



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA

**MODELOS CONJUNTOS DE ELECCIÓN
DE MODO Y HORA DE INICIO DEL
VIAJE CON DATOS COMBINADOS DE
PREFERENCIAS REVELADAS Y
DECLARADAS**

PEDRO LIZANA BESA

Tesis para optar al grado de
Magíster en Ciencias de la Ingeniería

Profesor Supervisor:
JUAN DE DIOS ORTÚZAR SALAS

Santiago de Chile, Enero, 2013

© 2013, Pedro Lizana



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA

**MODELOS CONJUNTOS DE ELECCIÓN
DE MODO Y HORA DE INICIO DEL
VIAJE CON DATOS COMBINADOS DE
PREFERENCIAS REVELADAS Y
DECLARADAS**

PEDRO LIZANA BESA

Tesis presentada a la Comisión integrada por los profesores:

JUAN DE DIOS ORTÚZAR SALAS

LUIS IGNACIO RIZZI CAMPANELLA

MARCELA MUNIZAGA MUÑOZ

JUAN DIXON ROJAS

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

Santiago de Chile, Enero, 2013

A mi familia, amigos y en especial a
Mimi y a mi perro por su apoyo
constante e incondicional.

AGRADECIMIENTOS

Quisiera comenzar agradeciendo a mi profesor supervisor, Juan de Dios, y a Julián Arellana. Ambos fueron mis guías en los momentos de incertidumbre, apoyándome constantemente y resolviendo todas mis inquietudes. No sólo en lo académico, sino también en lo personal, ya que me apoyaron al ciento por ciento en mi decisión de trabajar a la distancia y también en los momentos difíciles.

Me gustaría también agradecer al profesor Luis Rizzi por sus valiosos comentarios y apoyo durante mis últimos años en la universidad. Al profesor Hector Jorquera y a Luz María Velasco de Metro de Santiago S.A., por su voluntad a colaborar y suministrar información importante para el desarrollo de este trabajo. Además, a todos los funcionarios del Departamento de Ingeniería de Transporte y Logística, ya que de alguna u otra forma contribuyeron a que mi estadía en la universidad fuese una experiencia increíble. A todos los amigos que conocí estando en el Departamento, su amistad perdurará para siempre.

A mi familia y amigos de siempre por estar presentes todo el tiempo y ser un pilar fundamental en mi vida. Estoy inmensamente agradecido por todo su apoyo y motivación, en especial en la fase final de la tesis, por ayudarme a sobrepasar los problemas personales que surgieron repentinamente. A la Sra. María y su familia, que me acogieron en Curanipe, y definitivamente se convirtieron en una segunda familia para mí. A pesar de que procedemos de entornos culturales muy diferentes, me han enseñado muchas cosas que estando Santiago nunca hubiese aprendido, como por ejemplo, que la felicidad reside en las cosas simples.

A Mimi, que fue mi compañera de vida en los últimos tres años y medio. A pesar de que tristemente nuestra relación no prosiguió como esperábamos, estaré eternamente agradecido por los increíbles momentos juntos y por toda la felicidad entregada. Pero por sobre todo, por su apoyo incondicional, por haber dejado toda una vida de lado para embarcarse en una aventura, y esperar pacientemente a que yo terminará mis estudios. Y también por motivarme a hacer este Magíster. También agradecer a su familia, ya que sin ellos no hubiese sido posible vivir todas estas experiencias juntos. Y por último, a Ever, el perro que conocí estando en Curanipe, por demostrarme que ellos son los compañeros más fieles de los seres humanos y que no hace falta hablar para expresar sentimientos.

ÍNDICE GENERAL

	Pág.
DEDICATORIA.....	ii
AGRADECIMIENTOS.....	iii
RESUMEN.....	xiii
ABSTRACT.....	xiv
1 INTRODUCCIÓN	1
1.1 Motivación	1
1.2 Hipótesis, objetivos y alcances.....	3
1.3 Principales contribuciones y estructura del contenido	5
2 MARCO TEÓRICO.....	8
2.1 Antecedentes generales	8
2.1.1 <i>Representación del tiempo: continuo versus discreto</i>	13
2.1.2 <i>Tipos de datos utilizados para modelar EHIV</i>	19
2.1.3 <i>Estructura de modelos de elección discreta aplicados a EHIV</i>	20
2.1.4 <i>Modelación EHIV junto a otras decisiones</i>	23
2.2 Variabilidad del tiempo de viaje en contexto de EHIV	25
2.3 Modelos de EHIV en ausencia de demoras horarias (SD).....	27
2.4 Incorporación de variables climáticas	29
2.5 Principales conclusiones y limitaciones encontradas.....	30
3 GENERACIÓN DEL BANCO DE DATOS	32
3.1 Preferencias reveladas (PR)	33
3.1.1 <i>Niveles de servicio y demoras horarias</i>	34
3.1.2 <i>Generación de alternativas de modo y hora</i>	39
3.2 Preferencias declaradas (PD)	42
4 ANÁLISIS DEL BANCO DE DATOS	45
4.1 Información socioeconómica y de empleo.....	45
4.2 Banco de datos de preferencias reveladas (PR)	49
4.2.1 <i>Modo de transporte</i>	50
4.2.2 <i>Hora de inicio del viaje</i>	52
4.2.3 <i>Niveles de servicio y demoras horarias de elecciones</i>	54
4.3 Banco de datos de preferencias declaradas (PD)	57

5	MODELOS DE ELECCIÓN DE MODO Y HORA DE INICIO DEL VIAJE: BASE 15 MINUTOS.....	60
5.1	Análisis de resultados modelos base PR	60
5.1.1	<i>Flexibilidad horaria según disponibilidad de alternativas</i>	60
5.1.2	<i>Flexibilidad horaria como variable explicativa</i>	71
5.2	Análisis de resultados modelos base PD	74
5.2.1	<i>Modelos Logit Multinomial (MNL)</i>	75
5.2.2	<i>Modelos Logit Mixto (ML)</i>	81
5.3	Análisis de resultados modelos PR/PD	84
5.3.1	<i>Flexibilidad laboral según disponibilidad de alternativas</i>	84
5.3.2	<i>Flexibilidad laboral como variable explicativa</i>	88
6	MODELOS DE ELECCIÓN DE MODO Y HORA DE INICIO DEL VIAJE: BASE 30 y 60 MINUTOS.....	91
6.1	Resultados modelos mixtos PR/PD base 30 y 60 min	91
6.1.1	<i>Base 30 min</i>	91
6.1.2	<i>Base 60 min</i>	93
6.2	Comparación con modelos mixtos de PR/PD base 15 min.....	96
7	MODELOS EN FASE PREDICTIVA.....	102
7.1	Predicción de elecciones horarias observadas.....	102
7.2	Modelos sin variables de restricción de programación de actividades	105
7.3	Predicción de escenario hipotético: tarificación vial.....	106
8	CONCLUSIONES	114
8.1	Principales conclusiones	114
8.1.1	<i>Conclusiones metodológicas</i>	114
8.1.2	<i>Conclusiones de política pública</i>	117
8.2	Recomendaciones para futuras investigaciones	118
	BIBLIOGRAFIA.....	120
	A N E X O S.....	125
	ANEXO A: MODELOS DE ELECCION DISCRETA.....	126
	ANEXO B: RESULTADOS MODELOS BASE PR_15.....	131
	ANEXO C: RESULTADOS MODELOS BASE PD	132
	ANEXO D: RESULTADOS MODELOS PR/PD_15	134

ÍNDICE DE TABLAS

	Pág.
Tabla 3-1: Factores de corrección para tiempos de espera de modo bus	37
Tabla 4-1: Elección modal versus Disponibilidad modal	51
Tabla 4-2: Distribución de horas de inicio del viaje según tipo de desagregación	53
Tabla 4-3: Distribución de la duración de los viajes	54
Tabla 4-4: Costo total (operacional + peaje + estacionamiento) según tipo de modo	57
Tabla 4-5: Elección y niveles de servicio promedio base PD	58
Tabla 5-1: Resultados estimación modelos Modo-Hora MNL-1 base PR_15_1	63
Tabla 5-2: Resultados estimación modelos MNL-2 base PR_15_1	66
Tabla 5-3: Resultados estimación modelos MNL-3 base PR_15_1	67
Tabla 5-4: Valor subjetivo del tiempo y razones de compromiso modelos base PR_15_1	68
Tabla 5-5: Resultados estimación modelos MNL-1 base PR_15_2	72
Tabla 5-6: Valores subjetivos del tiempo y razones de compromiso modelo 4 base PR_15_2	75
Tabla 5-7: Resultados estimación modelos MNL base PD completa	77
Tabla 5-8: Resultados estimación modelos MNL banco de datos de PD depurado	79
Tabla 5-9: Resultados estimación modelos Logit Mixto (ML) base PD depurada	83
Tabla 5-10: Análisis de parámetros comunes PR/PD base PR_15_1	85
Tabla 5-11: Resultados modelos base PR/PD_15_1	86
Tabla 5-12: Análisis parámetros comunes PR/PD base PR_15_2	89
Tabla 5-13: Resultados modelos mixtos PR/PD base PR_15_2	90
Tabla 6-1: Resultados modelos mixtos PR/PD MNL base 30 minutos	92
Tabla 6-2: Resultados modelos mixtos PR/PD MNL y HL base 60 minutos	94
Tabla 6-3: Parámetros estructurales modelo 60_HL	96
Tabla 6-4: Resumen comparativo modelos mixtos PR/PD bases de 15, 30 y 60 minutos..	98
Tabla 7-1: Modelos mixtos de PR/PD en fase predictiva bases de 15, 30 y 60 min	107
Tabla B-1: Constantes específicas de modo (ASC_Modo) de modelos base PR_15_1 ...	131
Tabla C-1: Resultados estimación modelo MNL con variables de flexibilidad horaria base PD	132

Tabla C-2: Resultados estimación mejor modelo Logit Mixto (ML)	133
Tabla D-1: Resultados estimación modelos MNL-2 base PR/PD_15_1.....	134
Tabla D-2: Test LR modelos MNL PR/PD base PR/PD_15_1.....	135
Tabla D-3: Valores del tiempo y razones de compromiso modelos PR/PD base PR _15_1	136

ÍNDICE DE FIGURAS

	Pág.
Figura 3-1: Ubicación de los hogares (círculo) y trabajo (triángulo) de los encuestados ...	33
Figura 3-2: Densidad (pax/m ²) en línea 5 de metro dirección sur-norte	40
Figura 3-3: Ejemplo de situación de elección de primer experimento PD presentado a usuario de transporte público	43
Figura 4-1: Distribución porcentual del ingreso líquido mensual (ILM).....	47
Figura 4-2: Distribución de hora oficial de entrada al trabajo (<i>HOE</i>) y hora preferida de llegada al trabajo (<i>PAT</i>).....	48
Figura 4-3: Distribución de la flexibilidad laboral para llegar antes o después de la hora oficial de entrada al trabajo (<i>HOE</i>)	49
Figura 4-4: Distribución de las elecciones modales.....	51
Figura 4-5: Distribución de las horas de inicio y de llegada del viaje al trabajo	53
Figura 4-6: Histograma de disponibilidad de horarios de llegada al trabajo	54
Figura 4-7: Promedio y desviación estándar del tiempo de viaje según modo de transporte	55
Figura 4-8: Tiempo de caminata de acceso/egreso, tiempo de espera de acceso y tiempo de transbordo promedio según modo de transporte	56
Figura 6-1: Estructura jerárquica modelo 2_HL base PR/PD de 60 min.....	95
Figura 7-1: Elecciones de hora modeladas versus observadas modelo PR/PD base 15 minutos.....	103
Figura 7-2: Elecciones de hora modeladas versus observadas modelo PR/PD base 30 minutos.....	104
Figura 7-3: Elecciones de hora modeladas versus observadas modelos PR/PD base 60 minutos.....	104
Figura 7-4: Distribución horaria de demanda escenario tarificación vial modelos 15 min	108
Figura 7-5: Cambio porcentual partición modal escenario tarificación vial modelos 15 min	109

Figura 7-6: Distribución horaria de demanda escenario tarificación vial modelos 30 min	
.....	111
Figura 7-7: Distribución horaria de demanda escenario tarificación vial modelos 60 min	
.....	112

LISTADO DE ACRÓNIMOS

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

EHIV: Elección de hora de inicio del viaje

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*

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

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*

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

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

ACH: Auto-chofer

AAC: Auto-acompañante

CAM: Caminata

B: Bus

M: Metro

BM: Bus-metro

TX: Taxi

TXC: Taxi-colectivo

CombM: Combinación de metro con transporte privado

RESUMEN

Los factores que influyen la elección de hora de inicio del viaje son cada vez más relevantes dado el aumento en la aplicación de estrategias de gestión de demanda para mitigar la congestión. Extendiendo el trabajo de Arellana (2012), se estimaron modelos conjuntos de elección de modo y hora de inicio del viaje al trabajo combinando datos de preferencias reveladas (PR) y preferencias declaradas (PD). A partir de una encuesta tipo diario de viajes realizada a más de 500 trabajadores del Gran Santiago, se generó un banco de datos de PR consistente de nueve modos y hasta 11 períodos de tiempo.

Entre los resultados obtenidos, se destaca que los parámetros de las variables de tiempo de viaje, costo y costo dividido por la tasa salarial tomaron valores muy similares entre los modelos de PR y PD, resultando en modelos mixtos PR/PD superiores a los estimados en forma separada. Esto contrasta con los parámetros asociados a la demora por llegar temprano o tarde al trabajo (*SDE* y *SDL*), cuya diferencia se atribuye al contexto temporal de ambas elecciones (PR: decisión de largo plazo; PD: corto plazo) y al del experimento de PD (implementación de una política de tarificación vial y sistema flexible de horarios de entrada y salida al trabajo). Por otro lado, se encontró que la mejor forma de incluir la flexibilidad laboral en los modelos de elección de hora era a través de variables dicotómicas interactuando con los parámetros de *SDE* y *SDL*. También se destaca que a medida que disminuye la resolución temporal de las alternativas de hora en PR (al pasar de intervalos de tiempo de 15 min a 30 y 60 min) aumenta la bondad de ajuste del modelo, en parte por la menor correlación entre alternativas.

Finalmente, se comprobó que los modelos estimados podían ser utilizados para predecir variaciones temporales en la demanda por transporte y que el incremento en la correlación, al desagregar los períodos de tiempo, podía ser compensado por predicciones más precisas y desagregadas.

Palabras Clave: elección conjunta de hora de inicio del viaje y modo, modelos de elección discreta, modelación mixta PR/PD.

ABSTRACT

The factors influencing trip departure time are taking more importance in practice since urban congestion is increasingly being addressed by travel demand management (TDM) strategies. Extending the work done by Arellana (2012), we formulate and estimate a joint travel mode-departure time model for commuting trips, combining revealed preference (RP) and stated preference (SP) data. The data was gathered through a RP/SP/attitudinal survey applied to nearly 500 workers of the Santiago Metropolitan Area. The RP data consists of nine alternative modes and up to 11 time periods and the level-of-service data were obtained at unusual precision levels using GPS measurements.

Some relevant results are that the travel time, cost and cost divided by the wage rate coefficients were similar in both the RP and SP environments, yielding joint RP-SP models which are superior to those estimated with each data separately. The only parameters that differed between each data set were those associated with the schedule delay early (*SDE*) and late (*SDL*) variables from Small's scheduling model (Small, 1982). This was attributed to the different temporal perspectives between the RP choice (long term decision) and the SP choice (short term decision), and also due to the context presented in the SP experiment (implementation of a congestion charging policy and a flexible working hours scheme). On the other hand, we found that the best way to introduce scheduling flexibility into the models was by interacting the schedule delay terms (*SDE* and *SDL*) with dummy variables for the scheduling flexibility of each individual. It was also shown that when using time periods with a higher time resolution (15 min instead of 30 and 60 min time intervals), the model goodness of fit decreases, partly because of the higher correlation between alternatives.

Finally, it was shown that the estimated models can be used to explain and predict time of travel preferences, and that when using them for forecasting purposes the increase in correlation associated with using thinner time intervals can be offset by the more accurate and disaggregate predictions obtained.

Keywords: time-of-day models, joint trip departure mode choice models, discrete choice models, mixed RP/SP modelling.

1 INTRODUCCIÓN

1.1 Motivación

Los crecientes niveles de congestión experimentados en zonas urbanas a raíz de incrementos en la población y tasa de motorización, generan cuantiosas pérdidas de tiempo e inconvenientes para los usuarios del sistema de transporte; esto es particularmente cierto cuando los viajes se realizan en período punta. Los usuarios pueden responder a incrementos en los tiempos y costos de viaje cambiando su modo, ruta, y/o la hora de inicio de su viaje hacia períodos menos congestionados.

El fenómeno de redistribuir temporalmente los flujos a medida que las condiciones de tráfico se deterioran, se conoce como esparcimiento del horario punta (*peak spreading*) y es la segunda estrategia más común detrás de los cambios de ruta (Ortúzar y Willumsen, 2011) y bastante por delante de los cambios de modo (Hess *et al*, 2007). Esto se debe a que los niveles de tráfico vehicular y las velocidades varían considerablemente a lo largo de un día o inclusive en un mismo período del día, por lo que cambiar el horario de inicio de viaje puede ser una solución bastante atractiva, en especial cuando todas las rutas alternativas poseen niveles de congestión similares.

Para mitigar la congestión es cada vez más común la aplicación de estrategias de gestión de demanda de transporte (TDM por sus siglas en inglés, *Transportation Demand Management*). Ejemplos de estrategias TDM son la tarificación vial, horarios de trabajo flexibles, restricción vehicular y teletrabajo, o trabajo desde el hogar (Cambridge Systematics, 1999). En el caso particular de Santiago de Chile, la congestión y la contaminación atmosférica son dos grandes problemas aún no resueltos. Se ha intentado aminorar la contaminación atmosférica mediante una política de restricción vehicular según número de patente, pero su efectividad para reducir el tráfico vehicular es francamente dudosa (Fresard, 1998; de Grange y Troncoso, 2010; Cantillo y Ortúzar, 2012). En la actualidad, las autoridades están considerando implementar un sistema de tarificación vial en el centro de Santiago y en la zona conocida como “*Sanhattan*”, pues se prevé que la apertura del complejo de oficinas del proyecto

inmobiliario *Costanera Center*¹ pueda aumentar los niveles de congestión a niveles sin precedentes en una zona que ya se encuentra prácticamente saturada de tráfico vehicular a ciertas horas del día. Sin embargo, la respuesta de los usuarios del sistema de transporte ante políticas de este tipo no puede ser predicha de forma precisa por los modelos clásicos de planificación de transporte, pues en su mayoría utilizan factores fijos, generalmente derivados de Encuestas Origen Destino (EOD), para incorporar la dimensión temporal (Bhat, 1998a; Cambridge Systematics, 1999; Vovsha y Bradley, 2004).

Posiblemente la elección de la hora de iniciar un viaje no se realiza de forma aislada, sino que en conjunto con otras decisiones, por ejemplo, la elección de cuánto tiempo trabajar, la programación diaria de actividades, el modo de transporte y la ruta a tomar. No obstante, en el contexto de evaluar estrategias TDM u otras políticas (como aumentar la tarifa de transbordo hacia el Metro de Santiago desde los buses a fin de disminuir la demanda del Metro en hora punta y usar mejor la capacidad “ociosa” de estos), se estima que las elecciones más entrelazadas son las de modo y hora de inicio del viaje. Esto puede ser aún más cierto en una ciudad como Santiago, en la que compiten 12 o más modos, ya que es posible que el grado de sustitución entre alternativas de modo y la hora de inicio del viaje sea mayor; esto se debe a que, para un determinado horario de salida, los usuarios no están cautivos a unas pocas alternativas de modo. Asimismo, un modelo conjunto de elección de hora de inicio de viaje (EHIV) y modo, puede ser de gran utilidad para modelar con mayor precisión la distribución temporal y modal de emisiones móviles de contaminantes atmosféricos, ya que los modelos de calidad del aire usan esta información como dato de entrada y son sensibles a variables meteorológicas que varían considerablemente a lo largo del día (Steed y Bhat, 2000).

De esta manera, la motivación más importante de esta investigación es desarrollar modelos conjuntos de EHIV-modo basados en teorías apropiadas de comportamiento de los usuarios, que aporten a la comprensión de los principales factores que determinan

¹ Proyecto inmobiliario que comprende un centro comercial recientemente inaugurado, cuatro edificios de oficinas, y cerca de 3.000 estacionamientos.

sus elecciones temporales y modales, contribuyan al estado del arte y de la práctica en modelación de demanda de transporte, al combinar datos de preferencias reveladas (PR) y preferencias declaradas (PD), y que tengan el potencial de transformarse en una herramienta que facilite la toma de decisiones en el sector transporte.

1.2 Hipótesis, objetivos y alcances

Las hipótesis que busca dilucidar esta investigación son las siguientes:

- Un modelo de elección discreta conjunto de hora de inicio y modo de viaje, combinando datos de PR y PD, es superior a modelos estimados con cada tipo de dato en forma separada.
- Tal modelo puede ser una herramienta útil para predecir variaciones temporales en las decisiones de los usuarios del sistema de transporte.

En línea con las hipótesis planteadas, el objetivo general de esta investigación es desarrollar modelos conjuntos de elección de modo y EHIV para los viajes al trabajo, que sean estimados combinando datos de PR y PD y que se sustenten en la teoría de la utilidad aleatoria (Ortúzar y Willumsen, 2011).

Por ello, el primer objetivo específico de esta investigación es comparar modelos según tipo de datos empleados (PR y PR/PD) y estructura a utilizar, vale decir, Logit Multinomial (MNL), Logit Jerárquico (HL) y Logit Mixto (ML). Por un lado, al variar los tipos de datos empleados en la modelación es posible verificar si combinándolos se atenúan las desventajas de cada tipo de dato en particular (PR: alta correlación entre atributos; PD: menos confiable para hacer predicciones). Por otra parte, usar modelos con distinta estructura debiera permitir aclarar ciertos aspectos particulares de la modelación modo-EHIV, como por ejemplo, cual de ambas, modo o EHIV, va en la raíz de la estructura jerárquica del HL. Por otro lado, mediante el uso del ML se debiera poder detectar la presencia de *efecto panel* en las respuestas de PD, variaciones aleatorias en los gustos y/o si estructuras de correlación más flexibles son más adecuadas.

Estudiar y comprender los principales factores que influyen la decisión de escoger un modo de transporte y el horario de salida al trabajo es otro de los objetivos de

esta tesis. Para ello se busca incluir la mayor cantidad de variables explicativas en la modelación, algunas más innovadoras que otras, como la temperatura al interior del vehículo y la densidad de pasajeros en el metro. Además, se estimará y analizará la valoración de ciertos atributos, como valores del tiempo (caminata, espera, viaje y transbordo) u otros (llegar tarde o temprano al trabajo) por parte de la población.

Por último, se pretende analizar el rol de las constantes específicas (ASC) y de la longitud de los intervalos de tiempo considerados cuando el modelo se usa para predecir. Las ASC son fundamentales para utilizar el modelo en fase predictiva, pues permiten capturar el efecto de variables que generalmente no están disponibles en las grandes EOD de transporte.

Como ya se mencionó, el proceso de elección de los usuarios de un sistema de transporte es complejo, pues parece más razonable pensar que los individuos escogen de forma conjunta decisiones de programación de actividades (cuáles realizar en un día, por cuánto tiempo y en qué secuencia) y de localización (dónde realizarlas), en conjunto con la elección de modo y hora de inicio de los viajes. En contraste con los modelos clásicos de demanda de transporte que analizan sólo viajes, los modelos *basados en actividades* utilizan como unidad de análisis al *tour* (secuencia de viajes y actividades), pues así logran capturar de forma más adecuada la interrelación entre las decisiones comentadas (Vovsha y Bradley, 2004; Roorda *et al.*, 2008). La frecuencia y necesidad de realizar una actividad define un *tour* y la frecuencia con que este se realiza; en general los *tours* están basados en el hogar o trabajo (i.e. los viajes del *tour* suelen tener origen o destino en el hogar o trabajo) y se clasifican según la importancia relativa de la actividad (primario o secundario). En este contexto, las paradas intermedias (que no corresponden a la actividad principal) de un *tour* también se toman en consideración (Popuri *et al.*, 2008).

A pesar de que ambos bancos de datos (PR y PD) pueden usarse para modelar la elección de modo-EHIV en un contexto de *tours*, en esta tesis se considerará como unidad de análisis a los viajes al trabajo desde el hogar. Esto se debe a que si se considera el viaje de regreso del trabajo, por ejemplo, aumenta considerablemente la complejidad del modelo, ya que el número de alternativas crece fuertemente en un

escenario en que la cantidad de alternativas ya es alta (con nueve modos y 11 intervalos de tiempo se tienen 99 alternativas modo-hora para el viaje al trabajo). Por otro lado, aún no se implementa en Santiago un modelo de transporte basado en actividades para ser utilizado en la práctica y esto parece estar lejano de ocurrir. Así, el único elemento que se rescata de los modelos basados en actividades en esta investigación, es que la cantidad de paradas intermedias en el viaje al trabajo se usa como variable explicativa del modelo, de forma similar a lo propuesto por Habib (2011; 2012).

1.3 Principales contribuciones y estructura del contenido

Esta investigación busca contribuir en varios aspectos al estado del arte y de la práctica de modelación de preferencias temporales de los viajeros. En primer lugar, sólo se encontraron dos estudios en la literatura que combinaban datos de PR y PD en el contexto de elección de horario de viaje. En el primero, Borjesson (2008) estimó modelos conjuntos de modo y EHIV, pero consideró la elección de modo solamente en los datos de PD. En el segundo, Tseng *et al.* (2011) modelan solamente EHIV y no consideran la elección de modo. De esta manera, este sería el primer trabajo en que se modela simultáneamente EHIV y modo, con ambas elecciones presentes en cada tipo de datos. Considerar la elección de modo en ambos tipos de datos es fundamental en el contexto de elecciones temporales, ya que los datos de PD permiten evaluar el efecto en el modo escogido de políticas hipotéticas, como la tarificación vial en Santiago, mientras que los datos de PR permiten replicar las elecciones observadas en la realidad y contienen información acerca de todos los modos disponibles (en general las encuestas PD, en el contexto de EHIV, sólo presentan una alternativa de cambio de modo).

Por otra parte, sólo se encontró un estudio que comparaba modelos con alternativas de tiempo de distinta longitud (Hess *et al.*, 2007). Estudiar este fenómeno es de vital importancia si se desea emplear modelos de elección discreta para EHIV, pues uno de los inconvenientes atribuidos a estos modelos en este sentido es que la “discretización” del tiempo queda a criterio del modelador (Bhat y Steed, 2002; Habib *et al.*, 2009). La tesis busca aportar resultados empíricos que permitan facilitar y normalizar esta decisión, al estimar modelos de EHIV con intervalos de 15, 30 y 60 min

para los datos de PR, en que los niveles de servicio corresponderán al promedio para cada longitud de intervalo; cabe notar que Hess *et al.*, 2007, en cambio, usan datos de PD y, por ende, los niveles de servicio se mantienen iguales independiente de la longitud del intervalo. La tesis también busca analizar el efecto de variables que no han sido incluidas anteriormente en estos modelos. Estas son la temperatura, que en Santiago sobrepasa comúnmente los 30°C en verano (fecha cuando se realizó la encuesta) y la densidad del Metro, que pueden producir incomodidad al trasladarse.

Finalmente, en la última década se ha estudiado la forma funcional de los modelos EHIV para su uso en fase predictiva, enfatizando el rol que cumplen las constantes específicas (Hess *et al.*, 2005; Hess *et al.*, 2007; Ben-Akiva y Abou-Zeid, 2012). Esta tesis también busca aportar al estado de la práctica en modelación de transporte al estimar modelos de EHIV-modo para ser usados en fase predictiva, comparando los resultados predichos para un escenario hipotético cuando se varía la longitud del intervalo y las variables explicativas del modelo.

El resto del documento se organiza de la siguiente manera: en el capítulo 2 se presenta el marco teórico relevante para esta investigación, describiendo los conceptos en que se fundan los modelos de EHIV y discutiendo aspectos metodológicos relativos a la formulación de estos modelos, como el tipo de datos y estructura de modelo a usar, la representación del tiempo (continuo o discreto), las variables que influyen esta elección, y el rol de las ASC, entre otros. Luego, en el capítulo 3 se presenta brevemente la metodología utilizada para recolectar los datos y el posterior proceso de construcción de los bancos de datos utilizados, con énfasis en los datos de PR. En el capítulo 4 se analizan ambos bancos de datos, específicamente en cuanto a características socioeconómicas, y elecciones modales y de hora inicio de viaje de los encuestados. Posteriormente, en el capítulo 5 se presentan los resultados y el análisis de la estimación de los modelos usando intervalos de tiempo de 15 min; estos modelos varían según tipo de dato empleado (PR y PR/PD) y estructura del modelo (MNL, HL, ML). En el capítulo 6 se presentan los resultados de los mejores modelos PR/PD al usar intervalos de tiempo de 30 y 60 min, y, haciendo uso de los resultados de las secciones anteriores, en el capítulo 7 se estiman modelos a ser usados en fase predictiva, que son puestos a

prueba al comparar sus predicciones ante un simple escenario hipotético. Finalmente, en el capítulo 8 se presentan las conclusiones de la tesis y las recomendaciones para futuras investigaciones en esta área.

2 MARCO TEÓRICO

A continuación se exponen los principales elementos teóricos y conceptuales que enmarcan esta investigación. En particular, se describe y discute: (i) los factores que en general han incidido en la elección horaria, (ii) la representación del tiempo como continuo o discreto, (iii) los tipos de datos utilizados (PR, PD o PR/PD), las estructuras discretas empleadas y la modelación de EHIV junto a otras decisiones, (iv) la inclusión de la variabilidad del tiempo de viaje en un contexto de preferencias temporales, (v) los modelos de EHIV en fase predictiva y por último, (vi) la incorporación de variables climáticas en modelos de comportamiento respecto a viajes.

2.1 Antecedentes generales

En la literatura sobre modelos de demanda de transporte, la elección de la hora de inicio de viaje ha sido investigada muy poco en comparación a otras decisiones, como la elección de modo (Bhat, 1998a). El resultado de esto es que, por mucho tiempo, los modelos clásicos (de cuatro etapas) en la práctica no han tenido (muchos aún no la tienen) una componente explícita que permita modelar adecuadamente las elecciones temporales. El enfoque práctico predominante ha consistido en usar factores fijos para distribuir la demanda diaria en distintos períodos del día. Estos factores, que equivalen a la razón entre los viajes realizados en una hora o período (punta mañana, punta tarde, fuera de punta) y el total diario de viajes, son exógenos al modelo de transporte clásico e independientes de la congestión experimentada en la red vial; también se pueden aplicar en distintas etapas del modelo clásico de transporte, vale decir, entre generación y distribución, distribución y partición modal, partición modal y asignación, o post asignación. Si se aplican previamente a la etapa de asignación, en general se obtienen de las EOD diferenciando por propósito de viaje; mientras que si son aplicados post asignación, se pueden derivar de conteos de tráfico (Cambridge Systematics, 1999). Los principales problemas de esta metodología son que: (i) no es posible evaluar estrategias TDM que tengan impactos en las elecciones temporales, (ii) se asume que la elección de hora de inicio de viaje es independiente de las características socioeconómicas, lo que es un supuesto muy fuerte y poco probable que se de en la realidad, y (iii) no permite considerar

los potenciales cambios en la hora de inicio de los viajes al ocurrir variaciones no uniformes (a lo largo del día) en los niveles de servicio entre los años de estimación y predicción (Bhat, 1998a).

Otras metodologías prácticas (esto es, que operan dentro del marco convencional de modelación de cuatro etapas) más innovadoras, que surgieron en su momento, son los llamados modelos de extensión del horario punta (*peak spreading*). Estos modelos permiten abordar el problema que la demanda proyectada exceda la capacidad de ciertos arcos durante las horas punta, distribuyendo el exceso de demanda a los *shoulders* (o horarios adjuntos) de la hora punta (Cambridge Systematics, 1999). La principal limitación de estos modelos es que se concentran en los efectos y no en el proceso de decisión de los usuarios (Rand Europe, 2004), y por ende, al igual que la práctica de usar factores fijos, no son aptos para evaluar estrategias TDM (Cambridge Systematics, 1999).

La brecha entre el estado del arte y de la práctica se ha reducido en los últimos años debido a que el fenómeno ha sido investigado en mayor profundidad haciendo referencia - directa o indirectamente - a los aportes originales de Vickrey (1969) y Small (1982). Vickrey (1969) fue el primero en explorar las preferencias temporales de los usuarios del sistema de transporte al desarrollar un modelo conceptual de EHIV que consistía de un arco o cuello de botella (oferta), en el cual los viajeros al trabajo decidían el horario en que viajar (demanda) y el equilibrio oferta-demanda se determinaba explícitamente. El horario de viaje escogido puede ser distinto a la hora oficial de inicio del trabajo, por el deseo de viajar a horarios menos congestionados, y la congestión se generaba siempre que la demanda fuera mayor que la capacidad del cuello de botella.

Por su parte, Small (1982) transforma el problema continuo de escoger la hora de inicio de viaje en uno discreto y propone el *Scheduling Model* (SM), formulación que es la más conocida y sobre la cual se basan la mayoría de los estudios del área. El compromiso (*trade-off*) básico del SM de Small (y también del modelo de Vickrey) es entre la desutilidad de llegar al trabajo antes o después de un horario preferido y los mayores costos en términos de viajar a esa hora. Un viajero que llega al trabajo exactamente a su horario preferido no tiene desutilidad por concepto de demora o

penalidad horaria (SD , *Schedule delay*), pero puede experimentar un tiempo de viaje considerablemente mayor. Por otro lado, llegar antes de lo deseado genera desutilidad porque implica partir antes de lo deseado del hogar o tiempo perdido (o usado menos productivamente), y como llegar tarde al trabajo puede tener repercusiones más serias, se espera una penalidad mayor que por llegar temprano. De esta forma, el SM se resume en la siguiente función de utilidad (Small, 1982):

$$V_i = \beta_{TV} TV_i + \beta_{SDE} SDE_i + \beta_{SDL} SDL_i + \gamma D_L \quad (2.1)$$

donde

$$SDE_i = \text{Max}(-SD_i, 0) \quad (2.2)$$

$$SDL_i = \text{Max}(0, SD_i) \quad (2.3)$$

$$SD_i = DT_i + TV_i - PAT \quad (2.4)$$

$$D_L = \begin{cases} 1 & \text{si } SDL_i > 0 \\ 0 & \text{si } SDL_i = 0 \end{cases} \quad (2.5)$$

en que V_i es la utilidad representativa de la alternativa de hora de inicio de viaje i ; TV_i es el tiempo de viaje al salir del origen en la alternativa de hora i ; SD_i (*Schedule delay*) indica la demora horaria en unidad de tiempo; DT_i (*Departure time*) la hora de salida del lugar de origen (normalmente el hogar) y PAT la hora preferida de llegada (*Preferred arrival time*) al destino (trabajo en este caso); SDE_i y SDL_i denotan las demoras horarias (en unidades de tiempo) si el individuo llega temprano (*early*) o tarde (*late*) al destino, respectivamente; y D_L es una variable ficticia de penalidad adicional por llegar tarde, que refleja una discontinuidad en la función SD_i . Se espera que todos los coeficientes de la ecuación (2.1) sean negativos por la desutilidad que generan sus respectivas variables y, asimismo, que $\beta_{SDL} < \beta_{SDE}$ por la mayor penalidad que se otorga a llegar tarde al trabajo en comparación con arribar antes.

Cabe mencionar que no importa si se modela la hora de salida del origen o la hora de llegada al destino, siempre que no exista congestión no anticipada, pues por experiencia los usuarios pueden predecir el cambio en el horario de llegada producido

por un cambio en el horario de salida (de Jong *et al.*, 2003). Asimismo, las variables de demora horaria (*SD*) pueden definirse respecto a la hora preferida de llegada (*PAT*) o a la hora preferida de salida (*PDT*, *Preferred departure time*), independiente de si se modela elección de hora de salida o de llegada. Sin embargo, parece razonable usar *PAT* para viajes sensibles al horario de llegada, como el viaje hacia el trabajo, y *PDT* para viajes sensibles al horario de salida, como el viaje de regreso del trabajo (Ben-Akiva y Abou-Zeid, 2012). En cuanto a los horarios preferido *PAT* o *PDT*, algunos estudios han utilizado el horario oficial de inicio del trabajo (Small, 1982), otros el horario de salida o llegada observado en la encuesta (Rand Europe, 2004), y hay estudios en los que se pregunta directamente a los encuestados estos valores (Tseng *et al.*, 2011) o indirectamente, como Borjesson (2008) que pregunta por el horario de salida que escogería la persona en ausencia de congestión.

Otro aspecto que cabe destacar es que el *SM* no considera la variable costo, puesto que, en general, el costo operacional de un vehículo varía en forma despreciable con el horario de salida (Small, 1982; Borjesson, 2008), sumado a que la gran mayoría de los sistemas de transporte en el mundo no tienen en práctica un esquema de tarificación vial. Por esto, la mayoría de los modelos EHIV estimados a partir de datos de PR excluyen la variable costo, en contraste con experimentos de PD en los cuales es posible incorporar, por ejemplo, esquemas hipotéticos de cobro por congestión.

Otras variables que han sido incorporadas en modelos EHIV son del tipo socioeconómicas a nivel individual (género, edad y nivel educacional) y a nivel del hogar (ingreso familiar, número de residentes y de hijos, número de autos, etc.), variables relacionadas al empleo u ocupación (tipo, cantidad de horas trabajadas, flexibilidad laboral y subsidio de estacionamiento por parte del empleador), atributos del origen y destino del viaje (densidad urbana y una variable ficticia si destino está en el centro de la ciudad), características del viaje o *tour* (propósito, frecuencia y número de paradas intermedias), de la ruta (por ejemplo, si cruza un puente) o del modo (auto como chofer o acompañante, en modelos EHIV para usuarios de auto). Todas estas variables se han ingresado tanto en forma aditiva o interactuando con otras variables, ya que permiten diferenciar a la muestra e identificar las restricciones sociales y económicas de

los individuos (por ejemplo, un padre que debe llevar a sus hijos al colegio antes del trabajo). Asimismo, otras variables de nivel de servicio que también se han incluido son la tarifa, tiempo de espera y caminata, y la comodidad cuando la elección de hora está relacionada con un modo de transporte público (ver por ejemplo Bianchi *et al.*, 1998) o si se modela de forma conjunta con la elección de modo (Habib *et al.*, 2009; Habib, 2012).

En el caso de modelos EHIV basados en el *SM* y que usan *tours* como unidad de análisis, se agregan y/o modifican variables en la función de utilidad, pues éstos deben ser capaces de determinar simultáneamente la elección del horario de llegada al destino, el horario de salida de éste y la duración de la actividad llevada a cabo en el destino. De Jong *et al.* (2003), Rand Europe (2004) y Arellana (2012) modelan la hora de llegada y la duración de la actividad (lo que determina implícitamente el horario de salida de ésta) al incluir en las funciones de utilidad variables de penalidad por duraciones mayores o menores a una duración preferida (de forma similar a los *SD* del *SM*) y niveles de servicio que corresponden a la suma de los trayectos de ida y de vuelta.

En relación al propósito del viaje o *tour*, los estudios se enfocan, en general, en un tipo de propósito: (i) compras en Bhat (1998a; 1998b) y Bhat y Steed (2002); (ii) viajes no laborales (Tringides *et al.*, 2004), y (iii) viajes al trabajo (Small, 1982; Hendrickson y Plank, 1984; Arnott *et al.*, 1994; Habib *et al.*, 2009; Habib, 2012). Cuando se modelan múltiples propósitos, lo que sucede generalmente en el contexto de estudios prácticos, la tendencia ha sido estimar modelos de forma separada para cada propósito (de Jong *et al.*, 2003; Vovsha y Bradley, 2004; Abou Zeid *et al.*, 2006). En general, la mayoría de los estudios se han enfocado en los viajes basados en el trabajo (hacia o desde el trabajo) pues este propósito es el más importante en los horarios punta (en Santiago, el 70% de los viajes de la punta mañana son por motivo trabajo, DICTUC, 2006), y la congestión, un fenómeno especialmente relevante en las horas punta, es uno de los factores más influyentes en la elección de la hora de inicio del viaje (de Jong *et al.*, 2003).

La flexibilidad laboral también ha recibido un tratamiento dispar en la literatura. Algunos trabajos estiman modelos de forma separada para trabajadores con y sin flexibilidad laboral (Rand Europe, 2004; Hess *et al.*, 2005; Borjesson, 2008), otros la

incluyen interactuando con las variables de *SD* (Small, 1982; de Jong *et al.*, 2003), mientras que recientemente Arellana (2012) y Arellana *et al.* (2012c) la incorporan a través de una variable latente. Otra alternativa, que no se ha llevado a cabo en ninguno de los estudios revisados, es considerar que la disponibilidad de alternativas de hora de cada individuo depende de su flexibilidad laboral, acotando el conjunto de alternativas disponibles de los individuos que tengan baja flexibilidad laboral (esto siempre que el modelo este basado en alternativas de tiempo discretas).

En cuanto a estudios de elección horaria en Chile, el único que se asemeja a la presente investigación fue el realizado por Bianchi *et al.* (1998). El estudio consistió de un experimento de PD a usuarios del Metro de Santiago con el fin de explorar el impacto temporal en la demanda del Metro de implementar una nueva política de cobro que diferenciaba la tarifa según el período; el objetivo de la nueva política era reducir la congestión del Metro en la punta mañana a fin de posponer la compra de nuevo equipamiento. Otros estudios revisados guardan una relación más bien indirecta con modelos EHIV, como por ejemplo Munizaga *et al.* (2008), que derivan y calibran un modelo conjunto de asignación de tiempo y modo basado en un enfoque microeconómico, pero cuyo fin no es determinar el horario de inicio de las actividades.

Como ya fue mencionado, a partir de los aportes de Vickrey (1969) y Small (1982) se han estimado diversos modelos EHIV, ya sea con propósitos académicos o prácticos. Sin embargo, en la literatura persiste el debate sobre la representación del tiempo como continuo o discreto y, además, existe mucha diversidad en torno al tipo de dato empleado (PR y/o PD), a la estructura de los modelos de elección discreta (MNL, HL, etc.) y a la modelación conjunta de EHIV con otras decisiones, tales como elección de modo o ruta, o la programación de actividades. Estos temas se presentan en mayor detalle en las secciones 2.1.1. a 2.1.4.

2.1.1 Representación del tiempo: continuo versus discreto

La forma más común de especificar modelos de elección de la hora de inicio de viaje ha sido a través de una representación discreta del tiempo; es decir, dividir el tiempo, que es una variable esencialmente continua, en un número finito de intervalos. De esta manera, la situación de elección que enfrenta el individuo es la de escoger un

período, de entre todos los disponibles, en el cual comenzar el viaje (o llegar al destino). Small (1982) fue uno de los pioneros en “discretizar” el tiempo en modelos EHIV. Las principales razones que da para ello son: 1) si hay cobros en períodos punta, posiblemente algunos individuos elegirán los extremos de los períodos de cobro (solución esquina al problema microeconómico propuesto); 2) la pendiente de la curva que relaciona los tiempo de viaje con la longitud del período no puede ser medida con precisión (los niveles de servicio generalmente se encuentran promediados para cada intervalo de tiempo) y 3) las personas tienden a aproximar el horario que declaran a 5, 10 o 15 min.

Otra razón para justificar el uso del enfoque discreto es que se puede utilizar como herramienta de modelación a los modelos de elección discreta basados en la teoría de la utilidad aleatoria (TUA), ya que han sido extensivamente investigados (Ortúzar y Willumsen, 2011), son fáciles de estimar (hay una variedad de programas disponibles en el mercado, algunos gratuitos) y también son relativamente sencillos de acoplar a un modelo clásico de cuatro etapas.

De forma muy resumida, la TUA postula que las personas toman decisiones racionales, poseen información completa y eligen - dentro de un conjunto discreto de alternativas - aquella que le reporta más utilidad (sopesando las ventajas y desventajas de los distintos atributos que describen a las alternativas). El modelador trata a la utilidad como una variable aleatoria, pues no es capaz de observar todos los atributos considerados por los individuos; por esto, a la componente determinística u observada por él (que es generalmente una función lineal en parámetros a ser estimados) le asigna un término de error aleatorio. Dependiendo de la distribución de este componente de error, se pueden generar distintos modelos; por ejemplo, si el error es independiente e idénticamente distribuido (IID) Gumbel, se genera el modelo logit multinomial (MNL). Más detalles sobre la teoría de la utilidad aleatoria y los modelos de elección discreta relevantes a este estudio se explicitan en la sección 2.1.3 y en el Anexo A.

A pesar de las ventajas del enfoque discreto, este presenta algunos inconvenientes que surgen principalmente de la necesidad de “discretizar” una variable continua. La principal desventaja es la existencia de correlación entre los períodos horarios debido a

las variables no observadas o error aleatorio (de Jong *et al.*, 2003). Esta correlación puede ser más alta al emplear intervalos de tiempo pequeños, como, por ejemplo, de 5 o 15 min. No obstante, es posible acomodar esta correlación empleando estructuras de modelación discreta más flexibles que el MNL, como el logit mixto (ML, ver siguiente sección para más detalles). Sin embargo, incluso usando este tipo de estructuras no se pueden resolver los siguientes problemas (Bhat y Steed, 2002): (i) la elección del tamaño de los intervalos de tiempo queda a discreción del modelador; (ii) los tiempos cercanos al límite de dos períodos contiguos se asumen como diferentes cuando en realidad son casi idénticos, y (iii) los límites temporales de estrategias TDM, como la tarificación vial, deben concordar con los intervalos de tiempo definidos durante la estimación del modelo (por ejemplo, si se desea tarificar de 7:45 a 9:30, es necesario que ambas horas sean el límite de al menos un intervalo). En relación a (i), sólo se encontró un estudio que contrastaba modelos con alternativas de tiempo de distinta longitud (Hess *et al.*, 2007), en el cual comparaba intervalos de 15 minutos, una hora, y cinco períodos (punta mañana, fuera punta, etc.). Más investigación es necesaria en esta línea, en la que además se considere explícitamente el compromiso entre menos correlación entre alternativas y la pérdida de información al agregar niveles de servicio y elecciones. Para ello es fundamental contar con variables de nivel de servicio que se puedan agregar a distintas resoluciones temporales.

Una alternativa al enfoque discreto es modelar EHIV considerando el tiempo como variable continua; para esto existen diversos enfoques. El primero es una extensión y adaptación del trabajo de Vickrey (1969), denominada *Equilibrium Scheduling Theory (EST)*, en la cual el equilibrio entre la demanda y oferta del sistema de transporte en relación a las preferencias temporales, se determina de forma explícita y determinística. Uno de los principales modelos teóricos del *EST* es el desarrollado por Arnott *et al.* (1990; 1994); este asume, por el lado de la oferta, un cuello de botella (arco) en el cual las demoras son determinísticas y proporcionales al largo de la cola en el arco; por el lado de la demanda, se basa en la función de utilidad (o costo) del *SM* de Small, considerando viajeros homogéneos y heterogéneos en cuanto a los parámetros de la función de utilidad y horarios preferidos de llegada (*PAT*). Una aplicación práctica de

EST a redes de transporte operacionales es el modelo HADES (*Heterogeneous Arrival and Departure Times based on EST*) que se puede combinar con *software* más tradicional de asignación de viajes (van Vuren *et al.*, 1999). Sin embargo, existe consenso en que los modelos basados en *EST* funcionan adecuadamente para cambios pequeños (5 a 10 min) en la elección horaria, y que posiblemente HADES sea la última etapa de desarrollo de *EST* (de Jong *et al.*, 2003).

El segundo enfoque continuo, más reciente, se basa en los modelos de duración de riesgo (*hazard duration*), que describen la probabilidad de que una actividad finalice condicional en el tiempo que ya lleva realizándose (*hazard rate*); así, la hora de inicio del viaje es igual a la hora en que la actividad pre-viaje termina su duración. La tasa de riesgo se especifica como una función dependiente de una distribución base y de ciertas variables explicativas (nivel de servicio, socioeconómicas, etc.). Asimismo, hay dos tipos de modelos de esta clase, que difieren en la forma de acomodar el efecto de las variables explicativas: (i) modelo de riesgo proporcional, en que la tasa de riesgo es igual a la multiplicación de un término dependiente del tiempo (distribución base) y otro dependiente sólo de las variables explicativas (Hensher y Mannering, 1994), y (ii) modelo de riesgo acelerado, en el cual las variables explicativas pueden acelerar o desacelerar la tasa de riesgo (Lee y Timmermans, 2007). Se ha estimado modelos EHIV usando tanto riesgo proporcional (Wang, 1996; Bhat y Steed, 2002) como riesgo acelerado (Habib *et al.*, 2009; Habib, 2012), con las ventajas de no requerir intervalos discretos de tiempo y permitir estimar tanto la hora de inicio de las actividades como la duración de estas. No obstante, los modelos de riesgo acelerado permiten capturar de forma más realista el comportamiento dinámico de las elecciones temporales (Lee y Timmermans, 2007).

A pesar de que los resultados de los modelos de riesgo son promisorios, no existe consenso acerca de su utilización en la práctica, debido a que su estimación es compleja (no hay disponibilidad de *software* en paquetes econométricos tradicionales) y a que su aplicación consumiría considerables cantidades de tiempo y recursos computacionales (Arellana, 2012). Otro inconveniente de estos modelos es que sólo proveen una aproximación estadística para modelar las elecciones temporales, pero carecen de una

teoría de comportamiento individual (Ettema *et al.*, 2007; Lemp y Kockelman, 2010), lo que limita la interpretación económica de sus parámetros y el cálculo de medidas de bienestar.

El enfoque continuo, no obstante, no se limita a los modelos de *EST* y de riesgo. Otra formulación ha sido desarrollada por Ettema y colegas (Ettema y Timmermans, 2003; Ettema *et al.*, 2004; 2007). El enfoque propone que la EHIV de las personas se determine maximizando una función de utilidad con dos componentes: una relativa a la programación de sus actividades (en qué momento y por cuánto tiempo realizar una actividad) y otra asociada a la desutilidad de los viajes (tiempo y costo de viaje). La componente de actividad, en el modelo más avanzado, se ve influenciada por tres factores: (i) preferencias temporales individuales (i.e. en qué momento realizar una determinada actividad), representadas por funciones de utilidad marginal continuas (con forma de campana o siguiendo la distribución Cauchy); (ii) efecto saciedad, derivada de utilidades marginales decrecientes y representado por una función logarítmica que depende de la duración acumulada de la actividad, y (iii) restricciones en la programación de actividades del tipo *Schedule Delay (SD)* del *SM* de Small (1982), como por ejemplo, la hora de inicio del trabajo. La primera componente permite representar el hecho que las personas prefieren un período del día dónde realizar sus actividades, aquel en que la utilidad marginal debida a una unidad de tiempo extra de participación en una actividad está en su cota máxima; esta utilidad marginal disminuye a medida que la persona se mueve de ese período. De esta forma, los costos de programar actividades (*Scheduling costs*) no se derivan de los viajes, sino que de las actividades. Por ejemplo, un mayor tiempo de viaje implica menos tiempo dedicado a una actividad en el período de mayor utilidad marginal. Un aspecto positivo de estos modelos es que permiten modelar la EHIV en conjunto con la programación diaria de actividades (inclusive es posible modelar simultáneamente la elección de modo). Sin embargo, una limitación del enfoque es que requiere datos con un nivel de precisión muy alto, como la programación de actividades detallada de todos los individuos, incluyendo las horas de comienzo y termino de todas las actividades realizadas (Habib *et al.*, 2009).

Una cuarta formulación, más reciente, que emplea el tiempo como continuo es la elaborada por Habib (2011). Al igual que en los modelos de riesgo, este enfoque considera que la asignación de tiempo en el hogar, antes del trabajo, define la hora de salida al trabajo, pero a diferencia de dichos modelos el enfoque está basado completamente en la teoría de maximización de la utilidad aleatoria bajo restricciones de tiempo (24 horas para trabajar y realizar otras actividades). En esta línea, Lemp y Kockelman (2010) estiman un modelo logit continuo de EHIV mediante métodos bayesianos.

Por último, es interesante mencionar el enfoque propuesto por Vovsha y Bradley (2004) que, en estricto rigor, no es discreto ni continuo sino que más bien híbrido discreto. Este modelo emplea una estructura discreta que opera a nivel de *tour* (combinación de viajes basados en el hogar), en que la utilidad total depende de dos de los siguientes tres componentes (uno no se considera por ser una combinación lineal de los otros): hora de salida del hogar, hora de llegada al hogar y duración de la actividad del *tour*. En ambos componentes la utilidad depende de variables de “cambio continuo” (*continuous-shift variables*), es decir, variables explicativas que son constantes en el tiempo (por ejemplo, variables sociodemográficas y de uso de suelo) y que se multiplican por la hora de inicio y por la duración de la actividad (que son continuas). Esta multiplicación es sustentada en modelos de riesgo. Al emplear una estructura discreta (resolución temporal de una hora) con funciones de utilidad dependientes del tiempo como variable continua (por eso el nombre de híbrido discreto), este modelo combina las ventajas de un modelo de riesgo (estructura parsimoniosa con pocos parámetros que permiten cualquier resolución temporal, inclusive tratar el tiempo como continuo) con las ventajas de uno de elección discreta (flexibilidad en la especificación y facilidad de estimación y aplicación). Otra de sus ventajas es que se puede estimar usando datos convencionales provenientes de diarios de actividad (Ettema *et al.*, 2007), y ya ha sido estimado para su aplicación en un modelo operacional de demanda de transporte.

2.1.2 Tipos de datos utilizados para modelar EHIV

Los primeros estudios de EHIV utilizaron datos de preferencias reveladas (PR), pues se enfocaban principalmente a viajes al trabajo y hacían uso de datos recopilados en el contexto de grandes estudios de transporte (Arellana, 2012). Posteriormente, dada la escasez de datos de PR de buena calidad (por el alto costo que implica recolectarlos), la tendencia ha sido utilizar datos de preferencias declaradas (PD) para modelos de EHIV (Tseng *et al.*, 2011). A esto se debe sumar que cuando se utiliza el enfoque de modelación discreta, los datos de PR suelen presentar correlación entre alternativas y atributos (Borjesson, 2008), y a veces insuficiente variación en los niveles de algunos atributos en términos de obtener parámetros bien estimados (Ortúzar y Willumsen, 2011). Asimismo, en muchos casos se desea evaluar el efecto de políticas que no están en operación en la actualidad, como la tarificación vial, y esto no es posible a partir de datos de PR.

Las desventajas anteriores se pueden superar mediante datos provenientes de experimentos de PD bien diseñados. Sin embargo, los datos de PD presentan el inconveniente de ser menos confiables para realizar predicciones, a diferencia de los datos de PR, pues éstos contienen información sobre las elecciones efectivamente realizadas por los individuos. Por ejemplo, Borjesson (2008), que mezcla datos de PR y PD en un contexto de elección de modo y hora, encontró que los datos de PD reflejaban que los individuos preferirían cambiar el horario de su viaje por sobre el modo ante un cobro hipotético por congestión; sin embargo, una vez aplicada la tarificación vial lo que sucedió en la realidad fue que los usuarios de auto se cambiaron principalmente al transporte público sin cambios significativos en los horarios en que se realizaban los viajes.

En el último tiempo ha ocurrido un repunte en el uso de datos de PR, en parte gracias a la aplicación de metodologías menos intensivas en recursos para estimar la distribución temporal de los tiempos de viaje. Por ejemplo, Abou-Zeid *et al.* (2006) y Popuri *et al.* (2008), determinaron la distribución temporal de velocidades a partir de regresiones que utilizaban información tradicionalmente recopilada en encuestas de transporte y niveles de servicio generados por un modelo clásico de transporte

(velocidades en punta mañana y fuera de punta). Tseng *et al.* (2011), por su parte, hacen uso de información proveniente de cámaras viales para estimar los tiempos de viaje (aunque suplementan sus datos con información GPS especialmente recolectada para su estudio).

Una manera de mitigar las desventajas de cada tipo de dato consiste en estimar modelos que combinan ambas fuentes de datos (PR/PD). Aunque existe una larga tradición de usar datos de PR y PD en la estimación de modelos de elección discreta (Ben-Akiva y Morikawa, 1990; Louviere *et al.*, 2000; Cherchi y Ortúzar, 2002; Bhat y Castelar, 2002), en el contexto de elección de la hora de inicio de viaje sólo hay tres estudios a la fecha: Borjesson (2008), Tseng *et al.* (2011) y Arellana (2012). Borjesson (2008) estimó modelos conjuntos de modo y EHIV, pero sólo consideró la elección de modo en los datos de PD. Por otro lado, Tseng *et al.* (2011) sólo modelan EHIV y no consideran la elección de modo. Por su parte, Arellana (2012) modela la elección de modo-hora en el contexto de *tours* basados en el trabajo (viaje de ida y vuelta del trabajo), con elección de modo sólo en PR y modo-hora en PD. Algunos aspectos interesantes que se han explorado al combinar datos de PR y PD son la diferencia de escala (o varianza) entre ambos tipos de datos, la correlación entre las alternativas de PR y PD (efecto panel), el efecto inercia (tendencia a escoger la alternativa de PR en el experimento de PD) y la heterogeneidad en los gustos (Bhat y Castelar, 2002; Ortúzar y Willumsen, 2011).

2.1.3 Estructura de modelos de elección discreta aplicados a EHIV

La primera estructura utilizada para modelar EHIV mediante modelos de elección discreta fue el logit multinomial (MNL), debido a su simpleza y facilidad de estimación (Small, 1982; Hendrickson y Plank, 1984). Sin embargo, el principal inconveniente del MNL es que asume que las alternativas de elección son independientes entre sí, lo que no es muy realista considerando que en este caso corresponden a períodos de tiempo. Esta correlación se hace aún más evidente cuando los intervalos de tiempo son pequeños.

Una estructura más general que el MNL y que permite acomodar correlación entre alternativas es el logit jerárquico (HL). Bhat (1998a), Hess *et al.* (2007) y Bajwa (2008) han usado el HL para modelar en forma conjunta la elección de modo-EHIV, concluyendo que efectivamente el HL supera al MNL. No obstante, una limitación del HL es que la magnitud de la correlación debe ser la misma para todas las alternativas correlacionadas (o pertenecientes a un mismo nido). Esto no parece ser adecuado en el contexto de elecciones temporales, pues se esperaría que períodos de tiempo contiguos presentaran una mayor correlación que períodos de tiempo más lejanos. Es por ello que Small (1987) desarrolló el modelo de “valor extremo generalizado con alternativas ordenadas” (OGEV por sus siglas en inglés). Este permite estimar un parámetro de correlación para cada par de alternativas, que depende de la “cercanía” entre ellas, presentando su valor máximo para alternativas contiguas (para pares de alternativas muy alejadas entre sí, la correlación es prácticamente despreciable, lo que implica independencia como en el MNL); el caso más simple ocurre cuando se utiliza un solo parámetro de correlación para alternativas contiguas.

Bhat (1998a) aplicó con éxito una estructura OGEV para un modelo de modo-EHIV, consistente de un MNL en la elección de modo (nido superior) y un OGEV en la elección de hora (nido inferior), que consideraba correlación (adicional a la del HL) para alternativas adyacentes de un mismo nido (es decir, las alternativas de hora que no eran adyacentes se comportaban como un HL). Entre sus resultados se destaca que la estructura MNL-OGEV supera a los modelos HL y MNL, y que a su vez el HL es superior al MNL. Esta superioridad no sólo se explicaba por la mayor bondad de ajuste del modelo, sino porque el HL y MNL evalúan incorrectamente los efectos de medidas de gestión como la tarificación vial. Small (1987) y Steed y Bhat (2000) también estimaron modelos OGEV para la EHIV; sin embargo, el primero encontró que el OGEV no era estadísticamente superior al HL y al MNL, y el segundo prefirió el MNL por sobre el OGEV, debido a que en este último el parámetro de correlación presentó una inconsistencia teórica.

Otra desventaja del modelo HL es no permitir que alternativas de distintos nidos se encuentren correlacionadas, lo que se puede realizar con el modelo *cross-nested logit*

(CNL). Bajwa *et al.* (2008) comparan el CNL con el HL en el contexto de una elección de modo-EHIV, encontrando que el modelo CNL es superior.

Una estructura más general que todas las mencionadas es el modelo probit, ya que permite postular una matriz de covarianza arbitraria cuya estimación es bastante sencilla en situaciones de elección con sólo dos alternativas (probit binario o bivariado); Tringides *et al.* (2004) estimaron con éxito un modelo probit binario para la EHIV. Sin embargo, para situaciones de elección de tres o más alternativas (probit multinomial o MNP), la estimación del modelo es compleja pues la función de verosimilitud involucra integrales múltiples.

Finalmente, el modelo logit mixto (ML) de componentes de error (ECL, ver anexo A) es el más flexible de todos los modelos de elección discreta, pues puede aproximar con un nivel de exactitud notable (McFadden y Train, 2000) a todas las estructuras antes descritas (MNP, CNL, HL, OGEV, etc.), permitiendo construir la estructura de correlación que el modelador desee. Asimismo, permite capturar el efecto panel o de múltiples respuestas por individuo en un experimento PD al especificar que el componente de error aleatorio asociado a su matriz de covarianza arbitraria sea igual para todas las situaciones de elección de una misma persona. Por esto ha sido utilizado en diversos estudios de EHIV (Bhat, 1998b; de Jong *et al.*, 2003; Rand Europe, 2004; Hess *et al.*, 2007; Bajwa *et al.*, 2008; Borjesson, 2008). El resultado general de estos estudios es que el ECL supera a otras estructuras como el HL y que usar este último modelo en este caso puede conducir a conclusiones incorrectas ante cambios en los niveles de servicio (Bhat 1998b).

Por su parte, el ML de parámetros aleatorios (MLPA) permite capturar heterogeneidad en los gustos al postular coeficientes individuales que tienen una cierta distribución en la población. Bajwa *et al.* (2008), Borjesson (2008) y Tseng *et al.* (2011) postularon parámetros aleatorios con una distribución normal para los coeficientes de demora horaria (*SDE* y *SDL*), concluyendo que estas variables efectivamente exhibían heterogeneidad en los gustos. Otra ventaja del MLPA es que, al igual que el ECL, permite acomodar el efecto panel. Sin embargo, algunas limitaciones son que la distribución de los parámetros (normal, log-normal, etc.) se asume *a priori* y puede que

no replique adecuadamente la distribución en la población. Además, algunas distribuciones como la normal, pueden dar lugar a parámetros individuales con signos inconsistentes con la teoría microeconómica (ver la discusión en Sillano y Ortúzar, 2005). En cuanto a comparaciones entre el ECL y el MLPA, Bajwa *et al.* (2008) encontraron que un modelo de coeficientes aleatorios que utilizaba un HL como base (*mixed nested logit*) era superior al ECL, pues permitía acomodar heterogeneidad en los gustos y correlación entre alternativas.

A pesar de que modelos más flexibles, como el ML, permiten representar de forma más adecuada el proceso de elección de las personas, su aplicación para fines prácticos es más compleja y demandante de recursos computacionales (pues requieren simulación) en contraste con modelos más simples como el MNL o HL. Esto ha tenido como consecuencia que los principales modelos EHIV utilizados en la práctica sean el MNL (Rand Europe 2004; Vovsha y Bradley, 2004; Popuri *et al.*, 2008) y el HL (Rand Europe 2004; Hess *et al.*, 2007).

2.1.4 Modelación EHIV junto a otras decisiones

En la literatura existen diversos trabajos que han estudiado la elección de ruta y EHIV de forma conjunta, pues es de esperar que ambas decisiones estén relacionadas. Mannering (1989) y Khattak *et al.* (1995) desarrollaron modelos que determinaban las elecciones temporales en conjunto con la elección de ruta a nivel individual. Por otra parte, la *Equilibrium Scheduling Theory (EST)* descrita anteriormente, combina la asignación agregada (esto es, no a nivel individual) a una red con la componente temporal de la demanda por transporte. Heydecker y Addison (2005) realizaron algo similar considerando asignación dinámica. La relación entre la elección de ruta y EHIV parece estar más entrelazada cuando se consideran políticas como la tarificación vial, pues las personas pueden utilizar diferentes rutas y/o viajar en distintos horarios para evadir este pago.

La elección de la hora de inicio de viaje también se puede ver afectada por la elección de destino. Por ejemplo, Kitamura *et al.* (1998) encontraron evidencia de la existencia de una componente temporal que afectaba la elección de destino para distintos

propósitos de viaje, al incluir variables ficticias que reflejaran las elecciones temporales dentro de la función de utilidad de un modelo de elección de destino. No obstante, esta relación parece ser la menos investigada en la literatura, debido a que la elección horaria como respuesta a la congestión es más bien de corto plazo (no así la elección de cuánto tiempo y en qué horas del día trabajar, pues se determina en un horizonte de tiempo mayor), mientras que la de destino es de mediano-largo plazo.

La presente investigación se enfoca en la elección conjunta de modo y hora de inicio de viaje. A pesar de que esta relación es la más investigada en la literatura, no existe consenso generalizado sobre la posición relativa de estas elecciones o, lo que es similar, a cuál de estas dos decisiones son más sensibles los individuos. Por ejemplo, algunos estudios concluyen que primero ocurre la elección de modo y luego la elección horaria, en una estructura de modelación jerárquica para viajes al trabajo (Hess *et al.*, 2007; Arellana *et al.*, 2012a) o de compras (Bhat 1998a). Por otro lado, Hendrickson y Plank (1984), a pesar de utilizar un modelo MNL (que implica que ambas decisiones se determinan de forma simultánea), mediante un análisis de elasticidades concluyen que los individuos son más sensibles a cambiar el horario de inicio de viaje que a modificar su modo de transporte (lo que significaría que en un modelo HL la elección de modo sobre la elección horaria sería la estructura jerárquica más adecuada). A su vez, Tringides *et al.* (2004) concluyen que para los viajeros que trabajan, la estructura más adecuada es aquella en que la EHIV precede a la elección de modo (no así para viajeros que no son trabajadores) en el contexto de viajes no laborales. Finalmente, de Jong *et al.* (2003) encuentran que en los *tours* hogar-trabajo, los viajeros son más elásticos a cambios de modo que a cambios de hora.

Un aspecto que cabe destacar en la literatura revisada, es que los modelos de elección de modo y EHIV basados en el enfoque discreto presentan pocas alternativas de modo y de hora. En general, los experimentos PD de modo-hora consideran una alternativa de cambio de modo y tres alternativas de hora (salir antes, después y en horario similar al observado), pues más alternativas pueden conducir a una sobrecarga cognitiva del encuestado (Ortúzar y Willumsen, 2011). En el caso de PR, esto se debe a que al agregar una nueva alternativa de modo u hora, el número total de alternativas

(modo-hora) crece potencialmente, y también porque no suele haber disponibilidad de datos para modelar el problema en forma más desagregada. Por ejemplo, se encontró que el máximo número de alternativas de modo-hora ascendía a 28 (cuatro modos y siete horarios en Hendrickson y Plank, 1984), seguido por 15 (tres modos y cinco períodos) en Bhat (1998a; 1998b).

Por último, las preferencias temporales no sólo se ven afectadas por la elección de modo, hora y ruta, sino que también por la programación diaria de actividades (Wang, 1996; Ettema *et al.*, 2004; 2007) y la duración del trabajo (Habib, 2012), pues el tipo y duración de las actividades (ocio, restrictivas, trabajo, etc.) que una persona debe realizar en un día, influyen en cómo las personas escogen las horas para realizar sus viajes.

2.2 Variabilidad del tiempo de viaje en contexto de EHIV

La variabilidad en los tiempos de viaje corresponde a la incertidumbre temporal que experimentan los usuarios del sistema de transporte al desplazarse en la red vial. En general, la incertidumbre en los tiempos de viaje se clasifica de acuerdo a tres categorías (Li *et al.*, 2010): (i) variabilidad día a día o entre días; (ii) variabilidad entre períodos u horas de un día, y (iii) variabilidad entre vehículos, que surge por hábitos de conducción y por la señalización de tránsito. La mayoría de los estudios en la literatura se enfoca en la variabilidad de los tiempos de viaje entre distintos días (i.e. viajes que se repiten todos los días no tienen igual tiempo de viaje), la cual se explica en gran medida por la naturaleza estocástica del tiempo de viaje, que a su vez es causada por la aleatoriedad de la demanda y oferta de transporte. Por ejemplo, Noland y Small (1995) dividen el tiempo de viaje en tres componentes: a flujo libre, congestionado (que depende de la hora en que se realiza el viaje) y un tercer componente, aleatorio, que se refiere a la variabilidad del tiempo de viaje debido a congestión no recurrente y que no es posible anticipar por parte de los viajeros (i.e. causada por accidentes o trabajos en las vías, entre otras causas).

Esta variabilidad en los tiempos de viaje genera costos adicionales a los viajeros y puede impactar otras decisiones como la elección de modo, ruta u hora, siendo ésta

última la respuesta más inmediata en presencia de incertidumbre (Li *et al.*, 2010). Es por ello que, en general, la variabilidad del tiempo de viaje se evalúa en el contexto de elección de modo, ruta o EHIV.

Existen dos enfoques principales para modelar la variabilidad en los tiempos de viaje que se describen brevemente a continuación (para más detalles ver las revisiones de Li *et al.*, 2010 y Carrion y Levinson, 2012). El primer enfoque corresponde a los modelos de media-varianza, en que se especifica tanto la media (o mediana) de la distribución de tiempo de viaje como su varianza (o desviación estándar, coeficiente de variación o, incluso, la resta entre el percentil 80 y percentil 50 de la distribución del tiempo de viaje) de forma aditiva en la función de utilidad. En este enfoque se asume que la variabilidad del tiempo de viaje produce desutilidad *per se*. Un aspecto positivo de los modelos de media-varianza, que se pueden estimar a partir de datos de PR y PD, es que permiten calcular un índice de confiabilidad (*reliability ratio*) igual a la razón entre el valor subjetivo de la variabilidad del tiempo de viaje y el valor subjetivo del tiempo de viaje. Cabe destacar que algunos estudios reportan que este índice es mayor que uno, es decir, se valora más la variabilidad del tiempo de viaje que el tiempo de viaje en sí (ver Li *et al.*, 2010).

La segunda forma de abordar el problema hace uso de una extensión del modelo *SM* de Small (1982), incorporando incertidumbre en los tiempos de viaje (Noland y Small, 1995). De esta forma, los individuos escogen aquella alternativa que maximiza su utilidad esperada (ver ecuación 2.6). En este caso, tal como fue mencionado previamente, se incurre en desutilidad al llegar a una hora distinta de la preferida:

$$E(U_i) = \beta_{TV} E(TV_i) + \beta_{SDE} E(SDE_i) + \beta_{SDL} E(SDL_i) + \gamma P_L \quad (2.6)$$

donde $E(TV_i)$ es el valor esperado de los tiempos de viaje presentados al encuestado; $E(SDE_i)$ y $E(SDL_i)$ son las demoras horarias esperadas y P_L es la probabilidad de llegar tarde. El subíndice i indica la alternativa, que puede corresponder a hora de salida o ruta (u otra), dependiendo del contexto (hora, ruta otro) del experimento. Los modelos basados en el *SM* se suelen estimar sólo usando datos de PD; sin embargo, un modelo

del tipo *SM* que contenga demoras horarias esperadas, no necesita adicionar directamente una variable que refleje la incertidumbre de los tiempos de viaje (Small *et al.*, 1999).

En cuanto al tipo de datos (PR o PD), en general parece ser más adecuado modelar la variabilidad del tiempo de viaje a través de datos de PD debido a la alta correlación entre el tiempo de viaje y su variabilidad en los datos de PR. En cuanto al diseño del experimento de PD, Li *et al.* (2010) concluyen que la forma más apropiada de presentar la variabilidad del tiempo de viaje es a través de presentar múltiples tiempos de viaje, ya que esto facilita la comprensión de este fenómeno por parte del encuestado y permite estimar ambos tipos de modelos (media-varianza y *SM*). Esto último es altamente deseable, ya que posibilita realizar una comparación de resultados.

Existen estudios que investigan empíricamente la variabilidad del tiempo de viaje a través de modelos EHIV; sin embargo, en la mayoría de ellos el foco principal del trabajo está puesto en la variabilidad y no en la elección horaria. Algunas excepciones son los trabajos realizados por Borjesson (2008) y Arellana *et al.* (2012a). Borjesson (2008) encuentra que la razón de confiabilidad cambia considerablemente entre un modelo de EHIV que sólo usa datos de PR (1.66) y otro que combina datos de PR y PD (0.74). Por su parte, Arellana *et al.* (2012a) estima una razón de confiabilidad superior a uno, lo que es consistente con otros estudios.

2.3 Modelos de EHIV en ausencia de demoras horarias (SD)

En general las grandes encuestas de movilidad que se utilizan para calibrar los modelos de planificación de sistemas de transporte, cómo las encuestas Origen y Destino (EOD) en Chile, no recopilan información acerca de los horarios preferidos de llegada (*PAT*). Esto, sumado a la dificultad de predecir los *PAT* al mismo nivel de precisión que en la base de estimación, es un impedimento para incorporar las demoras horarias (*SDE* y *SDL*) en modelos prácticos de elección horaria (Hess *et al.*, 2007). En ausencia de estas variables y bajo algunos supuestos, las constantes específicas (ASC) capturan los efectos de estas variables. Ben-Akiva y Abou-Zeid (2012) demuestran que modelar la hora de llegada para viajes sensibles a la llegada (por ejemplo, el viaje al trabajo) reduce

las demoras horarias, *SDE* y *SDL*, a constantes si se asume que los *PAT* son constantes para individuos pertenecientes a un mismo estrato socioeconómico (ídem para viajes sensibles a la hora de salida, como por ejemplo el viaje de retorno del trabajo). Esto requiere asumir de antemano cuáles viajes son sensibles al horario de salida y cuáles al de llegada. En general existe consenso que las restricciones impuestas por ciertas actividades (por ejemplo, hora de entrada y de salida del trabajo) determinan la sensibilidad de las personas en este ámbito.

No obstante, persisten otros problemas, como que el número de ASC aumenta a medida que se incrementan las alternativas horarias del modelo (esto es aún más evidente al modelar *tours*), lo que puede conducir a problemas de identificación e interpretación de los parámetros del modelo, aparte de un alto costo computacional. Para resolver estos problemas, Hess *et al.* (2005) remplazan el conjunto completo de ASC por ciertas formas funcionales (potencia, exponencial y empírica), a fin de replicar la distribución de las ASC en la muestra. Por otro lado, Abou-Zeid *et al.* (2006), Popuri *et al.* (2008), y Ben-Akiva y Abou-Zeid (2012), abordan el problema de cuan cíclica es la utilidad (la utilidad de viajar en un instante de tiempo t debe ser igual a la utilidad en $t+24$ horas para un ciclo de 24 horas), sustituyendo las ASC por funciones trigonométricas dependientes del tiempo. Entre las variables socioeconómicas que interactúan con estas constantes trigonométricas, se encuentran variables ficticias asociadas con el empleo (media jornada o tiempo completo), flexibilidad laboral, mujer con hijos, ingreso, y distancia del viaje.

Existen otros enfoques alternativos que buscan estimar las horas preferidas de salida (PDT) o llegada (*PAT*) cuando la información no está disponible. Kristoffersson y Engelson (2008) describen una metodología que, a partir de las elecciones de hora observadas y un modelo estimado previamente (que si incluye los términos *SDE* y *SDL*), permite estimar distribuciones agregadas de las horas preferidas para realizar el viaje. Brey y Walker (2011), en un contexto de elección de itinerario de vuelo, modelan la hora preferida de realizar un viaje como una variable latente en función de características individuales y del viaje.

2.4 Incorporación de variables climáticas²

La literatura acerca de la influencia de variables climáticas en el comportamiento de los usuarios del sistema de transporte es bastante escasa. Khattak y de Palma investigaron la propensión de las personas a cambiar de modo, ruta u hora del viaje en situaciones climáticas adversas, encontrando que la respuesta más frecuente a estos fenómenos eran cambios en la hora de realización de los viajes. Sabir *et al.* (2009) y Sabir (2011) estudiaron el efecto de variables meteorológicas en la elección de modo. El contexto del estudio parece propicio para analizar esta relación, pues el país donde se realizó, Holanda, presenta mucha variabilidad estacional en su clima, con temperaturas que varían desde 12°C bajo cero a más de 30°C y en que la partición modal de los modos más expuestos al ambiente (caminata y bicicleta) es bastante elevada (alrededor del 45% en conjunto) y presenta considerable variación estacional. En cuanto a los datos, combinaron información de encuestas de movilidad con información meteorológica a nivel horario. Las dos fuentes de datos se vincularon asignando, a cada viaje, las condiciones meteorológicas de la hora en que comenzó el viaje, y usando información correspondiente a la estación meteorológica más cercana a la comuna de origen. Entre las variables explicativas del modelo se encuentran la temperatura, precipitaciones y viento (todas estratificadas en categorías o variables ficticias), y variables ficticias adicionales para controlar por efectos estacionales (verano, otoño, etc.) y por períodos del día (períodos punta y fuera de punta). Entre los resultados destacables se encuentran que las personas prefieren el automóvil y el transporte público a temperaturas muy bajas (menores a 0°C), y que la preferencia por la bicicleta aumenta a medida que la temperatura sube. Sin embargo, ambos estudios (Sabir *et al.*, 2009; Sabir, 2011) presentan debilidades en la modelación pues, a excepción de la distancia, no consideran ninguna variable de nivel de servicio (tiempo de viaje, espera, caminata o costo) y los coeficientes de todas las variables son específicos a cada modo, lo que dificulta su

² En estricto rigor son variables de tiempo porque clima es el patrón acumulado del tiempo (temperatura, humedad, presión) en una zona geográfica a lo largo de muchos años, mientras que el tiempo se refiere a las condiciones atmosféricas en un determinado momento. Sin embargo, para efectos de no confundir al lector (tiempo meteorológico u tiempo de la hora del reloj) se usa la palabra clima para referirse al tiempo meteorológico.

interpretación. Como a falta de variables de nivel de servicio es difícil aislar el verdadero efecto de las variables climáticas en la elección de modo, a pesar que les interesa evaluar la variación estacional en la elección de modo, las variaciones que se producen, en parte por variables como la temperatura y precipitaciones, el modelo estimado no es capaz de reproducir adecuadamente la partición modal observada (sólo reproduce la variación estacional en la partición modal).

2.5 Principales conclusiones y limitaciones encontradas

De las secciones recién descritas, es posible concluir que no existe consenso generalizado en torno a diversos aspectos de la modelación de EHIV. En primer lugar, la representación del tiempo como discreto o continuo no ha sido del todo resuelta, ya que a pesar del problema asociado a “discretizar” una variable continua, la gran mayoría de los modelos de elección horaria usados en la práctica emplean este enfoque.

En cuanto a los modelos que emplean el enfoque discreto, se han probado diversas estructuras de modelación y, en general, la estructura más adecuada depende del contexto, pues en algunos casos se desea sacrificar flexibilidad y capacidad explicativa (logit mixto) en pos de sencillez para usar el modelo en la práctica (MNL y HL). Dado el contexto y objetivos de este estudio, lo más adecuado parece ser estimar modelos más restringidos (MNL y HL) y compararlos con modelos más flexibles (ML de componentes de error y/o de parámetros aleatorios).

Por otro lado, a pesar de que la elección de modo es la que ha sido más investigada en conjunto con la elección horaria, no hay claridad respecto a cuál de las dos decisiones es más sensible para los usuarios de un sistema de transporte, ya que esto también parece diferir según contexto (ciudad, propósito del viaje, etc.). Por ello, esto se debe determinar empíricamente en cada contexto. Asimismo, se puede concluir que no existe un modelo conjunto de elección de modo y hora de inicio del viaje que combine ambas decisiones tanto en PR como en PD, y que a su vez disponga de un conjunto amplio de alternativas combinadas.

Una de las mayores críticas al enfoque discreto es que la longitud de los intervalos de tiempo a ser usados como alternativas queda a discreción del modelador.

Prácticamente ningún estudio (a excepción de Hess *et al.*, 2007) ha evaluado el efecto de variar el largo de las alternativas de tiempo en la bondad de ajuste y parámetros del modelo, a pesar de que en la literatura se han utilizado intervalos que van desde 10 min hasta períodos amplios como toda la punta mañana.

En cuanto a la variabilidad del tiempo de viaje, aparentemente lo ideal sería agregarla en las funciones de utilidad de PD usando un modelo de media-varianza como el *SM* extendido de Noland y Small (1995). De esta manera, es posible comparar resultados, y determinar la razón de confiabilidad y las demoras horarias esperadas (*ESDE* y *ESDL*).

Por último, sorprende que existan muy pocos estudios sobre la influencia de variables climáticas en el comportamiento de los usuarios de transporte. No obstante, esta carencia representa una interesante oportunidad de estudio, y el contexto de elección de modo-EHIV parece ser una buena instancia para examinar el impacto de este tipo de variables.

3 GENERACIÓN DEL BANCO DE DATOS

En la presente sección se describe brevemente el proceso de generación de los bancos de datos de preferencias reveladas (PR) y preferencias declaradas (PD) utilizadas para la estimación de modelos de elección de modo y hora de inicio del viaje en esta tesis. Se hace énfasis en la base de PR, pues una parte significativa del tiempo invertido en esta investigación fue destinado a colaborar en su elaboración.

Ambos bancos de datos se originan en una encuesta realizada a comienzos del año 2011 como parte de la tesis doctoral de Julián Arellana, en el Departamento de Ingeniería de Transporte y Logística de la Pontificia Universidad Católica de Chile (Arellana, 2012). La encuesta consistió de tres etapas: (i) entrevista personal, en que se recopiló información socioeconómica y del empleo, y la programación de actividades que los encuestados planeaban realizar el siguiente día laboral (de esta información se obtuvo el horario preferida de llegada o *PAT*); (ii) diario de viajes completado el día hábil siguiente a la realización de la primera etapa, y (iii) dos experimentos de PD, basados en las respuestas a las dos primeras etapas. Las etapas (i) y (iii) se realizaron en los lugares de trabajo de los encuestados mediante entrevistas presenciales, y la etapa (ii) se llevó a cabo completando un formulario web. Del total de 405 encuestados, 308 (76%) respondieron las tres etapas, mientras que 97 personas sólo respondieron las etapas (i) y (ii), es decir, no respondieron la componente de PD.

La Figura 3-1 muestra la ubicación de los hogares y lugares de trabajo de los encuestados. Como es posible observar, los hogares se encuentran distribuidos a lo largo de todo el Gran Santiago, en tanto, los trabajos se ubican principalmente en el centro de la ciudad (comuna de Santiago). La encuesta se enfocó en viajes al trabajo o laborales, pues estos representan cerca del 70% de los viajes en la hora punta mañana (DICTUC, 2006). Para más detalles acerca del diseño de la encuesta se recomienda revisar Arellana (2012) y Arellana *et al.* (2012a).

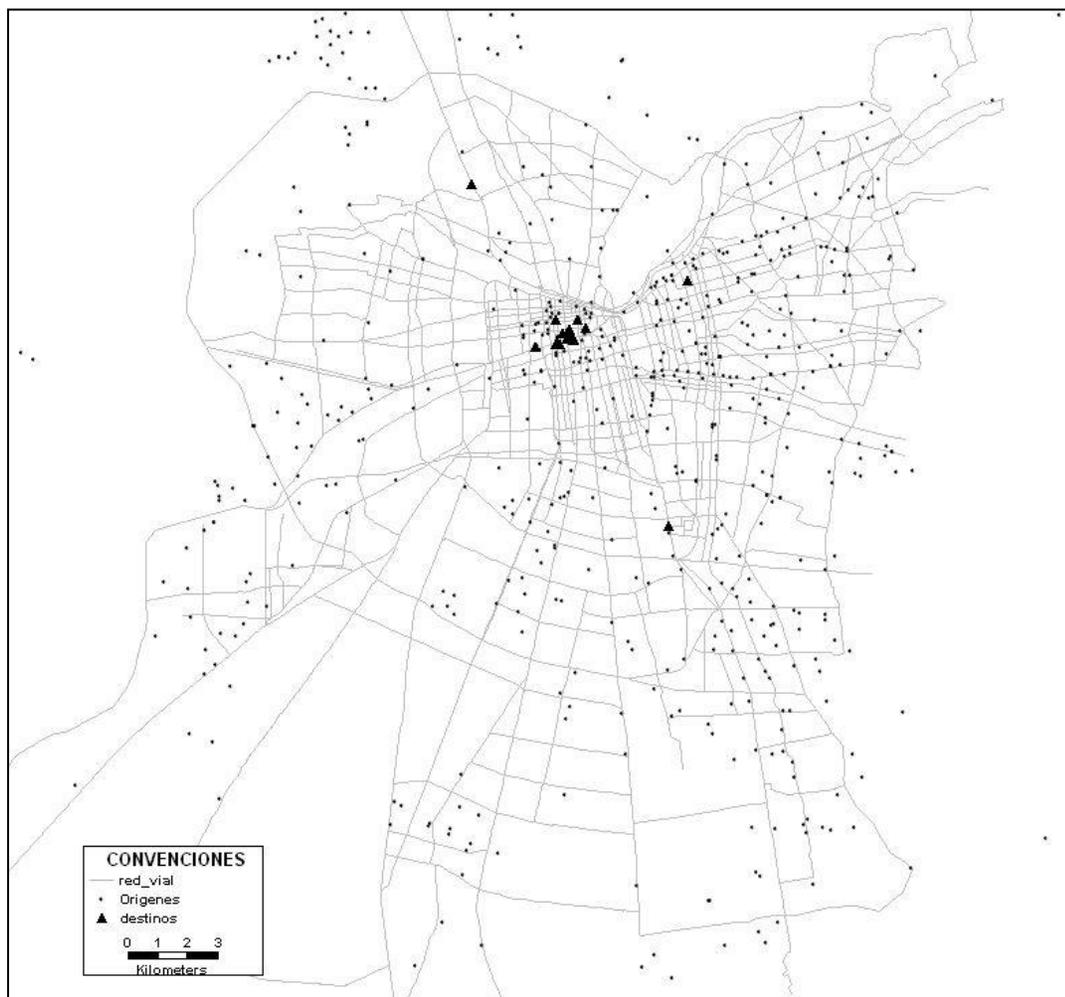


Figura 3-1: Ubicación de los hogares (círculo) y trabajo (triángulo) de los encuestados

Fuente: Arellana (2012)

3.1 Preferencias reveladas (PR)

El banco de datos de PR, correspondiente a modo-hora, fue confeccionado a partir del diario de viajes reportado en la etapa (ii) de la encuesta, pero considerando solamente el viaje de ida al trabajo.

Los principales desafíos de armar el banco de datos de PR consistieron en: (i) depurar, verificar y corregir las elecciones reportadas por los encuestados; (ii) generar los niveles de servicio para el período de tiempo en que se aplicó la encuesta (tercera semana de enero de 2011), para los 12 modos escogidos (auto-chofer, auto-acompañante, caminata-bicicleta, bus, metro, taxi, taxi-colectivo, auto-chofer-metro,

auto-acompañante-metro, bus-metro, taxi-metro y taxi-colectivo-metro), por arco de la red o ruta de transporte público (línea de Transantiago, metro o taxi-colectivo), para diferentes períodos del día y con diferente nivel de agregación temporal (15, 30 y 60 min) y (iii) generar las alternativas de modo y hora para posteriormente asignar la información de niveles de servicio de la red a todos los viajes reportados y a sus alternativas de hora-modo usando Visual Basic de Microsoft Excel.

3.1.1 Niveles de servicio y demoras horarias

A continuación se describen los datos usados como entrada y su procesamiento para generar los niveles de servicio de la red según tipo de modo y hora.

- Tiempo de viaje en vehículo (TVIA): La información de tiempo de viaje varía entre los modos que usan la red de transporte privado (auto, taxi, taxi-colectivo) y las redes de transporte público (Metro y Transantiago).

Los tiempos de viaje en la red de transporte privado se obtuvieron a partir de información de GPS recopilada especialmente para la encuesta. Vehículos instrumentados con dispositivos de GPS circularon por la red vial en variados períodos del día durante la semana en que las encuestas eran llevadas a cabo. La ciudad fue zonificada espacialmente usando las zonas ESTRAUS³ y mediante *software* especializado en redes de transporte y sistemas de información geográfica (TransCad y ArcView), se asignó la información de los GPS (posición geográfica y velocidad según período del día) a estas zonas. La velocidad de una zona y período de tiempo particular, se calculó como el promedio de todas las velocidades reportadas por todos los vehículos que circularon en ese período y zona específica. De esta manera, para cada zona se obtuvo una distribución temporal de velocidades, y a todos los arcos dentro de una determinada zona se les asignó la distribución de su zona. Luego, el tiempo de viaje en un arco se determinó usando la velocidad calculada y la distancia del arco.

Para las redes de transporte público, la información necesaria para obtener los tiempos de viaje fue entregada por Transantiago y Metro. En el caso del Metro, la empresa proporcionó directamente información sobre tiempos de viaje entre todas las

³ Modelo de planificación de sistemas de transporte utilizado en la ciudad de Santiago.

estaciones del sistema para un día completo de operación, desagregado en intervalos de 15 min. Por otro lado, Transantiago facilitó información sobre la posición y velocidad, cada 30 segundos (proveniente de los GPS instalados en los buses), de todos los recorridos de la ciudad durante la semana de la encuesta. Esta enorme cantidad de datos se procesó usando la metodología desarrollada por Arellana *et al.* (2012b), que tiene la particularidad de entregar como salida - para cada recorrido de bus - los tiempos de viaje entre paraderos consecutivos a cualquier nivel de agregación temporal (en esta tesis sólo se usaron 15, 30 y 60 min); esto difiere de otras metodologías que entregan velocidades y tiempos de viaje a un nivel de agregación espacial constante, puesto que la distancia entre paraderos de buses no es constante. Otra de las ventajas de esta metodología, es que se puede implementar usando *software* disponible gratuitamente (JavaScript y GoogleMaps); de hecho, el único *software* comercial usado fue MATLAB y Microsoft SQL Server, que también pueden ser fácilmente remplazados por otros lenguajes de programación gratuitos como Octave y MySQL (Arellana *et al.*, 2012b). Más detalles acerca de esta metodología en Arellana *et al.* (2012b).

- Tiempo de espera (TESP): En el caso del Metro, la empresa también proporcionó información de tiempos de espera en cada estación según línea y sentido por períodos de 15 min de un día laboral. Para los buses se utilizó la misma metodología de Arellana *et al.* (2012b), pues usa la información de intervalos de tiempo entre buses consecutivos que llegan a un paradero para estimar tiempos de espera al nivel de agregación temporal deseado. Sin embargo, la fórmula utilizada sólo considera la oferta de transporte y asume que la demanda, o la llegada de personas a un paradero, distribuye uniforme. Esto conlleva las siguientes limitaciones: (i) no se considera la posibilidad de que un bus se encuentre lleno y por ende la persona tenga que esperar al siguiente (restricción de capacidad); (ii) no se toma en consideración la probabilidad que a una persona le sirvan dos o más recorridos que pasan por el paradero (líneas comunes), y (iii) es posible que la distribución de llegadas no sea uniforme, sino que siga otra distribución (en especial al usar intervalos de tiempo amplios, por ejemplo 60 min), ya que dentro de un mismo intervalo pueden existir períodos de tiempo dónde más gente arribe al paradero debido a restricciones tales como hora de entrada al trabajo. Al no

considerar la restricción de capacidad se tendería a subestimar los tiempos de espera, mientras que al no considerar la posibilidad de que múltiples servicios puedan ser abordados por un pasajero se tiende a sobrestimar el tiempo de espera.

A fin de superar los problemas anteriores, los tiempos de espera por paradero obtenidos mediante la metodología de GPS fueron contrastados con mediciones reales realizadas por DICTUC en diversos paraderos de la ciudad en la misma fecha de la encuesta. Se estimó regresiones lineales entre el tiempo de espera de GPS y el medido por DICTUC para un mismo servicio, paradero y período del día, a fin de obtener un factor de corrección para los tiempos de espera de GPS. Los resultados de las regresiones se encuentran en la Tabla 3-1, desagregados por período (punta y fuera de punta), tipo de servicio (alimentador o troncal), y nivel de agregación temporal (15, 30 y 60 min). Como es posible apreciar, los factores de corrección fluctúan entre 0,3 y 0,7, reflejando el hecho que los tiempos de espera obtenidos vía GPS son consistentemente mayores que las mediciones reales. Esto puede significar que el supuesto que la distribución de llegadas es aleatoria no es correcto y/o que la posibilidad de que un individuo pueda abordar más de un recorrido es más preponderante que la restricción de capacidad de los buses. Esto último tiene sentido, puesto que las encuestas (y mediciones) fueron realizadas en verano, época del año en que los paraderos y buses no están sobrecargados de gente.

En el caso de los taxi-colectivos, el Ministerio de Transporte y Telecomunicaciones suministró información acerca de las frecuencias de todas las líneas del Gran Santiago, y el tiempo de espera se calculó como el inverso de la frecuencia.

Por último, y ante la ausencia de otra información más confiable, los tiempos de espera del modo taxi se calcularon usando promedios de las mediciones realizadas por DICTUC.

- Tiempo de caminata (TCAM): El tiempo de caminata de las alternativas escogidas fue calculado a un nivel de precisión inusualmente alto mediante el software de sistemas de información geográfica (GIS) ArcView. Para las alternativas de caminata, bus y Metro, se usó un código de Javascript para GoogleMaps, también obteniéndose tiempos de caminata a un nivel de exactitud muy alto. En el caso del Metro, también se

agregaron tiempos de acceso y egreso al andén de la estación; en el bus ese nivel de detalle fue posible pues se contaba con la ubicación geográfica de todos los paraderos de la ciudad.

Tabla 3-1: Factores de corrección para tiempos de espera de modo bus

Intervalo	Tipo de Servicio	Troncal		Alimentador	
	Periodo ⁴	Punta	Fuera de Punta	Punta	Fuera de Punta
15	Coficiente	0,39	0,60	0,56	0,63
	R ²	0,61	0,61	0,76	0,67
	Observaciones	540	697	128	175
30	Coficiente	0,34	0,47	0,39	0,47
	R ²	0,60	0,61	0,72	0,61
	Observaciones	395	458	97	112
60	Coficiente	0,38	0,53	0,45	0,56
	R ²	0,70	0,69	0,68	0,69
	Observaciones	256	400	54	88

Fuente: Elaboración Propia.

Para el resto de los modos se utilizó el tiempo de caminata obtenido de la asignación en TransCad, a excepción de las alternativas auto-chofer, auto acompañante y taxi, en los que se usó el promedio de los tiempos de caminata de las elecciones reportadas (según motivo del viaje, comuna de destino y si la persona tenía estacionamiento en el lugar de trabajo), pues TransCad no entregaba esta información.

- Tiempo de transbordo (TTRA): Este tiempo se calculó como la suma del tiempo de caminata y espera de una combinación entre diferentes modos (por ejemplo, transbordo de auto-chofer a Metro) o de combinaciones dentro de un mismo modo (por ejemplo, transbordo entre líneas de Metro). Así, es posible investigar si existen diferencias de percepción en los usuarios entre un transbordo privado-público (por ejemplo, auto a Metro) y un transbordo netamente público (por ejemplo, bus a Metro). Cabe mencionar que los tiempos de espera y caminata de acceso/egreso no se incluyen

⁴ Punta: punta mañana y punta tarde, fuera de punta: punta mediodía, punta nocturna y fuera de punta.

en esta variable, sino que se crearon otras dos variables adicionales: tiempo de espera y tiempo de caminata de acceso/egreso.

- Demoras horarias (*SDE* y *SDL*): Las demoras horarias, o penalidades por llegar antes o después del horario preferido de llegada al trabajo (*PAT*), se calcularon usando las ecuaciones (2.2) y (2.3).

Asimismo, se creó una variable ficticia que busca medir el efecto de llegar tarde al trabajo, independiente de cuantos minutos se llegue atrasado. La variable se denominó *D_TARDE*, y toma valor uno si es que la persona llegaba después del horario oficial de entrada al trabajo (*HOE*), o cero en caso contrario. Esto difiere del modelo *SM* de Small, ya que en general éste usa *PAT* y no *HOE* para definir esta variable ficticia.

- Número de paradas intermedias (*PARAD_INT*): Dado que se tenía información respecto a todos los viajes de los encuestados (no sólo el viaje al trabajo), se creó una variable asociada a la cantidad de paradas intermedias que una persona realizó en su viaje de ida al trabajo. El principal motivo de estas paradas es ir a buscar/dejar a alguien/algo, e imponen una restricción en la programación de actividades.

- Costo (*COST*): El costo del modo auto-chofer se determinó como la suma de: (i) un costo de operación vehicular (gasolina y mantenimiento de \$85/km, cuyo valor fue entregado por DICTUC) que se multiplicó por la longitud de la ruta para obtener el costo de operación total; (ii) peajes dependientes de la hora en el caso de usar autopistas urbanas, y (iii) costo de estacionamiento reportado por el encuestado.

En el caso de transporte público masivo, se asignaron las tarifas de Transantiago y del Metro, esta última según la hora a la que se accedió a la estación ya que la tarifa de este modo varía según horario (horarios punta, valle y bajo). Los costos adicionales, por concepto de transbordar de bus a Metro, también fueron considerados.

Para el taxi-colectivo se estimó un costo dependiente de la distancia, replicando así el mecanismo de cobro de este modo. Por último, el costo del taxi se calculó como un monto por iniciar el servicio (\$250) más un monto por kilómetro (\$500/km).

- Densidad de Metro (*DENSM*): La densidad en los vagones de Metro se mide como la cantidad de pasajeros por metro cuadrado (pax/m^2). La empresa entregó esta información para la fecha de la encuesta, desagregada por línea, sentido y períodos de 15

min, la cual fue posteriormente agregada a los intervalos de tiempo más amplios (30 y 60 min). Asimismo, si una persona utilizó múltiples líneas de Metro, la densidad final se calculó como el promedio ponderado entre la densidad y tiempo de viaje en cada línea.

En la Figura 3-2, se muestra la densidad de la Línea 5 dirección sur-norte (Vicente Valdés a Pudahuel⁵) según distintos niveles de agregación temporal (15, 30 y 60 min). Esto tiene por objeto ilustrar como se pierde información – la curva se vuelve más escalonada o menos suave – a medida que se usan intervalos de tiempo más amplios. Cabe recordar que uno de los objetivos de esta tesis es investigar las ventajas y desventajas de usar alternativas de tiempo de diferente longitud: a medida que aumenta la longitud de los intervalos de tiempo se tienen menos alternativas y en teoría menos correlación entre ellas, pero la información usada para la estimación del modelo es menos precisa.

Por último, cabe mencionar que como la encuesta se realizó en enero, época del año en que algunos trabajadores y estudiantes están de vacaciones, los valores de densidad utilizados para generar el banco de datos (y presentados en la Figura 3-2) son más bajos que en otros meses del año.

- Temperatura (TEMP): La temperatura en Santiago en la fecha de la encuesta, se obtuvo de registros correspondientes a tres estaciones de medición de la calidad del aire (Pudahuel, Parque O'Higgins y Las Condes). En base a estos datos se calculó la temperatura promedio semanal (para la semana de la encuesta) para cada estación y para todos los períodos del día usando intervalos de 15, 30 y 60 min; luego, a cada viaje se le asignó la temperatura de la estación más cercana a su coordenada de origen.

3.1.2 Generación de alternativas de modo y hora

Mientras que las rutas (transporte privado, público y caminata) de las elecciones modales se generaron de acuerdo a las respuestas de los encuestados, las alternativas de modo auto (chofer y acompañante), taxi, taxi-colectivo y combinaciones de estos tres con Metro, se crearon mediante TransCad usando como criterio la ruta de menor tiempo de viaje. Las alternativas de transporte público masivo (bus, Metro y bus-Metro) para

⁵ La actual extensión de la Línea a 5 a Maipú no estaba en operación en la fecha de la encuesta.

cada viaje reportado, se buscaron en la página web de Transantiago (www.transantiago.cl) eligiendo las alternativas de menor tiempo de viaje.

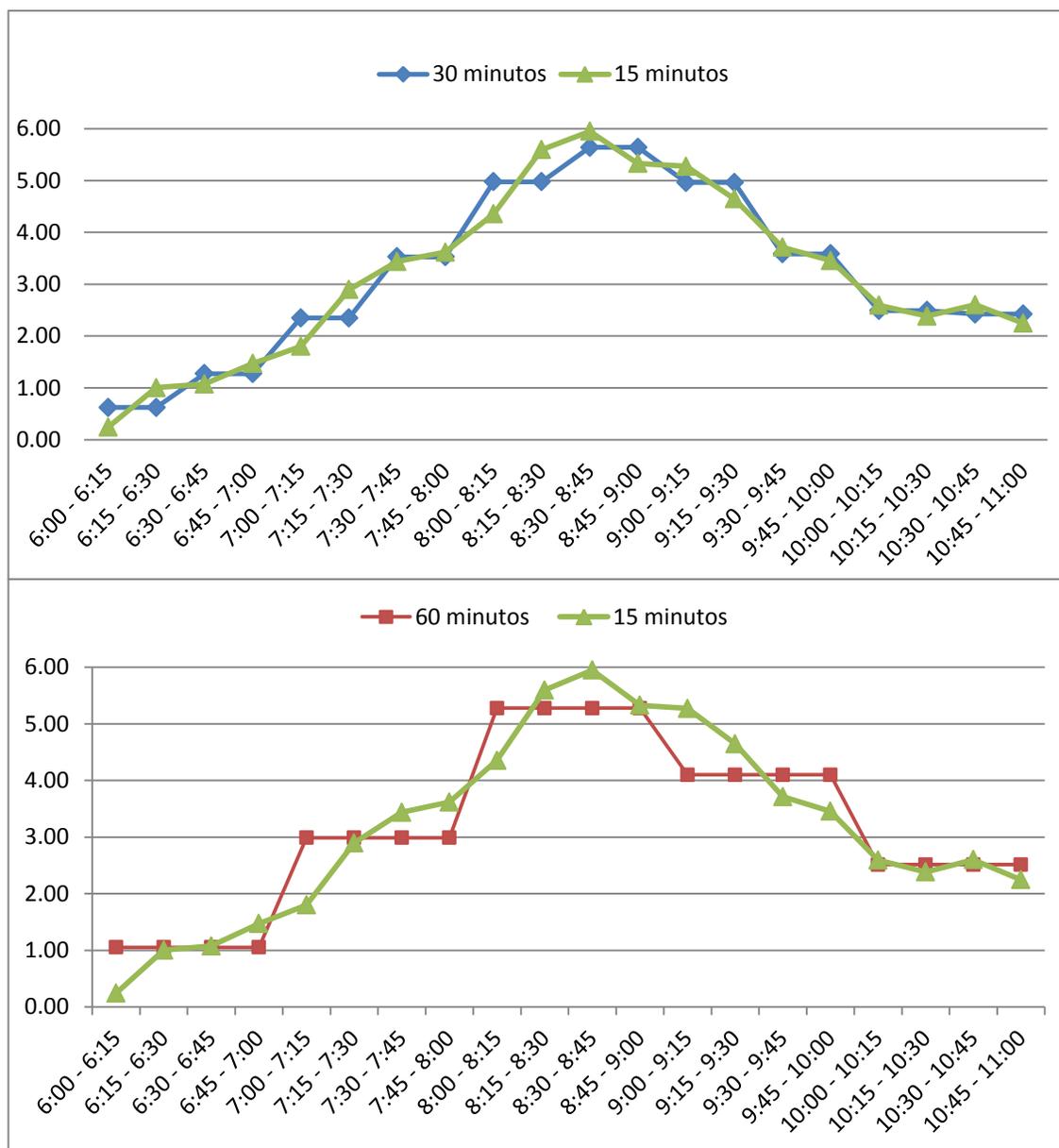


Figura 3-2: Densidad (pax/m²) en línea 5 de metro dirección sur-norte

Fuente: Elaboración propia en base a información de Metro de Santiago

Las alternativas de caminata se crearon con la aplicación JavaScript de GoogleMaps, definiendo como cota máxima, para que este modo estuviese disponible,

que la distancia de la ruta fuese menor a 4,5 km y tomando en consideración que personas de distinto género y edad tienen diferentes velocidades de caminata.

En relación a las alternativas de hora de inicio del viaje, se consideraron tres bancos de datos: desagregado (PR_15), intermedio (PR_30) y agregado (PR_60). Para el banco de datos más desagregado se consideraron ocho períodos de 15 min entre las 7:00 y 9:00 de la mañana, resultando en un total de 11 alternativas de hora (5:00 – 7:00, intervalos de 15 min, 9:00 – 11:00 y 11:00 – 14:00), mientras que la intermedia considera cuatro intervalos de 30 min entre las 7:00 y 9:00 (y un total de siete alternativa horarias), y la agregada considera dos intervalos de una hora entre las 7:00 y 9:00, con un total de cinco alternativas de hora. Se consideró que era mejor mantener fijos los intervalos fuera del período 7:00 y 9:00, debido a que pocas personas iniciaban sus viajes en esos intervalos y desagregarlos hubiese complejizado el problema.

Para crear las alternativas de hora se tomó como referencia la hora de inicio del viaje escogida por cada encuestado, y a ésta se le sumó y restó 15, 30 o 60 min (dependiendo del banco de datos a generar) todas las veces que fuese necesario, hasta cubrir el espectro completo de alternativas. Para crear el banco de 15 min, por ejemplo para una persona que inició su viaje a las 8:50, la alternativa de hora entre 9:00 – 11:00 se determinó como iniciar el viaje a las 9:05 (desechando salir a las 9:20, 9:35, etc.) y la alternativa 11:00 – 14:00 quedó definida como iniciar el viaje a las 11:05, también descartándose horarios posteriores. Siguiendo el mismo ejemplo, para los bancos de datos de 30 y 60 min, la alternativa 9:00 – 11:00 quedó definida como salir a las 9:20 y 9:50, respectivamente.

Con el fin de disminuir el número de alternativas y la complejidad del problema, los modos auto-chofer-Metro, auto-acompañante-Metro, taxi-Metro y taxi-colectivo-Metro se agruparon en uno sólo, llamado “combinación-Metro” (se dejó fuera a bus-Metro pues posee características diferentes y una alta partición modal). Al codificar los datos de esta alternativa compuesta para cada individuo (cuando era parte del conjunto de alternativas disponibles), se le asignó el nivel de servicio de la alternativa individual que primero estuviese disponible, de acuerdo al siguiente orden jerárquico: 1) auto-chofer-Metro, 2) taxi-colectivo-Metro, 3) auto-acompañante-Metro y 4) taxi-Metro. De

esta forma el número de modos se estableció en nueve y el número de alternativas totales (modo y hora) resultó ser 99, 63 y 45 para los bancos de datos de 15, 30 y 60 min, respectivamente.

Para simplificar el proceso de asignar niveles de servicio de la red a las rutas de los modos escogidos y alternativos, cada ruta fue codificada como una secuencia de etapas; estas podían corresponder a caminata (acceso, egreso o transbordo) o a tiempo de viaje en vehículo, y estaban compuestas por un origen, destino (paradero o estación de transporte público, hogar, trabajo, u arco de la red), y la hora en que comenzaba la etapa. De esta forma se logró asignar el tiempo de viaje, espera y costo de forma dinámica, ya que la hora se actualizaba luego de transcurrida cada etapa.

3.2 Preferencias declaradas (PD)

Como ya se señaló, se presentó dos experimentos de PD a cada encuestado con el fin de evaluar el efecto de dos estrategias de gestión de demanda de transporte que no están en operación actualmente en Santiago (tarificación vial y sistema de horarios flexibles para la entrada y salida del trabajo). Del total de 405 encuestados originales, 308 respondieron la componente PD; los individuos que habían caminado o usado la bicicleta para ir al trabajo y aquellos que cambiaban de modo entre los viajes de ida y regreso del trabajo, fueron excluidos para simplificar el diseño experimental (Arellana, 2012).

El primer experimento de PD estaba enfocado a los viajes de ida hacia al trabajo en la punta mañana. A cada encuestado se le presentaron cinco situaciones hipotéticas de elección, y cada una estaba compuesta por cuatro alternativas y sus respectivos atributos. Tres de estas alternativas correspondían a viajar en el mismo modo, pero en diferentes horarios: (i) hora similar a la reportada en la componente PR; (ii) salir antes, y (iii) salir después de la hora reportada; la cuarta alternativa consistía en viajar en otro modo, pero en un horario similar al escogido. Para aquellas personas que habían viajado en transporte público, el modo alternativo era transporte privado en caso de tener automóvil y licencia de conducir, o transporte público tipo taxi-colectivo en caso contrario; para los que escogieron transporte privado, el modo alternativo era transporte público.

Los atributos de cada alternativa eran el tiempo de viaje normal (similar al experimentado por el encuestado en la realidad), su hora de inicio (a partir de ambos se calculaba la hora de llegada al trabajo), un tiempo de viaje mayor que ocurría máximo una vez a la semana (y provee una medida de la variabilidad de los tiempos de viaje en la red), el costo monetario (en pesos) y la comodidad del viaje; este último atributo sólo aparecía en el caso de alternativas de transporte público y consideraba la ocupación del vehículo y si se debía viajar de pie o sentado. La variabilidad del tiempo de viaje se presentó de esa manera, y no de otras formas más comunes en la literatura (por ejemplo, mediante distintos tiempos de viaje, cada uno con una probabilidad asociada), para evitar que los encuestados se centrasen en esa variable (efecto observado en grupos focales), pues el objetivo prioritario del experimento era estudiar las elecciones horarias (Arellana, 2012). La Figura 3-3 muestra un ejemplo de situación de elección presentada a un usuario de transporte público. De ella se pueden apreciar los distintos atributos que componen las cuatro alternativas del experimento PD, de las cuales el encuestado debía escoger sólo una.

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"/>



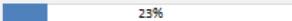



Figura 3-3: Ejemplo de situación de elección de primer experimento PD presentado a usuario de transporte público

Fuente: Arellana (2012)

El segundo experimento estaba enfocado a los *tours* compuestos por los viajes de ida y regreso al trabajo. Como en esta tesis se modelan viajes al trabajo y no *tours*, este segundo experimento no se utilizó en nuestro banco de datos de PD.

El diseño experimental realizado por Arellana (2012) hizo los siguientes aportes a la literatura: (i) incluyó la dependencia entre los niveles de los distintos atributos de una misma alternativa (el valor del costo y tiempo de viaje dependen de la hora de inicio del viaje), y (ii) toma en cuenta el problema que las variables presentadas a los encuestados no son las mismas que se utilizan en la modelación. Para acomodar todo esto (y, además, lograr un diseño “a la medida” basado en el viaje reportado en los datos de PR), se debió realizar un diseño eficiente optimizado en dos etapas. Para más detalles acerca del diseño de las componente PD revisar Arellana *et al.* (2012a).

4 ANÁLISIS DEL BANCO DE DATOS

En el presente capítulo se resumen las estadísticas descriptivas de ambos bancos de datos: preferencias reveladas (PR) y preferencias declaradas (PD). Luego de depurar los viajes que presentaban inconsistencias o que tenían una duración muy corta por la cercanía entre el origen y el destino, el banco de datos de PR quedó compuesto por 468 observaciones, ya que al total de 405 encuestados originales se agregaron observaciones provenientes de la encuesta piloto. Por otra parte, el banco de datos de PD contiene 1.505 observaciones (301 individuos por cinco situaciones de elección cada uno), al descartarse las respuestas de siete individuos que habían reportado horarios preferidos de llegada al trabajo (*PAT*) fuera de su rango de flexibilidad en el horario de llegada al trabajo. Asimismo, hay 28 individuos que se encuentran en el banco PD, pero no así en el de PR.

4.1 Información socioeconómica y de empleo

El análisis a continuación se realizó a partir de las 468 personas del banco de datos de PR. El promedio de edad de los encuestados es 39 años y su distribución es: 21% menores de 30 años, 66% entre 30 y 55 años, y el 13% restante mayores a 55 años. Esto significa que una mayoría de los encuestados se encuentra en plena etapa laboral. En cuanto al género, este se reparte equitativamente entre mujeres (49,6%) y hombres (50,4%). Por otro lado, de las 354 (76%) personas que poseen licencia de conducir, 89% tienen uno o más autos en el hogar, lo que implica en un principio que un 67% del total de encuestados tiene disponible el modo auto-chofer.

Los hogares están compuestos en promedio por 3,3 personas, lo que está línea con el valor reportado en la encuesta suplementaria de ingresos 2009 del INE⁶ (3,4); sólo un 17% de los encuestados tiene un grupo familiar de cinco o más personas. El número promedio de autos por hogar es de 1,1, mientras que el promedio de licencias de conducir asciende a 1,7; por ende, en promedio existen más licencias que autos por

⁶ Disponible en:
www.ine.cl/canales/menu/publicaciones/calendario_de_publicaciones/pdf/01_10_10/completaingresos09.pdf

hogar. Ahora, la disponibilidad de auto de un hogar está sujeta al número de licencias en éste; por ejemplo, si un hogar tiene cinco autos, pero sólo una licencia de conducir, entonces la cantidad de autos disponibles para ser usados es sólo uno; en el caso contrario, si un hogar tiene un solo auto y cinco licencias, como puede ser usado por cinco personas su disponibilidad real es menor que uno. Por estos motivos se creó la siguiente variable:

$$AUTLIC = \text{Min} \left\{ \frac{\text{Número de autos por hogar}}{\text{Número de licencias por hogar}}, 1 \right\} \quad (4.1)$$

El promedio de la variable AUTLIC es de 0,57, lo que es consistente con el hecho de que existen más licencias que autos por hogar. Por otra parte, el 22% de las personas debe acompañar a otras personas del hogar a otros lugares, lo que implica una restricción en su programación diaria de actividades.

En relación a los niveles de escolaridad los datos muestran que una sola persona no completó la educación media, 9% sólo finalizó la educación secundaria, 20% culminó estudios superiores de índole técnica, la cantidad de egresados de universidades asciende a un 64% y el 7% restante tiene estudios de postgrado. Llama la atención el alto porcentaje de encuestados con estudios universitarios (y de postgrado), revelando que esta muestra no es representativa de los habitantes de la ciudad. Esto puede estar relacionado con cómo se escogieron los centros de empleo donde se realizaron las encuestas. Estos tuvieron como característica que Julián Arellana, alguno de sus profesores supervisores o la Universidad Católica, tuvieran relación directa con al menos un miembro de su departamento administrativo (Arellana, 2012), aunque los encuestados finales fueron seleccionados aleatoriamente dentro de cada centro de empleo. Para más detalles acerca del proceso de selección de la muestra se recomienda al lector revisar Arellana (2012).

Los altos niveles de escolaridad también se traducen en mayores ingresos. La media y desviación estándar del ingreso líquido mensual por persona (ILM) es de \$1.041.763 y \$558.820, respectivamente. El primer valor es casi tres veces mayor que el

ingreso medio mensual que recibe un santiaguino⁷, mientras que el segundo demuestra que existe mucha varianza entre los ingresos de los encuestados. En la Figura 4-1 es posible observar que casi el 50% de los ingresos se concentra en el tramo entre 400 mil y un millón de pesos mensuales, y que también existe un porcentaje considerable (más del 20%) que gana entre 1,5 y 2 millones de pesos mensuales. Esto último claramente no se ajusta a la realidad de la ciudad, por lo que se puede afirmar con certeza que esta es una muestra de ingresos medios-altos.

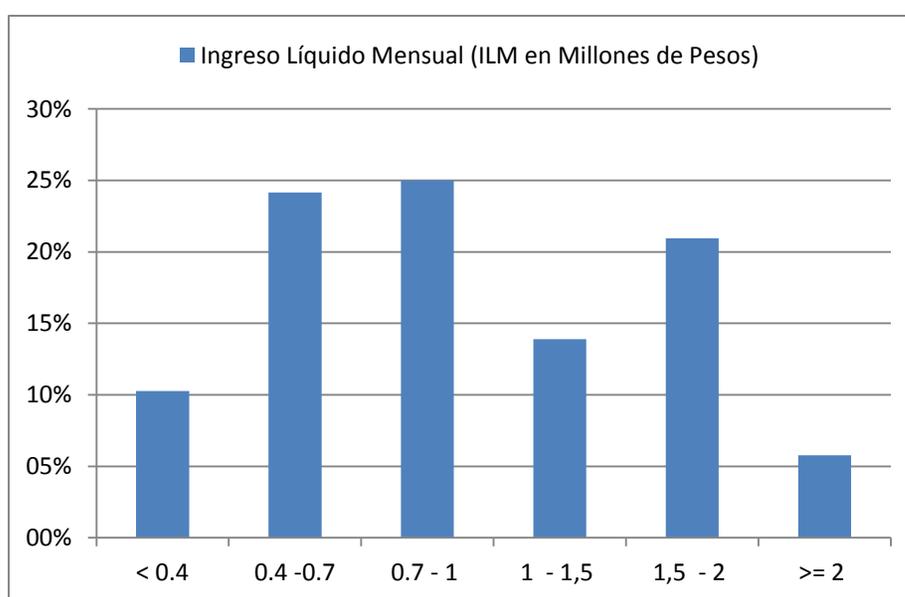


Figura 4-1: Distribución porcentual del ingreso líquido mensual (ILM)

Fuente: Elaboración Propia

El promedio de horas trabajadas por semana es 45,6, con una desviación estándar de sólo cuatro horas. Esto implica que la mayoría de los encuestados trabaja a tiempo completo y con turnos de trabajo entre 8 y 10 horas diarias. A partir de las horas de trabajo semanales es posible calcular la tasa salarial (w), esto es, el ingreso por hora de trabajo. La media de la tasa salarial es 3.683 (\$/hora), lo que supera la remuneración

⁷ Según la encuesta suplementaria de ingresos 2009 del INE, el ingreso promedio mensual de todas las personas ocupadas en la Región Metropolitana, para el último trimestre de 2009, era \$ 378.910 (en el caso de los asalariados este valor asciende a \$ 398.127). Datos disponibles en: www.ine.cl/canales/menu/publicaciones/calendario_de_publicaciones/pdf/01_10_10/completaingresos09.pdf

horaria a nivel nacional (2.730 \$/hora)⁸, y su desviación estándar es \$2.044, lo que reafirma la variación existente en los ingresos.

Por otra parte, la Figura 4-2, muestra la distribución de las horas oficiales de entrada al trabajo (*HOE*) y las horas preferidas de llegada al trabajo (*PAT*) declarada por los encuestados. De esta figura es posible apreciar que la distribución de *HOE* y *PAT* es bastante similar, lo que implica que las personas prefieren llegar cerca de la hora oficial en que comienza el trabajo. Asimismo, tanto *HOE* como *PAT* se concentran entre 8:30 y 9:00 (8:30 pertenece al intervalo 8:30-8:45), aunque hay cerca de 50 individuos que deberían y prefieren llegar a las 8:00. Así, existen algunas personas que comienzan a trabajar muy temprano por la mañana (cerca de las 6:00) y otros a la hora de almuerzo (13:30).

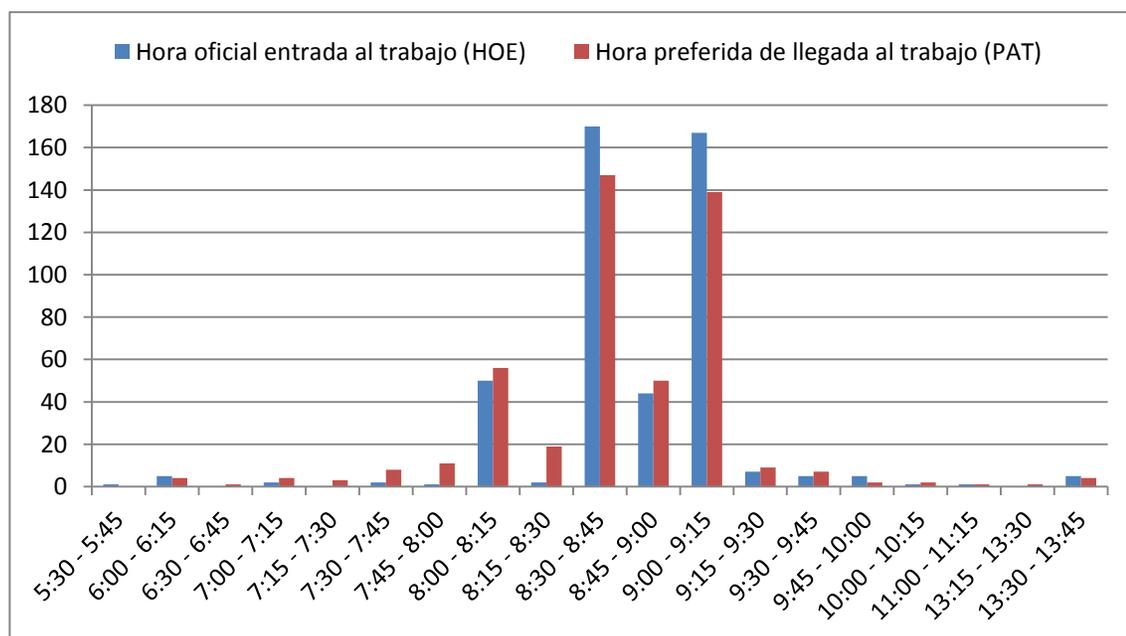


Figura 4-2: Distribución de hora oficial de entrada al trabajo (*HOE*) y hora preferida de llegada al trabajo (*PAT*).

Fuente: Elaboración propia

⁸ Valor para Diciembre de 2010 del informe “Remuneraciones y Costos Medios por Hora 2010” que el INE publica anualmente. Disponible en: www.ine.cl/canales/chile_estadistico/mercado_del_trabajo/remuneraciones/anuarios_estadisticos/pdf/RemuneracionesMediasyCostosMedios_informe_anual_2010.pdf

La flexibilidad laboral, en este caso definida como los minutos que es posible llegar antes o después de la hora oficial de entrada al trabajo (*HOE*), es de suma importancia para los modelos de elección horaria; muestra qué tanto poder de decisión efectivamente tiene un trabajador para atrasar o adelantar su viaje al trabajo. La Figura 4-3 muestra como se distribuye este indicador en la muestra; notar que existe un porcentaje considerable de la muestra que no puede llegar más de 15 min temprano o tarde al trabajo. Por otro lado, notar también que menos del 20% de los individuos puede llegar dos o más horas tarde o temprano. Esto demuestra la rigidez laboral horaria que tiene gran parte de los encuestados, aunque se puede notar que existe mayor flexibilidad para llegar temprano que tarde (más del 50% puede llegar hasta 30 min temprano, en comparación al 37% para llegar tarde).

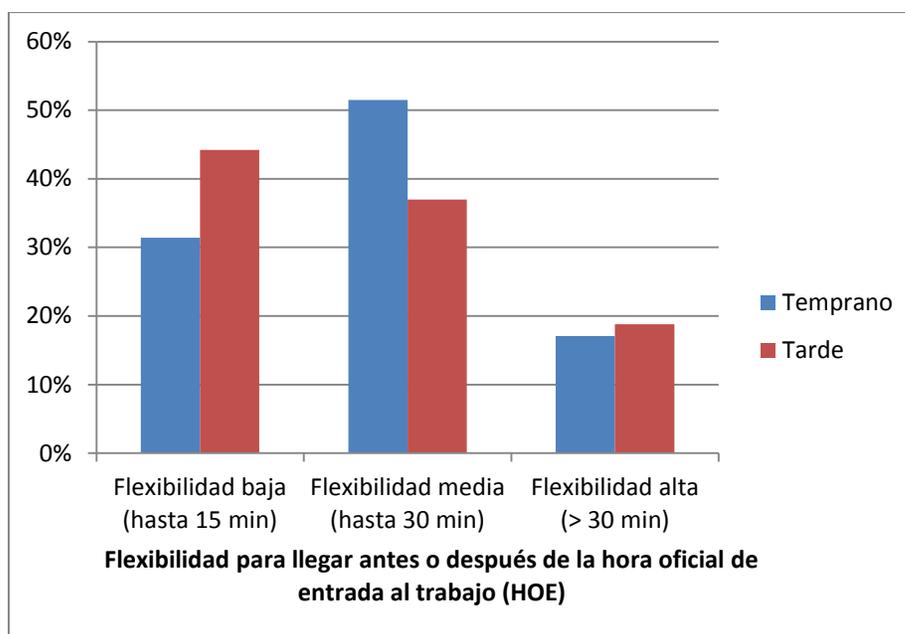


Figura 4-3: Distribución de la flexibilidad laboral para llegar antes o después de la hora oficial de entrada al trabajo (*HOE*)

Fuente: Elaboración propia

4.2 Banco de datos de preferencias reveladas (PR)

A continuación se describen las alternativas escogidas por los encuestados (modo de transporte y hora de inicio del viaje), y sus correspondientes variables de nivel de

servicio (tiempo de viaje, caminata, etc.) y de programación de actividades (*SD* o demoras horarias).

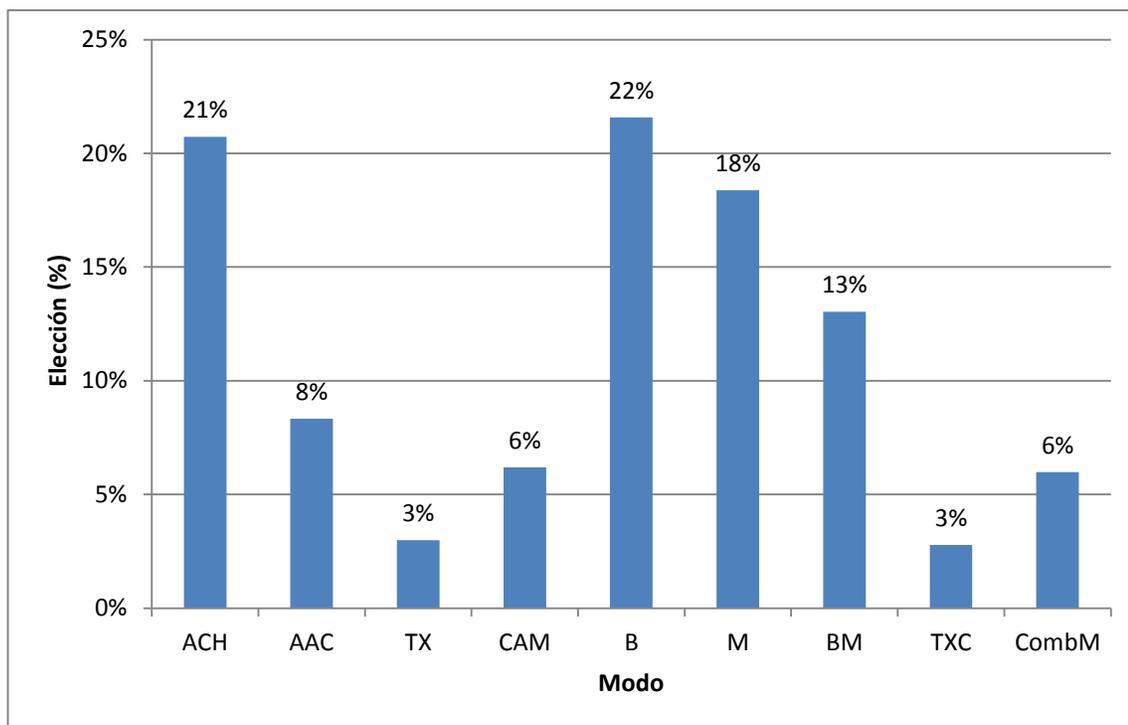
4.2.1 Modo de transporte

La distribución de las elecciones de modo, o partición modal, es un elemento de suma importancia para la toma de decisiones en planificación de transporte. Grandes inversiones en infraestructura (por ejemplo, una nueva línea de Metro) o en operación de sistemas de transporte público (como mayor frecuencia de transporte público), se hacen con el fin de que los usuarios de público no se cambien al transporte privado.

La Figura 4-4 muestra la partición modal de los encuestados. En el transporte privado predomina el modo “auto como chofer” con un 21%, seguido por “auto-acompañante” (8%), mientras que sólo 3% viaja al trabajo en taxi. Asimismo, más del 50% de los encuestados usa transporte público masivo (bus, Metro y combinación bus-Metro), en tanto sólo 3% viaja en taxi-colectivo. Por último, el 6% de las personas camina al trabajo y otro 6% usa alguno de los modos que combinan transporte privado o público no masivo (i.e. taxi-colectivo) con Metro. Estos resultados son consistentes con la partición modal observada en Santiago post-Transantiago (Yañez *et al.*, 2010).

En general interesa analizar la relación existente entre la partición modal y la disponibilidad de los distintos modos pues ésta condiciona las elecciones. Por ejemplo, gran parte de la decisión de una persona cautiva del bus, esto es, en que sus únicas alternativas disponibles son bus y/o bus-metro, se determina por su falta de alternativas. En la Tabla 4-1 se puede observar el cociente entre las personas que eligen un modo y su disponibilidad, que entrega un indicio del impacto de la disponibilidad en la elección final. El Metro es el modo más escogido cuando está disponible; esto parece lógico, pues es un modo que a pesar de estar congestionado en los períodos punta, tiene una confiabilidad (en tiempos de viaje y de espera) muy alta. Su baja disponibilidad (la más baja junto al modo caminata) se debe a su accesibilidad, ya que para poder usar Metro hay que vivir y trabajar cerca de una estación de Metro. Otros modos que tienen cocientes relativamente altos son el auto-chofer y la caminata, seguidos por el bus-Metro y, en menor medida, el bus. Los modos que son poco atractivos, a pesar de su

relativamente alta disponibilidad, son el auto como acompañante, taxi, taxi-colectivo, y el modo combinación con Metro.



ACH: auto-chofer; AAC: auto-acompañante; CAM: caminata; B: bus; M: Metro; BM: bus- Metro; TX: taxi; TXC: taxi-colectivo; y CombM: modos que realizan combinación con Metro.

Figura 4-4: Distribución de las elecciones modales

Fuente: Elaboración propia

Tabla 4-1: Elección modal versus Disponibilidad modal

Modo	Elección (%)	Disponibilidad (%)	Elección/Disponibilidad (%)
ACH	21	67	31
AAC	8	60	14
TX	3	84	4
CAM	6	19	33
B	22	96	22
M	18	31	59
BM	13	51	26
TXC	3	38	7
CombM	6	60	10

Fuente: Elaboración propia

La elección modal no sólo se ve afectada por variables de nivel de servicio y disponibilidad de modos, sino que también depende de la flexibilidad que cada modo otorga para realizar diversas actividades durante un día. En este caso, de las 10 personas que realizan paradas intermedias en el viaje hacia al trabajo, nueve utilizan el modo auto como chofer y una sola usa Metro. Este resultado es esperable ya que el modo auto-chofer permite mayor flexibilidad al no depender de otras personas (como el modo autoacompañante) ni es necesario tomar diversas rutas de transporte público para llegar a múltiples destinos.

4.2.2 Hora de inicio del viaje

En la Figura 4-5 se presenta la distribución de horas de inicio y llegada de los viajes. Ambas distribuciones tienen una cota máxima, están desfasadas por la duración de los viajes y, aunque ambas tienen una forma similar, la curva de llegadas es más suave que la de salidas, en especial en el intervalo entre 7:15 y 7:45 donde ésta última sufre una discontinuidad. Asimismo, ninguna de las dos parece ser perfectamente asimétrica respecto a su valor máximo.

La mayoría de los encuestados inicia su viaje entre las 7:30 y 8:30 (el máximo de salidas ocurre entre las 8:00 y 8:15), en tanto, las horas más comunes de llegada al trabajo son entre las 8:00 y 9:00, con el máximo de llegadas entre las 8:30 y 8:45. Según estos resultados, el horario punta mañana (PM) sería entre las 7:30 y 9:00, lo que es consistente con las curvas de flujo vehicular derivadas de conteos de tráfico, en que el período PM generalmente se establece entre las 7:00 y 9:00.

La distribución de las horas de inicio de los viajes cambia al usar intervalos de tiempo de distinta longitud. En la Tabla 4-2 se encuentra esta distribución para los tres bancos de datos con distintos intervalos de tiempo que se generaron: 15, 30 y 60 min. El de 30 min parece tener una distribución más asimétrica y suave que el de 15 min, lo que se puede deber a que las personas tienden a aproximar los valores reportados. En otras palabras, una persona que salió del hogar a las 7:07, pudo haber reportado las 7:00, 7:05 o 7:10; no poder acomodar errores en la aproximación de las respuestas es otro de los

problemas asociados a utilizar intervalos de tiempo muy desagregados (además de la correlación entre alternativas).

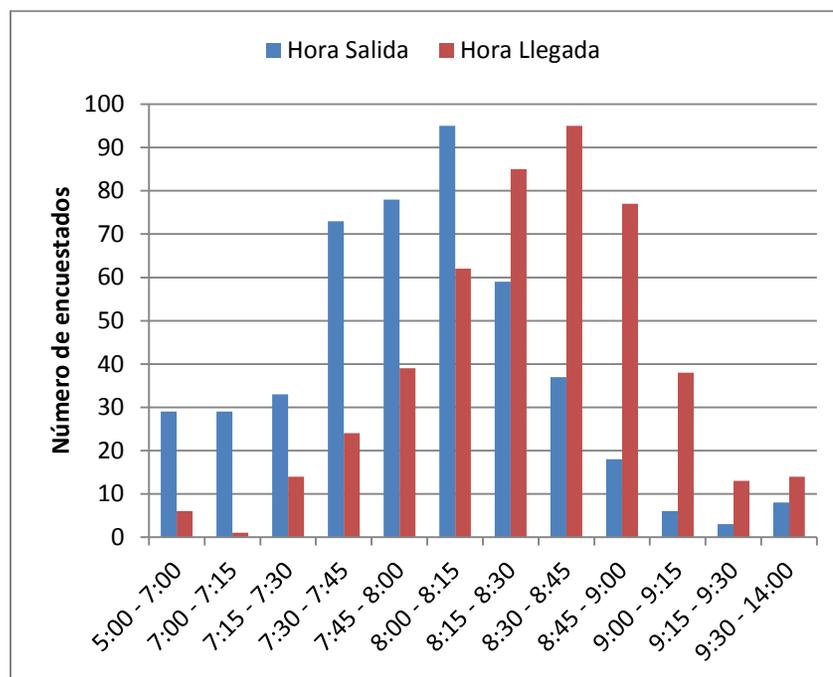


Figura 4-5: Distribución de las horas de inicio y de llegada del viaje al trabajo

Fuente: Elaboración propia

Tabla 4-2: Distribución de horas de inicio del viaje según tipo de desagregación

15 min		30 min		60 min	
5:00 - 7:00	6%	5:00 - 7:00	6%	5:00 - 7:00	6%
7:00 - 7:15	6%	7:00 - 7:30	13%	7:00 - 8:00	46%
7:15 - 7:30	7%	7:30 - 8:00	32%	8:00 - 9:00	45%
7:30 - 7:45	16%	8:00 - 8:30	33%	9:00 - 11:00	3%
7:45 - 8:00	17%	8:30 - 9:00	12%	11:00 - 14:00	1%
8:00 - 8:15	20%	9:00 - 11:00	3%		
8:15 - 8:30	13%	11:00 - 14:00	1%		
8:30 - 8:45	8%				
8:45 - 9:00	4%				
9:00 - 11:00	3%				
11:00 - 14:00	1%				

Fuente: Elaboración propia

Como entre los objetivos de esta tesis se encuentra estudiar el efecto de incorporar la flexibilidad laboral de dos formas distintas (vía la disponibilidad de alternativas o como variable explicativa del modelo), la Figura 4-6 muestra la distribución de la disponibilidad de horarios de llegada al trabajo utilizando intervalos de tiempo de 15 min. Estos valores se calcularon usando la información de la hora oficial de entrada al trabajo (*HOE*) y de la flexibilidad para llegar antes o después de *HOE*. Como era de esperar, más del 70% de los encuestados puede llegar al trabajo entre las 8:00 y 9:15 y menos del 20% puede hacerlo antes de las 7:30 o después de las 10:00.

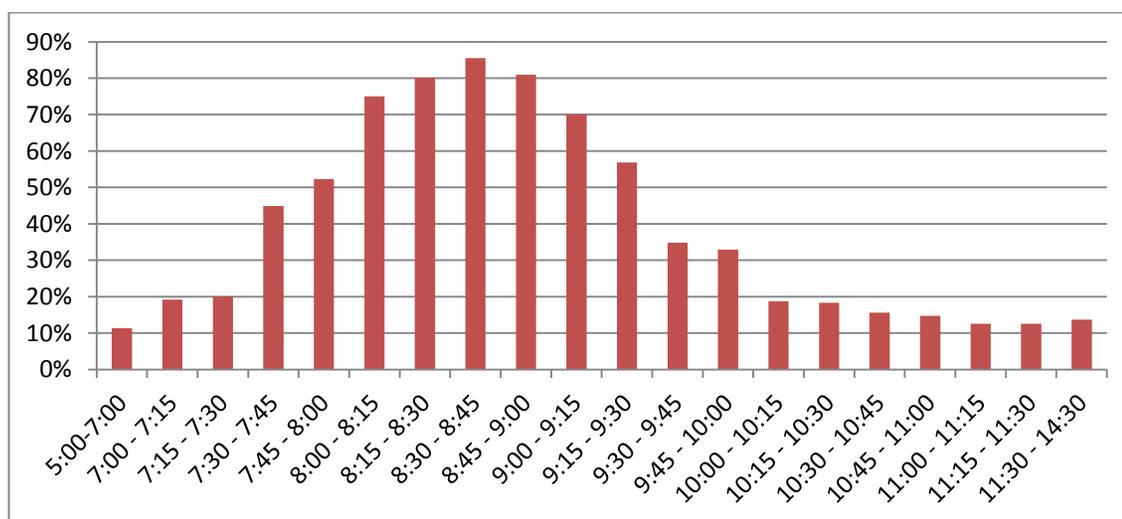


Figura 4-6: Histograma de disponibilidad de horarios de llegada al trabajo

Fuente: Elaboración propia

4.2.3 Niveles de servicio y demoras horarias de elecciones

En primer lugar, tal como se aprecia en la Tabla 4-3, la mayoría de los viajes tiene una duración total entre 15 min y una hora (tan sólo 8% son mayor a una hora y 9% menor a 15 min). Cabe mencionar que los resultados presentados en este apartado se confeccionaron usando el banco de datos de 15 min.

Tabla 4-3: Distribución de la duración de los viajes

< 15 min	15-30 min	30-45 min	45-60 min	1- 1.5 horas	>1.5 horas
9%	32%	32%	20%	7%	1%

Fuente: Elaboración propia

La Figura 4-7 muestra que el tiempo de viaje promedio de auto-acompañante (AAC), auto-chofer (ACH) y bus (B) es de 25 min con una desviación estándar entre 10 y 15 min. Por otro lado, los modos bus-Metro (BM) y combinación con Metro (CombM) promedian un tiempo de viaje de 30 min pero menor varianza; finalmente, los valores promedio para Metro (M), taxi (TX) y taxi-colectivo (TXC) son 17, 12 y 21 min respectivamente.

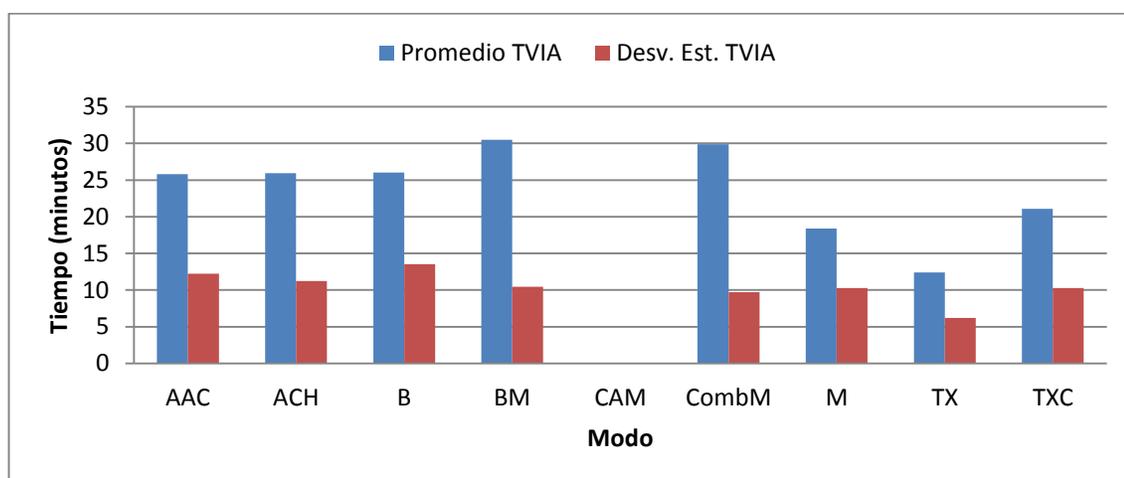


Figura 4-7: Promedio y desviación estándar del tiempo de viaje según modo de transporte

Fuente: Elaboración propia

En la Figura 4-8 se presenta el promedio de los tiempos de caminata (acceso y egreso), el tiempo de espera de acceso y el tiempo de transbordo para todos los modos. Los tiempos de caminata son consistentemente mayores que los de espera y transbordo, y los modos caminata (CAM), Metro y taxi-colectivo presentan los mayores valores (16, 14 y 11 min respectivamente). Esto era esperable, pues el modo caminata sólo involucra tiempo de caminata, mientras que los otros dos modos tienen menor cobertura espacial y se debe caminar más para acceder a ellos. Los usuarios de bus y bus-Metro caminan en promedio 9 min y el tiempo de caminata del modo auto chofer (en promedio 4 min) se debe a la necesidad de caminar desde dónde se estaciona el vehículo hacia el trabajo (en general, pocas oficinas cuentan con estacionamiento directamente debajo de su edificio en el centro de Santiago).

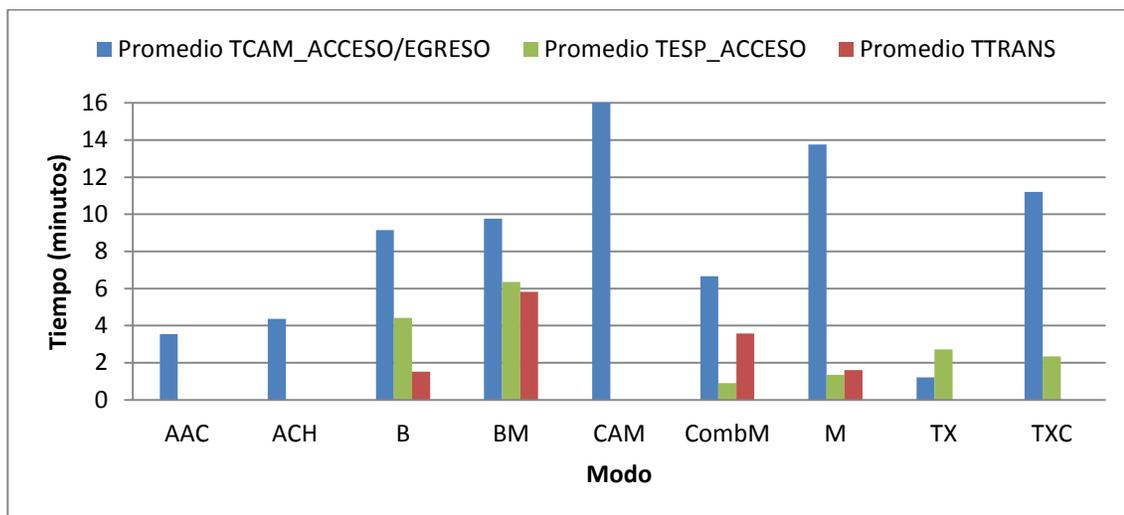


Figura 4-8: Tiempo de caminata de acceso/egreso, tiempo de espera de acceso y tiempo de transbordo promedio según modo de transporte

Fuente: Elaboración propia

Como era esperable, los modos bus y bus-Metro son los que tienen mayor tiempo de espera promedio (4 y 6 min, respectivamente), mientras que el tiempo de espera en el resto de los modos es bastante bajo. En el caso de los que usan Metro, esto se debe a la alta frecuencia y regularidad de su frecuencia, en contraste con los modos asociados al bus.

Respecto a los tiempos de transbordo, los modos bus-Metro y combinación con Metro son los que tienen mayor tiempo de transbordo promedio. Dado que el tiempo de transbordo está compuesto por la suma del tiempo de caminata y espera, estos resultados son consistentes (para transbordar al Metro desde otro modo se debe caminar). Este no es el caso para los modos bus y Metro; en el primero muchas veces se puede tomar el siguiente recorrido en el mismo paradero donde se egresa del recorrido anterior, mientras que las distancias de transbordo entre líneas de Metro no son altas. También influye la cantidad de transbordos que se realizan en cada modo. En bus y Metro se realizan respectivamente 0,2 y 0,6 transbordos en promedio, lo que refleja que las rutas escogidas son bastante directas ente orígenes y destinos; en contraste, en los modos bus-Metro y combinación con Metro se hacen respectivamente 1,7 y 1,5 transbordos en

promedio. Estos valores también son un indicio de que las personas ponderan negativamente tener que transbordar, en especial tener que hacerlo más de una vez.

En cuanto al costo total de viaje, como era de esperar los modos más costosos son el auto-chofer y taxi (ambos más de \$2.000 por viaje) y en menor medida el taxi-colectivo (ver Tabla 4-4). Los valores para el transporte público masivo son consistentes con las tarifas de Transantiago y el Metro de Santiago.

Tabla 4-4: Costo total (operacional + peaje + estacionamiento) según tipo de modo

AAC	ACH	B	BM	CAM	CombM	M	TX	TXC
\$ 228	\$ 2.311	\$ 520	\$ 594	\$ -	\$ 1.298	\$ 598	\$ 2.514	\$ 1.015

Fuente: Elaboración propia

La densidad promedio experimentada por los usuarios del Metro es de 2,2 pax/m², con un mínimo cercano a cero (viaje muy temprano en la mañana) y un máximo de 5,8 pax/m² (viaje en hora punta, en línea y sentido congestionado). La temperatura promedio al momento de comenzar el viaje fue 19,2 °C, con rango entre 12,8 y 29,5 °C.

En relación a las demoras horarias (*SD*), en promedio las personas llegan 10,2 minutos antes (*SDE*) y 3,7 minutos después (*SDL*) que el horario preferido de llegada al trabajo (*PAT*). Estos valores cambian si sólo se consideran dentro del promedio a las personas que llegaron más de 5 min antes (para *SDE*) o después (para *SDL*) de su *PAT*: el promedio de *SDE* es 20,2 y el de *SDL* es de 15,5. Estos valores son también bastante razonables y demuestran que las personas efectivamente escogen llegar cerca de sus horarios preferidos de llegada, ya sea por falta de flexibilidad laboral, penalidad por llegar tarde, u otro tipo de razones. Por último, 147 personas llegaban después del horario oficial de entrada al trabajo, lo que representa al 31% de la muestra.

4.3 Banco de datos de preferencias declaradas (PD)

Interesa analizar cómo se distribuyen las elecciones y niveles de servicio promedio según alternativa (ver Tabla 4-5). En general las personas tienden a escoger una alternativa similar a la escogida en la realidad (56% de las elecciones); esto se conoce como efecto hábito o inercia. Asimismo, si no escogen la alternativa de salir a la misma

hora, prefieren salir antes (20%), luego cambiar de modo (14%) y en último lugar salir después (10%).

Tabla 4-5: Elección y niveles de servicio promedio base PD

Alternativa	Elección (%)	Promedio				
		Tiempo de viaje usual (min)	SDE (min)	SDL (min)	Costo (\$)	Tiempo de viaje una vez a la semana (min)
Salir hora actual	56	50,6	4,9	10,0	891	61,7
Salir hora antes	20	41,5	57,7	0,8	782	50,4
Salir hora después	10	48,2	0,1	62,9	776	58,6
Cambiar modo	14	49,3	8,4	9,2	2022	60,5

Fuente: Elaboración propia

El tiempo de viaje promedio fluctúa entre los 40 y 50 min según alternativa, pero el valor correspondiente a la alternativa similar a la escogida en PR (“Salir hora actual” o “actual” de ahora en adelante) es mayor que el tiempo de viaje (en vehículo) promedio del banco de PR. Esto se puede deber a que el tiempo de viaje en PD, por motivos de simpleza, se asoció con la duración total del viaje, ya que presentar más atributos, como el tiempo de espera, caminata o transbordo complejizaba radicalmente el experimento. Asimismo, el tiempo de viaje promedio de las alternativas “salir antes” y “salir después” es menor al tiempo de viaje de las alternativas “salir a la hora actual” y “cambio de modo”. El tiempo de viaje una vez a la semana, que es un *proxi* de la variabilidad en los tiempos de viaje experimentados en la red de transporte es, en general, entre 21 y 23% mayor que el tiempo de viaje durante el resto de la semana.

Los valores promedio de *SDE* y *SDL* de la alternativa actual, son bastante similares a los de la alternativa escogida en PR. Además, los valores de *SDE* y *SDL* de las alternativas “salir antes” y “salir después”, son bastante altos (más de una hora); esto significa que esas alternativas implicaban cambios considerables en las horas de inicio y llegada al trabajo.

Respecto al costo, las alternativas “salir antes” y “salir después” tienen un costo menor que salir a la hora actual; esto es consistente con uno de los objetivos del

experimento: evaluar el efecto de la tarificación vial. Para la alternativa “cambio de modo”, el costo promedio asciende a más de \$2.000.

Por último, el 39% de los encuestados llega tarde al trabajo (después de *HOE*) de acuerdo a las alternativas que escogieron en el experimento de PD. Este valor es bastante similar al 31% del banco de datos de PR, lo que puede ser un indicio que quizá muchas personas le otorgaron bastante peso relativo a esta variable al momento de escoger en el experimento PD.

5 MODELOS DE ELECCIÓN DE MODO Y HORA DE INICIO DEL VIAJE: BASE 15 MINUTOS

A continuación se presentan y analizan los resultados de estimar modelos conjuntos de elección de modo y hora de inicio del viaje a partir de las bases de preferencias reveladas, con intervalos de 15 min (PR_15), y preferencias declaradas (PD) descritas en la sección anterior. De acuerdo a la metodología recomendada (Louviere *et al.*, 2000), en primer lugar, se estimaron modelos sólo con la base de PR, luego con la base de PD, y finalmente modelos combinando ambos tipos de datos, usando el *software* Biogeme (Bierlaire, 2003).

5.1 Análisis de resultados modelos base PR

Se generaron dos bancos de datos de PR, que difieren en la forma de incorporar la flexibilidad en los horarios de llegada al trabajo en el modelo de elección discreta. La flexibilidad horaria es fundamental para los modelos de elección de hora de inicio del viaje, pues es una medida de la capacidad de las personas para adelantar o atrasar sus viajes. La práctica más común en la literatura es estimar distintos modelos para individuos con distintas flexibilidades horarias; sin embargo, esto no se puede hacer con el banco de datos disponible dada su limitada cantidad de observaciones. El primer banco de datos (base PR_15_1) incorpora la flexibilidad horaria limitando el conjunto de alternativas horarias disponibles de cada individuo (no se encontró estudios anteriores en la literatura que realicen esto). La segunda base (base PR_15_2), considera que todos los individuos tienen disponibles todos los períodos de inicio del viaje, y la flexibilidad se incorpora haciendo interactuar las variables de demora horaria (*SD*) con variables ficticias asociadas al grado de flexibilidad horaria (baja, media y alta).

5.1.1 Flexibilidad horaria según disponibilidad de alternativas

Los primeros modelos estimados con el banco de datos base PR_15_1 fueron del tipo MNL, buscándose encontrar las variables que mejor explicasen las elecciones de las personas, tanto en términos estadísticos como en su consistencia con la teoría microeconómica. Cabe destacar que un modelo del tipo MNL implica escoger de manera simultánea el modo de transporte y la hora de inicio del viaje.

Las variables de nivel de servicio que se probaron inicialmente en las funciones de utilidad de las 99 alternativas de modo-hora (nueve modos por 11 períodos) fueron los tiempos de viaje, espera, caminata, acceso (caminata + espera), y transbordo, el costo de cada modo, la densidad en el Metro - sólo para los modos Metro (M) y bus-metro (BM) - y variables ficticias asociadas a la temperatura ambiente cuando comenzaba el viaje. Estas variables ficticias se agregaron sólo en el caso de los modos caminata (CAM) y Metro, dado que caminar (o ir en bicicleta) al trabajo cuando hace mucho calor puede ser más desagradable por la transpiración. En el caso del Metro, además de poca circulación de aire en sus vagones, la temperatura en andenes y vagones tiende a ser mayor que en la superficie, y esta incomodidad se amplifica a medida que el tren va más cargado. El hecho que Metro S.A. haya recientemente incorporado a su flota coches con aire acondicionado, refleja la importancia estratégica que otorga la compañía a reducir la incomodidad producida por las altas temperaturas en época de verano.

Las variables que contabilizan las restricciones en la programación de actividades incluidas en los modelos, fueron las demoras horarias temprano (*SDE*) y tarde (*SDL*), una variable ficticia que se activa al llegar tarde al trabajo, y el número de paradas intermedias previas en el viaje al trabajo (ésta último sólo en el modo auto-chofer). También se probaron variables socioeconómicas, como *AUTLIC* (en los modos auto-chofer y combinación-Metro), e interacciones entre éstas y variables de nivel de servicio, como el costo dividido por la tasa salarial (*Costw*). Esta última variable, muy empleada en la literatura, permite acomodar el efecto de que mientras mayor ingreso tiene un individuo, menos importancia le otorga al costo del viaje.

Se espera un signo negativo para los coeficientes de todas las variables antes descritas, pues provocan desutilidad, a excepción del número de paradas intermedias y *AUTLIC*, ya que ambas deberían hacer al modo auto-chofer más atractivo.

Como el número de alternativas modo-hora es muy alto (99) y no todas fueron escogidas, en vez de utilizar constantes específicas para cada una, sólo se especificaron constantes específicas modales (*ASC_Modo*), fijándose en cero la del modo auto-chofer. De esta forma el modelo MNL reproduce la partición modal observada. No se incluyó constantes específicas para cada período de inicio del viaje, porque los términos *SDE* y

SDL capturan su efecto. Esto, a la vez, implica que los modelos no necesariamente reproducirán las proporciones de mercado observadas en cuanto a hora de inicio del viaje.

La Tabla 5-1 muestra los resultados de la estimación de modelos conjuntos de modo y hora de inicio del viaje con las variables mencionadas en los incisos anteriores. Los resultados incluyen la log-verosimilitud en convergencia, el índice rho-cuadrado ajustado, el estimador puntual de cada coeficiente y el test-t robusto de cada parámetro; este último permite analizar si el parámetro es significativamente distinto de cero⁹.

El modelo 1 de la Tabla 5-1 incluye sólo las variables consideradas como clave o de política. Como es posible observar, la mayoría de estas variables son significativamente distintas de cero, a excepción del tiempo de espera y la variable ficticia que se activa si una persona llega después de la hora oficial de entrada al trabajo (*D_TARDE*). El tiempo de espera tiene un impacto bastante pequeño, contrariamente a lo esperado, ya que en general es percibido más negativamente que el tiempo de viaje. La variable *D_TARDE* no es significativa al 95% de confianza, pero se mantiene por ser considerada una variable de política.

Los tiempos de caminata y de transbordo son altamente significativos. Un minuto de tiempo de caminata (o de transbordo) produce aproximadamente la misma desutilidad que 3,5 (u 8) min de tiempo de viaje en vehículo. El coeficiente del costo también es bastante significativo, e incluirlo dividido por la tasa salarial (modelos 2) da resultados muy similares. Por otro lado, los coeficientes de demora horaria por llegar antes (*SDE*) y después (*SDL*) son muy significativos, lo que indica su importancia para modelar las preferencias temporales de los usuarios del sistema de transporte. Además, como era de esperar, llegar tarde (*SDL*) es percibido más negativamente que llegar temprano (*SDE*).

⁹ Como el valor crítico del test-t es 1,96 (para el 95% de confianza) o 1,645 (para el 90%), un estadígrafo t mayor a 1,645 en valor absoluto implica que en el 90% de los casos el parámetro es mayor (o menor, dependiendo del signo esperado) que cero. El test-t robusto, en comparación al test-t normal, tiene la ventaja de ser robusto ante especificaciones incorrectas acerca de la distribución de los errores de los parámetros, por ejemplo, asumir homocedasticidad o igual varianza en toda la muestra (White, 1980).

Tabla 5-1: Resultados estimación modelos Modo-Hora MNL-1 base PR_15_1

Modelos Base PR_15_1				
Coefficiente	1		2	
	Valor	Test-t robusto	Valor	Test-t robusto
<i>Niveles Servicio</i>				
Tiempo de viaje	-0.0208	-1.78	-0.0236	-1.86
Tiempo de caminata	-0.0734	-5.44	-0.0786	-5.52
Tiempo de espera	-0.00648	-0.15	-	-
Tiempo de transbordo	-0.169	-6.05	-0.167	-5.97
Costo	-0.000648	-5.44	-	-
Temperatura media (20°C-25°C)*	-	-	-0.0863	-0.34
Temperatura alta (>=25°C)*	-	-	-1.11	-1.16
Densidad Metro**	-	-	-0.0898	-1.56
<i>Restricciones de Programación</i>				
Demora horaria por llegar temprano (<i>SDE</i>)	-0.0349	-9.57	-0.0358	-9.56
Demora horaria por llegar tarde (<i>SDL</i>)	-0.073	-8.72	-0.0736	-8.6
Llega tarde al trabajo (<i>D_TARDE</i>)	-0.218	-1.45	-0.213	-1.41
# paradas intermedias***	-	-	3.18	3.27
<i>Socioeconómicas</i>				
<i>AUTLIC</i>	-	-	0.946	2.24
<i>Interacciones</i>				
Costo/tasa salarial (<i>Costw</i>)	-	-	-1.45	-5.2
Log-verosimilitud	-1146.63		-1135.29	
rho-cuadrado ajustado	0.239		0.244	
# parámetros	16		20	
# observaciones	468		468	

(*): Variable incluida sólo en las funciones de utilidad de los modos CAM y M; (**): Variable sólo incluida en modos M y BM; (***): Variable sólo incluida en modo ACH.

Fuente: Elaboración propia

En el modelo 2 se incluyen más variables explicativas. El parámetro de *AUTLIC* es significativo y con signo positivo, tal como se esperaba. El número de paradas intermedias también tiene un efecto positivo en el modo auto, permitiendo incorporar la flexibilidad que entrega el modo auto-chofer para realizar diversas paradas en un viaje. Por otra parte, la densidad del Metro efectivamente genera desutilidad, aunque sólo es estadísticamente distinta de cero al 80% de confianza. Asimismo, y como era de esperar, temperaturas altas (mayores a 25°C) generan bastante mayor desutilidad que temperaturas entre 20°C y 25°C (nivel medio). No obstante, ninguna de las dos variables

ficticias incorporadas son significativamente distintas de cero, y la variable asociada a temperaturas de nivel medio tiene un efecto mínimo.

Que la temperatura y la densidad en Metro no sean tan significativas se puede deber a diversos factores. En primer lugar, a pesar de que la temperatura del Metro está en su nivel máximo en el verano, la congestión no está en su nivel máximo en esa estación del año. Segundo, muchas personas prefieren viajar hacinadas y acaloradas en el Metro que en un bus relativamente vacío y con un poco más de ventilación, debido a la confiabilidad (en regularidad y variabilidad del tiempo de viaje) que ofrece el servicio de Metro; por esto, puede que su efecto real no sea tan significativo en comparación con otros medios de transporte público y también es posible que estas variables sean más preponderantes para algunos usuarios de auto, que ven al transporte público con cierto recelo debido a la incomodidad (relativa) de viajar en él. Tercero, la comodidad que ofrece un modo es una variable de carácter intangible y que depende de múltiples atributos (calor, olor, nivel de ocupación, nivel de peligrosidad, nivel de ruido, etc.). Al incorporarla de forma desagregada (i.e. usando diversos atributos), puede que no sea significativa; sin embargo, puede que sí lo sea si se utiliza una sola variable como *proxi* de comodidad (por ejemplo, (i) viajar sentado y temperatura normal; (ii) viajar parado, apretado y con mucho calor, etc.). Sin embargo, para captar su real efecto esa variable de comodidad debería estar en todos los modos; lamentablemente, para crearla hubiese sido necesario información que no se tenía disponible, como la diferencia de temperatura entre un vagón de Metro y en la superficie, el nivel de ocupación de los buses, o si el auto utilizado por el individuo tenía aire acondicionado. Por último, y en cuarto lugar, es perfectamente posible que la mayoría de las personas no den mucha importancia relativa a la temperatura al momento de escoger su modo o la hora de inicio del viaje.

La Tabla 5-2 presenta los resultados obtenidos luego de introducir algunas modificaciones a las funciones de utilidad anteriores (llamado modelo 3). La primera fue separar el coeficiente del transbordo: a los modos bus, Metro, bus-Metro y taxi-colectivo se les asignó un coeficiente (transbordo público) y al modo combinación con Metro se le

asignó otro (transbordo público-privado¹⁰). La segunda, es la forma de introducir el parámetro del costo a fin de evaluar el efecto de la especificación de esta variable en las valoraciones monetarias del tiempo. La tercera modificación es que dado que el tiempo de espera no fue significativo, este se sumó al tiempo de caminata, formando la variable tiempo de acceso.

Al comparar el modelo 3 de la Tabla 5-2 con los modelos 1 y 2 de la Tabla 5-1, se aprecia que existe una ganancia considerable en log-verosimilitud al separar el coeficiente de los transbordos entre modos de transporte público y los transbordos entre transporte público y privado. De hecho, un transbordo público-privado es percibido seis veces más negativamente que un transbordo público-público (ver la discusión de este tema en Navarrete y Ortúzar, 2012). Este es un resultado sumamente interesante, y explicaría en gran medida la baja elección del modo combinación-Metro. Asimismo, en comparación a los modelos 1 y 2, la percepción del tiempo de transbordo público-público se asemeja más a la del tiempo de acceso (que es la suma de los tiempos de caminata y espera, pero tiene un coeficiente muy similar al del tiempo de caminata). Finalmente, al separar por tipo de transbordo el coeficiente del tiempo de viaje aumenta (en valor absoluto), al igual que su significancia.

En cuanto a la forma de introducir el costo, se observa que al incluir simultáneamente el costo tanto de forma lineal (Costo) como dividido por la tasa salarial (Costw) ambos coeficientes son estadísticamente significativos. Este resultado llama en principio la atención ya que ambas variables debieran estar altamente correlacionadas y, por ende, incluirlas juntas podría afectar negativamente al modelo. La explicación de esta especificación se detalla más adelante en el análisis de los valores monetarios del tiempo (Jara-Díaz *et al.*, 2012).

¹⁰ Se le denomina transbordo público-privado pues, a pesar que el modo taxi-colectivo-Metro es parte del conjunto de modos combinación-Metro, involucra un transbordo desde alguno de los modos individuales - que tienen características similares entre sí y pueden asociarse a las del transporte privado (por ejemplo, baja ocupación del vehículo) - al modo Metro, que tiene otro tipo de características más típicas del transporte público masivo (por ejemplo, alta ocupación).

Tabla 5-2: Resultados estimación modelos MNL-2 base PR_15_1

Coeficiente	Modelos Base PR_15_1	
	Valor	Test-t robusto
Niveles Servicio		
Tiempo de viaje	-0.0374	-2.33
Tiempo de acceso	-0.071	-5.26
Tpo. transbordo público	-0.103	-4.04
Tpo. transbordo público-privado	-0.642	-3.67
Costo	-0.000373	-2.88
Restricciones de Programación		
Demora horaria por llegar temprano (<i>SDE</i>)	-0.0348	-9.59
Demora horaria por llegar tarde (<i>SDL</i>)	-0.074	-8.72
Llega tarde al trabajo (<i>D_TARDE</i>)	-0.217	-1.43
# paradas intermedias	3.48	3.34
Socioeconómicas		
<i>AUTLIC</i>	0.998	2.24
Interacciones		
Costo/tasa salarial (<i>Costw</i>)	-0.784	-3.31
Log-verosimilitud	-1120.43	
rho-cuadrado ajustado	0.254	
# parámetros	19	
# observaciones	468	

Fuente: Elaboración propia

Dado que las alternativas relacionadas con el período horario de salida no tienen la misma longitud, Ben-Akiva y Abou-Zeid (2012), recomiendan agregar un factor de corrección igual al logaritmo natural de la longitud del intervalo sobre el intervalo de mínima longitud (en este caso 15 min). Por ejemplo, para el período de 5:00 – 7:00, este factor toma el valor de logaritmo natural de ocho (120 partido por 15), y para el período 7:00 – 7:15, que es el de menor longitud, toma el valor de cero (logaritmo natural de uno). El coeficiente de esta variable no se estima y su valor se fija en uno. Los resultados de esto, que se muestran en la Tabla 5-3 (modelo 3_LN), permiten notar que aun cuando casi todos los parámetros aumentan su valor y su test-t robusto (a excepción de *Costw* y *AUTLIC*), la log-verosimilitud y el índice rho-cuadrado ajustado empeoran. Pareciese que los parámetros aumentan su valor absoluto para neutralizar la suma del factor de corrección. Es por ello que se decidió prescindir de este factor de corrección, y en el capítulo 7 se evalúa el efecto de incluirlo en los modelos predictivos.

Tabla 5-3: Resultados estimación modelos MNL-3 base PR_15_1

Modelos Base PR_15_1				
Coeficiente	3_LN		3_VarGustos	
	Valor	Test-t robusto	Valor	Test-t robusto
Niveles Servicio				
Tiempo de viaje	-0.0472	-2.87	-0.0366	-2.3
Tiempo de acceso	-0.0799	-5.81	-0.0709	-5.24
Tpo. transbordo público	-0.12	-4.81	-0.0603	-1.98
Tpo. transbordo público-privado	-0.663	-3.65	-0.636	-3.57
Restricciones de Programación				
Demora horaria por llegar temprano (<i>SDE</i>)	-0.0471	-13.38	-0.0293	-6.91
Demora horaria por llegar tarde (<i>SDL</i>)	-0.088	-10.48	-0.0662	-6.48
Llega tarde al trabajo (<i>D_TARDE</i>)	-0.32	-2	-0.2	-1.33
# paradas intermedias	3.47	3.8	3.43	3.82
Socioeconómicas				
<i>AUTLIC</i>	0.987	2.22	0.999	2.23
Interacciones				
Costo/tasa salarial (<i>Costw</i>)	-1.44	-5.1	-1.45	-5.11
<i>SDL</i> - Hombre			-0.0385	-3.08
<i>SDL</i> - Educación técnica			-0.0659	-2.23
<i>SDL</i> - Ingreso alto			0.0387	3.2
<i>SDE</i> - Ingreso medio alto			-0.0189	-2.6
Tpo. transbordo público - Edad mayor a 30			-0.0667	-1.72
Log-verosimilitud	-1177.01		-1106.11	
rho-cuadrado ajustado	0.218		0.261	
# parámetros	18		23	
# observaciones	468		468	

Fuente: Elaboración propia

En la misma Tabla 5-3 se presenta un modelo que considera variaciones sistemáticas en los gustos (3_VarGustos). Se intentó incluir interacciones entre los tiempos de viaje, acceso y transbordo, el costo, *SDE* y *SDL* con características sociodemográficas como la edad, sexo, nivel de escolaridad, si debe acompañar a alguien en su viaje al trabajo, y nivel de ingreso mensual. Las que resultaron significativas se muestran en el modelo 3_VarGustos. Se destaca que tanto los hombres como las personas que tienen educación técnica (por ende posiblemente un trabajo de índole técnica) son más reacios a llegar tarde, mientras que las personas de ingreso alto (ingreso mensual mayor a \$ 1,3 millones) no tienen tantos inconvenientes en llegar tarde. También, los resultados permiten deducir que a las personas de ingreso medio-alto

(entre \$ 800.000 y \$ 1,3 millones mensuales) les desagrada más tener que llegar antes de su horario preferido de entrada. Por último, a las personas de más de 30 años les resulta aún menos atractivo el tiempo incurrido en transbordar en transporte público.

La valoración monetaria de los distintos tiempos y las razones de compromiso que se derivan del *Scheduling Model* de Small (1982) se muestran en la Tabla 5-4. En este caso los modelos 3_Cost y 3_Costw corresponden al modelo 3, pero utilizando solamente las variables Costo y Costw, respectivamente (recordar que el modelo 3 incluye ambas variables en forma simultánea). De esta tabla llama la atención la sensibilidad de las valoraciones ante cambios en la especificación de las funciones de utilidad.

Tabla 5-4: Valor subjetivo del tiempo y razones de compromiso modelos base PR_15_1¹¹

Modelos Base PR_15_1					
	<i>1</i>	<i>3</i>	<i>3_Cost</i>	<i>3_Costw</i>	<i>3_LN</i>
Valor subjetivo (\$/min)					
Tiempo viaje	32	64	62	96	121
Tiempo caminata	113			-	-
Tiempo espera	10			-	-
Tiempo acceso	-	121	112	182	204
Tiempo transbordo	261			-	-
Tiempo transbordo público	-	176	163	263	307
Tiempo transbordo público-privado	-	1,096	1,040	1,633	1,696
Razón de compromiso					
<i>SDL/SDE</i>	2.1	2.1	2.1	2.1	1.9
<i>SDE/Tiempo viaje</i>	1.7	0.9	0.9	0.9	1.0
<i>SDL/Tiempo viaje</i>	3.5	2.0	1.9	2.0	1.9

Fuente: Elaboración propia

Por ejemplo, el valor del tiempo de viaje pasa de 32 \$/min en el modelo más simple (modelo 1) a 121 \$/min en el modelo 3_LN. Este cambio se debe principalmente a tres efectos: (i) la forma de introducir el costo; (ii) la desagregación del coeficiente de

¹¹ La valoración monetaria en los modelos que usan el parámetro costo/w, se determinó usando la tasa salarial (w) promedio de 3.683 \$/hora.

los transbordos (que aumenta la percepción negativa del tiempo de viaje y no la del costo), y (iii) la agregación del factor de corrección del logaritmo natural, que también hace que aumente (en valor absoluto) el parámetro del tiempo de viaje y no el del costo.

En relación a la forma de introducir el costo, se aprecia que al introducir la variable costo/tasa salarial (modelo 3_Costw) en lugar del costo en forma lineal (modelo 3_Cost), las valoraciones del tiempo aumentan casi en un 50%. Este efecto ha sido discutido recientemente por Jara-Díaz *et al.* (2012), quienes lo atribuyen a la varianza existente en los ingresos.

Los autores, a partir de una muestra sintética, demuestran que a medida que aumenta el coeficiente de variación (razón entre la desviación estándar y la media) del ingreso de la muestra, la valoración del tiempo también aumenta al usar la variable costo dividido por ingreso. En cambio, el valor del tiempo permanece invariante al utilizar el costo en forma lineal. La muestra generada por los autores tiene un ingreso medio de \$ 500 mil y su coeficiente de variación va desde cero (ingresos iguales para todos) hasta 0,3. En este último caso, la valoración del tiempo al usar la variable costo/ingreso es cerca de 154,9 \$/min, mientras que al usar sólo la variable costo es de 15,8 \$/min. En el banco de datos de esta tesis el ingreso medio es cercano al millón de pesos, con un coeficiente de variación de 0,54 (ídem para la tasa salarial), por lo que no sorprenden los resultados obtenidos.

Jara-Díaz *et al.* (2012) también proponen una forma de corregir el efecto antes descrito, ya que usar solamente el costo/ingreso puede inducir a errores al evaluar políticas públicas. La solución es agregar simultáneamente el costo de forma lineal y dividido por el ingreso (o tasa salarial), como se hizo en el modelo 3. Al hacer esto se puede comprobar que efectivamente las valoraciones disminuyen, pero también que la bondad de ajuste del modelo mejora. Esto se puede atribuir a la alta varianza en los ingresos, ya que a mayor variación en los ingresos, menor es la correlación entre las variables costo y costo/w.

Excluyendo a los modelos 3_Costw y 3_LN, el valor del tiempo de viaje fluctúa entre 30 y 65 \$/min, lo que se encuentra dentro del rango de valores reportados en Chile, considerando que esta es una muestra de ingreso medio-alto.

Otro resultado destacable es que la valoración del tiempo de transbordo entre modos de transporte público es casi 1,5 veces mayor que el tiempo de acceso/egreso; además, en el caso del transbordo público-privado este valor es cercano a 9. Esto confirma que a las personas les desagrada más el tiempo incurrido en transbordar que en acceder a un modo. Llama la atención que un minuto transbordando hacia Metro, desde un modo privado o taxi-colectivo, sea valorado en alrededor de \$1.000. Si consideramos que el tiempo promedio de transbordo de la alternativa combinación-Metro es de 3,6 min (ver Tabla 4-8), un transbordo de este tipo, completo, estaría valorado en promedio en \$ 3.600. A pesar de que este valor parece sumamente elevado, en parte se podría explicar porque hay personas que están dispuestas a pagar alrededor de \$ 4.000 diarios por estacionar su auto directamente en el lugar de trabajo en el centro de Santiago, en lugar de pagar muy poco o nada por estacionar cerca de una estación de Metro para luego combinar. Otra razón podría ser la alta disponibilidad del modo combinación-Metro (60%), lo que contrasta con la menor disponibilidad de los modos individuales (11% para los modos auto-chofer-Metro y auto-acompañante-Metro, 25% para taxi-Metro y 42% para taxi-colectivo-Metro).

También se trató de incluir en los modelos la variable de número de transbordos en lugar de la variable de tiempo de transbordo, sin embargo, la bondad de ajuste del modelo era mucho menor. La valoración (independiente del tipo de transbordo) en este caso, era de \$ 550 por transbordo. Si se incluía ambas variables simultáneamente, para controlar por el número de transbordos, el signo del coeficiente de esta última variable resultaba contrario al esperado, posiblemente debido a la alta correlación entre ambas variables.

Luego de estimados los modelos MNL, se procedió a estimar estructuras más flexibles y generales, como el logit jerárquico (HL) y el logit mixto (ML). En el caso del HL, se intentaron dos tipos de estructuras: una en que se agrupó en un nido a todas las alternativas de modo para cada alternativa de hora (i.e. elección de hora en jerarquía superior y modo en jerarquía inferior; 11 nidos) y otra en la que postulaba correlación entre las alternativas de hora (i.e. modo jerarquía superior y hora en la inferior; nueve nidos). Los resultados no fueron satisfactorios, ya que sólo uno (de un total de nueve u

11, dependiendo de la estructura) de los parámetros estructurales resultó significativamente menor a uno (un nido para el modo bus y un nido para la alternativa de hora 5:00 – 7:00).

En referencia al ML, se estimaron modelos con parámetros aleatorios (distribución normal) para las variables tiempo de viaje, *SDE* y *SDL*, pero sólo la media (valor del coeficiente: -0,0935) y desviación estándar (valor del coeficiente: -0,0371) de la variable *SDL* resultaron significativas. Esto era de esperar, ya que de acuerdo el modelo de variaciones sistemáticas en los gustos (modelo 3_VarGustos de la tabla 5-3), la variable *SDL* era la que presentaba mayor heterogeneidad. No obstante, el modelo con este parámetro aleatorio resultó inferior, en términos de bondad de ajuste, al modelo con variaciones sistemáticas en los gustos (esto se ha encontrado en otras experiencias, ver por ejemplo, Grisolia y Ortúzar, 2010).

Por último, en la Tabla B-1 del ANEXO B, se presentan las ASC de modo para los modelos presentados en las tablas 5-1, 5-2 y 5-3. Las ASC agrupan todos los atributos no explicados por las variables del modelo, por lo que ASC poco significativas implican que las variables del modelo explican bastante bien la variación observada en las elecciones de modo. En este caso, la ASC del modo auto-chofer (ACH) se fijó en cero.

5.1.2 Flexibilidad horaria como variable explicativa

A continuación se detallan los resultados de la estimación de modelos a partir de la base que considera que todos los individuos tienen disponibles todos los períodos de inicio del viaje (base PR_15_2). En este caso, la flexibilidad para modificar la hora de inicio del viaje se incorpora mediante interacciones entre las variables *SDE* y *SDL* y variables ficticias relacionadas con el grado de flexibilidad horaria. Se definió tres categorías de flexibilidad laboral en concordancia con la información que entregaron los encuestados: (i) flexibilidad alta: si la persona puede llegar más de 30 minutos tarde y temprano al trabajo, (ii) flexibilidad media: si puede llegar entre 15 y 30min temprano y tarde al trabajo, y (iii) flexibilidad baja: si la persona puede llegar como máximo 15 minutos temprano y tarde al trabajo (ver Figura 4-3).

La Tabla 5-5 contiene resultados selectos de la estimación de modelos conjuntos de elección de modo y hora de inicio del viaje a partir de esta segunda base de datos. De ella se pueden desprender varios resultados. Primero, la gran mayoría de las variables explicativas no sufre grandes modificaciones con respecto a los modelos de la sección anterior, a excepción lógicamente de los términos *SDE* y *SDL*, y que ahora la variable ficticia llega tarde al trabajo (*D_TARDE*) tiene mayor significancia estadística.

Tabla 5-5: Resultados estimación modelos MNL-1 base PR_15_2

Coeficiente	Modelos Base PR_15_2			
	4		5	
	Valor	Test-t robusto	Valor	Test-t robusto
Niveles Servicio				
Tiempo de viaje	-0.0383	-2.35	-0.0386	-2.37
Tiempo de caminata				
Tiempo de acceso	-0.0721	-5.26	-0.0725	-5.29
Tiempo de transbordo				
Tpo. transbordo público	-0.102	-3.97	-0.102	-3.98
Tpo. transbordo público-privado	-0.646	-3.69	-0.646	-3.7
Costo	-0.00037	-2.81	-0.0004	-2.82
Restricciones de Programación				
Demora horaria por llegar temprano (<i>SDE</i>)	-0.077	-11.34	-0.0539	-16.28
Demora horaria por llegar tarde (<i>SDL</i>)	-0.101	2.77	-0.0814	-10.99
Tarde al trabajo (<i>D_TARDE</i>)	-0.38	-2.46	-0.367	-2.33
# paradas intermedias	3.53	3.27	3.52	3.27
Socioeconómicas				
<i>AUTLIC</i>	0.942	2.08	0.956	2.13
Interacciones				
Costo/tasa salarial (<i>Costw</i>)	-0.807	-3.28	-0.805	-3.31
<i>SDE</i> - Flexibilidad horaria alta	0.0405	4.4		
<i>SDE</i> - Flexibilidad horaria media	0.0251	3.31		
<i>SDL</i> - Flexibilidad horaria alta	0.0409	2.77		
<i>SDL</i> - Flexibilidad horaria media	0.0198	1.67		
Log-verosimilitud	-1195.18		-1208.16	
rho-cuadrado ajustado	0.346		0.342	
# parámetros	23		18	
# observaciones	468		468	

Fuente: Elaboración propia.

Segundo, como era de esperar existen diferencias de percepción en las variables de demora horaria (*SDE* y *SDL*) dependiendo de la flexibilidad horaria reportada por los

individuos. En general, aquellas personas con alta flexibilidad para adelantar o atrasar sus viajes tienen menor valoración (en valor absoluto) de los términos *SDE* y *SDL*, respectivamente, en comparación a aquellos con flexibilidad media y flexibilidad baja.

En tercer lugar, es deseable comparar los modelos de la sección anterior con los de esta para examinar que forma de incorporar la flexibilidad horaria es más adecuada. Dado que no es posible comparar directamente la log-verosimilitud en convergencia de ambos tipos de modelos, porque el banco de datos no es el mismo, se comparan los índices rho-cuadrado ajustado¹². Por ejemplo, este índice toma el valor 0,254 para el modelo 3 de la Tabla 5-2, mientras que su símil de esta sección (modelo 4) asciende a 0,346. Si a esto se suma que las interacciones de flexibilidad con los parámetros de *SDE* y *SDL* son significativas y con signos y magnitudes esperadas, y que el aporte de la variable ficticia asociada a llegar tarde también aumenta, se concluye preliminarmente que es más adecuado incorporar la flexibilidad como variable explicativa. La conclusión es preliminar, porque para dar una respuesta definitiva se deben contabilizar también los resultados de la estimación mixta PR/PD.

El modelo 4 fue el que presentó el mejor ajuste; este consideraba la especificación propuesta por Jara-Díaz *et al.* (2012) consistente en introducir conjuntamente el costo y el costo dividido por la tasa salarial. Es interesante mencionar que al quitar las interacciones entre las variables de flexibilidad y demora horaria, los coeficientes de *SDE* y *SDL* (ver modelo 5 de Tabla 5-5) son consistentemente más altos (en valor absoluto) que en los modelos anteriores (ver tablas 5-1 y 5-2). Esto es esperable, ya que al haber más alternativas de hora disponibles (y estar más alejadas, en tiempo, de la alternativa escogida), mayor peso se da a estas variables. Esto demuestra, además, que usar las interacciones de flexibilidad es importante, pues si se excluyen para algunas personas se estaría sobrevalorando el efecto de las variables *SDE* y *SDL* (al considerar disponibles períodos que, en realidad, no lo están).

¹² El índice rho-cuadrado ajustado no es un índice de bondad de ajuste tan apropiado para comparar modelos como el R^2 de regresión lineal, pero ante la ausencia de otros índices pareció el más razonable de utilizar.

También se estimó modelos que incluyen las interacciones de *SDE* y *SDL* con la flexibilidad horaria, usando el banco de datos de la sección anterior (base PR_15_1). Los resultados no fueron buenos, con algunos coeficientes no significativos y signos contrarios a lo esperado. De esta forma, se concluye que la flexibilidad debe incorporarse en una de las dos formas, pero no de ambas simultáneamente.

La Tabla 5-6 contiene las valoraciones monetarias y razones de compromiso para el modelo 4. Se puede apreciar que las valoraciones monetarias son muy similares a los resultados de la sección anterior (valor subjetivo del tiempo de viaje alrededor de 65 \$/min). En cuanto a las razones, se derivaron tres conjuntos de valores de acuerdo a la magnitud de la flexibilidad horaria. La razón *SDL/SDE* es bastante constante, con un valor entre 1,3 y 1,7, lo que difiere de los resultados de la sub-sección anterior (cerca de 2). Como era de esperar, las razones *SDL/Tiempo de viaje* y *SDE/Tiempo de viaje* son mayores para personas con baja flexibilidad y disminuyen a medida que aumenta la flexibilidad. De esta forma, se comprueba que personas con distintos niveles de flexibilidad perciben distinto un minuto tarde o temprano en comparación a un minuto de tiempo de viaje, en contraste a lo encontrado en el modelo 3 de la sub-sección 5.1.1 (ver Tabla 5-4), en el cual todos los individuos tienen la misma percepción relativa entre *SDE* (*SDL*) y el tiempo de viaje. Esto reafirma la decisión de decantarse por la flexibilidad como variable explicativa. No obstante, una potencial desventaja de acomodar la flexibilidad de esta manera es que se limita la cantidad de otras interacciones con las variables *SDE* y *SDL* (por ejemplo, con el ingreso), ya que tendrían que ser interacciones triples.

5.2 Análisis de resultados modelos base PD

A continuación se detallan los resultados de la estimación de modelos conjuntos de elección de modo y hora de inicio del viaje al trabajo, a partir del banco de datos de preferencias declaradas (PD). Los modelos varían según estructura (MNL y ML, tanto de parámetros aleatorios como de componentes de error), la forma de incorporar la variabilidad en los tiempos de viaje, las variables explicativas incluidas y, por último, si se consideran o no todas las observaciones. Esto último se debe a que Arellana (2012)

detectó que 46 personas contestaron la encuesta siguiendo un comportamiento aparentemente no compensatoria (i.e. escogiendo en base a sólo un atributo) y/o que podían presentar valoraciones contra intuitivas (por ejemplo, valoración positiva del tiempo de viaje).

Tabla 5-6: Valores subjetivos del tiempo y razones de compromiso modelo 4 base PR_15_2

Modelos Base PR_15_2	
4	
Valor subjetivo (\$/min)	
Tiempo viaje	65
Tiempo caminata	-
Tiempo acceso	122
Tiempo transbordo	-
Tiempo transbordo público	173
Tiempo transbordo público-privado	1,097
Razones de compromiso	
Flexibilidad baja:	
SDL/SDE	1.3
SDE/Tvia	2.0
SDL/Tvia	2.6
Flexibilidad media:	
SDL/SDE	1.6
SDE/Tvia	1.4
SDL/Tvia	2.1
Flexibilidad alta:	
SDL/SDE	1.6
SDE/Tvia	1.0
SDL/Tvia	1.6

Fuente: Elaboración propia

De esta forma, el banco de PD inicial, compuesta por 1.505 observaciones (301 individuos), se reduciría a 1.275 observaciones (255 individuos). A continuación se presentan los resultados para los modelos MNL.

5.2.1 Modelos Logit Multinomial (MNL)

En primer lugar se estimó modelos MNL utilizando el banco de datos PD completo (1.505 observaciones). Inicialmente, las funciones de utilidad variaron de

forma de acomodar la incertidumbre en los tiempos de viaje, como se presenta a continuación:

$$V_i = \theta_{Tviaje} Tvia_i + \theta_{var\ via} VarTvia_i + \theta_{SDE} SDE_i + \theta_{SDL} SDL_i + \theta_{D_Tarde} D_Tarde_i + \theta_{Costw} Costw_i + ASC_i \quad (5.1)$$

$$V_i = \theta_{ETT} ETT_i + \theta_{ESDE} ESDE_i + \theta_{ESDL} ESDL_i + \theta_{Costw} Costw_i + \theta_{D_Tarde} D_Tarde_i + ASC_i \quad (5.2)$$

donde i se refiere a las cuatro alternativas del experimento PD: salir antes, después, y a una hora similar a la hora observada (en el mismo modo escogido), o cambiar de modo (y salir a una hora similar a la observada).

En la primera formulación, descrita en ecuación (5.1), la variabilidad del tiempo de viaje se agregó de forma aditiva y mediante una variación porcentual, tal como lo explicita la siguiente ecuación:

$$VarTvia = (Tvia_{una\ vez\ a\ la\ semana} - Tvia_{usual}) / Tvia_{una\ vez\ a\ la\ semana} \quad (5.3)$$

Esto tenía la ventaja de disminuir la correlación con la variable de tiempo de viaje, además de ser consistente con la formulación de la variabilidad del tiempo de viaje en bus usada en los modelos de PR (ver ecuación 5.1).

En la segunda formulación (ecuación 5.2), la incertidumbre en los tiempos de viaje se incorporó mediante la esperanza o media del tiempo de viaje (ETT o *expected travel time*, ver ecuación 5.4) y los términos $ESDE$ y $ESDL$ (*expected schedule delays*). Estos últimos se calcularon de la misma forma que SDE y SDL , sólo que se usó ETT en lugar del tiempo de viaje.

$$ETT = 0,2 * Tvia_{una\ vez\ a\ la\ semana} + 0,8 * Tvia_{usual} \quad (5.4)$$

La modelos 1 y 1b de la Tabla 5-7 corresponden a los resultados de la primera y segunda formulación, respectivamente. Como se ve, no hay mucha diferencia entre ambos, salvo que los parámetros $ESDE$ y $ESDL$ son un poco mayores (en valor absoluto)

y más significativos que los coeficientes de *SDE* y *SDL*; además, en el modelo 1b la variable ficticia asociada a llegar tarde al trabajo es menos significativa. Ante esto, y para ser consistentes metodológicamente con los modelos de PR (en que se tiene *SDE* y no *ESDE*), se prefirió utilizar solamente la primera formulación. El modelo 1 además incluye una variable ficticia que se activa cuando en las alternativas de transporte público se ofrece la posibilidad de viajar sentado. Su coeficiente es significativo y su signo es positivo, lo que es razonable pues la mayoría de las personas prefiere ir sentada por comodidad.

Tabla 5-7: Resultados estimación modelos MNL base PD completa

Coeficiente	Modelos Base PD			
	<i>I</i>		<i>Ib</i>	
	Valor	Test-t robusto	Valor	Test-t robusto
Nivel de Servicio				
Tiempo de viaje	-0.0172	-3.41		
Tiempo de viaje esperado (<i>ETT</i>)			-0.014	-2.87
Incertidumbre % en tiempo de viaje	-1.25	-1.83		
Comfort - Ir sentado	0.311	3.04		
Restricciones de Programación				
Demora horaria por llegar temprano (<i>SDE</i>)	-0.0205	-7.91		
Demora horaria por llegar tarde (<i>SDL</i>)	-0.0224	-7.46		
<i>ESDE</i>			-0.0214	-8.18
<i>ESDL</i>			-0.0248	-8.44
Llega tarde al trabajo (<i>D_TARDE</i>)	-0.266	-2.52	-0.151	-1.43
Interacciones				
Costo/tasa salarial (<i>Costw</i>)	-1.76	-5.77	-1.93	-6.18
ASC				
Salir antes	-0.671	-4.68	-0.598	-4.17
Salir después	-0.871	-5.18	-0.783	-4.72
Cambio modo	-1.15	-14.09	-1.14	-14.17
Log-verosimilitud	-1563.17		-1568.39	
rho-cuadrado ajustado	0.246		0.244	
# parámetros	10		8	
# individuos	301		301	
# observaciones	1505		1505	
Valor subjetivo tiempo viaje (\$/min)	36		27	
Razones de compromiso				
<i>SDL/SDE</i>	1.1		1.2	
<i>SDE/Tvia</i>	1.2		1.5	
<i>SDL/Tvia</i>	1.3		1.8	

Fuente: Elaboración propia

Luego se estimó modelos excluyendo a los 46 individuos que aparentemente no presentaban comportamiento compensatorio (banco de datos depurado). Los resultados de estos modelos se encuentran en la Tabla 5-8. En términos de bondad de ajuste (índice rho-cuadrado ajustado), estos nuevos modelos son levemente superiores a los estimados con el banco de datos completo. Los valores subjetivos del tiempo también aumentan al sacar las observaciones con problemas. Esto sugiere que algunas personas con comportamiento aparentemente no compensatorio habrían escogido considerando sólo levemente, o incluso descartando, el atributo tiempo de viaje. No obstante, los valores del tiempo estimados con ambos bancos de datos (completo y depurado) están dentro del rango de valores obtenidos comúnmente en estudios para Chile (por ejemplo, Munizaga *et al.*, 2008 obtienen un valor subjetivo de reducción del tiempo de viaje de 3,5 USD/hora, lo que equivale a alrededor de 30 \$/min).

Otro resultado interesante, es que los valores correspondientes a los datos depurados (entre 50 y 60 \$/min) son bastante similares a los obtenidos en los modelos de PR (por ejemplo, 65 \$/min en el modelo 4 de la Tabla 5-6). En cuanto a las razones de compromiso, prácticamente no hay diferencias de percepción entre llegar antes y después del horario preferido de llegada (*SDE* y *SDL*), lo que contrasta con lo encontrado en PR (el parámetro de *SDL* era casi el doble que el de *SDE*). Por otro lado, mientras que en los modelos derivados de los datos de PD depurados la percepción del tiempo de viaje es mayor que las demoras horarias (*SDE* y *SDL*), en la base de PR ocurre lo contrario, ya que generalmente la desutilidad asociada a un minuto de tiempo de viaje equivale a un minuto temprano y medio minuto tarde.

Tabla 5-8: Resultados estimación modelos MNL banco de datos de PD depurado

Coeficiente	Modelos Base PD			
	2		3	
	Valor	Test-t robusto	Valor	Test-t robusto
Nivel de Servicio				
Tiempo de viaje	-0.0323	-5.31	-0.0291	-4.87
Incertidumbre % en tiempo de viaje	-0.948	-1.23	-1.01	-1.28
Comfort - Ir sentado	0.334	2.93	0.357	2.53
Costo	-0.00038	-1.9		
Restricciones de Programación				
Demora horaria por llegar temprano (<i>SDE</i>)	-0.0223	-7.95	-0.02	-5.27
Demora horaria por llegar tarde (<i>SDL</i>)	-0.0223	-7.04	-0.0388	-4.91
Llega tarde al trabajo (<i>D_TARDE</i>)	-0.302	-2.59	-0.273	-2.34
Interacciones				
Costo/tasa salarial (<i>Costw</i>)	-0.968	-1.76	-1.82	-4.65
<i>SDE</i> - Edad entre 30 y 55 años			-0.00706	-2.17
<i>SDE</i> - Edad mayor a 55 años			-0.00931	-2.24
<i>SDL</i> - Edad entre 30 y 55 años			-0.00725	-2
<i>SDE</i> - Ingreso medio bajo			0.00921	2.72
<i>SDE</i> - Ingreso alto			0.00637	1.91
<i>SDL</i> - Ingreso medio bajo			0.0187	2.12
<i>SDL</i> - Ingreso medio alto			0.0266	3.32
<i>SDL</i> - Ingreso alto			0.0206	2.55
ASC				
Salir antes	-0.659	-4.3	-0.687	-4.42
Salir después	-0.873	-4.88	-0.866	-4.81
Cambio modo	-1.43	-14.48	-1.43	-14.55
Log-verosimilitud	-1302.42		-1288.93	
rho-cuadrado ajustado	0.257		0.261	
# parámetros	11		18	
# individuos	255		255	
# observaciones	1275		1275	
Valor subjetivo tiempo viaje (\$/min)	50		59	
Razones de compromiso				
<i>SDL/SDE</i>	1.0			
<i>SDE/Tvia</i>	0.7			
<i>SDL/Tvia</i>	0.7			

Fuente: Elaboración propia

El modelo 3 de la Tabla 5-8 incorpora una serie de variaciones sistemáticas en los gustos. A partir de estas interacciones se puede inferir que a las personas entre 30 y 55 años les desagrada tener que salir del hogar más temprano o tarde que lo usual. Esto se puede atribuir al hecho que, en general, las personas en este rango de edad o ya son o están comenzando a ser jefes de familia; por lo tanto, tener que modificar la hora de inicio de viaje no es tan sencillo debido a restricciones personales (por ejemplo, estar más tiempo con los hijos antes de que vayan al colegio). Asimismo, a las personas con 55 o más años, les causa mayor desutilidad tener que iniciar antes su viaje al trabajo. Esto no es tan intuitivo, pero puede deberse a que las personas de mayor edad tienen, en general, una rutina de trabajo establecida. Sin embargo, si esto fuese completamente cierto, sería esperable que percibieran aún más negativamente el tener que llegar más tarde, lo que no se detectó en los resultados. Esto quizás se explica porque las personas de esta edad ya inician sus viajes temprano en la mañana, por lo que iniciarlos más temprano aún, no es una alternativa muy atractiva.

Por otra parte, del mismo modelo 3 se desprende que a las personas de ingreso mayor a \$ 400.000 (esto incluye a las categorías de ingreso medio-bajo, medio-alto y alto) no parece causarles tanta desutilidad llegar tarde en comparación a las personas de ingreso bajo (menor a \$ 400.000). Es posible que esto sea el resultado de cómo distintas personas enfrentaron la encuesta: las de ingreso bajo, que generalmente tienen menos autonomía para tomar decisiones respecto a la jornada laboral, pueden haber asumido que todos los días se atrasaba el viaje (i.e. una decisión rutinaria), y por ende, que era menos factible para ellos. En la sección correspondiente a los resultados de modelos mixtos PR/PD hay una discusión más detallada respecto a las distintas perspectivas temporales que pudo haber tenido el experimento PD.

A fin de comparar con los resultados de PR, se estimó un modelo que incorporaba la flexibilidad laboral como una variable explicativa, tal como se realizó en la sección 5.1.2. Los resultados de este modelo están disponibles en la Tabla C-1 del ANEXO C. Como es posible apreciar, todas las interacciones entre los parámetros de *SDE* y *SDL* y las variables de flexibilidad, son positivas como era de esperar; sin embargo, su magnitud y significancia es bastante baja. Por ende, es posible que las personas hayan

respondido sin considerar a cabalidad la flexibilidad horaria que estipularon en la componente de PR. Esto es consistente con el contexto que se le presentó a los encuestados en el experimento de PD, ya que se contemplaba la implementación de un sistema flexible de horarios de entrada y salida del trabajo (una discusión más detallada se presenta en la sección 6.2). Por otro lado, resulta interesante notar que mientras en PR la flexibilidad horaria fue un factor decisivo al momento de introducir heterogeneidad al modelo, en PD lo fue el ingreso.

Un último aspecto interesante de resaltar es el signo y significancia de las constantes específicas. Según los valores de las tablas 5-7 y 5-8, al aislar el efecto de las variables explicativas, a las personas les desagrada menos tener que adelantar su viaje que atrasarlo, pero en último lugar se encuentra cambiar de modo. Esto es, los encuestados serían más sensibles a cambios de horario que a cambios de modo, lo que se puede deber a la presencia de hábito o inercia (Yañez y Ortúzar, 2009).

5.2.2 Modelos Logit Mixto (ML)

Una desventaja de los modelos estimados a partir de datos de PD es que pueden contener sesgos porque una misma persona responde varias situaciones de elección (lo que se conoce como *efecto panel*). Para controlar por este efecto es necesario estimar modelos con estructuras más avanzadas y flexibles, como por ejemplo el Logit Mixto de componentes de error (ECL).

La forma de introducir el componente de error que incorpora el efecto panel no es trivial, ya que distintas formas pueden llevar a distintos resultados. A fin de comparar, en esta investigación se emplearon dos formas: (i) efecto panel de la forma más simple y común, que es agregar un componente de error en todas las alternativas menos una, y (ii) efecto panel según especifica Arellana (2012), que consiste en agregar a cada alternativa una multiplicación de componentes de error, como la que se muestra a continuación (siguiendo la notación de Biogeme):

$$SIGMA * ZERO [SIGMA_i] \quad (5.5)$$

donde *ZERO* es la media del componente de error (fijada en cero) y *SIGMA_i* es la desviación estándar (de una normal en este caso) correspondiente la alternativa *i*, que se

fija en uno para no ser estimada y considera la misma instancia aleatoria (*draw*) para todas las respuestas de un mismo individuo, incorporando así el efecto panel; *SIGMA* es la desviación estandar transversal a todas las alternativas y se debe estimar.

La primera forma tiene la desventaja de generar correlación indeseada entre las alternativas del modelo (Ortúzar y Willumsen, 2011, pag. 308), en cambio la segunda evita esto, ya que en ninguna alternativa el componente de error toma el mismo valor. En este caso, los componentes de error según (i) se agregaron de forma de inducir correlación entre las alternativas que no consideraban cambio de modo. Cabe destacar que se intentó estimar, sin éxito, un modelo HL con una estructura de árbol similar a la estructura de correlación anterior.

Los modelos ECL_1 y ECL_2 de la Tabla 5-9 utilizan ambas formas de introducir las componentes de error. Es fácil ver que la log-verosimilitud del modelo ECL_2, que acomoda el efecto panel según el enfoque discutido por Arellana (2012), es bastante superior a la de los modelos MNL de la sub-sección anterior y a la del modelo ECL_1. Además, en el modelo ECL_2 aumenta la significancia y el valor absoluto del coeficiente del tiempo de viaje y de las demoras horarias (*SDE* y *SDL*), asemejándose a los coeficientes estimados en PR; sin embargo, disminuye el de la incertidumbre en los tiempos de viaje. Esto, a su vez, implica que las razones entre los parámetros de *SDE* y *SDL*, y el del tiempo de viaje, aumenten.

También se estimó un modelo Logit Mixto de parámetros aleatorios (ML) para las variables tiempo de viaje, costo/w, *SDE* y *SDL*, especificando una distribución normal para cada uno, y usando el mismo *draw* para cada individuo (no observación); esto permite tratar adecuadamente el efecto panel. Los resultados de este modelo también se presentan en la Tabla 5-9 (ML_1), apreciándose que la media de los coeficientes aleatorios es consistentemente más alta que en los otros modelos¹³.

¹³ Esto se debe al mayor factor de escala (implícito en la estimación de cualquier modelo logit) debido a que se resta de la varianza del error Gumbel IID variabilidad que es ahora tomada por los parámetros aleatorios; este efecto, esperable cuando existe heterogeneidad de gustos, es discutido por Sillano y Ortúzar (2005).

Tabla 5-9: Resultados estimación modelos Logit Mixto (ML) base PD depurada

Coeficiente	Modelos Base PD					
	<i>ECL_1</i>		<i>ECL_2</i>		<i>ML_1</i>	
	Valor	Test-t robusto	Valor	Test-t robusto	Valor	Test-t robusto
<i>Nivel de Servicio</i>						
Tiempo de viaje	-0,0314	-3,31	-0,0406	-3,42	-0,0558	-4,31
Incertidumbre % en tiempo de viaje	-1,45	-1,8	-1,01	-0,95	-1,83	-1,79
<i>Restricciones de Programación</i>						
Demora horaria por llegar temprano (<i>SDE</i>)	-0,023	-6,18	-0,036	-6,38	-0,0568	-8,2
Demora horaria por llegar tarde (<i>SDL</i>)	-0,023	-5,96	-0,0394	-7,49	-0,0683	-7,78
Llega tarde al trabajo (<i>D_TARDE</i>)	-0,348	-1,95	-0,467	-1,97	-0,431	-1,91
<i>Interacciones</i>						
Costo/tasa salarial (<i>Costw</i>)	-2,01	-3,2	-2,64	-2,55	-6,18	-4,56
<i>ASC</i>						
Salir antes	-0,586	-2,77	-1,1	-3,41	-0,718	-2,73
Salir después	-0,796	-3,56	-1,3	-4,04	-0,944	-3
Cambio modo	-2,23	-8,7	-2,36	-9,02	-2,01	-9,17
<i>Componente de error de efecto panel</i>						
Sigma	-1,67	-7,11	1,89	13,05		
<i>Desviación estándar parametro aleatorio</i>						
Tiempo de viaje					0,0897	4,47
<i>SDE</i>					0,0436	5,99
<i>SDL</i>					-0,0506	-7,76
<i>Costw</i>					-6,51	-3,78
Log-verosimilitud	-1280,55		-1085,62		-1128,55	
rho-cuadrado ajustado	0,27		0,38		0,354	
# parámetros	10		10		13	
# individuos	255		255		255	
# observaciones	1275		1275		1275	
# Draws	1000		1000		1000	
Valor subjetivo tiempo viaje (\$/min)	58		57		33	
Razones de compromiso						
<i>SDL/SDE</i>	1,0		1,1		1,2	
<i>SDE/Tvia</i>	0,7		0,9		1,0	
<i>SDL/Tvia</i>	0,7		1,0		1,2	

Fuente: Elaboración propia

Sin embargo, su desviación estándar es muy alta, en especial para el tiempo de viaje y costo/w, en los cuales es hasta mayor que la media. Esto revela que en la muestra hay una alta heterogeneidad en los gustos, pero también implica que para algunas

personas estos parámetros van a resultar positivos, lo que sería contra intuitivo (ver la discusión en Sillano y Ortúzar, 2005).

En términos de bondad de ajuste, el modelo de componentes de error ECL_2 es superior a este modelo de parámetros aleatorios. Se observa también que las constantes específicas (ASC) de los modelos de la Tabla 5-9 son menos significativas que las ASC de los modelos MNL, en especial para la alternativa cambio de modo. Esto implica que una parte considerable de las preferencias no explicadas por ese modelo, y que eran capturadas por la constante específica de alternativa de cambio de modo, se pueden atribuir al efecto panel.

Por último, se estimó un modelo ML que incorpora el efecto panel de la misma forma que el modelo ECL_2 de la Tabla 5-9, pero agrega las variables Costo y Costow en las funciones de utilidad. Los resultados se muestran en la Tabla C-2 del ANEXO C.

5.3 Análisis de resultados modelos PR/PD

A continuación se presentan los resultados de los modelos mixtos de PR/PD usando ambos bancos de datos de PR de 15 min (PR_15_1 y PR_15_2) y el banco de datos de PD sin considerar individuos lexicográficos.

5.3.1 Flexibilidad laboral según disponibilidad de alternativas

Previo a la estimación de modelos mixtos de PR/PD se debió determinar cuáles parámetros serán considerados “comunes” entre las funciones de utilidad de PR y PD. Para ello se debe hacer una comparación entre los coeficientes obtenidos en modelos estimados con los datos de PR y PD en forma separada (Louviere *et al.*, 2000), a fin de hacer una estimación preliminar del factor de escala que relaciona las varianzas en ambos ambientes (el sustento teórico de este tema se presenta en el ANEXO A).

En la Tabla 5-10 se muestran los coeficientes de los mejores modelos PR y PD (MNL y ML) presentados en las secciones anteriores y la razón entre los parámetros estimados en PD y PR; esta división entre parámetros entrega un indicio del valor que podría tomar el factor de escala.

Tabla 5-10: Análisis de parámetros comunes PR/PD base PR_15_1

	θ_{PD}	θ_{PD}	θ_{PR}	θ_{PD}/θ_{PR}	θ_{PD}/θ_{PR}
	Modelo 2	Modelo ECL_3	Modelo 3	MNL	ML
Tiempo Viaje	-0.032	-0.047	-0.037	0.9	1.2
Costo	-0.00038	-0.00064	-0.00037	1.0	1.7
Costw	-0.968	-0.826	-0.784	1.2	1.1
<i>SDE</i>	-0.022	-0.038	-0.035	0.6	1.1
<i>SDL</i>	-0.022	-0.038	-0.074	0.3	0.5
<i>D_Tarde</i>	-0.302	-0.474	-0.217	1.4	2.2

Fuente: Elaboración propia

En el caso de los modelos MNL, los coeficientes de PR y PD de las variables tiempo de viaje, costo y Costow son bastante similares entre sí, mientras que los parámetros de *SDE* y *SDL* son menores en PD que en PR, en especial *SDL*. Estos resultados sugieren que los parámetros “comunes” debiesen ser los tres primeros. Por otro lado, en el caso del ML, los parámetros con razones más similares son el tiempo de viaje, Costow y la demora horaria por llegar temprano (*SDE*).

El primer modelo mixto PR/PD estimado fue un MNL que consideró todos los parámetros (excepto *SDE* y *SDL*) como comunes en PR y PD, según los resultados de la Tabla 5-10. De estos resultados destaca que el factor de escala sería estadísticamente igual a uno, implicando que los errores de los datos de PR y PD tendrían similar varianza. Asimismo, se corrobora que los parámetros de *SDE* y *SDL* son diferentes entre PR y PD. La diferencia de los parámetros de demora horaria (*SDE* y *SDL*) entre PR y PD se puede deber a diversos factores. En primer lugar, a diferencias temporales entre las situaciones de elección de PR y PD: mientras la primera es el resultado de un proceso de adaptación de largo plazo (cuántas horas trabajar, dónde trabajar, etc.), la elección en PD puede ser una decisión con un horizonte de tiempo menor. En otras palabras, los individuos pueden haber preferido cambiar la hora de inicio para las condiciones de tráfico presentadas en el experimento PD, pero no estarían dispuestos a cambiar la hora a la que comienzan su viaje al trabajo de forma permanente, por motivos laborales (debe trabajar más en la tarde para compensar, etc.) o personales (menos tiempo disponible con la familia en la mañana, etc.).

Tabla 5-11: Resultados modelos base PR/PD_15_1

Modelos Base PR/PD_15_1				
Coefficiente	I		I_ECL	
	Valor	Test-t robusto	Valor	Test-t robusto
Niveles Servicio				
Tiempo de viaje - PR/PD	-0,0306	-4,03	-0,0299	-3,29
Tiempo de acceso - PR	-0,0699	-5,33	-0,0697	-5,32
Tpo. transbordo público - PR	-0,105	-4,13	-0,105	-4,14
Tpo. transbordo público-privado - PR	-0,627	-3,8	-0,626	-3,78
Costo - PR/PD	-0,00037	-3,22	-0,000379	-3,17
Comfort - Ir sentado - PD	0,307	2,22		
Incertidumbre en tiempos de viaje - PD	-0,891	-1,17	-0,69	-0,89
Restricciones de Programación				
SDE - PR	-0,0351	-9,95	-0,0354	-9,81
SDE - PD	-0,0205	-3,88	-0,0252	-3,21
SDL - PR	-0,0732	-9,54	-0,0726	-9,21
SDL - PD	-0,0204	-3,39	-0,0279	-2,8
Llega tarde al trabajo - PR/PD	-0,256	-2,71	-0,282	-2,3
# paradas intermedias - PR	3,44	3,25	3,44	3,24
Socioeconómicas				
AUTLIC - PR	1,01	2,26	1,01	2,27
Interacciones				
Costo/tasa salarial (Costw) - PR/PD	-0,821	-3,57	-0,793	-3,28
Factor de escala PR/PD				
σ PR / σ PD *	1,09	0,32	1,45	0,96
Componente de error de efecto panel				
Sigma			1,30	3,06
ASC - PD				
Salir antes	-0,602	-2,97	-0,749	-2,35
Salir después	-0,812	-3,17	-0,912	-2,61
Cambio modo	-1,31	-3,79	-1,6	-2,9
Log-verosimilitud	-2423,09		-2205,03	
rho-cuadrado ajustado	0,256		0,323	
# parámetros	27		27	
# observaciones	1743		1743	

* test-t respecto a 1

Fuente: Elaboración propia

Este argumento es similar al propuesto por Borjesson (2008) para explicar la diferencia encontrada entre los parámetros de demora horaria (*SDE* y *SDL*) en sus modelos de PR y PD.

El segundo argumento, que está relacionado con el primero, es la posibilidad de que los parámetros de *SDE* y *SDL* en los modelos de PD estén influenciados

implícitamente por el contexto de la encuesta PD. El experimento de PD contemplaba que en el trabajo de los encuestados se había implementado un sistema flexible de horarios de entrada y salida del trabajo (además de un esquema de tarificación vial), por ende, los parámetros *SDE* y *SDL* capturan esta mayor flexibilidad. En tercer lugar, es posible que influya el hecho que en PR hay más alternativas de hora que en PD. Es esperable que mientras más alternativas de hora estén disponibles y más alejadas se encuentren de la alternativa escogida, más negativo será el efecto de las variables de *SDE* y *SDL*.

Más resultados sobre los modelos MNL mixtos PR/PD se encuentran en la Tabla D-1 del ANEXO D. El modelo 2 es similar al modelo 1 de la Tabla 5-11, pero también considera los parámetros *SDE* y *SDL* como “comunes” en PR y PD. Se destaca que la log-verosimilitud del modelo 2 es menor a la del modelo 1 y que el factor de escala disminuye a 0,49. Por su parte, el modelo 3 incorpora las mismas variaciones sistemáticas en los gustos discutidas en la sub-sección 5.1.1 en las funciones de utilidad de PR. Finalmente, en el mismo ANEXO D se comprueba que el modelo 1 mixto de PR/PD es superior a los modelos estimados con cada tipo de dato en forma separada.

También se estimó un modelo mixto PR/PD del tipo ML con componentes de error, incorporándose el efecto panel sólo en las respuestas de PD (ver modelo 1_ECL en la Tabla 5-11). Para que el efecto panel también incorpore la respuesta de PR es necesario agregar otros 99 componentes de error, aumentando considerablemente el tiempo y recursos computacionales necesarios para su estimación (a modo de referencia, la estimación del modelo 1_ECL tardó 2 días y 8 horas en un computador con un procesador Core i5 de 2.5 GHz). Es por ello que esta última estimación no se realizó, recomendándose como trabajo futuro para extender esta tesis. En cuanto a los resultados del modelo 1_ECL, se aprecia que el componente de error es bastante significativo y que la bondad de ajuste del modelo mejora bastante; por ende, se corrobora la presencia de efecto panel. Asimismo, aumenta el factor de escala, implicando que los datos de PR tienen mayor varianza, aun cuando el factor sigue siendo estadísticamente igual a 1. El valor absoluto de los coeficientes *SDE* y *SDL* en PD también aumenta, aunque siguen siendo bastante menores que sus símiles en PR.

Por último, en la Tabla D-3 del ANEXO D se muestran las valoraciones monetarias y razones de compromiso entre las variables *SDE*, *SDL* y tiempo de viaje para los modelos de la Tabla 5-11.

5.3.2 Flexibilidad laboral como variable explicativa

En esta sub-sección se presentan los modelos mixtos de PR/PD utilizando la base de PR que incorpora la flexibilidad laboral como una variable explicativa (PR_15_2). Una comparación de los parámetros de modelos de PR y PD estimados de forma independiente se muestra en la Tabla 5-12. De ella se puede desprender que los parámetros del tiempo de viaje, costo, *Costow* y de la variable ficticia asociada a llegar tarde, son muy similares entre PR y PD. Por otro lado, los parámetros asociados a las demoras horarias difieren aún más entre PR y PD que en el caso de la sub-sección anterior. De esta manera, se consideraron sólo como parámetros comunes el tiempo de viaje, costo, *Costow* y *D_Tarde*.

Los resultados de la estimación mixta de PR/PD para modelos de elección de modo y hora de inicio del viaje del tipo MNL y ML, se detallan en la Tabla 5-13. Se destaca que el factor de escala también es estadísticamente igual a uno, implicando que las varianzas de los errores de PR y PD son muy similares. Los parámetros de *SDE* y *SDL* en PD son los mismos para todos los individuos, mientras que, como era esperable, en PR las personas con alta flexibilidad horaria son las que menos valoran las demoras horarias, seguido por aquellos con flexibilidad media y en último lugar las personas con flexibilidad horaria baja. Asimismo, se comprueba que efectivamente las variables de *SDE* y *SDL* son distintas en PR y PD, inclusive para las personas con flexibilidad alta en PR.

Al realizar el test de razón de verosimilitud (LR) para comparar modelos mixtos de PR/PD, se obtiene un estadígrafo de 0,63 para el modelo 5. Este número es menor que el valor de la distribución χ^2 con cuatro grados de libertad al 95% de confianza (9,49), por lo que el modelo mixto de PR/PD es preferible a los modelos estimados en forma separada.

Tabla 5-12: Análisis parámetros comunes PR/PD base PR_15_2

	θ_{PD}	θ_{PR}	θ_{PD}/θ_{PR}
	Modelo 5	Modelo 4 Base PR_15_2	
Tiempo Viaje	-0.029	-0.038	0.8
Costo	-0.00044	-0.00037	1.2
Costw	-0.93	-0.81	1.2
<i>SDE</i>	-0.022	-0.077	0.3
<i>SDL</i>	-0.027	-0.101	0.3
<i>D_Tarde</i>	-0.31	-0.38	0.8
<i>SDE</i> - Flexibilidad alta	0.0003	0.041	0.0
<i>SDE</i> - Flexibilidad media	0.0019	0.025	0.1
<i>SDL</i> - Flexibilidad alta	0.0011	0.041	0.0
<i>SDL</i> - Flexibilidad media	0.0035	0.020	0.2

Fuente: Elaboración propia

El modelo 5_ECL corresponde al modelo 5 anterior, con la diferencia que en las funciones de utilidad de PD se agrega un componente de error para capturar el efecto panel. Se comprueba su presencia pues la varianza del componente de error es significativa y aumenta considerablemente la bondad de ajuste del modelo. Asimismo, los coeficientes son similares al modelo MNL, con la excepción de que aumentan levemente los valores de las demoras horarias y el factor de escala, pero disminuye la significancia de la variable incertidumbre en los tiempos de viaje. Los indicadores rho-cuadrado ajustado tanto de los modelos MNL como ML de esta sub-sección (0,304 para el MNL), son consistentemente superiores a los de la sub-sección anterior (alrededor de 0,25 para el MNL). A esto se le debe agregar que, de acuerdo a los resultados del test de razón de verosimilitud (LR) para modelos mixtos de PR/PD, existe menos diferencia entre el modelo mixto de PR/PD y los estimados en forma independiente al considerar la flexibilidad como variable explicativa. Además, al agregar la flexibilidad se obtienen diferentes razones de compromiso para el *Scheduling Model*, pues esta permite introducir heterogeneidad según distintos niveles de flexibilidad horaria (estos valores se discuten en el siguiente capítulo). De esta forma, se puede concluir que la forma más adecuada de incorporar la flexibilidad horaria es mediante una variable explicativa y no restringiendo el conjunto de alternativas disponibles.

Tabla 5-13: Resultados modelos mixtos PR/PD base PR_15_2

Modelos Base PR/PD_15_2				
Coefficiente	4		4_ECL	
	Valor	Test-t robusto	Valor	Test-t robusto
<i>Niveles Servicio</i>				
Tiempo de viaje - PR/PD	-0.0318	-4.04	-0.0342	-3.51
Tiempo de acceso - PR	-0.071	-5.33	-0.0714	-5.34
Tpo. transbordo público - PR	-0.104	-4.05	-0.103	-4.03
Tpo. transbordo público-privado - PR	-0.632	-3.82	-0.638	-3.8
Costo - PR/PD	-0.000381	-3.26	-0.000376	-3.15
Comfort - Ir sentado - PD	0.326	2.22	0.438	2.33
Incertidumbre en tiempos de viaje - PD	-0.933	-1.17	-0.227	-0.28
<i>Restricciones de Programación</i>				
SDE - PR	-0.076	-11.35	-0.0767	-11.33
SDL - PR	-0.102	-10.79	-0.101	-10.31
SDE - PD	-0.0218	-3.87	-0.0283	-3.29
SDL - PD	-0.0212	-3.4	-0.0286	-2.85
Llega tarde al trabajo - PR/PD	-0.319	-2.96	-0.366	-2.65
# paradas intermedias - PR	3.5	3.17	3.51	3.21
<i>Socioeconómicas</i>				
AUTLIC - PR	0.947	2.09	0.946	2.09
<i>Interacciones</i>				
Costo/tasa salarial (Costw) - PR/PD	-0.85	-3.55	-0.829	-3.34
SDE - Flexibilidad horaria alta - PR	0.0401	4.4	0.0404	4.4
SDE - Flexibilidad horaria media - PR	0.0248	3.31	0.025	3.32
SDL - Flexibilidad horaria alta - PR	0.0416	2.76	0.0411	2.74
SDL - Flexibilidad horaria media - PR	0.0203	1.69	0.02	1.67
Tpo. transbordo público - Edad mayor a 30				
Tpo. acceso - Sexo (Hombre)				
<i>Factor de escala PR/PD</i>				
σ PR / σ PD *	1.04	0.15	1.32	0.78
<i>Componente de error efecto panel</i>				
Sigma (1000 draws)			1.43	3.17
Log-verosimilitud	-2497.91		-2273.37	
rho-cuadrado ajustado	0.304		0.365	
# parámetros	31		32	
# observaciones	1743		1743	

Fuente: Elaboración propia.

6 MODELOS DE ELECCIÓN DE MODO Y HORA DE INICIO DEL VIAJE: BASE 30 y 60 MINUTOS

En este capítulo se detallan los resultados de los modelos mixtos de PR/PD al usar los bancos de datos de PR de 30 y 60 min. La base de 30 min contiene cuatro períodos de tiempo entre las 7:00 y 9:00 am: (i) 7:00-7:30; (ii) 7:30-8:00; (iii) 8:00-8:30 y (iv) 8:30-9:00. La base de 60 min considera sólo dos períodos: (i) 7:00-8:00 y (ii) 8:00-9:00. El resto de los períodos de tiempo permanecen fijos (5:00-7:00, 9:00-11:00, 11:00-14:00). De esta forma, el número de alternativas combinadas de modo-hora asciende a 63 (nueve modos por siete períodos) en la base de 30 min y a 45 (nueve modos por cinco períodos) en la de 60 min. Además, ambos bancos de datos incorporan la flexibilidad para atrasar o retrasar la hora de inicio de los viajes como variable explicativa, en concordancia con los resultados del capítulo anterior. Por su parte, el banco de datos de PD no considera a los individuos con comportamiento no compensatorio. Por último, cabe recordar que en las bases de 30min y 60min se utilizan niveles de servicios promediados a 30 minutos y 1 hora, respectivamente.

6.1 Resultados modelos mixtos PR/PD base 30 y 60 min

A continuación se presentan los valores de los coeficientes y test-t asociados a la estimación mixta de PR/PD usando ambas bases (30 y 60 min), mientras que en la siguiente sección se realiza un análisis comparativo entre estos modelos y aquellos estimados a partir de la base de 15 min.

6.1.1 Base 30 min

La Tabla 6-1 muestra los resultados del mejor modelo mixto de PR/PD del tipo MNL estimado a partir de la base de 30 min. Por razones de espacio no se incluyen entre los resultados las constantes específicas de modo ni las constantes específicas de PD.

Tabla 6-1: Resultados modelos mixtos PR/PD MNL base 30 minutos

Coeficiente	30_MNL	
	Valor	Test-t robusto
<i>Niveles Servicio</i>		
Tiempo de viaje - PR/PD	-0.0321	-4.07
Tiempo de acceso - PR	-0.0561	-4.66
Tpo. transbordo público - PR	-0.0738	-3.28
Tpo. transbordo público-privado - PR	-0.548	-2.85
Costo - PR/PD	-0.000388	-3.27
Comfórt - Ir sentado - PD	0.338	2.21
Incertidumbre en tiempos de viaje - PD	-0.96	-1.16
<i>Restricciones de Programación</i>		
SDE - PR	-0.0822	-10.88
SDL - PR	-0.0946	-11.11
SDE - PD	-0.0226	-3.82
SDL - PD	-0.0217	-3.38
Llega tarde al trabajo - PR/PD	-0.37	-2.96
# paradas intermedias - PR	3.36	2.96
<i>Socioeconómicas</i>		
AUTLIC - PR	0.935	2.09
<i>Interacciones</i>		
Costo/tasa salarial (Costw) - PR/PD	-0.854	-3.46
SDE - Flexibilidad horaria alta - PR	0.0419	4.34
SDE - Flexibilidad horaria media - PR	0.027	3.29
SDL - Flexibilidad horaria alta - PR	0.0345	2.63
SDL - Flexibilidad horaria media - PR	0.0147	1.37
<i>Factor de escala PR/PD</i>		
σ PR / σ PD *	1.00	0.00
Log-verosimilitud	-2217.49	
rho-cuadrado ajustado	0.343	
# parámetros	31	
# observaciones	1743	

* test-t respecto a 1.

Fuente: Elaboración propia

En este modelo todas las variables son significativas y con el signo esperado, destacándose los siguientes resultados. En primer lugar, al comparar los valores de la Tabla 6-1 con los modelos mixtos de PR/PD del banco de datos de 15 min (Tabla 5-13), se observa que en el banco de datos de 30 min los valores de los parámetros de *SDE* y

SDL (asociados a flexibilidad horaria baja) son más cercanos entre sí. Esto significa que para las personas con poca flexibilidad para modificar la hora de inicio del viaje al trabajo, la penalización por minuto en que se llega tarde o temprano al trabajo, es más similar. Este cambio también se puede producir a raíz de la menor correlación entre alternativas.

Segundo, el factor de escala sigue siendo estadísticamente igual a uno, y al agregar el factor de corrección de logaritmo natural, la magnitud de los coeficientes aumenta; sin embargo, disminuye la bondad de ajuste del modelo. Esto es similar a lo descrito en la sección 5.1. Por último, se intentó estimar modelos del tipo HL, pero al igual que lo sucedido en el capítulo anterior, estos no eran superiores a los modelos MNL.

6.1.2 Base 60 min

Los resultados del mejor modelo mixto para el banco de datos de 60 min se presentan en la Tabla 6-2. En el primer modelo (60_MNL), del tipo MNL, se consideran todas las variables que resultaron ser estadísticamente significativas, con la excepción de la interacción de *SDL* con la variable ficticia de flexibilidad horaria media; esto implica que personas con flexibilidad baja y media para atrasar el viaje penalicen de forma similar un minuto tarde al trabajo.

La tercera columna de la Tabla 6-2 contiene los resultados de la estimación del mejor modelo logit jerárquico (60_HL). En éste, las funciones de utilidad son las mismas que para el modelo 60_MNL, pero se introduce correlación en los errores aleatorios de acuerdo a la estructura jerárquica ilustrada en la Figura 6-1. Esto implica que las alternativas de hora de los modos auto-chofer, auto-acompañante, bus, bus-Metro, Metro y caminata, están correlacionadas. En otras palabras, en el caso de estos modos las personas son más sensibles a cambios de hora que a cambios de modo, lo que concuerda con lo encontrado en PD al analizar la magnitud de las constantes específicas (ASC) de los modelos estimados. El resto de las alternativas (taxi, taxi-colectivo y combinación-metro) se comporta como un MNL (i.e. son independientes entre sí).

Tabla 6-2: Resultados modelos mixtos PR/PD MNL y HL base 60 minutos

Coeficiente	<i>60_MNL</i>		<i>60_HL</i>	
	Valor	Test-t robusto	Valor	Test-t robusto
<i>Niveles Servicio</i>				
Tiempo de viaje - PR/PD	-0.0286	-4.03	-0.0277	-4.3
Tiempo de acceso - PR	-0.0472	-4.42	-0.0417	-4.32
Tpo. transbordo público - PR	-0.0524	-2.67	-0.0539	-3.04
Tpo. transbordo público-privado - PR	-0.361	-2.21	-0.36	-2.19
Costo - PR/PD	-0.000347	-2.97	-0.000354	-3.14
Comfort - Ir sentado - PD	0.292	2.2	0.291	2.24
Incertidumbre en tiempos de viaje - PD	-0.838	-1.16	-0.826	-1.15
<i>Restricciones de Programación</i>				
<i>SDE</i> - PR	-0.0777	-9.58	-0.0538	-7.17
<i>SDL</i> - PR	-0.0771	-12.06	-0.0496	-6.99
<i>SDE</i> - PD	-0.0196	-3.79	-0.0197	-3.96
<i>SDL</i> - PD	-0.0195	-3.35	-0.0193	-3.44
Llega tarde al trabajo - PR/PD	-0.267	-2.52	-0.299	-2.87
# paradas intermedias - PR	3.16	2.81	3.15	2.85
<i>Socioeconómicas</i>				
<i>AUTLIC</i> - PR	1.03	2.35	1	2.37
<i>Interacciones</i>				
Costo/tasa salarial (<i>Costw</i>) - PR/PD	-0.806	-3.19	-0.772	-3.31
<i>SDE</i> - Flexibilidad horaria alta - PR	0.0399	4.29	0.0291	4.03
<i>SDE</i> - Flexibilidad horaria media - PR	0.0216	2.45	0.0171	2.61
<i>SDL</i> - Flexibilidad horaria alta - PR	0.0203	1.91	0.0152	2.14
<i>SDL</i> - Flexibilidad horaria media - PR	0.00172	0.2	0.00426	0.71
<i>Factor de escala PR/PD</i>				
σ PR / σ PD *	1.14	0.46	1.13	0.45
Log-verosimilitud				
rho-cuadrado ajustado	-1982.77		-1968.48	
# parámetros	0.383		0.385	
# observaciones	31		37	
	1743		1743	

* test-t respecto a 1.

Fuente: Elaboración propia.

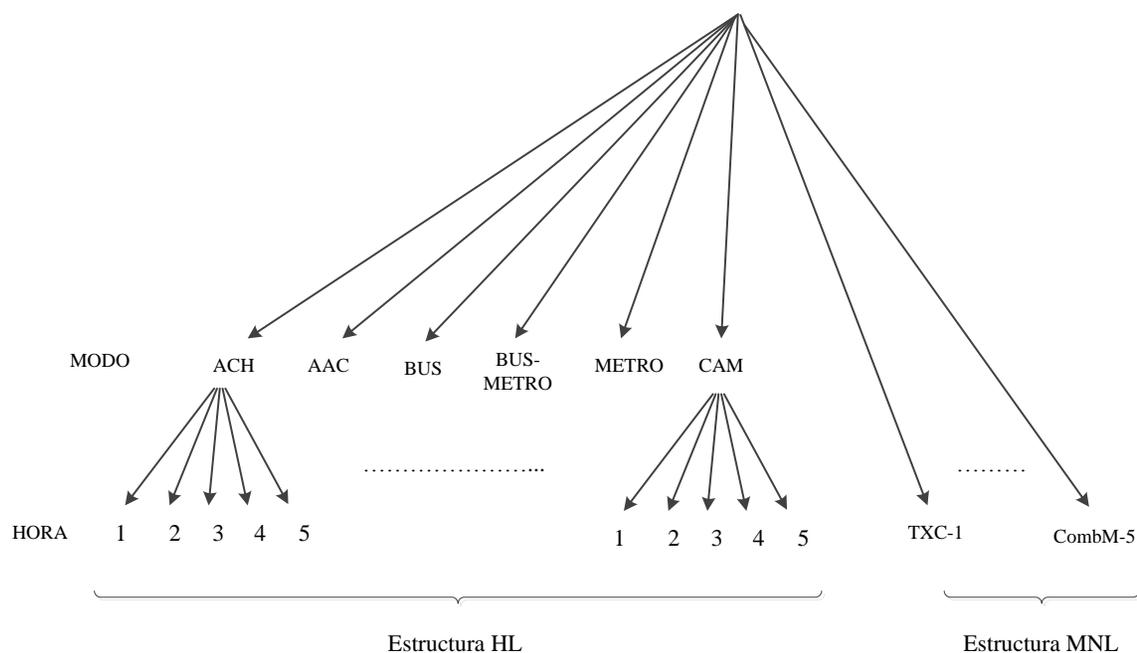


Figura 6-1: Estructura jerárquica modelo 2_HL base PR/PD de 60 min

Fuente: Elaboración propia

Los parámetros estructurales resultantes de estimar la estructura jerárquica de la Figura 6-1 se presentan en la Tabla 6-3. Biogeme entrega como salida el parámetro estructural de los nidos inferiores (λ), puesto que normaliza en uno al parámetro estructural del nido superior (β). Los parámetros λ de todos los modos son estadísticamente mayores a 1, resultando en valores de ϕ ($\phi = \beta/\lambda$) menores que uno, como requiere la teoría. Se postuló también una estructura jerárquica en que la elección de hora precedía a la elección de modo; sin embargo, en este caso ningún parámetro λ resultó ser significativamente mayor que uno.

Un resultado importante es que al incorporar la estructura jerárquica de la Figura 6-1 (modelo 60_HL), tanto los coeficientes puros de las demoras horarias *SDE* y *SDL* como sus interacciones con las variables de flexibilidad (a excepción del parámetro *SDL*-flexibilidad media) disminuyen en alrededor de un 30%. Esto implica que asumir que todas las alternativas son independientes entre sí (i.e. un MNL) conlleva una sobrevaloración de las penalizaciones por minuto de llegada tarde o temprano al trabajo.

Un aspecto que llama la atención del resultado anterior, es que a diferencia de lo acontecido en los modelos estimados a partir de las bases de 15 y 30 min, en que sólo uno o dos parámetros estructurales eran significativos, en este caso el modelo HL es claramente superior al MNL. Esto es algo contra intuitivo, porque en modelos con mayor desagregación temporal se esperaría una mayor correlación entre alternativas de tiempo. Así, es posible que la estructura de error del modelo sea más simple o acotada al utilizar intervalos de tiempo agregados, y se complejiza a medida que se desagrega el tiempo. En este caso, sería necesario emplear otras estructuras de correlación más flexibles. En la siguiente sección se discute una forma de realizar esto.

Tabla 6-3: Parámetros estructurales modelo 60_HL

<i>Nido</i>	λ	Test-t robusto*	$\phi=1/\lambda$
Auto-acompañante	1,7	1,78	0,58
Auto-chofer	1,6	2,01	0,62
Bus	2,6	2,6	0,39
Bus-metro	2,1	2,41	0,47
Caminata	2,6	2,36	0,38
Metro	2,0	2,54	0,50

* test-t respecto a 1

Fuente: Elaboración propia

Por último, interesa verificar que los modelos mixtos de PR/PD estimados en este capítulo sean superiores a los modelos estimados con cada tipo de dato en forma separada. Para ello se recurre al test LR para modelos mixtos de PR/PD, tomando como referencia el modelo 30_MNL para la base de 30 min y 60_HL para la base de 60 min. Para ambos modelos se obtiene un estadígrafo de 2, el cual es menor que el valor χ^2 para cuatro grados de libertad al 95% de confianza (9,49), por lo que se confirma – nuevamente - que los modelos mixtos de PR/PD son preferibles a los estimados en forma separada.

6.2 Comparación con modelos mixtos de PR/PD base 15 min

Una de las principales críticas a la utilización de modelos de elección discreta para modelar la hora de inicio de los viajes, es que la elección de la longitud de los períodos de tiempo queda a discreción del modelador. Por otra parte, la principal ventaja de usar

alternativas de tiempo más agregadas es que se reduce el número de alternativas y la correlación entre ellas; sin embargo, se pierde información valiosa, tanto para la estimación del modelo como para su posterior uso en fase predictiva. Esta información comprende tanto las elecciones de las personas como los niveles de servicio utilizados en la estimación del modelo. Por ejemplo, al usar períodos de una hora, para el modelo es lo mismo una persona que sale del hogar a las 7:05 que otra que sale a las 7:55. Además, en ambos casos se usan los niveles de servicio promediados para esa hora, cuando puede que existan diferencias importantes entre salir a las 7:05 y a las 7:55. Es por ello que tiene especial interés comparar el ajuste estadístico, grado de correlación entre las alternativas de hora, indicadores microeconómicos (valores subjetivos del tiempo y razones de compromiso), y predicciones realizadas por modelos mixtos de PR/PD estimados a partir de las bases de 15, 30 y 60 min. Los tres primeros aspectos se discuten en esta sección, mientras que el análisis de los distintos modelos en fase predictiva se presenta en el capítulo 7.

La Tabla 6-3 muestra un resumen comparativo, donde las primeras dos columnas corresponden a los modelos de 15 min del tipo MNL y ML de componentes de error, respectivamente. La tercera columna corresponde al modelo MNL de intervalos de 30 min, en tanto, la cuarta y la quinta son los modelos MNL y HL de 60 min, respectivamente. Cabe destacar que todos estos modelos incorporan la flexibilidad horaria como variable explicativa.

Como se aprecia en la Tabla 6-3, para los modelos MNL la bondad de ajuste - representada por el indicador rho-cuadrado ajustado ($\bar{\rho}^2$) - aumenta a medida que los períodos de tiempo son más amplios. Es más, los modelos MNL y HL de la base de 60 min son superiores al modelo ML de la base de 15 min.

Tabla 6-4: Resumen comparativo modelos mixtos PR/PD bases de 15, 30 y 60 minutos

Mmodelos mixtos PR/PD					
	<i>15_MNL</i>	<i>15_ECL</i>	<i>30_MNL</i>	<i>60_MNL</i>	<i>60_HL</i>
Rho-cuadrado ajustado ($\bar{\rho}^2$)	0,304	0,365	0,343	0,383	0,385
Valor subjetivo (\$/min)					
Tiempo viaje	52	57	52	51	50
Tiempo acceso	117	120	91	84	75
Tiempo transbordo público	171	173	120	93	96
Tiempo transbordo público-privado	1.041	1.070	891	643	644
Ratios PR					
Flexibilidad baja:					
<i>SDL/SDE</i>	1,3	1,3	1,2	1,0	0,9
<i>SDE/Tvia</i>	2,4	2,2	2,6	2,7	1,9
<i>SDL/Tvia</i>	3,2	3,0	2,9	2,7	1,8
Flexibilidad media:					
<i>SDL/SDE</i>	1,6	1,6	1,4	1,3	1,2
<i>SDE/Tvia</i>	1,6	1,5	1,7	2,0	1,3
<i>SDL/Tvia</i>	2,6	2,4	2,5	2,6	1,6
Flexibilidad alta:					
<i>SDL/SDE</i>	1,7	1,7	1,5	1,5	1,4
<i>SDE/Tvia</i>	1,1	1,1	1,3	1,3	0,9
<i>SDL/Tvia</i>	1,9	1,8	1,9	2,0	1,2
Ratios PD					
<i>SDL/SDE</i>	1,0	1,0	1,0	1,0	1,0
<i>SDE/Tvia</i>	0,7	0,8	0,7	0,7	0,7
<i>SDL/Tvia</i>	0,7	0,8	0,7	0,7	0,7

Fuente: Elaboración propia

Una explicación de lo anterior, es que al usar intervalos de tiempo más agregados, además de existir menos correlación entre alternativas, también resulta más fácil (para el modelo) replicar las elecciones realizadas por cada individuo, por lo que aumenta la bondad de ajuste.

También disminuyen, progresivamente, las valoraciones monetarias del tiempo de acceso y transbordo al pasar de intervalos de 15 min a 30 y 60 min. Este cambio es mayor para el tiempo de transbordo, con reducciones entre 55 y 60% al pasar de los modelos de 15 min a los de 60 min. Las valoraciones del tiempo de acceso (\$75/min) y transbordo, tanto público-público (\$96/min) como público-privado (\$644/min), del

modelo HL de 60 min parecen ser más razonables que las de los modelos de 15 min, que se consideran elevadas. Por otro lado, la valoración del tiempo de viaje permanece estable en torno a los \$50/min, producto de que los parámetros del tiempo de viaje, costo y costo/w son relativamente constantes.

También interesa analizar el impacto en las razones de compromiso que caracterizan al *Scheduling Model*. De acuerdo a los resultados de la Tabla 6-3, la razón entre llegar un minuto tarde y un minuto temprano al trabajo (SDL/SDE) en PR, decrece a medida que se usan períodos de tiempo más agregados, independientemente del nivel de flexibilidad disponible para modificar la hora de inicio del viaje. Es decir, un minuto tarde se penaliza de forma más similar, e incluso menor, que cada minuto que se llega temprano al trabajo. Esto último se asemeja más a lo encontrado en PD.

En cuanto a las valoraciones relativas del tiempo de viaje y las demoras horarias (SDE y SDL), no se percibe un patrón evidente al variar la longitud de los intervalos de tiempo de las alternativas de PR, aunque en todos los modelos presentados se cumple que aquellas personas con flexibilidad horaria baja penalizan más un minuto tarde o temprano que un minuto de tiempo de viaje en comparación a individuos con mayor flexibilidad. No obstante, en el caso del modelo HL de la Figura 6-1 ocurre un cambio considerable, puesto que las razones SDE/T_{via} y SDL/T_{via} decrecen consistentemente entre un 30% y 40% en relación al modelo MNL. Como ya fue mencionado, esto se puede atribuir a que en el modelo HL es posible incorporar una fuente de correlación entre alternativas.

Asimismo, siguen existiendo diferencias para las razones SDE/T_{via} y SDL/T_{via} entre las funciones de utilidad de PR y PD, aunque se acortan en el caso del modelo HL. Estas diferencias nuevamente se pueden atribuir a las diferentes perspectivas temporales entre las situaciones de elección de PR (largo plazo) y PD (corto plazo). En PR, el tener que modificar de forma permanente la hora de inicio del viaje puede implicar menos tiempo en el hogar o para realizar otras actividades en la mañana o la tarde. Por otro lado, en el experimento PD algunas personas quizá estaban dispuestas a cambiar su viaje sólo de manera esporádica y no en forma regular dado los niveles de congestión presentados; por ende, se obtiene una mayor valoración relativa del tiempo de viaje. Otra

explicación plausible y complementaria a la anterior (tal como se mencionó en la subsección 5.3.1), es que dentro del contexto del experimento de PD se contempló la implementación de un sistema flexible de horarios de entrada y salida del trabajo (además de un esquema de tarificación vial). En este, los encuestados podían modificar la hora de inicio del viaje (i.e. no tomaban en cuenta la flexibilidad horaria declarada en PR), pero se les imponía la condición de deber trabajar la misma cantidad de horas semanales y que, además, debían considerar sus restricciones personales (i.e. realizar actividades con la familia) al momento de escoger. Esto último es reafirmado por el hecho que las interacciones entre las demoras horarias (*SDE* y *SDL*) y el nivel de flexibilidad horaria reportado, que en PR si funcionaron, en PD no fueron significativas (la diferencia de estos parámetros entre PR y PD se puede apreciar en la Tabla 5-12). Es más, en el modelo HL de la Tabla 6-3, las razones de compromiso *SDE/Tvia* y *SDL/Tvia* para personas con flexibilidad alta son similares a las razones de PD. En otras palabras, las personas con una alta capacidad para modificar la hora de inicio de sus viajes penalizarían (en PR) llegar un minuto tarde o temprano al trabajo de manera similar que el total de los individuos en PD.

Dado que en los modelos de 15 y 30 min no fue posible incorporar exitosamente una estructura de correlación mediante el modelo HL, se postuló una estructura más flexible. Esta consistió en agregar a todas las funciones de utilidad de PR, un componente de error (distribuido normal) proporcional a la diferencia de tiempo, en valor absoluto, entre la alternativa correspondiente y la alternativa escogida, tal como se detalla en las siguientes ecuaciones (según notación en Biogeme):

$$TIMEDIFF_i * ZERO [SIGMA] \quad (6.1)$$

$$TIMEDIFF_i = Hora_Inicio_i - Hora_Inicio_{Eleccion} \quad (6.2)$$

donde *i* corresponde a la alternativa de hora. Esto es similar a lo realizado por de Jong *et al.* (2003), sólo que ellos lo hicieron para modelos estimados a partir de datos de PD (y en nuestro caso es para datos de PR). Para dos alternativas de hora cualesquiera, *i* y *j*, mientras más amplia sea la diferencia $TIMEDIFF_i - TIMEDIFF_j$, menor debiese ser la correlación entre el par de alternativas. Esto se traduce en que a mayor varianza del

componente de error (*SIGMA*), menor es el grado de substitución o correlación entre alternativas de hora.

Lamentablemente, los modelos propuestos en el inciso anterior no convergieron (se alcanzó el máximo número de iteraciones del algoritmo de optimización). Para solucionar esto se podría aumentar el máximo número de iteraciones o emplear otro algoritmo de optimización. Sin embargo, dado que en general Biogeme tardó tres días en que el modelo llegase al máximo número de iteraciones (el número de alternativas del modelo es muy alto, variando entre 45 y 99), aplicar las soluciones propuestas está fuera del alcance de esta tesis, y se propone como futuro tema para extender esta investigación. Otra alternativa para medir la correlación, podría ser utilizar modelos más generales que el HL y que tampoco requieran técnicas de simulación, como por ejemplo el modelo OGEV.

Como conclusión de este capítulo, se destaca que al usar intervalos de tiempo más desagregados, el modelo pierde capacidad explicativa y existe una tendencia a sobrevalorar el tiempo de acceso y transbordo. Sin embargo, no fue posible medir empíricamente el impacto de la correlación al desagregar los períodos de tiempo, algo que se propone como futura investigación.

7 MODELOS EN FASE PREDICTIVA

En este capítulo se discute la capacidad predictiva de los modelos estimados en los capítulos anteriores. En particular, interesa investigar tres fenómenos. Primero, dado que estos modelos no cuentan con constantes específicas asociadas a los períodos horarios, ellos no son necesariamente capaces de reproducir las elecciones de hora observadas en la muestra de estimación; por ello, interesa evaluar esta capacidad. En segundo lugar, se estimó modelos sin las variables de restricción de programación de actividades (*SDE*, *SDL*, variable ficticia de llegar tarde al trabajo, y número de paradas intermedias en el viaje al trabajo), ya que en general las grandes encuestas de movilidad no recopilan esta información y tampoco se cuenta con predicciones de ellas. Por último, se propuso un escenario hipotético de tarificación vial y se comparó las predicciones de los modelos con y sin variables de restricción de programación de actividades, al igual que modelos de intervalos de diferente longitud (15, 30 y 60 min). A partir de estos últimos resultados se entregan sugerencias y recomendaciones para los tomadores de decisión que estén evaluando la posibilidad de implementar un esquema de tarificación vial en Santiago.

Se reconoce que lo ideal para realizar las predicciones hubiese sido contar con una base de validación, o en su defecto, una base de datos simulada. Lo primero no se hizo porque la cantidad de observaciones disponibles para la estimación era muy pequeña, mientras que lo segundo se consideró fuera del alcance de esta tesis, y se sugiere como una extensión de esta.

7.1 Predicción de elecciones horarias observadas

A continuación se presentan las predicciones de elección de periodos de hora de inicio del viaje para los modelos de 15, 30 y 60 min estimados en los capítulos anteriores. Los modelos representativos para cada longitud de intervalo son: modelo 5 de la Tabla 5-13 para intervalos de 15 min; modelo MNL_30 de la Tabla 6-1 para la base de 30 min; y modelos MNL_60 y HL_60 de la Tabla 6-2 para periodos de 60 min. Se utilizó la técnica de enumeración muestral (Ortúzar y Willumsen, 2011) para obtener predicciones agregadas.

La Figura 7-1 muestra una comparación entre las elecciones predichas y las observadas para la base de 15 min. Aunque existen diferencias entre ambas, en general el modelo reproduce razonablemente las elecciones observadas, encontrándose una diferencia de hasta un 20% en el periodo 8:15-8:30. Otro aspecto a destacar, es que la curva modelada tiene una forma más suave (o menos discontinua) que la curva observada.

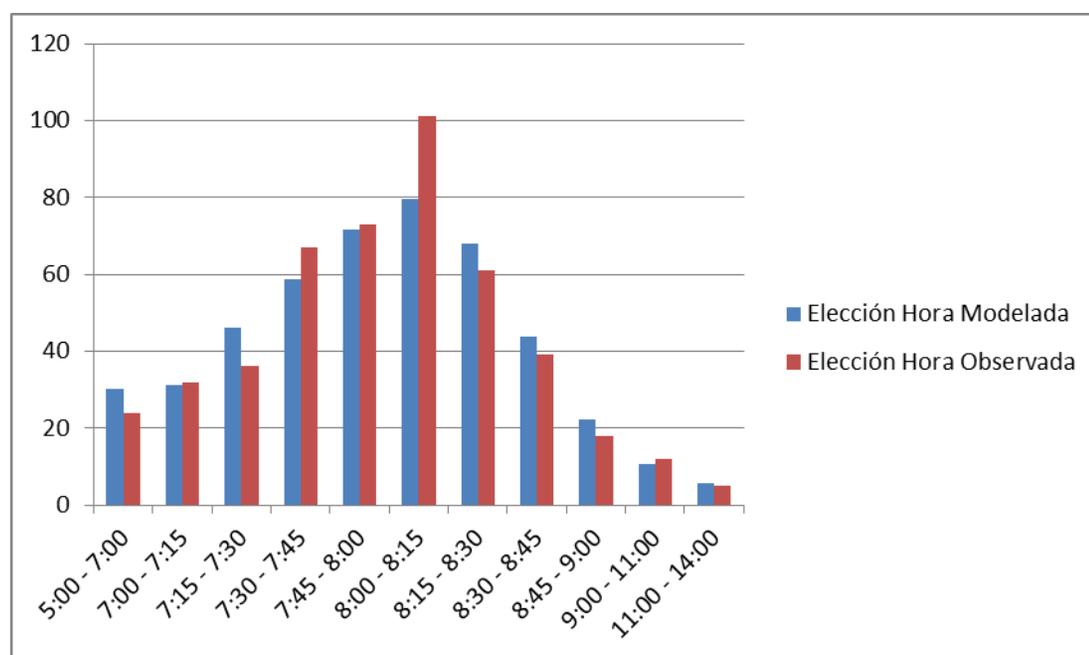


Figura 7-1: Elecciones de hora modeladas versus observadas modelo PR/PD base 15 minutos

Fuente: Elaboración propia

Por su parte, la Figura 7-2 muestra la misma información solo que para la base de 30 min. En este caso, las predicciones parecen ser más acertadas que para el modelo de 15 min; los períodos que presentan mayores diferencias son los intervalos 5:00-7:00 y 8:30-9:00, con diferencias de 30% y 15%, respectivamente.

Las predicciones para los modelos MNL y HL de 60 min se muestran en la Figura 7-3. El aumento en el acierto de las predicciones es pequeño en comparación a la base de 30 min, aunque para el intervalo de 9:00-11:00 la precisión de las predicciones disminuye. Asimismo, el modelo HL es levemente superior al modelo MNL.

Como era de esperar, los modelos con mayor bondad de ajuste dieron lugar a mejores predicciones, aunque la ganancia al pasar del banco de 30 min al de 60 min es marginal. Además, se destaca que las variables asociadas a la restricción de programación de actividades son capaces de capturar las preferencias temporales de los usuarios del sistema de transporte.

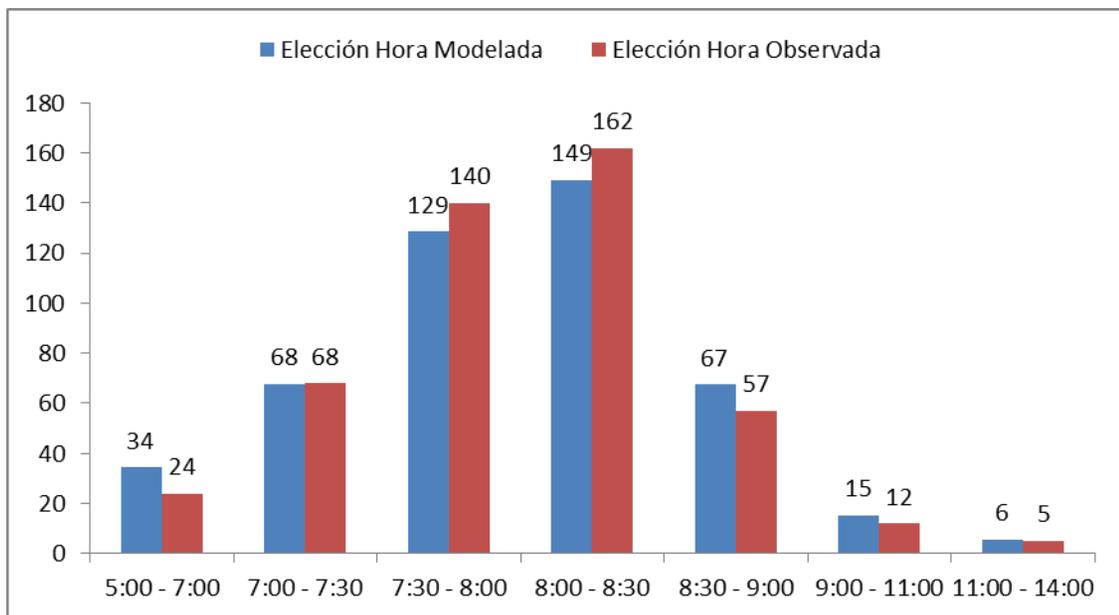


Figura 7-2: Elecciones de hora modeladas versus observadas modelo PR/PD base 30 minutos

Fuente: Elaboración propia.

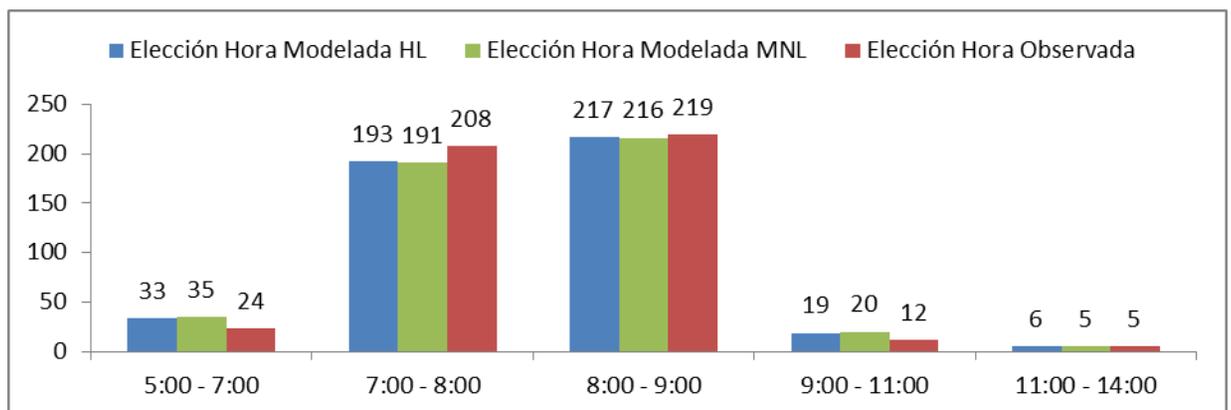


Figura 7-3: Elecciones de hora modeladas versus observadas modelos PR/PD base 60 minutos

Fuente: Elaboración propia.

7.2 Modelos sin variables de restricción de programación de actividades

En esta sección se presentan los resultados de la estimación de modelos de elección de modo y hora, sin incluir las variables de restricción de programación de actividades. En ausencia de estas variables, se agrega constantes específicas (ASC) de hora que permitan captar las preferencias agregadas por viajar en un determinado periodo. Ben-Akiva y Abou Zeid (2012) demuestran que, bajo ciertos supuestos, las preferencias temporales son constantes según estrato sociodemográfico y, por ende, pueden ser capturadas al hacer interactuar a las constantes específicas con variables socioeconómicas. Un supuesto hecho por los autores es que para viajes más sensibles a la hora de llegada que a la hora de inicio (como por ejemplo el viaje al trabajo, pues la restricción es la hora de entrada al trabajo), se debiera modelar la hora de llegada y no la hora de inicio del viaje. Esto no se cumple en esta investigación, ya que se modeló la hora de comienzo del viaje al trabajo y no la de llegada al destino, debido a que eso permitió simplificar considerablemente el cálculo de niveles de servicio y la generación de los bancos de datos. Los autores mencionan que para corregir este problema sería necesario interactuar las constantes específicas con variables de duración del viaje. Sin embargo, dada la limitada cantidad de observaciones del banco de datos de PR, sería difícil obtener parámetros representativos para las interacciones entre todas las constantes específicas con variables socioeconómicas y de duración del viaje. Es por ello que se estimó modelos que utilizaban solo las constantes específicas de hora, sin considerar las interacciones antes mencionadas.

Cabe mencionar que en los modelos de esta sección se agregó el factor de corrección (logaritmo natural) propuesto por Ben-Akiva y Abou-Zeid (2012). Al agregarlo, la bondad de ajuste no varió pero las constantes específicas (ASC) de hora de los periodos de mayor resolución temporal (por ejemplo, 7:00-7:15 en comparación a 5:00-7:00) aumentaron su valor.

La Tabla 7-1 muestra los resultados de los modelos de elección de modo y hora en fase predictiva, destacándose dos aspectos. Primero, la bondad de ajuste es bastante menor que en los modelos que incluyen variables de restricción de programación de

actividades. Esto era esperable, ya que al no considerar estas variables el modelo pierde capacidad explicativa y sólo captura las preferencias temporales agregadas de la muestra (en lugar de las preferencias individuales). Lo segundo, es que las constantes específicas de hora son muy significativas, en especial para los periodos más escogidos por los encuestados, lo que contrasta con la menor significancia de las constantes de modo. Esto último se debería a que la elección de modo está explicada, en mayor medida, por las variables de nivel de servicio que por las variables de restricción de programación de actividades.

7.3 Predicción de escenario hipotético: tarificación vial

A fin de comparar las predicciones efectuadas por los diversos modelos estimados en estas tesis, se postuló un escenario hipotético de tarificación vial. En éste, a las alternativas de hora en que finaliza el viaje (i.e. hora de inicio + duración del viaje) en los periodos 7:00-8:00 y 8:00-9:00, para el modo auto-chofer, se agregó una tarifa fija de \$2.000 y \$3.000, respectivamente. Para cada longitud de intervalos de tiempo en estudio (15, 30 y 60 min) se realizaron tres predicciones: (i) modelo con los parámetros de *SDE* y *SDL* de PR (denominado *SD-PR*); (ii) modelo con los parámetros de *SDE* y *SDL* de PD, deflactados por el factor de escala (denominado *SD-PD*); y (iii) modelo sin las variables de restricción de programación de actividades (denominado *ASC*). Se decidió considerar un modelo que utilizaba los parámetros de *SDE* y *SDL* de PD, porque en la actualidad no hay un esquema de tarificación vial en Santiago y justamente dentro de los objetivos del experimento de PD se encontraba evaluar una política de este tipo.

La Figura 7-4 muestra las predicciones realizadas por los tres modelos anteriores (*SD-PR*, *SD-PD* y *ASC*) utilizando intervalos de 15 min. A modo de comparación se presentan las elecciones observadas en la encuesta, esto es, sin la aplicación del esquema de tarificación vial.

Tabla 7-1: Modelos mixtos de PR/PD en fase predictiva bases de 15, 30 y 60 min

Coeficiente	Modelos PR/PD Fase Predictiva					
	15		30		60	
	Valor	Test-t robusto	Valor	Test-t robusto	Valor	Test-t robusto
Niveles Servicio						
Tiempo de viaje - PR/PD	-0,0323	-4,23	-0,0313	-4,27	-0,0283	-4,45
Tiempo de acceso - PR	-0,0633	-5,14	-0,0511	-4,59	-0,0387	-4,02
Tpo. transbordo público - PR	-0,0956	-3,82	-0,0712	-3,24	-0,0506	-2,7
Tpo. transbordo público-privado - PR	-0,629	-3,73	-0,542	-2,85	-0,358	-2,2
Costo - PR/PD	-0,000378	-3,29	-0,000371	-3,27	-0,000353	-3,2
Comfort - Ir sentado - PD	0,326	2,23	0,316	2,24	0,291	2,25
Incertidumbre en tiempos de viaje - PD	-1,04	-1,3	-1,01	-1,29	-0,931	-1,27
Restricciones de programación						
SDE - PD	-0,0202	-3,81	-0,0197	-3,84	-0,0184	-3,93
SDL - PD	-0,0232	-3,49	-0,0227	-3,51	-0,0215	-3,55
Socioeconómicas						
AUTLIC - PR	1,03	2,46	1	2,48	0,986	2,51
Interacciones						
Costo/tasa salarial (Costw) - PR/PD	-0,834	-3,58	-0,819	-3,58	-0,793	-3,57
Factor de escala PR/PD						
σ PR / σ PD *	1,04	0,15	1,06	0,22	1,13	0,45
ASC_Hora - PR						
5:00 - 7:00	2,07	4,20	2,10	4,27	2,16	4,40
7:00 - 7:15	4,43	9,22				
7:15 - 7:30	4,56	9,55				
7:30 - 7:45	5,19	11,19				
7:45 - 8:00	5,28	11,42				
8:00 - 8:15	5,59	12,21				
8:15 - 8:30	5,08	10,92				
8:30 - 8:45	4,60	9,69				
8:45 - 9:00	3,81	7,53				
9:00 - 11:00	1,30	2,45	1,31	2,46	1,31	2,45
11:00 - 14:00	0		0		0	
7:00 - 7:30			4,54	9,81		
7:30 - 8:00			5,28	11,61		
7:00 - 7:30			5,42	11,93		
8:30 - 9:00			4,33	9,27		
7:00 - 8:00					5,00	11,06
8:00 - 9:00					5,02	11,11
Log-verosimilitud	-2835,44		-2557,81		-2314,45	
rho-cuadrado ajustado	0,21		0,244		0,282	
# parámetros	33		29		27	
# observaciones	1743		1743		1743	

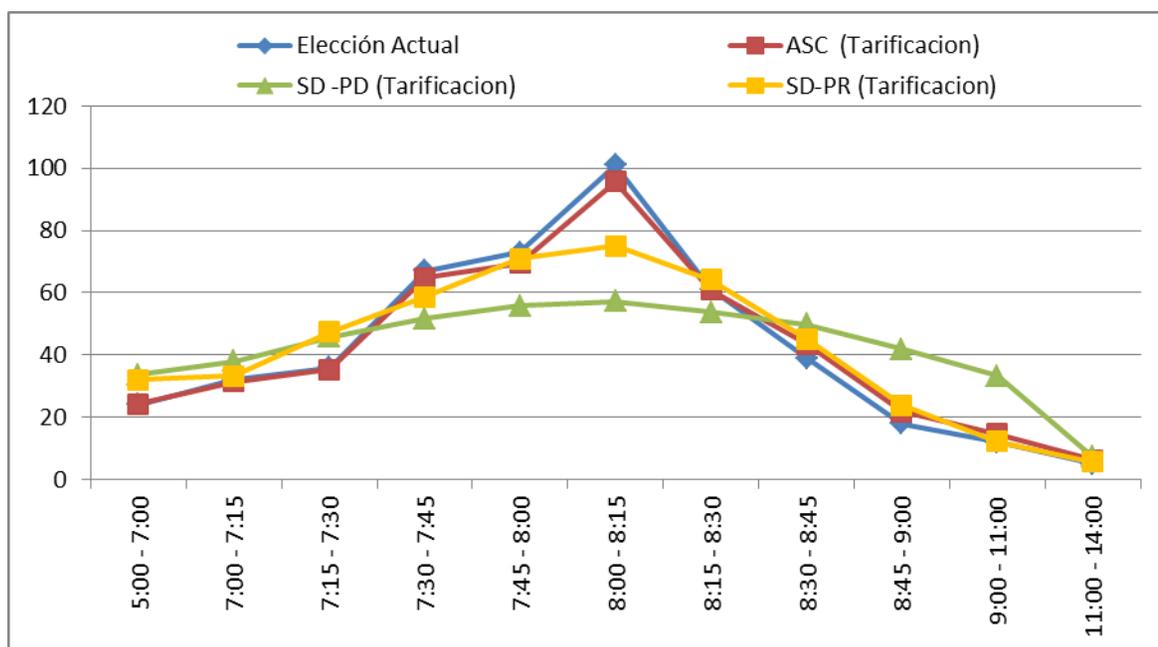


Figura 7-4: Distribución horaria de demanda escenario tarificación vial modelos 15 min

Fuente: Elaboración propia

De este gráfico es posible rescatar diversos aspectos: en primer lugar, el modelo que usa los parámetros de *SDE* y *SDL* provenientes de las funciones de utilidad de PD (*SD-PD*), presenta una curva más plana, en la cual el tráfico se redistribuye más equitativamente a lo largo del tiempo, y justamente uno de los objetivos de una política de tarificación vial es redistribuir temporalmente el tráfico. Asimismo, en comparación con las otras curvas, en la de *SD-PD* más personas aparecerían respondiendo a una política de cobro por congestión al viajar en los períodos posteriores de la mañana. Esto era esperable, pues en las funciones de utilidad de PD la penalización por llegar un minuto tarde era similar a la de llegar un minuto más temprano al trabajo.

Segundo, las predicciones realizadas por el modelo sin las variables de restricción de programación de actividades (ASC) no difieren mucho de la situación actual. Esto quiere decir que el modelo sin estas variables no es muy sensible a cambios en la hora de inicio del viaje ante cambios en los niveles de servicio. Esto era esperable, pues en este modelo no hay suficientes variables que expliquen la elección horaria.

Por último, la curva del modelo *SD-PR* presenta una curva más plana que la actual y que la curva ASC, pero con un período punta más pronunciado que su contraparte *SD-*

PD. Esto también era esperable, dado que los parámetros de *SDE* y *SDL* en PR eran más negativos que los de PD; por ende, en PR habría menos gente dispuesta a realizar grandes modificaciones en la hora de inicio del viaje que en PD.

También es de interés investigar los cambios en la partición modal al implementar estas tarifas. La Figura 7-5 muestra la diferencia porcentual en la partición modal al aplicar la política de tarificación (usando intervalos de 15 min). Como se puede apreciar, la partición modal del modo auto-chofer (ACH) disminuye consistentemente entre un 40 y 50%, según modelo. Los modos que presentan un mayor aumento porcentual son el auto-acompañante, taxi y bus; Metro y caminata son los modos que menos aumentan, lo que se debe a que hay menos personas que no los escogen cuando están disponibles (i.e. la razón entre su elección y disponibilidad es bastante alta), y por lo tanto, menos gente tiene la posibilidad de cambiarse a ellos. Además, los cambios en la partición modal son menores en el modelo *SD*-PD en comparación con las otras dos curvas, lo que significa que en el modelo *SD*-PD más personas están dispuestas a cambiar la hora de inicio del viaje que el modo de transporte.

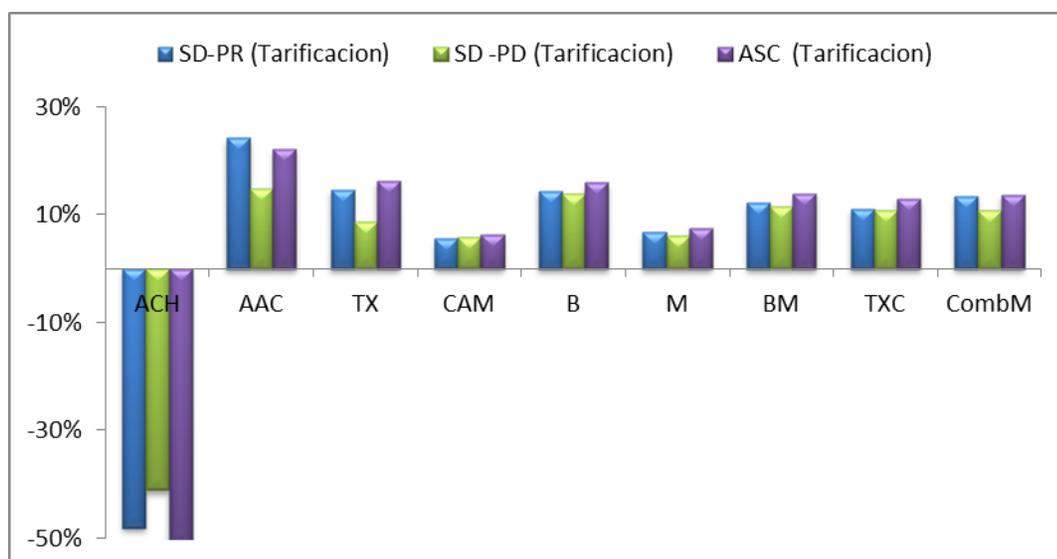


Figura 7-5: Cambio porcentual partición modal escenario tarificación vial modelos 15 min

Fuente: Elaboración propia

Los resultados recién mencionados podrían ser de bastante utilidad para las autoridades de Santiago que estén evaluando la posibilidad de tarificar en el barrio

caricaturizado como “*Sanhattan*” o en el centro de la ciudad. Para que la tarificación vial tenga el efecto deseado (disminuir el tráfico vehicular en ciertos períodos horarios), se sugiere implementar una serie de medidas complementarias.

En primer lugar, los centros de empleos deberían otorgar mayor flexibilidad a sus empleados para llegar tarde o temprano al trabajo. Como lo demuestra la curva *SD-PD* de la Tabla 7-4, bastantes personas estarían dispuestas a viajar en horarios menos congestionados de contar con mayor flexibilidad para modificar la hora de inicio del viaje. De no implementar las empresas un sistema de horarios de entrada y salida del trabajo más flexible, es posible que muchos automovilistas no cambien su hora de inicio del viaje, y el tráfico vehicular no disminuya como se esperaría. Además, las personas que se verían más perjudicadas con la tarificación vial serían los automovilistas con baja flexibilidad horaria, ya que se les obligaría a pagar la tarifa a pesar de que muchos de ellos estarían dispuestos a cambiar la hora en que viajan con tal de evadir el cobro.

En segundo lugar, tal como lo demuestra la Figura 7-5, es posible que existan cambios substanciales en la partición modal. Para que los modos de transporte sean capaces de asimilar la nueva demanda procedente de los automovilistas que se cambiarían, las autoridades deberían invertir en infraestructura para transporte público y otros modos (por ejemplo, vías exclusivas, ciclovías, etc.), aparte de mejoras operacionales (por ejemplo, aumentar la frecuencia de recorridos que lleguen al centro o a “*Sanhattan*”). De no invertir y ocurrir los cambios en la partición modal pronosticados, muy probablemente la calidad del servicio caería, y los más perjudicados serían los usuarios de transporte público.

Asimismo, de no aumentar la flexibilidad para llegar temprano o tarde al trabajo, es posible que ocurran mayores cambios en la partición modal (automovilistas con poca flexibilidad y alta aversión a pagar la tarifa preferirían cambiar de modo), por ende, se requerirían mayores inversiones en transporte público. Esto último es confirmado por el hecho que, en la Figura 7-5, el modelo *SD-PD* - que considera un sistema flexible de horarios de entrada y salida al trabajo - predice menos cambios en la partición modal que el modelo *SD-PR*, que no lo considera.

Las figuras 7-6 y 7-7 presentan la misma información que la Figura 7-4, pero para los modelos de intervalos de tiempo de 30 y 60 min, respectivamente. En el caso de 60 min, el modelo *SD-PR* corresponde al modelo 2_HL de la Tabla 6-2, mientras que el modelo *SD-PD* corresponde al modelo 2_MNL.

Al comparar las figuras 7-4, 7-6 y 7-7, se concluye que al agregar los intervalos de tiempo, menor es la diferencia entre las predicciones del modelo que utiliza los parámetros *SDE* y *SDL* de PR y el modelo ASC. En otras palabras, como resultado de agregar los períodos de tiempo, y por ende de utilizar predicciones temporales más agregadas, el modelo *SD-PR* se vuelve más insensible a cambios en la elección de hora de inicio del viaje. Asimismo, al utilizar intervalos de tiempo más amplios, en ambos modelos (*SD-PR* y ASC) existen menos cambios en la hora de inicio del viaje como producto de la implementación de la política de tarificación. Esto, en cierta medida, era esperable ya que al agregar el tiempo las alternativas de hora se encuentran más alejadas entre sí, y por ende resulta menos atractivo cambiar la hora de inicio del viaje.

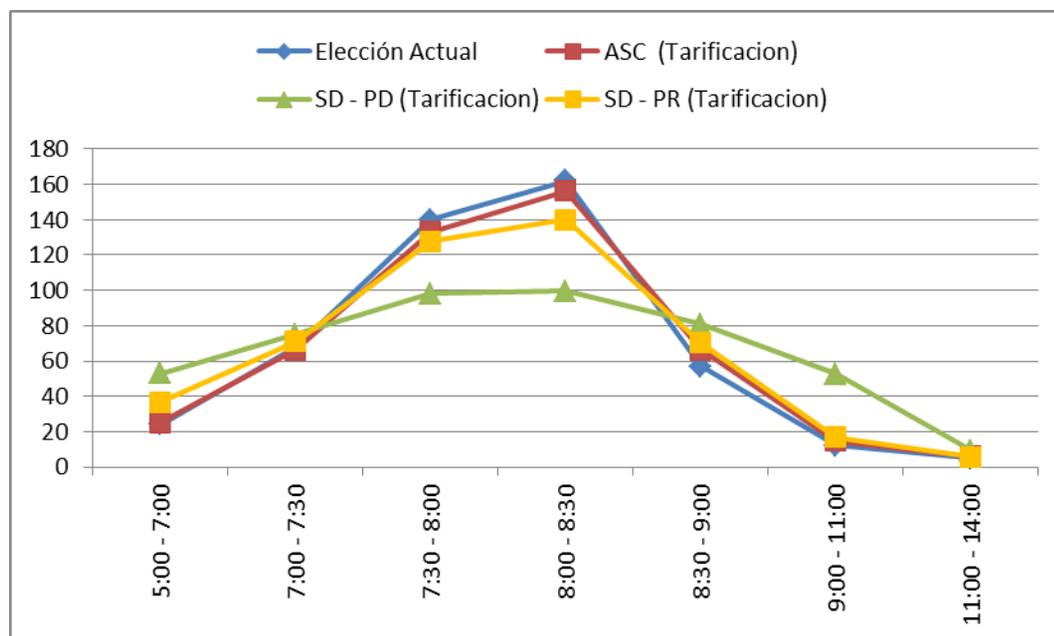


Figura 7-6: Distribución horaria de demanda escenario tarificación vial modelos 30 min

Fuente: Elaboración propia

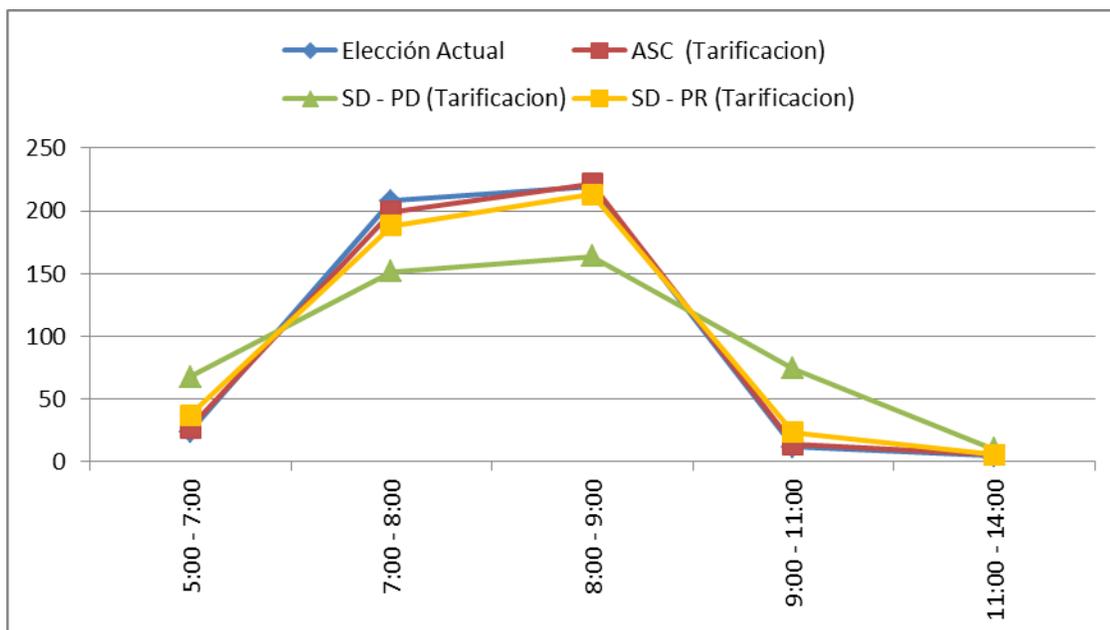


Figura 7-7: Distribución horaria de demanda escenario tarificación vial modelos 60 min

Fuente: Elaboración propia

Por su parte, el modelo *SD-PD* sigue mostrando la misma tendencia (una curva más plana) a raíz de la mayor flexibilidad horaria incorporada en los parámetros *SDE* y *SDL*. Además, la pérdida de información producto de agregar el tiempo es bastante menor que en los modelos *SD-PR* y *ASC*. Es por ello que si se va a usar los modelos *SD-PR* y *ASC* para predecir, se recomienda utilizar los que tengan mayor resolución temporal (intervalos de 15 min). En cambio si se va a usar el modelo *SD-PD*, se podría preferir utilizar los modelos de 30 o 60 min ya que son más parsimoniosos que el de 15 min y la pérdida de información es menor.

De esta sub-sección se puede concluir que la mayor sensibilidad para cambiar horarios se obtiene al usar los parámetros de *SDE* y *SDL* provenientes de los datos de *PD*. Esto posiblemente se debe a que el contexto del experimento de *PD* consideraba la aplicación de un esquema de tarificación vial y de un sistema de horarios flexible. Como los resultados anteriores lo demuestran, se debiera usar los parámetros *SDE* y *SDL* de *PD* para predecir un escenario de tarificación vial, pues en cierta medida es esperable que al momento de aplicarlo la flexibilidad horaria para llegar al trabajo aumente. Si se espera que la flexibilidad horaria no varíe en forma considerable, quizá sería adecuado

usar los parámetros de PR. Sólo en último lugar se debiera recurrir al modelo sin variables de restricción de programación de actividades, ya que este es el menos sensible a cambios en la hora de inicio del viaje.

No obstante lo anterior, los resultados discutidos en esta sub-sección deben ser validados utilizando otro banco de datos y lo más sencillo sería utilizar un banco de datos simulado.

8 CONCLUSIONES

8.1 Principales conclusiones

Las conclusiones de este estudio se pueden dividir en dos tipos: las primeras están relacionadas con aspectos metodológicos de la modelación de demanda de transporte, mientras que las del segundo tipo son sugerencias y recomendaciones de política pública en materia de transporte.

8.1.1 Conclusiones metodológicas

En primer lugar, se comprueban las dos hipótesis planteadas. Un modelo conjunto de elección de modo y hora, que mezcla datos de preferencias reveladas (PR) y declaradas (PD), es efectivamente superior a modelos estimados con cada tipo de dato en forma separada. Asimismo, este modelo mixto de PR/PD se puede utilizar para predecir las preferencias temporales de los usuarios del sistema de transporte. Si se cuenta con información acerca de las horas preferidas de llegada al trabajo (*PAT*) y de la hora oficial de entrada al trabajo (*HOE*), es posible usar los modelos con variables de restricción de la programación de actividades para predecir. Dependiendo del tipo de política a evaluar, se deben usar los parámetros de demora horaria de PD multiplicados por el factor de escala (tarificación vial o sistema flexible de horarios de entrada al trabajo) o los de PR en caso de evaluar cualquier otro tipo de política (como por ejemplo, la construcción de infraestructura vial). En caso de no contar con la información de *PAT* y *HOE*, se debiera usar los modelos que intercambian las variables de restricción de programación de actividades por constantes específicas para los periodos horarios. Sin embargo, como demuestran las predicciones, esto conlleva una pérdida en la capacidad predictiva del modelo, ya que sólo permite capturar las preferencias temporales agregadas y no las individuales.

Respecto a la forma de incluir la flexibilidad laboral dentro del modelo, se concluye que la mejor manera es a través de variables ficticias (flexibilidad baja, media y alta) interactuando con los parámetros asociados a las demoras horarias (*SDE* y *SDL*); esto permite capturar la desutilidad por llegar un minuto tarde o temprano al trabajo. Una forma alternativa de incorporarla, consistente en restringir el conjunto de

alternativas de hora disponibles, da como resultado modelos con menor bondad de ajuste y entrega sólo un valor para las razones de compromiso planteadas por el *Scheduling Model* (SM). En cambio, al usar la flexibilidad como variable explicativa se obtienen tres conjuntos de razones entre los parámetros de *SDE* (*SDL*) y el tiempo de viaje, para personas con flexibilidad horaria baja, media y alta respectivamente. Como era de esperar, en este caso se comprueba que las personas con flexibilidad alta son las que menos valoran llegar un minuto tarde o temprano al trabajo, seguidos por aquellos con flexibilidad media, y en último lugar las personas con flexibilidad baja.

La mejor especificación del modelo considera las siguientes variables de nivel de servicio: tiempo de viaje (común a PR y PD), tiempo de acceso, tiempo de transbordo (desagregado según si es un transbordo entre modos de transporte público o entre transporte privado y metro), costo y costo dividido por la tasa salarial (*Costw*). Estas últimas dos variables se ingresan de forma simultánea para estabilizar los valores monetarios del tiempo, ya que si sólo se utiliza *Costw* los valores monetarios aumentan considerablemente, y ambas variables son comunes a PR y PD. Asimismo, los parámetros de *SDE* y *SDL*, difieren bastante entre PR y PD; esto es atribuible, por un lado, a la diferencia temporal entre la elección de PR (largo plazo) y PD (corto plazo), y por otro a que al responder el experimento de PD, las persona posiblemente consideraban un mayor nivel de flexibilidad al actual, pues el contexto del experimento trataba de un sistema flexible de horarios de entrada y salida del trabajo. En general, en PR se obtiene que llegar un minuto tarde al trabajo es valorado más que llegar un minuto temprano, y ambos (minuto tarde y temprano) son más valorados que un minuto de tiempo de viaje. En PD, en cambio, llegar un minuto tarde es valorado en forma similar a llegar un minuto temprano, y el tiempo de viaje se valora más que ambos.

Respecto de la longitud de las alternativas de hora de PR (15, 30 y 60 min), se concluye que la bondad de ajuste de los modelos aumenta progresivamente al utilizar intervalos de tiempo más agregados y, además, disminuye la valoración del tiempo de transbordo y de acceso. Esto puede ser producto de dos fenómenos: (i) al agregar las alternativas de tiempo existe menor correlación entre las alternativas de hora, y (ii) al usar intervalos de tiempo más amplios aumenta el acierto al predecir las elecciones de

hora observadas (aunque esto último no es tan evidente al pasar de intervalos de 30 a 60 min).

Se trató de medir la correlación entre las alternativas de hora en PR a través de modelos HL, con una estructura jerárquica modo – hora; sin embargo, contrariamente a lo esperado, esto sólo funcionó en el modelo con intervalos de tiempo de 60 min. Una explicación es que quizás la estructura de correlación sea más clara (o simple) al utilizar intervalos de tiempo agregados y, a medida que estos se desagregan, la estructura se complejiza. Es por ello que también se trató de aplicar una estructura altamente flexible para comparar el impacto de la correlación en los modelos de 15, 30 y 60 min; sin embargo, los modelos no convergieron (se alcanzó el número máximo de iteraciones del algoritmo de optimización).

A pesar de que los modelos con mayor nivel de agregación temporal tienen mejor bondad de ajuste, se reconoce que existe una pérdida de información al hacer esta agregación, en especial al utilizar el modelo para predecir políticas como la tarificación vial. De hecho, de los resultados obtenidos al predecir un esquema de tarificación vial se puede concluir que para predecir una política de este tipo, es conveniente emplear modelos con intervalos de tiempo lo más desagregados posible si se utilizan los parámetros de *SDE* y *SDL* provenientes de PR (o si se utilizan los modelos que prescindan de ellos). En cambio, si se utilizan los parámetros de *SDE* y *SDL* provenientes de las funciones de utilidad de PD, la pérdida de información es menor y, por simpleza, se recomienda usar modelos con mayor agregación temporal.

Se puede concluir que el efecto de las variables *innovadoras* (temperatura y densidad del metro) no es tan significativo como el de otras variables explicativas. Asimismo, el factor de logaritmo natural especificado para corregir por alternativas de hora de diferente longitud se agregó solamente en los modelos que utilizan constantes específicas (ASC) de hora en lugar de las variables *SDE* y *SDL*, ya que en los modelos con las variables de *SDE* y *SDL* el efecto que generó su introducción fue aumentar el valor de los coeficientes y disminuir la bondad de ajuste del modelo. En cambio, al agregar el factor de corrección en los modelos con ASC de hora la bondad de ajuste

permaneció igual y también aumentó el valor de las ASC de hora de los períodos de mayor resolución temporal.

Finalmente, se comprobó la presencia de efecto panel en el experimento de PD a través de la formulación y estimación de modelos logit mixto de componentes de error. Estos modelos fueron claramente superiores a los modelos del tipo MNL, ya que la varianza del componente de error fue estadísticamente significativa; sin embargo, la mayor complejidad técnica y computacional de aplicar un modelo logit mixto en la práctica, hizo que se prefiriesen modelos del tipo MNL y HL para ser usados en fase predictiva.

8.1.2 Conclusiones de política pública

A partir del ejercicio de predecir el impacto de un esquema de tarificación vial es posible derivar una serie de conclusiones en términos de política pública. Si las autoridades de Santiago están evaluando implementar un esquema de cobro vial para mitigar el tráfico vehicular que hay en ciertas zonas y períodos del día en la capital, se sugiere aplicar una serie de medidas complementarias. Esto con el fin de que los usuarios del sistema de transporte capitalino tengan la mayor cantidad de alternativas de transporte disponibles (ya sea modo de transporte u horario de salida) para responder a la tarificación vial, y que a su vez estas alternativas sean de calidad acorde a un sistema moderno y eficiente.

La primera recomendación a las autoridades es que, en lo posible, procuren incentivar a las empresas a que flexibilicen sus políticas de horario de entrada y salida del trabajo. De no implementarse un esquema de este tipo, es posible que muchos automovilistas no cambien su hora de inicio del viaje y, por ende, que el tráfico vehicular no disminuya tanto como se esperaría; además se perjudicaría a aquellos automovilistas con baja flexibilidad horaria (que en muchos casos pueden coincidir con personas de menores ingresos), ya que se les obligaría a pagar la tarifa aun cuando muchos estén dispuestos a viajar en otras horas para evadir el cobro y/o la congestión. Un tipo de incentivo podría ser tributario, en que la empresa pueda descontar impuestos si es que cuenta con un esquema de horarios flexible.

Tal como fue predicho, es altamente probable que con un esquema de tarificación vial muchos automovilistas dejen de realizar su viaje en auto para hacerlo en otros modos de transporte. Por ello, la segunda sugerencia sería que las autoridades inviertan (gran parte de) los ingresos recaudados por este cobro en mejorar el transporte público, tanto a nivel de infraestructura como operacional (mayor frecuencia). Usar los fondos recaudados en mejorar el transporte público tiene la ventaja política de ser la medida de compensación más popular entre los usuarios del sistema de transporte en caso de implementarse la tarificación vial (ver Salata *et al.*, 2012). Asimismo, de no invertir en mejorar el transporte público masivo, es posible que el sistema no sea capaz de asimilar correctamente la nueva demanda generada y la calidad del servicio ofrecido decaería. La implementación de Transantiago es un buen ejemplo de la necesidad de anticipar correctamente los cambios, ya que la demanda del Metro se duplicó de un mes para otro, disminuyendo el nivel de confort que ofrecía Metro.

Por último, con los fondos recaudados también se debiera favorecer a aquellos medios de transporte que congestionan menos la red vial (y contaminan menos), como por ejemplo la bicicleta, construyendo infraestructura adecuada como buenas ciclovías. Un buen ejemplo podría ser la ciclovía que se piensa construir en la rivera del río Mapocho, la cual pasaría por las áreas consideradas a tarificar.

8.2 Recomendaciones para futuras investigaciones

Una de las extensiones naturales de esta investigación sería estimar modelos de elección de hora considerando, como unidad de análisis, *tours* basados en el viaje al trabajo en lugar de sólo viajes al trabajo. El principal desafío de esto es acotar lo más posible el conjunto de alternativas.

Otra recomendación sería estimar modelos con estructuras más complejas para las funciones de utilidad de PR, a fin de capturar adecuadamente la magnitud de la correlación entre alternativas de hora. Estos pueden ser del tipo logit jerárquico de nidos cruzados, OGEV u otras que no requieran técnicas de simulación y, además, tienen la ventaja de ser más simples de utilizar en la práctica. Otra posibilidad es intentar que

converjan los modelos logit mixto de componentes de error propuestos para medir la correlación entre alternativas de hora en esta tesis.

Otra línea de estudio podría ser investigar la posibilidad de estimar un modelo de elección de hora secuencial, esto es, primero escoger un intervalo de tiempo amplio (por ejemplo, 1 hora) y luego, dentro de este intervalo de una hora, escoger sub-intervalos de 15 o 30 min.

Una extensión más sencilla, consistiría en reestimar los modelos utilizados para la fase predictiva, considerando interacciones entre las constantes específicas de hora y variables socioeconómicas. Esto con el objeto de aumentar la capacidad explicativa y predictiva de estos modelos. Asimismo, realizar las mismas predicciones de la subsección 7.3, pero a partir de un banco de datos simulado, a fin de validar los resultados obtenidos.

Por último, está la posibilidad de estimar modelos de elección de modo-hora a partir de una gran encuesta de movilidad como la EOD. La ventaja de esto sería contar con más observaciones, pero posiblemente un nivel de precisión considerablemente menor. Los niveles de servicio por período del día se pueden determinar utilizando metodologías propuestas recientemente (por ejemplo, Ben-Akiva y Abou Zeid, 2012). En caso de no contar con información acerca de la hora preferida de llegada al trabajo (*PAT*), ésta se puede modelar como variable latente, o determinar mediante el enfoque de Kristofferson y Engelson (2008), que la derivan a partir de un modelo de elección de hora estimado previamente (como los de esta tesis). La otra posibilidad es sencillamente omitirla y usar en su lugar constantes específicas.

BIBLIOGRAFIA

- Abou-Zeid, M., Rossi, T. y Gardner, B. (2006). Modelling time-of-day choice in the context of tour- and activity based models. *Transportation Research Record*, **1981**, 42-49.
- Arellana, J. (2012). *Modelación de Elección de la Hora de Inicio de los Viajes*. Tesis de Doctorado en Ciencias de la Ingeniería. Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile, Santiago.
- Arellana, J., Daly, A.J., Hess, S., Ortúzar, J. de D. y 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., Daly, A.J., Hess, S., Ortúzar, J. de D. y Rizzi, L.I. (2012c). The flexibility of departure time for work based trips: an integrated choice and latent variable approach. *IATBR Conference 2012*. Toronto, Canadá.
- Arellana, J., Ortúzar, J. de D., Rizzi, L.I., y Zuñiga, F. (2012b). 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. y Lindsey, R. (1990). Departure time and route choice for the morning commute. *Transportation Research*, **24B**, 209-228.
- Arnott, R., de Palma, A. y Lindsey, R. (1994). Welfare effects of congestion tolls and heterogeneous commuters. *Journal of Transport Economics and Policy*, **28**, 139-161.
- Ben-Akiva, M. y Abou-Zeid, M. (2012). Methodological issues in modelling time-of-travel preferences. *Transportmetrica* (DOI:10.1080/18128602.2012.686532).
- Ben-Akiva, M. y Morikawa, T. (1990). Estimation of travel demand models from multiple data sources. *Proceedings 11th International Symposium on Transportation and Traffic Theory*. Yokohama, Japón.
- 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 modelling: formulation and application to travel mode and departure time choice. *Transportation Research*, **32B**, 455-466.
- Bhat, C.R. y Castelar, S. (2002). A unified mixed logit framework for modelling revealed and stated preferences: formulation and application to congestion pricing analysis in the San Francisco Bay area. *Transportation Research*, **36B**, 593-616.
- Bianchi, R., Jara-Díaz, S.R. y 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. *3rd Swiss Transport Research Conference*, Ascona, Suiza.

- Börjesson, M. (2008). Joint RP/SP data in a mixed logit analysis of trip timing decisions. *Transportation Research*, **44E**, 1025-1038.
- Brey, R. y Walker, J.L. (2011) Latent temporal preferences: an application to airline travel. *Transportation Research*, **45A**, 880–895.
- Cambridge Systematics Inc. (1999). *Time-of-Day Modelling Procedures: State-of-the-Practice, State-of-the-Art*. Publication DOT-T-99-01, FHWA, U.S. Department of Transportation, Washington, D.C.
- Cantillo, V.M. y Ortúzar, J. de D. (2012). Restricción vehicular según número de patente: réquiem para una política errónea. *Ingeniería de Sistemas*, **26**, 7-22.
- Carrasco, J.A. y Ortúzar, J. de D. (2002). Review and assessment of the nested logit model. *Transport Reviews*, **22**, 197-218.
- Carrion, C. y Levinson, D. (2012). Value of travel time reliability: a review of current evidence. *Transportation Research*, **46A**, 720-741.
- Cherchi, E. y Ortúzar, J.de D. (2002). Mixed RP/SP models incorporating interaction effects: modelling new suburban train services in Cagliari. *Transportation*, **29**, 371-395.
- De Grange, L. y Troncoso, R. (2010) Impacto de la restricción vehicular sobre los flujos en transporte urbano de Santiago. *Serie Informe Económico*, Libertad y Desarrollo, Santiago.
- De Jong, G., Daly, A.J., Pieters, M., Vellay, C., Bradley, M. y 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*. Departamento de Ingeniería de Transporte, Pontificia Universidad Católica de Chile, Santiago.
- Ettema, D., Ashiru, O. y Polak, J. (2004). Modelling 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. y Ashiru, O. (2007). Modelling the joint choice of activity timing and duration. *Transportation Research*, **41A**, 827-841.
- Ettema, D.F. y Timmermans, H.J.P. (2003). Modelling departure time choice in the context of activity scheduling behaviour. *Transportation Research Record*, **1831**, 39–46.
- Fresard, F. (1998). Efecto real de la restricción vehicular en Santiago de Chile. *Actas X Congreso Panamericano de Ingeniería de Tránsito y Transporte*, Santander, España.
- Grisolía, J.M. y Ortúzar, J. de D. (2010). Forecasting vs. observed outturn: studying choice in faster inter-island connections. *Transportation Research*, **44A**, 159-168.
- Habib, K.M.N., Day, N. y Miller E.J. (2009). An investigation of commuting trip timing and mode choice in the greater Toronto area: application of a joint discrete-continuous model. *Transportation Research*, **43A**, 639-653.

- Habib, K.M.N. (2011). A joint discrete-continuous model considering budget constraint for the continuous part: application in joint mode and departure time choice modelling. *Transportmetrica* (DOI:10.1080/18128602.2010.551525).
- Habib, K.M.N (2012). Modelling commuting mode choice jointly with work start time and duration. *Transportation Research*, **46A**, 33–47.
- Hendrickson, C., y Plank, E. (1984). The flexibility of departure times for work trips. *Transportation Research*, **18A**, 25-36.
- Hensher, D.A. y 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. y Hyman, G. (2007). 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.W. y Bierlaire, M. (2005). Functional approximations to alternative specific constants in time period choice-modelling. *Proceedings 16th International Symposium on Transportation and Traffic Theory*, University of Maryland, College Park.
- Jara-Díaz, S.R., Munizaga, M. y Guerra, R. (2012). Sensitivity of the value of time to the specification of income in discrete choice models. *International Journal of Transport Economics*, **39**, 239-254.
- Khattak, A.J., Schofer, J.L. y Koppelman, F.S. (1995). Effect of traffic information on commuters' propensity to change route and departure time. *Journal of Advanced Transportation*, **29**, 193–212.
- Kristoffersson, I. y Engelson, L. (2008). Estimating preferred departure times of road users in a real-life network. *Proceedings European Transport Conference*. Noordwijkerhout, Holanda.
- Lemp, J. y Kockelman, K. (2010). Empirical investigation of continuous logit for departure time choice with Bayesian methods. *Transportation Research Record*, **2165**, 59-68.
- Li, Z., Hensher, D.A. y 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. y Swait, J.D. (2000). *Stated Choice Methods: Analysis and Application*. Cambridge University Press, Cambridge.
- Mannering, F.L. (1989). Poisson analysis of commuter flexibility in changing routes and departure times. *Transportation Research*, **23B**, 53–60.
- McFadden, D. y Train, K. (2000) Mixed MNL models for discrete response. *Journal of Applied Econometrics*, **15**, 447-470.

- Munizaga, M.A., Jara-Díaz, S.R., Greeven, P. y Bhat, C. (2008). Econometric calibration of the joint time assignment-mode choice model. *Transportation Science*, **42**, 208-219.
- Navarrete, F. y Ortúzar, J. de D. (2012) Subjective valuation of the transit transfer experience: the case of Santiago de Chile. *Transport Policy* (en imprenta).
- Noland, R.B. y 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 y Willumsen L.G. (2011) *Modelling Transport*: Cuarta Edición, John Wiley and Sons, Chichester.
- Popuri, Y., Ben-Akiva, M. y Prousaloglou, K. (2008). Time-of-day modelling in a tour-based context: the Tel Aviv experience. *Transportation Research Record*, **2076**, 88-96.
- Rand Europe (2004). *PRISM West Midlands. Time of Day Choice Models*. Cambridge.
- Roorda, M.J, Miller, E.J. y Habib, K.M.N. (2008). Validation of TASHA: a 24-hour travel-activity micro simulation model. *Transportation Research*, **42A**, 360-375.
- Sabir, M. (2011). *Weather and Travel Behaviour*. Doctoral Thesis. Department of Spatial Economics, VU University, Amsterdam.
- Sabir, M., Koetse, M.J. y Rietveld, P. (2009). The Impact of weather conditions on mode choice: empirical evidence for The Netherlands. *17th Annual Conference European Association of Environmental and Resource Economists*. Amsterdam, Holanda.
- Salata, A., Ortúzar, J. de D. y Rizzi, L.I. (2012) Assessing the potential acceptability of road pricing in Santiago. *Transportation Research A* (submitted).
- Sillano, M. y Ortúzar, J. de D. (2005). Willingness-to-pay estimation with mixed logit models: some new evidence. *Environment and Planning*, **37A**, 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.
- Small, K.A., Noland, R., Chu, X. y Lewis, D. (1999). Valuation of travel-time savings and predictability in congested conditions for highway user-cost estimation. *NCHRP Report 431*. Transportation Research Board, US National Research Council, Washington, D.C.
- Steed, J.L. y Bhat, C.R. (2000). On modelling departure time choice for home-based social/recreational and shopping trips. *Transportation Research Record*, **1706**, 152-159.
- Train, K.E. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge.
- Tringides, C., Ye, X. y 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. y Verhoef, E. (2011). Discrete choice analysis for trip timing decisions of morning commuters: estimations from joint SP/RP-GPS data. *International Choice Modelling Conference 2011*. Leeds, Inglaterra.
- Van Vuren, T., Carmichael, S., Polak, J., Hyman, G. y Cross, S. (1999). Modelling peak spreading in continuous time. *European Transport Conference*, Cambridge, Inglaterra.
- Vickrey, W.S. (1969). Congestion theory and transport investment. *The American Economic Review*, **59**, 251-260.
- Vovsha, P. and Bradley, M. (2004). Hybrid discrete choice departure-time and duration model for scheduling travel tours. *Transportation Research Record*, **1894**, 46-56.
- Wang, J.J. (1996) Timing utility of daily activities and its impact on travel. *Transportation Research Part A*, **30**, 189–206.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, **48**, 817-838.
- Yañez, M.F., Mansilla, P. y Ortúzar, J. de D. (2010). The Santiago Panel: measuring the effects of implementing Transantiago. *Transportation*, **37**, 125-149.
- Yañez, M.F. y Ortúzar, J. de D. (2009) Modelling choice in a changing environment: assessing the shock effects of a new transport system. En S. Hess y A.J. Daly (eds.), *Choice Modelling: The States of the Art and of Practice*. 445-460, Emerald, Bingley.

ANEXOS

ANEXO A: MODELOS DE ELECCION DISCRETA

- **Modelo Logit Multinomial (MNL)**

Este es el modelo de elección discreta más simple y utilizado en la literatura; se deriva al asumir que el término de error aleatorio distribuye IID Gumbel. La probabilidad de elección de una alternativa j para un individuo q está dada por (Ortúzar y Willumsen, 2011):

$$P_{jq} = \frac{\exp(\lambda \cdot V_{jq})}{\sum_{i \in A(q)} \exp(\lambda \cdot V_{iq})} \quad (\text{A.1})$$

donde λ es un parámetro de escala del modelo que es inversamente proporcional a la varianza del error aleatorio y V_{jq} representa la utilidad representativa o determinística. En general V_{jq} se trata como una función lineal en los coeficientes y parámetros:

$$V_{jq} = \sum_k \theta_{jk} \cdot X_{jkq} \quad (\text{A.2})$$

en que X_{jkq} puede ser un atributo específico de la alternativa j (por ejemplo, nivel de servicio) o una característica socioeconómica del individuo q y θ_{jk} son los coeficientes a estimar.

Como es posible notar, estos coeficientes son constantes para todos los individuos y es común que también sean iguales o genéricos para todas las alternativas, es decir, sin subíndice j (θ_k). De esta forma θ_k representa la utilidad marginal del atributo k . Asimismo, dado que estos coeficientes se encuentran afectados por el parámetro de escala λ y que ambos no se pueden estimar por separado, se tiende a normalizar λ a 1.

Las ventajas del modelo MNL es que su estructura y estimación es parsimoniosa (su función de probabilidad es cerrada y por ende su estimación no requiere simulación), y si cada alternativa (a excepción de una, que se fija en cero para que el modelo sea estimable) tiene una constantes específica (ASC), el modelo reproduce las proporciones de mercado observadas en la muestra. Sin embargo, sus principales falencias son asumir que: (i) todas las alternativas son independientes entre sí; (ii) todas las alternativas tienen

igual varianza, lo que se conoce como homocedasticidad, y (iii) no admite variaciones en los gustos, ya que todos los individuos poseen el mismo parámetro. Esto último se puede corregir haciendo interactuar los atributos de cada alternativa con las características socioeconómicas de los individuos, lo que se conoce como variaciones sistemáticas en los gustos (Ortúzar y Willumsen, 2011, pag. 279).

- **Modelo Logit Jerarquico (HL)**

Este modelo, que es una versión generalizada del MNL, es capaz de acomodar cierto grado de correlación entre alternativas al utilizar una estructura jerárquica, en la que un conjunto de alternativas puede pertenecer a un nido. Su función de probabilidad también es cerrada y tiene la siguiente formula (Carrasco y Ortúzar, 2002):

$$P_j = \frac{\exp(\lambda_l \cdot V_{j/l})}{\sum_{i \in l} \exp(\lambda_l \cdot V_{i/l})} \cdot \frac{\exp\left(\beta \cdot \left[\frac{1}{\lambda_l} \cdot \text{Ln}\left\{\sum_{i \in l} \exp(\lambda_l \cdot V_{i/l})\right\}\right]\right)}{\sum_{m=1}^L \exp\left(\beta \cdot \left[\frac{1}{\lambda_m} \cdot \text{Ln}\left\{\sum_{i \in l} \exp(\lambda_l \cdot V_{i/l})\right\}\right]\right)} \quad (\text{A.3})$$

donde l es el subíndice de los nidos y $V_{j/l}$ es la utilidad representativa de la alternativa j perteneciente al nido l . Como requisito estructural del modelo se debe cumplir que $\beta \leq \lambda_l, \forall l$, y en el caso límite en que $\beta = \lambda_l, \forall l$ el modelo colapsa al MNL ya que no existe correlación entre alternativas.

Tanto el modelo HL como el MNL se pueden estimar usando *software* especializado, como por ejemplo, ALOGIT o Biogeme (Bierlaire, 2003); este último se encuentra disponible gratuitamente en la *web*. La técnica de estimación más común para ambos modelos es la de máxima verosimilitud.

- **Modelo Logit Mixto (ML)**

Este es el modelo de elección discreta más general y flexible, pues puede aproximar a cualquier otro modelo basado en la teoría de la utilidad aleatoria (McFadden y Train, 2000). De forma resumida, el modelo ML generalmente se subdivide en dos categorías dependiendo de su especificación. Por un lado, está el ML de parámetros aleatorios (MLPA), que permite que los parámetros θ varíen en la población al estimar una distribución de probabilidades para cada uno. De esta forma, su función de

probabilidad está determinada por la integral de un logit tradicional (MNL o HL) sobre la función de densidad de los parámetros (Train, 2009):

$$P_{jq} = \int \frac{\exp(\theta' \cdot X_{jq})}{\sum_{i \in A(q)} \exp(\theta' \cdot X_{iq})} \cdot f(\theta) d\theta \quad (\text{A.4})$$

donde θ' es el vector de parámetros que varían en la población y $f(\theta)$ es la función de densidad de estos parámetros aleatorios. Una desventaja es que esta distribución se debe determinar *a priori*, y podrían producirse parámetros con signo contrario a la teoría para algún porcentaje de la población (Sillano y Ortúzar, 2005).

Por otro lado, en el modelo ML de componentes de error (ECL), a las utilidades se agregan componentes aleatorios con media cero y varianza de una distribución de probabilidades determinada (generalmente normal o uniforme). De esta forma es posible generar correlación entre las múltiples respuestas de individuos, como en el caso de datos de PD, o generar correlación cruzada entre alternativas.

En teoría ambas especificaciones del ML, parámetros aleatorios o componentes de error, son equivalentes, ya que un modelo ECL puede aproximar a un ML de parámetros aleatorios. Asimismo, mediante ambos es posible acomodar el *efecto panel* o múltiples respuestas de un individuo.

Dado que la función de la probabilidad de elección no tiene una forma cerrada, el modelo ML en general se estima utilizando técnicas de máxima verosimilitud simulada (Train, 2009). En otras palabras, para el caso del ML de parámetros aleatorios, la ecuación A.4 se reemplaza por la ecuación A.5:

$$\hat{P}_{jq} = \frac{1}{R} \sum_{r=1}^R \frac{\exp(\theta' \cdot X_{jq})}{\sum_{i \in A(q)} \exp(\theta' \cdot X_{iq})}. \quad (\text{A.5})$$

donde R es el número de instancias (*draws*) de la simulación. Con Biogeme es posible estimar tanto el ML de parámetros aleatorios como el de componentes de error.

- **Factor de Escala en Modelos Mixtos PR/PD**

A continuación se demuestra por qué es necesario un factor de escala para mezclar distintos tipos de datos (PR y PD). Definiendo las funciones de utilidad de PR y PD de la siguiente forma (con σ las varianzas de los términos de error aleatorios):

$$U_i^{PR} = c_i^{PR} + \theta X_i^{PR} + \alpha Y_i^{PR} + \epsilon_i, \quad \epsilon_i \sim (0, \sigma_\epsilon^2)$$

$$U_i^{PD} = c_i^{PD} + \theta X_i^{PD} + \beta Z_i^{PD} + \eta_i, \quad \eta_i \sim (0, \sigma_\eta^2)$$

Y las funciones de probabilidad de PR y PD para un modelo MNL son las siguientes (con λ el factor de escala del modelo MNL):

$$P_i^{PR} = \frac{e^{\left(\lambda^{PR}(c_i^{PR} + \theta X_i^{PR} + \alpha Y_i^{PR})\right)}}{\sum_{A_j \in A^{PR}} e^{\left(\lambda^{PR}(c_j^{PR} + \theta X_j^{PR} + \alpha Y_j^{PR})\right)}}, \quad \forall A_i \in A^{PR}, \quad \lambda^{PR} = \frac{\pi}{\sqrt{6}\sigma_\epsilon}$$

$$P_i^{PD} = \frac{e^{\left(\lambda^{PD}(c_i^{PD} + \theta X_i^{PD} + \beta Z_i^{PD})\right)}}{\sum_{A_j \in A^{PD}} e^{\left(\lambda^{PD}(c_j^{PD} + \theta X_j^{PD} + \beta Z_j^{PD})\right)}}, \quad \forall A_i \in A^{PD}, \quad \lambda^{PD} = \frac{\pi}{\sqrt{6}\sigma_\eta}$$

Dado que los verdaderos parámetros poblacionales (θ) son únicos e independientes del tipo de dato y que se encuentran deflactados por los factores de escala, se postula lo siguiente:

$$\theta^{PR} = \lambda^{PR} \cdot \theta$$

$$\theta^{PD} = \lambda^{PD} \cdot \theta$$

De esta forma, el factor de escala (u) permite igualar las varianzas de PR y PD para que ambos tipos de datos puedan ser mezclados:

$$u = \theta^{PD} / \theta^{PR} = \lambda^{PD} / \lambda^{PR} = \sigma^{PR} / \sigma^{PD}$$

- **Test de Razón de Verosimilitud para Modelos Mixtos PR/PD**

El test de razón de verosimilitud para modelos mixtos PR/PD busca comparar modelos mixtos PR/PD con modelos estimados usando ambos tipos de datos de forma separada. El estadígrafo del test es el siguiente:

$$LR = -2\{l(\theta^{PR/PD}) - l(\theta^{PR}) - l(\theta^{PD})\} \sim \chi_k^2$$

donde k son la cantidad de parámetros comunes a PR y PD, $l(\theta^{PR/PD})$ es las log-verosimilitud del modelo mixto y $l(\theta^{PR})$ y $l(\theta^{PD})$ son las log-verosimilitudes de los modelos estimados de forma separada.

Si los k parámetros comunes son iguales, entonces LR distribuye chi-cuadrado con k grados de libertad. De esta forma, si $LR < \chi_k^2$ entonces se acepta la hipótesis nula de que el modelo mixto es equivalente a los modelos estimados de forma separada, por lo que se prefiere al modelo mixto.

ANEXO B: RESULTADOS MODELOS BASE PR_15

Tabla B-1: Constantes específicas de modo (ASC_Modo) de modelos base PR_15_1

	Modelos Base PR_15_1									
	<i>1</i>		<i>2</i>		<i>3</i>		<i>3_LN</i>		<i>3_VarGustos</i>	
	Valor	Test-t robusto	Valor	Test-t robusto	Valor	Test-t robusto	Valor	Test-t robusto	Valor	Test-t robusto
<i>ASC_Modo</i>										
AAC	-2.9	-7.61	-1.62	-3.58	-1.54	-3.33	-1.57	-3.41	-1.54	-3.33
ACH	0	-	0	-	0		0		0	
B	-1.28	-3.62	-0.0521	-0.13	0.263	0.61	0.279	0.64	0.277	0.64
BM	-0.693	-1.85	0.747	1.68	0.634	1.4	0.678	1.5	0.654	1.44
CAM	0.407	0.439	1.83	3.51	1.52	2.93	1.6	3.11	1.53	2.95
COMBM	-0.673	-1.89	9.74E-02	0.28	2.17	3.11	2.26	3.14	2.17	3.08
M	0.667	2.06	2.24	5	1.88	4.38	1.91	4.47	1.9	4.41
TX	-1.51	-4.4	-0.925	-2.13	-0.72	-1.62	-0.811	-1.84	-0.711	-1.59
TXC	-0.62	0.454	0.596	1.17	0.527	0.94	0.625	1.1	0.551	0.98

Fuente: Elaboración propia

ANEXO C: RESULTADOS MODELOS BASE PD

Tabla C-1: Resultados estimación modelo MNL con variables de flexibilidad horaria base PD

Coeficiente	Modelos Base PD	
	Valor	Test-t robusto
5		
Nivel de Servicio		
Tiempo de viaje	-0,0293	-4,89
Incertidumbre % en tiempo de viaje	-1,43	-1,88
Costo	-0,000437	-2,01
Restricciones de Programación		
Demora horaria por llegar temprano (<i>SDE</i>)	-0,0216	-6,99
Demora horaria por llegar tarde (<i>SDL</i>)	-0,0265	-8,05
Llega tarde al trabajo (<i>D_TARDE</i>)	-0,313	-2,71
Interacciones		
Costo/tasa salarial (<i>Costw</i>)	-0,932	-1,65
<i>SDE</i> - Flexibilidad horaria alta	0,000272	0,07
<i>SDE</i> - Flexibilidad horaria media	0,00187	0,62
<i>SDL</i> - Flexibilidad horaria alta	0,00106	0,23
<i>SDL</i> - Flexibilidad horaria media	0,00352	0,91
ASC		
Salir antes		
Salir antes o después	-0,703	-7,61
Salir después		
Cambio modo	-1,44	-14,6
Log-verosimilitud	-1307,05	
rho-cuadrado ajustado	0,253	
# parámetros	13	
# individuos	255	
# observaciones	1275	
Valor subjetivo tiempo viaje (\$/min)	42	

Fuente: Elaboración propia

Tabla C-2: Resultados estimación mejor modelo Logit Mixto (ML)

Coeficiente	Modelos Base PD	
	Valor	Test-t robusto
<i>Nivel de Servicio</i>		
Tiempo de viaje	-0,0466	-4,02
Incertidumbre % en tiempo de viaje	-0,121	-0,11
Costo	-0,00064	-1,86
Comfort - Ir sentado	0,511	2,42
<i>Restricciones de Programación</i>		
Demora horaria por llegar temprano (<i>SDE</i>)	-0,0378	-6,67
Demora horaria por llegar tarde (<i>SDL</i>)	-0,0384	-7,11
Llega tarde al trabajo (<i>D_TARDE</i>)	-0,474	-2
<i>Interacciones</i>		
Costo/tasa salarial (<i>Costw</i>)	-0,826	-0,74
<i>ASC</i>		
Salir antes	-1,14	-3,53
Salir después	-1,39	-4,22
Cambio modo	-2,35	-9,05
<i>Componente de error de efecto panel</i>		
Sigma	1,92	12,73
Log-verosimilitud	-1077,58	
rho-cuadrado ajustado	0,384	
# parámetros	12	
# individuos	255	
# observaciones	1275	
# Draws	1000	
Valor subjetivo tiempo viaje (\$/min)	54	
Razones de compromiso		
<i>SDL/SDE</i>	1,0	
<i>SDE/Tvia</i>	0,8	
<i>SDL/Tvia</i>	0,8	

Fuente: Elaboración propia

ANEXO D: RESULTADOS MODELOS PR/PD_15

Tabla D-1: Resultados estimación modelos MNL-2 base PR/PD_15_1

Modelos Base PR/PD_15_1				
Coefficiente	2		3	
	Valor	Test-t robusto	Valor	Test-t robusto
<i>Niveles Servicio</i>				
Tiempo de viaje - PR/PD	-0.0433	-4.52	-0.0305	-4.03
Tiempo de acceso - PR	-0.0737	-5.58	-0.07	-5.32
Tpo. transbordo público - PR	-0.101	-3.97	-0.0605	-2
Tpo. transbordo público-privado - PR	-0.665	-3.82	-0.628	-3.78
Costo - PR/PD	-0.000489	-3.88	-0.000363	-3.19
Densidad Metro - PR				
Comfort - Ir sentado - PD	0.725	2.91	0.307	2.22
Incertidumbre en tiempos de viaje - PD	-2.27	-1.37	-0.889	-1.17
<i>Restricciones de Programación</i>				
<i>SDE</i> - PR/PD	-0.0361	-10.81		
<i>SDE</i> - PR			-0.0292	-7.06
<i>SDE</i> - PD			-0.0205	-3.88
<i>SDL</i> - PR/PD	-0.06	-8.34		
<i>SDL</i> - PR			-0.0762	-7.68
<i>SDL</i> - PD			-0.0205	-3.39
Llega tarde al trabajo - PR/PD	-0.403	-3.18	-0.255	-2.72
# paradas intermedias - PR	3.59	3.23	3.51	3.39
<i>Socioeconómicas</i>				
<i>AUTLIC</i> - PR	0.99	2.11	0.99	2.21
<i>Interacciones</i>				
Costo/tasa salarial (Costw) - PR/PD	-0.918	-3.49	-0.849	-3.57
<i>SDL</i> - Hombre			-0.0378	-2.99
<i>SDL</i> - Ingreso Alto			0.0478	3.88
<i>SDE</i> - Ingreso Medio Alto			-0.0205	-2.8
Tpo. transbordo público - Edad Mayor a 30			-0.0681	-1.78
<i>Factor de escala PR/PD</i>				
σ PR / σ PD *	0.49	-8.67	1.09	0.32
<i>ASC - PD</i>				
Salir antes	-1.63	-4.93	-0.603	-2.97
Salir después	-1.12	-3.48	-0.813	-3.17
Cambio modo	-3.03	-6.81	-1.31	-3.79
Log-verosimilitud	-2435.19		-2408.06	
rho-cuadrado ajustado	0.253		0.26	
# parámetros	25		31	
# observaciones	1743		1743	

Fuente: Elaboración propia

Tabla D-2: Test LR modelos MNL PR/PD base PR/PD_15_1

Modelos Base PR/PD_15_1		
	<i>1</i>	<i>2</i>
<i>Log-verosimilitud</i>		
Modelo PR	-1120	-1120
Modelo PD	-1302	-1302
Modelo PR/PD	-2423	-2435
Razón de Verosimilitud (LR)	2	26
# Parámetros comunes (k)	4	6
χ^2 (k; 1-α=95%)	9,49	12,59
¿ LR < χ^2 ?	Sí	No

Fuente: Elaboración propia

Para verificar que los modelos mixtos de PR/PD sean superiores a los modelos mixtos estimados de forma separada, es necesario realizar el test de razón de verosimilitud (LR) que se indica en el ANEXO A (Ortúzar y Willumsen, 2011). La Tabla D-2 muestra los resultados de aplicar este test al modelo 1 de la Tabla 5-11 y al modelo 2 de la Tabla D-1. En el modelo 2 existe una diferencia significativa entre los modelos estimados en forma separada y el mixto PR/PD, y, por lo tanto, se prefieren los modelos estimados en forma separada. En cambio, en el modelo 1 no hay diferencia significativa entre ambos modelos, por ende, se prefiere el modelo mixto PR/PD. Esto además confirma que los parámetros de *SDE* y *SDL* son distintos en PR y PD.

Tabla D-3: Valores del tiempo y razones de compromiso modelos PR/PD base PR_15_1

Modelos mixtos PR/PD base PR_15_1		
	<i>I</i>	<i>I_ECL</i>
Valor subjetivo (\$/min)		
Tiempo viaje	52	51
Tiempo acceso	119	118
Tiempo transbordo público	179	178
Tiempo transbordo público-privado	1.066	1.061
Ratios		
PR		
<i>SDL/SDE</i>	2,1	2,1
<i>SDE/Tiempo viaje</i>	1,1	1,2
<i>SDL/Tiempo viaje</i>	2,4	2,4
PD		
<i>SDL/SDE</i>	1,0	1,1
<i>SDE/Tiempo viaje</i>	0,7	0,8
<i>SDL/Tiempo viaje</i>	0,7	0,9

Fuente: Elaboración propia