$C_{conductor}$	q_{cmc}	5,394 ¹ (5,9) ²	3,881 (4,3)	5,983 (5,4)	4,562 (1,9)
$C_{acompañante}$	q_{cma}	-0,8028 (-2,1)	-1,669 (-3,1)	-0,559 (-1,2)	-1,604 (-2,7)
$C_{conductor}$	q_{cmcc}	2,087 (4,0)	0,6468 (1,1)	2,333 (4,4)	0,9786 (1,3)
T_{viaje}	q_{iv}	-0,01463 (-1,5)	-0,07848 (-4,6)	-0,003733 (-0,2)	-0,07571 (-4,3)
$Coste$	q_c	-0,01623 (-2,8)	-0,01336 (-2,2)	-0,02129 (-2,2)	-0,01882 (-0,9)
$Coste\ al\ cuadrado$	q_{c^2}	0,1929e-4 (1,7)	0,1814e-4 (1,8)	0,2991e-4 (1,4)	0,2701e-4 (0,7)
$C_{aparc.}$	q_{cap}	-0,003678 (-2,7)	-0,002060 (-2,4)	-0,004041 (-2,8)	-0,00211 (-2,1)
T_{cam}	q_{tcam}	-0,06017 (-1,3)	-0,2409 (-2,9)	-0,05702 (-1,2)	-0,2576 (-2,4)
$Frecuencia$	q_f	0,09979 (2,1)	0,04236 (0,8)	0,1247 (1,9)	0,06815 (0,4)
$Comodidad\ Baja$	q_{cB}	-1,256 (-2,1)	-2,315 (-2,7)	-	-
$Comodidad\ estándar$	q_{cE}	-0,2826 (-1,0)	-0,981 (-2,0)	-	-
Tv_{cBaja}	$q_{iv_{cB}}$	-	-	-0,04272 (-1,2)	-0,09496 (-0,8)
$Tv_{cEstándar}$	$q_{iv_{cE}}$	-	-	-0,01047 (-0,8)	-0,03063 (-0,6)
$Factor\ de\ escala$	m	1,288 (2,1) [0,38] ³	0,8777 (2,7) [0,31]	0,8003 (1,2) [0,31]	0,5576 (0,7) [0,54]
$r^2(C)$		0,5298	0,6236	0,5265	0,6191
$l(\hat{q})$		-313,2192	-250,7391	-315,4613	-253,7344
$N^\circ\ de\ observaciones$		672	614	672	614

¹Valor del parámetro estimado.

²t estadístico

³test del factor de escala con respecto a uno.

**Tabla II.2: Porcentaje de variación de políticas individuales
con todos los individuos**

Modelo 1	Bus		
Políticas	C. Alta	C. Estándar	C. Baja
50% de Frecuencia	6,83	8,22	11,77
100% de Frecuencia	14,19	17,41	25,69
Tarjeta Valor	12,52	14,62	19,51
Tarjeta Cliente	26,67	31,68	44,09
-10% T. viaje en bus	6,74	7,31	8,54
50% de C. aparcamiento	64,26	73,33	103,98
Modelo 4	Bus		
Políticas	C. Alta	C. Estándar	C. Baja
50% de Frecuencia	10,76	14,24	20,32
100% de Frecuencia	14,79	19,57	29,05
Tarjeta Valor	10,85	12,92	15,93
Tarjeta Cliente	22,67	27,86	35,70
-10% T. viaje en bus	10,76	16,89	29,38
50% de C. aparcamiento	41,29	45,63	56,19

¹En porcentaje.

Tabla II.3: Porcentaje de variación para la combinación de políticas con todos los individuos-modelo 1.

Modelo 1	Bus		
Políticas	C. Alta	C. Estándar	C. Baja
50% F. + T. Valor	21,06	25,14	36,27
50% F. + T. Cliente	37,21	44,89	66,56
100% F. + T. Valor	30,10	36,37	55,64
100% F. + T. Cliente	48,24	58,80	92,06
50% F. + 50% Cap.	73,92	85,13	121,95
100% F. + 50% Cap.	84,07	97,81	142,77
50% Cap + T. Valor	78,92	91,05	130,85
50% Cap + T. Cliente	95,45	112,03	166,81
50% F. + T. Valor + 50% Cap.	90,56	105,59	155,31
50% F. + T. Cliente + 50% Cap.	109,40	129,82	198,92
100% F. + T. Valor + 50% Cap.	102,58	120,81	182,83
100% F. + T. Cliente + 50% Cap.	123,64	148,21	234,43
50% F. + T. Valor + 50% Cap+ Tv.	102,78	119,96	176,72
50% F. + T. Cliente + 50% Cap. + Tv.	121,98	144,86	222,37
100% F. + T. Valor + 50% Cap.+Tv.	115,26	135,89	205,93
100% F. + T. Cliente + 50% Cap.+ Tv.	136,69	163,97	259,73

¹En porcentaje.

Tabla II.4: Porcentaje de variación para la combinación de políticas con todos los individuos-modelo 4

Modelo 4	Bus		
Políticas	C. Alta	C. Estándar	C. Baja
50% F. + T. Valor	16,00	19,61	26,51
50% F. + T. Cliente	29,35	36,48	49,67
100% F. + T. Valor	21,52	26,59	37,90
100% F. + T. Cliente	36,42	45,39	64,41
50% F. + 50% Cap.	46,98	52,67	65,98
100% F. + 50% Cap.	53,03	60,25	76,97
50% Cap + T. Valor	53,40	60,18	74,73
50% Cap + T. Cliente	66,30	76,89	98,54
50% F. + T. Valor + 50% Cap.	60,57	69,13	88,15
50% F. + T. Cliente + 50% Cap.	75,29	88,25	116,19
100% F. + T. Valor + 50% Cap.	68,07	78,48	102,57
100% F. + T. Cliente + 50% Cap.	84,54	99,96	134,82
50% F. + T. Valor + 50% Cap+ Tv.	67,17	81,33	114,74
50% F. + T. Cliente + 50% Cap. + Tv.	82,06	100,93	144,96
100% F. + T. Valor + 50% Cap.+Tv.	74,96	91,17	130,34
100% F. + T. Cliente + 50% Cap.+ Tv.	91,62	113,17	164,93

¹En porcentaje