



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA

MODELOS DE ELECCIÓN DE LA HORA DE INICIO DE VIAJE

JULIÁN ALBERTO ARELLANA OCHOA

Tesis para optar al grado de Doctor en Ciencias de la Ingeniería.

Profesores Supervisores:

JUAN DE DIOS ORTÚZAR SALAS

LUIS IGNACIO RIZZI CAMPANELLA

Santiago de Chile, Agosto, 2012

© 2012, Julián Arellana



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA

MODELOS DE ELECCIÓN DE LA HORA DE INICIO DE VIAJE

JULIÁN ALBERTO ARELLANA OCHOA

Tesis presentada a la Comisión integrada por los profesores:

JUAN DE DIOS ORTÚZAR SALAS

LUIS IGNACIO RIZZI CAMPANELLA

SERGIO JARA DÍAZ

ALEJANDRO TUDELA ROMÁN

JUAN ENRIQUE COEYMANS AVARIA

STEPHANE HESS

CRISTIÁN VIAL EDWARDS

Para completar las exigencias del grado de
Doctor en Ciencias de la Ingeniería

Santiago de Chile, Agosto, 2012

AGRADECIMIENTOS

Han sido muchas las personas que de alguna manera han compartido conmigo y han contribuido al desarrollo de esta tesis. Si de repente olvido mencionar algún nombre, les pido me disculpen pues no es intencional y los que me conocen saben de mis despistes. De todas maneras, sepan todos que llevan un lugar especial en mi corazón.

Quiero iniciar agradeciendo a mis dos excelentes tutores de tesis, Juan de Dios y Luis. Es un orgullo haber trabajado con dos personas de tan altas cualidades. Muchas gracias por todo, por ser guías, por ser amigos, por ser padres. En especial, gracias por todos los consejos, por el apoyo en los momentos más difíciles, y por ser una fuente de inspiración constante en la realización de este trabajo. Definitivamente, todo lo que aprendí de ustedes a nivel académico y personal va mucho más allá de lo que pueda escribir aquí.

A mis co-supervisores en Inglaterra, Stephane Hess y Andrew Daly, quienes me acogieron con mucho cariño y dedicación durante mi visita a la Universidad de Leeds. Gracias por toda su ayuda y colaboración. A los miembros del Comité de Tesis, quienes me ayudaron a enfocar el trabajo en sus inicios y además entregaron valiosos aportes para el desarrollo de esta investigación.

A Doina Olarú, por ser una fuente de inspiración inconmensurable y por haberme dado el privilegio de conocerla y trabajar con ella. A Víctor Cantillo, por apoyarme en algunas etapas de la tesis, y acercarme a mi actual trabajo y a mi querida ciudad.

A mis compañeros y amigos de la Pontificia Universidad Católica de Chile (PUC), y del *Institute for Transport Studies* (ITS) de la Universidad de Leeds. Gracias por compartir conmigo esta maravillosa experiencia de vida. Especial gratitud con mis compañeros de oficina: Pao Light y Reinel en la PUC; Qian Fu (alias Mike), Ian, y Fahmi, en el ITS. Gracias por soportar mis vallenatos y mis ocurrencias. Definitivamente hicieron que todos los días valieran la pena. Les tengo un inmenso cariño, amigos!.

A todos quienes de una u otra forma participaron en la etapa de recolección de datos: a las empresas que nos permitieron encuestar a sus trabajadores; a los encuestados por su disposición de colaborar; a mis excelentes encuestadoras, quienes hicieron un muy buen trabajo, con mucho amor y dedicación; a la empresa Metro S.A., a Transantiago, y al Ministerio de Transporte, por su valiosa colaboración suministrando datos valiosos para el

cálculo de los niveles de servicio; al DICTUC, en especial a Francisco Fresard, Paula Iglesias, Iván Nuñez, y Angélica Fuentes, quienes colaboraron con información y con el préstamo de los equipos GPS para la medición de niveles de servicio; por último, a los conductores de taxi quienes accedieron a llevar el GPS e hicieron posible las mediciones de los niveles de servicio en la red de transporte privado. A Pedro Lizana, Karen Díaz, Francisca Alvarez, Pilar Quezada, John Ríos, Felipe Zúñiga y Pilar Walker quienes participaron en diversas etapas para la elaboración de la base de datos de la tesis. Gracias por el excelente trabajo de cada uno de ustedes, fue fundamental.

Al Departamento de Transporte y Logística de la PUC por haberme acogido y siempre brindarme una sonrisa: a los profesores y personal técnico; y en especial, al personal administrativo: Ann Lore, Pía, Carla, Kathy, Ceci María, y Juvenal.

A la Dirección de Investigación, Innovación y Postgrado de la Escuela de Ingeniería: en especial a Fernanda Kattan, Danisa Herrera y Debbie Meza, por apoyarme con todos los aspectos administrativos y también regalarme sonrisas.

A los patrocinadores que brindaron su apoyo económico para que este proyecto fuera posible: Universidad del Norte, COLCIENCIAS, LASPAU, Escuela de Ingeniería de la PUC, VRI de la PUC, Instituto Milenio en Sistemas Complejos de Ingeniería, FONDECYT (Proyecto 1100483) y el Proyecto TEMPO del *Institute of Transport Economics* (TOI) de Oslo.

A Chile y los chilenos, por haberme acogido durante el desarrollo de mi tesis y por permitirme conocer de cerca su linda cultura.

A mis nuevos amigos colombianos en Chile, quienes representaron una familia para mí y mi esposa, y siempre estuvieron cerca para que nuestra estancia fuera placentera. Un gusto compartir con cada uno de ustedes, espero verlos pronto.

Finalmente, deje de último quizás a los más importantes. A Dios, pues sin él no es posible lograr nada.

A la flaca, mi esposa, por su gran amor, soporte y sacrificio. Sin ella nada de esto hubiera sido posible. Las dificultades y la experiencia de vida en ambos países nos ayudo a crecer como familia.

A mi familia y amigos de siempre, por siempre estar para mí, confiar en mí.

A todos, mil gracias, los llevaré en mi corazón.

A Dios por permitirme vivir
esta maravillosa
experiencia. A mi esposa,
familia y amigos, por su
apoyo constante e
incondicional.

ÍNDICE GENERAL

Pág.

DEDICATORIA	ii
AGRADECIMIENTOS	ii
RESUMEN	xii
ABSTRACT	xiv
1. INTRODUCCIÓN	1
1.1 Motivación	1
1.2 Objetivos y alcances	2
1.3 Principales contribuciones	3
1.4 Estructura del contenido.....	5
2 ANTECEDENTES GENERALES	7
2.1 Enfoque de modelación tradicional	7
2.1.1 ¿Tiempo como variable discreta o como variable continua?	12
2.1.2 Tipos de datos para modelar la elección de hora.....	17
2.2 Relación de la elección de hora de inicio de viaje con otras elecciones de transporte	19
2.3 Factores que influyen la elección de hora de inicio de los viajes	22
2.4 Conclusiones (resumen de limitaciones encontradas)	28
3 DISEÑO Y METODOLOGIA PARA LA RECOLECCION DE DATOS	31
3.1 Actividades previas al diseño de la encuesta	31
3.2 Características generales del instrumento de diseño	33
3.3 Selección de la muestra y estrategia de recolección	35
3.4 Fuentes de información.....	37
3.4.1 Preferencias reveladas.....	37
3.4.2 Obtención de niveles de servicio en la red de transporte.....	40
3.4.3 Preferencias declaradas	44
3.4.4 Variables latentes.....	61
4 DESCRIPCION DE LA MUESTRA.....	66
4.1 Características socioeconómicas	66
4.2 Características de los empleos.....	67

4.3	Descripción de las elecciones de los encuestados	68
4.3.1	Modo de transporte	69
4.3.2	Hora de inicio de los viajes	70
5	MODELOS DE ELECCIÓN DE LA HORA DE INICIO DE VIAJE CON DATOS DE PD	72
5.1	Logit Multinomial (MNL).....	72
5.2	Incorporando heterogeneidad en los gustos	76
5.3	Efecto panel.....	79
5.3.1	Enfoque tradicional para incorporar el efecto panel.....	81
5.3.2	Enfoques alternativos para incorporar el efecto panel.....	81
5.3.3	Efecto panel heterogéneo	83
6	MODELOS HIBRIDOS DE ELECCIÓN DE LA HORA DE INICIO DE VIAJE.....	85
6.1	Formulación de modelos híbridos de elección discreta	85
6.2	Estimación de modelos híbridos.....	86
6.3	Modelos híbridos de elección de hora de inicio de viaje usando datos PD y variables latentes.....	90
7	MODELO CONJUNTO DE ELECCIÓN DE MODO Y HORA	96
7.1	Formulación del modelo conjunto modo - hora	96
7.2	Análisis de resultados del modelo conjunto modo y hora	100
8	CONCLUSIONES	105
9	BIBLIOGRAFÍA	111
10	A N E X O S	118
	ANEXO A: Listado de atributos presentados en encuesta cualitativa previa	119
	ANEXO B: Cartas y formularios utilizados en la encuesta	120
	ANEXO C: Estimaciones de los modelos con efecto panel.....	129
	ANEXO D: Estimaciones de los modelos híbridos	133

ÍNDICE DE TABLAS

Tabla 3-1: Estrategia de recolección de datos	38
Tabla 3-2: Resultados de las regresiones lineales para corregir tiempos de espera	44
Tabla 3-3: Nivel de atributos y valores de parámetros <i>a priori</i> en el diseño genérico	54
Tabla 3-4: Niveles de variación del tiempo de viaje usados en la segunda parte del diseño	57
Tabla 3-5: Valor del costo por periodo para la segunda parte del diseño del experimento 1 ...	58
Tabla 3-6: Valor de los parámetros <i>a priori</i> utilizados en la optimización final.....	60
Tabla 3-7: Resultados estimación usando experimento PD para viajes al trabajo	61
Tabla 3-8: Promedio de respuestas cuestionario de variables latentes (N = 405)	62
Tabla 3-9: Resultados seleccionados del modelo MIMIC propuesto (N = 405).....	64
Tabla 4-1: Estadísticos de viajes después del trabajo para la muestra ampliada.....	69
Tabla 4-2: Disponibilidades y elecciones de modo viajes basados en el hogar	69
Tabla 5-1: Resultados de estimación para modelo MNL conjunto	75
Tabla 5-2: Resultados de estimación para modelos conjuntos	77
Tabla 5-3: Resultados de estimación con datos de la muestra definitiva.....	80
Tabla 5-4: Bondad de ajuste para modelos que incluyen efecto panel	83
Tabla 5-5: Comparación de resultados estimación efecto panel heterogéneo y panel fijo	84
Tabla 6-1: Bondad de ajuste para modelos híbridos estimados.....	92
Tabla 6-2: Resultados más importantes del mejor modelo híbrido	94
Tabla 7-1: Alternativas consideradas en la componente de elección de modo para <i>tours</i> basados en el trabajo	98
Tabla 7-2: Constantes específicas incluidas en el modelo conjunto de modo y hora.....	101
Tabla 7-3: Modelo conjunto de elección de modo y hora.....	102
Tabla A-1: Atributos presentados en la encuesta preliminar por tipo de viaje	119
Tabla C-1: Constantes de los modelos con efecto panel.....	129
Tabla C-2: Coeficientes de los modelos para efecto panel	130
Tabla C-3: Constantes de los modelos estimados para comparar efecto panel fijo con efecto panel heterogéneo	131
Tabla C-4: Coeficientes de los modelos para comparar efecto panel fijo con efecto panel heterogéneo.....	132

Tabla D-1: Constantes de los modelos híbridos	133
Tabla D-2: Coeficientes de la componente de elección de hora - modelos híbridos.....	134

ÍNDICE DE FIGURAS

Figura 3-1: Ubicación de los hogares y destinos de los encuestados	36
Figura 3-2: Ejemplo de presentación del experimento PD enfocado en viajes hacia el trabajo para usuarios de transporte público	46
Figura 3-3: Ejemplo de la presentación del experimento PD enfocado a toures basados en el trabajo para usuarios de transporte público	46
Figura 3-4: Ilustración de alternativas en el experimento PD enfocado a tours.....	47
Figura 3-5: Esquema general del procedimiento de diseño utilizado	51
Figura 3-6: Estructura del modelo MIMIC propuesto	63
Figura 4-1: Horas de inicio de los viajes basados en el trabajo.....	71
Figura 6-1: Estructura de un modelo híbrido	85
Figura B-1: Plan de actividades para el día de la encuesta	124
Figura B-2: Registro general de salidas	125
Figura B-3: e-mail de contacto para llenar parte 2 – componente PR.....	126
Figura B-4: e-mail recordatorio Semana 1 para llenado de parte 2 PR – componente PR.....	127
Figura B-5: Anuncio de ganadores rifas e inicio parte 3 – componente PD.....	128

LISTA DE ACRÓNIMOS

TDM: estrategias y políticas para gestionar la demanda de transporte. De las iniciales en inglés de *Travel Demand Management*.

EOD: Encuesta Origen Destino

PR: Preferencias Reveladas

PD: Preferencias Declaradas

GPS: sistemas de posicionamiento global. De las iniciales en inglés *Global Positioning System*

MNL: Modelo logit simple. Del inglés *Multinomial Logit*

HL: Modelo logit jerárquico. De las iniciales en inglés *Hierarchical Logit*

ML: Modelo logit mixto. De las iniciales en inglés *Mixed Logit*

OGEV: Modelo de valor extremo generalizado. De las iniciales en inglés *Ordered Generalized Extreme Value*

VL: Variables Latentes

SM: Modelo de demora horaria. De las iniciales en inglés *Scheduling Model*

TT: Tiempo de viaje. De las iniciales en inglés *Travel Time*

SD: Demora horaria. De las iniciales en inglés *Schedule Delay*

SDE: Demora horaria por llegar temprano. De las iniciales en inglés *Schedule Delay Early*

SDL: Demora horaria por llegar tarde. De las iniciales en inglés *Schedule Delay Late*

COST: Costo

PTD: Penalidad por menor duración de una actividad. De las iniciales en inglés *Participation Time Decrease*

PTI: Penalidad por mayor duración de una actividad. De las iniciales en inglés *Participation Time Increase*

Time-diff: Incertidumbre asociada a la variabilidad en el tiempo de viaje

MAVC: Matriz asintótica de covarianza

MIMIC: Modelo de variables latentes con múltiples indicadores y múltiples causas. De las iniciales en inglés *Multiple Indicator Multiple Cause*

ASC: Constantes específicas por alternativa. De las iniciales en inglés *Alternative Specific Constant*

Ceteris Paribus: Expresión latina que significa manteniendo el resto constante

VSG: Variaciones sistemáticas en los gustos

LR: Test de razón de verosimilitud. De las iniciales en inglés *Likelihood Ratio*

MLHS: Número pseudo-aleatorio generado por el algoritmo *Modified Latin Hypercube Sampling*

ACH: Auto como chofer

AAC: Auto como acompañante

CAM: Caminata

B: Bus

M: Metro

BM: Combinación de bus y metro

TX: Taxi

TXC: Taxi colectivo

CombTP: Combinación de transporte público con transporte privado

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA

MODELOS DE ELECCIÓN DE LA HORA DE INICIO DE VIAJE

Tesis enviada a la Dirección de Investigación, Innovación y Postgrado en cumplimiento parcial de los requisitos para el grado de Doctor en Ciencias de la Ingeniería.

JULIAN ALBERTO ARELLANA OCHOA

RESUMEN

El estudio de la elección de la hora en que se realizan los viajes ha tomado creciente importancia últimamente debido al aumento en los niveles de congestión de muchas ciudades y a la necesidad de evaluar los efectos de implementar estrategias de gestión como la tarificación vial, que buscan manejar la creciente demanda por transporte en las ciudades.

El principal objetivo de esta investigación es desarrollar modelos avanzados de elección de la hora de inicio de viajes, que incluyan atributos observables y no observables con el fin de mejorar el entendimiento de este complejo proceso de elección discreta.

Para lograr este objetivo, se partió por proponer una metodología de recolección de los datos adecuados para modelar esta elección. Adicionalmente, se decidió centrar la investigación en el estudio de viajes al trabajo, pues éstos típicamente representan más de la mitad de los viajes que se realizan en las horas de mayor congestión. La elección de la hora de partida fue abordada a partir de diferentes enfoques de modelación que permitieron: (i) evaluar formulaciones alternativas de modelos que permitan incorporar heterogeneidad en los gustos y correlación entre individuos; (ii) estudiar la posibilidad de incorporar atributos no directamente observables en estos modelos; (iii) estimar modelos conjuntos de elección del horario de inicio y de la elección del modo de transporte para los viajes, usando información de distinta naturaleza.

Este documento parte describiendo la metodología de diseño de la encuesta propuesta, y luego presenta la formulación, estimación y análisis de una serie de modelos de elección discreta que permitieron estudiar el problema de interés.

Las características del viaje y otros atributos relacionados con el cumplimiento de los horarios del trabajo, demostraron tener un rol preponderante en la elección de la hora en que se realizan los viajes al trabajo.

La formulación de modelos flexibles, que permitan capturar adecuadamente la heterogeneidad en los gustos y la correlación entre individuos y alternativas, son necesarios para una adecuada comprensión de elecciones complejas como la modelada en esta tesis. Sin embargo, debe prestarse especial cuidado a la interpretación de sus resultados pues, como se sabe, en estos modelos complejos puede producirse confusión de distintos efectos. Así, modelos parsimoniosos que expliquen la heterogeneidad entre individuos mediante variaciones sistemáticas en los gustos e incorporen adecuadamente la correlación entre distintas respuestas individuales (efecto panel) podrían preferirse por su gran potencial y mayor simpleza para su posterior uso en fase de predicción.

La incorporación de atributos no observables, la utilización de diferentes fuentes de información, y la integración de la elección de modo en la especificación y estimación de modelos de elección de la hora de inicio de viaje, produjo mejoras en su capacidad explicativa al permitir representar de mejor manera el proceso de elección subyacente, a costa de una mayor complejidad. Se presume que estos modelos podrían tener mejor comportamiento en fase predictiva, por lo que deben dedicarse mayores esfuerzos a esta área de estudio.

Miembros de la Comisión de Tesis Doctoral

Juan de Dios Ortúzar Salas
Luis Ignacio Rizzi Campanella
Sergio Jara Díaz
Alejandro Tudela Román
Juan Enrique Coeymans Avaria
Stephane Hess
Cristian Vial Edwards

Santiago, Agosto de 2012

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
SCHOOL OF ENGINEERING

TRIP DEPARTURE TIME CHOICE MODELS

Thesis submitted to the Office of Research, Innovation and Graduate Studies in
partial fulfilment of the requirements for the Degree of Doctor in Engineering Sciences
by

JULIAN ALBERTO ARELLANA OCHOA

ABSTRACT

The study of departure time choice has become increasingly relevant due to the growing levels of congestion in many cities and the need to assess the effects of travel demand management strategies such as road pricing.

The main aim of this research is to use state of the art discrete choice models to develop advanced departure time choice models using both observable and unobservable attributes to better represent the intrinsic complexity of this choice process.

A methodology to collect adequate departure time choice data is proposed. Analysis is focused on work departure time choice decisions, because they typically represent more than half of the trips in rush hours. Different modelling approaches in this thesis were used for: (i) evaluating alternative model formulations allowing to capture taste heterogeneity and panel effects, (ii) formulating and estimating hybrid departure time choice models, (iii) estimating joint mode and departure time choice models combining different data sources.

This document firstly describes the proposed survey method, and then presents the formulation, estimation and analysis of several departure time choice models that allowed studying this subject in certain depth.

Models confirm that attributes such as cost, travel time, travel time variability, comfort, and other work schedule related attributes, play a key role in the travel to work departure time choice process.

The development of flexible models to incorporate taste heterogeneity and correlation effects is desirable for a better understanding of complex choices such as departure time choice. However, modellers should be aware when interpreting results that confounding effects may occur. Simpler models able to accommodate properly panel effects and taste heterogeneity via deterministic interactions may be preferred due to their potential benefits in forecasting.

The inclusion of attitudinal variables along the use of different data sources and the consideration of mode choice within a departure time choice model framework, led to improvements in explanatory power as these joint models can better represent the underlying complexity of the choice process. Therefore, greater efforts should be devoted into this area of study because these models are expected to achieve better predictions.

Members of the Doctoral Thesis Committee:

Juan de Dios Ortúzar Salas
Luis Ignacio Rizzi Campanella
Sergio Jara Díaz
Alejandro Tudela Román
Juan Enrique Coeymans Avaria
Stephane Hess
Cristian Vial Edwards

Santiago, August, 2012

1. INTRODUCCIÓN

1.1 Motivación

La congestión, en vías y sistemas de transporte utilizados diariamente por las personas, produce incomodidades y pérdidas de tiempo irre recuperables. Los inconvenientes son más evidentes en las grandes ciudades, pues cada vez se encuentran sujetas a mayores niveles de congestión debido al crecimiento de la población y del parque automotor.

El aumento de la congestión en ciudades ha hecho que las estrategias y políticas para gestionar la demanda de transporte, de ahora en adelante estrategias TDM por las iniciales en inglés de *Travel Demand Management*, sean cada vez más populares. El propósito de las estrategias y políticas TDM es modificar el comportamiento de viaje de las personas con el fin de redistribuir la demanda hacia algunos lugares, modos de transporte y/o periodos de tiempo menos congestionados. Ejemplos de estrategias TDM son la tarificación vial, la implementación de horarios de trabajo flexibles, y las restricciones de circulación por patente.

Ante la presencia de congestión o cuando se busca implementar políticas y/o estrategias TDM, interesa saber cómo reaccionan los usuarios frente a las elecciones de modo, hora y ruta de sus viajes. En la literatura se reconoce que ante cambios en la condición de los viajes causados por congestión o por la implementación de políticas TDM, la segunda alternativa para los viajeros, después del cambio de ruta, es el cambio del horario de inicio del viaje (Ortúzar y Willumsen, 2011), superando al cambio de modo de transporte (Bianchi *et al.*, 1998; Hendrickson y Plank, 1984; Hess *et al.*, 2007a).

A pesar de que la elección de hora de inicio del viaje es una de las determinantes principales de la distribución temporal y espacial de la demanda, históricamente ha recibido menos atención que las elecciones de modo y de ruta. No obstante, en los últimos años su estudio ha cobrado mayor importancia.

Modelar la elección de la hora en que los individuos realizan sus viajes durante el día es importante porque permite entender cómo planifican sus viajes en función de los niveles de servicio que experimentan, el modo de transporte elegido, y sus compromisos horarios. Entender cómo se realizan las elecciones de hora de viaje puede ayudar en la planificación de infraestructura de transporte para acoger adecuadamente la demanda

proyectada a diferentes horas del día, desarrollar modelos de tráfico que tengan en cuenta la reacción de las personas frente a la congestión (Tseng *et al.*, 2011), y evaluar la efectividad de las políticas de transporte urbano (Bhat y Steed, 2002).

La motivación más importante de esta investigación es desarrollar modelos de elección de la hora de inicio de viaje, que tengan potencial para ser aplicados en la práctica y constituyan una herramienta útil para la toma de decisiones en transporte. En esta tesis se pretende extender el estado del arte en modelación de elecciones discretas desarrollando modelos que incluyan atributos observables y no observables para describir, de la mejor forma posible, la complejidad de la elección que se desea estudiar.

1.2 Objetivos y alcances

El objetivo principal de esta investigación es formular y estimar modelos de elección de la hora de inicio de viaje que constituyan el estado del arte en el tema.

Con el fin de alcanzar este objetivo general, se plantearon una serie de objetivos específicos. El primero consiste en proponer una metodología de encuesta, a partir de datos de movilidad “clásicos”, para estimar modelos apropiados de elección del horario de inicio de viajes. Esto implica obtener información que se encuentra disponible usualmente en las encuestas origen-destino (EOD) de viajes e incluir otros atributos que se consideren relevantes para el proceso de elección estudiado. Específicamente, en la tesis se propone una metodología de encuesta basada en una EOD, que contempla recolectar datos de preferencias reveladas (PR), preferencias declaradas (PD) y percepciones y actitudes individuales con el fin de construir variables latentes.

El segundo objetivo específico es estudiar la elección de la hora de inicio de viaje, en el contexto de políticas de gestión de la demanda que no se encuentran en funcionamiento en la actualidad. En particular, en esta tesis se evalúan diferentes formulaciones para la hora de inicio de viaje a partir de datos de PD en escenarios que incorporan tarificación vial y la flexibilización en los horarios de trabajo. A pesar de que los datos recolectados son lo suficientemente flexibles como para permitir estudiar la elección de hora de inicio de viaje usando simultáneamente datos de PR y PD, los modelos de elección horaria reportados en este documento se basan en sólo datos de PD.

El tercer objetivo específico es incorporar variables latentes en los modelos de elección de la hora de inicio de viaje. En particular, se discute la inclusión de la flexibilidad en el horario de trabajo como un atributo no directamente observable, que depende de las características del empleo de cada individuo y es determinante en el estudio de esta elección en el contexto de viajes basados en el trabajo.

El cuarto objetivo específico consiste en estimar modelos conjuntos de elección de modo y hora de inicio, para viajes basados en el trabajo. En la tesis se discute el tema y se presenta un modelo simplificado de elección conjunta aprovechando la disponibilidad de datos de distinta naturaleza. La componente PR de la encuesta, se usó para estimar modelos de elección de modo de transporte y la componente PD para modelar la elección de la hora de inicio de viajes.

El último objetivo específico consiste en calcular y discutir la valoración de los diferentes atributos incorporados en los modelos estimados. La valoración de atributos tiene la particularidad de estar influenciada por el efecto que pueda tener la implementación de políticas TDM (vale decir, tarificación vial y flexibilización de horarios de trabajo) que no se encontraban operando al recolectar los datos para esta investigación.

Por último, es interesante aclarar que en la encuesta se recolectó información respecto a todos los viajes que realizaban los encuestados durante un día, pero los resultados presentados en la tesis se enfocan sólo en viajes basados en el trabajo; es decir, viajes que terminan o se inician en el lugar de trabajo. El estudio del horario de inicio de viajes no basados en el trabajo, y la influencia de otras actividades en la elección de la hora de partida de viajes basados en el trabajo, se dejan como temas de investigación futura.

1.3 Principales contribuciones

Respecto a la recolección y procesamiento de datos (objetivo específico 1):

- Una metodología de diseño novedosa que permite recolectar datos de diferente naturaleza para estudiar la elección de la hora de inicio de viaje. Esta metodología contempla la recolección de datos PR, PD e indicadores de percepción para la formación de variables latentes.
- Una metodología para obtener niveles de servicio, de alta precisión, en redes de transporte público a partir de datos GPS; los niveles de servicio utilizados

comúnmente en modelos de demanda de transporte suelen ser agregados y poco detallados. Estos niveles de servicio de alto nivel de precisión, fueron utilizados para conformar la base de datos PR e informar el diseño de la encuesta de PD.

- Una metodología para diseñar eficientemente experimentos de elección (datos de PD) relacionados con la hora de inicio de viajes, incorporando dependencia entre atributos y atributos como el confort y la variabilidad del tiempo de viaje. La metodología de diseño se aplicó para generar experimentos de PD enfocados en el viaje al trabajo y también para *tours* basados en el trabajo.

Respecto a la modelación:

- Discusión metodológica, relacionada con el segundo objetivo específico, acerca de las consecuencias y forma de incorporar el *efecto panel* (correlación entre respuestas individuales en ejercicios de PD) en el contexto de modelar el horario de inicio en viajes al trabajo. De esta discusión nació el concepto definido como “efecto panel variable”, que propone tratar la correlación generada por el hecho de tener múltiples respuestas de parte de un individuo como una función de sus características socioeconómicas, dando lugar a que las respuestas de cada persona estén más o menos correlacionadas entre sí.
- La incorporación de variables latentes para modelar la elección de la hora de inicio de viaje es otro aspecto novedoso de esta investigación. En particular, en la tesis se presenta la consideración del atributo *flexibilidad en los horarios de trabajo* como un concepto latente (objetivo específico 3).
- Desarrollo de un modelo conjunto de elección modo y hora de inicio de viaje en el contexto de *tours*. El hecho que la unidad de análisis sean *tours* en lugar de viajes, ya es una novedad en sí; sin embargo, otra particularidad del modelo propuesto es que utiliza datos de diferente naturaleza. La elección de modo es descrita a través de los datos de PR que fueron recolectados, mientras que la elección de hora es modelada a partir de los experimentos de PD (objetivo específico 4).

1.4 Estructura del contenido

Aparte de esta Introducción, el documento tiene otros siete capítulos. En el segundo se presentan los antecedentes generales que enmarcan el trabajo, partiendo por una contextualización de los conceptos en que se basan los modelos de horario de inicio de viajes, continuando con una breve discusión de algunos aspectos relevantes relacionados con la formulación de dichos modelos. En particular, se discuten aspectos metodológicos como el enfoque de modelación, tipos de modelos y datos a emplear, factores que pueden incidir en la elección, y relación de la hora de inicio de viaje con otras elecciones de transporte. El capítulo termina con un resumen de las principales limitaciones encontradas en la literatura.

En el tercer capítulo se presenta brevemente el diseño y la metodología para la recolección de la información utilizada en la tesis con el fin de modelar la elección de hora de partida de viaje, enfatizando el desarrollo de encuestas de preferencias reveladas y declaradas.

En el cuarto capítulo se describen algunos atributos de la muestra recolectada utilizando el instrumento descrito en el tercer capítulo. En particular, se muestran las características socioeconómicas de los encuestados, las características de los lugares de empleo donde se llevó a cabo las encuestas, y se analizan las elecciones de modo de transporte y hora en que los encuestados iniciaron sus viajes.

En el quinto capítulo se presenta la formulación teórica y la estimación práctica de modelos de elección de la hora de inicio de viaje, usando datos de preferencias declaradas. El análisis se inicia con un modelo logit simple, que asume parámetros genéricos para toda la población; luego se incluye heterogeneidad entre individuos a través de variaciones sistemáticas de gustos y posteriormente modelos más flexibles y complejos de tipo logit mixto. El capítulo finaliza con el análisis de la inclusión del efecto panel en los modelos de elección de hora de inicio de viaje formulados previamente.

En el sexto capítulo se describe la formulación y estimación de modelos de elección de la hora de inicio de viaje considerando variables latentes. En particular, se discute el efecto y la forma adecuada de incluir la flexibilidad en el horario de trabajo como una variable latente relevante en el proceso de elección estudiado.

En el séptimo capítulo se presenta una modelación conjunta de las elecciones de modo de transporte y hora de inicio en viajes al trabajo. Se describe la formulación y se discuten las implicancias de modelar en forma conjunta el modo de transporte y la hora de partida al considerar simultáneamente los viajes de llegada y salida del trabajo.

Finalmente, el octavo capítulo presenta las conclusiones y recomendaciones generales de esta investigación.

2 ANTECEDENTES GENERALES

El presente capítulo tiene el objetivo de entregar el marco teórico que sustenta la investigación. Inicialmente se entrega una breve motivación acerca de los modelos de elección de hora de inicio de viaje. Luego se presenta el enfoque de modelación tradicionalmente usado para estudiar la elección y se discuten algunos aspectos metodológicos importantes para su formulación. Posteriormente se describen algunos atributos considerados relevantes para la elección de hora, y se entrega información sobre la potencial relación de esta elección con otras decisiones de transporte.

La principal motivación de los modelos de elección del horario de viaje consiste en estimar cómo se distribuyen los desplazamientos de las personas a lo largo del día. Estos modelos contribuyen a entender cómo organizan sus viajes y actividades en el tiempo los viajeros (Hess *et al.*, 2007a), y constituyen un área de investigación en pleno desarrollo. Además, permiten vincular variables socioeconómicas de los individuos con su percepción de los niveles de servicio de los distintos modos de transporte.

Conocer las horas elegidas por los individuos para realizar sus viajes y los atributos que toman en cuenta para efectuar dicha elección, es crucial para cuantificar detalladamente la demanda por transporte a lo largo del día. El crecimiento de la congestión en las ciudades y la implementación de políticas TDM ha hecho que el cambio en la hora de inicio de viaje sea una opción popular para evitar transportarse por ciertos lugares en periodos congestionados, o para experimentar costos (monetarios o de tiempo de viaje) menores en períodos diferentes al horario que sería preferido si no existiera congestión.

2.1 Enfoque de modelación tradicional

La gran mayoría de los estudios de elección horaria reportados hasta el momento se sustentan en la teoría de la utilidad aleatoria (Domencich y McFadden, 1975; Ortúzar y Willumsen, 2011; Williams, 1977). La teoría de utilidad aleatoria considera que los tomadores de decisiones son racionales, poseen información completa y realizan una elección dentro de un conjunto discreto de alternativas disponibles, luego de compensar los pros y contras de una serie de atributos que describen las alternativas. La teoría considera

que la utilidad (U_{iq}) que obtiene un individuo q al elegir cierta alternativa i es “aleatoria” debido a que el analista no posee información perfecta y sólo es capaz de representar una parte de la utilidad experimentada por el individuo (V_{iq}). Para explicar las elecciones de los individuos que no parecen razonables, dada la diferencia entre los atributos que el analista puede observar y los que los individuos realmente consideran, el modelador adiciona un término de error (ε_{iq}) asociado a cada alternativa i , como se representa en la ecuación (1).

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

A partir de la función de distribución que el analista suponga para los términos de error ε_{iq} en (1), es posible generar diferentes modelos de elección. El modelo logit simple (Domencich y McFadden, 1975) o logit multinomial (MNL), se genera al asumir una función de distribución independiente e idénticamente distribuida Gumbel (también conocida como Weibull o Valor Extremo tipo I) para los términos de error ε_{iq} .

El MNL es el modelo más simple y popular. Considera que la probabilidad (P_{iq}) que el individuo q elija la alternativa i está dada por:

$$P_{iq} = \frac{\exp(\lambda \cdot V_{iq})}{\sum_j \exp(\lambda \cdot V_{jq})} \quad (2)$$

donde V_{iq} representa la parte de la utilidad observada por el analista. V_{iq} es función de una serie de k atributos observables (denotados por X) asociados a la alternativa i y/o al individuo q , y a un conjunto de k parámetros a estimar (denotados por θ) relacionados con cada atributo. j representa el conjunto de alternativas disponibles, donde $i \in j$. Si se asume una forma funcional lineal, como en (3), dichos parámetros (θ_{ik}) representan la utilidad marginal que percibe un individuo q al aumentar una unidad del respectivo atributo k .

$$V_{iq} = \sum_k \theta_{ik} \cdot X_{ikq} \quad (3)$$

Por otra parte, λ en (2) es conocido como el *factor de escala* del modelo y es inversamente proporcional a la desviación estándar (σ) de los términos de error ε_{iq} . El factor de escala generalmente no es identificable, ya que a partir de (2) es claro que todos los parámetros θ_{ik} están multiplicados por el factor de escala λ , haciendo imposible la identificación conjunta de ambos parámetros. Dado lo anterior, es necesario normalizar (fijar) el factor de escala para poder estimar parámetros θ_{ik} . Usualmente, se asume que $\lambda=1$

y como $\lambda = \frac{\pi}{\sqrt{6}\sigma}$ para un MNL, es importante tener en cuenta que los parámetros θ_{ik} estimados van a estar deflactados por las desviaciones estándar de los errores, que son desconocidas.

Debido al supuesto de identidad e independencia entre términos de error ε_{iq} , el MNL asume ausencia de correlación entre alternativas, igualdad en la varianza de los términos de error asociados a las alternativas (homocedasticidad), e inexistencia de heterogeneidad en los gustos de la población.

El modelo logit jerárquico (HL), es una generalización del MNL, que se basa también en la distribución Gumbel pero permite incluir correlación entre algunas de sus alternativas (Carrasco y Ortúzar, 2002; Daly y Zachary, 1978; Williams, 1977). La derivación y los supuestos detrás de este modelo pueden encontrarse en Ortúzar y Willumsen (2011).

La estructura de error y la formulación de los modelos MNL y HL permiten su rápida estimación. La estimación de estos modelos se realiza comúnmente mediante la técnica de máxima verosimilitud. Sin embargo, su formulación no es tan flexible como para acomodar estructuras complejas de correlación o heterogeneidad no observable en los gustos de la población. Si se desea incluir parte de la heterogeneidad en los gustos en modelos MNL y HL, se deben incluir interacciones entre variables socioeconómicas y los atributos de las alternativas, enfoque que se conoce como “variaciones sistemáticas de gustos” (Ortúzar y Willumsen, 2011).

Los modelos probit (Daganzo, 1979) y logit mixto (ML) son más flexibles que los anteriores y permiten reproducir estructuras complejas de correlación y heterogeneidad. El modelo probit, asume que los términos de error distribuyen Normal multivariado pero su estimación es tan compleja como la del ML y no se encuentra fácilmente disponible en paquetes de *software* comerciales. El ML (Cardell y Reddy, 1977; McFadden y Train, 2000; Ortúzar y Willumsen, 2011) permite asumir parámetros θ_{ikq} que pueden variar aleatoriamente en la población, a diferencia de los modelos ML y HL donde los parámetros θ_{ik} son los mismos para todos los individuos dentro de la población. El ML es un modelo que a través de la inclusión de parámetros aleatorios y/o componentes de error,

combinados con el término de error típico de la familia logit (es decir, aquel que distribuye Gumbel) puede generar cualquier estructura de correlación entre las alternativas (McFadden y Train, 2000).

El ML puede ser derivado bajo diversas especificaciones. La especificación más sencilla y popular se basa en la consideración de parámetros aleatorios. La función de utilidad del individuo q para la alternativa i puede escribirse como (4):

$$U_{iq} = \sum_k \theta'_{ikq} \cdot X_{ikq} + \varepsilon_{iq} \quad (4)$$

donde X_{ikq} son atributos asociados a la alternativa i y al individuo q , θ'_{ikq} es el parámetro asociado al atributo k y alternativa i , cuyo valor varía entre individuos q , representando sus gustos. ε_{iq} representa el término de error que distribuye iid Gumbel asociado a la alternativa i .

Por otra parte, una especificación de modelos ML menos utilizada, consiste en introducir *componentes de error* para, por ejemplo, crear correlaciones entre utilidades de distintas alternativas y/o capturar heterocedasticidad. En este caso, la función de utilidad del individuo q para la alternativa i puede escribirse como (5):

$$U_{iq} = \sum_k \theta'_{ikq} \cdot X_{ikq} + \nu'_{iq} + \varepsilon_{iq} \quad (5)$$

donde X_{ikq} son atributos asociados a la alternativa i y al individuo q . θ'_{ikq} representa el parámetro asociado al atributo k y alternativa i , cuyo valor puede variar entre individuos q o mantenerse constante. ν'_{iq} representa las *componentes de error*, pudiendo ser función de atributos de las alternativas, y depender del individuo q . ε_{iq} es el término de error iid Gumbel. Resulta entonces que ν'_{iq} y ε_{iq} definen la parte aleatoria de la utilidad U_{iq} .

Dado que los parámetros θ pueden variar en la población, la probabilidad de elección para el caso del modelo ML es la integral de probabilidades Logit tradicionales sobre la densidad de los parámetros considerados como aleatorios:

$$P_{iq} = \int \frac{\exp(V_{iq}(\theta))}{\sum_j \exp(V_{jq}(\theta))} \cdot f(\theta) d\theta \quad (6)$$

donde $f(\theta)$ se conoce como *distribución de mezclas (mixing distribution)*. Como el analista no conoce los valores de θ para cada individuo q , el analista debe asumir una función de distribución para los valores θ' y estimar los parámetros que describen tal distribución.

Si se consideran varios parámetros aleatorios y componentes de error en la formulación del ML, su probabilidad de elección contendrá múltiples integrales que no pueden resolverse analíticamente con facilidad. Actualmente, la técnica de máxima verosimilitud simulada es la más utilizada en la práctica para el cálculo de las probabilidades de modelos ML. En particular, la ecuación (7) representa la forma general de la técnica de máxima verosimilitud:

$$\hat{P}_{iq} = \frac{1}{R} \sum_{r=1}^R \frac{\exp(V_{iq})}{\sum_j \exp(V_{jq})} \quad (7)$$

donde \hat{P}_{iq} representa la probabilidad promedio, obtenida al computar la probabilidad R veces (*draws*) durante el procedimiento de simulación. Cuando se consideran parámetros aleatorios, se deben extraer entonces R valores provenientes de la función de distribución $f(\theta)$ supuesta por el analista.

Debido a que la estimación por máxima verosimilitud simulada del ML se encuentra disponible en los *softwares* comerciales de estimación más conocidos (BIOGEME, LIMDEP, ALOGIT), su uso ha ganado gran popularidad sobre el uso del probit en los últimos años. Sin embargo, la interpretación de los resultados del ML suele ser compleja (Sillano y Ortúzar, 2005; Train y Wilson, 2009) y debido a su extrema flexibilidad existe la posibilidad de que ciertos efectos se confundan (Amador *et al.*, 2008).

La mayoría de los modelos de elección de la hora de inicio de viaje reportados en la literatura son del tipo logit (de Jong *et al.*, 2003; Hess *et al.*, 2007a; Small, 1982; Tseng *et al.*, 2011) y usan intervalos discretos para representar el tiempo. Sin embargo, el supuesto de considerar que el tiempo puede dividirse en períodos discretos no es ampliamente aceptado debido a que el tiempo es, en esencia, una variable continua y no es evidente la forma en que estos intervalos deben formarse (Bhat, 1998b).

2.1.1 ¿Tiempo como variable discreta o como variable continua?

La manera más simple y tradicional de formular modelos de elección de la hora de inicio de viaje es dividir el tiempo en períodos discretos. La elección del individuo consiste, entonces, en elegir un período de tiempo para realizar su viaje entre una serie de alternativas discretas disponibles. Considerar el tiempo como una variable discreta permite utilizar modelos de elección discreta tradicionales que se encuentran implementados en *software* econométrico disponible en el mercado.

Debido a su simpleza y facilidad de estimación, el primer modelo de elección discreta utilizado para modelar la elección de horario de viaje fue el MNL. La estructura del MNL se ha propuesto para calcular la probabilidad de elegir cierto periodo de llegada al trabajo (Abkowitz, 1981; Small, 1982), y también para modelar la elección de la hora de inicio de viajes que se realizan inmediatamente después de salir del trabajo (Hendrickson y Plank, 1984). La principal limitación del MNL es que las alternativas de elección, en este caso periodos de hora continuos, estén correlacionadas. Sin embargo, podría existir una fuerte correlación entre viajes que se realicen al final y al inicio de períodos horarios consecutivos, o si los periodos horarios a elegir son pequeños. Por ejemplo, si se crean períodos de quince minutos, viajes a las 8:31 hrs y a las 8:30 hrs pertenecerían a períodos diferentes e independientes, y es difícil pensar que viajar a las 8:31 hrs sea muy diferente e independiente de hacerlo a las 8:30 hrs.

Como alternativa para considerar la existencia de correlación entre los períodos de tiempo y capturar de forma más flexible la potencial heterogeneidad de los gustos de los individuos respecto a esta elección, se han formulado modelos más complejos. En la literatura se reportan modelos de elección de horario de viaje basados en estructuras HL (Chin, 1990); valor extremo generalizado con alternativas ordenadas, OGEV (Bhat, 1998b; Small, 1987; Steed y Bhat, 2000), probit (Tringides *et al.*, 2004), y ML (Bhat, 1998a; Börjesson, 2008; de Jong *et al.*, 2003; Hess *et al.*, 2007b).

Chin (1990) formuló un modelo HL para estudiar la elección del horario de inicio de los viajes sujetos a congestión vial durante tres horas consecutivas en la mañana. Las alternativas formuladas en el modelo correspondían a 3 periodos consecutivos de 1 hora, que además fueron subdivididos en periodos consecutivos de 15 minutos. Aquellas

alternativas de 15 minutos dentro de un mismo periodo de 1 hora se consideraron correlacionadas. A pesar que el modelo HL de Chin (1990) representa un avance en cuanto a la inclusión de correlación entre alternativas, considerar que periodos consecutivos de 15 minutos que pertenezcan a dos periodos de hora diferentes (por ejemplo, 7:45-8:00 y 8:00-8:15) son independientes entre si, puede ser cuestionable.

Algunos autores han intentado capturar la potencial correlación entre periodos discretos consecutivos, utilizando modelos de alternativas ordenadas. En particular, Small (1987) fue el primero en intentarlo pero encontró que los resultados al estimar un modelo considerando la ordinalidad de las alternativas discretas mediante una estructura del tipo valor extremo generalizado (OGEV), no eran estadísticamente superiores a los entregados por un modelo MNL que tenía la misma configuración de alternativas (tamaño y cantidad de periodos discretos). Steed y Bhat (2000) tampoco obtuvieron resultados satisfactorios cuando usaron un modelo tipo OGEV dado que sus estimaciones, en el contexto de la elección de hora de inicio de viajes recreacionales, no fueron consistentes con la teoría de la utilidad aleatoria. Sin embargo, Bhat (1998b) concluyó que se justifica formular modelos de elección discreta con estructuras de error más complejas al contrastar los resultados de un modelo OGEV con un MNL, en el contexto de 4516 viajes de compras en San Francisco. Bhat (1998b) encontró que no considerar la correlación existente entre periodos consecutivos alternativos puede llevar a estimaciones sesgadas de los efectos de las variables de niveles de servicio que influyen en la elección de hora de inicio de los viajes, y subestimaciones en la estimación de la reducción de viajes en periodos punta.

Dado que el ML permite mayor libertad en la formulación de la estructura de correlación entre las alternativas discretas consideradas, diversos autores han modelado la elección de la hora de inicio de los viajes usando estos modelos (Bhat, 1998a; Börjesson, 2008; de Jong *et al.*, 2003; Hess *et al.*, 2007b). En particular, Bhat (1998a) reportó que el ML realiza mejores predicciones en el contexto de viajes recreacionales que los modelos más simples, sin embargo incurre en un costo computacional mayor que puede hacerlo inviable en la práctica.

A pesar de que la mayor parte de la literatura considera modelos que tratan el tiempo de manera discreta, también se encuentran modelos que consideran al tiempo como

variable continua. Los primeros intentos por hacer esto se encuentran en modelos teóricos relacionados con la elección de horario de viaje en una sola vía particular (Arnott *et al.*, 1990; Vickrey, 1969). Recientemente, se han utilizado modelos de tipo *Hazard* (Hensher y Mannering, 1994) para la elección de horario de viaje (Bhat y Steed, 2002; Day *et al.*, 2010; Lee y Timmermans, 2007; Wang, 1996). Los modelos *Hazard* permiten estimar tanto la duración de las actividades como el inicio o término de éstas, lo cual otorga teóricamente mayor flexibilidad en la estimación y mejor precisión en la predicción.

Los modelos *Hazard* se ocupan principalmente para estimar la duración de un evento o determinar el tiempo en que este termina. En el contexto de la elección de horario de inicio de un viaje desde la casa hacia el trabajo, se intenta estimar la hora en que se termina la actividad “*estar en casa*” con el inicio de un viaje al trabajo.

Sea T una variable aleatoria no negativa que representa la duración desde un punto de referencia hasta que inicie el evento. En relación a la elección del horario de inicio de viajes, T representa las horas de inicio de los viajes, expresado como la cantidad de minutos desde la medianoche en una población. Si se define $f(t)$ como la función de probabilidad de que finalice el evento en el instante de tiempo t , $F(t)$ es la función de distribución acumulada de la variable aleatoria T , que representa la probabilidad de que el evento ocurra en el instante t o antes. $F(t)$ puede escribirse entonces como (8):

$$F(t) = P(T \leq t) = \int_0^t f(t)dt \quad (8)$$

La probabilidad que el evento ocurra después del instante de tiempo t , se define como la función de supervivencia, $S(t)$, y se puede escribir como (9):

$$S(t) = P(T > t) = \int_t^{\infty} f(t)dt = 1 - F(t) \quad (9)$$

La tasa de riesgo (*Hazard rate*) representa la probabilidad de finalización instantánea de un evento o actividad entre el instante de tiempo t y el instante $t + dt$, condicional en que el evento no ha terminado antes que el instante de tiempo t . La tasa de riesgo puede representarse matemáticamente como (10):

$$\lambda(t) = \lim_{dt \rightarrow 0} \frac{P[t < T < t + dt | T > t]}{dt} \quad (10)$$

Si se considera dt como un periodo de tiempo infinitesimalmente pequeño, la probabilidad que el evento finalice precisamente en el periodo $t + dt$, viene dada por $\lambda(t) \times dt$. Teniendo en cuenta lo anterior, la tasa de riesgo puede describirse como (11):

$$\lambda(t) = \frac{f(t)}{[1-F(t)]} = \frac{f(t)}{S(t)} \quad (11)$$

De esta manera, la tasa de riesgo puede interpretarse como la probabilidad de que se termine el evento en el siguiente instante de tiempo respecto a un tiempo t . La forma que toma la tasa de riesgo tiene importantes implicancias en la duración estimada de los eventos por parte del modelo.

La formulación de la tasa de riesgo, por lo general, considera el efecto de variables exógenas que pueden depender o no del tiempo en que ocurre el evento. Por lo general, se considera que la tasa de riesgo puede descomponerse en dos, como se indica en (12):

$$\lambda(t, x, \beta) = \lambda_0(t) \varphi(t, x, \beta) \quad (12)$$

donde λ_0 representa una tasa de riesgo base, x es un vector de variables explicativas exógenas, y β corresponde a un vector de coeficientes a estimar.

En general, la formulación de la tasa de riesgo base es desconocida, y se debe asumir una distribución para describirla. Distribuciones paramétricas y no paramétricas son usadas comúnmente para describir la tasa de riesgo base. Entre las distribuciones paramétricas comunes se encuentran las distribuciones exponencial, weibull, log-logística, gompertz, log-normal, y gamma (ver Hensher y Mannering, 1994). Una aplicación de métodos no paramétricos para determinar la tasa de riesgo base puede encontrarse en Bhat y Steed (2002), quienes recomiendan usar una distribución no paramétrica para no realizar supuestos acerca de la forma de la tasa de riesgo base (como en el caso paramétrico) que pueden llevar a resultados inconsistentes.

Dependiendo de la forma funcional y las variables explicativas consideradas dentro de la función multiplicativa $\varphi(t, x, \beta)$, se obtienen diferentes efectos sobre la función tasa de riesgo. Si se considera que la única componente que depende del tiempo es la función de tasa de riesgo base (λ_0), la forma funcional de la tasa de riesgo se conoce como *riesgo proporcional*. Si las variables explicativas dentro de la función multiplicativa afectan directamente la escala de tiempo considerado en el modelo, la forma funcional de la tasa de

riesgo se conoce como *modelo de tiempo de fallo acelerado*. La especificación típica de la función multiplicativa en modelos de *riesgo proporcional* viene dada por (13), mientras que la de los *modelos de tiempo de fallo acelerado* viene dada por (14).

$$\phi(x, \beta) = e^{-\beta x} \quad (13)$$

$$\varphi(t, x, \beta) = t e^{-\beta x} \quad (14)$$

Lee y Timmermans (2007) sugieren que un *modelo de tiempo de fallo acelerado* es más adecuado para modelar las elecciones de hora de inicio de los viajes porque es esperable que las decisiones de la hora de viaje sean dinámicas debido a que el comportamiento de las personas no es proporcional. Adicionalmente, considerar la especificación descrita en (14) provee una interpretación conveniente de los coeficientes asociados a las variables explicativas, pues la ecuación (14) puede transformarse en un modelo log-lineal que puede escribirse como (15):

$$\ln(t) = \beta'x + \varepsilon \quad (15)$$

De esta manera, un incremento en una unidad de la *j-esima* variable explicativa x resulta en un incremento en la duración del evento en un β_j por ciento.

La estimación de los modelos de duración generalmente se realiza mediante máxima verosimilitud. La función de verosimilitud a maximizar puede escribirse en términos de la función de densidad de duración como (16):

$$L(\theta, \beta) = \prod_q f(t_q, \theta, x_q, \beta) \quad (16)$$

donde θ representa el vector de parámetros que caracteriza la función de distribución que caracteriza la tasa de riesgo base asumida.

Los resultados reportados por algunos autores (Bhat y Steed, 2002; Day *et al.*, 2010; Lee y Timmermans, 2007; Wang, 1996) quienes han estudiado la elección de la hora de inicio de los viajes mediante formulaciones tipo *Hazard* son prometedores, sin embargo aún no se cuenta con un consenso generalizado sobre la utilización de este tipo de modelos, ya que su sustento económico no ha sido desarrollado. Considerar el tiempo como una variable continúa elimina el problema de correlación entre períodos excluyentes y evidentemente es atractivo; sin embargo, su principal limitante es la interpretación económica y la complejidad de su estimación, que tampoco se encuentra disponible en

paquetes de *softwares* econométricos convencionales utilizados en la modelación de elecciones.

2.1.2 Tipos de datos para modelar la elección de hora

Los estudios de elección de la hora de inicio del viaje, comúnmente hacen uso de datos preferencias declaradas (PD) porque la información de preferencias reveladas (PR) necesaria para lograr buenas estimaciones de los parámetros incluidos en los modelos de elección de horario de viaje son muy difíciles de obtener (Hess *et al.*, 2007b; Tseng *et al.*, 2011), ya que requieren un procedimiento de recolección muy riguroso y costoso.

Los primeros modelos de elección de hora de inicio de viaje utilizaban datos de PR recopilados en el contexto de grandes estudios de demanda de transporte urbano y se enfocaban principalmente en los viajes al trabajo en horario punta. La tendencia en los últimos años ha sido utilizar datos de PD para cuantificar los compromisos entre diferentes variables relacionadas con la elección del horario, que son difíciles de obtener con datos de PR, principalmente porque algunos de los atributos de estos datos suelen estar altamente correlacionados o no se encuentran disponibles en grandes encuestas de movilidad. Sin embargo, una desventaja de los modelos de elección de hora estimados usando datos de PD es que sus predicciones pueden ser menos confiables.

Aunque la combinación de datos PR y PD para aprovechar las ventajas de cada tipo de datos tiene una larga tradición en estudios de transporte, este no es el caso para estudios donde se modela la elección de la hora de inicio de viaje. Sólo Börjesson (2008) y Tseng *et al.* (2011) han propuesto modelos de elección de horario de esta clase que usan simultáneamente datos de PR y PD.

Börjesson (2008) recolectó sus datos de PR y PD para evaluar el comportamiento de usuarios de auto ante la implementación de una prueba de tarificación vial en la ciudad de Estocolmo. Los datos PD se enfocaron en explorar los compromisos de usuarios de auto respecto a sus horas preferidas de inicio de los viajes, tiempo de viaje, variabilidad del tiempo de viaje y costo. La encuesta PD contaba con 8 situaciones de elección donde se presentaban 2 horarios de viaje alternativos, y 1 alternativa para cambiar de modo manteniendo la hora de viaje (no se mostraban los atributos de esta alternativa). En la

componente PR, la autora asumió que los individuos escogían viajar en uno de los 18 periodos de 15 minutos entre las 5:15 y 9:15 am.

Los resultados obtenidos por Börjesson (2008) sugieren que los datos PD son menos confiables que los PR para analizar y predecir las elecciones de hora de inicio de viajes. Börjesson (2008) argumenta que las respuestas observadas en los datos PR se relacionan con elecciones de largo plazo relacionadas con el hecho que el viaje que se investiga en su artículo es de tipo rutinario. Börjesson (2008) comenta que las predicciones realizadas a partir de su modelo PD no se ajustan a lo observado en la realidad, donde el resultado de la prueba de la tarificación vial indicó que las personas afectadas por la política prefirieron cambiar de modo de transporte en lugar de horario de inicio de sus viajes como lo sugería el modelo formulado a partir de sus datos PD. La evaluación de la prueba de tarificación vial no encontró evidencias que los usuarios de auto cambiaran considerablemente sus horarios de viaje.

Por otra parte, Tseng *et al.* (2011) recolectaron sus datos de PR y PD en el marco del proyecto *Spitsmijden* en Holanda, que consiste en remunerar económicamente a aquellos usuarios de auto que decidan cambiar sus horas de viaje desde las horas punta en la mañana a horas menos congestionadas. La encuesta PD consistió en la presentación de 10 situaciones de elección, con dos alternativas de elección donde el individuo necesitaba decidir si viajaba en horario punta o fuera de punta. Los atributos dentro de cada alternativa fueron la hora de inicio del viaje desde el hogar, posibles tiempos de viaje en diferentes arcos de la vía, posible hora de llegada al trabajo, y remuneración económica por viajar a determinadas horas. Los datos PR se obtuvieron a partir de mediciones GPS junto con observaciones con cámaras de tráfico en una vía específica, que capturaron vehículos viajando entre las 4:45 y 9:45 am. Los autores asumieron que los individuos observados escogieron su horario de inicio de viaje entre 20 posibles periodos alternativos de 15 minutos de duración.

Al igual que Börjesson (2008), Tseng *et al.* (2011) encontraron que los datos PR relacionados con la elección de la hora de inicio de los viajes tienen menor varianza que los datos PD. La potencial explicación de este resultado es que el comportamiento observado a través de los datos PR es el resultado de un proceso de elección a largo plazo,

que lleva a menores errores de respuesta. Además, los autores concluyen que las disposiciones al pago obtenidas a partir de los datos PD son mucho mayores que las obtenidas al usar datos PR, posiblemente porque los individuos responden estratégicamente a las encuestas PD, tienen otros incentivos para responder la encuesta, o no consideran todos los atributos presentados en el experimento de elección.

Debido a las pocas aplicaciones de datos PR-PD para estudiar la elección de hora de inicio de los viajes, y las potenciales ventajas que tiene la modelación con diferentes fuentes de datos, es de interés contar con mayor número de estudios que utilicen esta técnica de mezcla de datos para aprovechar sus bondades en el estudio de esta elección. A pesar que las conclusiones de Börjesson (2008) y Tseng *et al.* (2011) se encuentran en la misma línea, otras investigaciones son necesarias para extender y complementar sus resultados.

Otro tema que se mantiene inexplorado en los modelos de elección de hora de inicio de viaje es la inclusión de actitudes y percepciones individuales. La incorporación de variables latentes (VL), que permiten capturar actitudes y percepciones, ha ganado popularidad en los últimos años, aunque de nuevo, este no el caso en estudios de elección de la hora de inicio de viaje. Parece razonable esperar que incluir información de actitudes y percepciones en estos modelos mejore su poder explicativo. Así, la combinación de diferentes tipos de datos para formular y estimar modelos de elección de hora es uno de los temas que se abordará en esta tesis.

2.2 Relación de la elección de hora de inicio de viaje con otras elecciones de transporte

En la literatura no existe un consenso claro sobre qué etapa, dentro de un modelo de planificación estratégica, es la mayormente afectada por la elección de horario ni tampoco se conoce con certeza cuál debiera ser la posición adecuada de la elección de hora de inicio de los viajes dentro de un modelo estratégico.

Existen dos corrientes marcadas de estudio de la elección de la hora de inicio de viajes. La primera corriente de estudio se concentra en estudiar la componente de demanda del problema de elección de hora de inicio de viajes, considerando que los atributos de

nivel de servicio (tiempo de viaje, costo, y variabilidad del tiempo de viaje) se determinan de forma exógena. La segunda corriente, estudia la elección de horario de inicio de viajes en el contexto de un problema de equilibrio oferta-demanda donde los atributos de nivel de servicio se consideran endógenos al problema.

Dentro de la primera corriente, la formulación más conocida y sobre la que se basan la mayoría de estudios realizados es el *Scheduling Model* (SM) propuesto por (Small, 1982). El SM supone que los individuos eligen el periodo de inicio de sus viajes teniendo en cuenta el tiempo de viaje y la utilidad que les produce llegar temprano, tarde o a la hora, respecto a una hora deseada de llegada al destino.

Con el fin de caracterizar más ampliamente la demanda, la elección de la hora de inicio de los viajes se ha estudiado en conjunto con la elección de modo de transporte. Las elecciones de modo y hora de inicio de viaje ha sido simultáneamente abordadas utilizando diferentes enfoques de modelación (Bajwa *et al.*, 2008; Bhat, 1998a; Day *et al.*, 2010; Hess *et al.*, 2007a) que buscan principalmente estudiar la relación entre estas dos elecciones y determinar la sensibilidad de estas elecciones ante cambios en los niveles de servicio. En esta línea, varios estudios han concluido que los individuos prefieren cambiar sus horarios de inicio de viajes antes que modificar su modo de transporte acostumbrado (Arellana *et al.*, 2012a; Hendrickson y Plank, 1984; Hess *et al.*, 2007b). En particular, Hendrickson y Plank (1984) sugieren que la elección horaria debe considerarse previamente a la elección de modo pues la hora de partida está condicionada fuertemente por el nivel de congestión en las redes, la confiabilidad de los modos de transporte, y la disponibilidad de estacionamientos en el lugar de destino del viaje.

La elección de modo también puede estar afectada, entre otras, por el nivel de congestión en los modos disponibles y el costo de viajar a diferentes horas del día. Por ejemplo, en el contexto de actividades no obligatorias de recreación, el individuo puede decidir primero en que modo de transporte realizar su viaje, y de acuerdo a su elección de modo decidir la hora en que inicia su viaje.

Por otra parte, el trabajo seminal dentro de la segunda corriente de estudios que considera la elección de hora de inicio de viajes como un problema de equilibrio, es el desarrollado por Vickrey (1969). El modelo desarrollado por Vickrey (1969) es un modelo

determinístico donde un número fijo de personas viajan desde un origen a un destino, usando una sola ruta y atraviesan por un cuello de botella que se congestiona debido a que tiene capacidad limitada. Vickrey (1969), consideró que los usuarios son agentes microeconómicos homogéneos que buscan minimizar una función de costos determinada que depende del tiempo de viaje y la demora horaria respecto a un horario de referencia establecido. En el modelo, idealmente los viajeros escogerán viajar a sus horas preferidas de viaje; sin embargo, debido a la presencia de congestión en el cuello de botella, los tiempos de llegada de los viajeros no serán los preferidos. Dado que los usuarios son considerados homogéneos, la única posibilidad que el sistema se encuentre en equilibrio es que la utilidad total de todos los viajeros sea la misma para todos los horarios de inicio de viaje escogidos.

El trabajo de Vickrey (1969), ha sido extendido para incorporar diferentes supuestos de demanda, como heterogeneidad en los viajeros y diferentes funciones de costos (Arnott *et al.*, 1990), y para considerar la presencia de múltiples rutas para viajar desde un origen a un destino (concepto de red). Los modelos que consideran el problema de equilibrio oferta-demanda determinado por la elección de hora de inicio de viajes ante la presencia de una red de transporte, donde existen múltiples rutas para llegar a un destino, se conocen como asignación de tráfico dinámica con elección de hora de inicio de viajes (Ettema *et al.*, 2004; Heydecker y Addison, 2005). Este tipo de modelos, que considera simultáneamente las elecciones de hora y ruta, es especialmente relevante cuando existen políticas de tarificación en ciertas zonas, a las cuales se puede acceder por diferentes rutas y a diferentes horas del día.

La asignación de tiempo a las actividades y horas en que un individuo realiza sus viajes puede afectar también la elección de destino. Kitamura *et al.* (1998) encontraron alta correlación entre la disponibilidad del tiempo y la elección de destino, confirmando que ésta no sólo se encuentra gobernada por cuán rápido un individuo pueda viajar, sino también por la hora en que el individuo realiza el viaje, y el tiempo disponible para viajar y desarrollar otras actividades.

2.3 Factores que influyen la elección de hora de inicio de los viajes

Como se indicó en la sección 2.2, el SM de Small (1982) permite estudiar los diferentes factores que pueden influenciar la elección de hora de inicio de viaje. La importancia del SM radica en la inclusión de términos relacionados con la programación de las actividades dentro de las funciones de utilidad de un modelo de elección discreta. Los términos que capturan la programación de actividades, se conocen en la literatura como demoras horarias (SD, *Schedule Delays*), y son un desarrollo motivado por el trabajo de Vickrey (1969). Los términos de SD indican la cantidad de tiempo que una persona llega tarde o temprano a su destino respecto a una hora de llegada deseada.

Diversos estudios han comprobado que el SM permite establecer de forma exitosa los compromisos entre el tiempo de viaje y los términos asociados a la demora horaria. El SM se derivó a partir de un modelo de asignación de tiempo, que desde el punto de vista microeconómico consiste en modelar la demanda de transporte urbano mediante la teoría económica del consumo introduciendo tiempo y espacio como recursos.

Small (1982) definió cuatro tipos de bienes en su modelo: tiempo de ocio (denotado por l), tiempo de trabajo (denotado por W), tiempo de consumo (denotado por t), y un bien de consumo generalizado (denotado por G y expresado en términos monetarios). El tiempo de consumo se refiere a aquellas actividades que no proporcionan directamente utilidad a los individuos, y que son complementarias al consumo generalizado, al ocio y al trabajo. El viaje al trabajo puede considerarse como una actividad de consumo complementario, la cual se realiza a un cierto costo (denotado por c) y a una hora determinada del día (denotada por s). En el caso del viaje al trabajo, s puede referirse a la hora de inicio del viaje, la hora de finalización del mismo o a la hora de entrada a un área congestionada en particular.

Para incorporar las preferencias de realizar las actividades de consumo a diferentes horas del día, Small (1982) incluyó a s en la función de utilidad y consideró al costo c y al tiempo t como funciones de s (para tener en cuenta su variación respecto a la hora de inicio de viaje). La formulación de Small (1982) incorporó las limitaciones de horarios impuestas por los empleadores por intermedio de una restricción explícita de horarios (denotada por una función F).

El modelo de asignación de tiempo planteado por Small (1982) plantea que el individuo maximiza su utilidad escogiendo las variables G , l y W a determinado horario s , teniendo en cuenta una serie de restricciones relacionadas con los horarios y la disponibilidad total de dinero y tiempo. Matemáticamente, este modelo puede representarse como:

$$\text{Max } U(G, l, W, s) \quad (17)$$

sujeto a

$$G + c(s) = Y + wW \quad (18)$$

$$l + W + t(s) = T \quad (19)$$

$$F(s, W; w) = 0 \quad (20)$$

donde U representa la utilidad que obtiene el individuo debido al consumo de G , l , y W en un determinado horario s . La restricción expresada en la ecuación (18) se refiere a la disponibilidad total de dinero para el consumo de bienes; w representa la tasa salarial del individuo, que al multiplicarse por el tiempo de trabajo (W) entrega el ingreso producto del trabajo de la persona; Y representa otros ingresos fijos. La restricción expresada en la ecuación (19) corresponde a la disponibilidad de tiempo total de la persona (T), para distribuir en términos de tiempo de ocio (l), tiempo de trabajo (W) y tiempo de viaje a una hora determinada, $t(s)$. Los ingresos fijos Y , la tasa salarial w , y el tiempo total disponible T son variables exógenas.

El lagrangiano para el modelo planteado por Small (1982) puede escribirse como:

$$L = U + \lambda[Y + wW - G - c(s)] + \mu[T - l - W - t(s)] - \nu F(s, W; w) \quad (21)$$

donde λ , μ , y ν denotan los multiplicadores de Lagrange para las restricciones de ingreso, tiempo y de horario, respectivamente. Estos multiplicadores representan la variación de la función objetivo (es decir, la utilidad) evaluada en el óptimo, debido a una variación marginal en el valor de la correspondiente restricción. En este contexto, λ representa la utilidad marginal del ingreso, μ la utilidad marginal del tiempo y ν la utilidad marginal relacionada con la relajación de las restricciones de horario.

Si la función U se conoce analíticamente, el problema de maximización puede resolverse en términos del consumo del bien generalizado (G) y las asignaciones de tiempo de ocio (l) y tiempo de trabajo (W), condicionales en los horarios (s). En el contexto de un

viaje al trabajo, al remplazar G^* , l^* y W^* en la función de utilidad, se podría obtener la máxima utilidad para un horario (s) dado en que se inicia el viaje al trabajo. Por definición, esto sería la función de utilidad indirecta condicional (V_s) que comanda la elección de la hora de inicio de los viajes y puede escribirse como (22):

$$V_s = V(c(s), t(s), s) = U(G^*(s), l^*(s), W^*(s), s) \quad (22)$$

donde V_s denota a la utilidad evaluada en el óptimo; es decir, representa la máxima utilidad que un individuo puede obtener si inicia su viaje en s , optimizando su consumo del bien generalizado, y la asignación de tiempo para ocio y trabajo.

A pesar de que no se conoce analíticamente la función U , que debe ser maximizada de acuerdo a la elección de horarios discretos de inicio de viajes, es posible realizar una simplificación mediante una aproximación de primer orden como (23):

$$U = V_o(G, l, W, s) + \frac{\partial U}{\partial G} \cdot G + \frac{\partial U}{\partial l} \cdot l + \frac{\partial U}{\partial W} \cdot W \quad (23)$$

donde V_o representa una función constante que contiene atributos no incluidos dentro del modelo de asignación de tiempo y términos constantes relacionados con las variables G , l y W en los puntos de aproximación (condiciones óptimas de consumo). Las derivadas de la utilidad respecto a G , l y W pueden obtenerse derivando las condiciones de optimalidad de primer orden del problema de maximización (17-20), usando el lagrangiano descrito en (21). Trabajando con las derivadas de la utilidad respecto a G , l y W , y teniendo en cuenta las restricciones de ingreso (18) y tiempo (19), la ecuación (23) puede rescribirse como (24) y (25):

$$V_s = V_o(G, l, W, s) + \lambda [Y + wW - c(s)] + \mu [T - W - t(s)] + W \left[\mu - w\lambda - v \frac{\partial F}{\partial W} \right] \quad (24)$$

$$V_s = V_o'(G, l, W, s) - \lambda \cdot c(s) - \mu \cdot t(s) - v \cdot \frac{\partial F}{\partial W} \cdot W \quad (25)$$

Small (1982) propuso que la ecuación (25) se podía estimar mediante técnicas de elección discreta si se especificaba en forma conveniente. El problema de elección se enfoca, entonces, como la programación del viaje al trabajo usando un modo de transporte determinado, para llegar al trabajo dentro de un periodo de tiempo discreto definido como s . Si se supone que la red de transporte se encuentra sujeta a congestión, entonces el tiempo

de viaje a diferentes horas del día puede ser representado por una curva $t(s)$, para el individuo q .

Dado que la situación evaluada por Small (1982) no consideraba cobros por congestión y que los costos de operación de los vehículos varían relativamente poco con la congestión, la variación del costo respecto a las horas de inicio de los viajes puede ser despreciada. Es decir, $c(s) = 0$.

Con el fin de considerar las restricciones de horario relacionadas con la duración del trabajo y la programación del viaje al trabajo para llegar a un periodo discreto determinado s , Small (1982) definió la variable *demora horaria*, como la diferencia entre la hora de llegada en el período s y un horario de llegada preferido (que supuso igual a la hora oficial de inicio del trabajo, exógenamente determinada). Así, la demora horaria puede expresarse como (26):

$$SD_i = Hr \text{ llegada observada} - Hr \text{ llegada preferida} \quad (26)$$

Según Small (1982), llegar temprano ($SDE = SD < 0$) sugiere malgastar tiempo o usarlo en actividades menos productivas a las preferidas para esa hora, que disminuyen la utilidad de un individuo. Por otro lado, llegar tarde ($SDL = SD > 0$) tiene, para la mayoría de los trabajadores, repercusiones mas severas que llegar temprano. SDE_i y SDL_i son variables que pueden representarse matemáticamente como (27) y (28):

$$SDE_i = \text{Max}\{-SD_i, 0\} \quad (27)$$

$$SDL_i = \text{Max}\{0, SD_i\} \quad (28)$$

Small (1982) también postuló que el impacto del comportamiento de programar horarios cuando existe congestión podía estudiarse mediante el compromiso entre las restricciones de horario, $\frac{\partial F}{\partial W} \cdot W$, y el tiempo de viaje $t(s)$. De esta manera, utilizó la especificación (29) para representar las potenciales restricciones de horario que puede enfrentar un viajero:

$$\frac{\partial F}{\partial W} \cdot W = \beta_{SDE} SDE_i + \beta_{SDL} SDL_i + \delta_L d_L \quad (29)$$

donde d_L representa una penalidad extra si el individuo llega tarde a su trabajo, que se define como (30):

$$d_L = \begin{cases} 1 & \text{si } SDL_i > 0 \\ 0 & \text{si } SDL_i = 0 \end{cases} \quad (30)$$

Los coeficientes β_{SDE} , β_{SDL} y δ_L reflejan las restricciones en los horarios del viajero, que pueden depender de factores como horas de trabajo, estatus familiar, ocupación, elección de modo de transporte y política del empleador frente a la flexibilidad en las horas de trabajo. Small (1982), probó otras especificaciones diferentes para la ecuación (29) pero no obtuvo resultados estadísticamente significativos para los coeficientes β_{SDE} , β_{SDL} y δ_L .

Small (1982), también sugirió que la elección de horario puede depender de características no observables como la disponibilidad de contar con un lugar cómodo para leer el periódico o para tomar desayuno, la posibilidad de trabajar desde casa, el deseo de pasar más tiempo con la familia, o la necesidad de llevar a algunos miembros de la familia a la escuela o al trabajo. Estos atributos pueden ser capturados por una correcta especificación de $V_o(G,l,W,s)$ en (23) o mediante el término de error asociado a cada alternativa de hora (s), en la estimación discreta. Small (1982) no incluyó dichos atributos dentro de su formulación, por lo que supuso implícitamente que eran capturados por el término de error. La especificación de la función de utilidad del modelo propuesto por Small (1982), puede escribirse entonces como (31):

$$V_i = \beta_{TT} TT_i + \beta_{SDE} SDE_i + \beta_{SDL} SDL_i + \delta_L d_L \quad (31)$$

donde el subíndice i se refiere a la alternativa (periodo de hora de inicio del viaje, antes denotado por s) y V_i a su utilidad representativa; β son los parámetros a estimar y TT_i indica el tiempo de viaje del individuo al salir en el periodo i . En concordancia con la formulación presentada en (25), se asume que todos los coeficientes tienen signo negativo por la des utilidad que se produce al llegar al destino a un horario diferente al deseado.

En un sistema de transporte, por lo general, existe cierta variabilidad en los tiempos de viaje que está principalmente asociada a la operación misma del sistema o a la congestión presente. La variabilidad en los tiempos de viaje no sólo depende de la relación entre el volumen de tráfico y la capacidad de la infraestructura, sino también de la distribución del tráfico en los diferentes periodos del día (de Jong *et al.*, 2003). Los

individuos incorporan la variabilidad en los tiempos de viaje en sus decisiones de viaje basados en la percepción del nivel de congestión y la experiencia de viajes anteriores. Con el fin de incorporar la incertidumbre asociada a los sistemas de transporte, Noland y Small (1995) extendieron el SM inicial (Small, 1982) y definieron la utilidad esperada al salir a cierta hora del día como se muestra en la ecuación (32):

$$E(U_i) = \beta_{TT} E(TT_i) + \beta_{SDE} E(SDE_i) + \beta_{SDL} E(SDL_i) + \delta_L P_L \quad (32)$$

donde $E(TT_i)$ indica el valor esperado de los tiempos de viaje presentados al individuo cuando se incorpora incertidumbre, $E(SDE_i)$ y $E(SDL_i)$ los valores esperados de las demoras horarias si el individuo llega temprano o tarde, respectivamente; finalmente, P_L es la probabilidad asociada al número de veces que el individuo llegaría tarde según los tiempos de viaje presentados.

Los modelos de elección de la hora de partida de viajes no sólo deben incorporar la información de los horarios de cada viaje y los niveles de servicio e incertidumbre asociada a ellos, sino también la duración de las actividades (Ettema *et al.*, 2004). El tiempo de participación en las actividades que realiza una persona durante un día, así como la información descriptiva sobre sus orígenes-destinos, también es relevante pues condiciona la realización de viajes, el orden de participación en actividades y la elección de las horas de inicio de los viajes. Para estimar buenos modelos, idealmente debe conocerse la secuencia de actividades a realizar, los lugares dónde se realizarán, y la flexibilidad de horario con que se cuenta para llevar a cabo cada actividad. Para incorporar la duración de las actividades y eventualmente su secuencialidad, el enfoque de modelación no sólo debe considerar los viajes por sí solos. La realización de otras actividades impone restricciones sobre la elección del viaje, por lo que es ideal considerar *tours* o sucesión de varios viajes, que tengan en cuenta la relación entre las diferentes actividades durante el día.

Con la idea de incluir el efecto de la duración de las actividades en el destino, de Jong *et al.* (2003) y Hess *et al.* (2007b) presentan ejemplos de SM que incluyen penalidades explícitas relacionadas con una menor o mayor duración de las actividades. La función de utilidad genérica al incluir estas penalidades debido a la duración de las actividades (definidas como PTD_i y PTI_i en las ecuaciones (34) a (36)), se muestra en la ecuación (33):

$$V_i = \beta_{TT} TT_i + \beta_C costo_i + \beta_{SDE} SDE_i + \beta_{SDL} SDL_i + \beta_{PTD} PTD_i + \beta_{PTI} PTI_i \quad (33)$$

donde:

$$PTD_i = Max\{-PT_i, 0\} \quad (34)$$

$$PTI_i = Max\{0, PT_i\} \quad (35)$$

$$PT_i = Duración actividad observada - Duración actividad preferida \quad (36)$$

Otros aspectos que no se deben dejar de lado al modelar la elección de horario de viaje son las características socioeconómicas del individuo, la información sobre las actividades a realizar en los destinos y la información de otras elecciones que puedan interactuar con la elección de hora (por ejemplo, elección de ruta y de modo).

La información socioeconómica es importante para la modelación, no sólo para caracterizar la muestra, sino también para considerar las restricciones inherentes a las responsabilidades sociales y económicas de las personas (Bajwa *et al.*, 2008). Por ejemplo, si un padre debe llevar a y recoger a sus hijos de la escuela a determinadas horas, estas pueden influir en la elección de la hora en que el padre realiza sus viajes y determina cuanto duran sus otras actividades. Información socioeconómica, como la edad, sexo, ingreso familiar, ingreso individual, estatus marital, ocupación, nivel de escolaridad, número de niños en el hogar, horas de trabajo y flexibilidad de horario, son útiles en los modelos de elección de la hora de inicio de viaje.

Por último, no es claro cuáles factores no observables podrían influenciar el proceso de elección de la hora de inicio de los viajes y se necesita mayor investigación en esta línea. Sin embargo, es esperable que incluir actitudes y percepciones relacionadas con el proceso de elegir la hora de inicio mejoren el poder explicativo de estos modelos.

2.4 Conclusiones (resumen de limitaciones encontradas)

En la literatura se encuentran modelos de elección de la hora de inicio de viaje que consideran el tiempo como una variable continua o discretizan el tiempo en periodos de longitud determinada por el modelador. Si bien los modelos que consideran el tiempo como una variable continua tienen la ventaja de solucionar los problemas de correlación

entre alternativas y la definición de periodos de duración arbitraria, su uso no está aun totalmente difundido debido a dificultades en la interpretación económica de sus resultados y del enfoque de modelación. Sin embargo, la mayor desventaja de los modelos que consideran al tiempo como variable continua es la poca disponibilidad de *software* comercial que permita masificar su compleja estimación. En esta tesis, se considerará el tiempo como una variable discreta pues el enfoque de modelación está basado en la teoría de la utilidad aleatoria, la que tiene un sustento económico razonable, aceptado y además compatible con el resto de elecciones de transporte en el marco de un modelo de planificación estratégica. Adicionalmente, a pesar de los problemas inherentes a discretizar el tiempo, estos modelos continúan siendo los más usados en la práctica y la aplicación de técnicas avanzadas de elección discreta pueden ser beneficiosa en términos de capacidad explicativa y para avanzar hacia una formulación más adecuada de los modelos.

Los datos de PD y PR ofrecen diferentes ventajas y desventajas cuando son usados en la estimación de modelos de elección horaria. Los datos de PD son adecuados para cuantificar los compromisos entre ciertos atributos que podrían estar altamente correlacionados en el caso de PR; sin embargo, presentan la desventaja que podrían ser menos confiables en un contexto predictivo. Por su parte, a pesar de que los datos de PR podrían ser mejores en el contexto predictivo, tienen desventajas asociadas a su costo de recolección, y a la obtención e identificación de los atributos de las alternativas disponibles. Con el fin de aprovechar las ventajas de los datos de PR y PD, en esta tesis se plantea formular modelos combinando ambos tipos de datos. Adicionalmente, otra componente innovadora de esta tesis es la inclusión de actitudes y percepciones en los modelos de elección horaria, tema que prácticamente no ha sido abordado en la literatura relacionada con este tema.

La literatura tampoco es explícita respecto a las metodologías de recolección de datos más adecuadas para modelar la elección de hora de inicio de viaje. La estimación de este tipo de modelos usualmente requiere de datos específicos que no están disponibles fácilmente en grandes encuestas de movilidad (Ben-Akiva y Abou-Zeid, 2012). De hecho, ni siquiera existe un consenso respecto al diseño experimental más adecuado para generar experimentos de elección relacionados con la hora de inicio de viajes (Arellana *et al.*,

2012c; Tseng *et al.*, 2011). En esta tesis se propone una metodología de encuesta específica para recolectar datos de PR, PD y variables latentes (VL) que intenten capturar diferentes factores que potencialmente influyen en la decisión de la hora de inicio de viajes.

Por otra parte, desde el punto de vista de la demanda, el enfoque de modelación tradicional en el campo es el SM de Small (1982). Este modelo y sus extensiones, consideran típicamente atributos como demoras horarias, variables socioeconómicas, condiciones de viajes (tiempos, costos), variabilidad inherente a los tiempos de viaje, efecto de la duración de actividades y flexibilidad de horarios. En esta tesis, se investiga la inclusión de diferentes atributos que han sido relevantes en estudios previos de elección de horario en la especificación del SM. Principal atención se prestará a la forma más apropiada de incluir la flexibilidad en los horarios del trabajo y a la incidencia de esta variable en la modelación.

La tesis abordará los viajes al trabajo debido a su gran influencia en la programación de las actividades de un día laboral, y por ser realizados – comúnmente – durante las horas punta del día. Para considerar el efecto de la duración de la jornada de trabajo en la elección de horario, se consideran variables que permitan capturar dicho efecto y además considerar *tours* simples compuestos por dos viajes consecutivos basados en el trabajo (es decir, el viaje hacia el trabajo y el viaje desde el trabajo).

Por último, en la literatura se encuentra que la elección de horario de inicio de viaje puede afectar o ser afectada por otras elecciones de transporte. En particular, su relación con las elecciones de ruta y modo para un viaje determinado han sido las más estudiadas en el contexto de visión del problema como un equilibrio oferta-demanda. Dado que esta tesis estará enfocada mayormente en lograr una mejor caracterización de la demanda, desde el punto de vista de la elección de horario de viajes en el contexto de *tours* basados en el trabajo, se estimará un modelo conjunto de elección de hora y modo de transporte aprovechando la riqueza de los datos. Este modelo conjunto en el contexto de *tours*, es una contribución interesante de esta tesis ya que extiende el enfoque más común en la literatura.

3 DISEÑO Y METODOLOGIA PARA LA RECOLECCION DE DATOS

Este capítulo presenta un resumen del diseño y metodología para la recolección de los datos utilizados en el desarrollo de esta tesis. La información fue recolectada en Santiago durante el año 2011, teniendo en cuenta que algunos datos requeridos para modelar la hora de inicio de viaje no son parte de las grandes EOD con que usualmente se calibran los modelos estratégicos de transporte urbano.

La información requerida para modelar adecuadamente la elección de la hora de inicio de viaje suele ser muy detallada y difícil de recolectar. Idealmente se requieren datos precisos acerca de las horas preferidas de llegada a los destinos, del trabajo de los encuestados, del tiempo de participación en diversas actividades durante el día, y de sus actitudes y percepciones que puedan influir en la elección de la hora de inicio de sus viajes.

Teniendo en cuenta la cantidad de datos a recolectar y la diversidad de factores que pueden influir en la elección, se decidió obtener la información utilizando diferentes enfoques (PR, PD, y VL) y diversas técnicas de recolección. Las encuestas fueron aplicadas en diferentes centros de empleo en Santiago, la mayoría ubicados en el centro de la ciudad. La metodología para recolectar los datos se definió a partir de la realización de estudios cualitativos (grupos focales y opiniones de expertos) preliminares y la ejecución de encuestas piloto.

3.1 Actividades previas al diseño de la encuesta

Inicialmente, se decidió realizar una encuesta preliminar, vía internet, con el objetivo de profundizar el conocimiento acerca de los atributos o factores que podrían influenciar la elección del horario de inicio de viaje, y apoyar la definición de los atributos a ser utilizados en el diseño de la encuesta definitiva. La encuesta preliminar fue respondida por cerca de 250 personas que trabajaban y que posteriormente no participaron en la encuesta final. Los encuestados debían calificar, en una escala de 1 a 7, la importancia de una serie de atributos en sus elecciones de hora de inicio de viajes laborales y no laborales durante un día laboral (Arellana *et al.*, 2012b). El listado de los atributos presentados a los encuestados, según tipo de viaje, se presenta en el Anexo A.

Las calificaciones otorgadas por los individuos a cada uno de los atributos presentados, fueron analizados mediante modelos de ecuaciones estructurales usando el software MPLUS (Muthén y Muthén, 2007). Los resultados obtenidos muestran que la elección de la hora de inicio de viajes laborales y no laborales parece estar influenciada principalmente por la congestión, la programación de actividades a realizar, las condiciones de los viajes, y por factores externos que puedan afectarlos.

Las personas pueden percibir factores como la congestión, a través de atributos como tiempo de viaje, variabilidad del tiempo de viaje, e información externa acerca de la congestión en la red. La programación de actividades a realizar puede capturarse a través de atributos como las horas de inicio y fin de las actividades, la cantidad de horas dedicadas a cada actividad durante el día, y la secuencia de las actividades planeadas en cada día. Las condiciones de los viajes vienen representadas por atributos como su costo, confort, seguridad y modo de transporte a utilizar. Por último, factores externos como las condiciones climáticas, viajar sólo o acompañado, actitudes y percepciones individuales y otras restricciones impuestas por otras actividades o circunstancias del viaje, pueden afectar la elección de hora de inicio de viaje de las personas.

Otro resultado interesante del análisis de los datos recolectados en la encuesta preliminar, fue la importancia relativa, para cada tipo de viaje, de cada uno de los factores considerados en el estudio. La programación de las actividades durante el día parece ser más importante en el caso de la elección del horario de viajes relacionados con el trabajo que en el caso de viajes no laborales. Los viajes relacionados con el trabajo pueden tener restricciones importantes respecto a sus horas de inicio o de fin debido a las restricciones laborales (entrada o salida del trabajo) impuestas por los empleadores y por la flexibilidad de horario con que cuentan las personas. Los viajes no laborales tienden a ser más flexibles, pues usualmente se pueden llevar a cabo a diferentes horas del día o incluso ser efectuados en otro día, cuando las personas posean disponibilidad de tiempo (Arellana *et al.*, 2012b).

Durante el diseño de los instrumentos para la encuesta final, se tuvo en cuenta los factores que resultaron relevantes en la encuesta preliminar y se procuró incluir atributos que permitieran representar dichos factores en las componentes de PR, PD y VL de los

modelos. Para asegurar la relevancia de los atributos y la claridad de las preguntas en cada componente de la encuesta, se hizo una extensiva revisión bibliográfica y se solicitó la opinión de expertos en el área, quienes modificaron los instrumentos como consideraron pertinente.

Luego del juicio de expertos, la comprensión de los instrumentos de la encuesta fue puesta a prueba mediante la realización de tres grupos focales, integrados por personas con diferentes características socioeconómicas, que usaban diferentes modos de transporte, y trabajaban en diferentes centros de empleo ubicados en diversos lugares de la ciudad. Adicionalmente, los grupos focales fueron usados para conocer la opinión de potenciales encuestados acerca de la metodología de recolección a utilizar para las diferentes componentes de la encuesta y seleccionar el incentivo más atractivo para los encuestados.

Finalmente, las actividades previas de diseño culminaron con una encuesta piloto que incluyó todas las componentes de la encuesta, y utilizó la misma metodología de recolección que la encuesta final. La encuesta piloto se aplicó hasta que 30 personas completaran todas sus componentes. La experiencia con la encuesta piloto permitió calcular tasas de no respuesta esperadas e identificar preguntas problemáticas, con el fin de mejorar el diseño de la encuesta final.

3.2 Características generales del instrumento de diseño

Una de las principales restricciones para la recolección de datos específicos para la modelación de la hora de inicio de los viajes son sus altos costos en términos de tiempo y dinero. En particular, se consideró restricciones temporales y de presupuesto en el desarrollo de la metodología, la selección del tamaño de muestra, y la cantidad de datos a analizar y recolectar en las encuestas. En esta tesis, las encuestas se enfocaron principalmente a recolectar información para estudiar la elección de hora para los viajes basados en el trabajo, dado que estos representan cerca del 70% de los viajes totales en la hora punta de la mañana durante un día laboral (DICTUC, 2006). Además, se definió al individuo como unidad de muestreo (en lugar del hogar), debido a los mayores costos y complejidad asociados a la recolección de datos cuando se define al hogar como unidad muestral. De esta manera, la única condición requerida para participar en la encuesta fue que el individuo trabajara.

El objetivo final de la encuesta fue obtener información completa y robusta para estudiar las elecciones de hora de inicio de viaje a mínimo costo. Debido a la cantidad y al tipo de información requerida para lograr el objetivo de la encuesta, se decidió recolectar la información en tres etapas. Esto tiene como principal ventaja la disminución del cansancio y/o aburrimiento de los encuestados asociado a tener que responder encuestas demasiado largas. Además, ofrece la posibilidad de validar los datos entre las distintas etapas de aplicación de la encuesta, obtener información que tal vez no haya sido reportada en etapas anteriores, y personalizar algunas partes de la encuesta. Las principales desventajas de la recolección por etapas son: la potencial deserción de encuestados entre etapas; la posibilidad que se presenten cambios importantes en el sistema de transporte, y los mayores costos asociados a los encuestadores, quienes típicamente deben contactar a los encuestados varias veces.

La primera etapa de la encuesta consistió en una entrevista presencial en los lugares de trabajo de los encuestados, donde se recopiló información socioeconómica, del empleo, de los atributos que las personas consideraban relevantes para la elección de hora, y de la programación de las actividades que los encuestados planeaban realizar en el siguiente día hábil. La segunda etapa se realizó al siguiente día hábil después de la realización de la primera etapa y consistió en el llenado, vía web, de un diario de viajes utilizando un marco recordatorio de actividades (Ampt y Ortúzar, 2004). Finalmente, la tercera parte consistió en la realización presencial, también en los lugares de trabajo de los encuestados, de dos experimentos PD (basados en las respuestas a las dos primeras etapas) y un cuestionario para la posterior conformación de variables latentes, que contenía una pregunta para obtener información del ingreso de los encuestados.

La decisión de tomar la segunda etapa de la encuesta usando formularios web de auto llenado, se debió principalmente a restricciones de costo. No obstante, en los grupos focales se había llegado a la conclusión que las personas podían entender el formulario sin la ayuda de un encuestador y además preferían diligenciar el formulario cuando tuvieran tiempo disponible, a fin de tener mayor privacidad.

Por otra parte, las etapas 1 y 3 se realizaron mediante entrevistas presenciales por ser la metodología que brinda mayores tasas de respuesta (Ortúzar, 2006). Con el fin de

mejorar la tasa de respuesta, se utilizó como encuestadores a mujeres estudiantes muy bien entrenadas, y las entrevistas se efectuaron directamente en los lugares de trabajo de los encuestados. En particular, los centros de empleo donde se realizaron las encuestas tuvieron como característica que el autor, alguno de sus profesores supervisores o la Universidad Católica, tuvieran alguna relación directa con al menos un miembro de su departamento administrativo. Experiencias previas en Chile han demostrado que es posible obtener altas tasas de respuesta cuando se encuesta directamente en los lugares de trabajo luego de que la encuesta es recomendada por un jefe o algún miembro del departamento administrativo (Cantillo *et al.*, 2006; Caussade *et al.*, 2005; Hojman *et al.*, 2005). A pesar de que las encuestas se realizaron en pocos lugares de trabajo (con el fin de disminuir los costos de implementación), la Figura 3-1 muestra que los viajes al trabajo provienen de todas las comunas de Santiago.

3.3 Selección de la muestra y estrategia de recolección

Los participantes de la encuesta fueron seleccionados aleatoriamente dentro de los centros de empleo donde se aplicaron las encuestas. La cantidad total de personas a encuestar fue definido de acuerdo a la ecuación (37) (Smith, 1979):

$$n = \frac{s^2 \cdot Z_\alpha^2}{d^2} \quad (37)$$

donde s es la desviación estándar de la variable bajo estudio, d el error absoluto deseado y Z_α es el valor de la Normal estándar para un nivel de confianza α .

El objetivo de la encuesta fue obtener una muestra indicativa que permitiera obtener una distribución de viajes cada 15 minutos, similar a la reportada por la EOD para Santiago del año 2006 (DICTUC, 2006).

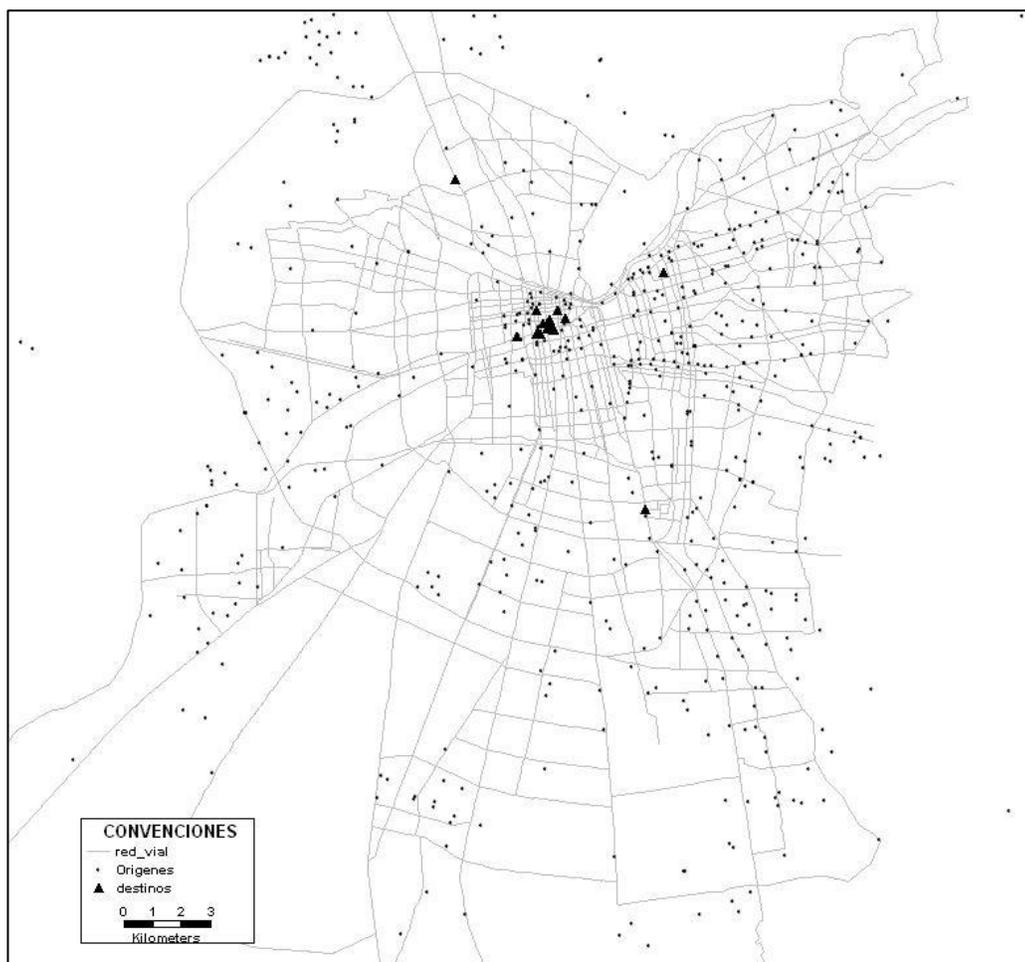


Figura 3-1: Ubicación de los hogares y destinos de los encuestados
Fuente: Elaboración propia

De acuerdo a esta EOD, durante el período crítico de 15 min se produjo el 13,75% del total de los viajes al trabajo en la hora punta mañana, considerando que el periodo punta de la mañana transcurre entre las 7:00 hrs y las 9:15 hrs. Suponiendo una distribución binomial (éxito es el período crítico, fracaso es lo contrario), la desviación estándar s de la partición horaria observada fue calculada como lo indica la ecuación (38):

$$s = \sqrt{p(1-p)} = \sqrt{0,1375(1-0,1375)} = 0,344 \quad (38)$$

Asumiendo un error absoluto deseado de 0,02 y un 90% para el nivel de confianza, el tamaño de la muestra para la encuesta debería ascender a 487 individuos. Sin embargo,

como es esperable una cierta tasa de deserción en encuestas con diferentes etapas de recolección (Stopher *et al.*, 2006), se decidió tomar una muestra de 600 individuos para la primera etapa. Como la tasa de no respuesta entre las fases 1 y 2 de la encuesta piloto fueron cercanas al 20%, el número final de encuestados se calculó como 20% mayor que el tamaño de muestra calculado según la ecuación (37).

La estrategia de recolección comprendió dos visitas a los encuestados (etapa 3 y 7), el llenado de un formulario vía web (etapa 5), y una serie de cartas recordatorias y de agradecimiento para mantener interesados a los participantes durante la encuesta (ver Tabla 3-1). Adicionalmente, se intentó mantener el mismo encuestador para cada participante y se incluyó a cada persona en la rifa de cinco premios de \$100.000 al término de cada visita. Cabe aclarar que los cinco premios de \$100.000 constituyeron el incentivo preferido en los grupos focales y la encuesta piloto.

3.4 Fuentes de información

La encuesta tiene cuatro fuentes de información: datos PR recolectadas entre los pasos 3 y 5 (Tabla 3-1), datos PD recolectados en el paso 7 (Tabla 3-1), datos para formar variables latentes recolectados en los pasos 3 y 7 (Tabla 3-1) y, por último, mediciones precisas de las variables de nivel de servicio.

3.4.1 Preferencias reveladas

El instrumento para la recolección de los datos PR se basó en la EOD del 2001 para el Gran Santiago (Ampt y Ortúzar, 2004). Su recolección se realizó en dos partes. La primera consideraba una entrevista personalizada realizada por un encuestador entrenado en el lugar de trabajo de cada encuestado, mientras que la segunda parte se recolectó usando un formulario de auto llenado implementado en SurveyMonkey®. Esta comenzaba con una pequeña introducción del proyecto, donde se precisaba el objetivo del estudio, quién lo estaba realizando, la descripción de la participación de los encuestados en el estudio, y la firma de un consentimiento informado (ver Anexo B).

Tabla 3-1: Estrategia de recolección de datos

Paso	Actividad	Contacto	Fecha
1	Carta de reclutamiento empresas	E-Mail	Noviembre 2010
1 ^a	Recordatorio 1 – Carta de reclutamiento empresas	E-mail	1 semana después de la etapa 1
1b	Recordatorio 2 – Carta de reclutamiento empresas	Teléfono	1 semana después de la etapa 1 ^a
2	Selección de la muestra y programación de primera visita	Correo electrónico	1 semana antes de la primera visita
3	Visita 1 (componte PR parte 1) – Charla motivacional u entrega del registro general de salidas	Entrevista personal	Diciembre – Febrero (1 día antes de la etapa 4)
4	Recordatorio para llenar el registro general de salidas y agradecer por haber completado la parte 1 de la componente PR	Correo electrónico	Diciembre – Febrero (día de viaje)
5	Envío de componente PR parte 2 a ser diligenciado via web. Motivación para completar PR parte 2	Correo electrónico	Diciembre – Febrero (1 día después de etapa 4)
5 ^a	Recordatorio 1 – completar parte 2 PR	Correo electrónico	1 semana después de etapa 4, si es necesaria
5b	Recordatorio 2 – completar parte 2 PR	Correo electrónico	2 semanas después de etapa 4, si es necesaria
5c	Recordatorio 3 – completar parte 2 PR	Correo electrónico	3 semanas después de etapa 4, si es necesaria
5d	Recordatorio 4 – completar parte 2 PR	Correo electrónico	4 semanas después de etapa 4, si es necesaria
6	Anuncio de los ganadores de las rifas, agradecimiento por completar PR parte 2 e información acerca de visita 2	Correo electrónico	Mayo
7	Visita 2 (componente PD + cuestionario de variables latentes)	Entrevista personal	Junio
8	Anuncio ganadores última rifa, agradecimiento por completar encuesta	Correo electrónico	Julio

Fuente: Elaboración propia

Luego de la introducción, se presentaron las siguientes secciones dentro del instrumento de encuesta de la primera etapa:

Información personal: La encuesta fue anónima y el uso de los datos se rotuló como confidencial, por lo que el encuestador asignó un número único de identificación, que fue

asociado al correo electrónico del encuestador para ser contactado en etapas subsecuentes de la encuesta. Adicionalmente, se recolectó información como edad, género, nivel de estudios, y posesión de licencia de conducir.

Información del hogar: Dirección (esquina más cercana), comuna, tamaño familiar, número y tipo de vehículos motorizados y no motorizados en el hogar, y número de miembros del hogar con licencia de conducir.

Información del empleo: Número de horas trabajadas a la semana, número de veces en la semana que trabaja en el mismo lugar, posibilidad de trabajar desde la casa, si debe trabajar una cantidad de horas fijas a la semana, posibilidad de trabajar una cantidad de horas diferentes cada día, facilidades provistas por el empleador (como estacionamiento liberado, otorgamiento de vehículo, subsidio de bencina, bus para transportar a los trabajadores de la empresa).

Factores que influyen la elección de hora: En esta sección se presentaron los mismos dos cuestionarios utilizados en la encuesta preliminar descrita en 3.1. Esta información no fue usada en la modelación presentada en esta tesis, pero se piensa ocupar para fines de segmentación de la muestra en trabajos futuros. La información también se está usando en un proyecto de colaboración con la Universidad de Western Australia acerca del estudio de las preferencias relacionadas con la elección de la hora de inicio de viajes laborales y no laborales, que se encuentra en etapa de desarrollo (Arellana *et al.*, 2012b).

Plan de actividades para el día de encuesta: En esta etapa se solicitó a los encuestados indicar las actividades que planeaban realizar el día en que llenarían el diario de viajes. Los encuestados debían indicar el modo y las horas de inicio y fin de cada actividad reportada. Las horas de inicio y fin reportadas se utilizaron para definir las horas preferidas de llegada e inicio de los viajes, y la duración preferida de las actividades en la encuesta de PD (ver Anexo B).

Al finalizar las secciones anteriores, los encuestadores explicaban la metodología a seguir para recopilar el resto de información en las restantes etapas de la encuesta. Además, entregaban el formulario “registro general de salidas” (ver Anexo B), para que los encuestados reportaran los viajes realizados durante el siguiente día hábil y lo utilizaran

para completar el formulario de auto llenado, que correspondía a la segunda parte de la componente PR.

La segunda parte de la componente PR comenzaba con un correo electrónico enviado a cada encuestado para recordarle el registro de sus viajes en el formulario “registro general de salidas”. Luego, al final del día o durante el día siguiente, se enviaba otro correo electrónico con el enlace al formulario de auto llenado implementado en SurveyMonkey®. Este formulario pedía reportar características detalladas de todas las etapas de los viajes realizados durante el día, usando el “registro general de salidas”, para asegurar buenos datos para el análisis de la elección de modo y hora (Ortúzar, 2006). Específicamente, las principales características solicitadas fueron: modo de transporte, hora de inicio y fin del viaje, tiempo de viaje, tiempo de espera, tiempo de acceso, costos, ruta elegida, origen, destino, y propósito del viaje. Adicionalmente, se recopiló información acerca de las horas oficiales de inicio y fin de la jornada laboral, la cantidad de minutos en que podían atrasar o adelantar los viajes relacionados con el trabajo, y la posibilidad de ser penalizados si adelantaban o atrasaban su hora de viaje.

También se envió correos electrónicos recordando el llenado del “registro general de salidas” y del formulario web con las características detalladas de cada viaje, a aquellos individuos que no completaron el formulario web de la segunda parte tras un período de cuatro semanas consecutivas. Se envió correos electrónicos recordando completar ambos formularios, con un día de separación entre ellos, para evitar problemas asociados al olvido de los detalles de los viajes cuando la diferencia entre los días de llenado de los formularios era muy grande.

3.4.2 Obtención de niveles de servicio en la red de transporte

Con el fin de enriquecer las observaciones de PR y generar algunos de los atributos utilizados en el diseño de los experimentos de PD, se obtuvo niveles de servicio con alta precisión para los distintos modos de transporte en la ciudad. Los niveles de servicio se obtuvieron mediante mediciones GPS usando vehículos instrumentados que circularon por las redes de transporte público y privado, simultáneamente a la recolección de los datos de PR. El uso de medidas precisas de nivel de servicio en la modelación de demanda por

transporte es muy poco común en la práctica (Daly y Ortúzar, 1990), y su obtención y uso en el contexto de modelos de hora y modo, es otro aporte de la presente tesis.

Los niveles de servicio sobre la red de transporte privado fueron obtenidos a través de mediciones directas realizadas por vehículos instrumentados transitando por toda la red a diferentes horas del día. Las rutas donde se realizaron las mediciones se determinaron teniendo en cuenta las vías utilizadas, y los orígenes y destinos reportados por los encuestados.

La información requerida para obtener los niveles de servicio en las redes de transporte público fue proporcionada por Metro y Transantiago S.A. Los niveles de servicio para el modo metro fueron procesados y entregados al mayor nivel de desagregación disponible en la empresa operadora. En el caso de los buses, los niveles de servicio fueron obtenidos a partir de las posiciones de todos los vehículos operando en Santiago cada 30 segundos. Debido a la enorme cantidad de datos disponibles y a que la red de buses en Santiago es densa, el manejo de los datos para el cálculo y procesamiento de niveles de servicio se constituyó en un proceso altamente desafiante y complejo, en términos de tiempo y costo.

Para procesar los datos disponibles, se desarrolló una metodología que permite obtener niveles de servicio para diferentes agregaciones de tiempo y espacio (Arellana *et al.*, 2012e). La metodología es altamente flexible pues permite procesar gran cantidad de datos en poco tiempo, y su implementación no requiere de un software SIG comercial.

A pesar de que existen algunos procedimientos en la literatura que permiten calcular niveles de servicio en redes de transporte público a partir de datos GPS, la metodología utilizada en esta tesis se desarrolló con el fin de obtener niveles de servicio durante todo el día a nivel de paraderos. El autor no conoce otro procedimiento que permita obtener la información al nivel de desagregación espacial y temporal solicitado. La metodología desarrollada recientemente por Cortes *et al.* (2011) permite monitorear las velocidades de los buses usando una grilla definida por un diagrama espacio-tiempo que se debe fijar para cada ruta dada antes de iniciar los cálculos. Si bien el procedimiento de Cortes *et al.* (2011) puede entregar los resultados a cualquier agregación espacial constante que se defina, no es muy claro cómo adaptar su procedimiento para reportar niveles de

servicio entre paraderos de buses consecutivos, que tienen la particularidad de estar separados por distancias que no son constantes entre sí.

La metodología desarrollada consta de seis etapas (Arellana *et al.*, 2012e):

- i) Etapa preliminar: En esta fase los datos se revisan para evitar inconsistencias y se organizan para su procesamiento. Debido a la cantidad de datos GPS a analizar, esta etapa debió realizarse en un manejador de bases de datos.
- ii) Definición de rutas y cálculo de distancias en ruta: Dado que no se contaba con las distancias en ruta entre paraderos de un mismo servicio, fue necesario su cálculo a partir de la ubicación de todos los paraderos de la ciudad y la información acerca de los servicios de buses que usan cada paradero. Esta información también fue proporcionada por Transantiago. Los cálculos de las distancias en ruta se desarrollaron con un código Javascript utilizado por el servicio de Google MapsTM.
- iii) Asignación de datos GPS a cada servicio: Esta etapa consistió en identificar los registros GPS de cada vehículo y vincularlos con un recorrido (servicio) específico dentro de la red. Esta etapa se desarrolló en un manejador de base de datos.
- iv) Reducción de datos: Esta etapa consistió en seleccionar sólo los registros GPS utilizados en el cálculo de niveles de servicio de cada recorrido en particular. La selección de los datos a utilizar se realizó siguiendo una serie de reglas determinísticas (Arellana *et al.*, 2012e) codificadas en un manejador de bases de datos. Se seleccionó un sólo registro GPS para cada tramo de ruta formado por dos paraderos consecutivos. En particular, el registro GPS seleccionado fue aquel que se encontraba más cerca, en términos de distancia, al paradero aguas abajo.
- v) Predicción de los tiempos de pasada de buses por cada paradero: Para esto se introdujo el concepto de velocidad de aproximación. Esta se definió como la diferencia de distancias acumuladas en ruta (calculadas mediante Google MapsTM) dividida por la diferencia entre las horas del día en que se tomó cada registro consecutivo seleccionado en la etapa de reducción de datos, para una ruta en particular. El tiempo de pasada por el paradero aguas abajo fue definido como la suma de la hora del día en que se tomó cada registro; y la distancia que existe entre el

registro GPS seleccionado y el próximo paradero aguas abajo, dividida por la velocidad de aproximación. Esta etapa fue desarrollada en MATLAB.

- vi) Generación, agregación y reporte de los niveles de servicio a nivel de paraderos: A partir de las predicciones de las horas de pasada de los vehículos por cada paradero, fue posible estimar los tiempos de viaje y espera por servicio al nivel de agregación temporal deseado. Dado que se contaba con registros GPS durante una semana del año, la variabilidad de los niveles de servicio para cada período del día se obtuvo usando los cálculos de nivel de servicio para cada día. Esta etapa fue codificada en MATLAB y los resultados generados en tablas de MS Excel.

Cabe aclarar que los tiempos de espera por paradero obtenidos mediante el procedimiento anterior, sólo tienen en cuenta la oferta de transporte e ignora las restricciones impuestas por la demanda de transporte, relacionadas con: limitaciones de capacidad que pueden existir si un vehículo se encuentra totalmente ocupado al llegar a un paradero; comportamiento del individuo, quien puede decidir no subirse al bus para esperar el siguiente o tomar una ruta de bus alternativa; tasas de llegada no uniformes en los paraderos. Para superar la limitación anterior, los tiempos de espera por paradero fueron corregidos usando mediciones reales en diversos paraderos de la ciudad realizadas por DICTUC, en la misma fecha en que se realizó las encuestas.

El procedimiento de corrección de los tiempos de espera consistió en buscar un modelo de regresión entre los tiempos de espera, obtenidos a partir de las mediciones con GPS, y los tiempos de espera medidos por DICTUC, para el mismo paradero, servicio, y período del día. Se estimó regresiones lineales usando datos de paraderos que contaban con información de ambas fuentes para el mismo período del día. Los resultados de las regresiones se presentan en la Tabla 3-2, teniendo en cuenta que los datos se agruparon por tipo de servicio (alimentador o troncal) y período del día (período punta y fuera de punta).

Los resultados indican el factor de corrección aplicado a los tiempos de espera obtenidos por la metodología utilizada en esta tesis, los cuales fluctúan entre 0,3 y 0,7, de acuerdo al tipo de servicio y el período del día que se está evaluando. Los tiempos de espera obtenidos a partir de las mediciones de GPS son consistentemente mayores que los tiempos de espera medidos en campo, pues la llegada de los individuos al paradero puede

no seguir la distribución aleatoria supuesta en el computo de la metodología; además, las mediciones en campo consideran la posibilidad que un individuo aborde otro recorrido de transporte público si este le sirve a la persona y pasa antes por el paradero. Cabe recordar que las mediciones se efectuaron en el mes de Enero, cuando no hay presencia de escolares, por lo que las condiciones de congestión en el transporte público en hora punta, no son tan extremas y los participantes no se ven muy afectados por restricciones de capacidad al momento de abordar un bus.

Tabla 3-2: Resultados de las regresiones lineales para corregir tiempos de espera

Tipo de Servicio	Troncal		Alimentador	
	Punta	Fuera de punta	Punta	Fuera de punta
Coefficiente	0,388	0,598	0,563	0,631
R²	0,606	0,612	0,761	0,667
Observaciones	540	697	128	175

Fuente: Elaboración propia

3.4.3 Preferencias declaradas

La componente PD de la encuesta tiene el objetivo de incorporar el efecto, sobre las elecciones de modo y hora, de dos estrategias TDM que no se encuentran implementadas en la actualidad en Santiago. En particular, las dos estrategias consideradas son la tarificación vial y la implementación de un sistema de flexibilidad de horario para la entrada y salida del trabajo. La componente PD no fue aplicada a las personas que realizaban sus viajes caminando o en bicicleta, ni tampoco a las personas que reportaban viajes basados en el trabajo cambiando de transporte público a transporte privado (o viceversa) en sus diferentes etapas (en este último caso, porque la complejidad del diseño de PD aumentaba demasiado su complejidad).

La primera parte de la componente PD fue informativa, buscando entregar información de contexto indicando las razones para considerar estrategias TDM en Santiago, cuáles estrategias podrían ser empleadas y cuáles comportamientos se estudiarían a partir de sus respuestas. Además, dado que la realización de la componente PD fue mucho después que la componente PR, también se presentaron datos relacionados con los

viajes reportados por los encuestados con el fin de minimizar el olvido de las condiciones de viaje, en las cuales se basaron los experimentos de PD.

La segunda parte fue la presentación secuencial de dos experimentos de PD. El primero contenía cinco situaciones de elección y estaba enfocado a los viajes al trabajo realizados en la punta mañana. El segundo contenía ocho situaciones de elección y estaba enfocado al *tour* simple compuesto por el viaje de ida hacia el trabajo y el viaje de salida del trabajo. La presentación de ambos experimentos contenía cinco columnas. En la primera se entregaba información acerca de los atributos y en las cuatro restantes se presentaban cuatro opciones de elección por viaje (ver figuras 3-2 y 3-3). Entre la segunda y la cuarta columnas, las opciones correspondían a viajar en el mismo modo reportado, pero a diferentes horas: (i) la hora reportada por el encuestado (alternativa de referencia), (ii) antes y (iii) después de la hora reportada. La quinta columna siempre correspondió a viajar en un modo de transporte diferente, llegando al destino a una hora muy similar a la reportada. Para los usuarios de transporte privado este modo alternativo fue el transporte público, mientras que para el caso de usuarios de transporte público fue transporte privado, si lo tenían disponible, o un servicio de transporte público tipo taxi colectivo, en caso de no poder viajar en auto.

Es importante mencionar que la posición de las alternativas de elección de hora dentro de la pantalla, fue fijada aleatoriamente para evitar hábito en las elecciones de las personas.

Dado que el primer experimento se enfocaba sólo en los viajes hacia el trabajo, las cuatro opciones de elección representaban las cuatro alternativas disponibles para el individuo.

En el caso del segundo experimento, dado que se presentaban los viajes hacia y desde el trabajo, y que el objetivo principal del experimento era estudiar la elección de hora, las alternativas de elección se fijaron en diez. Nueve de ellas correspondían a viajar en el mismo modo, pero a horas diferentes. La última alternativa ofrecía la posibilidad de cambiar de modo; no obstante, si esta se escogía, debía hacerse en ambos viajes (Polak y Jones, 1994).

Situación de elección: 2	Opción A	Opción B	Opción C	Otro medio de transporte
Hora de inicio del viaje	7:06	8:21	9:20	8:25
Tiempo de viaje usual hacia el trabajo (Hora de llegada usual al trabajo)	50 (7:55)	59 (9:19)	41 (10:01)	53 (9:18)
Una vez a la semana su viaje demoraría (Hora de llegada a su trabajo)	60 (8:05)	74 (9:34)	54 (10:13)	64 (9:28)
Comodidad del viaje	Vehículo lleno, viaja de pie	Vehículo lleno, viaja de pie y con frecuencia debe esperar siguiente	Vehículo a medio llenar, viaja de pie	
Costo adicional del viaje (\$)	\$ 493	\$ 527	\$ 476	\$ 1,500
¿Cuál opción escogería para viajar?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

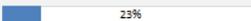





Figura 3-2: Ejemplo de presentación del experimento PD enfocado en viajes hacia el trabajo para usuarios de transporte público
Fuente: Elaboración propia

Situación de elección: 2	Opción A	Opción B	Opción C	Otro medio de transporte
Hora de inicio del viaje	7:21	8:11	9:06	8:15
Tiempo de viaje usual hacia el trabajo (Hora de llegada al trabajo)	32 (7:53)	54 (9:04)	38 (9:43)	45 (9:00)
Una vez a la semana su viaje demoraría (Hora de llegada al trabajo)	42 (8:03)	62 (9:12)	42 (9:47)	56 (9:11)
Comodidad del viaje de ida al trabajo	Vehículo lleno, viaja sentado	Vehículo lleno, viaja de pie y con frecuencia debe esperar siguiente vehículo	Vehículo a medio llenar, viaja sentado	
Costo adicional del viaje al trabajo (\$)	\$ 527	\$ 561	\$ 493	\$ 1,500
¿Cuál opción escogería para viajar hacia el trabajo?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hora de salida del trabajo	17:00	18:00	18:45	18:10
Tiempo de viaje usual después del trabajo (Hora de llegada al destino)	40 (17:39)	56 (18:56)	48 (19:33)	44 (18:54)
Una vez a la semana su viaje demoraría (Hora de llegada al destino)	51 (17:51)	65 (19:04)	53 (19:37)	55 (19:05)
Comodidad del viaje después del trabajo	Vehículo lleno, viaja sentado	Vehículo lleno, viaja de pie y con frecuencia debe esperar siguiente vehículo	Vehículo a medio llenar, viaja de pie	
Costo adicional del viaje después del trabajo (\$)	\$ 434	\$ 527	\$ 561	\$ 1,200
¿Cuál opción escogería para viajar después del trabajo?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>






Figura 3-3: Ejemplo de la presentación del experimento PD enfocado a toures basados en el trabajo para usuarios de transporte público
Fuente: Elaboración propia

La descripción del número de alternativas disponibles en el experimento PD enfocados en el *tour* simple basado en el trabajo se ilustra en la Figura 3-4.

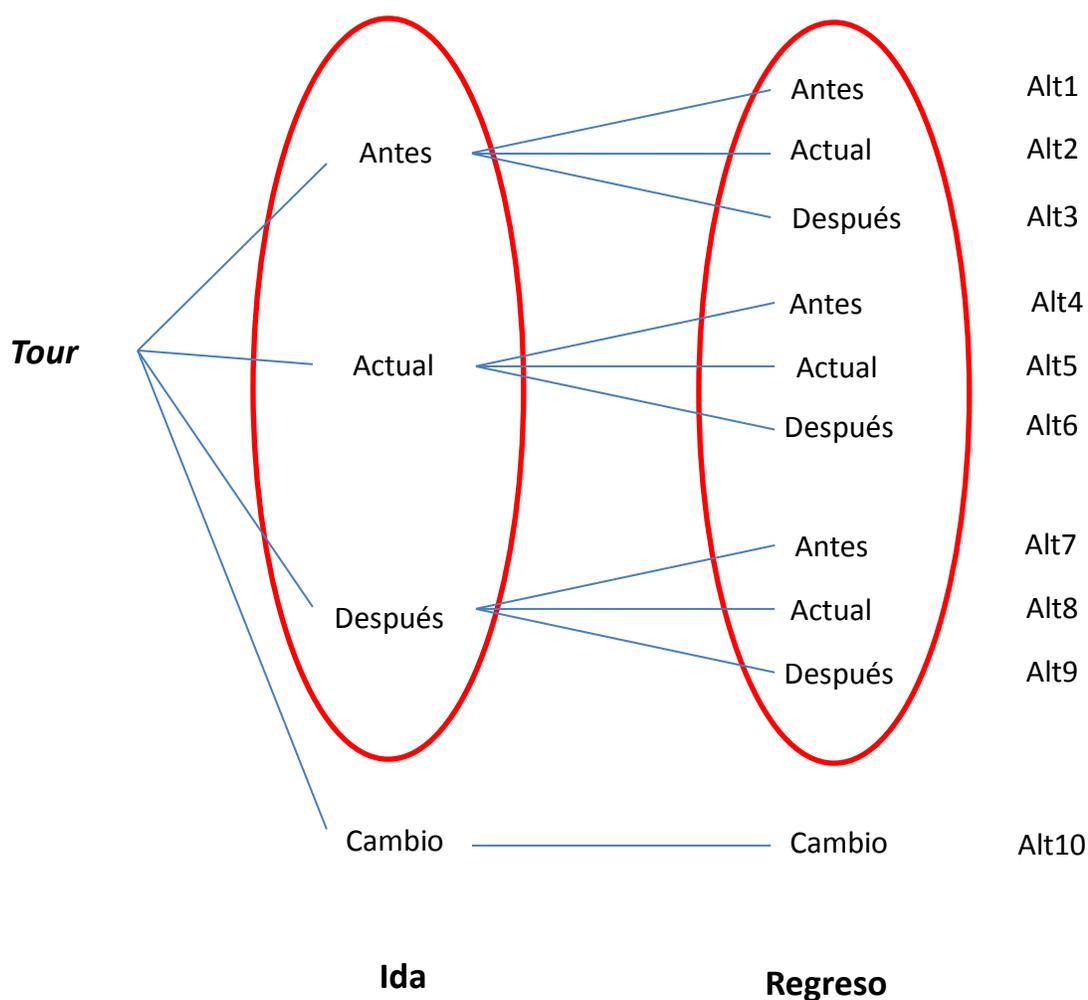


Figura 3-4: Ilustración de alternativas en el experimento PD enfocado a tours
Fuente: Elaboración propia

3.4.3.1 Experimentos de elección de la hora de inicio de viajes

A pesar de que el enfoque predominante para modelar la elección de horario de inicio de viaje son las encuestas de PD, no existe consenso respecto al diseño más adecuado a utilizar para recolectar datos para estimar el SM.

En estudios previos, las situaciones de elección se han generado con diferentes tipos de diseño, desde técnicas estándar como diseños ortogonales en diferencias (Börjesson, 2008) y diseños fraccionales factoriales (Bajwa *et al.*, 2008; Saleh y Farrell, 2005; Tseng y Verhoef, 2008), hasta diseños más complejos que combinan técnicas

ortogonales con diseños manuales para acomodar bastantes atributos con sus respectivos niveles (de Jong *et al.*, 2003; Ettema *et al.*, 2007; Hess *et al.*, 2007b).

Los diseños experimentales usados hasta el momento tienden a simplificar la complejidad de la tarea de elección. Dentro de los problemas principales que dificultan el desarrollo de experimentos de elección para SM se encuentran: (i) la dependencia de algunos niveles de ciertos atributos con los niveles de otros atributos del diseño; (ii) las variables dentro del modelo a estimar no corresponden exactamente a los mismos atributos presentados a los encuestados (Koster y Tseng, 2010); (iii) para garantizar el realismo del experimento, el diseño debe ser personalizado y basarse en las características específicas de los viajes de cada encuestado.

La dependencia entre atributos del diseño no sólo es la que se presenta cuando los niveles de un atributo en la alternativa j pueden generarse a partir de un cierto nivel del mismo atributo en la alternativa i . Esta dependencia ha sido estudiada anteriormente por Rose *et al.* (2008), y produce diseños que se conocen como *pivoteados* a partir de alternativas de referencia. La dependencia que se consideró en el diseño utilizado en esta tesis, es la que ocurre cuando un atributo dentro de la alternativa j depende del nivel de otro atributo de la misma alternativa j , que a su vez es también parte del diseño (Arellana *et al.*, 2012a; Arellana *et al.*, 2012c).

El tiempo de viaje, su variabilidad y el costo, son función de la hora en que se realiza el viaje; pero ésta, a su vez, es otro atributo que debe variar dentro del diseño. Si se asigna diferentes niveles de variación a ciertos atributos y se ignora la interdependencia entre ellos, pueden generarse situaciones irreales o incluso absurdas para los encuestados.

Para ilustrar el problema anterior, consideremos el caso de generar un experimento con sólo dos atributos, tiempo de viaje y hora de inicio del viaje. Supongamos que los niveles del primero son: 20, 30 y 40 min y que las horas de inicio serían las 7:00, 8:00 y 10:00 am. Además, se sabe que el periodo punta está entre las 7:45 y las 8:45 am. Entonces, al generar el diseño podrían presentarse situaciones de elección con alternativas que muestren que viajar a las 8:00 am tardaría 20 min, mientras que viajar a las 7:00 am demoraría 40 min. Lo anterior sería contra intuitivo y podría generar un problema si el

encuestado decide no tomar en serio la encuesta porque los niveles presentados no le parecen realistas.

Mediante la inclusión de restricciones, en un procedimiento de diseño en varias etapas, es posible controlar la interdependencia entre atributos de un experimento. Esto es de suma importancia para el realismo y el compromiso del participante en la encuesta, ya que debe impedirse a toda costa que el individuo encuentre situaciones de elección contra intuitivas o simplemente poco probables en la realidad.

Por otra parte, formular modelos de elección que utilicen variables obtenidas de forma indirecta a partir de los atributos presentados a los encuestados, caso común en un SM, es un problema desde el punto de vista de la generación del diseño, ya que la relación entre los atributos presentados en el experimento y las variables modeladas no es uno a uno, y en el contexto de un diseño eficiente es muy importante definir cuales atributos se debe intentar optimizar y cómo realizar la equivalencia entre ellos. Este problema había sido señalado anteriormente por Koster y Tseng (2010), pero no ha sido tratado en profundidad en la literatura de diseño de encuestas de PD.

En el contexto de un diseño eficiente (Rose y Bliemer, 2008), la información *a priori* con que se puede contar está relacionada con los términos que finalmente se estimarán en el modelo (por ejemplo, la demora horaria por llegar temprano o tarde) y esto permite inferir que éstos son los términos que se debiera optimizar. Intuitivamente, el diseño debería generarse controlando los atributos del modelo y luego utilizar las ecuaciones que permitan determinar los niveles a presentar al encuestado.

En un SM sencillo, sin variabilidad del tiempo de viaje y sólo dentro del contexto de un viaje, tal vez no sea un problema mayor optimizar estos niveles de los atributos porque las ecuaciones que los vinculan son sencillas; sin embargo, al incluir incertidumbre en el SM, los niveles de variación de los atributos del modelo pueden no encontrarse en la misma escala (unidades) que los niveles presentados a los encuestados. Por esta razón, en modelos complejos hay que tener cuidado en cuanto al nivel de atributos a optimizar, pues pueden generarse variaciones no deseadas en los niveles presentados a los encuestados (a veces muy grandes, o a veces dentro del rango de indiferencia de las personas).

El diseño utilizado en esta tesis representa una alternativa para tratar el problema de equivalencia entre atributos y evitar variaciones no deseadas en los niveles presentados a los encuestados. La idea es optimizar el diseño controlando inicialmente los niveles de los atributos a presentar al encuestado en lugar de definir los niveles de los atributos en el modelo. En este caso, los parámetros *a priori* de los atributos a presentar se obtienen de los parámetros *a priori* de los atributos del modelo usando las ecuaciones que permiten relacionarlos.

Por último, con el fin de aumentar el realismo de la encuesta y lograr que el individuo recuerde el contexto de la situación de elección que enfrenta usualmente, es deseable presentar una alternativa igual o muy similar a la elección reportada originalmente por el individuo. Dotar al encuestado de un contexto general dentro del ejercicio, le permite incorporar restricciones no necesariamente tenidas en cuenta por el modelador y que pueden influir en su comportamiento real. En conclusión, incluir una alternativa de referencia puede otorgarle mayor relevancia al ejercicio. Además, en la literatura se ha encontrado que las encuestas de PD que se generan a partir de una respuesta de PR previa, suelen presentar mayor eficiencia en el diseño (Caussade *et al.*, 2005; Rose *et al.*, 2008; Train y Wilson, 2009).

3.4.3.2 Descripción del procedimiento de diseño experimental

Para generar situaciones de elección en los dos tipos de experimentos de PD se utilizó el procedimiento esquematizado en la Figura 3-5.

a) Características preliminares del diseño

Se decidió realizar un diseño eficiente de los experimentos de elección basado en el modelo SM. En la segunda etapa se utilizó el software de diseño NGENE (www.choice-metrics.com) y la tercera etapa se programó usando macros en Visual Basic para Excel (dado que los paquetes de *software* generalmente disponibles no permiten incluir restricciones de dependencia entre atributos de la misma alternativa).

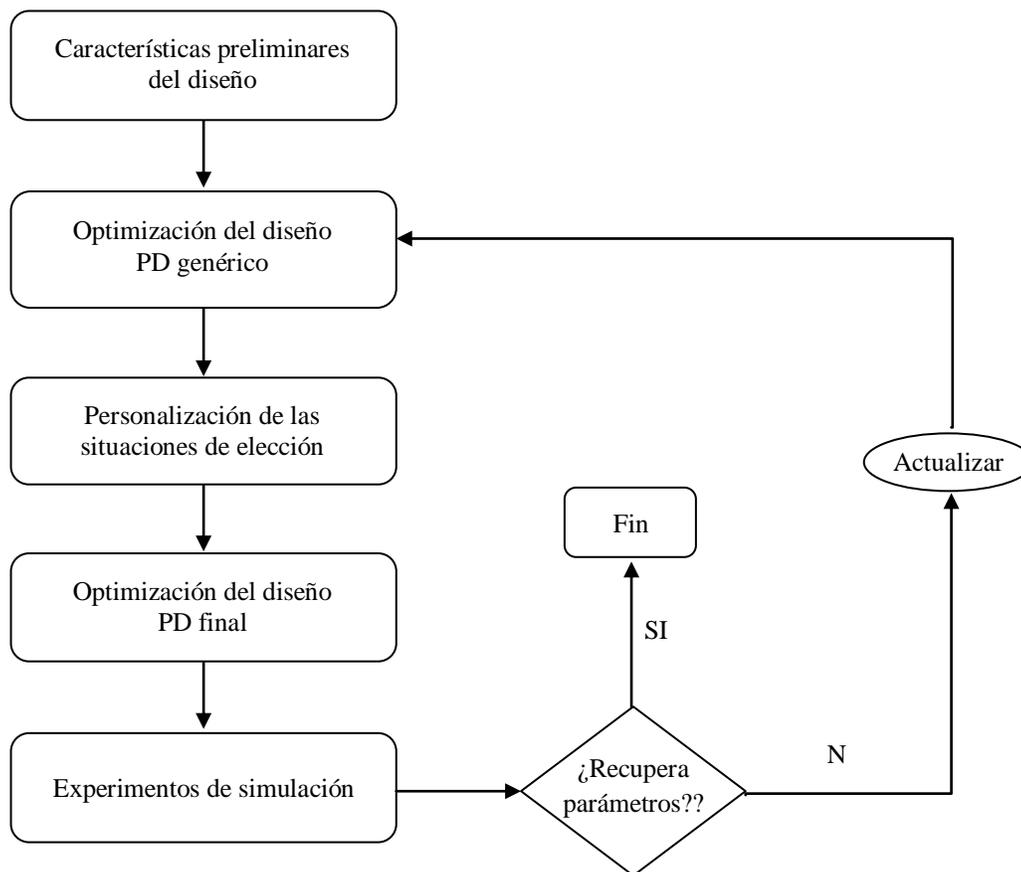


Figura 3-5: Esquema general del procedimiento de diseño utilizado
Fuente: Elaboración propia

Contexto del experimento y definición de atributos a presentar

Antes de presentar las situaciones de elección a los encuestados, se decidió entregar una descripción del experimento con el fin de crear un contexto común. Así, al principio de la encuesta de PD se presentaba la siguiente descripción contextual:

Debido al crecimiento de la población y del parque automotor en la Región Metropolitana, la congestión en el Metro, los buses, y la producida por los autos que circulan en la ciudad es cada vez mayor. De hecho, se pronostica que dentro de 10 años la duración promedio de los viajes será dos o tres veces mayor que la duración actual.

Por estos motivos, estamos realizando una investigación sobre las preferencias de usuarios del sistema de transporte santiaguino, con miras a evaluar diferentes

alternativas de gestión que permitan mitigar los efectos de la congestión esperada en la ciudad a futuro.

En esta encuesta, le presentaremos una serie de situaciones de elección (basadas en los viajes reportados por Ud. anteriormente), en que Ud. debe elegir la opción que le parezca más atractiva en cada caso, teniendo en cuenta que el gobierno ha decidido implementar un sistema de cobro a los automovilistas que circulen por el centro de la ciudad en horarios de congestión.

Además, se indicaba a las personas que tenían la opción de modificar sus horas de entrada y de salida de trabajo, siempre y cuando cumplieran con un número de horas de trabajo fijas y preestablecidas a la semana.

Los atributos considerados en el experimento asociado al viaje de ida hacia el trabajo, fueron: hora de inicio del viaje, tiempo de viaje, variabilidad del tiempo de viaje, comodidad del tiempo de viaje y costo. Es importante resaltar que se decidió presentar la variabilidad del tiempo de viaje como el mayor tiempo de viaje posible para ese viaje, durante algún día de la semana a esa hora, en lugar de otras presentaciones comunes (ver Hensher *et al.*, 2011), porque se consideró prioritario el compromiso entre atributos para elegir la hora de viaje y el modo.

En grupos focales, se encontró que las personas podían manejar formatos de presentación de la variabilidad del tiempo de viaje utilizados previamente en la literatura (ej: cinco tiempos de viaje alternativos). Sin embargo, cuando en el experimento se presentaba la variabilidad de esa forma, las personas tendían a fijarse mayormente en ese atributo dándole menos importancia a los restantes. Así, si bien en el contexto de un estudio sobre valoración de la incertidumbre del tiempo de viaje resultaría adecuado incluir un enfoque con cinco tiempos de viaje (y sus probabilidades asociadas), por ejemplo, en el contexto de un ejercicio de elección horaria se consideró mejor introducir la variabilidad mediante un atributo más simple que no distrajera al encuestado del foco del experimento.

Restricciones y relaciones de dependencia entre atributos

En el diseño de este experimento se consideró que los atributos costo y tiempo de viaje dependían de la hora en que se realizaban los viajes. Como interesaba mantener la

presentación de los atributos de la forma más simple, se decidió calcular los niveles del tiempo de viaje como un porcentaje promedio de la variación observada en las mediciones realizadas sobre la red de transporte de Santiago.

Adicionalmente, haciendo uso de NGENE, se incluyó restricciones para impedir la dominancia entre alternativas durante la segunda etapa del experimento.

Atributos que se optimizan primero

En este caso, en el diseño genérico se propuso optimizar primero a los cambios de los atributos hora de inicio del viaje, comodidad y variabilidad del tiempo de viaje.

Niveles y coeficientes a priori de los atributos

Para generar el diseño genérico de cambios relativos respecto a un punto de referencia, se decidió utilizar cinco niveles por atributo. Los valores de los coeficientes *a priori* para el costo, tiempo de viaje y comodidad, se obtuvieron a partir de estudios previos realizados en Chile (MIDEPLAN-SECTRA, 2005; Rizzi y Limonado, 2008; Yañez *et al.*, 2009). Los valores de los coeficientes *a priori* para las demoras horarias fueron obtenidos a partir de los valores de la confiabilidad del tiempo de viaje (β_{SDE}/β_{TT} y β_{SDL}/β_{TT}) reportados en estudios internacionales (Li *et al.*, 2010); por este motivo, se decidió incluir incertidumbre en los valores de todos los parámetros *a priori* (enfoque Bayesiano). Los niveles de los atributos y el rango de valores *a priori* utilizados, se muestran en la Tabla 3-3.

Dado que la mayoría de los valores de los coeficientes *a priori* reportados en la Tabla 3-3 se refieren a cambios en los atributos y no a los atributos en sí, los valores provenientes de la literatura se “adaptaron” multiplicando el valor *a priori* por valores medios de referencia dentro de la muestra.

Número de bloques y de situaciones de elección

Luego de realizar una prueba piloto, considerando la duración del experimento y comparando el número de situaciones de elección con la de otros estudios similares, se decidió que los encuestados no debieran enfrentar más de 13 situaciones de elección (cinco en el primer experimento y ocho en el segundo).

Tabla 3-3: Nivel de atributos y valores de parámetros *a priori* en el diseño genérico

Atributo	Nivel					Parámetro <i>a priori</i>	
	1	2	3	4	5	Min	Max
Cambio del tiempo de viaje*	1	1,05	1,1	1,15	1,2	-7	-0,7
Cambio hora de inicio del viaje: alternativa actual y cambio de modo	-10	-5	0	5	10	-	-
Cambio hora de inicio del viaje: alternativa viaja antes	-30	-45	-60	-75	-90	0,01	0,24
Cambio hora de inicio del viaje: alternativa viaja después	30	45	60	75	90	-0,36	-0,015
Variabilidad del tiempo de viaje transporte privado	0,1	0,15	0,2	0,25	0,3	-9	-0,3
Variabilidad del tiempo de viaje transporte público	0,15	0,2	0,25	0,3	0,35	-9	-0,3
Cambio del costo*	0,7	0,85	1	1,15	1,3	-0,3	-0,017
Comodidad	0,7	0,85	1	1,10	1,2	-1,5	-0,85

Fuente: Elaboración propia

Utilizando NGENE, se generó un diseño genérico, común para el primer experimento, a partir de 50 situaciones de elección y diez bloques. Las situaciones de elección del segundo experimento se obtuvieron a partir de un diseño genérico común con 40 situaciones de elección y cinco bloques.

b) Proceso de optimización para obtener el diseño PD genérico

Dado que cada encuestado iba a enfrentar diferentes situaciones de elección, el enfoque propuesto en esta etapa consistió en generar un diseño inicial donde los niveles de los atributos (en su mayoría) representaban cambios relativos (porcentuales), no absolutos, respecto a un punto de referencia. Este diseño se definió como genérico y puede ser común a todos los encuestados, si se desea; alternatively, se puede crear un diseño diferente si se cuenta con diversos estratos predefinidos e identificados dentro de la muestra.

*Notar que los atributos tiempo de viaje y costo están fijos entre alternativas para el diseño genérico, pues no se optimizan en la segunda etapa debido a su relación de dependencia respecto a la hora de inicio del viaje. Los niveles que se presentan en estos casos, son variaciones del nivel base que se incorporan con el fin de variar los escenarios entre situaciones de elección.

En esta etapa se optimizó los niveles de los atributos que condicionaban los niveles de otros atributos y también los niveles de aquellos atributos que no tenían relaciones de dependencia. Así, las actividades desarrolladas en esta etapa fueron:

- Definir el criterio de eficiencia para evaluar los diseños: se escogió el D_b -error;
- Definir el criterio de parada del procedimiento: este se fijó como 60 min sin encontrar un diseño superior al correr NGENE en un computador Core 2 Duo de 1,86 MHz;
- Seleccionar un diseño PD candidato aleatoriamente o mediante heurísticas: se utilizó el procedimiento SWAP que viene predeterminado en NGENE;
- Calcular probabilidades y matriz de covarianza asintótica (MAVC), dados los atributos de diseño y coeficientes *a priori*;
- Calcular criterios de eficiencia para evaluar el diseño
- Volver a escoger otro diseño PD candidato hasta que se cumpla el criterio de parada.

c) Personalización de las situaciones de elección

Esta etapa se realizó en Microsoft Excel usando los diseños genéricos de la etapa anterior y la información entregada por el encuestado en las dos primeras partes de la encuesta. Pasar de los niveles de cambio relativo por atributo (que se encuentran en el diseño genérico de la etapa anterior) a los niveles de atributos personalizados, de acuerdo a la situación de elección de cada encuestado, fue el objetivo de la etapa descrita a continuación.

El primer paso consistió en asignar los niveles de cada atributo, para cada situación de elección, a cada encuestado. Cada bloque se asignó aleatoriamente y se buscó repetir su uso un número similar de veces en la muestra total.

Niveles de atributos no optimizados en el diseño genérico y restricciones de dependencia

El cálculo del tiempo de viaje para alternativas que no implican cambios considerables de modo ni hora, se basó en el valor reportado por los encuestados en las dos primeras partes de la encuesta. El tiempo de viaje para la alternativa *cambio de modo* se basó en el valor obtenido al dividir la distancia que recorrería el encuestado si usará el

modo alternativo y el promedio de la velocidad medida en la red de transporte durante la recolección de datos, para la hora en que se iniciaba el viaje según la alternativa correspondiente. El cálculo del tiempo de viaje para las alternativas con cambios considerables de hora, fue optimizado dentro del diseño y los valores se pivotearon a partir del tiempo de viaje de la alternativa donde se usaba el mismo modo y se viajaba a una hora muy similar a la reportada. Los niveles de variación del tiempo de viaje para diferentes horas, fue determinado a partir de mediciones en GPS en campo que permitieron obtener los niveles de servicio sobre la red, a diferentes horas del día.

Para calcular los tiempos de viaje base y sus niveles de variación, se introdujo una restricción respecto a la hora de inicio del viaje. En el caso de la alternativa que llegaba al lugar de destino a una hora muy cercana a la reportada, la restricción fue diferente al caso de las otras dos alternativas (donde el individuo llega a su destino antes o después que la hora reportada). Para la alternativa *cambio de modo* y salir a la hora base reportada, se tuvo en cuenta que al aplicar el nivel de variación de la hora de inicio del viaje, el individuo podría terminar realizando su viaje en un periodo del día diferente al reportado (notar que el periodo puede cambiar porque a la hora del día reportada se le adiciona el nivel del atributo “cambio en la hora del día”, obtenido del diseño genérico para cada alternativa). Los periodos del día considerados para calcular el tiempo de viaje de usuarios de auto, eligiendo estas alternativas, fueron: horario punta (entre 7:30 y 9:00 hrs), fuera de punta (antes de las 6:30 hrs, y después de las 10 hrs), y periodo valle (entre 6:30 y 7:30 hrs, y 9:00-10:00 hrs).

El tiempo de viaje base, entonces, se afectó por un valor igual a uno si la persona no cambiaba de periodo del día, por un valor menor a uno si la alternativa implicaba que la persona viajaría en un periodo menos congestionado, y por un valor mayor a uno si se le ofrecía viajar en un periodo de mayor congestión al reportado. Finalmente, el valor que aparecía en la encuesta correspondía al tiempo base afectado por el factor de la hora del día y multiplicado por el nivel del escenario de cambio de tiempo de viaje en el diseño genérico. Para el cálculo del tiempo de viaje de las alternativas que implicaban cambios de hora considerables, se definieron cinco niveles de variación que dependían de la hora del día en que se realizarían los viajes. La restricción de hora evalúa en qué hora del día se

propone realizar el viaje, y toma los cinco niveles de variación correspondientes a viajar a esa hora.

Adicionalmente, se consideró diferentes niveles de variación del tiempo de viaje para estas alternativas de acuerdo a la duración del viaje realizado por el encuestado. Por esta razón, viajes con duración menor a 50 min, tuvieron diferentes niveles de variación del tiempo de viaje respecto a viajes de mayor duración. No es lo mismo proponer un cambio de 50% en un viaje de 10 min que un cambio de 50% en un viaje de 120 min. Un resumen de los cinco niveles de variación en los tiempos de viaje se muestra en la Tabla 3-4.

Tabla 3-4: Niveles de variación del tiempo de viaje usados en la segunda parte del diseño

Alt	Niveles	Duración < 50 min					Duración > 50 min				
		1	2	3	4	5	1	2	3	4	5
Antes	TT1	0,80	0,85	0,90	0,95	1,00	0,90	0,92	0,94	0,96	0,98
	TT2	0,70	0,75	0,80	0,85	0,90	0,80	0,85	0,90	0,95	1,00
	TT3	0,55	0,60	0,65	0,70	0,75	0,70	0,80	0,85	0,90	0,95
	TT4	1,02	1,04	1,06	1,08	1,10	1,02	1,04	1,06	1,08	1,10
	TT5	1,00	1,05	1,10	1,15	1,20	1,00	1,05	1,10	1,15	1,20
Después	TT1	0,80	0,85	0,90	0,95	1,00	0,90	0,92	0,94	0,96	0,98
	TT2	0,70	0,75	0,80	0,85	0,90	0,80	0,85	0,90	0,95	1,00
	TT3	0,55	0,6	0,65	0,70	0,75	0,70	0,80	0,85	0,90	0,95
	TT4	1,05	1,1	1,15	1,20	1,25	1,02	1,04	1,06	1,08	1,10
	TT5	1,25	1,3	1,35	1,40	1,45	1,00	1,05	1,10	1,15	1,20

Fuente: Elaboración propia

Así, para calcular el tiempo de viaje de estas alternativas se debió seleccionar un nivel en el procedimiento de optimización del diseño y luego multiplicarlo por el tiempo de viaje de la alternativa base (es decir, la alternativa de viajar en el mismo modo y a una hora similar a la reportada).

El cálculo del costo por viaje se hizo a partir del valor personalizado del atributo “hora de inicio del viaje” para cada alternativa. Dentro del ejercicio, los encuestadores indicaban a los participantes que si eran usuarios de transporte privado, el costo indicado se refería al valor adicional a pagar por concepto de tarificación vial. Por otra parte, si el participante era usuario de transporte público, este valor correspondía a la tarifa completa del viaje.

Para el atributo costo, no se realizaron variaciones de niveles por simplicidad y para evitar situaciones poco realistas. De acuerdo a un esquema de cobro por periodos del día, como el que actualmente rige en la ciudad para el sistema de transporte público, sería poco creíble encontrar situaciones de elección donde la persona tenga dos alternativas de viaje dentro de un mismo período del día con costo de viaje diferente. Este atributo pudo haber sido optimizado aplicando un procedimiento similar al del tiempo de viaje, pero adicionando una restricción de diferencia de costos para la selección de su nivel de variación. La optimización de este atributo no se hizo por considerar que ya se iba a obtener una variación suficiente del mismo, al definir un valor fijo a mostrar dependiendo de la hora de inicio de los viajes.

El procedimiento seguido para asignar los costos consistió en multiplicar la variación fija entre alternativas, incluida dentro del diseño genérico, por el costo correspondiente al periodo del día en que viajaría la persona según lo propuesto en la alternativa y lo indicado en la Tabla 3-5.

Tabla 3-5: Valor del costo por periodo para la segunda parte del diseño del experimento 1

Hora	-6:30	6:30- 7:00	7:00- 7:30	7:30- 8:00	8:00- 8:30	8:30- 9:00	9:00- 9:30	9:30- 10:00	10:00- 10:30	10:30-
T. Privado	500	800	1000	1200	1500	1500	1200	1000	800	500
T. Público	510	560	620	660	660	620	580	560	540	510

Fuente: Elaboración propia

Otras restricciones para aumentar el realismo del diseño

Luego de efectuar la personalización inicial, y en conjunto con la definición de los niveles de los atributos para el segundo proceso de optimización, se definieron tres restricciones adicionales para aumentar el realismo del diseño.

En el experimento 1, la primera restricción incluida se refiere al atributo “hora de inicio de los viajes” para la alternativa que implicaba salir antes de la hora reportada hacia el trabajo. Se consideró que no sería atractivo, para una persona que reportaba iniciar su viaje antes de las 7:30 hrs, presentarle una alternativa de viaje que implicara tener que salir antes de las 6:00 hrs. Por esta razón, a estas personas se les incluyó una corrección en la

hora de inicio del viaje (ecuación 39), de tal manera que los cambios a presentar, sólo para esta alternativa, fueran menores a los propuestos en la Tabla 3-3:

$$Hr\ inicio_B = Hr\ reportada + Cambio\ de\ hr_B \cdot \frac{(Hr\ inicio_A - Hr\ referencia)}{90} \quad (39)$$

donde el subíndice B se refiere a la alternativa de salir antes, el subíndice A a la alternativa de salir a una hora similar a la actual, $Cambio\ de\ hr$ se refiere al nivel del atributo dentro del diseño genérico y $Hr\ referencia$ corresponde a la hora mínima fijada para que el individuo inicie su viaje. En esta tesis, la hora de referencia utilizada fue las 6:10 am.

Las dos restricciones restantes corresponden a la especificación de umbrales de diferencia permitidos entre los atributos de diferentes alternativas (Ortúzar y Rodríguez, 2002), y fueron aplicadas al tiempo de viaje y a las horas de inicio de viaje entre diferentes alternativas en una misma situación de elección. Diferencias de tiempo u hora menores a cinco minutos tienden a ser consideradas irrelevantes por los individuos. Por esta razón, si se encontraba una diferencia menor a este umbral (respecto al atributo de la alternativa de referencia), el atributo de la alternativa no referencial se especificaba con una diferencia de al menos cinco minutos en valor absoluto.

d) Proceso de optimización para obtener el diseño PD final

Este proceso fue programado en Visual Basic para Excel y, en líneas generales, es muy similar al realizado en la segunda etapa, pero con dos diferencias. La primera se relaciona con los atributos que varían al optimizar el diseño, que son diferentes a los de la etapa anterior. La segunda tiene que ver con la matriz de covarianza utilizada. En esta parte se deriva y se optimiza una matriz de covarianza a partir del total de datos en la muestra, teniendo en cuenta los atributos del modelo a estimar obtenidos a partir de los atributos personalizados para cada encuestado. Es decir, aquí no se está optimizando un diseño común para todos los individuos, sino que se está optimizando un diseño que contiene tantas filas como situaciones de elección por encuestado multiplicada por el número de estos.

El criterio de comparación entre diseños en esta fase continuó siendo el el D_b -error medio, utilizando 150 realizaciones *Halton* (Ortúzar y Willumsen, 2011; Press *et al.*, 1992). Los parámetros *a priori* utilizados se muestran en la Tabla 3-6 y el criterio de parada fue fijado en 30 min si no se encontraba un mejor diseño.

Tabla 3-6: Valor de los parámetros *a priori* utilizados en la optimización final

Atributo	Tiempo de viaje	SDE	SDL	Costo	Comodidad
Max	-0,012	-0,0072	-0,0144	-0,00017	-0,00038
Min	-0,12	-0,24	-0,36	-0,003	-0,00666

Fuente: Elaboración propia

e) Experimentos de simulación

Tras obtener los diseños, estos se sometieron a un procedimiento de simulación donde se verificó si era posible recuperar el rango de parámetros *a priori* utilizados para optimizar los diseños generados para el proyecto. Cabe destacar que la simulación se realizó con la muestra total que participó en el diseño, variando solamente los coeficientes a recuperar y simulando las elecciones de cada encuestado. En el 100% de los casos, se pudo constatar que las muestras generadas permitían recuperar los parámetros de diseño. Si no se hubiera logrado estimar parámetros dentro del rango de los coeficientes *a priori* usados en el proceso de optimización y simulación de elecciones, se debería haber vuelto a repetir el procedimiento desde el paso (b).

3.4.3.3 Resultados generales para comprobar la bondad del diseño

La Tabla 3-7 muestra los resultados correspondientes a la estimación, utilizando BIOGEME (Bierlaire, 2003), de un modelo MNL usando sólo los datos para el experimento PD de viajes al trabajo. Al analizar sus valores se puede concluir que el procedimiento de diseño fue aparentemente exitoso, ya que todos los parámetros estimados se encuentran dentro del rango de incertidumbre utilizada en el procedimiento de diseño (reportados en la Tabla 3-6). Adicionalmente, todos los coeficientes presentan las magnitudes y signos esperados.

Tabla 3-7: Resultados estimación usando experimento PD para viajes al trabajo

Parámetro	Valor estimado	Test-t
ASC alternativa antes	-0,436	-3,02
ASC alternativa después	-0,813	-5,01
ASC alternativa cambio de modo	-1,182	-13,75
TVIA (tiempo de viaje)	-0,012	-2,36
ESDE (penalidad por llegar temprano)	-0,021	-8,04
ESDL (penalidad por llegar tarde)	-0,026	-8,56
COSW (costo dividido por tasa salarial)	-0,032	-8,17

Fuente: Elaboración propia

El valor del tiempo de viaje obtenido, al operar con estos coeficientes, es cercano a $\$24/\text{min}^1$ lo que está en línea con valoraciones del tiempo de viaje utilizadas normalmente en Chile.

3.4.4 Variables latentes

Con el objetivo de incorporar variables latentes en los modelos de elección, al finalizar la componente PD de la encuesta se recolectó el ingreso personal del encuestado y se agregó una sección con un cuestionario especialmente diseñado para capturar actitudes y percepciones de los individuos relacionadas con la elección de la hora de inicio de sus viajes. El cuestionario de VL también fue aplicado a aquellas personas que no fueron seleccionadas para responder la componente PD.

3.4.4.1 Cuestionario de variables latentes

En este cuestionario se solicitó a los encuestados evaluar ocho afirmaciones de acuerdo a su nivel de acuerdo o desacuerdo respecto a cada una de ellas. La escala de calificación fue de 1 a 7, donde 1 indicaba muy en desacuerdo y 7 muy de acuerdo. La Tabla 3-8 resume los resultados de esta sección, es decir, presenta los promedios y desviaciones estándar para las ocho afirmaciones.

Usando un análisis factorial exploratorio se identificaron dos factores a partir de las ocho afirmaciones contenidas en el cuestionario de VL.

¹ La tasa salarial promedio para los individuos encuestados era 62,91 \$/min.

Las primeras cuatro afirmaciones se agruparon en un factor que se denominó “actitud hacia la puntualidad”. Las afirmaciones contenidas en ese factor, tienen mayor calificación en promedio que las cuatro afirmaciones restantes, que se agruparon en otro factor identificado como “disposición hacia los cambios en las condiciones de viaje”. La anterior tendencia sugiere que los encuestados podrían estar más preocupados por ser puntuales en sus trabajos. Sin embargo, al observar las respuestas directas a los indicadores dentro del factor, se observa que a los encuestados les preocupa más la incertidumbre asociada a los tiempos de viaje que cumplir a cabalidad con sus horarios de entrada y salida del trabajo.

Tabla 3-8: Promedio de respuestas cuestionario de variables latentes (N = 405)

Afirmación	Media	σ
1. Me molesta no saber con exactitud cuánto durará mi viaje	6,7	0,7
2. Es muy importante para mi llegar a tiempo al trabajo	5,4	1,7
3. Es muy importante para mi salir del trabajo a una hora fija	5,7	2,2
4. Por lo general, me preocupo más por cumplir los horarios en la jornada de la mañana que en horas de la tarde	5,3	2,1
5. Estoy dispuesto a viajar antes o después que mi hora usual con tal de reducir mi tiempo de viaje	5,2	1,9
6. No me molestaría tener que pagar más con tal de viajar a mis horas usuales de viaje	4,4	2,1
7. Me agradaría variar la cantidad de trabajo durante diferentes días con tal de viajar en horas menos congestionadas	4,8	2,1
8. No me molestaría tener que pagar más para viajar más cómodo o durante menos tiempo	5,4	1,8

Fuente: Elaboración propia

3.4.4.2 Estructura del modelo de variables latentes

Basado en los resultados del análisis factorial exploratorio, se propuso un modelo de variables latentes (Bollen, 1989) siguiendo la estructura de un modelo MIMIC (*Múltiple Indicator Múltiple Cause*). Las VL en el modelo MIMIC son explicadas, a través de ecuaciones estructurales, por características socioeconómicas y de empleo de los individuos; al mismo tiempo, las VL explican los indicadores de percepción a través de las

ecuaciones de medición. La estructura del modelo MIMIC propuesto se muestra en la Figura 3-6.

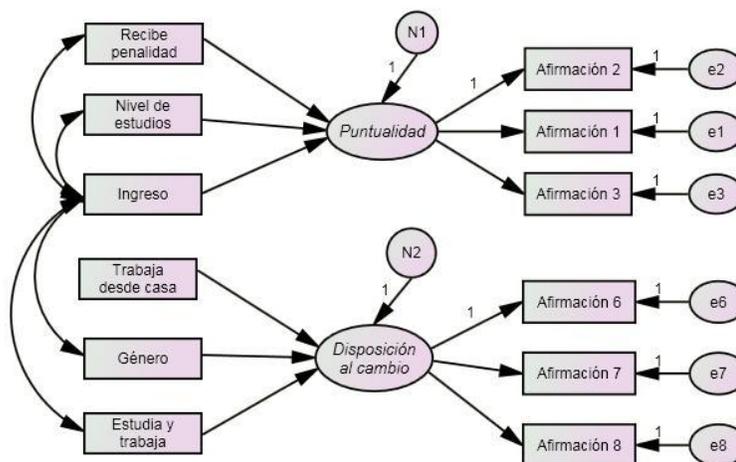


Figura 3-6: Estructura del modelo MIMIC propuesto
Fuente: Elaboración propia

Específicamente, las VL fueron observadas a través de seis indicadores. Las afirmaciones 1, 2 y 3 se asociaron a la variable puntualidad; mientras que las afirmaciones 6, 7 y 8 representan la disposición al cambio en las condiciones de viaje. Además, se hizo la hipótesis que variables como el nivel de educación, el ingreso y la posibilidad de ser penalizados por sus empleadores (si atrasan o adelantan sus viajes relacionados con el trabajo), podían influenciar la puntualidad de los encuestados. Por otra parte, se propuso que la disposición a aceptar cambios en las condiciones de los viajes variara de acuerdo al género, posibilidad de trabajar desde la casa, y la condición de ser estudiante y trabajador a la vez.

3.4.4.3 Resultados

Se utilizó una muestra de 405 individuos para estimar el modelo MIMIC descrito anteriormente mediante máxima verosimilitud. El modelo fue normalizado al fijar dos parámetros que forman parte de sus ecuaciones de medición (Raveau *et al.*, 2010). Algunos resultados seleccionados de la estimación del modelo se reportan en la Tabla 3-9.

Tabla 3-9: Resultados seleccionados del modelo MIMIC propuesto (N = 405)

Variable		Causa	Estimación	test t
Disposición al cambio	<---	Estudia y trabaja	0,396	1,973
Disposición al cambio	<---	Género	0,366	2,665
Disposición al cambio	<---	Trabaja desde casa	-0,607	-1,714
Puntualidad	<---	Ingreso	-0,002	-2,319
Puntualidad	<---	Nivel de estudios	-0,222	-1,771
Puntualidad	<---	Recibe penalidad	-0,075	-2,181
Afirmación 8	<---	Disposición al cambio	0,977	3,781
Afirmación 7	<---	Disposición al cambio	1,087	3,741
Afirmación 6	<---	Disposición al cambio	1,000	Fijo
Afirmación 2	<---	Puntualidad	1,000	Fijo
Afirmación 3	<---	Puntualidad	2,011	3,017
Afirmación 1	<---	Puntualidad	0,715	1,680
Parámetros estimados			28	
Grados de libertad			50	
χ^2			65,166	
p(χ^2)			0,073	
$\chi^2/df (< 2)$			1,303	
GFI (cerca de 1)			0,973	
RMSEA (<0,05)			0,027	

Fuente: Elaboración propia

Los índices de bondad de ajuste para el modelo son satisfactorios. A pesar de tener más de 400 observaciones en la muestra, el estadístico Chi-cuadrado no es significativo, lo que sugiere que el modelo no puede reproducir consistentemente la matriz de covarianza de la muestra. Los otros test de bondad de ajuste reportados en la tabla, superan los umbrales recomendados para indicar que el modelo es parsimonioso y adecuado estadísticamente.

Los parámetros de las ecuaciones estructurales relacionados con las variables *Trabaja desde casa*, y *Nivel de estudios*; y el parámetro de la ecuación de medición que relaciona la variable latente *puntualidad* con la *afirmación 1*, son significativos a un nivel de confianza del 90%. El resto de los parámetros son significativos a un 95% de confianza

Los resultados muestran que las mujeres de la muestra y las personas que estudian y trabajan a la vez, tienden a estar más dispuestas a aceptar cambios respecto a las condiciones de sus viajes. Personas con mayores ingresos en la muestra y/o con estudios de postgrado, tienden a estar menos preocupados de ser puntuales respecto a los horarios de

sus trabajos, ya que probablemente son personas con cargos más altos y mayor grado permitido de flexibilidad en los horarios de los viajes.

4 DESCRIPCION DE LA MUESTRA

En el presente capítulo se entrega una descripción estadística de la muestra recolectada mediante el instrumento de encuestas descrito en el capítulo 3. Un total de 405 personas, incluyendo 30 participantes de la encuesta piloto, completaron todas las etapas de la encuesta. El 76% de las personas (308 encuestados) completaron tres componentes de la encuesta (PR/PD/VL), mientras que el 24% restante sólo respondió las componentes PR y VL de la encuesta porque no fueron elegibles para responder la componente PD. Las estadísticas descriptivas acerca de las características socioeconómicas y del empleo que se reportan en este capítulo se calcularon usando la muestra de 405 personas.

4.1 Características socioeconómicas

La edad de los participantes oscila entre los 19 y 78 años, con una edad promedio de 40 años. El 47% de los participantes son mujeres, y sólo el 7% de las personas encuestadas cuenta con estudios de postgrado; 83% posee títulos de técnicos o profesionales, y el 10% restante posee a lo más estudios secundarios completos.

La composición familiar de los encuestados es diversa. La muestra contiene personas que viven solas (47 personas) y dos encuestados reportaron vivir con más de 10 personas en sus hogares. La mayoría de encuestados viven con su familia, y las familias tienen una composición promedio de 3,7 habitantes/hogar, sin considerar a aquellos que viven solos.

Respecto al acceso a auto, que puede ser una variable influyente en la elección del modo de transporte, el 76% de los encuestados posee licencia de conducción y cuenta con al menos un auto en el hogar. Sólo el 12% del total de encuestados reportó no tener posibilidades de acceder al auto en los viajes que inician en el hogar, pues ninguno de los miembros del hogar poseía licencia de conducir.

Adicionalmente, en las elecciones de modo y hora puede ser relevante el ingreso de los participantes. En la encuesta se solicitó a cada encuestado indicar su ingreso líquido mensual. El rango de ingresos varía entre \$ 230.000 y \$ 4.500.000, con un promedio de \$ 1.045.206 entre los encuestados. El 14% de los encuestados reportó ingresos inferiores a \$ 400.000; el 25% ingresos entre \$ 400.000 y \$ 800.000; el 27% ingresos entre \$ 800.000 y

\$ 1.200.000; el 28% ingresos entre \$ 1.200.000 y \$ 2.000.000; mientras que el 6% restante, ingresos superiores a \$ 2.000.000.

4.2 Características de los empleos

Las características de los empleos pueden tener importancia en la elección la hora de inicio de los viajes al trabajo. La mayoría de los trabajadores encuestados laboran jornadas completas de 45 horas a la semana. Menos del 1,5% de ellos reportó trabajar una cantidad inferior a 40 horas a la semana, y cerca del 18% reportó trabajar 50 horas a la semana o más.

Consultados sobre la importancia de la cantidad de horas trabajadas por día, sólo el 7% de los encuestados reportó que en su empleo no tenía importancia la cantidad de horas que trabaja al día. Otra característica interesante de la muestra, es que menos del 4% tiene la posibilidad de trabajar desde el hogar. Además, los viajes basados en el trabajo pueden considerarse rutinarios para la muestra, debido a que sólo una persona indicó que trabajaba menos de cuatro días en el lugar donde fue encuestado.

El 46% de los encuestados declaró no recibir penalizaciones en sus centros de empleo si llegaban tarde en la mañana al trabajo, o si se retiraban antes de la hora de finalización oficial del trabajo. Por otra parte, sólo el 34% del total de participantes reportó recibir incentivos de algún tipo (principalmente económicos) por trabajar horas extras.

Por otro lado, el 30% de los encuestados reportó no tener disponibilidad de adelantar su hora de llegada al trabajo, el 55% indicó que podría adelantarla hasta en 30 min, y el 15% restante en más de 30 min. El 37% de los encuestados no tiene la intención de retrasar su hora de llegada al trabajo; mientras que el 46% contemplaría hacerlo hasta en un período de 30 min más tarde. Sólo un 17% podría retrasar su viaje al trabajo por la mañana en más de 30 min.

Al parecer, las personas tienen mayor disposición a permanecer en el trabajo que a irse antes de que el trabajo finalice oficialmente. El 40% de los encuestados no tiene la disponibilidad de adelantar su viaje de salida del trabajo, mientras que sólo el 26% indica que no consideraría salir mas tarde del trabajo. Casi el 70% del total de la muestra se encuentra dispuesto a trabajar por al menos 30 min más que la hora oficial de salida del trabajo.

Por último, al tener información acerca de la cantidad de horas trabajadas a la semana por cada persona, es posible calcular la tasa salarial de los encuestados para fines de comparación. La tasa salarial de esta muestra oscila entre \$ 13/min y \$ 281/min, con un promedio de \$ 61,6/min.

4.3 Descripción de las elecciones de los encuestados

A pesar que los encuestados trabajan en lugares similares, la distribución de las horas de inicio de los viajes basados en el trabajo es amplia y los modos de transporte utilizados son diversos. Aunque el viaje al trabajo es de tipo rutinario, las personas no siempre realizan sus viajes a la misma hora ni utilizan el mismo modo de transporte. En contraposición, el viaje de salida del trabajo es algo más complejo pues su propósito no siempre es el mismo y por ende su calificación de rutinario no es clara. Precisamente, debido a las condiciones especiales de estos viajes es que resulta interesante su estudio. Es esperable que la variabilidad, en términos del modo de transporte y hora de inicio, del viaje luego del trabajo sea mayor que la del viaje hacia el trabajo. Además, las elecciones de hora y modo para el viaje de regreso del trabajo podrían estar relacionadas con las mismas elecciones en el viaje de ida. Con el fin de aumentar el tamaño de muestra en el análisis de modo y hora, se decidió aprovechar la información proveniente de la componente PR diligenciada por algunos encuestados que no contestaron las componentes PD y VL de la encuesta. Así, la descripción estadística de las elecciones de modo y hora se realizará utilizando una muestra de 478 personas.

Los propósitos de los viajes de salida del trabajo de los encuestados tienen alta variabilidad. La mayoría de personas en la muestra vuelve a su casa luego de trabajar (71%); sin embargo, otros encuestados realizan actividades que van desde compras, hasta recreación, deporte, y salud (ver Tabla 4-1). En promedio, la duración del viaje luego del trabajo es menos prolongada que la realizada en horas de la mañana. La duración promedio de los viajes al trabajo reportados por los encuestados es de 35 min, mientras que la duración de los viajes después del trabajo es pocos minutos menor, en promedio. Los viajes de salud o estudio que se realizan luego del trabajo, tienden a ser más cortos que el resto de viajes durante el día. Al parecer, las personas prefieren realizar sus chequeos de salud o estudios cerca del trabajo.

Tabla 4-1: Estadísticos de viajes después del trabajo para la muestra ampliada

Propósito	Duración del viaje (min)			No. Viajes	%
	Promedio	Mínimo	Máximo		
Volver a casa	37	1	82	339	71
Compras	28	1	61	38	8
Recreación/Deporte	24	3	67	26	5
Ver algo/a alguien	23	7	45	19	4
Buscar/dejar a alguien	28	5	56	16	3
Chequeos de salud	19	2	47	13	3
Trámites personales	24	7	44	8	2
Estudio	12	7	19	7	1
Buscar/dejar algo	26	12	41	5	1
Comer/tomar algo	31	10	55	4	1
Por trabajo	25	16	36	3	1
Total Viaje Salida Trabajo	34	1	82	478	

Fuente: Elaboración propia

4.3.1 Modo de transporte

Los participantes utilizaron en total 12 modos de transporte; sin embargo, los modos que implicaban combinaciones con transporte público se agruparon en uno solo dando un total de nueve modos de transporte para el análisis (Tabla 4.2).

Tabla 4-2: Disponibilidades y elecciones de modo viajes basados en el hogar

Modo	Elegido		Disponible		Elegido/Disponible
	Total	%	Total	%	
1 Auto chofer	198	21	612	64	32%
2 Metro	195	20	321	34	61%
3 Bus	186	19	910	95	20%
4 Bus-Metro	120	13	459	48	26%
5 Caminata	78	8	149	16	52%
6 Auto acompañante	77	8	564	59	14%
7 Combinación					
T. Público	48	5	600	63	8%
8 Taxi colectivo	28	3	367	38	8%
9 Taxi	26	3	813	85	3%

Fuente: Elaboración propia

La disponibilidad de diferentes modos es una variable clave en la estimación de modelos de partición modal. El modo más disponible para las personas encuestadas es el

Bus (disponible para cerca del 95% de los individuos). Debido a la duración promedio de los viajes y a que las personas provienen de diferentes partes de Santiago, comúnmente de lugares alejados de su trabajo, entre los modos menos disponibles se encuentran la caminata, el Metro y el Taxi-colectivo.

Es interesante destacar que más de la mitad de los viajes se realiza en transporte público, y que cerca del 75% de estos viajes corresponden a Metro o Bus sin combinación intermodal (modos puros), con lo cual sólo el 25% de los viajes en transporte público requiere al menos un transbordo entre modos diferentes. Los modos de transporte motorizados menos utilizados fueron, en conjunto, el Taxi y el Taxi colectivo.

En la última columna de la Tabla 4-2 se presenta una medida de la atractividad de cada modo. Este indicador ha sido calculado como la razón entre la cantidad de personas que escoge el modo y el total de personas que lo tiene disponible (valor entre cero y uno). La caminata es muy utilizada por quienes viven cerca de su lugar de trabajo, aunque son pocos los que residen suficientemente cerca como para considerar factible realizar sus viajes a pie.

Entre los modos motorizados el de mayor preferencia, en caso de estar disponible, es el Metro. De los modos restantes, el Auto como chofer presenta mayor elección y preferencia por parte de los encuestados. Por otra parte, el Bus y el Taxi son los modos menos preferidos a pesar de encontrarse usualmente disponibles.

4.3.2 Hora de inicio de los viajes

Los viajes basados en el trabajo reportados en la encuesta se extienden a lo largo del día como se muestra en la Figura 4-1. Los encuestados viajan mayoritariamente al trabajo entre 7:00 y 9:00 hrs, siendo el período que termina a las 8:15 hrs, aquel en que se inician más viajes durante la punta mañana. El patrón de horas de inicio de viajes reportado es consistente con las horas oficiales de entrada a los trabajos donde se realizaron las encuestas.

Usualmente, la hora de inicio del viaje hacia el trabajo depende de la hora en que se desea llegar al trabajo y la duración del viaje. La hora de llegada al trabajo está ligada a las potenciales disposiciones y acciones del empleador, y es el dato comúnmente utilizado para el cálculo de atributos relevantes en la modelación de horarios.

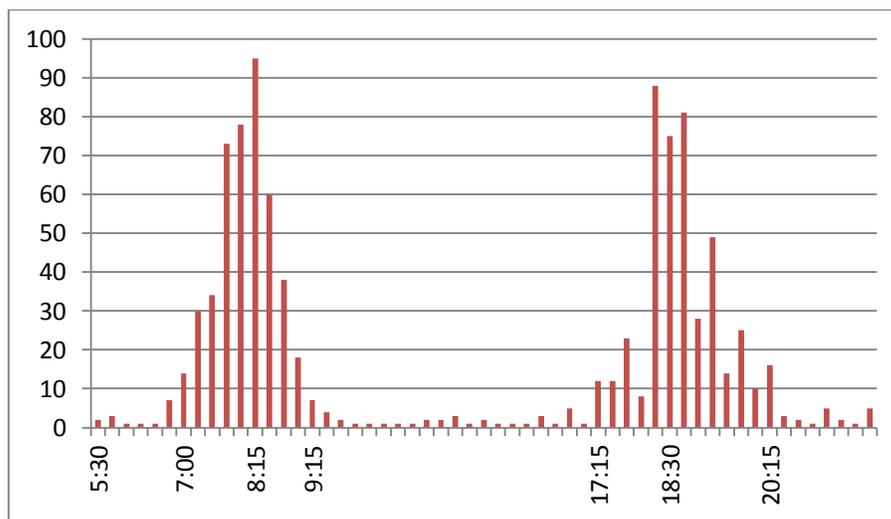


Figura 4-1: Horas de inicio de los viajes basados en el trabajo
Fuente: Elaboración propia

En particular, la hora de llegada al trabajo se usa para obtener las demoras horarias o penalidades por llegar tarde o temprano respecto a un horario preferido de llegada. La distribución de esta variable sigue un patrón similar al que se muestra en la Figura 4-1 para la punta mañana, con la diferencia que la llegada de los viajes se concentra mayormente entre 8:30 y 9:00 hrs.

Los viajes de salida del trabajo siguen una distribución similar a los que tienen lugar en punta mañana. La diferencia es que la hora de salida desde el origen se considera más relevante que la hora de llegada al destino. Los viajes de salida del trabajo, por la tarde, se realizan principalmente entre 18:15 y 19:30 hrs.

5 MODELOS DE ELECCIÓN DE LA HORA DE INICIO DE VIAJE CON DATOS DE PD

Los principales desarrollos respecto a la modelación de la elección de la hora de inicio de viaje, se han realizado usando datos de PD (ver sección 2.1.2). Los datos de PD ofrecen la oportunidad de tener múltiples observaciones por individuo a un costo menor que el de los datos de PR; además permiten incluir en el experimento de elección atributos difíciles de medir/controlar en un contexto de PR y evaluar políticas que en la actualidad no se encuentran implementadas.

En el presente capítulo se discute la modelación de la elección de horario usando los dos experimentos de PD presentados en la sección 3.4.3. Además de la formulación del modelo y la inclusión de variables relevantes, se discuten algunos aspectos metodológicos relacionados con la incorporación de la heterogeneidad en los gustos y el tratamiento del efecto panel en su aplicación a modelos de elección de la hora de inicio de viaje.

5.1 Logit Multinomial (MNL)

Con el fin de aprovechar la configuración de la componente de PD, se formuló inicialmente un MNL que permitiera capturar diferencias de escala entre los dos experimentos de elección aplicados. La conveniencia de modelar conjuntamente los dos experimentos de elección no se presenta en esta tesis, pero fue analizada mediante un test de razón de verosimilitud, donde se comparó el modelo conjunto (permitiendo diferencias de escala entre los dos experimentos) versus los modelos correspondientes a los experimentos 1 y 2, pero estimados por separado. El resultado del test arrojó que la hipótesis nula que el modelo conjunto es preferible sobre los dos modelos por separado, no puede ser rechazada a un nivel de confianza del 95% (Arellana *et al.*, 2012a).

Para modelar la elección de horario en el contexto del experimento enfocado en los viajes al trabajo, se planteó la siguiente formulación genérica:

$$V_i = ASC_i + \beta_{TT} TT_i + \beta_{Time_diff} Time_diff_i + \beta_C cost_i + \beta_{SDE} SDE_i + \beta_{SDL} SDL_i \quad (40)$$

donde el subíndice i se refiere a las cuatro alternativas del experimento basado en viajes, ASC_i representa las constantes asociadas a cada alternativa. Dado que existen cuatro opciones en el experimento para viajes al trabajo, se decidió fijar en cero la constante de la alternativa de viajar a la hora actual y estimar las tres restantes.

TT_i indica el tiempo de viaje del individuo y $Cost_i$ el costo del viaje al iniciar su viaje en el periodo i . SDE_i y SDL_i se refieren a las demoras horarias si el individuo llega temprano o tarde, definidas previamente en las ecuaciones (26) a (28).

El parámetro β_{Time_diff} representa la incertidumbre asociada a la variabilidad del tiempo de viaje, capturada a través de la diferencia entre el tiempo “usual” de viaje presentado y aquel que sucede “una vez a la semana”, según la ecuación (41):

$$Time_{diff} = \frac{(Tviaje_{una\ vez\ a\ la\ semana} - Tviaje_{usual})}{Tviaje_{una\ vez\ a\ la\ semana}} \quad (41)$$

Para el experimento basado en *tours*, la formulación incluyó una penalidad relacionada con la duración de las actividades (de Jong *et al.*, 2003; Hess *et al.*, 2007b) que relaciona explícitamente los viajes de ida hacia y salida desde el trabajo. La formulación del experimento basado en *tours* se muestra en la ecuación (42):

$$V_i = ASC_i^{ida} + ASC_i^{regreso} + \beta_{TT} TT_i + \beta_{Time_diff} Time_diff_i + \beta_C cost_i + \beta_{SDE}^{ida} SDE_i^{ida} + \beta_{SDL}^{ida} SDL_i^{ida} + \beta_{PTD} PTD_i + \beta_{PTI} PTI_i \quad (42)$$

en que PTD_i y PTI_i , se refieren a las penalidades por la duración del trabajo y fueron definidas en las ecuaciones (34) a (36). Tal como en la formulación del experimento enfocado en viajes al trabajo, las constantes para cada alternativa se representan por los parámetros ASC_i . En este experimento se definió cuatro ASC para representar las alternativas en el viaje de ida y cuatro para el viaje de regreso. Luego de fijar las ASC asociadas a la alternativa de viajar a la hora reportada en la ida y al regreso, la combinación de las tres constantes restantes, para la ida y regreso, en todas las alternativas, permitió capturar las preferencias generales de salir antes, después o cambiar de modo en cada etapa del *tour*. Cabe aclarar que los atributos TT , $Time_diff$, y $Cost$ se calculan para ambas etapas del *tour*, mientras que los atributos SDE y SDL sólo se computan para el viaje de ida. Los

valores de SDE y SDL para el viaje de retorno no deben incluirse en la formulación, porque el efecto de un potencial cambio de horario en el viaje de retorno, es en parte capturado por los parámetros PTD_i y PTI_i . Nuevamente, los subíndices i representan una de las 10 alternativas disponibles.

Usando las formulaciones presentadas sobre una muestra de 308 encuestados (que corresponde al total de personas que completaron las encuestas PD) se estimó un modelo MNL mediante un código desarrollado especialmente para este fin en OX (Doornik, 2007). Para considerar los dos experimentos de elección en un sólo modelo MNL, se fijó el parámetro de escala asociado a las alternativas del experimento basado en *tours*, y se estimó el parámetro de escala del otro experimento.

Los resultados del modelo se presentan en la Tabla 5-1. Todos los coeficientes estimados tienen signo esperado y son significativos a 95% de confianza. En general, si los encuestados debieran cambiar hora de inicio de viaje, preferirían adelantar su viaje en lugar de atrasarlo, *ceteris paribus*. Las constantes estimadas sugieren que los encuestados prefieren salir antes de sus hogares en lugar de cambiar de modo de transporte, pero prefieren cambiar de modo antes que salir tarde hacia sus trabajos. Adicionalmente, las constantes para la componente del *tour* sugieren que las personas preferirían realizar cambios de modo u hora en los viajes después del trabajo antes que cambiar el modo o la hora de inicio del viaje hacia el trabajo.

Al evaluar la relación de los coeficientes estimados respecto al tiempo de viaje (*razones de compromiso*), se encuentra que la mayoría está en línea con lo reportado anteriormente en la literatura (específicamente las de SDE y SDL son consistentes con los reportados por Li *et al.*, (2010) y de Jong *et al.*, (2003). El valor de compromiso de la penalidad por llegar tarde (SDL) sugiere que la valoración de un minuto adicional, cuando se está retrasado respecto a la hora preferida, es aproximadamente el doble del valor de un minuto adicional en el tiempo de viaje. Pero las razones de compromiso respecto a las penalidades asociadas a cambiar el largo de la jornada laboral (coeficientes PTD y PTI) son mayores a los reportados por Hess *et al.* (2007b) y de Jong *et al.* (2003); sugiriendo que si las personas de la muestra santiaguina pudieran cambiar libremente la cantidad de horas trabajadas en el día manteniendo una cantidad de horas fijas trabajadas a la semana,

preferirían quedarse más tiempo en el trabajo durante determinados días para tener la posibilidad de reducir el largo de la jornada laboral en un día específico.

Tabla 5-1: Resultados de estimación para modelo MNL conjunto

	Modelo conjunto	
	Estimación	test t robusto
ASC		
Cambio de modo	-0,786	-6,95
Salir antes – experimento viaje	-0,387	-3,02
Salir antes viaje ida – experimento <i>tour</i>	-0,657	-3,99
Salir antes viaje regreso – experimento <i>tour</i>	-0,516	-4,01
Salir después – experimento viaje	-0,614	-4,33
Salir después viaje ida – experimento <i>tour</i>	-1,622	-8,49
Salir después viaje regreso – experimento <i>tour</i>	-1,108	-7,41
Log-Cost (β_C)	-0,668	-5,99
Demora horaria por llegar temprano (β_{SDE})	-0,013	-5,30
Demora horaria por llegar tarde (β_{SDL})	-0,016	-5,36
Tiempo de viaje (β_{TT})	-0,008	-2,05
Incertidumbre en tiempos de viaje (β_{time_diff})	-0,767	-3,75
Penalidad por disminuir horario de trabajo (β_{PTD})	-0,013	-5,67
Penalidad por aumentar horario de trabajo (β_{PTI})	-0,006	-3,89
Factor de escala – sobre el experimento de viajes	1,517	11,65
Razones de compromiso		
SDE / TT		1,625
SDL / TT		2,000
Time_diff / TT		1,324
PTD / TT		1,625
PTI / TT		0,750
Número de parámetros		17
Número de observaciones		4068
Número de individuos		308
Log-verosimilitud en convergencia		-6.225,65

Fuente: Elaboración propia

El resultado anterior está en línea con lo reportado en el análisis estadístico de la muestra (sección 4.2), donde se indicaba que sus miembros parecían tener mayor disposición a permanecer en la oficina que irse antes de que su horario de trabajo finalizara oficialmente.

Por último, al operar con el parámetro β_{time_diff} para obtener la valoración de reducir un minuto adicional asociado a la incertidumbre del tiempo de viaje y dividirlo por el parámetro del tiempo de viaje, se obtiene una medida similar a la confiabilidad del tiempo

de viaje. Este valor sugiere que la reducción de un minuto asociado a la incertidumbre del tiempo de viaje es valorado aproximadamente 30% más que la reducción de un minuto en el tiempo usual de viaje, indicando (correctamente) que las personas prefieren sistemas de transporte con menor variabilidad del tiempo de viaje.

5.2 Incorporando heterogeneidad en los gustos

A pesar de que los resultados obtenidos en el modelo estimado en la sección anterior pueden considerarse satisfactorios, se decidió buscar un modelo con mayor poder explicativo. El modelo anterior supone que los gustos son homogéneos en la población. Con el fin de refinar el modelo, se incorporó la posibilidad de capturar la heterogeneidad en los gustos a través de dos enfoques: (i) la inclusión de variaciones sistemáticas en los gustos (VSG) y (ii) la posibilidad que los parámetros tengan una distribución aleatoria en la población. Al incorporar parámetros aleatorios para capturar heterogeneidad, la estructura del modelo deja de ser un MNL para convertirse en un ML, que se estima usualmente por máxima verosimilitud simulada.

La formulación y estimación de modelos que intentan capturar la heterogeneidad en los gustos también se realizó en OX. Las variaciones sistemáticas de gustos que resultaron significativas, correspondieron a interacciones entre los atributos de los experimentos PD y características socioeconómicas o de los empleos de las personas. Los resultados de los modelos propuestos en esta sección, se muestran en la Tabla 5-2.

En términos de bondad de ajuste, el modelo ML tiene una log-verosimilitud en convergencia muy superior a la del MNL. De hecho al aplicar un test de razón de verosimilitud, se puede demostrar estadísticamente que el ML es superior al MNL (LR=3201,04; $p = 0,000$). Por otro lado, todos los coeficientes estimados tienen el signo esperado y son significativos a un 95% de confianza, a excepción del nivel bajo del atributo *confort* en el modelo MNL y algunas interacciones en el modelo ML. Las variaciones sistemáticas de gustos pierden significancia cuando son incluidas junto a los parámetros aleatorios en el modelo ML. Lo anterior podría deberse a que los parámetros aleatorios capturan mejor la heterogeneidad en la muestra. Las constantes de los modelos estimados siguen una tendencia similar a la observada en el modelo MNL de la sección anterior.

Tabla 5-2: Resultados de estimación para modelos conjuntos

	MNL		ML	
	Estimación	test t	Estimación	test t
ASC				
Cambio de modo	-0,823	-15,70	-2,549	-16,73
Salir antes – experimento viaje	-0,407	-4,71	-0,960	-4,77
Salir antes viaje ida – experimento <i>tour</i>	-0,6717	-7,45	-0,600	-4,06
Salir antes viaje regreso – experimento <i>tour</i>	-0,4882	-5,84	-0,799	-6,75
Salir después – experimento viaje	-0,625	-5,80	-1,444	-5,99
Salir después viaje ida – experimento <i>tour</i>	-1,580	-13,60	-1,722	-9,17
Salir después viaje regreso – experimento <i>tour</i>	-1,114	-11,93	-1,549	-11,65
Atributos				
Log-Cost (β_C)	-0,689	-11,15	-1,554	-8,89
Tiempo de viaje (β_{TT})	-0,011	-5,40	-0,048	-7,93
Demora horaria por llegar temprano (β_{SDE})	-0,012	-8,14	-0,056	-12,57
Demora horaria por llegar tarde (β_{SDL})	-0,013	-7,60	-0,082	-12,08
Incertidumbre en tiempos de viaje (β_{time_diff})	-0,787	-4,44	-1,669	-5,70
Penalidad por disminuir horario de trabajo (β_{PTD})	-0,012	-7,93	-0,035	-8,10
Penalidad por aumentar horario de trabajo (β_{PTI})	-0,007	-5,52	-0,013	-5,83
Confort – De pie (β_{COM3})	-0,038	-0,69	-0,256	-2,31
Confort – De pie, vehículo lleno y a veces hay que esperar por el siguiente vehículo (β_{COM5})	-0,113	-2,09	-0,585	-5,30
Interacciones				
TT – Penalidad por llegar tarde en la mañana	0,014	4,04	0,065	6,72
SDE – Nivel de educación	-0,013	-3,66	-0,018	-1,73
SDL – Nivel de educación	-0,023	-3,78	-0,037	-2,11
SDE – Incentivo por llegar temprano al trabajo	-0,008	-3,64	-0,021	-3,93
SDL – Penalidad por salir antes del trabajo	-0,005	-2,82	0,009	1,40
PTD – Posibilidad de trabajar desde casa	0,024	5,23	0,020	2,39
PTI – Posibilidad de trabajar desde casa	0,012	3,97	0,007	1,12
Desviación estándar				
Log-Cost (S_C)	-	-	2,916	13,93
Tiempo de viaje (S_{TT})	-	-	0,120	18,17
Demora horaria por llegar temprano (S_{SDE})	-	-	-0,049	-13,75
Demora horaria por llegar tarde (S_{SDL})	-	-	-0,063	-13,62
Penalidad por disminuir horario de trabajo (S_{PTD})	-	-	0,040	9,73
Penalidad por aumentar horario de trabajo (S_{PTI})	-	-	-0,020	-13,83
Factor de escala				
Factor de escala – sobre el experimento de viajes	1,464	14,12	0,672	16,71
Resumen de la estimación				
Número de parámetros	25		31	
Número de observaciones	4068		4068	
Número de individuos	308		308	
Número de realizaciones	1		100	
Log-verosimilitud en convergencia	-6167,87		-4657,35	

Fuente: Elaboración propia

Sin embargo, el modelo ML entrega un valor mucho más negativo para la constante del cambio de modo, sugiriendo que las personas prefieren cambiar de hora antes que

cambiar de modo en condiciones *ceteris paribus*. Esto es consistente con lo reportado por de Jong *et al.* (2003) y Hess *et al.* (2007b). Finalmente, un resultado interesante del modelo ML tiene que ver con el valor del factor de escala, que es menor a uno, en contraste con el valor (mayor que uno) obtenido para el MNL. Esto sugiere que la consideración de parámetros aleatorios dentro de las funciones de utilidad disminuye la magnitud del término de error Gumbel, por lo que se obtienen parámetros estimados de mayor magnitud (ver la discusión en Sillano y Ortúzar, 2005).

A simple vista, los modelos estimados hasta el momento parecerían satisfactorios; sin embargo, la magnitud de las desviaciones estándar en el ML y de algunos parámetros de interacción en el MNL, sugieren que algunas personas podrían valorar en forma contra intuitiva (respecto a la teoría económica) algunos atributos del modelo. Por ejemplo, según el modelo MNL las personas que son penalizadas por llegar tarde al trabajo tendrían una valoración positiva del tiempo de viaje, lo que no parece económicamente correcto.

Teniendo en cuenta los resultados anteriores, la tarea se centró en identificar individuos que, potencialmente al menos, pudieran tener comportamiento contra-intuitivo. Para esto, se decidió primero explorar las respuestas individuales utilizando reglas determinísticas que permitieran detectar encuestados con comportamiento potencialmente lexicográfico. Luego, a partir de los parámetros poblacionales estimados con el modelo ML, se infirió la distribución de los parámetros individuales siguiendo el procedimiento Bayesiano descrito por Train (2003).

Inferir la distribución de los parámetros individuales puede resultar útil en la identificación de individuos que no tienen en cuenta cierto atributo en un experimento de elección o aquellos que realizan sus elecciones siguiendo otro paradigma de elección (por ejemplo, eliminación por aspectos).

Estudios previos que utilizan técnicas Bayesianas han mostrado que la cantidad de individuos con valoración contra-intuitiva de algún atributo es mucho menor que el valor esperado sugerido por los parámetros poblacionales. De hecho, estos estudios reportan que si bien algunos individuos efectivamente tienen parámetros individuales con signo contra-intuitivo, éstos no son estadísticamente distintos de cero (Godoy y Ortúzar, 2009; Sillano y Ortúzar, 2005).

Luego de analizar cuidadosamente las respuestas de los encuestados con el fin de identificar individuos potencialmente lexicográficos, y aplicar la inferencia Bayesiana para identificar otros individuos que pudieran exhibir algún problema de esta naturaleza, se llegó a una muestra final de 262 encuestados con comportamiento claramente compensatorio. Cabe destacar que la inferencia Bayesiana de las distribuciones de los parámetros individuales a partir de las estimaciones del ML, se realizó mediante un código escrito por el autor en OX utilizando 20.000 realizaciones de *Modified Latin Hypercube Sampling* - MLHS (Hess *et al.*, 2006).

Usando la información reportada por estos 262 individuos se estimó nuevamente modelos MNL con variaciones sistemáticas de gustos y modelos ML con parámetros aleatorios; sus resultados se presentan en la Tabla 5-3. En términos de bondad de ajuste, el ML sigue siendo muy superior a los modelos con estructura de error más sencilla; por ejemplo, la log-verosimilitud en convergencia para el modelo MNL, estimado con esta nueva muestra, es de -5.307,44 mientras que las correspondientes a los modelos MNL con variaciones sistemáticas de gustos y ML, son de -5.162,62 y -4.055,85, respectivamente.

La diferencia de log-verosimilitudes existente entre el modelo MNL con interacciones y el modelo ML, sugiere que la heterogeneidad aleatoria o algún efecto que el MNL no captura (como el efecto panel), deben ser importantes en la bondad de ajuste.

Por otro lado, los parámetros estimados están en línea con las estimaciones realizadas previamente. Con la nueva muestra, todos los parámetros del modelo MNL presentan el signo adecuado y a pesar que las desviaciones estándar de los parámetros poblacionales estimados en el ML vuelven a sugerir la existencia de individuos con valoración inconsistente de ciertos atributos, este no es el caso pues el valor esperado de sus parámetros individuales no es significativamente distinto de cero. Esto se encuentra en línea con lo reportado por Sillano y Ortúzar (2005).

5.3 Efecto panel

Dado que los modelos MNL anteriores no consideran que la muestra posee múltiples respuestas por cada individuo (lo que se conoce como “efecto panel”), se incluyó una componente de error para intentar capturar la correlación asociada a las múltiples respuestas de cada individuo.

Tabla 5-3: Resultados de estimación con datos de la muestra definitiva

	MNL		ML	
	Estimación	test t	Estimación	test t
ASC				
Cambio de modo	-1,1239	-8,5	-3,0413	-9,0
Salir antes – experimento viaje	-0,4685	-3,0	-1,0791	-3,6
Salir antes viaje ida – experimento <i>tour</i>	-0,5580	-3,3	-0,6658	-2,9
Salir antes viaje regreso – experimento <i>tour</i>	-0,5007	-3,7	-0,7153	-4,1
Salir después – experimento viaje	-0,8845	-4,8	-1,6617	-4,3
Salir después viaje ida – experimento <i>tour</i>	-1,6119	-7,8	-1,6748	-5,7
Salir después viaje regreso – experimento <i>tour</i>	-1,1107	-7,1	-1,8237	-7,8
Atributos				
Log-Cost (β_C)	-1,5800	-8,2	-2,4160	-8,0
Tiempo de viaje (β_{TT})	-0,0240	-4,6	-0,0341	-4,1
Demora horaria por llegar temprano (β_{SDE})	-0,0097	-2,4	-0,0581	-5,7
Demora horaria por llegar tarde (β_{SDL})	-0,0184	-1,8	-0,0838	-6,9
Incertidumbre en tiempos de viaje (β_{time_diff})	-2,3849	-3,9	-2,9020	-4,8
Penalidad por disminuir horario de trabajo (β_{PTD})	-0,0270	-2,3	-0,0373	-6,6
Penalidad por aumentar horario de trabajo (β_{PTI})	-0,0127	-3,0	-0,0111	-3,8
Confort – De pie (β_{COM3})	-0,1354	-1,7	-0,4042	-2,8
Confort – De pie, vehículo lleno y a veces hay que esperar por el siguiente vehículo (β_{COM5})	-0,2515	-2,9	-0,8921	-5,3
Interacciones				
TT – Penalidad por llegar tarde en la mañana	0,0152	1,6	-	-
SDE – Nivel de educación	-0,0128	-1,8	-	-
SDL – Nivel de educación	-0,0236	-2,1	-	-
SDE – Estudia y trabaja	0,002	5,4	-	-
SDL – Edad	0,0003	2,9	-	-
PTD – Posibilidad de trabajar desde casa	0,0151	4,3	-	-
PTI – Posibilidad de trabajar desde casa	0,0098	2,8	-	-
Desviación estándar				
Log-Cost (σ_C)	-	-	-2,9701	-7,4
Tiempo de viaje (σ_{TT})	-	-	-0,1086	-10,4
Demora horaria por llegar temprano (σ_{SDE})	-	-	0,0513	10,6
Demora horaria por llegar tarde (σ_{SDL})	-	-	0,0764	9,2
Penalidad por disminuir horario de trabajo (σ_{PTD})	-	-	0,0355	7,8
Penalidad por aumentar horario de trabajo (σ_{PTI})	-	-	-0,0259	-9,8
Factor de escala				
Factor de escala – sobre el experimento de viajes	1,2486	12,0	0,6232	11,5
Resumen de la estimación				
Número de parámetros	25		23	
Número de observaciones	3.464		3.464	
Número de individuos	262		262	
Número de realizaciones MLHS	200		200	
Log-verosimilitud en convergencia	-5.162,62		-4.055,48	

Fuente: Elaboración propia

En el caso de los modelos ML no se incluyen términos de error adicionales porque al considerar parámetros aleatorios, también se captura el efecto panel. Considerar parámetros aleatorios (es decir, heterogeneidad aleatoria en los gustos) implica que cada

individuo tiene asociado un parámetro específico que se mantiene constante entre todas sus observaciones y alternativas, pero puede diferir del parámetro correspondiente a otro individuo. Cabe destacar que recientemente Hess y Rose (2009) propusieron una estructura de modelación que permite incluso capturar la variación de gustos entre observaciones de un mismo individuo.

5.3.1 Enfoque tradicional para incorporar el efecto panel

El método tradicional para incorporar la correlación tipo panel en la práctica, consiste en adicionar una componente de error a $(J-1)$ alternativas, siendo J el número total de alternativas disponibles, a fin de que no existan problemas con su identificación (Walker, 2001). Desafortunadamente, esta forma tradicional de incorporar el efecto panel puede conducir a resultados sesgados, ya que es necesario elegir arbitrariamente una alternativa de referencia que no contenga la componente de error, y esto modifica sensiblemente la estructura de error implícita en el modelo; en efecto, al adicionar una componente de error común a $J-1$ alternativas, implícitamente se genera un efecto de correlación entre ellas.

En conclusión, la metodología tradicional adopta implícitamente una estructura tipo logit jerárquico, debido a la correlación entre alternativas, pero esta no es exactamente la estructura que se desea modelar. Así, incluir una componente de error en $J-1$ alternativas no es recomendable.

5.3.2 Enfoques alternativos para incorporar el efecto panel

Con el fin de evitar inducir correlación entre alternativas debido a la adición de una componente de error en $J-1$ de ellas, Yañez *et al.* (2011) propusieron una solución práctica para capturar el efecto panel. La diferencia con el enfoque tradicional radica en la forma como se definen las $J-1$ alternativas en que se incorpora el término de error de panel. Yañez *et al.* (2011) plantean que el problema de correlación sistemática del enfoque tradicional puede evitarse, parcialmente, si se varía aleatoriamente la alternativa de referencia donde no se coloca la componente de error. Incluso, sugieren que la variación de la alternativa de referencia se realice a nivel de observaciones y no a nivel individual.

El problema del procedimiento anterior radica precisamente en su carácter estocástico, pues al variar aleatoriamente la alternativa en que se coloca la componente de error, se desconoce la estructura de error real que se está estimando. Para capturar adecuadamente el efecto panel, la componente de error debe incluirse en todas las alternativas a fin de evitar una estructura indeseada de correlación entre ellas. A su vez, para evitar problemas de identificación, se propone aprovechar la técnica de máxima verosimilitud simulada y extraer realizaciones aleatorias o cuasi-aleatorias diferentes para cada alternativa. Si se extrae diferentes realizaciones por alternativa, provenientes de una misma función de distribución estandarizada con media cero, y se multiplica dicha realización - para cada alternativa (que seguramente será diferente) - por el parámetro común a estimar (usualmente la desviación estándar de la distribución), se soluciona el problema de identificabilidad. Así, si se supone que la componente de error para capturar efecto panel sigue una distribución Normal con media cero y una desviación estándar a estimar, la distribución supuesta para el término de error sería equivalente a tener una componente Normal estándar multiplicada por la desviación estándar a estimar. Para evitar problemas de identificabilidad, cuando se realice la estimación por máxima verosimilitud simulada se deberá extraer diferentes realizaciones aleatorias (o cuasi-aleatorias) por alternativa desde la distribución Normal estándar.

El efecto panel fue incluido en los modelos a continuación siguiendo esta metodología. Para su implementación, se modificó el código en OX utilizado para estimar los modelos anteriores. Cabe destacar, que la metodología también puede implementarse en paquetes de estimación comerciales como BIOGEME. Un resumen de la bondad de ajuste de los modelos incluyendo efecto panel se presenta en la Tabla 5-4; la muestra con que se estimaron estos modelos fue la misma (262 individuos, 3.464 observaciones) que se había utilizado anteriormente.

Por último, cabe destacar que los parámetros de los coeficientes correspondientes a estos modelos se encuentran en línea con las estimaciones anteriores y no se presentan en el texto de esta tesis por motivos de espacio y simplicidad (ver Anexo C).

Tabla 5-4: Bondad de ajuste para modelos que incluyen efecto panel

Modelo	MNL	MNL + panel	MNL + panel + vsg	ML
Número de parámetros	17	18	26	23
Realizaciones MLHS	1	200	200	200
Log-verosimilitud	-5.307,44	-4.123,43	-4.095,48	-4.055,85

vsg: variaciones sistemáticas de gustos.

Fuente: Elaboración propia

5.3.3 Efecto panel heterogéneo

El efecto panel tradicionalmente se introduce para capturar la correlación generada por la presencia de múltiples respuestas de un mismo individuo dentro de una muestra. El concepto que se desarrolla en esta sección, está relacionado con la forma más adecuada de considerar ese efecto de correlación.

Históricamente, la correlación generada por las múltiples respuestas de un individuo se ha considerado constante para todos los individuos o grupos de individuos de la muestra. En cambio, en esta tesis se hace la hipótesis que la correlación debida a la multiplicidad de respuestas de un individuo no sea igual para todos los individuos de la muestra.

La variación en la especificación del efecto panel, se denominó “efecto panel heterogéneo”, debido a que considera que la correlación observada entre las múltiples respuestas de un individuo puede determinarse como función de las características propias del individuo o de un grupo en particular. Así, este efecto panel heterogéneo se basa en el concepto que diversos grupos de personas pueden ser más o menos consistentes en sus respuestas y, por tanto, la correlación observada en sus respuestas puede diferir.

En particular, en esta tesis se propone una especificación del efecto panel basada en la hipótesis que la varianza estimada para el término de error, incorporada en las funciones de utilidad, debiera ser función de las características socioeconómicas de cada individuo, como la edad y nivel de estudios (43):

$$\sigma_{panel} = K + \delta_{edad} \cdot edad + \delta_{estudios\ de\ postgrado} \cdot D_{postgrado} \quad (43)$$

donde K , δ_{edad} y $\delta_{estudios\ de\ postgrado}$ son parámetros a estimar, y $D_{postgrado}$ es una variable *dummy* que indica si la persona posee estudios de postgrado; finalmente, la variable *edad* indica la cantidad de años del participante.

Los resultados obtenidos al estimar un modelo con la especificación de efecto panel heterogéneo propuesto, fue comparada con la del mismo modelo incluyendo el efecto panel de la forma tradicional (vale decir, como constante entre los individuos). Un resumen de los resultados más relevantes de este ejercicio se muestra en la Tabla 5-5. Una tabla con los parámetros estimados durante este ejercicio, los cuales se encuentran en línea con las estimaciones anteriores, se pueden encontrar en el Anexo C.

Un test de razón de verosimilitud aplicado al resultado de ambos modelos, indica que la inclusión del efecto panel heterogéneo mejora significativamente la bondad de ajuste del modelo (LR=6,72; $\chi^2_{(0,05)} = 5,99$). Los resultados del modelo que incluye el efecto panel heterogéneo, sugieren que las respuestas de las personas con mayor edad tienden a estar más correlacionadas. Por el contrario, las respuestas de los encuestados que cuentan con estudios de postgrado tienden a estar menos correlacionadas; esto permitiría presumir que estas personas, quizás, realizaron la encuesta con mayor prisa.

Tabla 5-5: Comparación de resultados estimación efecto panel heterogéneo y panel fijo

Parámetro	Modelo		Panel fijo		Panel heterogéneo	
	Estimación	test t	Estimación	test t	Estimación	test t
K (constante)	2,035	22,71	1,770	5,60	0,008	1,61
δ_{edad}	-	-	-0,503	-2,59		
$\delta_{estudios\ de\ postgrado}$						
Número de parámetros	35		37			
Realizaciones MLHS	200		200			
Log-verosimilitud	-4.049,45		-4.046,09			

Fuente: Elaboración propia

Los resultados obtenidos al aplicar el denominado efecto panel heterogéneo son interesantes; no obstante su aplicación en otro contexto sería útil para evaluar la real potencialidad de la especificación. Adicionalmente, la inclusión de otras variables socioeconómicas y atributos relacionadas con el compromiso de la persona respecto a la encuesta (por ejemplo, tiempo entre respuestas), podrían ser de utilidad para mejorar la especificación de los modelos.

6 MODELOS HIBRIDOS DE ELECCIÓN DE LA HORA DE INICIO DE VIAJE

6.1 Formulación de modelos híbridos de elección discreta

En los últimos años, los modelos de elección discreta han evolucionado hacia la incorporación explícita de variables latentes (Walker, 2001). Las variables latentes (VL), que suelen estar fuertemente asociadas a actitudes y percepciones, son factores que pueden incidir fuertemente en el proceso de elección, pero tienen la característica de no ser directamente cuantificables dada su intangibilidad o subjetividad. Los modelos que incorporan VL se conocen comúnmente como “modelos híbridos” y su aplicación en diversas áreas del transporte ha ganado gran popularidad en los últimos años (Ben-Akiva *et al.*, 2002; Bolduc *et al.*, 2008); sin embargo, este no ha sido el caso en estudios relacionados con la hora de inicio de viaje (Arellana *et al.*, 2012d).

Los modelos híbridos están compuestos por dos sub-modelos principales: un modelo de elección discreta y un modelo de variables latentes (Figura 6-1).

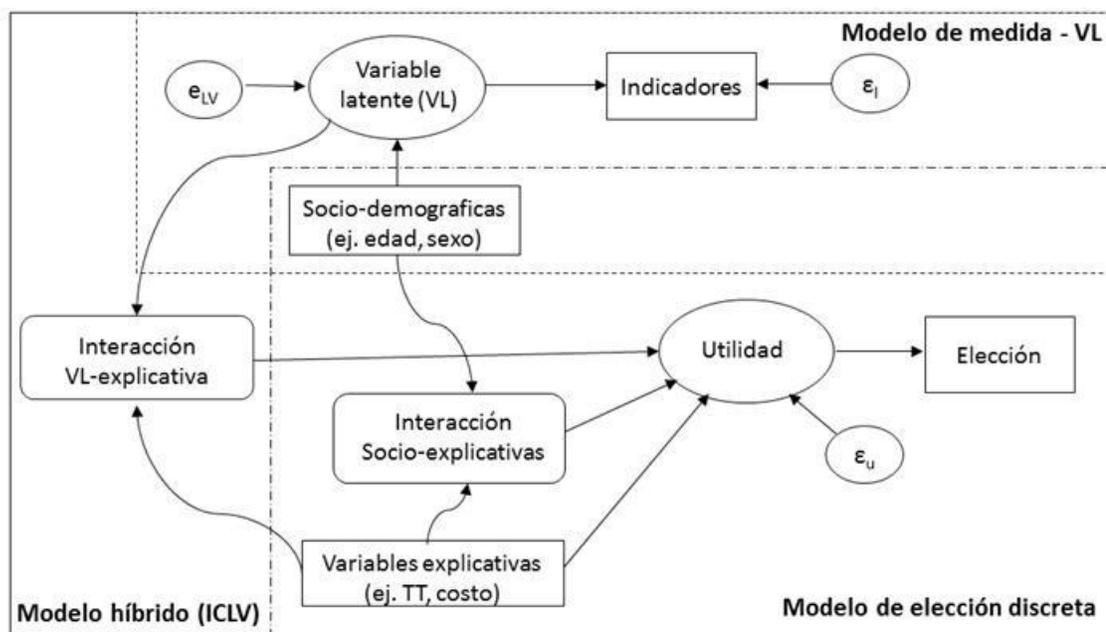


Figura 6-1: Estructura de un modelo híbrido
Fuente: Elaboración propia

El modelo de VL, a su vez, es comúnmente expresado como una estructura MIMIC (Bollen, 1989) con dos componentes: (i) una serie de *ecuaciones estructurales* (44) que establecen que la variable latente η es función de características sociodemográficas s ; y una serie de *ecuaciones de medición* (45) que establecen que la variable latente η puede reflejar una serie de indicadores y . Los indicadores pueden ser continuos, binarios o variables categóricas.

$$\eta_{ilq} = \sum_r \alpha_{ilr} \cdot s_{irq} + v_{ilq} \quad (44)$$

$$y_{ipq} = \sum_l \gamma_{ilp} \cdot \eta_{ilq} + \zeta_{ipq} \quad (45)$$

donde el índice i representa una alternativa, l una variable latente, q un individuo, r una variable explicativa y p un indicador. Por otro lado, α_{ilr} y γ_{ilp} son parámetros a estimar, y v_{ilq} y ζ_{ipq} son componentes de error que tienen media cero y desviación estándar a determinar. Como las variables latentes (η_{ilq}) no son conocidas, el proceso de estimación debe considerar ambas ecuaciones en forma conjunta (Bollen, 1989).

El modelo de elección discreta dentro de un modelo híbrido sólo se diferencia de un modelo de elección discreta tradicional por la presencia de las variables latentes η en la especificación de la utilidad representativa de las alternativas. Las VL pueden ser incluidas directamente en las funciones de utilidad o a través de interacciones con otras variables explicativas del modelo X (k indexa las variables explicativas X). La especificación de la utilidad representativa de las alternativas en un modelo híbrido, puede escribirse como (46), donde θ_{ik} , β_{ilq} , τ_{ilkq} son parámetros a estimar:

$$V_{iq} = \sum_k \theta_{ikw} \cdot X_{ikq} + \sum_l \beta_{ilq} \cdot \eta_{ilq} + \sum_{l,k} \tau_{ilkq} \cdot \eta_{ilq} \cdot X_{ikq} \quad (46)$$

6.2 Estimación de modelos híbridos

Los modelos híbridos pueden ser estimados usando máxima verosimilitud o técnicas Bayesianas (Alvarez-Daziano y Bolduc, 2011). La estimación por máxima verosimilitud es la más popular y puede realizarse en forma secuencial o simultánea.

En el método secuencial, primero se debe estimar el modelo MIMIC y su resultado es luego incorporado en el modelo de elección discreta, como si las VL fueran unas variables explicativas adicionales. Por otra parte, el método simultáneo estima de forma conjunta los dos modelos y en teoría obtiene estimaciones consistentes y eficientes (Ben-Akiva *et al.*, 2002).

En esta tesis, los modelos híbridos se estimaron siguiendo el enfoque simultáneo de máxima verosimilitud, codificado especialmente por el autor en el software OX (Doornik, 2007). Para estimar el modelo en forma simultánea, un requisito fundamental es plantear la función de verosimilitud conjunta del modelo de elección y del modelo de VL.

El modelo de VL propuesto en esta tesis siguió la estructura de un modelo MIMIC, con la particularidad que sus ecuaciones de medición se formularon y estimaron siguiendo la estructura de un modelo logit ordinal. Utilizar la estructura de un modelo logit ordinal para la componente de medición del modelo de variables latentes (Daly *et al.*, 2012) constituye un enfoque más sofisticado y realista que el enfoque lineal tradicional, porque permite reconocer la ordinalidad de las respuestas relacionadas con los indicadores.

Siguiendo la literatura de modelos ordinales (Greene y Hensher, 2010), acá se plantea que cada resultado discreto j reportado en el indicador y , asociado al individuo q , es el resultado de cierto mecanismo de censura aplicado sobre la suma de la VL y una componente de error continua, como se expresa en (47) y (48):

$$y_{pq} = 0 \text{ si } umbral_{p0} < y_{pq}^* \leq umbral_{p1}; y_{pq} = 1 \text{ si } umbral_{p1} < y_{pq}^* \leq umbral_{p2};$$

$$\dots \quad ; y_{pq} = j \text{ si } umbral_{pj} < y_{pq}^* \leq umbral_{p(j+1)} \quad (47)$$

$$y_{pq}^* = d_p \eta_q + \zeta_{pq} \quad (48)$$

Puesto que en la aplicación de este capítulo, la variable latente es a nivel individual, se eliminó el subíndice i de la ecuación anterior; además, como solo se tendrá una variable latente, también se eliminó el subíndice l a fin de simplificar la notación. Si se desea definir una estructura logit ordinal, ζ_{pq} asume una función de distribución logística, independiente de d_p , donde el parámetro d_p representa el impacto de la VL en la explicación del indicador y_p .

El modelo logit ordinal estima la probabilidad que la VL se encuentre entre ciertos umbrales crecientes a estimar, y produzca cierto resultado j para el indicador y . Esta probabilidad puede ser escrita como (49) y (50):

$$Pr\{y_{pq} = j | \eta_q\} = Pr(\zeta_{pq} \leq umbral_{pj} - d_p \eta_q) - Pr(\zeta_{pq} \leq umbral_{p(j-1)} - d_p \eta_q) \quad (49)$$

$$Pr\{y_{pq} = j | \eta_q\} = \frac{\exp(umbral_{pj} - d_p \eta_q)}{1 + \exp(umbral_{pj} - d_p \eta_q)} - \frac{\exp(umbral_{p(j-1)} - d_p \eta_q)}{1 + \exp(umbral_{p(j-1)} - d_p \eta_q)} \quad (50)$$

Para que el modelo sea identificable, los parámetros d_p y las varianzas de los términos de error ζ_{pq} se fijaron en uno, y los límites de los umbrales se fijaron entre $-\infty$ y $+\infty$ (es decir, $umbral_{q0} = -\infty$, y $umbral_{q(j+1)} = \infty$). Teniendo en cuenta lo anterior, la función de verosimilitud para la componente de medición puede expresarse como (51):

$$L = \prod_q \prod_p \left[\sum_j \left(\frac{\exp(umbral_{pj} - d_p \eta_q)}{1 + \exp(umbral_{pj} - d_p \eta_q)} - \frac{\exp(umbral_{p(j-1)} - d_p \eta_q)}{1 + \exp(umbral_{p(j-1)} - d_p \eta_q)} \right) \cdot D_{pqj} \right] \quad (51)$$

donde $D_{pqj} = 1$ cuando $y_{pq} = j$, y cabe destacar que al reemplazar (34) en (41), se puede obtener la expresión de la función de verosimilitud para el modelo de VL.

Por otra parte, la función de verosimilitud de la componente de elección discreta, cuando se incorpora la variable latente en la especificación de la utilidad sistemática y se tiene en cuenta que existen t observaciones por individuo, puede escribirse como (52) en el caso de una estructura MNL, y como (53) en el caso de una estructura ML (β_q representa los parámetros que se consideran aleatorios):

$$L_{MNL} = \prod_q \left\{ \int_{\eta} \left[\prod_t \left(\frac{\exp(V_{iqt})}{\sum_{i'} \exp(V_{i'qt})} \right) \right] \cdot dF(\eta) \right\} \quad (52)$$

$$L_{ML} = \prod_q \left\{ \int_{\beta} \int_{\theta} \int_{\eta} \left[\prod_t \left(\frac{\exp(V_{iqt})}{\sum_{i'} \exp(V_{i'qt})} \right) \right] \cdot dF(\eta) \cdot dF(\theta) \cdot dF(\beta) \right\} \quad (53)$$

En ambas ecuaciones debe integrarse sobre el dominio de la variable latente y en la segunda ecuación también debe integrarse sobre el dominio de variación de β_q . Dado que el propósito del modelo híbrido es modelar simultáneamente las elecciones de los individuos y las respuestas a los indicadores, se debe maximizar una función de verosimilitud conjunta ((54), para una estructura MNL y (55), para una estructura ML):

$$L_{ICLV-MNL} = \prod_q \left\{ \int_{\eta} \left[\prod_t \left(\frac{\exp(V_{iqt})}{\sum_{i'} \exp(V_{i'qt})} \right) \right] \cdot \left[\prod_p \left(\sum_j \left(\frac{\exp(\text{umbral}_{pj} - d_p \eta_q)}{1 + \exp(\text{umbral}_{pj} - d_p \eta_q)} - \frac{\exp(\text{umbral}_{p(j-1)} - d_p \eta_q)}{1 + \exp(\text{umbral}_{p(j-1)} - d_p \eta_q)} \right) \cdot D_{pqj} \right) \right] \cdot dF(\eta) \right\} \quad (54)$$

$$L_{ICLV-ML} = \prod_q \left\{ \int_{\beta} \int_{\theta} \int_{\eta} \left[\prod_t \left(\frac{\exp(V_{iqt})}{\sum_{i'} \exp(V_{i'qt})} \right) \right] \cdot \left[\prod_p \left(\sum_j \left(\frac{\exp(\text{umbral}_{pj} - d_p \eta_q)}{1 + \exp(\text{umbral}_{pj} - d_p \eta_q)} - \frac{\exp(\text{umbral}_{p(j-1)} - d_p \eta_q)}{1 + \exp(\text{umbral}_{p(j-1)} - d_p \eta_q)} \right) \cdot D_{pqj} \right) \right] \cdot dF(\eta) \cdot dF(\theta) \cdot dF(\beta) \right\} \quad (55)$$

6.3 Modelos híbridos de elección de hora de inicio de viaje usando datos PD y variables latentes

Como se indicó en la sección 2.2, no es claro qué factores no observables podrían influenciar el proceso de elección de la hora de inicio de viaje. No obstante, teniendo en cuenta otras aplicaciones de modelos híbridos en transporte, es esperable que incluir VL relacionadas con este proceso de elección represente una mejora en el poder explicativo de los modelos bajo estudio.

En particular, la motivación de esta sección es mostrar los resultados de una aplicación² del enfoque de modelación híbrido - descrito en las dos secciones anteriores - en el contexto de elegir la hora de inicio de viaje.

El autor no conoce aplicaciones previas de VL en este contexto. Específicamente, se presentará un modelo que define la *flexibilidad de los horarios de la jornada laboral* como un concepto latente influenciado por variables socioeconómicas y del empleo del individuo, y observado a través de una serie de indicadores relacionados con la cantidad de minutos disponibles por los individuos para adelantar o atrasar sus viajes.

La flexibilidad en los horarios de trabajo se ha considerado un atributo relevante desde los primeros estudios de elección de la hora de inicio de viaje. (Abkowitz, 1981; Hendrickson y Plank, 1984). Usualmente, la flexibilidad en los horarios ha sido incorporada como una variable binaria que indica si la persona tiene o no un horario de trabajo flexible (Senbil y Kitamura, 2004); o como una variable de estrato para segmentar la muestra, estimando modelos diferentes para usuarios con o sin flexibilidad (Börjesson, 2008; Hess *et al.*, 2007b).

El objetivo del modelo que se presenta en esta sección, es evaluar el efecto de la flexibilidad en las elecciones de la hora de inicio de viaje, considerándola como un

² Las variables latentes *puntualidad* y *disposición a los cambios*, encontradas en el modelo MIMIC preliminar, discutido brevemente en la sección 3.4.4, no son parte del modelo presentado en esta sección. Su inclusión en el contexto de elección horaria no será discutido en esta tesis por restricciones de tiempo y espacio, y su incorporación será objeto de trabajo futuro.

concepto latente que puede tener una percepción diferente por parte de los individuos. Para cumplir el objetivo planteado, se planteó una estructura de modelación híbrida.

La *flexibilidad en los horarios de trabajo* se definió como un concepto latente relacionado con las características del empleo de cada encuestado (56). Específicamente, se hizo la hipótesis que la flexibilidad podría estar determinada por atributos como la posibilidad de trabajar desde la casa (tr_encasa), y por la posibilidad de que el individuo sea penalizado si llega tarde al trabajo en la mañana o si se retira más temprano que la hora oficial de salida en la tarde. Para esto se definieron dos variables diferentes a partir de las posibles penalizaciones que podrían recibir los encuestados. Si una persona es penalizada si llega tarde en la mañana y también si sale temprano por la tarde ($Penal_AM_y_PM$), se esperaría que tuviera menos flexibilidad en su horario de trabajo. Si esta persona es penalizada sólo en alguno de los dos casos ($Penal_AM_o_PM$), entonces es esperable que tenga un grado de flexibilidad mayor que en el caso anterior:

$$Flexibilidad = g_{tr_encasa} \cdot tr_encasa + g_{penAMyPM} \cdot Penal_AM_y_PM + g_{penAMoPM} \cdot Penal_AM_o_PM + v \quad (56)$$

donde g_{tr_encasa} , $g_{penAMyPM}$, y $g_{penAMoPM}$ son parámetros a estimar, y se asume que v sigue una distribución Normal estándar para hacer identificable al modelo.

Se usaron cuatro variables como indicadoras del grado de flexibilidad de las personas; éstas contenían información acerca de la cantidad máxima de minutos que un encuestado tiene disponible para adelantar o atrasar sus viajes basados en el trabajo. Específicamente se utilizó: (i) la cantidad de minutos disponibles para adelantar su viaje al trabajo; (ii) la cantidad de minutos disponibles para retrasar su viaje al trabajo; (iii) la cantidad de minutos disponibles para adelantar su viaje de salida del trabajo, y (iv) la cantidad de minutos disponibles para retrasar su viaje de salida del trabajo. Los individuos debían indicar cuál de las cinco categorías era la que mejor representaba su disponibilidad para adelantar o retrasar sus viajes: 0 min, 15 min, 30 min, 60 min y 2 horas o más.

La formulación del modelo de elección discreta corresponde a las mejores especificaciones presentadas en la sección 5.3.2. La VL *flexibilidad* fue incorporada en interacción con los atributos más relevantes (tiempo de viaje, demoras horarias y

penalizaciones por duración de actividades). Cabe destacar que los parámetros que representan la utilidad marginal de las interacciones de las variables explicativas con la variable latente, son específicos a cada individuo debido a la presencia del componente de error ν en la ecuación estructural, y hacen que la estimación del modelo deba ser realizada por máxima verosimilitud simulada.

El resumen del ajuste de los modelos híbridos estimados se muestra en la Tabla 6-1. En todos los casos se utilizó la misma muestra (262 individuos, 3.464 observaciones), la misma cantidad de realizaciones MLHS (200), y todos los modelos incorporan una componente de error para capturar el efecto panel.

Tabla 6-1: Bondad de ajuste para modelos híbridos estimados

Modelo híbrido	MNL + panel+VL	MNL + panel + vsg + VL	ML + VL
Número total de parámetros	42	50	47
Parámetros en DCM	23	31	28
Log-verosimilitud total	-5.447,63	-5.417,24	-5.422,98
Log-verosimilitud DCM	-4.031,65	-4.008,83	-4.043,61
Modelo base (sin VL)	MNL + panel	MNL + panel + vsg	ML
Número de parámetros	18	26	23
Log-verosimilitud base	-4.123,43	-4.095,48	-4.055,85

DCM = Modelo de elección discreta.

Fuente: Elaboración propia

La log-verosimilitud de la componente de elección discreta del modelo híbrido, mejora de forma estadísticamente considerable respecto al modelo base (que no contiene la VL) en todos los modelos estimados. Al aplicar un test de razón verosimilitud entre la componente de elección discreta de todos los modelos híbridos y los modelos base (sin variable latente), se concluye que considerar la variable *flexibilidad* como un concepto latente e incluirla dentro del modelo de elección discreta, mejora significativamente la capacidad explicativa de los modelos.

Sorprendentemente, el mejor modelo en términos de bondad de ajuste es el que captura la heterogeneidad de la muestra a través de variaciones sistemáticas de gustos (VSG) y la interacción de la VL con las variables explicativas del modelo. Además, el modelo con peor log-verosimilitud en su componente de elección discreta, es el ML, sugiriendo que el efecto panel no capturado cabalmente por los parámetros aleatorios

especificados. No obstante, la omisión de la VL *flexibilidad* es mejor capturada por el modelo ML, pues los parámetros aleatorios capturan parte de su efecto.

La conclusión más importante del análisis de la bondad de ajuste de los modelos estimados, es que las variables latentes pueden jugar un rol muy significativo en la explicación de la heterogeneidad en los gustos entre personas.

Los resultados de estos modelos también pueden ser tomados como evidencia en contra del uso indiscriminado de modelos con estructuras altamente flexibles (por ejemplo, el logit mixto o el probit) cuya interpretación no siempre es sencilla. En efecto, si bien los modelos complejos, con estructuras de error altamente flexibles, generalmente ofrecen mejores resultados en cuanto a ajuste, la interpretación de sus resultados puede tornarse compleja y los efectos que capturan tienden a estar confundidos entre si. Usando esta muestra en particular, el modelo ML con la mayoría de sus parámetros aleatorios (y considerando la VL) no es superior en términos de bondad de ajuste a un modelo que captura el efecto panel, incorpora variaciones sistemáticas de gustos y una variable latente con racionalidad teórica. Además, pensando en función de su potencial de aplicación a futuro, el modelo que intenta explicar la heterogeneidad mediante variaciones sistemáticas de gustos (definidas con un buen sentido teórico) posiblemente sea más valioso que un modelo ligeramente más poderoso, pero que obtiene su beneficio de la flexibilidad de su estructura.

Adicionalmente, la conveniencia de implementar un enfoque híbrido de modelación fue testado con otra prueba de razón de verosimilitud, que comparó los valores de la log-verosimilitud total del mejor modelo (Modelo MNL + panel + VSG + VL) con la suma de las log-verosimilitudes de las componentes de elección y de VL estimadas por separado (57). Este test es similar a un test de razón de verosimilitud cuando se mezclan datos de distinta naturaleza (Ortúzar & Willumsen, 2011):

$$LR = -2 \{ LL_{DCM} + LL_{VL} - LL_{\text{modelo híbrido}} \} \sim \chi^2_{k,95\%} \quad (57)$$

donde k es el número de parámetros adicionales en la componente de elección discreta debido a la inclusión de la variable latente. En este caso el cálculo anterior da:

$$LR = -2\{-4.095,5 - 1.327,3 - (-5.417,2)\} = 11,2$$

y dado que el estadístico LR es mayor que el valor crítico de χ^2 para un nivel de confianza del 95% y cinco grados de libertad (11,07), se concluye que la estimación conjunta efectivamente vale la pena.

Todos los coeficientes estimados dentro del sub-modelo de elección (Anexo D) tienen los signos correctos, la mayoría son significativos al 95% de confianza y están en línea con las estimaciones de los restantes modelos estimados durante esta tesis.

En la Tabla 6-2 se presenta un resumen de los coeficientes estimados para la componente estructural del sub-modelo de VL del modelo híbrido, los parámetros correspondientes a las interacciones de otros atributos con la variable latente, y el rango de variación de los razones de compromiso de las variables más relevantes para el mejor modelo estimado y todos los individuos.

Tabla 6-2: Resultados más importantes del mejor modelo híbrido

Modelo Estructural	Estimación	Test t robusto
Posibilidad de trabajar desde casa	1,3001	5,3
Penalidad por llegar tarde o retirarse temprano	-0,2287	-3,5
Penalidad por llegar tarde y retirarse temprano	-0,3561	-4,7
Interacciones en la componente DCM	Estimación	Test t robusto
Flexibilidad – Tiempo de viaje	-0,0447	-4,4
Flexibilidad – SDE	0,0155	1,9
Flexibilidad – SDL	0,0359	2,8
Flexibilidad – PTD	0,0108	2,0
Flexibilidad – PTI	0,0033	0,9
Rango de variación – razones de compromiso	Sin flexibilidad	Con flexibilidad
SDE / Tiempo de viaje	0,73 - 1,94	0,32 - 3,33
SDL / Tiempo de viaje	0,43 - 1,98	0,21 - 3,25
PTD / Tiempo de viaje	0,46 - 1,60	0,20 - 1,60
PTI / Tiempo de viaje	0,14 - 0,42	0,06 - 0,46

SDE: penalidad por arribo temprano; SDL: penalidad por arribo tardío; PTD: penalidad por trabajar menos tiempo que el horario exigido; PTI: penalidad por trabajar más tiempo que el horario exigido.

Fuente: Elaboración propia

Los parámetros de la componente estructural del modelo de VL sugieren, correctamente, que las personas con posibilidad de trabajar desde su casa tienen mayor flexibilidad, mientras que las personas que son penalizadas de alguna manera por llegar tarde al trabajo y/o retirarse antes que finalice la jornada laboral, tienen menor flexibilidad.

Como era de esperarse, las personas que pueden ser penalizadas si llegan tarde y si se retiran antes de su trabajo, tienen menos flexibilidad aun que las personas que podrían ser penalizadas sólo en uno de los dos casos.

Los resultados del modelo también indican que las personas con mayor flexibilidad en sus horarios de trabajo tienden a ser más sensibles a aumentos en los tiempos de viaje, dado que su valoración del tiempo de viaje es mayor. Además, el modelo sugiere que a mayor flexibilidad aumenta la disposición de las personas a modificar sus horarios de viaje. En otras palabras, entre más grande sea la flexibilidad respecto a los horarios de salida y entrada del trabajo, las personas se preocupan menos por llegar a la hora pues les produce menos “desutilidad” llegar o partir fuera de los horarios oficiales de la empresa.

Por último, el rango de variación de las razones de compromiso para todos los atributos se encuentra en línea con lo reportado por Li *et al.* (2010), en su artículo de revisión, donde reportan resultados provenientes de diferentes estudios (principalmente evaluando confiabilidad del tiempo de viaje). Al incluir la flexibilidad en el cálculo de las distintas medidas, los rangos de variación aumentan. Por ejemplo, para las personas con menor flexibilidad en los horarios de trabajo, cada minuto adicional respecto al horario preferido de llegada es tres veces más valorado que la reducción de un minuto del tiempo de viaje.

7 MODELO CONJUNTO DE ELECCIÓN DE MODO Y HORA

Las elecciones de modo y hora de inicio de los viajes son dos decisiones de transporte que, junto a la elección de ruta, definen la distribución temporal de la demanda. Desarrollar modelos que simultáneamente consideren las elecciones anteriores es necesario para completar una evaluación adecuada de los posibles comportamientos de las personas cuando realizan sus viajes sobre redes congestionadas, o se implementan estrategias de gestión de la demanda tal como la tarificación vial.

Este capítulo tiene como finalidad presentar un primer enfoque de modelación conjunta de las elecciones de modo y hora en el contexto de *tours* basados en viajes al trabajo. La motivación para utilizar *tours* en lugar de viajes, está relacionada con la posibilidad de incluir restricciones de horarios y de duración de las actividades en el contexto de elección de horario, y con la posibilidad de modelar la potencial elección de diferentes modos de transporte a diferentes horas del día, en viajes que se basan en el trabajo pero tienen diferentes propósitos.

Adicionalmente, a pesar que la mezcla de datos de preferencias declaradas y reveladas es una práctica común en el área de transporte, su aplicación en el contexto del estudio de la elección de hora de salida de viajes es poco común (Börjesson, 2008; Tseng *et al.*, 2011). El enfoque de modelación que se presenta en este capítulo utiliza datos de diferente naturaleza con el fin de capturar las preferencias relacionadas con la elección de modo y hora en el contexto de *tours* simples basados en el trabajo.

7.1 Formulación del modelo conjunto modo - hora

Un problema común, cuando se utilizan datos de PD para la estimación de modelos de elección, es la incertidumbre asociada a la posibilidad que los individuos actúen de la forma que han declarado hacerlo, en la realidad, al enfrentarse a una situación similar a la planteada hipotéticamente en el experimento de elección. Los experimentos de PD, sin embargo, son una excelente fuente de información para capturar los compromisos entre diferentes atributos. Los datos de PR superan el problema de las PD porque representan el comportamiento real de los individuos, pero en algunos casos obtener parámetros significativos de algunos atributos es complejo debido a la presencia de correlaciones entre

atributos o factores no observados. La mezcla de datos de PR y PD permite aprovechar las bondades de ambos tipos de datos y su aplicación en estudios de transporte es cada vez más común.

El modelo presentado en esta sección se basa en la teoría desarrollada para mezclar datos de diferente naturaleza (Louviere *et al.*, 2000). Su finalidad es adicionar, a la componente de PD que captura las preferencias de elección de hora de inicio de viaje, una componente explícita de elección de modo de transporte usando datos de PR. La idea es aprovechar las bondades de los últimos en un contexto de elección de modo, y establecer la escala de las respuestas obtenidas en la componente de PD, considerando la homogeneidad en los gustos de cada persona que responda a ambas componentes de la encuesta.

En una primera aproximación, se propone una estructura MNL para modelar la componente PR de elección de modo. Es esperable que otros modelos, con estructuras de error más flexibles (es decir, que permitan incorporar, por ejemplo, correlación entre alternativas), representen de mejor forma la elección a costa de un incremento en la complejidad de formulación y estimación. Dado que el modelo considera la elección de modo en las etapas del *tour* basado en el trabajo, una de sus bondades debiera ser capturar la posibilidad de cambio de modo entre los diferentes viajes dentro del *tour*. Así, la cantidad total de alternativas consideradas en esta componente de elección corresponde a todas las combinaciones de modo posibles entre los viajes de ida hacia el trabajo y de salida del trabajo. Para determinarlas, se consideró nueve alternativas para cada tipo de viaje: auto como chofer (ACH), auto como acompañante (AAC), caminata (CAM), bus (B), metro (M), combinación de bus y metro (BM), taxi (TX), taxi colectivo (TXC), y combinación de transporte público con transporte privado (CombTP).

Para definir el número de alternativas totales a utilizar en la modelación, se tuvo en cuenta que los viajes que definen el *tour* son consecutivos; por tanto, algunas combinaciones de modos (entre viajes hacia y desde el trabajo) no son factibles o los encuestados no las reportaron como alternativas utilizadas. Teniendo en cuenta lo anterior, se definió 51 alternativas en total para la elección de modo en los *tours* basados en el trabajo (Tabla 7-1).

Tabla 7-1: Alternativas consideradas en la componente de elección de modo para *tours* basados en el trabajo

Alternativa	Modo ida	Modo regreso	Alternativa	Modo ida	Modo regreso
1	ACH	ACH	27	AAC	B
2	M	M	28	TXC	M
3	B	B	29	CAM	M
4	BM	BM	30	TXC	CAM
5	CAM	CAM	31	TX	CAM
6	AAC	AAC	32	M	AAC
7	B	BM	33	CombTP	AAC
8	CombTP	CombTP	34	M	BM
9	B	M	35	CombTP	B
10	BM	B	36	AAC	BM
11	TX	TX	37	M	TXC
12	BM	CAM	38	AAC	TX
13	AAC	M	39	M	B
14	TXC	TXC	40	CAM	B
15	B	CAM	41	TXC	B
16	M	CAM	42	B	TX
17	CombTP	M	43	TX	CombTP
18	CombTP	BM	44	BM	CombTP
19	BM	M	45	TXC	AAC
20	B	CombTP	46	TX	AAC
21	B	AAC	47	BM	TXC
22	AAC	TXC	48	CombTP	TX
23	B	TXC	49	CAM	BM
24	M	CombTP	50	AAC	CombTP
25	TX	BM	51	CAM	TX
26	BM	AAC			

ACH: Auto como chofer; AAC: auto como acompañante; CAM: caminata; B: bus; M: metro; BM: combinación de bus y metro; TX: taxi; TXC: taxi colectivo; y CombTP: combinación de transporte público con transporte privado.

Fuente: Elaboración propia

Dentro de la formulación de la utilidad sistemática para estas 51 alternativas, se consideró atributos como el tiempo de transbordo, tiempo de espera, relación entre número de autos y número de licencias en el hogar (sólo incluida en la alternativa ACH), tiempos de viaje y costos. Cabe aclarar, que se usó una transformación logarítmica del costo porque también utilizada en la formulación de la componente elección de horario de partida y también porque generó un mejor ajuste del modelo.

Para no generar 50 constantes específicas (ASC) en la componente de elección de modo, cuya interpretación sería difícil, se decidió utilizar sólo dos constantes por modo, la primera asociada al viaje de ida hacia el trabajo ($ASC_{modo_{ida}}$) y la segunda al viaje de salida del trabajo ($ASC_{modo_{regreso}}$). Con fines de identificación, se fijó como referencia a

las ASC asociadas a viajar en bus a la ida y regreso del trabajo (el bus es el modo con mayor disponibilidad en la muestra para cada tipo de viaje, ver Tabla 4-2).

De esta manera, la función de utilidad sistemática para la alternativa j puede escribirse de forma genérica como (58):

$$\begin{aligned}
 U_j = & ASC_{modo_ida} + ASC_{modo_regreso} + \beta_{TT} \cdot tiempo\ de\ viaje_j + \beta_{ICOST} \cdot \log(\text{costo}_j) \\
 & + \beta_{Transb} \cdot tiempo\ de\ transbordo_j + \beta_{espera} \cdot tiempo\ de\ espera_j \\
 & + \beta_{autlic} \cdot relación\ entre\ autos\ y\ licencias + e_j
 \end{aligned}
 \tag{58}$$

La formulación de la componente PD, de elección de hora, no se discutirá en detalle en esta sección pues el enfoque es análogo al de los modelos MNL sin variaciones sistemáticas de gustos que consideran efecto panel (ver capítulo 5).

Cabe señalar que los atributos tiempo de viaje y costo se supusieron comunes a las alternativas de PR y PD. Esto implica que su valoración es la misma por parte de un mismo individuo (existe homogeneidad de gustos hacia el tiempo de viaje y costo).

La combinación de las dos componentes de elección (modo y hora) en el modelo conjunto planteado, también permite estimar factores de escala asociados a las diferentes elecciones y tipos de datos utilizados ($THETA_i$). Como el modelo considera ambas elecciones simultáneamente, la cantidad total de alternativas es 65; este número se obtiene de sumar las 51 alternativas de las componentes de modo y 14 alternativas de hora (4 alternativas del experimento PD basado en viajes, ver Figura 3-2, y 10 alternativas del experimento PD basado en *tours*, ver Figura 3-4).

La formulación de las utilidades sistemáticas para todas las alternativas consideradas dentro del modelo conjunto se puede escribir en términos generales como en las ecuaciones (59) a (61):

$$\begin{aligned}
 V_j^{modo} = & THETA1 \cdot (ASC_{modo_ida} + ASC_{modo_regreso} + \beta_{TT} \cdot tiempo\ de\ viaje_j \\
 & + \beta_{ICOST} \cdot \log(\text{costo}_j) + \beta_{transb} \cdot tiempo\ de\ transbordo_j \\
 & + \beta_{espera} \cdot tiempo\ de\ espera_j + \beta_{autlic} \cdot relación\ entre\ autos\ y\ licencias),
 \end{aligned}$$

donde $1 \leq j \leq 51$ (59)

$$\begin{aligned}
V_i^{PD2-hora} = & THETA2 \cdot (ASC_i^{ida} + ASC_i^{regreso} + \beta_{TT} \cdot tiempo\ de\ viaje_i \\
& + \beta_{iCOST} \cdot \log(costo_i) + \beta_{Time_diff} \cdot Time_diff_i + \beta_{SDE}^{ida} \cdot SDE_i^{ida} \\
& + \beta_{SDL}^{ida} \cdot SDL_i^{ida} + \beta_{PTD} \cdot PTD_i + \beta_{PTI} \cdot PTI_i),
\end{aligned}$$

donde $52 \leq i \leq 61$ (60)

$$\begin{aligned}
V_k^{PD1-hora} = & THETA3 \cdot (ASC_k + \beta_{TT} \cdot tiempo\ de\ viaje_k + \beta_{iCOST} \cdot \log(costo_k) \\
& + \beta_{Time_diff} \cdot Time_diff_k + \beta_{SDE}^{ida} \cdot SDE_k^{ida} + \beta_{SDL}^{ida} \cdot SDL_k^{ida}),
\end{aligned}$$

donde $62 \leq k \leq 65$ (61)

7.2 Análisis de resultados del modelo conjunto modo y hora

La estimación del modelo conjunto se realizó en BIOGEME tardando aproximadamente 32 horas en un computador Quad Core a 2.40 GHz y 3.24 GB de memoria RAM. Los resultados muestran que todos los parámetros estimados tienen el signo esperado y la mayoría de ellos son estadísticamente significativos. La muestra total con que se estimó el modelo fue de 500 individuos, dado que se aprovechó los datos de algunas personas que no respondieron la encuesta PD. Se incluyó efecto panel en la componente PD del modelo y se estimó mediante 200 realizaciones MLHS. El factor de escala fijado fue el correspondiente al experimento enfocado en la elección de hora, para *tours* basados en el trabajo.

La Tabla 7-2 muestra el resultado de las estimaciones de todas las constantes incluidas en las diferentes componentes del modelo. El valor de las constantes incluidas en la componente de hora del modelo conjunto, se encuentran en línea con la estimación de las mismas en el caso de modelos MNL para sólo la componente PD. Ellas sugieren que si se implementara una medida de manejo de la demanda, como la tarificación vial, a los individuos les disgustaría cambiar la hora de inicio o el modo que usualmente utilizan para realizar sus viajes basados en el trabajo.

Por otro lado, *ceteris paribus*, los encuestados muestran una mejor disposición a adelantar sus viajes que a retrasarlos o cambiar de modo de transporte. Una disposición similar es observada respecto a cambiar de modo o retrasar sus viajes basados en el trabajo.

Tabla 7-2: Constantes específicas incluidas en el modelo conjunto de modo y hora

Constantes	Valor	Test t robusto
Componente PD de hora		
Cambio de modo	-1,520	-8,4
Salir antes – experimento viaje	-0,850	-2,9
Salir antes viaje ida – experimento <i>tour</i>	-0,768	-3,2
Salir antes viaje regreso – experimento <i>tour</i>	-0,666	-3,0
Salir después – experimento viaje	-1,430	-4,2
Salir después viaje ida – experimento <i>tour</i>	-1,530	-6,2
Salir después viaje regreso – experimento <i>tour</i>	-1,200	-5,1
Componente PR de modo		
<i>Ida y regreso</i>		
Auto como chofer (ACH)	-0,197	-0,1
<i>Ida</i>		
Auto como acompañante (AAC)	-5,660	-5,4
Bus – Metro (BM)	1,410	2,3
Caminata (CAM)	-2,340	-2,0
Combinación con TP (CombTP)	0,975	1,2
Metro (M)	3,400	3,6
Taxi colectivo (TXC)	2,090	2,1
Taxi (TX)	-2,350	-2,0
<i>Regreso</i>		
Auto como acompañante (AAC)	-5,520	-6,1
Bus – Metro (BM)	2,310	3,1
Caminata (CAM)	1,360	1,3
Combinación con TP (CombTP)	0,644	0,8
Metro (M)	5,160	4,2
Taxi colectivo (TXC)	2,680	2,4
Taxi (TX)	-2,330	-2,1

Fuente: Elaboración propia

Al analizar las constantes correspondientes a la componente de modo del modelo conjunto se observa que, *ceteris paribus*, los encuestados prefieren utilizar el transporte público para sus viajes basados en el trabajo. Este resultado, era en parte esperable ya que casi el 70% de las personas de la muestra usa transporte público en al menos uno de sus viajes basados en el trabajo (ver Tabla 4-2). Algunas constantes específicas a cada modo en los viajes de ida y regreso no son estadísticamente significativas entre si, por lo que podría considerarse que *ceteris paribus* el efecto de la elección de modo a diferentes horas del día es el mismo. Modelos futuros pueden incluir constantes genéricas para aquellos modos donde los usuarios no establezcan diferencias notorias en la percepción del modo para los viajes de ida y regreso del trabajo.

Tabla 7-3: Modelo conjunto de elección de modo y hora

Coefficientes	Valor	Test t robusto
<i>Comunes</i>		
Tiempo de viaje	-0,017	-3,4
log – costo	-3,010	-8,8
<i>Factores de escala</i>		
SIGMA	1,770	21,0
THETA1	0,348	5,1
THETA2 (experimento <i>tours</i>)	1,000	Fijado
THETA3 (experimento viajes)	0,833	9,7
<i>Componente de hora</i>		
Penalidad horaria por llegar antes (SDE)	-0,022	-5,9
Penalidad horaria por llegar después (SDL)	-0,028	-6,0
Disminución en horas de trabajo (PTD)	-0,017	-4,9
Incremento en horas de trabajo (PTI)	-0,007	-2,8
Incertidumbre tiempo de viaje (Timediff)	-1,750	-3,7
<i>Componente de modo</i>		
Tiempo de transbordo	-0,414	-4,1
Tiempo de espera	-0,215	-2,0
Relación entre autos y licencias	4,760	2,6
<i>Razones de compromiso</i>		
SDE / Tiempo de viaje		1,31
SDL / Tiempo de viaje		1,66
PTD / Tiempo de viaje		0,99
PTI / Tiempo de viaje		0,41
Tiempo de transbordo / Tiempo de viaje		8,43
Tiempo de espera / Tiempo de viaje		4,38

Fuente: Elaboración propia

En la Tabla 7.3 se presentan los restantes parámetros del modelo conjunto de modo y hora. Estos permiten deducir que, al parecer, a las personas no les agrada la idea de caminar hacia el trabajo si se encuentran a distancias considerables de estos; sin embargo, este modo de transporte no es visto con tan malos ojos en el viaje de salida del trabajo. En parte esto puede deberse a que, en esta base, aproximadamente el 30% de los viajes después del trabajo tienen un propósito diferente a volver a casa.

Cuando la localización de los destinos está relativamente cercana al hogar, caminar se vuelve una opción atractiva. En modelos futuros, considerar el propósito del viaje de

vuelta y la distancia a recorrer, pueden ser variables que mejoren la capacidad explicativa de los modelos.

Los coeficientes asociados a los atributos de la componente de elección de la hora de inicio de viaje son también consistentes con los estimados anteriormente con los datos de PD solos. La componente de error que captura el efecto panel debido a las múltiples respuestas de un mismo individuo, es altamente significativo, lo cual también se encuentra en línea con lo reportado en el capítulo 5.

Un resultado llamativo se obtiene al analizar la magnitud de los factores de escala respecto al experimento PD basado en *tours*. El modelo conjunto sugiere que la varianza entre las respuestas del experimento más simple (viajes), es mayor que las respuestas obtenidas en el experimento complejo (*tours*); sin embargo, esto podría explicarse porque en la componente de elección de horario basada en *tours* se tienen más variables explicativas asociadas, en particular las que se refieren a penalidades debidas a la variación de la duración de las actividades. Por su parte, el factor de escala del modelo PR (de elección de modo) es menor que los factores de escala asociados a la componente de elección de hora, sugiriendo que esta última es una componente relevante del análisis del comportamiento de la demanda individual si se desea evaluar estrategias para el manejo de la demanda.

Las razones entre los parámetros estimados y el coeficiente del tiempo de viaje (razones de compromiso) sugieren que cada minuto adicional que implique llegar tarde o temprano al trabajo, es más oneroso que un minuto adicional del tiempo de viaje. En particular, llegar tarde es peor valorado que llegar temprano al lugar de trabajo, en línea con resultados previos (Hess *et al.*, 2007b; Small, 1982). Por otro lado, llegar un minuto antes cuando se va retrasado se valora 66% más que una reducción de un minuto en el tiempo usual de viaje. La tendencia de las razones de compromiso relacionadas con las penalidades asociadas a la duración del trabajo, se encuentran en línea con lo reportado en capítulos previos de esta tesis.

Por último, un resultado muy interesante de la componente de elección de modo es la posibilidad de comparar las razones de compromiso de los tiempos de transbordo y espera, con los de la componente de elección de hora. Un minuto adicional de espera o

transbordo, es peor valorado que llegar un minuto tarde al trabajo o tener que cambiar las horas de trabajo en el día. De hecho, reducir en un minuto el tiempo de transbordo es equivalente a reducir ocho minutos el tiempo de viaje usual, y reducir el tiempo de espera en un minuto adicional es equivalente a reducir cuatro minutos el tiempo de viaje usual.

8 CONCLUSIONES

Esta tesis se enfoca en el desarrollo de modelos avanzados de elección de la hora de inicio de viajes, concentrándose en particular en *tours* de ida y vuelta al trabajo. En el documento se presenta una metodología que permite recolectar datos de diferente naturaleza para estudiar la elección de la hora de inicio de viaje. Usando los datos recolectados, se formularon, estimaron y discutieron diferentes modelos que permiten:

- evaluar las preferencias individuales frente a diferentes atributos relacionados con la elección de horario de inicio de viaje, en un contexto de viajes al trabajo y de *tours* basados en el trabajo
- estudiar la forma más adecuada de incorporar la heterogeneidad de gustos en este tipo de modelos
- probar diferentes formulaciones para tratar el efecto panel
- incorporar la flexibilidad en el horario de trabajo como un concepto latente, y
- estudiar conjuntamente las elecciones de modo y hora de partida de viajes en el contexto de *tours* basados en el trabajo utilizando datos de diferente naturaleza.

Con respecto a la recolección y procesamiento de datos:

La metodología propuesta en esta tesis permite recolectar, mediante un novedoso procedimiento de varias etapas, datos de preferencias reveladas, preferencias declaradas y percepciones para construir variables latentes, con el fin de estudiar en forma específica la elección de la hora de inicio de viaje.

Una característica interesante de la metodología propuesta, consiste en el uso de diferentes métodos de recolección como parte de un mismo procedimiento. En particular, los datos se recopilaron mediante formularios de auto-llenado en línea y entrevistas personalizadas. Además, el proceso de recolección incluyó una serie de contactos virtuales y personales (vía email y visitas a los lugares de trabajo), el uso de diferentes tipos de recordatorios (emails y formularios llamativos), y el diseño y utilización de herramientas para incrementar la tasa de respuesta, minimizar la fatiga y disminuir la pérdida de interés hacia la encuesta (realización de rifas de dinero, cartas de agradecimiento, y uso del mismo entrevistador, para cada encuestado, durante las diferentes etapas del proceso).

A pesar que el procedimiento de recolección de datos es costoso en términos de tiempo (asociado a obtener los datos y planificar el proceso completo), la encuesta no tuvo un costo monetario demasiado elevado. De hecho, el costo de cada respuesta de preferencias reveladas (sin incluir el tiempo del autor de esta tesis) se estima en US\$ 21,14 mientras que el costo de cada respuesta de preferencias declaradas fue de aproximadamente US\$ 9,66.

La componente PR también se considera una experiencia exitosa, en que se aplicó un diario de viajes (diseñado en un marco basado en actividades) que fue completado en un proceso de dos etapas; primero, una entrevista asistida por computador y luego un formulario de auto-llenado vía web.

Otra característica interesante de la encuesta de PR fue el uso de variables de nivel de servicio obtenidas con muy alta precisión (en lugar de utilizar datos promedio para hora punto o aún más agregados como es la práctica usual). Para esto se desarrolló un procedimiento de seis pasos que permite obtener niveles de servicio usando Google MapsTM, un manejador de bases de datos y un *software* de programación. Algunas características llamativas del procedimiento son que puede ser implementado a bajo costo en términos de requerimientos de *software*, que entrega resultados relativamente rápidos en comparación con otras metodologías disponibles en la literatura, y que sus resultados pueden ser analizados y desplegados a cualquier nivel de agregación espacio - temporal. Los niveles de servicio de alta precisión conseguidos, fueron utilizados en esta tesis para potenciar la calidad de la base de datos de PR e informar el diseño de la encuesta PD.

Los cuestionarios incluidos en las encuestas, para recolectar información detallada acerca de los empleos y acerca de la programación de actividades de los encuestados, fueron claves para obtener información relevante como las horas preferidas de inicio/fin de los viajes al trabajo y las restricciones en cuanto a la jornada laboral de cada persona. El detalle de la información recolectada permitió definir indicadores útiles para la definición de una variable latente, denominada *flexibilidad en los horarios de trabajo*, que permitió estimar modelos híbridos de elección de la hora de inicio de viaje.

El diseño eficiente de los datos de PD recolectados en esta tesis es totalmente nuevo, ya que permite incorporar dependencia entre atributos de una misma alternativa en

experimentos de elección, utilizando un procedimiento de optimización del diseño en dos etapas. Además, el procedimiento permite incluir restricciones de dominancia e incorporar umbrales mínimos de variación entre los niveles de los atributos de diferentes alternativas, con el fin de no presentar comparaciones irrelevantes y aumentar el realismo del experimento.

Como desventajas del procedimiento de encuestas desarrollado en esta tesis, se puede señalar que habría sido deseable contar con un mayor tamaño de muestra; sin embargo, la cantidad de datos recolectados permitió la estimación de modelos avanzados con buenos índices de bondad de ajuste. Además, los patrones observados para la hora de inicio de viaje son similares a los reportados en la EOD de Santiago para el 2006, y los individuos provienen de prácticamente todas las comunas de Santiago.

Con respecto a la formulación y estimación de modelos de elección de la hora de inicio de viaje:

Durante el desarrollo de la tesis se evaluaron diferentes formas funcionales relativas a la inclusión de atributos considerados relevantes. El análisis realizado es robusto, dado que para el caso de datos de PD se analizaron los compromisos entre diferentes atributos en dos contextos: (i) viajes hacia el trabajo y (ii) *tours* definidos por los viajes de ida y salida del trabajo. Los resultados permitieron comprobar que atributos tradicionales en modelación horaria, como el costo, tiempo de viaje, demoras horarias, y penalidades asociadas a la duración del trabajo, eran efectivamente relevantes en el estudio de esta elección en Santiago. En esta tesis, además, se incluyó explícitamente atributos como el *confort* y la variabilidad del tiempo de viaje, los cuales evidenciaron ser también relevantes en la elección de horario de partida de viajes. En particular, la variabilidad del tiempo de viaje se incluyó dentro del diseño de forma poco convencional (la forma típica de considerar la variabilidad del tiempo de viaje presenta cinco posibles tiempos de viaje con sus probabilidades de ocurrencia), pero - a pesar de ello - la estimación arrojó resultados en línea con lo reportado en la literatura donde se evalúa la confiabilidad de los tiempos de viaje.

La tesis también presenta una discusión metodológica interesante acerca del efecto y la forma de incorporar el *efecto panel* (correlación debida a la presencia de múltiples

respuestas de un individuo) en el contexto de la elección en estudio. Se realizó una serie de comentarios tendientes a identificar las falencias de los métodos tradicionales para incorporar el efecto panel y se presentó una forma de incorporarlo que no genera el indeseado efecto de correlación entre alternativas. De acuerdo a los resultados de esta tesis, el efecto panel juega un rol clave en la mejora del ajuste de los modelos, si se especifica correctamente.

Otra contribución de esta tesis se generó a partir de la discusión del efecto panel y se refiere al concepto definido como “efecto panel variable”. Este propone que la correlación generada por el hecho de tener múltiples respuestas para cada individuo, puede expresarse como función de las características socioeconómicas de las personas. Esto es, el modelo con efecto panel variable permite que la correlación entre las respuestas individuales varíe de acuerdo a las características de cada individuo, lo que puede determinar que las respuestas de cada uno estén mas o menos correlacionadas entre sí.

Otro aspecto discutido durante la tesis, se refiere a la forma de incorporar la heterogeneidad. Los resultados evidenciaron que cuando se especifica un modelo con parámetros aleatorios para capturar heterogeneidad (tipo logit mixto), las mejoras en el ajuste son evidentes, pero su interpretación debe ser cuidadosa debido a que pueden existir, como se ha encontrado en ocasiones anteriores, confusión de efectos. En particular, se encontró que parte importante de la mejora del ajuste del modelo era precisamente debido al efecto panel. Además, resultó interesante notar que un modelo con parámetros fijos bien especificado, permitiendo adecuadas variaciones sistemáticas de gustos, puede incluso superar una especificación más compleja de parámetros aleatorios; con la enorme ventaja de su simpleza en términos de uso futuro en etapas de predicción.

La incorporación de la variable latente *flexibilidad en los horarios de trabajo* en modelos de elección de la hora de inicio de viaje, bajo un enfoque de modelación híbrido, es otro aspecto novedoso de esta investigación. La incorporación de la variable latente resultó relevante para explicar la heterogeneidad en la valoración de atributos como el tiempo de viaje, las demoras horarias, y las penalidades por cambio en las duraciones de las horas trabajadas. La *flexibilidad en los horarios de trabajo* se definió como un concepto latente, que depende de las características del empleo y que se puede observar a

través de la cantidad de minutos que una persona puede adelantar o retrasar sus viajes. Se considera que el enfoque presentado en esta tesis, es teóricamente superior a la tradicional definición de flexibilidad como una variable *dummy* que sólo indica si la persona tiene flexibilidad o no.

Por último, se desarrolló un modelo conjunto de elección de modo y hora de inicio de viaje en el contexto de *tours* y mezclando datos de diferente naturaleza. Aunque el modelo desarrollado no recoge toda la complejidad en torno a la modelación de las elecciones consideradas, es una aproximación que permite reconocer la relación existente entre las elecciones de modo y hora de inicio del viaje. El hecho de plantear el modelo utilizando la teoría de mezcla de datos y que la elección de modo sea descrita a través de datos PR, mientras que la elección de hora sea modelada a partir de los experimentos PD, representa una ventaja para una comparación (teniendo en cuenta las posibles diferencias de escalas) de los atributos que describen las dos elecciones.

Finalmente, en relación a potenciales futuras líneas de investigación, se plantea aplicar el nuevo concepto de *efecto panel variable* sobre otras muestras, para validar los resultados obtenidos en esta tesis, e identificar otras variables que puedan determinar una mayor o menor correlación entre las múltiples respuestas de un individuo.

Respecto a la captura de la heterogeneidad, se desea comparar los mejores modelos estimados con modelos de elección tipo logit mixto considerando otros supuestos respecto a la distribución de sus parámetros; en esta tesis el modelo de parámetros aleatorios, fue estimado asumiendo que los parámetros en la población seguían una distribución Normal.

Respecto a la inclusión de variables latentes en modelos de elección de la hora de inicio de viaje, se plantea -primero- revisar a fondo las consideraciones de identificabilidad del modelo planteado en la tesis (que incluye la flexibilidad como un concepto latente). Adicionalmente, se piensa refinar este modelo a través de la posible incorporación de otras variables latentes que fueron potencialmente identificadas: *puntualidad* y *disposición al cambio*, y que podrían resultar relevantes en el proceso de elección.

En referencia a la modelación conjunta, es claro que deben plantearse estructuras de modelación que permitan incorporar correlación entre las alternativas de modo, e incluso incorporar estructuras más complejas en la componente de elección de hora de inicio del

viaje. Adicionalmente, un análisis de elasticidades cruzadas sería útil para evaluar el efecto en la elección de modo u hora cuando se varía un atributo en alguno de los componentes de elección.

Otras direcciones futuras de investigación podrían enfocarse a la aplicación de estos modelos en un contexto predictivo, usando una nueva muestra para ello. Además, los modelos podrían extenderse hacia otros contextos de aplicación: por ejemplo, la elección de itinerario de viajes interurbanos, en transporte aéreo y terrestre; y la aplicación del enfoque de modelación en el contexto de la elección de paquetes turísticos cuando se toman vacaciones.

9 BIBLIOGRAFÍA

- Abkowitz, M.D. 1981. An analysis of the commuter departure time decision. *Transportation*, 10, 283-297.
- Alvarez-Daziano, R. & Bolduc, D. 2011. Incorporating pro-environmental preferences toward green automobile technologies through a Bayesian hybrid choice model. *Transportmetrica*, DOI: 10.1080/18128602.2010.524173.
- Amador, F.J., González, R.M. & Ortúzar, J. de D. 2008. On confounding preference heterogeneity and income effect in discrete choice models. *Networks and Spatial Economics*, 8, 97-108.
- Ampt, E. & Ortúzar, J. de D. 2004. On best practice in continuous large-scale mobility surveys. *Transport Reviews*, 24, 337-363.
- Arellana, J., Daly, A.J., Hess, S., Ortúzar, J. de D. & Rizzi, L. I. 2012a. Developing surveys for the study of departure time choice: a two-stage efficient design approach. *Transportation Research Record (en imprenta)*.
- Arellana, J., Olarú, D. , Ortúzar, J. de D. & Rizzi, L. I. 2012b. All travellers care about the same when deciding the time of travel? A multivariate analysis of departure time preferences. *Working paper*. Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile.
- Arellana, J., Ortúzar, J. de D. & Rizzi, L. I. 2012c. Diseño experimental de PD con dependencia entre niveles de atributos: Aplicación a modelos de elección de horario de viajes. *Revista Ingeniería de Transporte (en imprenta)*.
- Arellana, J., Ortúzar, J. de D. & Rizzi, L. I. 2012d. Survey data to model time-of-day choice: methodology and findings. In: Zmud, J., Lee-Gosselin, M., Munizaga, M.A. & Carrasco, J.A. (eds.) *Transport Survey Methods: Scoping the Future While Staying on Track*. Bingley: Emerald.
- Arellana, J., Ortúzar, J. de D., Rizzi, L.I. & Zuñiga, F. 2012e. Obtaining public transport level-of-service measures using in-vehicle GPS data and freely available GIS web-based tools. *Conference on Advanced Systems for Public Transport - CASPT 12*. Santiago, Chile.
- Arnott, R., de Palma, A. & Lindsey, R. 1990. Departure time and route choice for the morning commute. *Transportation Research*, 24B, 209-228.
- Bajwa, S., Bekhor, S., Kuwahara, M. & Chung, E. 2008. Discrete choice modeling of combined mode and departure time. *Transportmetrica*, 4, 155-177.

- Ben-Akiva, M. , Walker, J., Bernardino, A.T., Gopinath, D.A., Morikawa, T. & Polydoropoulou, A. 2002. Integration of choice and latent variable models. *In: Mahmassani, H.S. (ed.) Perpetual Motion: Travel Behaviour Research Opportunities and Challenges*. Amsterdam: Pergamon.
- Ben-Akiva, M. & Abou-Zeid, M. 2012. Methodological issues in modelling time-of-travel preferences. *Transportmetrica*, DOI: 10.1080/18128602.2012.686532.
- Bhat, C. R. 1998a. Analysis of travel mode and departure time choice for urban shopping trips. *Transportation Research*, 32B, 361-371.
- Bhat, C.R. 1998b. Accommodating flexible substitution patterns in multi-dimensional choice modeling: formulation and application to travel mode and departure time choice. *Transportation Research*, 32B, 455-466.
- Bhat, C.R. & Steed, J.L. 2002. A continuous-time model of departure time choice for urban shopping trips. *Transportation Research*, 36B, 207-224.
- Bianchi, R., Jara-Díaz, S.R. & Ortúzar, J. de D. 1998. Modelling new pricing strategies for the Santiago Metro. *Transport Policy*, 5, 223-232.
- Bierlaire, M. 2003. BIOGEME: A free package for the estimation of discrete choice models. *Proceedings of the 3rd Swiss Transportation Research Conference*. Ascona.
- Bolduc, D., Boucher, N. & Alvarez-Daziano, R. 2008. Hybrid choice modeling of new technologies for car choice in Canada. *Transportation Research Record*, 2082, 63-71.
- Bollen, K.A. 1989. *Structural equations with latent variables*. Chichester, John Wiley & Sons, Ltd.
- Börjesson, M. 2008. Joint RP-SP data in a mixed logit analysis of trip timing decisions. *Transportation Research*, 44E, 1025-1038.
- Cantillo, V., Heydecker, B.G. & Ortúzar, J. de D. 2006. A discrete choice model incorporating thresholds for perception in attribute values. *Transportation Research*, 40B, 807-825.
- Cardell, N.S. & Reddy, B. 1977. A multinomial logit model which permits variations in tastes across individuals. *Working Paper*. Charles River Associates, Boston.
- Carrasco, J.A. & Ortúzar, J. de D. 2002. Review and assessment of the nested logit model. *Transport Reviews*, 22, 197-218.
- Caussade, S., Ortúzar, J. de D., Rizzi, L.I. & Hensher, D.A. 2005. Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research*, 39B, 621-640.
- Cortés, C.E., Gibson, J., Gschwender, A., Munizaga, M. & Zuñiga, M. 2011. Commercial bus speed diagnosis based on GPS-monitored data. *Transportation Research*, 19C, 695-707.

- Chin, A.T.H. 1990. Influences on commuter trip departure time decisions in Singapore. *Transportation Research*, 24A, 321-333.
- Daganzo, C.F. 1979. *Multinomial Probit: The Theory and its Applications to Demand Forecasting*. New York, Academic Press.
- Daly, A.J., Hess, S., Patruni, B., Potoglou, D. & Rohr, C. 2012. Using ordered attitudinal indicators in a latent variable choice model: A study of the impact of security on rail travel behaviour. *Transportation*, 39, 267-297.
- Daly, A.J. & Ortúzar, J. de D. 1990. Forecasting and data aggregation: theory and practice. *Traffic Engineering and Control*, 31, 632-643.
- Daly, A.J. & Zachary, S. 1978. Improved multiple choice models. In: Hensher, D.A. & Dalvi, M.Q. (eds.) *Determinants of Travel Choice*. Saxon House, Westmead.
- Day, N., Habib, K.N. & Miller, E.J. 2010. Analysis of work trip timing and mode choice in the Greater Toronto Area. *Canadian Journal of Civil Engineering*, 37, 695-705.
- de Jong, G., Daly, A.J., Pieters, M., Vellay, C., Bradley, M. & Hofman, F. 2003. A model for time of day and mode choice using error components logit. *Transportation Research*, 39E, 245-268.
- DICTUC 2006. Encuesta de movilidad 2006 de Santiago. Santiago: Departamento de Ingeniería de Transporte, Pontificia Universidad Católica de Chile.
- Domencich, T. & McFadden, D. 1975. *Urban Travel Demand: A Behavioural Analysis*. North-Holland, Amsterdam.
- Doornik, J.A. 2007. Object-oriented matrix programming using Ox. 3rd ed. London: Timberlake Consultants Press and Oxford: www.doornik.com.
- Ettema, D., Ashiru, O. & Polak, J. 2004. Modeling timing and duration of activities and trips in response to road-pricing policies. *Transportation Research Record*, 1894, 1-10.
- Ettema, D., Bastin, F., Polak, J. & Ashiru, O. 2007. Modelling the joint choice of activity timing and duration. *Transportation Research*, 41A, 827-841.
- Godoy, G. & Ortúzar, J. de D. 2009. On the estimation of mixed logit models. In: Inweldi, P.O. (ed.) *Transportation Research Trends*. New York: Nova Science Publishers.
- Greene, W.H. & Hensher, D.A. 2010. *Modeling ordered choices: A primer*. New York, Cambridge University Press.
- Hendrickson, C. & Plank, E. 1984. The flexibility of departure times for work trips. *Transportation Research*, 18A, 25-36.

- Hensher, D.A. & Mannering, F.L. 1994. Hazard-based duration models and their application to transport analysis. *Transport Reviews*, 14, 63-82.
- Hess, S., Daly, A.J., Rohr, C. & Hyman, G. 2007a. On the development of time period and mode choice models for use in large scale modelling forecasting systems. *Transportation Research*, 41A, 802-826.
- Hess, S., Polak, J., Daly, A.J. & Hyman, G. 2007b. Flexible substitution patterns in models of mode and time of day choice: new evidence from the UK and the Netherlands. *Transportation*, 34, 213-238.
- Hess, S. & Rose, J.M. 2009. Allowing for intra-respondent variations in coefficients estimated on repeated choice data. *Transportation Research*, 43B, 708-719.
- Hess, S., Train, K. & Polak, J. 2006. On the use of a Modified Latin Hypercube Sampling (MLHS) method in the estimation of a Mixed Logit Model for vehicle choice. *Transportation Research*, 40B, 147-163.
- Heydecker, B.G. & Addison, J.D. 2005. Analysis of Dynamic Traffic Equilibrium with Departure Time Choice. *Transportation Science*, 39, 39-57.
- Hojman, P., Ortúzar, J. de D. & Rizzi, L.I. 2005. On the joint valuation of averting fatal and severe injuries in highway accidents. *Journal of Safety Research*, 36, 377-386.
- Kitamura, R., Chen, C. & Narayanan, R. 1998. Traveller destination choice behaviour: effects of time of day, activity duration, and home location. *Transportation Research Record*, 1645, 76-81.
- Koster, P. & Tseng, Y.Y. 2010. Stated choice experimental designs for scheduling models. In: Hess, S. & Daly, A. (eds.) *Choice Modelling: The State-of-the-Art and the State-of-Practice*. Bingley: Emerald.
- Lee, B. & Timmermans, H.J.P. 2007. A latent class accelerated hazard model of activity episode durations. *Transportation Research*, 41B, 426-447.
- Li, Z., Hensher, D.A. & Rose, J.M. 2010. Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence. *Transportation Research*, 46E, 384-403.
- Louviere, J.J., Hensher, D.A. & Swait, J.D. 2000. *Stated Choice Methods: Analysis and Application*. Cambridge, University Press, Cambridge
- McFadden, D. & Train, K. 2000. Mixed MNL models for discrete response. *Journal of applied Econometrics*, 15, 447-470.
- MIDEPLAN-SECTRA 2005. Análisis y actualización del modelo ESTR AUS. Santiago: SECTRA.

- Muthén, L.K. & Muthén, B.O. 2007. Mplus user's guide. Sixth edition. Los Angeles, CA: Muthén & Muthén.
- Noland, R.B. & Small, K.A. 1995. Travel-time uncertainty, departure time choice, and the cost of morning commutes *Transportation Research Record*, 1493, 150-158.
- Ortúzar, J. de D. 2006. Travel survey methods in Latin America. In: Stopher, P. & Stecher, C. (eds.) *Travel Survey Methods: Quality and Future Directions*. Oxford: Elsevier.
- Ortúzar, J. de D. & Rodríguez, G. 2002. Valuing reductions in environmental pollution in a residential location context. *Transportation Research*, 7D, 407-427.
- Ortúzar, J. de D. & Willumsen, L.G. 2011. *Modelling Transport*, Cuarta Edición. Chichester, John Wiley and Sons.
- Polak, J. & Jones, P. 1994. A tour-based model of journey scheduling under road pricing. *73rd Annual Meeting of the Transportation Research Board*. Washington, DC.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P. 1992. *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edition. Cambridge, Cambridge University Press.
- Raveau, S., Alvarez-Daziano, R., Yáñez, M.F., Bolduc, D. & Ortúzar, J. de D. 2010. Sequential and simultaneous estimation of hybrid discrete choice models: some new findings. *Transportation Research Record*, 2156, 131-139.
- Rizzi, L.I. & Limonado, J.P. 2008. Integración y medición de los costos externos de la congestión mediante el uso de encuestas de preferencias declaradas. *XV Congreso Panamericano Ingeniería Tránsito y Transporte*. Cartagena (Colombia).
- Rose, J.M. & Bliemer, M.C.J. 2008. Stated Preference Experimental Design Strategies. In: Hensher, D.A. & Button, K.J. (eds.) *Handbook of Transport Modelling*. Segunda edición. Amsterdam: Elsevier.
- Rose, J.M., Bliemer, M.C.J., Hensher, D.A. & Collins, A.T. 2008. Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research*, 42B, 395-406.
- Saleh, W. & Farrell, S. 2005. Implications of congestion charging for departure time choice: Work and non-work schedule flexibility. *Transportation Research*, 39A, 773-791.
- Senbil, M. & Kitamura, R. 2004. Reference points in commuter departure time choice: A prospect theoretic test of alternative decision frames. *Intelligent Transportation Systems*, 8, 19-31.
- Sillano, M. & Ortúzar, J. de D. 2005. Willingness-to-pay estimation with mixed logit models: some new evidence. *Environment and Planning A*, 37, 525-550.

- Small, K.A. 1982. The scheduling of consumer activities: work trips. *The American Economic Review*, 72, 467-479.
- Small, K.A. 1987. A discrete choice model for ordered alternatives. *Econometrica*, 55, 409-424.
- Smith, M.E. 1979. Design of small sample home interview travel surveys. *Transportation Research Record*, 701, 29-35.
- Steed, J.L. & Bhat, C.R. 2000. On modelling departure time choice for home-based social/recreational and shopping trips. *Transportation Research Record*, 1706, 152-159.
- Stopher, P., Wilmot, C., Stecher, C. & Alsnih, R. 2006. Household travel surveys: proposed standards and guidelines. In: Stopher, P. & Stecher, C. (eds.) *Travel Survey Methods: Quality and Future Directions*. Oxford: Elsevier.
- Train, K. 2003. *Discrete Choice Methods with Simulation*, Segunda edición. Cambridge, University Press, Cambridge.
- Train, K. & Wilson, W. 2009. Monte Carlo analysis of SP-off-RP data. *Journal of Choice Modelling*, 2(1), 101-117.
- Tringides, C., Ye, X. & Pendyala, R.M. 2004. Departure-time choice and mode choice for non-work trips: alternative formulations of joint model systems. *Transportation Research Record*, 1898, 1-9.
- Tseng, Y.Y., Koster, P., Peer, S., Knockaert, J. & Verhoef, E. 2011. Discrete choice analysis for trip timing decisions of morning commuters: estimations from joint SP/RP-GPS data. *International Choice Modelling Conference*. Leeds.
- Tseng, Y.Y. & Verhoef, E. 2008. Value of time by time of day: A stated-preference study. *Transportation Research*, 42B, 607-618.
- Vickrey, W. 1969. Congestion theory and transport investment. *The American Economic Review*, 59, 251-260.
- Walker, J. 2001. *Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures and Latent Variables*. Ph.D. Thesis, MIT.
- Wang, J.J. 1996. Timing utility of daily activities and its impact on travel. *Transportation Research*, 30A, 189-206.
- Williams, H.C.W.L. 1977. On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning*, 9A, 285-344.
- Yañez, M.F., Cherchi, E., Heydecker, B.G. & Ortúzar, J. de D. 2011. On the treatment of repeated observations in panel data: efficiency of mixed logit parameter estimates. *Networks and Spatial Economics*, 11, 393-418.

Yañez, M.F., Cherchi, E., Ortúzar, J. de D. & Heydecker, B.G. 2009. Inertia and shock effects on mode choice panel data: implications of the Transantiago implementation. *12th International Conference on Travel Behaviour Research*. Jaipur (India).

10 ANEXOS

**ANEXO A: LISTADO DE ATRIBUTOS PRESENTADOS EN ENCUESTA
CUALITATIVA PREVIA**

El cuestionario cualitativo preliminar recopiló la evaluación de la importancia de una serie de atributos en las elecciones de hora de los viajes laborales y no laborales, realizados en días laborales. Luego de contestar unas pocas preguntas socioeconómicas, los individuos calificaron la influencia de los atributos sobre la elección de la hora de inicio de los viajes. Primero evaluaron los atributos pensando en un viaje laboral, y luego evaluaron los atributos pensando en un viaje no laboral dentro de un día hábil. El encuestado debía calificar los atributos que se muestran en la tabla A-1, en una escala 1 a 7 (donde 1 indicaba menor influencia y 7 mayor influencia).

Tabla A-1: Atributos presentados en la encuesta preliminar por tipo de viaje

Atributo	Viaje	
	Laboral	No laboral
1 Horas a trabajar en el día	X	X
2 Información previa sobre congestión	X	X
3 Tiempo de viaje	X	X
4 Costo del viaje	X	X
5 Actividades que debe realizar después	X	X
6 Si debe viajar con alguien	X	X
7 Hora en la que pasa el bus de la empresa	X	X
8 Seguridad durante el viaje	X	X
9 Actividades que debe realizar antes	X	
10 Hora en que inicia su actividad	X	X
11 Hora en que termina su actividad	X	
12 Modo de transporte en el que viajará	X	X
13 Congestión en el viaje	X	X
14 Congestión en los vehículos de transporte público	X	X
15 Disponibilidad de estacionamientos en el destino	X	X
16 Comodidad durante el viaje	X	X
17 Es de día o está oscuro	X	X
18 Cantidad de trabajo en el día	X	
19 Condiciones climáticas (lluvia, frío, calor)	X	X
20 Hora que se encuentra menos lleno el destino		X

Fuente: Elaboración propia

ANEXO B: CARTAS Y FORMULARIOS UTILIZADOS EN LA ENCUESTA

En este anexo se presentan las cartas y formularios usados para la recolección de los datos. Los cuestionarios de las partes 1 y 2 de la componente de preferencias declaradas no se muestran en este anexo por motivos de espacio. La persona interesada, favor contactar al autor.

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE DEPARTAMENTO DE INGENIERÍA DE TRANSPORTE Y LOGÍSTICA

CARTA DE INVITACIÓN

Santiago, Noviembre de 2010

De mi consideración,

Como parte de un proyecto de Investigación financiado por el Fondo Nacional de Desarrollo Científico y Tecnológico (Proyecto FONDECYT 1100483), estamos desarrollando una tesis doctoral sobre el tema “Modelos de elección del horario de inicio de viajes”. Nuestro objetivo es lograr formular y estimar modelos de comportamiento que permitan predecir la hora en que las personas inician sus viajes durante los días hábiles de la semana, dependiendo de diferentes características del sistema de transporte. Este es un tema cada vez más importante en planificación urbana, dadas las dificultades crecientes experimentadas por los ciudadanos debido a la congestión.

Para llevar a cabo la investigación, es necesario recopilar abundante información relativa a los viajes que realizan las personas durante el día. Para esto, contemplamos la realización de encuestas a personas que realicen viajes al trabajo en forma relativamente rutinaria durante la semana. Dichas encuestas serán coordinadas en terreno por el estudiante de Doctorado en Ciencias de la Ingeniería, Julián Arellana (Fono: 354 17 96; Celular 09 - 93201867; email: jlarella@uc.cl), y considerarán un máximo tres breves visitas a cada encuestado en días diferentes.

La presente tiene por objeto solicitar a Ud. la autorización para encuestar a parte de su personal, en el establecimiento u oficina a su cargo. La encuesta será coordinada personalmente por Julián y contará con el apoyo de encuestadores debidamente capacitados e identificados.

Esperando una favorable acogida, se despide atentamente,

Juan de Dios Ortúzar Salas
Profesor Titular
Departamento de Ingeniería de Transporte y Logística
Pontificia Universidad Católica de Chile
jos@ing.puc.cl

FORMULARIO DE CONSENTIMIENTO INFORMADO

Yo,....., doy mi consentimiento para participar en las encuestas que forman parte del proyecto de investigación “Modelos de Elección de Horario de Inicio de Viajes”, llevada a cabo por la Facultad de Ingeniería de la Pontificia Universidad Católica de Chile

Al dar mi consentimiento, yo reconozco que:

1. Se me han explicado todos los procedimientos y el tiempo requerido para participar en la encuesta, y toda pregunta sobre el proyecto ha sido respondida a mi entera satisfacción.
2. He leído la Declaración de Información para el Participante y se me ha ofrecido la oportunidad de examinar toda la información sobre mi participación en el proyecto.
3. Entiendo que mi participación es estrictamente confidencial y que ninguna información que revele mi identidad será utilizada en modo alguno.
4. Entiendo que mi participación en esta encuesta es completamente voluntaria – no estoy bajo ninguna presión para participar ni entregar mi consentimiento.
5. Entiendo que puedo retirarme de la encuesta en cualquier momento, sin que ello afecte mi relación con el investigador(a) ahora o en el futuro.
6. Entiendo que si no quisiera continuar cualquier información que pude haber dado al entrevistador hasta ese momento será destruida.

Firma:

Fecha:

DECLARACION DE INFORMACION PARA EL PARTICIPANTE

Proyecto: Modelos de Elección de Horario de Inicio de Viajes

(1) ¿De qué se trata el estudio?

El objetivo del estudio es formular y estimar modelos de elección de la hora en que los trabajadores inician sus viajes en días laborales.

(2) ¿Quién está desarrollando el estudio?

Coordinador de la investigación: Julián Arellana, Estudiante de Doctorado, Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile. Vicuña Mackenna 4860, Macul, Santiago. Fono: (562) 354-1796, Fax: (562) 553-0281. Correo electrónico: jlarella@uc.cl

Director de la investigación: Juan de Dios Ortúzar, Profesor Titular, Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile. Vicuña Mackenna 4860, Macul, Santiago. Fono: (562) 354-4822, Fax: (562) 553-0281. Correo electrónico: jos@ing.puc.cl.

(3) ¿Qué involucra mi participación en el estudio?

Su participación involucra participar en una encuesta que tiene tres etapas. En ésta se le harán preguntas sobre los viajes realizados en un día laboral determinado. Se le preguntará sobre el motivo de los viajes, medios de transporte utilizados, flexibilidad horaria y otras características de los mismos. También se le solicitará información socio-económica (edad, género, nivel de educación, etc.). La información recabada será almacenada sin vinculación alguna a su nombre, a fin de resguardar completamente la confidencialidad de sus datos.

(4) ¿Cuánto tiempo dura mi participación?

En promedio, la encuesta completa toma aproximadamente 30 minutos.

(5) ¿Puedo retirarme anticipadamente del ejercicio?

Su participación es completamente voluntaria – usted no está bajo obligación alguna de acceder a participar. Incluso si accede, puede retirarse en cualquier momento sin ningún tipo de multa o sanción.

(6) ¿Alguien más conocerá los resultados?

Con la información disponible se harán estudios cuyos resultados son de carácter agregado; es decir, se conocerán patrones de comportamiento por grupos de individuos, pero no será posible asociar resultados a los nombres de las personas participantes en el estudio (que además no se conocerán, pues no serán guardados). Los resultados, a nivel agregado, serán publicados en documentos de trabajo, en revistas científicas y/o en documentos de divulgación. En la publicación de los resultados, se garantiza el anonimato de los participantes del estudio.

(7) ¿Me beneficiará en algo la participación en este estudio?

No es posible prometer beneficio personal alguno a partir de este estudio, aparte de participar (si lo desea) en 5 rifas de \$100.000 cada una; del punto de vista social, más que individual, esperamos que los resultados del estudio contribuyan a mejorar las condiciones de los desplazamientos en la ciudad a futuro.

(8) ¿Puedo contarle a otras personas sobre el estudio?

Usted puede contarles a otras personas sobre el estudio y sobre su participación en él sin ningún problema.

(9) ¿Qué debo hacer si necesito más información?

Cuando haya leído esta información, la persona a cargo de administrar la entrevista responderá cualquier pregunta que usted pueda tener. Si desea profundizar en algún aspecto del estudio, Ud. puede contactar directamente al Prof. Juan de Dios Ortúzar, al (02) 354-4822 o por correo electrónico a: jos@ing.puc.cl.

(10) ¿Qué pasa si tengo alguna queja o inquietud?

Cualquier persona con inquietudes o quejas sobre el desarrollo de una investigación puede ponerse en contacto con el Dr. Pedro Bouchon, Director de Investigación y Postgrado de la Escuela de Ingeniería, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Santiago, al teléfono (02) 354-4076 o por correo electrónico a la dirección: pbouchon@ing.puc.cl.

Ud. puede conservar esta hoja informativa

Información actividades programadas para su próximo día hábil

ID asignado por el encuestador

Respecto a las actividades que Ud. planea realizar su próximo día hábil, indique:
(Conteste sólo las actividades que tiene planeadas realizar)

Actividad

	Hora de inicio	Hora de fin	Medio de transporte
Trabajar	<input type="text"/>	<input type="text"/>	<input type="text"/>
Ir de compras	<input type="text"/>	<input type="text"/>	<input type="text"/>
Buscar, ver o encontrarse con alguien	<input type="text"/>	<input type="text"/>	<input type="text"/>
Realizar trámites personales	<input type="text"/>	<input type="text"/>	<input type="text"/>
Estudiar	<input type="text"/>	<input type="text"/>	<input type="text"/>
Recreación (ir a cine, espectáculos, hacer deporte, etc)	<input type="text"/>	<input type="text"/>	<input type="text"/>
Relacionadas a la salud (ir al médico, odontología)	<input type="text"/>	<input type="text"/>	<input type="text"/>
Buscar o dejar algo	<input type="text"/>	<input type="text"/>	<input type="text"/>
Otra cosa	<input type="text"/>	<input type="text"/>	<input type="text"/>

Escriba cualquier comentario adicional acerca de sus actividades programadas

33%

Siguiente

Figura B-1: Plan de actividades para el día de la encuesta
Fuente: Elaboración propia

Investigación financiada por el Fondo Nacional de Investigaciones Científicas y Tecnológicas (FONDECYT), su colaboración es muy importante para nosotros. Cualquier inquietud comunicarse con Julián Arellano (jarell@uccl)

Utilice este registro para anotar los datos principales de sus salidas durante su Día de Encuesta. Después entréguelo al encuestador cuando lo visite luego del Día de Encuesta.

En el respaldo encontrará un ejemplo de cómo llenarlo



registro general de salidas

→ Recuerde registrar también sus viajes de vuelta

Partí a las	Fui a	Me fui en	Llegué a las	Dirección
8:36	Trabajar	Metro <small>Nº de Pasajero: Líneas 1 y 2</small>	9:44	Montt con Alfidinate
3:20	Hacer un trámite	Micro <small>Nº de Pasajero: 301</small>	3:41	Santa Rosa con Matta
4:18	Volví a mi trabajo	Micro <small>Nº de Pasajero: 506</small>	4:50	Montt con Alfidinate
6:42	Comprar	Camionando <small>Nº de Pasajero:</small>	6:47	Montt con Prat
7:03	Volví a la casa	Metro <small>Nº de Pasajero: Líneas 1 y 2</small>	7:42	Manquehue 300
		<small>Nº de Pasajero:</small>		

→ Recuerde registrar también sus viajes de vuelta

Partí a las	Fui a	Me fui en	Llegué a las	Dirección
		<small>Nº de Pasajero:</small>		
		<small>Nº de Pasajero:</small>		
		<small>Nº de Pasajero:</small>		
		<small>Nº de Pasajero:</small>		
		<small>Nº de Pasajero:</small>		
		<small>Nº de Pasajero:</small>		

→ Recuerde registrar también sus viajes de vuelta

Partí a las	Fui a	Me fui en	Llegué a las	Dirección
		<small>Nº de Pasajero:</small>		
		<small>Nº de Pasajero:</small>		
		<small>Nº de Pasajero:</small>		
		<small>Nº de Pasajero:</small>		
		<small>Nº de Pasajero:</small>		
		<small>Nº de Pasajero:</small>		

Figura B-2: Registro general de salidas
Fuente: Elaboración propia

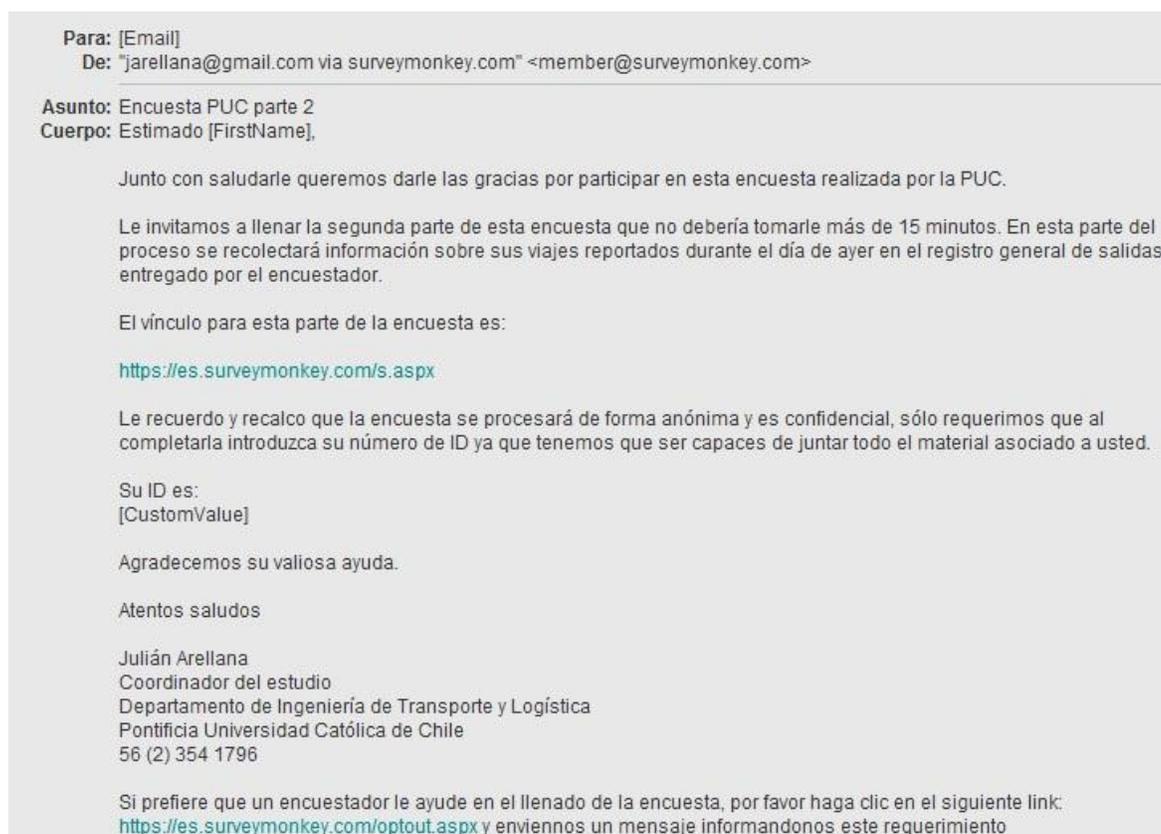


Figura B-3: e-mail de contacto para llenar parte 2 – componente PR
Fuente: Elaboración propia

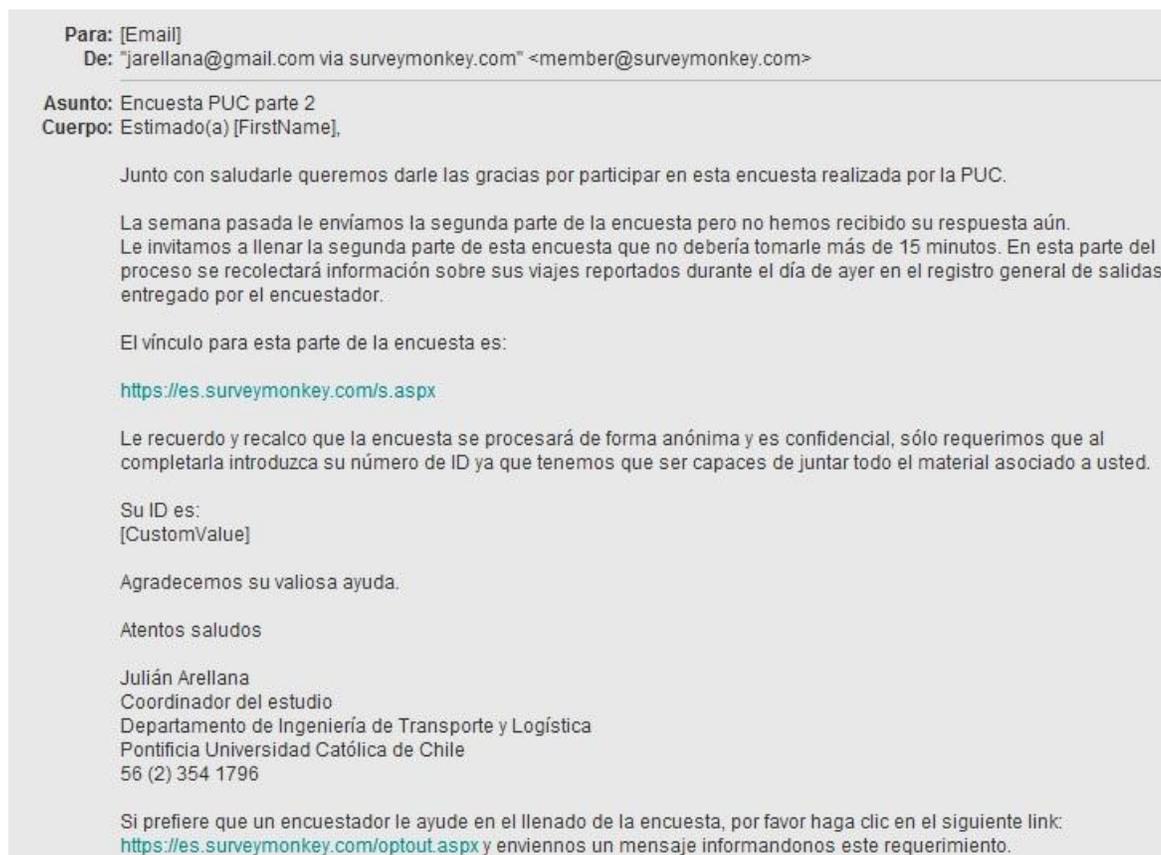


Figura B-4: e-mail recordatorio Semana 1 para llenado de parte 2 PR – componente PR
Fuente: Elaboración propia

Estimado participante:

En mi calidad de coordinador del proyecto “Modelos de Elección del Horario de Inicio de Viajes”, quiero agradecer nuevamente su participación en las dos primeras etapas del proceso de encuestas contempladas en la investigación.

En particular, me permito informarle que luego de procesar las respuestas de todos los participantes en las dos primeras etapas del proyecto, hemos realizado las cinco rifas prometidas de \$100.000 cada una.

Debido a la política de confidencialidad suscrita con cada uno de ustedes, no nos era permitido divulgar el nombre de los ganadores. Sin embargo, luego de contactarlos y pedirles permiso, ellos han aceptado que sus nombres e instituciones de trabajo sean incluidos en el presente comunicado. Los afortunados ganadores son:

ID	Nombre	Institución
1033		Ministerio de Obras Públicas
1141		SODIMAC
2113		Ministerio de Vivienda y Urbanismo
2144		SODIMAC
6063		METRO

Actualmente, nos encontramos próximos a iniciar la tercera etapa del proceso de encuestas, por lo que en el transcurso de las próximas tres semanas deberá estar recibiendo la visita de nuestro grupo de encuestadoras, con una encuesta de preferencias específicamente diseñada para Ud. a partir de la información suministrada en las etapas anteriores.

Como Ud. desgraciadamente no ganó en las cinco primeras rifas realizadas, como una muestra de gratitud para quienes continúan colaborando con el proyecto, hemos decidido realizar **4 nuevas rifas de \$100.000 cada una** entre las personas que respondan la tercera parte de las encuestas.

Su continuidad en esta última etapa es vital para nosotros y sus respuestas imprescindibles para el buen éxito del proyecto.

Les saluda cordialmente,

Julián A. Arellana O.

Julián Arellana O.
 Coordinador del Proyecto “Modelos de elección de horario de inicio de viajes”
 Estudiante de Doctorado, Pontificia Universidad Católica de Chile
 Email: jlarella@uc.cl

Figura B-5: Anuncio de ganadores rifas e inicio parte 3 – componente PD
 Fuente: Elaboración propia

ANEXO C: ESTIMACIONES DE LOS MODELOS CON EFECTO PANEL

Tabla C-1: Constantes de los modelos con efecto panel

Coeficiente	MNL		MNL+panel		MNL+panel+vsg		ML	
	Valor	Test t	Valor	Test t	Valor	Test t	Valor	Test t
Salir antes viaje ida – experimento <i>tour</i>	-0.5915	-3.5	-0.8462	-3.9	-0.7660	-2.9	-0.6658	-2.9
Salir antes viaje regreso – experimento <i>tour</i>	-0.4538	-3.3	-0.7175	-2.8	-0.7295	-2.9	-0.7153	-4.1
Salir antes – experimento viaje	-0.4292	-3.0	-0.9770	-3.2	-1.0209	-2.6	-1.0791	-3.6
Salir después viaje ida – experimento <i>tour</i>	-1.5766	-8.0	-1.7561	-6.7	-1.7899	-6.5	-1.6748	-5.7
Salir después viaje regreso – experimento <i>tour</i>	-1.0998	-6.7	-1.5868	-6.5	-1.5610	-6.4	-1.8237	-7.8
Salir después – experimento viaje	-0.8195	-4.8	-1.7379	-4.3	-1.8300	-4.3	-1.6617	-4.3
Cambio de modo	-1.1118	-8.5	-1.9806	-10.8	-1.9846	-8.2	-3.0413	-9.0

Fuente: Elaboración propia

Tabla C-2: Coeficientes de los modelos para efecto panel

Coeficiente	MNL		MNL+panel		MNL+panel+vsg		ML	
	Valor	Test t	Valor	Test t	Valor	Test t	Valor	Test t
Tiempo de viaje	-0.0209	-4.9	-0.0375	-7.0	-0.0395	-6.9	-0.0341	-4.1
Penalidad por arribo temprano (SDE)	-0.0164	-5.9	-0.0269	-6.6	-0.0103	-0.9	-0.0581	-5.7
Penalidad por arribo tardío (SDL)	-0.0157	-5.0	-0.0266	-6.6	-0.0238	-2.1	-0.0838	-6.9
Penalidad por trabajar menos tiempo (PTD)	-0.0150	-5.6	-0.0181	-5.2	-0.0180	-5.0	-0.0373	-6.6
Penalidad por trabajar mas tiempo (PTI)	-0.0067	-3.5	-0.0074	-2.6	-0.0072	-2.5	-0.0111	-3.8
Incertidumbre en tiempos de viaje	-1.3522	-3.9	-2.0237	-4.0	-2.0192	-3.8	-2.9020	-4.8
Log (costo)	-0.9069	-8.0	-1.544	-8.6	-1.5528	-7.5	-2.4160	-8.0
Ir de pie	-0.0816	-1.1	-0.1915	-1.9	-0.2068	-2.0	-0.4042	-2.8
Ir de pie y a veces no puede abordar	-0.2468	-2.9	-0.5666	-4.1	-0.5793	-4.4	-0.8921	-5.3
Interacciones								
Tiempo de viaje – Penalidad	-	-	-	-	0.0177	1.9	-	-
SDE – Postgrado	-	-	-	-	-0.0137	-1.5	-	-
SDE – Estudia y trabaja	-	-	-	-	-0.0202	-1.7	-	-
SDE – Penalidad	-	-	-	-	0.0097	1.7	-	-
SDL – Postgrado	-	-	-	-	-0.0436	-0.9	-	-
PTD – Postgrado	-	-	-	-	-0.0188	-1.6	-	-
PTI – Postgrado	-	-	-	-	-0.0126	-2.3	-	-
Desviación estándar								
Tiempo de viaje	-	-	-	-	-	-	-0.1086	-10.4
SDE	-	-	-	-	-	-	0.0513	10.6
SDL	-	-	-	-	-	-	0.0764	9.2
PTD	-	-	-	-	-	-	0.0355	7.8
PTI	-	-	-	-	-	-	-0.0259	-9.8
Log (costo)	-	-	-	-	-	-	-2.9701	-7.4
Efecto panel y Factor de escala								
Efecto panel	-	-	2.0556	23.5	2.0441	23.4	-	-
Factor de escala – sobre el experimento de viajes	1.3384	12.1	0.8308	12.9	0.8237	11.8	0.6232	11.5

Fuente: Elaboración propia

Tabla C-3: Constantes de los modelos estimados para comparar efecto panel fijo con efecto panel heterogéneo

Coeficiente	Panel fijo		Panel heterogéneo	
	Valor	Test t	Valor	Test t
Salir antes viaje ida – experimento <i>tour</i>	-0.7551	-3.6	-0.7663	-3.7
Salir antes viaje regreso – experimento <i>tour</i>	-0.6845	-3.0	-0.6890	-3.0
Salir antes – experimento viaje	-0.9915	-3.3	-0.9937	-3.3
Salir después viaje ida – experimento <i>tour</i>	-1.7711	-7.1	-1.7801	-7.1
Salir después viaje regreso – experimento <i>tour</i>	-1.5297	-6.3	-1.5196	-6.2
Salir después – experimento viaje	-1.7382	-5.3	-1.6951	-5.2
Cambio de modo	-1.8918	-9.0	-1.8777	-9.3

Fuente: Elaboración propia

Tabla C-4: Coeficientes de los modelos para comparar efecto panel fijo con efecto panel heterogéneo

Coeficiente	Panel fijo		Panel heterogéneo	
	Valor	Test t	Valor	Test t
Tiempo de viaje	-0.0382	-7.0	-0.0385	-7.2
Penalidad por arribo temprano (SDE)	-0.0156	-2.3	-0.0152	-2.2
Penalidad por arribo tardío (SDL)	-0.0321	-2.6	-0.0293	-2.3
Penalidad por trabajar menos tiempo (PTD)	-0.0395	-1.8	-0.0390	-1.7
Penalidad por trabajar mas tiempo (PTI)	-0.0181	-3.1	-0.0179	-3.1
Incertidumbre en tiempos de viaje	-2.0069	-3.8	-2.0044	-3.9
Log (costo/tasa salarial)	-48.519	-3.8	-48.452	-3.9
Ir de pie	-0.2183	-2.0	-0.2169	-2.0
Ir de pie y a veces no puede abordar	-0.5356	-4.1	-0.5394	-4.1
Interacciones más relevantes				
Tiempo de viaje – Penalidad	0.0171	2.0	0.0173	2.0
SDE – Estudia y trabaja	-0.0154	-2.4	-0.0157	-2.3
SDE – Penalidad	0.0099	2.0	0.0104	2.1
SDE – Postgrado	-0.0110	-1.3	-0.0101	-1.0
SDL – Postgrado	-0.0410	-1.8	-0.0332	-1.9
SDL – Tasa salarial	-0.5028	-1.4	-0.5043	-1.4
PTD – Edad	-0.0006	-2.2	-0.0006	-2.1
PTD – Postgrado	-0.0208	-1.9	-0.0163	-1.2
PTD – Trabaja en casa	0.0365	3.0	0.0372	3.0
PTI – Edad	0.0002	1.8	0.0002	1.8
PTI – Postgrado	-0.0127	-2.2	-0.0117	-1.4
PTI – Trabaja en casa	0.0123	1.6	0.0145	2.3
Efecto panel y Factor de escala				
Efecto panel - Constante	2.0352	22.7	1.7702	5.6
Efecto panel - Edad	-	-	0.0078	1.6
Efecto panel - Estudios de postgrado	-	-	-0.5035	-2.6
Factor de escala – sobre el experimento de viajes	0.8434	15.4	0.8540	15.4

Fuente: Elaboración propia

ANEXO D: ESTIMACIONES DE LOS MODELOS HÍBRIDOS

Tabla D-1: Constantes de los modelos híbridos

Coefficiente	MNL+panel		MNL+panel+vsg		ML	
	Valor	Test t	Valor	Test t	Valor	Test t
Salir antes viaje ida – experimento <i>tour</i>	-1.0484	-4.8	-0.9306	-2.9	-0.5028	-2.1
Salir antes viaje regreso – experimento <i>tour</i>	-0.5543	-3.1	-0.7135	-2.6	-0.8336	-4.5
Salir antes – experimento viaje	-1.2351	-3.9	-1.0678	-3.1	-0.9622	-3.2
Salir después viaje ida – experimento <i>tour</i>	-1.6685	-5.5	-1.6643	-5.0	-1.8229	-6.3
Salir después viaje regreso – experimento <i>tour</i>	-1.5690	-6.4	-1.5722	-5.4	-1.5396	-7.6
Salir después – experimento viaje	-1.8842	-4.5	-1.6755	-2.8	-1.5344	-4.5
Cambio de modo	-2.1562	-7.9	-2.0492	-5.7	-2.5429	-8.1

Fuente: Elaboración propia

Tabla D-2: Coeficientes de la componente de elección de hora - modelos híbridos

Coeficiente	MNL+panel		MNL+panel+vsg		ML	
	Valor	Test t	Valor	Test t	Valor	Test t
Tiempo de viaje	-0.0375	-7.0	-0.0427	-2.5	-0.0413	-6.0
Penalidad por arribo temprano (SDE)	-0.0269	-6.6	-0.0499	-6.2	-0.0647	-7.5
Penalidad por arribo tardío (SDL)	-0.0266	-6.6	-0.0496	-5.8	-0.0691	-8.5
Penalidad por trabajar menos tiempo (PTD)	-0.0181	-5.2	-0.0197	-6.5	-0.0421	-5.6
Penalidad por trabajar mas tiempo (PTI)	-0.0074	-2.6	-0.0061	-2.2	-0.0156	-4.3
Incertidumbre en tiempos de viaje	-2.0237	-4.0	-2.1482	-3.6	-2.7324	-4.9
Log (costo)	-1.5445	-8.6	-1.6368	-7.4	-2.1947	-6.7
Ir de pie	-0.1915	-1.9	-0.2262	-1.3	-0.3322	-2.6
Ir de pie y a veces no puede abordar	-0.5666	-4.1	-0.5272	-2.6	-0.8288	-4.7
Interacciones						
Tiempo de viaje – Penalidad	-	-	0.0191	1.7	-	-
SDE – Postgrado	-	-	-0.0331	-1.3	-	-
SDE – Estudia y trabaja	-	-	0.0188	3.5	-	-
SDE – Penalidad	-	-	0.0055	1.5	-	-
SDL – Postgrado	-	-	0.0055	1.5	-	-
PTD – Postgrado	-	-	0.0188	-1.6	-	-
PTI – Postgrado	-	-	-0.0037	-1.1	-	-
Desviación estándar						
Tiempo de viaje	-	-	-	-	0.0524	5.3
SDE	-	-	-	-	0.0594	6.2
SDL	-	-	-	-	-0.0678	-11.6
PTD	-	-	-	-	-0.0374	-7.6
PTI	-	-	-	-	0.0271	11.0
Log (costo)	-	-	-	-	2.4181	6.5
Efecto panel y Factor de escala						
Efecto panel	2.0556	23.5	2.0441	23.4	-	-
Factor de escala – sobre el experimento de viajes	2.0556	23.5	1.8537	12.5	-	-

Fuente: Elaboración propia