



**UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA CIVIL**

**ANÁLISIS DE PATRONES DE VIAJE UTILIZANDO DATOS MASIVOS DE
TRANSPORTE PÚBLICO**

**TESIS PARA OPTAR AL GRADO DE MAGISTER EN CIENCIAS DE LA
INGENIERIA, MENCIÓN TRANSPORTE**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL

MARGARITA MAGDALENA AMAYA TORRES

PROFESORA GUÍA:
MARCELA MUNIZAGA MUÑOZ

MIEMBROS DE LA COMISIÓN:
CRISTIAN GUEVARA CUE
FRANCISCO MARTINEZ CONCHA
CAROLINA PALMA ALVARADO

Este trabajo ha sido parcialmente financiado por FONDEF D10I-1002 Y FONDECYT
1120288

SANTIAGO DE CHILE
2013

**RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE: Ingeniera Civil y al grado de
Magíster en Ciencias de la Ingeniería, mención
Transporte
POR: Margarita Magdalena Amaya Torres
FECHA: 20 de diciembre de 2013
PROFESORA GUÍA: Marcela Munizaga Muñoz**

ANÁLISIS DE PATRONES DE VIAJE UTILIZANDO DATOS MASIVOS DE TRANSPORTE PÚBLICO

Los cambios experimentados el año 2007 en el Sistema de Transporte Público de Santiago, Transantiago, modificaron la percepción de los ciudadanos respecto de la importancia del transporte en su calidad de vida, logrando que la clase política comprendiera la necesidad de asignar recursos técnicos y monetarios a este complejo sistema. Entre los cambios introducidos, la implementación de nuevas tecnologías de pago tuvo como efecto colateral el almacenamiento de cuantiosa y valiosa información que facilita el estudio del mismo sistema.

Este trabajo de tesis utiliza dichos datos para generar información sobre conductas recurrentes y patrones de viaje de los usuarios de Transantiago con el objetivo de desarrollar herramientas que permitan al planificador comprender de mejor manera el comportamiento de la demanda, y por lo tanto, aumentar la información disponible para el diseño de la oferta de transporte.

Esta tesis estudia la frecuencia de viaje, permanencia de las tarjetas en el tiempo, variables de calidad de servicio y perfiles de uso de tiempo. Por la gran cantidad de datos disponibles es posible analizar separadamente a las tarjetas de estudiantes, concluyendo que sus conductas de viaje son diferentes a las de tarjetas *bip!* tradicionales.

Uno de los problemas de los datos de transacciones, almacenados producto del pago del pasaje, es la falta de información socioeconómica de los usuarios. El principal aporte de este trabajo es el desarrollo de una metodología para estimar la zona de residencia de una tarjeta en base a condiciones espaciales y temporales de sus transacciones, que permite suplir en parte esa falencia, dada la naturaleza segregada de la localización residencial.

Utilizando esta información se presentan indicadores de calidad de servicio, como son tiempo promedio de viaje y trasbordo, según zona de residencia en la ciudad. Además se analiza la asignación de tiempo a actividades con el mismo nivel de segregación espacial.

Los resultados cuantifican el impacto de la elección de zona de residencia en el tiempo de viaje y trasbordo y las diferencias de comportamiento y asignación de tiempo a actividades. Las principales conclusiones de este trabajo son que una gran parte de la demanda está conformada por tarjetas que realizan viajes con poca frecuencia; que la distribución de viajes de la tarjeta estudiante agrupa usuarios con diferentes patrones de viaje; que existen diferencias en el perfil de uso de tiempo de los días lunes a jueves respecto del día viernes; y que las personas que habitan en las zonas Centro y Oriente de Santiago tienen tiempos de viaje menores y asignan más tiempo al hogar y a otras actividades distintas de trabajo y estudio.

Agradecimientos

A Domingo y Margarita, mis padres, gracias por tantos sacrificios para darnos a mis hermanos y a mí mucho más que educación. Gracias por enseñarme el valor de lo colectivo y por hacerme responsable de mis acciones. Gracias porque cuando se tiene la certeza del apoyo incondicional, se vive, se crece y se aprende con otra perspectiva.

Agradezco a mis hermanos, porque una infancia llena de juegos, sencillez y complicidad es fuente inagotable de alegría. A mi numerosa familia le agradezco tantas historias, consejos, preocupación y cariño.

A cada uno de mis profesores, gracias por no limitarse a entregar conocimientos. Aprecio sus experiencias, dedicación, ejemplos y sabiduría. Hago mención explícita a la Profesora Marcela por celebrar mis aciertos, comprender mis dificultades y porque valoro profundamente sus muchos esfuerzos por potenciar esta investigación mediante el trabajo en equipo.

A mis amigos de la época del liceo les agradezco rodearme de honestidad, seguridades y proyectos de futuro. A mis compañeros de primer año les debo la maravillosa sensación de pertenencia a una generación sin prejuicios. A mis queridos amigos civiles les agradezco por construir y encarnar mi concepto preferido de amistad. A los transportistas les agradezco la compañía durante este largo proceso, amistades sinceras y el común interés por el bienestar social. A mis compañeros de ideales, gracias por cambiar resignación y desesperanza por la promesa de un mundo mejor.

A ti, Camilo, de entre tantas cosas que podría agradecerte sólo mencionaré la más importante: hacerme inmensamente feliz.

Tabla de contenido

1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	2
1.3. Estructura de la tesis	3
2. Revisión bibliográfica	4
2.1. Potencial de los datos masivos de transporte público	4
2.2. Análisis de patrones de comportamiento y variabilidad temporal	6
2.3. Análisis de caracterización del sistema de transporte	10
2.3.1. Tiempo de viaje	10
2.3.2. Tiempo de trasbordo	11
2.3.3. Número de trasbordos	12
2.3.4. Tiempo de acceso	13
2.3.5. Perfil de carga	13
2.4. Modelos de fidelidad	14
2.5. Síntesis y conclusiones	14
3. Análisis de las bases de datos	16
3.1. Descripción del sistema de transporte y procesamiento de datos	16
3.1.1. Operación	16
3.1.2. Viaje y etapa	17
3.1.3. Características relevantes	17
3.1.4. Método de estimación de bajada	18
3.1.5. Método de estimación de propósito de viaje	19
3.2. Datos disponibles y procesamiento de la información	20
3.3. Análisis preliminar de las bases de datos de transacciones	22
3.3.1. Análisis de tarjetas tipo adulto	22
3.3.2. Análisis de tarjetas tipo estudiante	27
3.4. Permanencia de la tarjeta	31
3.5. Síntesis y conclusiones	33
4. Método de estimación de zona de residencia	35
4.1. Zonificación	35
4.2. Estimación de zona de residencia	36

4.2.1.	Metodología	36
4.2.2.	Resultados	39
4.2.3.	Validación exploratoria del método asignación de zona de residencia	40
4.3.	Síntesis y conclusiones	41
5.	Resultados	43
5.1.	Patrón de viajes diarios	43
5.1.1.	Patrón de viajes diarios de tarjetas tipo adulto	44
5.1.2.	Patrón de viajes diarios de tarjetas tipo estudiante	48
5.2.	Patrón de transacciones entre dos cortes temporales	52
5.3.	Tiempo de viaje en vehículo y tiempo de trasbordo	54
5.3.1.	Tiempo de viaje en vehículo y de trasbordo de tarjetas tipo adulto	55
5.3.2.	Tiempo de viaje en vehículo y de trasbordo de tarjetas tipo estudiante	61
5.4.	Perfil de uso de tiempo en Santiago	66
5.4.1.	Comparación de perfiles de uso de tiempo entre zonas	71
5.5.	Síntesis y conclusiones	80
6.	Conclusiones y futuras líneas de investigación	83
6.1.	Síntesis y conclusiones	83
6.2.	Líneas de investigación futura	85
7.	Referencias Bibliográficas	86
A.	Patrón de viajes diarios (código c++)	89
A.1.	Tarjeta Adulto	89
A.2.	Tarjeta Estudiante	89
B.	Patrón de transacciones entre 2 cortes temporales (tabla datos)	90
C.	Patrón de transacciones entre 2 cortes temporales (código c++)	91
D.	Tiempo de viaje en vehículo y tiempo de trasbordo (código c++)	92
D.1.	Tarjeta Adulto	92
D.2.	Tarjeta Estudiante	92
E.	Comparación de perfiles de uso de tiempo entre zonas (código c++)	93
	Tabla de contenido	

Índice de figuras

2.1.	Variabilidad en origen del primer viaje del día	7
2.2.	Patrón de actividades según sector de residencia, día laboral	9
2.3.	Patrón de actividades lunes a jueves	10
3.1.	Metodología de Estimación de Bajada	18
3.2.	Distribución de transacciones sep-08, tarjetas tipo adulto	23
3.3.	Distribución de transacciones ago-09, tarjetas tipo adulto	23
3.4.	Distribución de transacciones jun-10, tarjetas tipo adulto	24
3.5.	Distribución de transacciones abr-11, tarjetas tipo adulto	24
3.6.	Distribución de transacciones abr-12, tarjetas tipo adulto	24
3.7.	Comparación de distribución de transacciones, tarjetas tipo adulto	26
3.8.	Distribución de tarjetas tipo adulto	26
3.9.	Distribución de transacciones sep-08, tarjetas tipo estudiante	27
3.10.	Distribución de transacciones ago-09, tarjetas tipo estudiante	28
3.11.	Distribución de transacciones jun-10, tarjetas tipo estudiante	28
3.12.	Distribución de transacciones abr-11, tarjetas tipo estudiante	28
3.13.	Distribución de transacciones abr-12, tarjetas tipo estudiante	29
3.14.	Comparación de distribución de transacciones, tarjetas tipo estudiante	30
3.15.	Distribución de tarjetas tipo estudiante	31
3.16.	Distribución de tarjetas según número de cortes temporales	32
3.17.	Distribución de tarjetas según diferencia entre primer y último corte tempral	32
4.1.	Agregación de Santiago en seis zonas	36
4.2.	Método de estimación de zona de residencia	38
4.3.	Distribución de los radios de circunferencia	39
5.1.	Distribución de viajes totales, tarjetas tipo adulto	45
5.2.	Distribución de viajes, tarjetas tipo adulto, un viaje diario	45
5.3.	Distribución de viajes, tarjetas tipo adulto, dos viajes diarios	46
5.4.	Distribución de viajes, tarjetas tipo adulto, tres viajes diarios	47
5.5.	Distribución de viajes, tarjetas tipo adulto, cuatro viajes diarios	47
5.6.	Distribución de viajes, tarjetas tipo adulto, cinco viajes diarios	48
5.7.	Distribución de viajes totales, tarjetas tipo estudiante	49
5.8.	Distribución de viajes, tarjetas tipo estudiante, un viaje diario	49
5.9.	Distribución de viajes, tarjeta tipo estudiante, dos viajes diarios	50

5.10. Distribución de viajes, tarjetas tipo estudiante, tres viajes diarios	51
5.11. Distribución de viajes, tarjetas tipo estudiante, cuatro viajes diarios	51
5.12. Distribución de viajes, tarjetas tipo estudiante, cinco viajes diarios	52
5.13. Patrón de transacciones semanales entre abril-2011 y abril-2012	53
5.14. Tiempo de viaje en vehículo y desviación estándar, tarjeta adulto	55
5.15. Tiempo de trasbordo y desviación estándar, tarjeta adulto	56
5.16. Comparación días hábiles de variación del tiempo de viaje, tarjeta adulto . . .	57
5.17. Comparación días hábiles de variación del tiempo de trasbordo, tarjeta adulto	57
5.18. Tiempo promedio de viaje en vehículo y trasbordo, tarjeta adulto, por zona .	59
5.19. Tiempo de viaje en vehículo y desviación estándar, tarjeta estudiante	61
5.20. Tiempo de trasbordo y desviación estándar, tarjeta estudiante	62
5.21. Comparación días hábiles de variación del tiempo de viaje, tarjeta estudiante	62
5.22. Comparación días hábiles de variación del tiempo de trasbordo, tarjeta estu- diante	63
5.23. Tiempo promedio de viaje en vehículo y trasbordo, tarjeta estudiante, por zona	64
5.24. Comparación de % de tarjetas en actividad hogar entre días laborales	66
5.25. Comparación de % de tarjetas en actividad trabajo entre días laborales	67
5.26. Comparación de % de tarjetas en actividad estudio entre días laborales	68
5.27. Comparación de % de tarjetas en actividad otros entre días laborales	68
5.28. Comparación de % de tarjetas en actividad viaje entre días laborales	69
5.29. Perfil de uso de tiempo, lunes-jueves	69
5.30. Perfil de uso de tiempo, viernes	70
5.31. Perfil de uso de tiempo, sábado	70
5.32. Perfil de uso de tiempo, domingo	71
5.33. Perfil de uso de tiempo actividad hogar por zona, día laboral	72
5.34. Comparación de perfiles de uso de tiempo actividad hogar por zona, día laboral	72
5.35. Perfil de uso de tiempo actividad hogar por zona, fin de semana	73
5.36. Comparación de perfiles de uso de tiempo actividad hogar por zona, fin de semana	73
5.37. Perfil de uso de tiempo actividad trabajo por zona, día laboral	73
5.38. Comparación de perfiles de uso de tiempo actividad trabajo por zona, día laboral	74
5.39. Perfil de uso de tiempo actividad trabajo por zona, fin de semana	74
5.40. Comparación de perfiles de uso de tiempo actividad trabajo por zona, fin de semana	75
5.41. Perfil de uso de tiempo actividad estudio por zona, día laboral	75
5.42. Comparación de perfiles de uso de tiempo actividad estudio por zona, día laboral	75
5.43. Perfil de uso de tiempo actividad estudio por zona, fin de semana	76
5.44. Comparación de perfiles de uso de tiempo actividad estudio por zona, fin de semana	76
5.45. Perfil de uso de tiempo actividad otros por zona, día laboral	77
5.46. Comparación de perfiles de uso de tiempo actividad otros por zona, día laboral	77
5.47. Perfil de uso de tiempo actividad otros por zona, fin de semana	78

5.48. Comparación de perfiles de uso de tiempo actividad otros por zona, fin de semana	78
5.49. Perfil de uso de tiempo actividad viaje por zona, día laboral	79
5.50. Comparación de perfiles de uso de tiempo actividad viaje por zona, día laboral	79
5.51. Perfil de uso de tiempo actividad viaje por zona, fin de semana	79
5.52. Comparación de perfiles de uso de tiempo actividad viaje por zona, fin de semana	80

Índice de tablas

3.1. Información de las bases de datos disponibles	21
3.2. Procesos aplicados a las bases de datos	22
3.3. Información de transacciones para tarjeta tipo adulto	23
3.4. Información de transacciones para tarjeta tipo estudiante	27
4.1. Validación exploratoria del método de estimación de zona de residencia	41
5.1. Información del patrón de viajes diarios para tarjeta tipo adulto	44
5.2. Información del patrón de viajes diarios para tarjeta tipo estudiante	48
5.3. Variación del número de transacciones abril2011-abril2012	54
5.4. Tiempo promedio de viaje en vehículo y trasbordo, tarjetas tipo adulto	56
5.5. Tiempo de viaje en vehículo y de trasbordo por zona, tarjetas tipo adulto	60
5.6. Tiempo promedio de viaje en vehículo y trasbordo, tarjetas tipo estudiante	62
5.7. Tiempo de viaje en vehículo y de trasbordo por zona, tarjetas tipo estudiante	65
5.8. Número de observaciones por zona y tipo de día	71

Capítulo 1

Introducción

1.1. Motivación

Durante décadas, los investigadores en el área de transporte han estudiado las características de los modos de viaje y las conductas de viaje de usuarios bajo la premisa de escasez de datos. Se han hecho importantes esfuerzos en las áreas de la estadística y la econometría para utilizar los datos provechosamente y con la menor pérdida de información. Sin embargo, esta premisa ha comenzado a romperse con las nuevas tecnologías.

Con el objetivo central de mejorar la operación al interior de los buses, se han diseñado dispositivos que permiten pagar el pasaje de transporte sin afectar la concentración del operador del bus ni utilizar más personal. Como efecto colateral, estos dispositivos almacenan datos requeridos para validar el pago del pasaje. Estos datos se han transformado en información valiosa y, por sobre todo, cuantiosa para la investigación.

En el caso particular de Santiago de Chile, su sistema de Transporte Público almacena abundantes datos provenientes de las transacciones de pago realizadas por los usuarios al utilizar el sistema, además de la información obtenida de las emisiones de los GPS con que cuentan los buses. Esta significativa cantidad de información hoy está disponible y registra datos de todos los usuarios de tarjetas inteligentes de pago de manera continuada. La base de datos contiene información de la utilización del transporte público por parte de los pasajeros y presenta la posibilidad de analizar su comportamiento y patrones de viaje.

La literatura pionera en el tema describe los potenciales usos de este tipo de bases de datos en la estimación de variables relevantes para un sistema de transporte. Por ejemplo, Bagchi y White (2004) señalan el tipo de información disponible para análisis: la regularidad con que un usuario utiliza su tarjeta, información de viajes de los usuarios en grandes periodos de tiempo, fecha y hora de utilización de los servicios de la red, información del lugar de subida del viaje y en algunos casos del lugar de bajada, caracterización de los usuarios en los casos que existen tarjetas personalizadas e información sobre los montos de carga y pago. En particular, con la información descrita se podría determinar qué zonas de la ciudad presentan los mayores tiempos de viaje, espera, excesivo número de trasbordos, entre otras variables de

calidad de servicio.

Posteriormente, la literatura se ha enfocado en el estudio de variables de calidad de servicio y regularidad espacial/temporal de los viajes. Se plantea estudiar la estructura de la demanda, su comportamiento y necesidades para entender las causas de sus fluctuaciones y mejorar la predicción temporal y espacial de la misma. Estas mejoras son herramientas cruciales para que oferta y demanda coincidan en lugar, momento y cantidad. En particular, la información anterior permite

Uno de los objetivos centrales de este trabajo es la búsqueda de comportamientos de viaje sistemáticos y conductas recurrentes que mejoren el conocimiento del planificador respecto a las características de la demanda de transporte. En esta tesis se propone utilizar los datos de transacciones de tarjetas de pago para realizar estos análisis, además de entregar información sobre la forma de viaje de los usuarios del sistema de transporte público de Santiago.

1.2. Objetivos

El objetivo principal de esta tesis es analizar los datos de transacciones de *Transantiago* en busca de patrones de viaje y conductas comunes entre usuarios, que permitan comprender y conocer mejor el comportamiento de los usuarios del sistema de transporte público de Santiago. Para llevar esto a cabo se desarrollan dos tipos de análisis: (1) aquellos que permiten analizar datos entre cortes temporales y (2) aquellos que permiten analizar los datos en busca de diferencias en el comportamiento entre zonas de la ciudad, en un mismo corte temporal.

Los objetivos específicos son:

- Explorar y analizar las bases de datos de transacciones de *Transantiago* para cinco períodos de tiempo de una semana cada uno.
- Generar rutinas que permitan analizar los datos de transacciones como viajes independientes, como viajes realizados por un mismo usuario y que permitan comparar cortes temporales.
- Buscar características de viaje comunes entre usuarios.
- Identificar perfiles o tipos de usuario por sus patrones de viaje recurrentes.
- Estudiar variabilidad en el origen del viaje para un mismo usuario y desarrollar una metodología para predecir la localización de su hogar.
- Concluir sobre los patrones de viaje de usuarios en la ciudad de Santiago y su variación en el tiempo.

1.3. Estructura de la tesis

El siguiente capítulo describe los potenciales usos de los datos masivos de transporte público en la planificación. Se exponen trabajos con información relativa a patrones de viaje y estudios de variabilidad y regularidad de comportamiento de la demanda. La revisión concluye con una sección dedicada a estudiar la lealtad de los usuarios al sistema de transporte.

En el capítulo 3 se expone una síntesis de las características generales de *Transantiago*, sistema de transporte público de Santiago, del cual provienen las bases de datos analizadas en esta tesis. A continuación se describe la información contenida en las bases de datos. El capítulo culmina con un análisis preliminar de esta información.

Las conclusiones desprendidas de los capítulos 2 y 3 dan soporte a la propuesta de qué características de los datos estudiar en busca de patrones de viaje de los usuarios del sistema de transporte. El capítulo 4 presenta el método desarrollado para predecir la zona de residencia de una tarjeta a nivel agregado. En el capítulo 5 se muestran los resultados obtenidos y los patrones de viaje encontrados. En el capítulo 6 se presentan las conclusiones generales más importantes de la tesis junto con la descripción de las líneas futuras de investigación.

Al final de la tesis se encuentran los anexos que incluyen, entre otras cosas, los códigos de las rutinas utilizadas.

Capítulo 2

Revisión bibliográfica

La literatura asociada a datos masivos de transporte público, obtenidos de las transacciones realizadas con tarjetas inteligentes de pago, describe la gran riqueza de estos datos en cuanto a la cantidad de información que se puede extraer de ellos. Debido a que la información obtenida no puede usarse directamente, se han desarrollado nuevas metodologías que permiten extraer la información necesaria para su posterior análisis. Estas nuevas metodologías han enriquecido los datos, permitiendo extraer información que en un comienzo no se consideraba disponible.

En esta revisión bibliográfica se sistematiza, revisa y analiza críticamente la literatura que utiliza datos masivos de transporte público para establecer patrones de viaje y caracterizar el comportamiento de los usuarios. La revisión comienza señalando las características y el potencial de los datos masivos de transporte público en general, además de las diferencias respecto de encuestas convencionales, todos ellos temas descritos en la literatura pionera en esta área.

Luego se describen los avances en la literatura en el análisis de datos masivos de transporte, en dos grupos: análisis de patrones de comportamiento y variabilidad temporal, y análisis de caracterización del sistema de transporte. El primero, análisis de patrones de comportamiento y variabilidad temporal, muestra los avances en análisis de datos donde se observan ciertos patrones y conductas de viaje y su variación en el tiempo. El segundo, análisis de caracterización del sistema de transporte, muestra el tipo de trabajo descrito en la literatura para entender y evaluar los sistemas de transporte público en función de las variables que afectan al usuario.

Finalmente, se incluye un capítulo dedicado a la fidelidad de usuarios al sistema de transporte, entendiendo que esta información aporta a conocer el lapso temporal y frecuencia de uso de la tarjeta de pago llamada *bip!*.

2.1. Potencial de los datos masivos de transporte público

Los primeros trabajos que utilizan información de datos masivos de transporte público y describen su potencial como herramienta complementaria a las tradicionales encuestas

de recolección de datos mediante diarios de viajes o similares, son Bagchi y White (2004, 2005). En su trabajo describen los alcances de la información proveniente del registro de transacciones:

- La información de cada viaje es almacenada con un número de serie único asociado a cada tarjeta. Por tanto, cada viaje puede ser atribuido a una tarjeta dada.
- Se obtiene datos para un gran volumen de pasajeros, dependiendo del nivel de penetración de este medio de pago. Todos los viajes realizados con otro medio de pago no pueden agruparse por usuario ni asociarse al dueño de alguna tarjeta.
- Aunque es posible que un usuario no realice todos sus viajes pagando con su tarjeta y se pierda cierta información, se podrá obtener información continua de los viajes de cada individuo, por periodos de tiempo que superan por mucho a las tradicionales encuestas de viajes de un día o una semana de duración.
- Idealmente, cada viaje realizado con la tarjeta de pago debería registrar el punto de subida y bajada del modo de transporte. Sin embargo, típicamente se exige acercar la tarjeta al validador sólo en la subida a un modo de viaje, por lo que se pierde la información del lugar y hora de bajada.
- En cada transacción se guarda fecha, hora y detalles del modo o servicio utilizado.
- En el caso de tarjetas personalizadas, es posible tener información de los usuarios, como lugar de residencia, género y edad, información que puede ser relevante para la investigación y planificación.
- Se cuenta con información del lugar de recarga de los usuarios y los tipos de tarifa pagada.
- No es posible capturar información del propósito de viaje.
- No es posible obtener información de la percepción o satisfacción de los pasajeros.

Considerando todos estos puntos, Bagchi y White (2004) señalan que las encuestas tradicionales de viajes y los datos masivos de transporte público juegan un rol complementario. Ambas herramientas difieren en la información que son capaces de proporcionar y los aspectos en que uno aventaja al otro, derribando la hipótesis de que la información de tarjetas de pago acabaría desplazando a las encuestas tradicionales de viajes.

Posteriormente, Utsunomiya et al. (2006) señalan potenciales usos de la información, agrupados según áreas de estudio, como son: planificación de servicios y estudio del mercado, predicción de la demanda de viajes, operaciones, políticas tarifarias, marketing y publicidad. En el primer punto, planificación de servicios y estudio del mercado, se describe brevemente

las posibilidades de estudio de calidad de servicio, cobertura y conductas de viaje de los pasajeros a nivel temporal y espacial.

Más recientemente, Pelletier et al. (2010) resumen el uso de la información de tarjetas inteligentes como medio de pago de transporte en una extensa revisión bibliográfica. En ella describen los tipos de tarjeta y su forma de uso (tarjetas con y sin contacto), la manera en que se almacena la información, los dispositivos involucrados en el funcionamiento de sistemas de pago automático, las ventajas y desventajas de utilizar tarjetas inteligentes de pago en distintas áreas (como la de seguridad y confidencialidad), experiencias en marketing y una síntesis de los trabajos avocados al área de transporte.

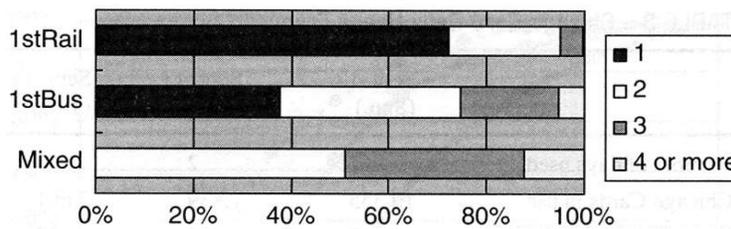
2.2. Análisis de patrones de comportamiento y variabilidad temporal

La mayoría de los reportes de la literatura en datos masivos de transporte público corresponden a información obtenida del análisis de un breve período de tiempo (usualmente 7 días consecutivos). Sin embargo, cuando existen datos de largos periodos de tiempo, se puede ampliar la gama de estudio, observando variabilidad en el tiempo, cambios estacionales y regularidad.

Uno de estos casos es el mostrado por Morency et al. (2007), quienes analizan datos de viajes de 10 meses consecutivos. Una primera aproximación es observar el número de viajes por semana, según tipo de tarjeta, y su evolución en el tiempo. Por ejemplo, observan una muy importante baja en el número de transacciones para todas las tarjetas salvo adulto mayor, durante la primera semana de agosto, la cual coincide con un período de vacaciones. Con esta información pueden observar como varía el patrón de viajes de los tipos de usuarios en vacaciones y sobre todo, pueden cuantificar con precisión el descenso de la demanda en este periodo según para cada tipo de tarjeta.

Una mirada diferente es mostrada por Utsunomiya et al. (2006), quienes pese a tener pocos días observados estudian la variabilidad espacial del lugar de origen del primer viaje de cada día, en día laboral, para todos los usuarios, en un lapso de tiempo determinado. Uno de sus resultados es que existen fuertes diferencias en el promedio de la variabilidad de origen del viaje entre zonas de la ciudad. Al clasificar a los usuarios según tipo de modo de transporte utilizado (sólo metro, sólo bus, metro y bus) observan diferencias en variabilidad de origen del viaje. Por ejemplo, los usuarios de metro son más propensos a utilizar la misma estación cada día (sobre el 70%), respecto de los usuarios de bus donde menos del 40% utiliza la misma parada en su primer viaje del día, como se muestra a continuación:

Figura 2.1: Variabilidad en el origen del primer viaje del día según modo de transporte. Fuente: Utsunomiya et al. (2006).



En el mismo trabajo, Utsunomiya et al. muestran las diferencias en elección de ruta del primer viaje del día entre usuarios de distintas zonas de la ciudad, según el número de rutas usadas. Se observa que las zonas cercanas al centro muestran mayor variabilidad en la elección de ruta, lo que se condice con una red más densa y mejores alternativas de viaje en ese lugar.

Por su parte, Morency et al. (2006) cuentan con datos de 277 días consecutivos. En un enfoque más cercano a conocer la forma de viajar del individuo, analizan la variabilidad del comportamiento de los usuarios clasificándolos con un índice propuesto, llamado *Activity rate* (AR); el cual se calcula como el número de días en que la tarjeta fue validada para abordar un bus dividido en el número de días de observación de la tarjeta (último día de subida – primer día de subida).

Con este índice clasifican a los usuarios según el nivel de utilización de la tarjeta y el número total de días de observación. Se encuentra que pocas tarjetas se acercan a tener un 100 % de utilización, salvo aquellas que fueron usadas pocos días, por ejemplo, dos días consecutivos y nada más. El usuario más común parece ser aquel que viaja de manera infrecuente.

En el caso de las tarjetas con un alto AR, Morency et al. (2006) estudian el comportamiento de viaje en cuanto la regularidad temporal y espacial por tarjeta. Para ello se agrupa transacciones según lugar y hora del día. Luego, se observa la proporción de viajes realizados en los mismos horarios y lugares, entregando un índice de regularidad del comportamiento de las tarjetas. Lógicamente, el estudio de la regularidad requiere de usuarios que utilizan la tarjeta con cierta frecuencia.

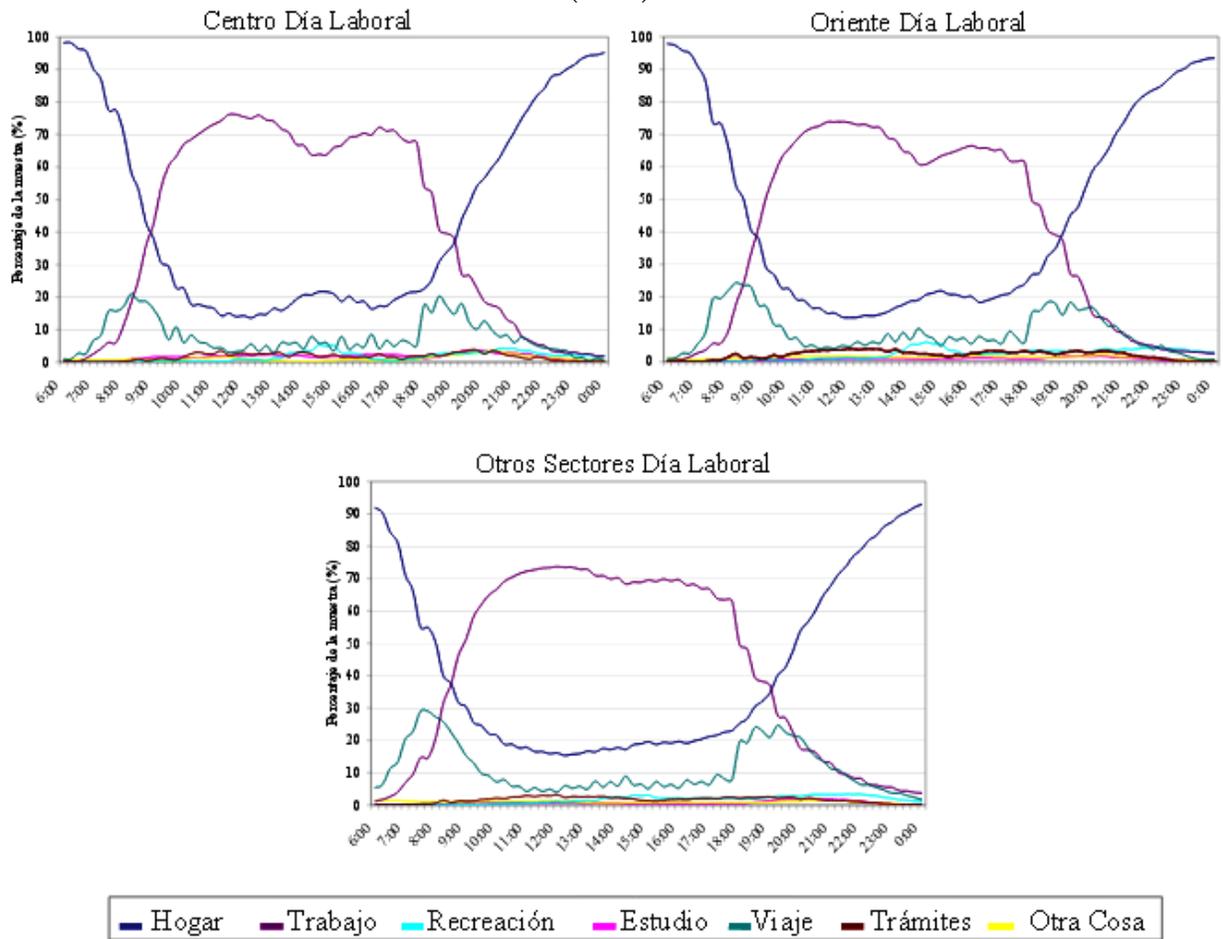
Agard et al. (2009) utilizan minería de datos para analizar los datos de transacciones y estudiar conductas de viaje. Proponen generar n grupos donde reunir a aquellos usuarios que tengan características similares, de manera tal que cada grupo contenga a aquellos individuos más parecidos entre sí y que cada grupo sea lo más diferente posible de los otros grupos. El resultado son cuatro grupos de diferentes comportamientos, y por tanto, tipos de usuarios. El primer grupo está compuesto principalmente por trabajadores y se caracteriza por alta regularidad en la hora que se realizan los viajes, concentrados en punta mañana y punta tarde. Un segundo grupo, también regular, concentra los viajes en punta mañana y punta mediodía. El tercer grupo encierra a adultos y estudiantes con poca regularidad, pero uso

constante de la tarjeta. Un último grupo reúne a quienes usan la tarjeta con poca frecuencia. Esta caracterización de las tarjetas permite conocer e identificar los perfiles de usuarios del sistema, pero sobre todo, cuantificarlos y saber qué proporción de ellos genera variabilidad en la demanda. Por tanto, la regularidad y frecuencia se observan conjuntamente como factores a considerar en la planificación de la oferta de transporte.

En la misma línea, Ma et al. (2013) observan la primera y última transacción del día y su variabilidad temporal y espacial. Con esta información agrupan a los individuos en 5 grupos según su regularidad, encontrando que los usuarios regulares tienen tiempos promedio de viaje diarios mayores por pasajero y que realizan hasta 6 veces más viajes por día que los viajeros más irregulares.

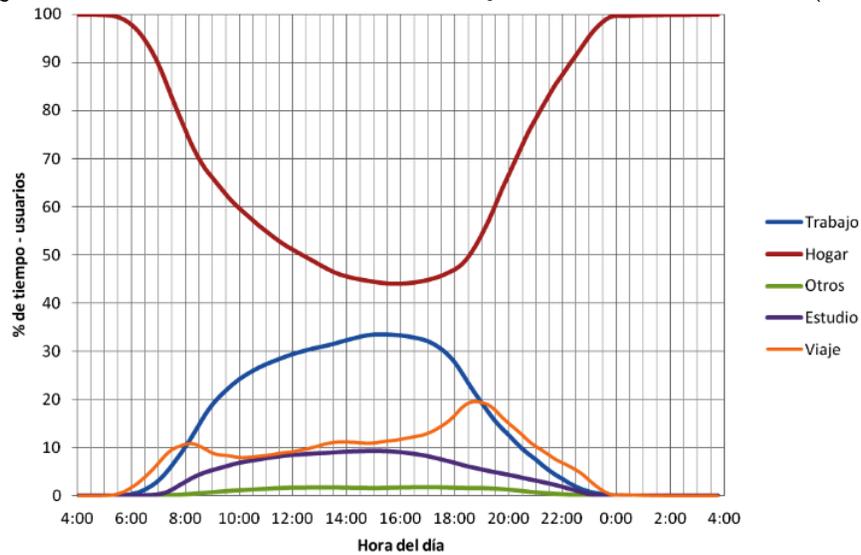
En todos los casos presentados anteriormente se analiza el comportamiento de los usuarios en cuanto a sus viajes y el patrón de conductas asociadas a movilización. La Figura ?? muestra el patrón de actividades según sector de residencia para la ciudad de Santiago. Esta sustancial contribución fue desarrollada por Olgún et al. (2009), quienes utilizan los datos de una encuesta Origen-Destino (que describe la secuencia de viajes de un periodo determinado) para analizar conductas de la población y asignación de tiempo, más allá de lo relativo a viajes. La ventaja de esta encuesta es que se posee información socioeconómica detallada de los individuos y el propósito de viaje. Con esto se construye la asignación de tiempo de cada individuo y posteriormente el patrón de actividades por propósito, tipo de día, zona de la ciudad, género y rango de edad.

Figura 2.2: Patrón de actividades según sector de residencia, día laboral. Fuente: Olgúin et al. (2009).



Una adaptación de lo anterior, utilizando datos masivos de transporte público, es propuesta por Devillaine (2012) para las ciudades de Santiago y Gatineau (ver Figura ??). Devillaine desarrolla un método que construye las cadenas de viaje de un individuo, distinguiendo entre puntos de transferencia y destino de viaje, y que asigna un propósito estimado a cada viaje basándose en reglas de carácter temporal y espacial. Adicionalmente, en el caso de la ciudad de Gatineau, existe información del uso de suelos que posibilita refinar la asignación de propósito de cada viaje. Los datos permiten asignar el tiempo de los usuarios a cinco propósitos posibles: Hogar, Trabajo, Estudio, Otros y Viaje. Sin embargo, Devillaine no puede analizar los datos desagregando según alguno de los criterios descritos por Olgúin et al., debido a la falta de información socioeconómica de éstos. En particular, no es posible analizar la información según zona de residencia. La siguiente figura muestra el patrón de actividades construido por Devillaine en base a la información de transacciones:

Figura 2.3: Patrón de actividades lunes a jueves. Fuente: Devillaine (2012).



Otro trabajo que busca obtener información adicional a partir de las tarjetas de pago es realizado por Ordoñez y Erath (2013), quienes se basan en información de transacciones, encuestas a hogares y datos de uso de suelos para determinar la distribución espacial y temporal de lugares de trabajo.

2.3. Análisis de caracterización del sistema de transporte

Aunque el presente trabajo no tiene como objetivo estudiar directamente la calidad de servicio, ésta es una de las áreas más exploradas en la literatura. Es de particular interés para este trabajo conocer cómo se miden las distintas variables de calidad de servicio utilizando datos masivos, para buscar posibles patrones de viaje en ellas. A continuación se presentan las variables más observadas, agrupadas según tipo de variable.

2.3.1. Tiempo de viaje

El tiempo de viaje es una de las variables más relevantes en la evaluación de calidad de servicio de un sistema de transporte público. Existen numerosas maneras de observar el tiempo de viaje usando datos pasivos para diagnosticar y tomar medidas de mejoramiento del sistema de transporte. Una de ellas es mostrada por Park et al. (2008), quienes analizan los tiempos de viaje para la ciudad de Seúl. Muestran la distribución de los tiempos de viaje según modo de transporte, encontrando importantes diferencias entre éstos. Luego, Park et al. presentan información sobre cómo varía el número y el tiempo promedio de viaje según la hora de inicio del mismo. Como resultado, observan cierta estabilidad en los tiempos, con aumentos levemente superiores al 10% en punta mañana, donde el número de pasajeros triplica a los de horario valle. Es decir, el aumento de pasajeros, en promedio, no causa grandes aumentos en el tiempo de viaje. Las dos formas en que Park et al. observan el tiempo de viaje

permiten detectar problemáticas espaciales o temporales que se registren sistemáticamente en el sistema de transporte.

Una mirada distinta al análisis de tiempos de viaje es realizada por Jang (2010). Para ello genera mapas donde colorea las zonas de la ciudad según el tiempo requerido para desplazarse desde cada una de las zonas de la ciudad a una zona en particular. Para ejemplificar, se muestra la distribución de tiempo de viaje para ir al centro de Seúl. Si la velocidad sólo dependiera de la distancia a un punto, se esperaría ver anillos del mismo tono concéntricos a la zona de destino. Sin embargo, Jang observa zonas contiguas, que deberían pertenecer al mismo anillo, con considerables diferencias de tiempo de viaje. Por tanto, los mapas contruidos por Jang permiten detectar zonas específicas que son especialmente mal servidas o con algún problema de accesibilidad.

Continuando con lo anterior, Jang aporta con otra forma de visualización que consiste en colorear cada zona del mapa de Seúl con una escala asociada a la diferencia de tiempo de viaje si se utiliza metro o bus, para llegar al centro de la ciudad. Con esto, el mapa muestra en qué modo conviene viajar a una cierta zona, como el centro de la ciudad, según el lugar de origen del viaje.

Como se describe en los cuatro casos anteriores, la misma variable visualizada y analizada de distintas formas, entrega información que permite corregir y medir el desempeño del sistema de transporte público en términos de accesibilidad, problemas en ciertos modos o servicios y respuesta ante cambios previsibles de la demanda.

2.3.2. Tiempo de trasbordo

El tiempo de trasbordo es otra de las variables principales al evaluar calidad de servicio. Chu y Chapleau (2008) estudian la distribución de los tiempos de transferencia, los cuales se definen como la diferencia entre el tiempo de bajada estimado de una transacción de un usuario respecto al tiempo de subida de la transacción siguiente del mismo usuario. Se observa que sobre el 50 % de las observaciones con transferencias tienen tiempos de transferencia menores a los siete minutos, por transferencia, mientras que el 80 % de las observaciones tiene tiempos de transferencia inferiores a 18 minutos. Pese a los buenos números en promedio, el trabajo no se refiere a la cola de la distribución (20 %), ni a su ubicación espacial o problemas sistemáticos en el trasbordo.

Seaborn et al. (2009) utilizan los datos de tarjetas de pago para calibrar el umbral de tiempo desde el cual dos transacciones consecutivas dejan de ser etapas de un viaje y se convierten en dos viajes distintos. Los resultados son validados con una encuesta. Encuentran que el umbral de tiempo que determina con mayor precisión si dos transacciones son etapas de un solo viaje o si son dos viajes distintos, depende de los modos entre los cuales se transfiere y los ordenan de menor a mayor como sigue: metro-bus (20 min), bus-metro (35 min) y bus-bus (45 min). Este análisis tiene sentido en los sistemas de transporte que no requieren

validación en el punto de bajada y puede ser un gran aporte a determinar con mayor precisión el lapso de tiempo en que dos transacciones deben ser consideradas viajes distintos.

2.3.3. Número de trasbordos

El número de trasbordos por viaje y la proporción de viajes sin trasbordos son algunos índices de cómo se estructura el transporte de una ciudad, por ejemplo, en cuanto a indicar en qué medida se privilegia el uso de líneas directas respecto de otro tipo de estructuras. En este ámbito, Jang (2010) cuantifica el número de viajes realizados durante siete días consecutivos, clasificándolos según las combinaciones de modos utilizados. En el caso de la ciudad de Seúl, hasta cuatro trasbordos son permitidos en un viaje. No se considera los trasbordos al interior del metro, pues estos trasbordos no requieren validación y por tanto no queda registro de ellos. Pese a la cantidad de trasbordos permitidos, más de un 68 % de las transacciones observadas son viajes de una etapa (sin trasbordo), sobre el 26 % de las transacciones son viajes de dos etapas y un 5 % son viajes de tres o más etapas.

En un análisis similar al de Jang, Utsunomiya et al. (2006) trabajan con datos de la ciudad de Chicago. Observan con qué frecuencia se producen viajes de una o más etapas, según la combinación de modos utilizados, pero sólo para el primer viaje del día. Destacan que sobre un 85 % de los viajes correspondientes al primer viaje del día, son viajes sin transferencias.

Profundizando en esto, Utsunomiya et al. observan la frecuencia de los tipos de cadenas diarias de viaje según combinaciones de modos utilizados a lo largo del día. Por ejemplo, la cadena de transacciones “BMMB” significa que una tarjeta realizó las siguientes transacciones en un día de viaje: Bus, Metro, Metro, Bus. Existen cadenas simétricas y asimétricas, como “MMB”. Aunque las cadenas simétricas de viaje son el comportamiento más común, destaca un 30 % de viajes no simétricos, correspondientes a transacciones únicas en el día realizadas en metro o en bus. El resultado confirma lo mostrado por Morency et al. (2006): existe un considerable número de tarjetas de las cuales existe registro apenas uno o dos días de la muestra. Este análisis es una primera aproximación a entender las formas de viaje y plantea la posible utilización de modos no integrados como complemento. El estudio más profundo de este fenómeno requiere modificar el trabajo de Utsunomiya et al. y observar la distribución temporal de los viajes de una misma tarjeta en un día, en cuanto a número de viajes realizados en ese día, más que por modo utilizado.

En un esfuerzo por visualizar espacialmente la ubicación de los trasbordos, Hofmann y O'Mahony (2005) muestran el número de transferencias entre pares origen-destino de zonas de manera tridimensional, nombrando a cada zona con una letra. Este análisis permite establecer el número de trasbordos por zona, las zonas de la ciudad donde éstos se concentran y dónde es necesario tomar medidas para reducir elevados tiempos de trasbordo. Incluso, esta información es útil para desarrollar mejoras en muestreos para encuestas de viaje, al proporcionar con detalle los puntos con mayores y menores transferencias.

Enriqueciendo el trabajo previo, Chu y Chapleau (2008) muestran la distribución geográfica del total de transacciones de un día, lo que permite un análisis espacial de las transacciones. Para cada parada se especifica la cantidad de transacciones de no transferencia (origen de viaje) y las transacciones de transferencia.

Un análisis similar a este y el anterior, pero desagregado por horario, ayudaría a establecer qué zonas de la ciudad son de tipo residencial, comercial o simplemente puntos de transferencia clave para los viajes.

2.3.4. Tiempo de acceso

Por las características propias de los datos masivos de transporte público, la información disponible comienza en el momento de abordar al modo de transporte. No es posible tener información del tiempo de espera o acceso previos a la primera etapa de viaje. Una excepción a este caso se tiene cuando las tarjetas de pago son personalizadas y existe información adicional que permite enriquecer el análisis, como muestra Utsunomiya et al. (2006), donde las tarjetas están asociadas a una dirección de facturación, que usualmente es el lugar de residencia del individuo. Con esta información se realizan histogramas de distancia estimada de acceso para los modos bus y metro. Los resultados muestran que existe una mayor disposición a caminar para acceder a estaciones de metro que a paraderos de buses. Este resultado es asociado a la mayor calidad del modo metro respecto del bus y a que tiene una red menos densa de estaciones respecto de la red de paraderos de buses. Así mismo, se encuentran diferencias significativas en la distribución de distancia de acceso al modo metro según zona de residencia. Las distancias son menores en las zonas con mayor densidad de red. No ocurre lo mismo con el modo bus, donde la red tiene una densidad con menor variación.

Tanto en el caso de tiempo de acceso, como en otras cualidades del transporte, es razonable pensar que existan diferencias en el comportamiento entre diferentes zonas de residencia de la ciudad, por lo que se explorará este punto.

2.3.5. Perfil de carga

Una de las áreas menos reportadas es el estudio de perfiles de carga en buses. Para estudiarlos se requiere información adicional a las transacciones, como el itinerario de los buses en funcionamiento o la posición de cada bus en el tiempo. Además, es necesario contar con el punto de bajada (o la estimación de este punto) de cada transacción. Finalmente, sólo es posible obtener buenos perfiles si existe un registro adecuado de todos los medios de pago que permiten utilizar un modo. Cuando no existe una alta penetración en el uso de tarjetas inteligentes y no se tiene información precisa de qué vehículo de transporte abordan quienes usan otros medios de pago, no existe forma de calcular perfiles de carga (solo pueden estimarse). Chu y Chapleau (2008) determinan el perfil de carga a nivel de paraderos y de buses. Con esto se puede observar la variabilidad de la demanda y los efectos de salirse del itinerario previsto para el bus.

2.4. Modelos de fidelidad

Otra línea de investigación con datos de tarjetas inteligentes de pago estudia el concepto de lealtad de los individuos al uso de la red de transporte público. Este es el caso del trabajo presentado por Trépanier y Morency (2010), quienes desarrollan un modelo de regresión lineal que predice el tiempo que un individuo se mantiene utilizando el transporte público. El modelo corresponde a una regresión lineal múltiple donde la variable dependiente es el número de meses transcurridos entre la primera y la última transacción realizadas por una tarjeta, y las variables independientes (convertidas en variables tipo dummy) son: mes en que se comenzó a usar la tarjeta de pago, año en que se comenzó a usar la tarjeta de pago, zona de la ciudad en que se localiza el hogar y tipo de tarjeta.

Este modelo es bastante básico, puesto que no incorpora variables de calidad de servicio del transporte. Además, considera que el año de ingreso explica la permanencia en el sistema, lo que entrega como resultado que una persona que comenzó a usar el sistema al comienzo del estudio es más leal que uno que ingreso después. Sería más adecuado pensar que aquellos que ingresaron antes al sistema, llevan más tiempo en él y por eso han permanecido más tiempo usándolo. Parece ser una variable endógena del modelo.

Otro modelo que estudia la fuga de pasajeros de transporte público es mostrado por Bass et al. (2010). A diferencia del caso anterior, Bass et al. cuentan con información proveniente de una encuesta, la cual considera datos de cuatro cortes temporales. Se trata de un modelo de elección discreta tipo Logit Jerárquico que predice si un usuario de transporte público en un período t continuará en el sistema en el período $t+1$ (no migra del transporte público) o si se cambiará a un modo de transporte privado (migra del transporte público). A diferencia del modelo de Trépanier y Morency, la permanencia en el sistema depende de los parámetros de calidad de servicio de los modos y de variables socioeconómicas del individuo, como: tiempo de viaje, tiempo de caminata, tiempo de espera, número de transferencias, edad, ingreso, entre otros.

2.5. Síntesis y conclusiones

En las primeras secciones se hace una revisión del potencial de los datos masivos de transporte público, señalando las ventajas y desventajas descritas en la literatura respecto a los datos obtenidos de encuestas tradicionales. Con esto en consideración, se revisa la bibliografía según dos enfoques: estudio de patrones y variabilidad y análisis de variables que caracterizan los viajes utilizando datos masivos. Ambas miradas se integran en el sentido de exponer de qué manera se han analizado dichas variables hasta ahora, para posteriormente aplicar sobre ellas procedimientos en la búsqueda de conductas comunes o patrones, con el fin de adquirir mayor conocimiento de la forma de viajar de los usuarios.

Aunque las tarjetas inteligentes de pago llevan algunos años en uso, los requerimientos computacionales y desarrollo de metodologías para extraer información han demorado el aná-

lisis. Es probable que por esto existan muy pocos trabajos de estudio de patrones utilizando estos datos, y que aquellos que existen se avoquen a una o dos variables cuando más.

Finalmente, se incluye una sección dedicada a la fidelidad de usuarios al sistema de transporte. La permanencia de los usuarios en el sistema es relevante porque delimita el alcance de estudiar variaciones de conducta en el tiempo. Además, la permanencia de una tarjeta en el tiempo y la calidad de servicio percibida, podrían predecir el cambio de un usuario de transporte público a otro modo.

Los trabajos presentados sugieren variadas ideas para enfrentar la búsqueda de patrones de viaje para los datos de Santiago. Entre ellas se encuentran observar como se distribuyen las tarjetas según alguna medida de frecuencia de uso. Otro punto es observar diferencias entre tipos de tarjetas. En este caso, existen solo dos tipos: estudiante y adulto. También se observan indicios de posibles diferencias entre usuarios que habitan en diferentes zonas de la ciudad, como muestran Utsunomiya et al. (2006) y Olguín et al. (2009). Por último, la señal de alta rotación de tarjetas en otros sistemas de transporte sugiere chequear este comportamiento y verificar su importancia con los datos disponibles.

El estudio y análisis de conductas de viaje y patrones de actividades desagregados por zona de residencia son el núcleo de este trabajo.

Capítulo 3

Análisis de las bases de datos

El presente capítulo explica brevemente las características de funcionamiento del sistema de transporte público de Santiago, llamado *Transantiago*, sistema de transporte del cual provienen los datos. A continuación se describen dos procesos aplicados previamente a los datos para obtener la información requerida para este trabajo: método de estimación de bajada y método de asignación de propósito de viaje. El primero estima el lugar de bajada de una transacción con el principal objetivo de obtener el destino de los viajes; el segundo asigna un propósito de viaje en base al tipo de tarjeta y una serie de reglas.

Luego se detalla la información contenida en cinco bases de datos correspondientes a cinco semanas de transacciones no consecutivas. Se procede a analizar la información considerando diferencias entre tipos de tarjetas, frecuencia de uso y nivel de permanencia de una tarjeta entre cortes temporales, como se sugiere en el capítulo dos. No se realizan análisis por zona de residencia debido a que esta información no está disponible en las bases de datos.

Se espera que estas variables muestren diferencias sistemáticas que entreguen indicios de posibles patrones de comportamiento.

3.1. Descripción del sistema de transporte y procesamiento de datos

Para los casos en que sólo hay validación de la tarjeta en la subida, la literatura describe distintos métodos de estimación de bajada, con diferencias propias de la diversidad de sistemas de transporte. Entre ellos se cuentan los desarrollados por: Hoffman y O'Mahony (2005), Trépanier et al. (2007), Zhao et al. (2007), entre otros. Aquí se explica en detalle el método desarrollado por Munizaga y Palma (2012), que es el utilizado para estimar el lugar y hora de bajada de las transacciones de las bases de datos que se utilizarán posteriormente. Antes de explicar el método, se describe brevemente la operación del sistema y algunos conceptos.

3.1.1. Operación

El sistema de transporte público de Santiago, Transantiago, opera con un sistema electrónico de recolección de tarifa. Para esto se utilizan tarjetas inteligentes, con las cuales los

usuarios pueden pagar la tarifa acercando la tarjeta a un dispositivo llamado validador, que a través de una señal electromagnética detecta la tarjeta próxima y determina el monto a pagar. La transacción de pago almacena información que puede utilizarse para estudiar el sistema de transporte.

A diferencia de otros sistemas de transporte, se requiere acercar la tarjeta al validador cada vez que se sube al vehículo, pero no al momento de bajar. Por tanto, no se almacena información de la bajada. Para obtener información del lugar de bajada de cada transacción se utiliza un método que requiere los datos almacenados de transacciones y de las emisiones de los GPS que se encuentran a bordo de los buses.

3.1.2. Viaje y etapa

Los sistemas de transporte público integrados tarifariamente se caracterizan por permitir que los usuarios utilicen más de un medio de transporte para realizar un mismo viaje, pagando una tarifa que crece menos que proporcionalmente con el número de medios de transporte utilizados en el viaje.

En este sentido, se entiende como etapa de viaje al intervalo comprendido entre la subida y la bajada por parte de un usuario a un vehículo de transporte. Por tanto, un viaje desde un origen a un destino, con un propósito determinado, puede estar comprendido por una o más etapas y los intervalos entre estas etapas. Como los trasbordos entre líneas dentro del metro no requieren validación, para este modo se considera que una etapa abarca la utilización de una o más líneas de manera consecutiva, mientras no se realice otra validación.

Con la base de datos de transacciones, la base de datos de emisiones GPS y el método de estimación de bajada en condiciones ideales, se cuenta con información de la etapa en cuanto a: número identificador de la tarjeta (ID), coordenadas geográficas de subida y bajada, tiempo de subida y bajada, modo de transporte, recorrido, costo de la transacción, entre otros.

3.1.3. Características relevantes

Hay algunas características importantes de mencionar respecto al sistema de transporte y recolección de tarifa. El pago al interior de los buses sólo puede efectuarse con la tarjeta inteligente. En el caso del metro, existe la opción de pagar con la tarjeta o comprar boletos que sólo sirven en este modo. Para acceder al sistema de integración tarifaria y utilizar ambos modos combinadamente sin duplicar el valor del pasaje, se debe utilizar la tarjeta inteligente. Este hecho redundante en una alta penetración en el uso de las tarjetas inteligentes en la población, con un 97 % de los viajes realizados pagando por este medio (Beltrán et al., 2012). Debido a que la tarjeta no es personalizada, una misma tarjeta podría ser prestada y usada por varias personas en distintos momentos.

Otro punto importante es la evasión. Este concepto corresponde a la utilización de modos de transporte público sin realizar validación o pago. La evasión puede observarse en una o

varias etapas del viaje, introduciendo sesgos en cuanto al lugar de origen o destino del viaje, entre otros problemas. La información disponible actualmente no permite determinar los niveles de evasión desagregadamente ni corregir los datos.

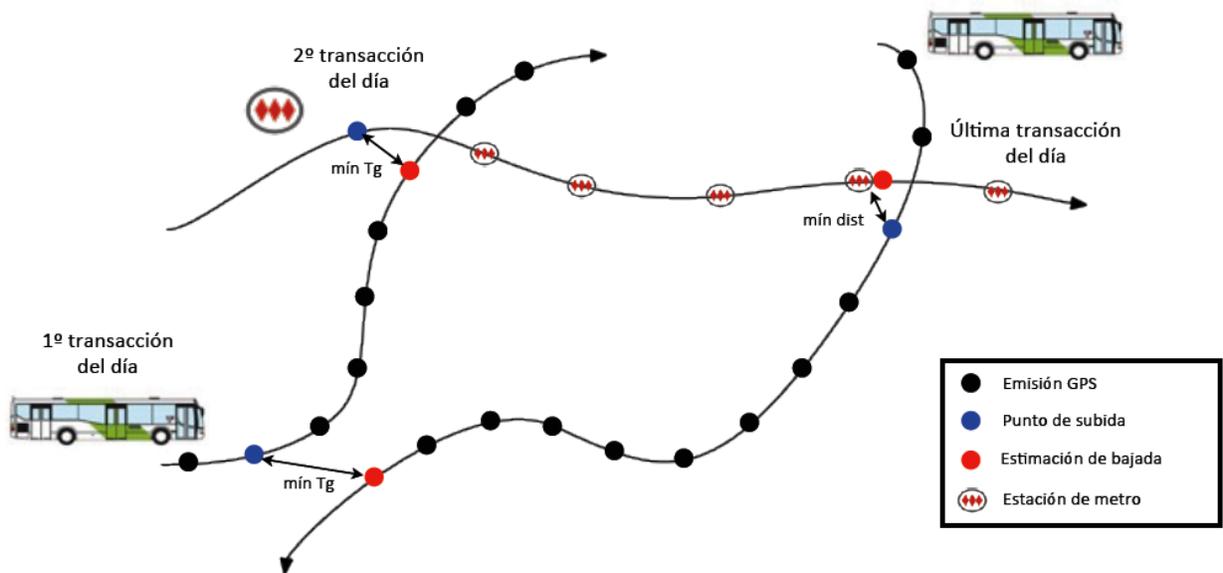
La información disponible para aplicar el método de estimación de bajada corresponde a la base de datos de transacciones con información de cada validación realizada y a la base de datos de emisiones GPS con información regular de la posición de los buses en funcionamiento.

3.1.4. Método de estimación de bajada

La metodología desarrollada por Munizaga y Palma (2012) para estimar la bajada de una transacción en modo bus se basa en observar el lugar de subida de la siguiente transacción y estimar cuál es el posible punto de bajada más cercano a esa transacción.

La metodología se ilustra en la Figura ?? y a continuación se describe el procedimiento.

Figura 3.1: Metodología de Estimación de Bajada.



Supongamos que para una misma tarjeta se tiene dos transacciones consecutivas j y $j + 1$, siendo la primera en bus (j). Como se conoce el servicio asociado a esa transacción, para estimar la bajada de la transacción j se analiza la trayectoria que realiza el individuo a bordo del bus y se identifica el paradero dentro de esta trayectoria, cuyo tiempo generalizado respecto del punto en que se realiza la siguiente transacción $j + 1$, es mínimo. Para esto se considera un umbral de distancia caminable d_{maxw} . El punto que minimiza la función de tiempo generalizado es el punto estimado como bajada de la transacción j . Este punto de bajada queda especificado por sus coordenadas geográficas (x_{b_j}, y_{b_j}) y un tiempo de bajada t_{b_j} asociado al momento en que el vehículo transitaba por el punto señalado como bajada.

Para determinar el destino de la última transacción de un día, se observa el punto de origen de la primera transacción del mismo día o bien la primera del día siguiente.

Para el caso del modo metro, la metodología es similar, pero se consideran como puntos factibles de bajada las estaciones de la red de metro en vez de los paraderos. En lugar de minimizar la función de tiempo generalizado para encontrar el punto de bajada, se busca la estación de metro que se encuentra a menor distancia del paradero de subida de la transacción siguiente, lo que se muestra en la Figura ???. Como los trenes de metro no poseen GPS, el tiempo de bajada se obtiene considerando el tiempo de subida y la suma de los tiempos de viaje entre estaciones. En caso de existir más de una ruta para realizar el mismo viaje en metro, se determina la ruta que escogió el usuario mediante el algoritmo de rutas mínimas Dijkstra y se asume que se utiliza la ruta de menor costo.

El sistema además cuenta con paraderos cerrados, denominados zonas paga, a los que se accede validando la tarjeta inteligente al ingresar a la parada, por lo cual no se valida tarjeta a bordo del bus. Esto redundante en que no se conoce con certeza el vehículo abordado, y por tanto se desconoce el servicio y ruta que realizó el usuario. Sin embargo, se asume que el bus abordado corresponde a alguna de las líneas que se detienen en esa zona paga. Antes de estimar las coordenadas y tiempo de bajada, se identifica las líneas comunes (Chriqui y Robillard, 1975). El procedimiento para identificar las líneas comunes consiste en determinar los recorridos asociados a la zona paga que transitan en alguna parte de la ruta cerca de la siguiente transacción $j + 1$ y buscar la estrategia óptima; es decir, el conjunto de líneas que minimiza la esperanza del tiempo generalizado a la posición de la siguiente transacción. Con la información de qué servicio podría haber abordado, se supone que el usuario subió al primer bus observado en el paradero con posterioridad a la hora en que se realizó la transacción, de entre los servicios que son líneas comunes.

Silva (2012) utiliza información de encuestas origen-destino para validar el método, encontrando que la estimación de bajada y destino de actividades es correcta en el 80% de los casos.

3.1.5. Método de estimación de propósito de viaje

Con el método de estimación de bajada aplicado a las transacciones se conocen los horarios de viaje de usuarios de transporte público. El método propuesto para establecer el propósito de estos viajes fue desarrollado por Devillaine (2012). Se trata de una metodología para distinguir viajes y actividades que se basa en condiciones espaciales, temporales y patrones de conducta. Algunas de las condiciones para establecer que una ventana de tiempo corresponde a una actividad que separa dos viajes son: ventanas de tiempo de trasbordo superiores a 30 minutos, utilización del mismo servicio-sentido consecutivamente y escoger rutas excesivamente largas en comparación a la distancia euclidiana entre origen y destino.

Una segunda parte de la metodología asigna propósito a las actividades identificadas. Los posibles propósitos son: *Trabajo, Estudio, Hogar y Otros*. *Trabajo y Estudio* corresponden a

las principales actividades del día según tipo de tarjeta, con una duración de al menos 2:00 horas en día laboral y de al menos 5:00 horas en fines de semana. El propósito *Hogar* se asigna en el lapso de tiempo que va desde la última transacción de un día hasta la primera transacción del día siguiente. El propósito *Otros* se asigna a cualquier actividad que no esté comprendida entre las anteriores, es decir, actividades que se realicen fuera del hogar con una duración inferior a las 2:00 horas en día laboral e inferior a 5:00 horas en fines de semana.

3.2. Datos disponibles y procesamiento de la información

Cada vez que un usuario de transporte público acerca su tarjeta inteligente de pago (o tarjeta *bip!*) a un validador, el detalle de la transacción es almacenado en forma automática, independiente de si la transacción se realiza a bordo de un bus, en una zona paga o en una estación de metro. Cada transacción o grupo de transacciones pasa por varias rutinas, entre las que se encuentran la metodologías de estimación de bajada y la de asignación de propósito, descritas en la sección anterior. A continuación se describen las variables más relevantes para este trabajo, disponibles por cada transacción:

- ID: número identificador asociado a cada tarjeta *bip!*.
- Pago: valor monetario pagado por el usuario para utilizar el medio de transporte, que se descuenta automáticamente de la tarjeta.
- Escolar: variable $\{0,1\}$ que define si la transacción fue realizada con una tarjeta tipo escolar o adulto.
- Número de etapa: entero correlativo asignado según ordenamiento temporal de cada transacción.
- Número de viaje: entero correlativo asignado a una o varias etapas que conforman un mismo viaje, de acuerdo a una serie de reglas temporales y espaciales. Variable estimada.
- Tipo de operador: define si la transacción se realizó en metro o en bus, señalando si es tipo troncal o alimentador.
- Servicio: nombre de la estación de metro o servicio del bus donde se realizó la transacción, obtenido utilizando la patente del vehículo y tablas de asignación de servicio proporcionadas por la Coordinación de Transantiago.
- Sentido (ida o retorno): sentido del bus donde se realizó la transacción, obtenido en base a la patente del vehículo de las tablas de asignación de servicio proporcionadas por Transantiago.
- Tiempo de subida: fecha y hora en que se produce la validación de la tarjeta.
- Coordenadas de subida (x,y): coordenadas UTM donde se realiza la transacción.

- Lugar de subida: código del paradero, zona paga o estación de metro donde se aborda el vehículo, obtenido de cruzar información de las coordenadas de subida con información de distribución de paraderos y recorridos proporcionada por Transantiago.
- Zona de subida: zona de la ciudad donde se encuentra el lugar de subida. Se utiliza la zonificación EOD 777 (ver Sectra, 2001).
- Comuna de subida: comuna donde se encuentra el lugar de subida.
- Tiempo de bajada: fecha y hora estimada de bajada del vehículo por cada transacción. Esta estimación se obtiene a partir del procedimiento explicado en la sección 3.1.4.
- Coordenadas de bajada (x,y): Corresponde a las coordenadas UTM donde se estima que descendió el individuo. Esta estimación se obtiene a partir del procedimiento explicado en la sección 3.1.4.
- Lugar de bajada: código del paradero, zona paga o estación de metro más cercana a las coordenadas de bajada, obtenido de cruzar información de las coordenadas de bajada con información de distribución de paraderos y recorridos proporcionada por Transantiago.
- Zona de bajada: zona de la ciudad donde se encuentra el lugar de bajada estimado. Se utiliza la zonificación EOD 777 (ver Sectra, 2001).
- Comuna de bajada: comuna donde se encuentra el lugar de bajada estimado.
- Tiempo de viaje: lapso de tiempo que va desde el instante de subida al instante de bajada estimado.
- Tiempo de trasbordo: lapso de tiempo que va desde el instante de bajada estimado de una transacción el instante de subida de la transacción siguiente, si ésta pertenece al mismo viaje.
- Propósito: tipo de actividad para la cual se realiza el viaje. Se observa sólo en la última transacción de cada viaje. Variable estimada.

La información disponible corresponde a cinco bases de datos (BD) que contienen el detalle de las transacciones realizadas durante cinco semanas (siete días) no consecutivas. El detalle del lapso de tiempo de cada base de datos o corte temporal se encuentra a continuación:

Tabla 3.1: Información de las bases de datos disponibles.

Nombre BD	Día inicio	Día Fin	Año
Sep-08	01 septiembre	07 septiembre	2008
Ago-09	17 agosto	23 agosto	2009
Jun-10	31 mayo	06 junio	2010
Abr-11	11 abril	17 abril	2011
Abr-12	16 abril	22 abril	2012

Como no es posible ligar la información de transacciones de los boletos de viaje a usuarios específicos o a una cadena de viajes, sólo se utiliza la información de transacciones de la tarjeta inteligente. Por otra parte, las transacciones son registradas cada vez que un usuario acerca su tarjeta inteligente a un validador de pago, por lo que no incluyen información de evasión.

Debido a los requerimientos computacionales y altos costos de tiempo de procesamiento, no todas las bases de datos contienen la misma información. Existen tres niveles de información en las bases de datos utilizadas en este trabajo. El primer nivel corresponde al preproceso que incorpora información como coordenadas geográficas de la transacción, modo de transporte, servicio, operador, paradero de subida, entre otros datos necesarios para aplicar el proceso de estimación de bajada. El segundo nivel es el proceso que incorpora la estimación de hora y lugar de bajada de cada transacción, además de los tiempos de etapa, trasbordo, caminata, espera, distancia de viaje en ruta, entre otros. El último nivel de proceso asigna propósito a los viajes.

La siguiente tabla muestra el nivel en que se encuentra cada base de datos:

Tabla 3.2: Procesos aplicados a las bases de datos.

Nombre BD	Preproceso	Estimación de Bajada	Propósito de viaje
Sep-08	Si	No	No
Ago-09	Si	No	No
Jun-10	Si	Si	No
Abr-11	Si	Si	Si
Abr-12	Si	Si	Si

3.3. Análisis preliminar de las bases de datos de transacciones

A continuación se realiza un análisis preliminar de las bases de datos correspondientes a cinco cortes temporales, cada uno de siete días consecutivos. El análisis se realiza separadamente para tarjetas tipo adulto y tarjetas tipo estudiante, por tener comportamientos muy dispares entre sí. Se incluyen características básicas, distribución temporal de las transacciones y frecuencia en el uso de la tarjeta.

3.3.1. Análisis de tarjetas tipo adulto

La Tabla ?? muestra el número de tarjetas tipo adulto con distinto número identificador y el total de transacciones realizadas con éstas en cada corte temporal. En una semana de viajes se observa sobre 2,5 millones de tarjetas distintas. Por su parte, las transacciones realizadas con este tipo de tarjeta en una semana bordean los 25 millones, lo que promedia diez transacciones por tarjeta por semana. Un descenso importante en el número de transacciones por tarjeta se observa en los dos últimos cortes temporales. Entre las posibles causas para este fenómeno está la creación y modificación de recorridos que reducen el número de

transacciones por viaje, descenso en la demanda debido sustitución del transporte público por modos alternativos y un aumento en la evasión que conlleva menor cantidad de transacciones registradas.

Tabla 3.3: Información de transacciones para tarjeta tipo adulto.

Tarjeta Adulto	Sep-08	Ago-09	Jun-10	Abr-11	Abr-12
N° de tarjetas	2.633.190	2.527.434	2.821.098	2.670.059	2.593.145
N° de transacciones	25.981.216	25.487.130	28.457.318	25.507.110	23.747.958
N° transacciones/N° tarjetas	9,87	10,08	10,09	9,55	9,16

Las siguientes figuras muestran la distribución temporal de las transacciones realizadas durante los siete días de la semana de cada corte temporal, en tramos de 15 minutos, para la tarjeta tipo adulto. Las transacciones fueron agrupadas según hora de inicio de la etapa.

Figura 3.2: Distribución de transacciones sep-08, tarjetas tipo adulto.

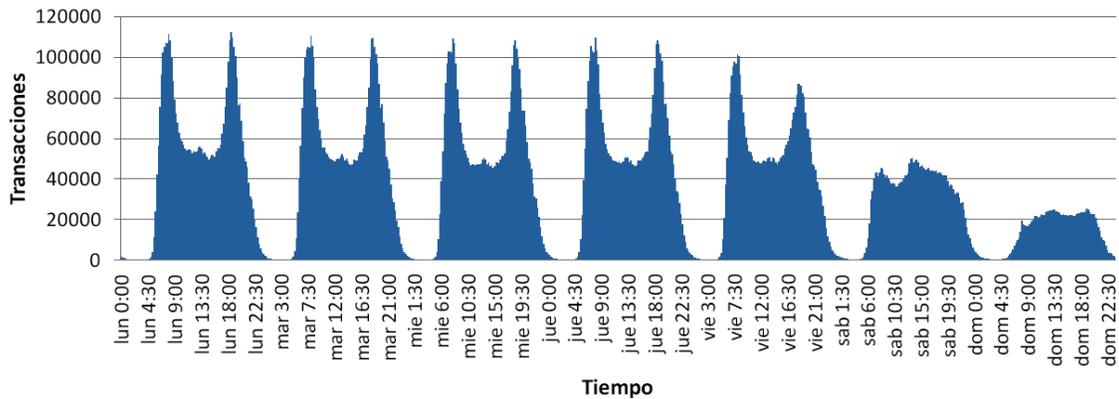


Figura 3.3: Distribución de transacciones ago-09, tarjetas tipo adulto.

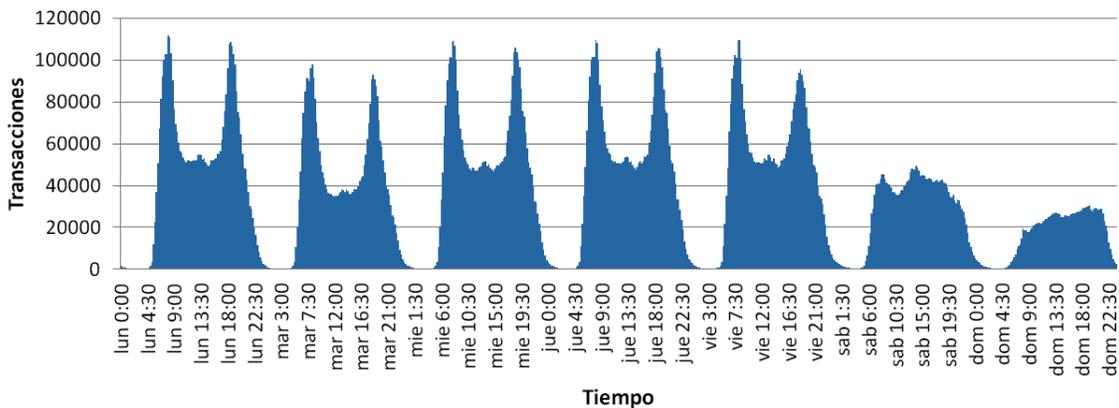


Figura 3.4: Distribución de transacciones jun-10, tarjetas tipo adulto.

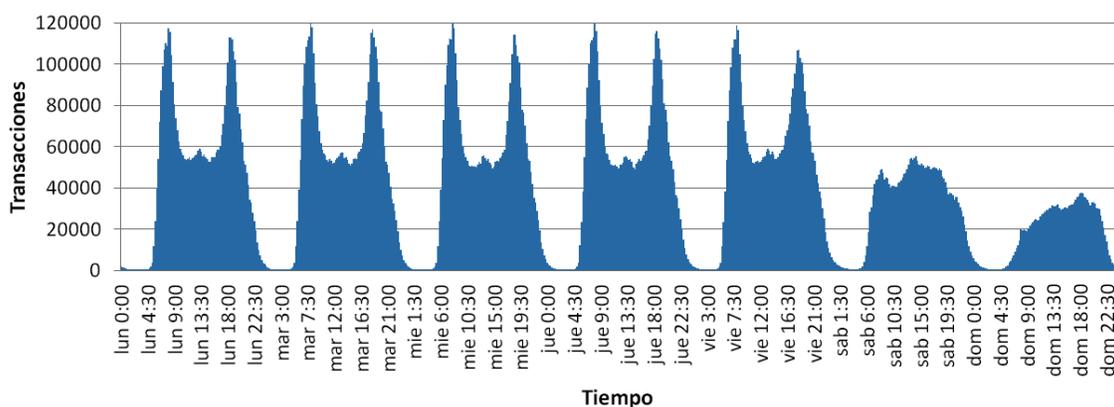


Figura 3.5: Distribución de transacciones abr-11, tarjetas tipo adulto.

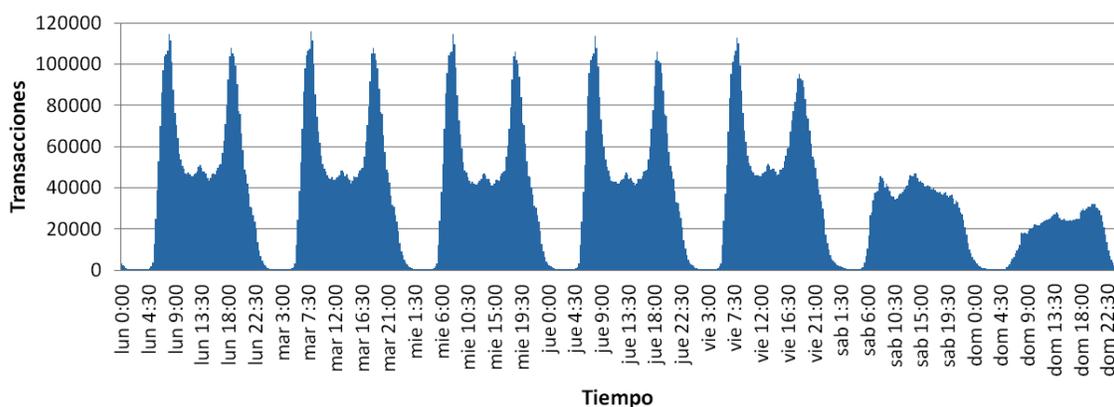
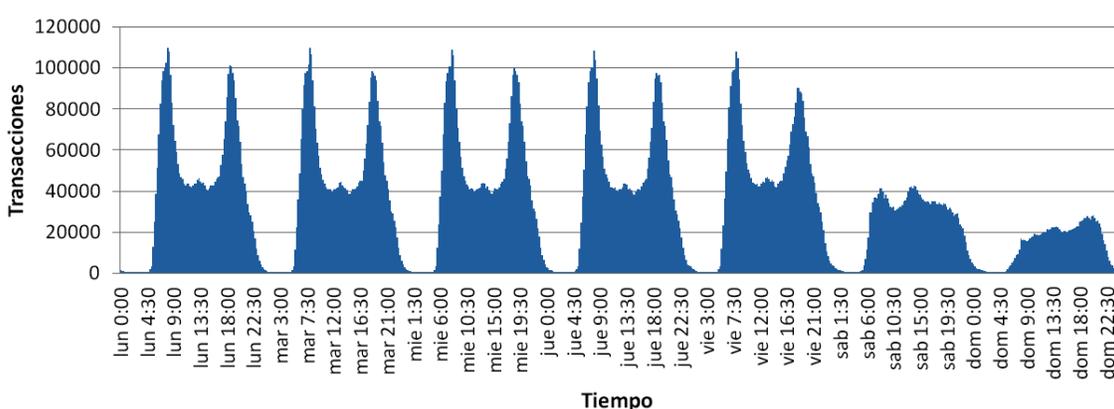


Figura 3.6: Distribución de transacciones abr-12, tarjetas tipo adulto.



La mayoría de los cortes temporales muestra un comportamiento similar en el número y distribución de transacciones de lunes a jueves. La excepción es agosto-2009, donde el día martes se observa un descenso significativo en el número de transacciones, coincidiendo con un día de lluvia. Aunque por muy poco, la punta mañana es superior a la punta tarde en can-

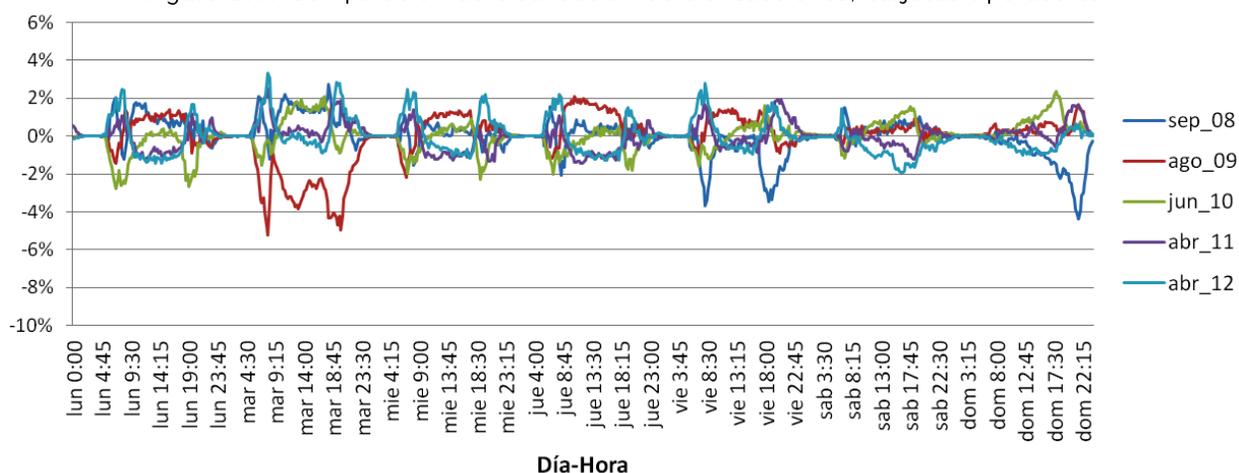
tividad de transacciones acumuladas en los 15 minutos más cargados. Los máximos ocurren entre 8:00 y 8:15 hrs y entre 18:30 y 18:45 hrs, respectivamente. La punta medio día tiene su máximo entre las 13:15 y las 13.30 hrs. En todos los cortes temporales se aprecia diferencias en la punta tarde del día viernes, la cual es menos pronunciada y más dispersa respecto a los otros días laborales. La distribución de transacciones en los días de fin de semana es bastante estable en el tiempo, con algunas diferencias en el día domingo.

Por otra parte, el horario con menos transacciones ocurre en alguno de los dos lapsos temporales de 15 minutos que se encuentran entre las 3:30 y las 4:00 am, dependiendo del corte temporal. Si se desea observar ciclos de un día de viaje para un usuario, lo más apropiado es considerar que el corte ocurre en este horario, puesto que el corte oficial de media noche no corresponde al mínimo de transacciones diarias.

Para detectar si algún corte temporal tiene diferencias marcadas en la forma de la distribución de las transacciones en algún periodo, se propone normalizar las distribuciones de transacciones dividiendo por el total de transacciones observadas en cada corte temporal. Luego, se promedian los 5 cortes temporales, lo que entrega un corte promedio de referencia. Posteriormente se compara cada semana de datos normalizada con el corte de referencia y se observa si existen diferencias importantes respecto al promedio.

Al realizar este proceso se obtiene la Figura ??, que permite comparar las distribuciones de transacciones. Se puede ver que la mayor diferencia entre distribuciones ocurre el día martes de la distribución de agosto-2009, que como ya se señaló fue un día de lluvia. Sus transacciones, en proporción, están muy por debajo del promedio. Otras alteraciones se observan los días viernes y domingo de septiembre-2008. Al mirar esta distribución se observa que efectivamente el día viernes la campana de viajes es inferior en proporción a los otros días. El día domingo, en cambio, no parece tener proporcionalmente menos datos, pero si es el único corte temporal que muestra una distribución diferente de transacciones en ese día, por lo que se aparta del promedio.

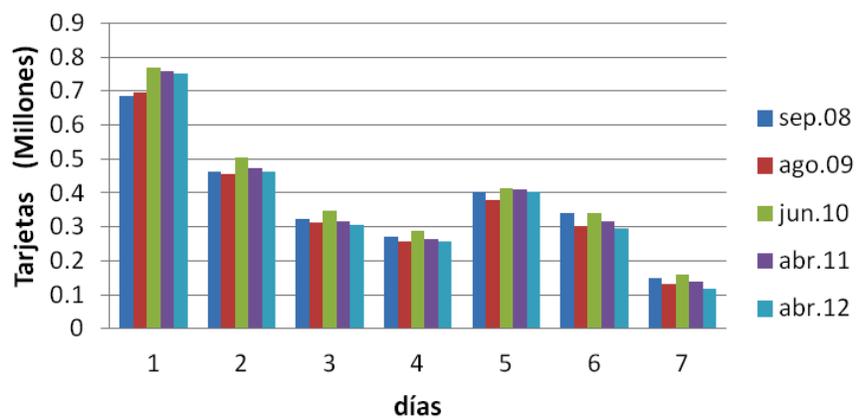
Figura 3.7: Comparación de distribución de transacciones, tarjetas tipo adulto.



Por otra parte, las distintas puntas de las distribuciones revelan comportamientos de viaje. La punta mañana se asocia típicamente al viaje con dirección al trabajo. Al medio día hay una pequeña punta de viajes de almuerzo u otras actividades. La punta tarde se asocia al viaje de retorno al hogar. Se plantea profundizar en el estudio de este patrón de comportamiento, examinando por separado el primer, intermedio y último viaje del día, para observar como distribuyen estos viajes y posibles cruces entre campanas.

Varios autores estudian el nivel de utilización de tarjetas inteligentes según el número de días de la semana que se usó para determinar la frecuencia de viaje y variabilidad espacial/temporal. A continuación se presenta una gráfica donde se muestra cómo distribuye el número de tarjetas según cuántos días de la semana fue utilizada:

Figura 3.8: Distribución de tarjetas tipo adulto según número de días de utilización en cada corte temporal.



Se observa que la mayor cantidad de tarjetas se utiliza uno o dos días de la semana. Las siguientes mayorías corresponden a usuarios que ocupan la tarjeta durante cinco o seis

días de la semana. Aunque se aprecian ciertas fluctuaciones leves entre cortes temporales, se distingue un comportamiento transversal a todos los cortes. Si se considera dos grupos de usuarios, frecuentes e infrecuentes, la separación entre ambos tipos de usuarios está dada por el mínimo local situado en 4 viajes a la semana. En este trabajo se considerará usuarios frecuentes a quienes viajan cuatro o más días de la semana y usuarios ocasionales a quienes viajan tres o menos días de la semana. Los usuarios frecuentes permitirán realizar un estudio más detallado de los patrones de comportamiento.

3.3.2. Análisis de tarjetas tipo estudiante

La Tabla ?? muestra el número de tarjetas tipo estudiante con distinto número identificador y el total de transacciones realizadas con éstas en cada corte temporal. En una semana de viajes se tiene más de 600 mil tarjetas diferentes. Por su parte, las transacciones van desde los 8,5 a poco más de 10 millones, lo que promedia entre 13 y 15 transacciones por tarjeta por semana. La tarjeta estudiante tiene una mayor utilización en promedio respecto de la tarjeta adulto, lo que concuerda con la obligatoriedad de la escolaridad básica y media, la prohibición de conducir para menores de edad y el pasaje rebajado en enseñanza media y superior.

Tabla 3.4: Información de transacciones para tarjeta tipo estudiante.

Tarjeta Estudiante	Sep-08	Ago-09	Jun-10	Abr-11	Abri-12
N° de tarjetas	646.613	693.234	645.750	651.080	695.319
N° de transacciones	8.532.715	9.705.433	9.661.631	9.855.979	10.249.881
N° transacciones/N° tarjetas	13,20	14,00	14,96	15,14	14,74

Las figuras ?? a ?? muestran la distribución temporal de las transacciones realizadas durante los siete días de la semana de cada corte temporal, en tramos de 15 minutos, para la tarjeta tipo estudiante. Las transacciones fueron agrupadas según hora de inicio de la etapa.

Figura 3.9: Distribución de transacciones sep-08, tarjetas tipo estudiante.

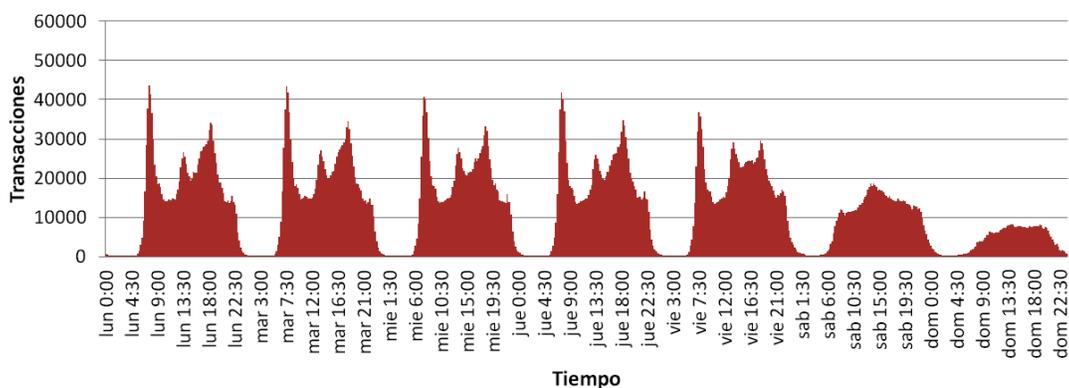


Figura 3.10: Distribución de transacciones ago-09, tarjetas tipo estudiante.

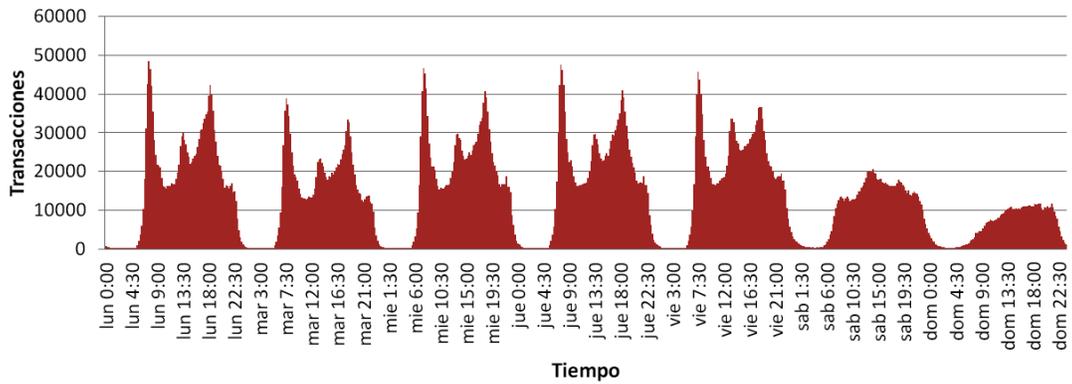


Figura 3.11: Distribución de transacciones jun-10, tarjetas tipo estudiante.

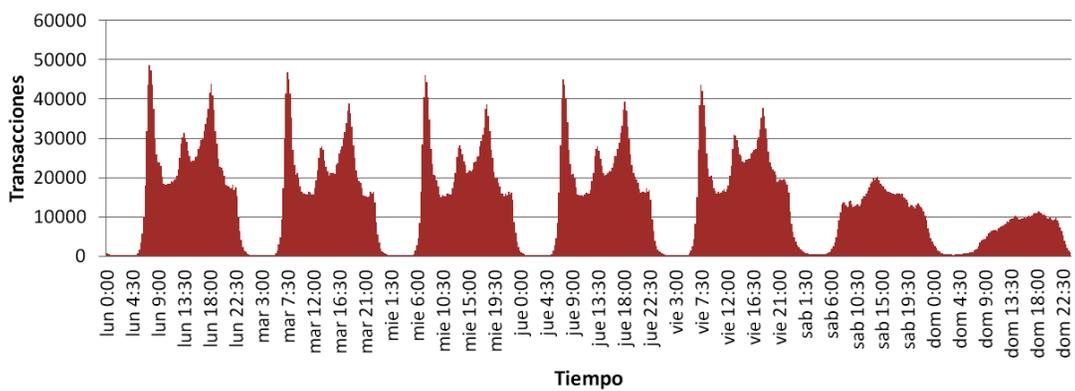


Figura 3.12: Distribución de transacciones abr-11, tarjetas tipo estudiante.

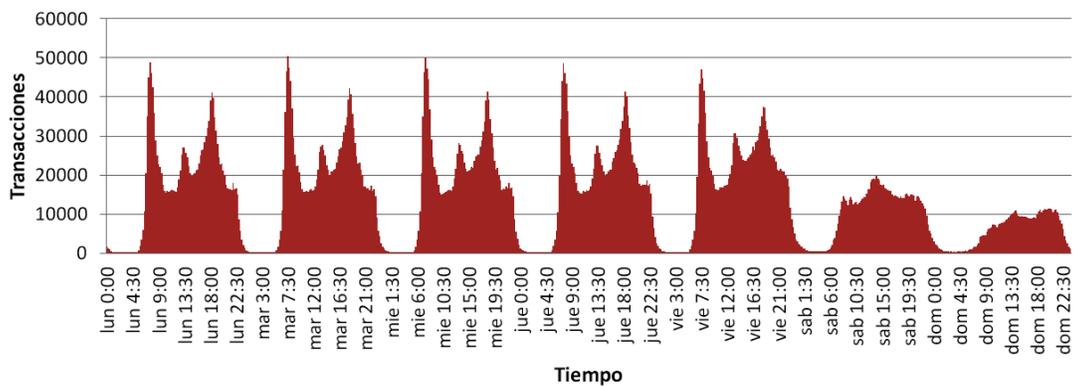
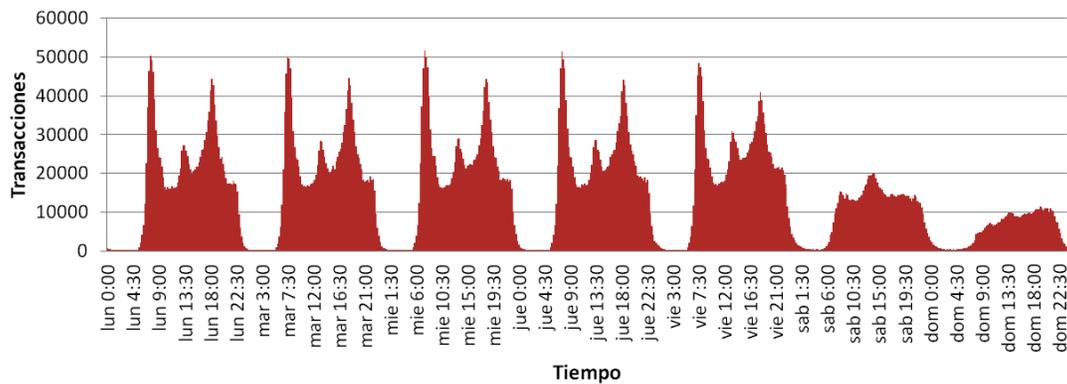


Figura 3.13: Distribución de transacciones abr-12, tarjetas tipo estudiante.



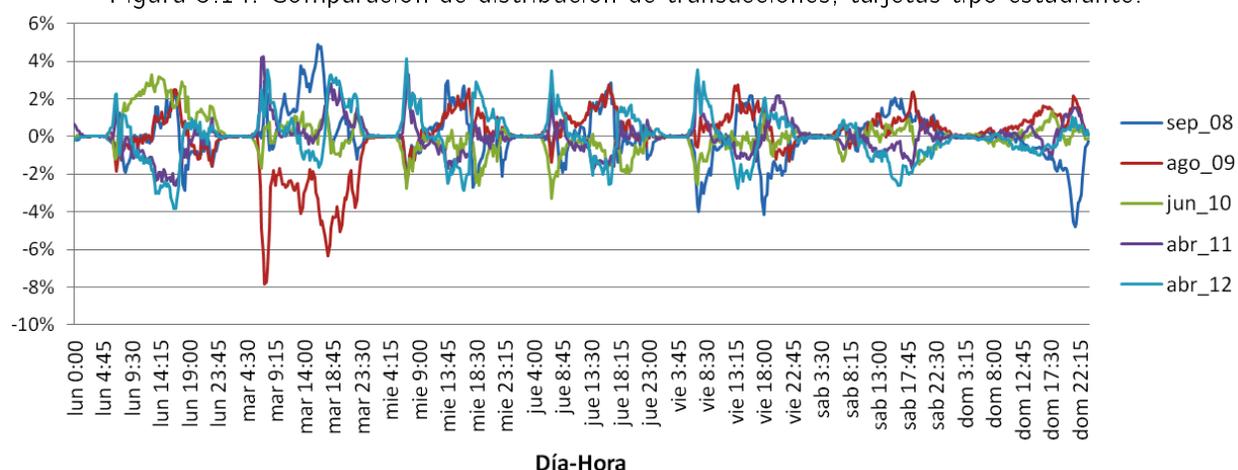
En general, la distribución de transacciones para tarjeta tipo estudiante es distinta a la observada en tarjetas tipo adulto. Esto puede explicarse por diferentes jornadas de estudio (horario diurno y vespertino) y por la combinación de actividades de estudio con trabajos de medio tiempo de este tipo de usuarios. La distribución para estudiantes no muestra simetría en torno al medio día como en el caso anterior. Se aprecia una pronunciada punta mañana, proporcionalmente mucho más fuerte que la observada en la tarjeta adulto. A continuación aparece una punta medio día también de mayor altura que en la tarjeta adulto, para culminar con una punta tarde mucho menos pronunciada y más extendida. Se observa un pequeño realce en la noche durante todos los días laborales, probablemente atribuible al viaje de retorno de quienes estudian en horario vespertino. Esta punta ocurre típicamente entre las 21:45 y las 22:15 hrs.

La mayoría de los cortes temporales muestra un comportamiento similar en el número y distribución de transacciones de lunes a jueves. La excepción es agosto-2009, donde el día martes se observa un descenso significativo en el número de transacciones, coincidiendo con un día de lluvia. En todos los cortes temporales se aprecia diferencias el día viernes: la punta medio día es más pronunciada y la campana asociada a la punta tarde comienza antes, es menos pronunciada y más dispersa. Los máximos en punta mañana y punta tarde ocurren entre 7:30 y 7:45 hrs y entre 18:15 y 18:30 hrs, respectivamente. La punta medio día tiene su máximo entre las 13:15 y las 13.45 hrs.

La distribución de transacciones en los días de fin de semana es bastante estable en el tiempo, con algunas disimilitudes el día domingo.

Para observar diferencias entre distribuciones se repite el procedimiento anterior de normalizar el número de transacciones de cada corte temporal en base al total de transacciones realizadas, promediar los cortes temporales y luego compararlos respecto a la distribución promedio.

Figura 3.14: Comparación de distribución de transacciones, tarjetas tipo estudiante.

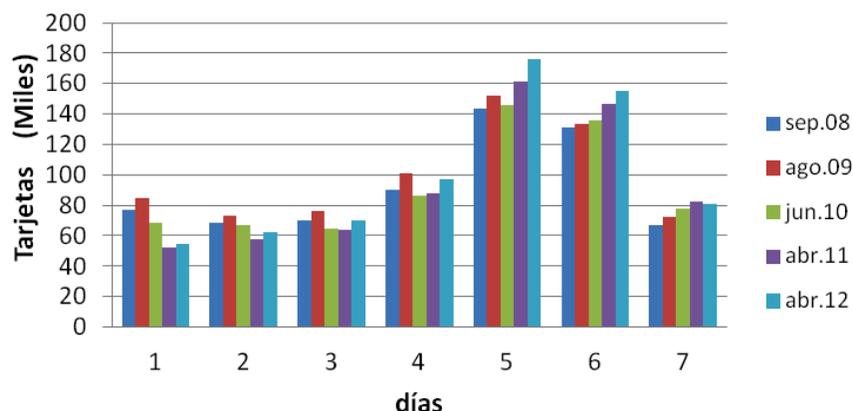


Se observa un resultado similar al de la tarjeta adulto, donde la mayor diferencia ocurre el día martes de agosto-2009 en que cayeron precipitaciones.

La diferencia en la forma de distribución de transacciones para tarjetas tipo estudiante plantea que las campanas de viaje son distintas a las de la tarjeta adulto en cuanto a horario y número de viajes diarios. Se propone profundizar en el estudio de este patrón de comportamiento, para observar como distribuyen estos viajes y posibles cruces entre campanas.

La distribución del número de tarjetas, según cantidad de días de la semana que es utilizada se muestra en la figura ???. Se puede ver que la distribución es completamente diferente para usuarios tipo estudiante respecto de lo observado en usuarios tipo adulto. En este caso, las primeras mayorías corresponden a tarjetas que viajan (de mayor a menor) cinco, seis o cuatro veces a la semana. Es decir, el usuario más común es el que utiliza su tarjeta frecuentemente. Lógicamente, esto responde a las características de la mayoría de los usuario de esta tarjeta que por su rango etario deben recibir educación obligatoriamente.

Figura 3.15: Distribución de tarjetas tipo estudiante según número de días de utilización en cada corte temporal.



3.4. Permanencia de la tarjeta

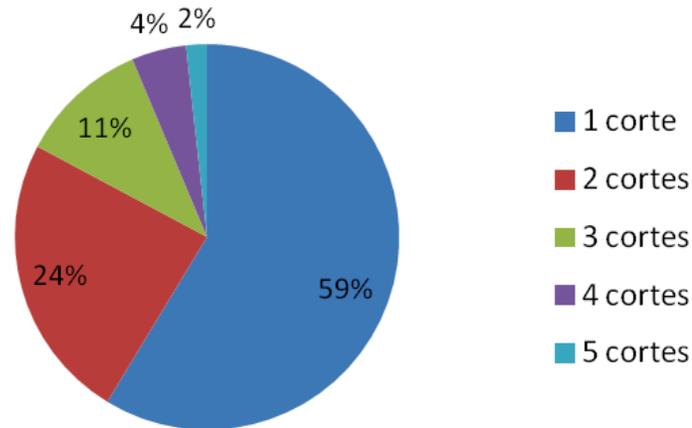
Como se vio en el capítulo 2, la literatura describe casos en que una parte importante de las tarjetas observadas son utilizadas durante breves lapsos de tiempo. Debido a que los datos de transacciones obtenidos para este trabajo no son temporalmente continuos, es imposible determinar la fecha exacta en que una tarjeta aparece y desaparece del sistema. Sin embargo, si este fenómeno se replica en las bases de datos disponibles, es importante determinar qué proporción de tarjetas aparece en uno o más cortes temporales.

Por lo anterior, se cuantifica la permanencia de las tarjetas en el tiempo, es decir, durante cuánto tiempo una tarjeta permanece en uso en término de cortes temporales. Este análisis no se realiza separadamente entre tarjetas tipo estudiante y adulto debido principalmente a una razón: las tarjetas tipo estudiante en un cierto corte temporal, pueden transformarse en tarjetas tipo adulto en un corte temporal posterior, generando una tercera categoría de análisis de tarjetas que cambian de tipo.

La figura ?? considera como 100 % la suma (sin repeticiones) del total de tarjetas (de cualquier tipo) con diferentes identificadores pertenecientes a al menos uno de los cinco cortes temporales. Para ejemplificar, si una tarjeta realiza transacciones en agosto-2009, abril-2011 y abril 2012, entonces esa tarjeta se cuenta una única vez y pertenece al grupo de tarjetas que aparecen en tres cortes temporales.

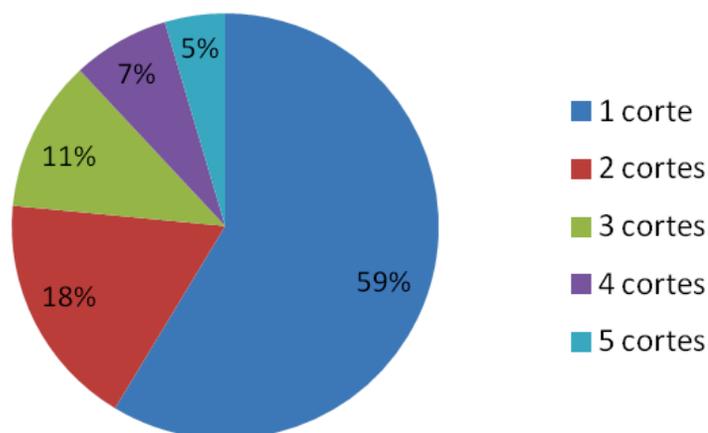
El total de tarjetas distintas observadas en los cinco cortes temporales es 9.948.849. El 59 % de las tarjetas aparece únicamente en uno de ellos, en contraste con el 1.75 % del total de tarjetas que aparecen en los cinco cortes temporales.

Figura 3.16: Distribución de tarjetas según número de cortes temporales en los que una tarjeta está presente.



Cabe señalar que el gráfico anterior considera que una tarjeta pertenece al grupo de n cortes temporales si ésta tiene transacciones estrictamente en n cortes temporales, sean o no consecutivos. Si una tarjeta aparece en n cortes temporales no consecutivos, como por ejemplo: agosto-2009, abril-2011 y abril-2012, se podría decir que la tarjeta esta activa durante cuatro cortes temporales, pese a que no se observan transacciones en la semana de junio-2010. Se llama *permanencia de una tarjeta* al número de cortes temporales transcurridos entre el primer y último corte en que una tarjeta registra transacciones, ambos inclusive, independiente de si existen registros en los cortes intermedios. Esta información se muestra en la figura ??:

Figura 3.17: Distribución de tarjetas según diferencia entre primer y último corte temporal en que una tarjeta está presente.



Como no es posible estimar el punto exacto de entrada y salida de una tarjeta del sistema con datos de cortes semanales, no es posible establecer la vida útil de la tarjeta. Sin embargo,

al contrastar los gráficos anteriores que exponen dos formas distintas de observar los mismos datos, sus diferencias permiten obtener una importante conclusión: hay un número de tarjetas que están presentes en dos o más cortes temporales y que viajan de manera tan infrecuente que pueden pasar desapercibidas en lapsos de tiempo de una semana. Un ejemplo de esto es que un 2 % las tarjetas aparece estrictamente en cinco cortes temporales (ver figura ??), pero en cambio un 5 % de las tarjetas utiliza su tarjeta en el mismo lapso de tiempo aunque no viaje en los cinco cortes temporales (ver figura ??). Este hecho establece que para estudiar los patrones de comportamiento de usuarios ocasionales se requiere datos de periodos de tiempo mayores a una semana de datos.

Es importante señalar que este análisis no diferencia entre usuarios por nivel utilización de la tarjeta. Si entre un corte temporal y el siguiente el número de transacciones aumenta o cae fuertemente para una misma tarjeta, el análisis no hace distinción entre los usuarios que aumentan, mantienen o reducen sus viajes. Solo se toma como variable usar o no la tarjeta. Se propone estudiar esta variable, es decir, si se observa alguna tendencia al aumento o disminución del uso de la tarjeta entre cortes temporales.

3.5. Síntesis y conclusiones

En este capítulo se presenta una breve descripción de las características del sistema de transporte del cual provienen los datos. Además se explica los procesos computacionales aplicados a los datos para obtener información que no se extrae directamente del registro de transacciones. Luego se expone la información contenida en las bases de datos, el número de bases de datos disponible y el detalle de los periodos de tiempo observados. Posteriormente se describe brevemente el orden en que se aplicaron los distintos procesos y en qué etapa de proceso se encuentra cada una de las bases de datos.

Posteriormente se procede a analizar las bases de datos entre cortes temporales, por tipo de tarjeta, en los siguientes puntos: cantidad de tarjetas, cantidad de transacciones, distribución temporal de las transacciones y frecuencia de uso de la tarjeta.

Se concluye el capítulo con la clasificación de las tarjetas según el tiempo permanencia en el sistema de transporte en términos del número de cortes temporales en que se observa al menos una validación.

De los resultados de este capítulo se confirma la necesidad de realizar análisis separados por tipo de tarjeta, debido a las evidentes diferencias de comportamiento entre tarjetas tipo estudiante y adulto. También se plantea la importancia observar el comportamiento según número de viajes diarios, considerando que la tarjeta adulto muestra una distribución de viajes bastante conocida (típicamente un viaje con origen en el hogar durante la mañana, un viaje con destino el hogar en punta tarde y un viaje al medio día en algunos casos) a diferencia de la tarjeta de estudiante que presenta una distribución menos explorada.

Por otra parte, se concluye que en el caso adulto el usuario más común es aquel que viaja de manera infrecuente (utiliza la tarjeta tres o menos días de la semana) a diferencia del caso de los estudiantes que en su mayoría son usuarios frecuentes (utilizan la tarjeta cuatro o más días a la semana). Es importante notar que en el caso de la tarjeta estudiante parece existir una única distribución de usuarios centrada en el punto donde la tarjeta tiene alta frecuencia de uso, mientras que para la tarjeta adulto se observan dos campanas, cada una centrada en alta y baja frecuencia de uso, lo que indicaría la existencia de dos tipos de usuarios desde el punto de vista de la frecuencia.

Se concluye que para estudiar comportamiento de usuarios infrecuentes se requiere datos en lapsos de tiempo mayores a siete días, puesto que muchas de estas tarjetas presentan transacciones separadas por periodos de tiempos superiores a una semana. Si suponemos que los usuarios que viajan de forma infrecuente son responsables de parte importante de la demanda variable por transporte, se vuelve crucial cuantificarlos y estudiar su comportamiento para efectos de planificación de la oferta.

Finalmente, se propone observar los datos en busca de tendencias de aumento o reducción del uso de la tarjeta en el tiempo, puesto que el análisis de permanencia no distingue entre estos casos. Este análisis tienen sentido en la búsqueda de tendencias de alza o descenso de uso de transporte público.

A partir de estas conclusiones se proponen y construyen las metodologías que permitan estudiar con mayor profundidad los fenómenos encontrados.

Capítulo 4

Método de estimación de zona de residencia

Una de las conclusiones del análisis descrito en el capítulo 2 es que existen diferencias sustantivas en calidad de servicio al comparar zonas de una misma ciudad. Se postula que no solo existen diferencias de calidad de servicio, sino también en los patrones de viaje, y que estas diferencias se aprecian según zona de residencia. En el presente capítulo se propone la principal contribución de este trabajo, una metodología para identificar en qué zona de la ciudad habita el usuario de una determinada tarjeta. Este dato es un primer paso en la incorporación de información socioeconómica a la base de transacciones, lo que permite ampliar el espectro de análisis factibles de desarrollar a partir de las transacciones *bip!*. Es importante señalar que luego de implementar el método no hay costos asociados a obtener información de la zona de residencia de las tarjetas, lo que es sustantivamente más barato que el costo de ocupar la herramienta alternativa (encuestas) para obtener el lugar de residencia para esta cantidad de datos.

Para comenzar, se presenta una descripción de la zonificación que será utilizada para realizar la estimación y luego se describe el método propuesto para identificar la zona de residencia. A continuación se muestran los resultados y lo que se denomina *validación exploratoria* del método.

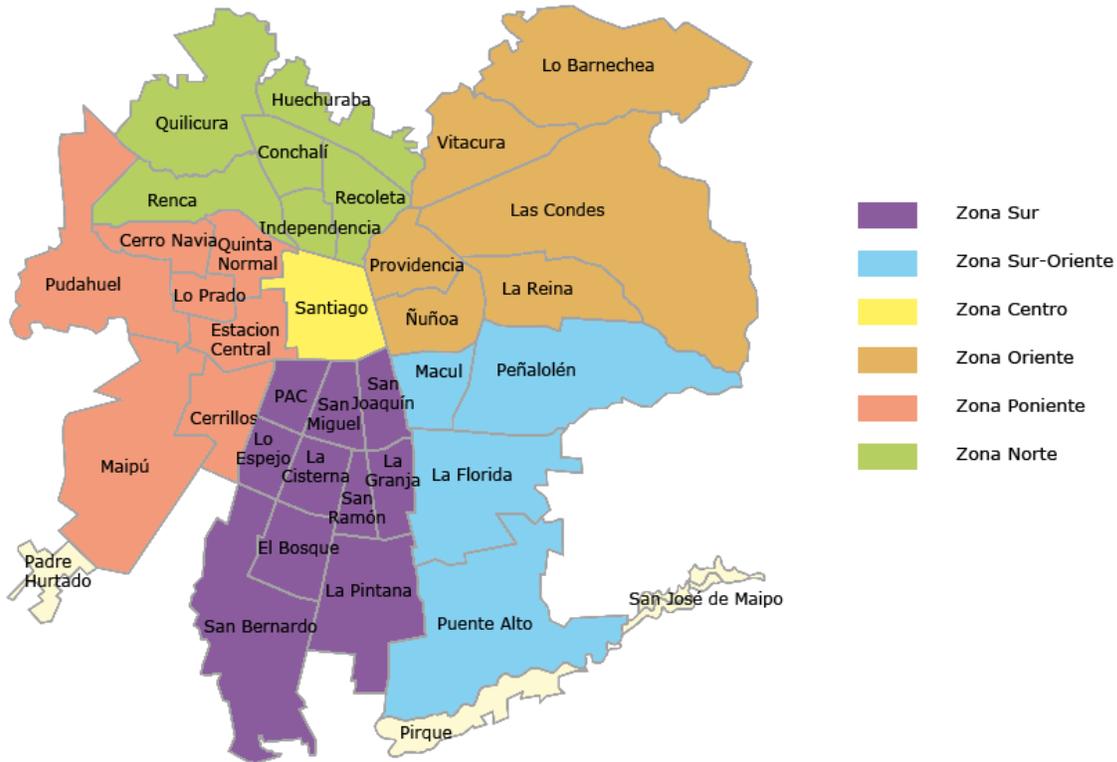
4.1. Zonificación

El sistema de transporte público *Transantiago* está presente en 34 de las 37 comunas que componen el Gran Santiago. Una de las características de Santiago es la marcada segregación espacial de sus habitantes según ingreso, como lo muestran Jara-Díaz et al. (2011). En su trabajo agregan la ciudad de Santiago en seis grandes zonas, utilizando la misma zonificación de la encuesta Origen-Destino del 2001: Centro, Norte, Sur, Poniente, Oriente y Suroriente; encuentran diferencias de más de 350 % en el ingreso promedio per cápita entre dos zonas de la ciudad. También se observan grandes diferencias en el número de vehículos por persona y la tasa de utilización de transporte público. Basándose en esta información y en lo expuesto en los dos capítulos anteriores, se propone conservar la zonificación sugerida por Jara-Díaz

et al. (2011) en busca de diferencias en los patrones de viaje entre zonas de la ciudad.

La siguiente imagen muestra las comunas del Gran Santiago agrupadas por colores según la zonificación señalada:

Figura 4.1: Agregación de Santiago en seis zonas.



4.2. Estimación de zona de residencia

4.2.1. Metodología

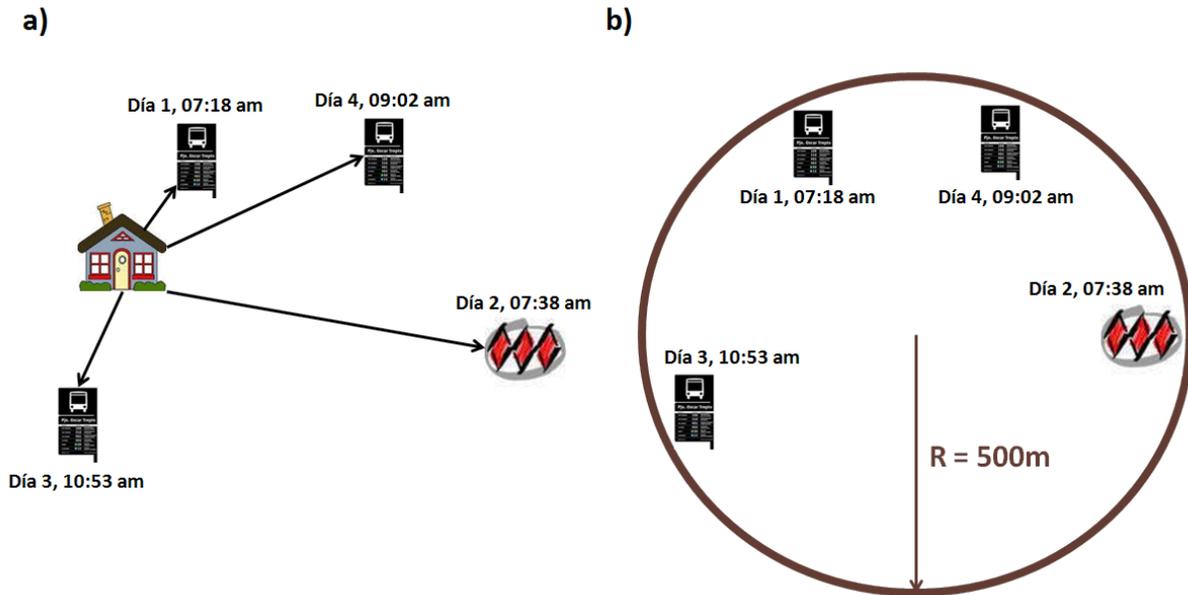
Se plantea estudiar diferencias en los patrones de comportamiento de usuarios según zona de residencia, considerando la zonificación expuesta. Para ello es relevante conocer la zona de residencia de los usuarios de Transporte Público. Como esta información no está disponible en las bases de datos de transacciones, se propone una metodología para establecer a qué zona pertenece cada individuo. Es importante señalar que aunque esta zonificación es bastante agregada y el método permite obtener información más precisa del lugar de residencia, no se presentarán resultados para tamaños de zona más pequeños. Esto porque se espera validar el método con una muestra estadísticamente significativa como es la que se obtendrá de la próxima encuesta Origen-Destino y porque no existe suficiente información en zonificaciones más desagregadas con las que se puedan comparar los resultados expuestos en el capítulo 5.

Esta metodología se basa en el supuesto de que el principal modo de transporte de un usuario frecuente es el transporte público. Por tanto, se espera que la mayor parte de sus viajes de realicen en este medio y que la primera transacción de cada día sea cercana al lugar de residencia del individuo. Si el conjunto de las primeras transacciones de cada día tienen origen en un mismo entorno cercano, se considera que el lugar de residencia de la tarjeta se encuentra en el entorno de este conjunto de transacciones. En caso de que la distancia entre las transacciones no sea caminable, el método no realiza estimación. La metodología propuesta se describe a continuación:

- Se selecciona a aquellas tarjetas con viajes observados entre las 4:00 am y 12:00 am durante al menos cuatro de los siete días de la semana. Esta condición se asocia a ser usuario frecuente, como se describe en el capítulo de Análisis de datos.
- Para cada tarjeta que cumple la condición anterior, se selecciona la primera transacción de cada día. En consecuencia, para cada tarjeta se tiene al menos cuatro (y hasta siete) transacciones.
- Se selecciona la coordenada (x_i, y_i) , correspondiente al punto donde se realizó la primera transacción del día i .
- Se calcula la distancia entre las coordenadas espaciales de todas las transacciones, primeras en cada día, (x_i, y_i) .
- Si la distancia entre las transacciones más lejanas es menor o igual a D , se establece que el individuo vive en el entorno de ese grupo de transacciones.
- Se determina a qué comuna pertenecen (mayoritariamente) las transacciones, y con esto se asigna cada tarjeta a una de las seis grandes zonas de Santiago.

Para clarificar, la siguiente figura ?? (a) muestra la primera transacción diaria de un usuario en una semana de datos, dentro de la ventana horaria 4:00 am - 12:00 am, en un esquema donde se supone conocido el lugar de residencia. En este caso, el usuario tiene transacciones en cuatro días de la semana en la ventana horaria requerida y por tanto es considerado un usuario frecuente de transporte público. Sin embargo, la posición del hogar no es información conocida, solo se conoce la ubicación espacial de las transacciones, como se muestra en la figura ?? (b). El método supone que las primeras transacciones de cada día se encuentran a una distancia caminable del hogar. Para descartar tarjetas que tengan sus primeras transacciones de cada día excesivamente lejos entre sí (razones para esto podrían ser que la persona usó un modo no integrado antes de realizar la primera transacción o que pasó la noche en otro lugar que no corresponde a su hogar) se exige que el diámetro de la circunferencia que agrupa a las transacciones no exceda la distancia D . Para este trabajo se considerará D igual a 1000m. Se escoge este tamaño de circunferencia porque la vivienda no necesariamente se encuentra central entre las transacciones y se considera como distancia caminable hasta 1000m, lo que corresponde al diámetro de una circunferencia de radio 500m.

Figura 4.2: Método de estimación de zona de residencia: a) Esquema de ubicación del hogar y transacciones; b) Información observada y radio de circunferencia.



Es relevante señalar que se realizaron pruebas para D menor a 1000m de distancia entre las transacciones más lejanas de una tarjeta. Lógicamente, al aumentar la distancia se logra incrementar el número de tarjetas a las que se estima zona de residencia. No se consideran distancias mayores (por ejemplo 1200m) debido a una restricción impuesta por el método de estimación de bajada, método que se aplica previamente a los datos para obtener el destino de cada viaje. El método de estimación de bajada considera como distancia caminable hasta 1000m. Si la bajada estimada de una transacción y la subida de la transacción siguiente se encuentran a una distancia mayor, el método asume que se utilizó otro modo de transporte (distinto a la caminata y el transporte público) para trasladarse, por lo cual se trunca el viaje. Para mantener coherencia con este método precedente, se ha considerado mantener la restricción.

El motivo para seleccionar la primera transacción entre 4:00 y 12:00 hrs responde a lo siguiente: En primer lugar, una cantidad importante de tarjetas realiza la última transacción de cada "día de viajes" después de media noche. Si se considera utilizar la primera transacción de cada día desde media noche, lo que se obtiene es un punto espacialmente cercano a la última actividad realizada por el individuo el día anterior, pues la transacción corresponde al retorno al hogar y no al origen del hogar. En todos los cortes temporales en que se estudio la distribución de transacciones, se observa que los horarios con menor número de transacciones se ubican entre las 3:15 y las 4:00 am. Por tanto, se considera que el "día de viajes" de una tarjeta comienza a las 4:00 am y termina a las 3:59 am del día siguiente. En segundo lugar, una cantidad importante de tarjetas registran el primer viaje del día en punta tarde. Es posible que esto ocurra, pero también es posible que su primer viaje del día haya sido

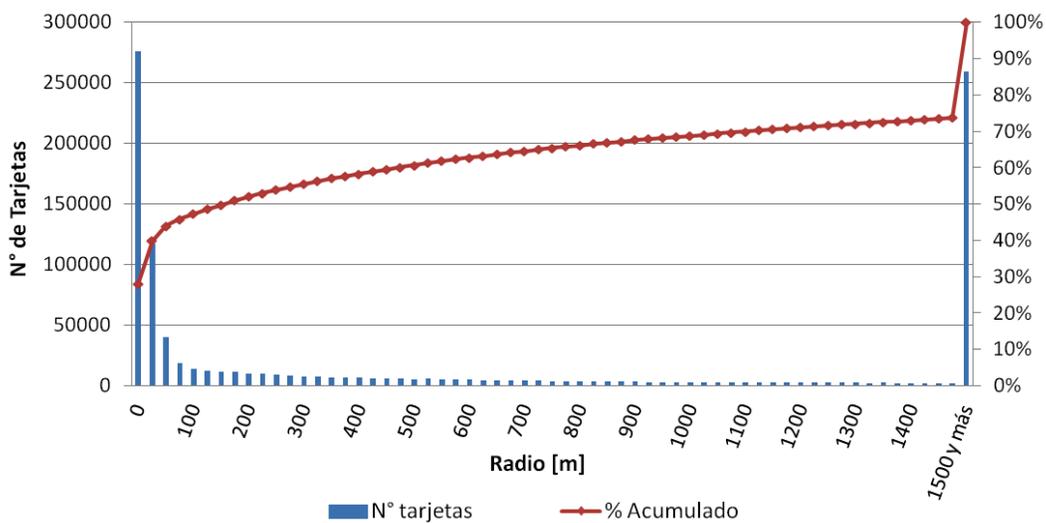
realizado en otro modo, como auto-acompañante o colectivo, y por ello no exista registro de sus viajes. Por esto se utiliza las 12:00 hrs (previo a la punta medio día) como límite horario para realizar el primer viaje del día. Por otra parte, para tarjetas que solo registren viajes fuera de este horario (es decir, entre 12:00 pm y 4:00 am del día siguiente) se considera implícitamente que siempre hacen el primer viaje del día en otro modo distinto al transporte público.

Es importante señalar que los horarios establecidos en la metodología y el diámetro de la circunferencia que es utilizado para determinar el lugar de residencia son valores arbitrarios para los cuales no se cuenta con mediciones que permitan calibrarlos adecuadamente, pero están basado en el análisis de los datos y consistencia con el resto de la metodología.

4.2.2. Resultados

La metodología descrita fue aplicada a la base de datos temporalmente más reciente, correspondiente a la tercera semana de abril de 2012. De un total de 2.593.145 tarjetas presentes en esa semana de datos, 985.085 observciones (el 38 %) cumplen con el número de viajes diarios requeridos en la ventana horaria especificada (al menos cuatro). De estas tarjetas, un 60 % (591.976) tienen radios de circunferencia de hasta 500m que circunscriben a todas las transacciones seleccionadas. Este universo de tarjetas permite observar patrones de comportamiento por zona. La figura ?? muestra la distribución de estos radios.

Figura 4.3: Distribución de los radios de circunferencia que circunscriben las transacciones de cada individuo. Semana de abril 2012.



Las tarjetas con radios inferiores a 25m están asociadas mayoritariamente a quienes utilizaron la misma estación de metro en las primeras transacciones de cada día, razón por la cual la primera columna presenta una fuerte concentración de tarjetas. En este trabajo se ha estimado residencia a la suma de tarjetas de las barras que se encuentran hacia la izquierda de la barra correspondiente a 500m de radio. Las barras ubicadas a la derecha corresponden

al total de tarjetas sin estimación.

4.2.3. Validación exploratoria del método asignación de zona de residencia

Con el objetivo de explorar la validez del método de asignación de zona de residencia, se realizó una validación exploratoria utilizando información de un grupo de 55 voluntarios, quienes accedieron a entregar su número de tarjeta y comuna de residencia en el contexto de una validación más general de la metodología (Silva, 2012).

De los 55 voluntarios, solo 15 tenían transacciones en al menos cuatro días de la semana entre las 4:00 y las 12:00 hrs. La Tabla ?? muestra el ID de la tarjeta y la comuna y zona de residencia declaradas por el individuo, más la comuna y zona de residencia predichas por el método. La última columna contiene el radio de la circunferencia que circunscribe a las transacciones de los individuos que cumplen con la condición de número de transacciones en la semana (al menos cuatro), pero no necesariamente con la condición de relación de distancia entre las transacciones. Las transacciones que no cumplen esta condición incluyen un asterisco (*). Por otra parte, por cada transacción se registra la coordenada (x, y), y por tanto comuna y zona en que se realiza la validación. Para predecir lugar de residencia (comuna o zona) el método asigna al individuo el lugar en que se ubiquen el total o la mayoría de sus transacciones seleccionadas.

Los resultados de este proceso se presentan en la Tabla ?. Como se observa, la predicción de comuna y zona de residencia tiene un 100 % de aciertos en los casos en que se cumple la condición de radio de circunferencia de hasta 500m e incluso acierta en casos de radios mayores. Además, hay 100 % de aciertos en la zona de residencia, incluyendo los casos en que no se cumple la condición para el radio de la circunferencia. Este resultado sugiere que probablemente el radio de circunferencia exigido es bastante estricto. Sin embargo, el diámetro de la circunferencia es consecuente con la metodología de estimación de bajada descrita en el capítulo 2, en cuanto a lo que se considera distancia caminable (1000m).

Tabla 4.1: Resultados de la verificación exploratoria del método de estimación de zona de residencia.

ID tarjeta	Comuna Residencia	Zona Residencia	Predicción Comuna	Predicción Zona	Radio
xxxxx8324	Santiago	Centro	Santiago	Centro	595*
xxxxx3978	Maipú	Poniente	Maipú	Poniente	8990*
xxxxxx3067	Cerro Navia	Poniente	Cerro Navia	Poniente	57
xxxxxx0052	Maipú	Poniente	Maipú	Poniente	70
xxxxxx8036	Las Condes	Oriente	Las Condes	Oriente	26
xxxxxx6564	Pudahuel	Poniente	Maipú	Poniente	2075*
xxxxxx4740	Estación Central	Poniente	Estación Central	Poniente	23
xxxxxx9275	Maipú	Poniente	Estación Central	Poniente	10200*
xxxxxx5770	Huechuraba	Norte	Huechuraba	Norte	105
xxxxxx7452	Maipú	Poniente	Lo Prado	Poniente	2868*
xxxxxx9844	San Miguel	Sur	San Miguel	Sur	0
xxxxxx3364	Maipú	Poniente	Maipú	Poniente	5702*
xxxxxx4628	Las Condes	Oriente	Las Condes	Oriente	217
xxxxxx4764	La Florida	SurOriente	La Florida	SurOriente	8688*
xxxxxx4876	La Florida	SurOriente	La Florida	SurOriente	484

Por supuesto, esta validación exploratoria no tiene validez estadística, pero es una primera aproximación a cómo se comporta el método expuesto.

4.3. Síntesis y conclusiones

En este capítulo se presenta la zonificación escogida. Corresponde a una zonificación utilizada previamente por Olgún et al. (2009) y Jara-Díaz et al. (2011) para el estudio de perfiles de uso de tiempo. Posteriormente se describe la metodología de predicción de zona de residencia, basada en la posición geográfica de la primera transacción de cada día de una semana de viajes. Luego se presentan los resultados del modelo en cuanto a cómo se comporta el método al imponer las condiciones de predicción. Finalmente se presentan los resultados de una validación exploratoria con un grupo reducido de personas que accedieron a entregar la información requerida.

Se concluye que a un 23 % del total de tarjetas se le puede predecir el lugar de residencia con una semana de datos consecutivos, debido a que la mayoría de las tarjetas corresponde a usuarios poco frecuentes, es decir, con insuficientes viajes como para adjudicarles una zona de residencia.

Los resultados de la validación exploratoria son positivos, pues hubo 100 % de acierto en la predicción. Se estima que por el tamaño de zona utilizado, es posible que la condición de distancia exigida sea demasiado estricta y que calibrando el radio podría aumentar el porcentaje de tarjetas con predicción. Sin embargo, se optó por mantener el radio utilizado ante la falta de una muestra de calibración y también por las condiciones impuestas por el método de estimación de bajada. Es importante señalar que esta validación exploratoria no tiene validez estadística porque el grupo de usuarios encuestados es extremadamente reducido.

En cuanto a la metodología de estimación de bajada se debe considerar varios aspectos. En primer lugar, el radio utilizado para predecir zona de residencia debería calibrarse y depender del tamaño de las zonas a las cuales se asignará residencia. Es evidente que el nivel de acierto en la predicción cambiará con esta variable. Por otra parte, la metodología no parece tener límite en cuanto a ampliar el tamaño de las zonas, pero si en cuanto a reducirlo. Suponiendo que se ha calibrado la distancia caminable para un viaje de transporte público, esta distancia es el límite de precisión del método. El supuesto detrás de la metodología es que una persona accede al transporte público en modo caminata y esta característica es la que permite presumir que vive en una superficie cercana a sus transacciones. Utilizar zonas de dimensiones menores a la distancia caminable es incongruente con este mismo argumento. Pese a esto, si se utiliza zonas de tamaño similar a lo que se considera distancia caminable, será difícil establecer a qué zona pertenece un individuo cada vez que sus transacciones se repartan en los límites entre dos o más zonas, o se aproximen a los límites de una zona. Este es el caso de la zonificación EOD 777, donde varias zonas tienen al menos una de sus dimensiones de tamaños comparables o menores a 1000m.

Aunque el problema de asignar zona a las tarjetas con transacciones en los límites zonales no es exclusivo de zonificaciones de menor tamaño, en zonificaciones mayores se reduce fuertemente. Para resolver definitivamente esta dificultad se podría utilizar zonas dinámicas, es decir, zonas que no tienen límites preestablecidos, en las cuales se calcule el centroide de las transacciones de una tarjeta y se considere como zona de residencia un radio de influencia propio de cada tarjeta, en vez de la zona estática donde se ubican la mayoría de las transacciones.

Otra consideración relevante es la heterogeneidad del acceso al transporte en la ciudad. Es probable que la distancia caminable varíe entre zonas de la ciudad según la densidad de accesos a la red y otras variables, por lo que puede calibrarse más de un radio de circunferencia según lugar de la ciudad. Sin embargo, se requiere calibrar previamente la distancia caminable por zona de la ciudad para poder realizar este tipo de análisis.

Finalmente, contar con una muestra de datos de más de una semana permitiría mejorar la precisión de la predicción, aumentar el porcentaje de tarjetas a las cuales se les puede predecir zona de residencia y llevar la predicción a otros ámbitos como estimar el lugar de trabajo o estudio de cada usuario.

Capítulo 5

Resultados

Este capítulo muestra los patrones de comportamiento de viaje observados en uno o más cortes temporales. Los comportamientos explorados provienen del análisis de datos mostrado en el capítulo tres, el cual se inspira en los trabajos presentados en la revisión bibliográfica.

Para no redundar, los análisis que no tienen por objetivo observar la evolución temporal se presentan para un único corte de tiempo. Se ha escogido la semana de datos más reciente (abril 2012) para exponer el resultado más cercano a la situación actual.

En cuanto a la metodología de estimación de zona de residencia, la información socioeconómica obtenida permite desagregar los resultados por esta variable para observar posibles diferencias en los patrones de viaje de quienes habitan distintos sectores de la ciudad. Sin embargo, para algunos de los patrones de comportamientos explorados no se muestra resultados por zona porque por construcción la metodología genera sesgo en ciertos análisis. Esto se explicará en detalle más adelante.

En este capítulo se estudia el patrón de viajes diarios de las tarjetas (número de viajes por día); el patrón de transacciones entre dos cortes temporales; los tiempos de viaje en vehículo y trasbordo por zona de residencia; y el perfil de uso de tiempo de Santiago, también por zona de residencia.

5.1. Patrón de viajes diarios

Como se mostró en los capítulos dos y tres, la frecuencia de viaje es una de las características del uso de la tarjeta que permite diferenciar comportamientos de viaje entre usuarios. Cuantificar el número de viajeros según sus horarios y frecuencia de viaje es información crucial para la planificación y predicción de la demanda por transporte público.

A continuación se presenta la distribución diaria de viajes de las tarjetas, donde los usuarios son clasificados según el número de viajes diarios. Se muestra resultados para tarjetas que viajan desde una hasta cinco veces en un día (viajes, no transacciones). La unidad de observación es el día-tarjeta, por lo que una misma tarjeta puede pertenecer a distintas ca-

tegorías en diferentes días, pero en un mismo día solo puede aparecer en una de las categorías.

Este análisis no se realiza por zona de residencia, debido al sesgo que el método de estimación de bajada induce en la distribución horaria de los viajes. Como la metodología requiere seleccionar usuarios con al menos un viaje entre las 4:00 y 12:00 am, los usuarios que realizan su primer viaje fuera de este horario son excluidos, modificando las campanas de distribución temporal de los viajes al descartar a quienes no cumplen esta condición. Por tanto, el comportamiento mostrado en las dos siguientes subsecciones se realiza para el total de tarjetas de la base de datos abril 2012. Los resultados se muestran separadamente entre tarjetas tipo adulto y estudiante.

5.1.1. Patrón de viajes diarios de tarjetas tipo adulto

En esta sección se presenta el patrón de viajes diarios de las tarjetas tipo adulto observado en la semana de datos de abril-2012. La Tabla ?? muestra la clasificación de las tarjetas según número de viajes por día para esa semana de datos. Existen tantas observaciones como días-tarjetas. Por ejemplo, si una tarjeta muestra transacciones en tres días de la semana, habrá tres observaciones para ese número de tarjeta. Para cada día se cuenta el número de viajes realizados y cada observación se incluye en alguna de las categorías que aparecen en la primera columna de la tabla.

De los más de 16 millones de viajes registrados en una semana para la tarjeta adulto, el comportamiento más común es realizar dos viajes diarios (48,3 %). Se destaca que un 16 % de los viajes corresponde a tarjetas que realizan sólo un viaje diario. Este comportamiento puede explicarse por la utilización de modos no integrados al transporte público, evasión de pago del pasaje, entre otras razones. La información disponible actualmente no permite cuantificar qué proporción de los viajes se debe a una u otra causa.

Tabla 5.1: Información del patrón de viajes diarios para tarjeta tipo adulto.

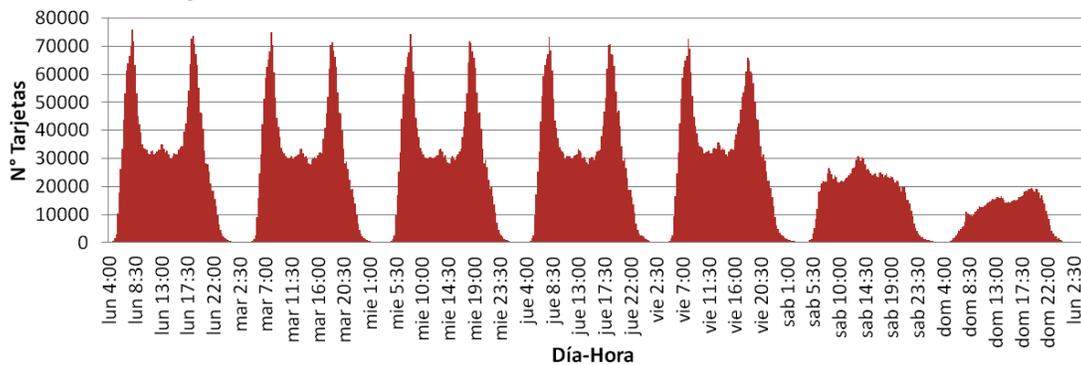
Nº de viajes diarios	Nº de observaciones	Nº de viajes semanales	% del total de viajes
1	2.617.575	2.617.575	16,0
2	3.951.576	7.903.152	48,3
3	1.064.895	3.194.685	19,5
4	392.446	1.569.784	9,6
5	116.290	581.450	3,6
Acumulado		15.866.646	97,1
Total de viajes		16.343.346	
Total de tarjetas		2.593.145	

La Figura ?? muestra la distribución horaria del total de viajes (100 %). La distribución muestra que punta mañana y punta tarde son bastante simétricas, salvo el día viernes en que la punta tarde comienza antes y tiene mayor dispersión. La punta mañana tiene su punto más alto entre las 8:00 y 8:15 hrs. y la punta tarde entre las 18:15 y 18:30 hrs. Se observa

una pequeña punta al medio día en torno a las 13:00 hrs. El fin de semana el total de viajes diarios se reduce notoriamente y es mucho más estable a lo largo del día. El día sábado se observa el punto más alto en el cuarto de hora que comienza a las 13:15, mientras que el domingo ocurre a las 19:00 hrs.

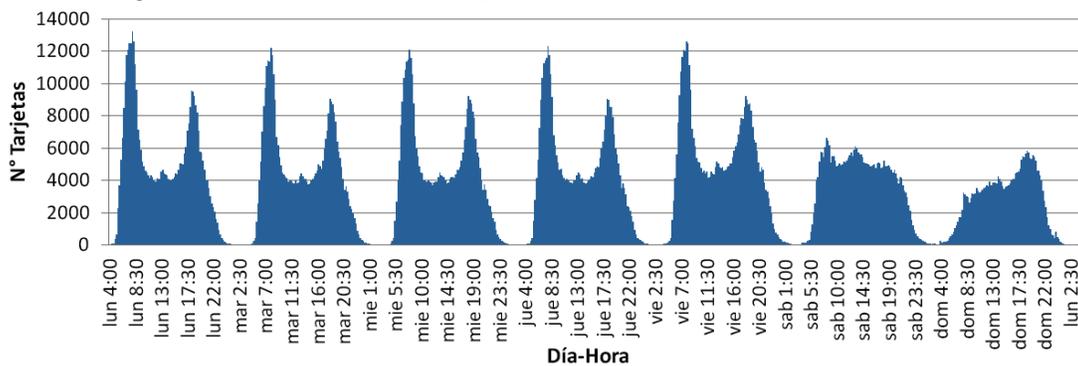
Esta distribución del total de viajes se utilizará como base para ser comparada con las distribuciones desagregadas por número de viajes, según la clasificación de uno a cinco viajes diarios (equivalentes en suma al 97,1 % del total de viajes en una semana). En el eje horizontal se indica la hora del día en tramos de 15 minutos y en el eje vertical el número de tarjetas que inicia su viaje en ese lapso.

Figura 5.1: Distribución de viajes totales, tarjetas tipo adulto.



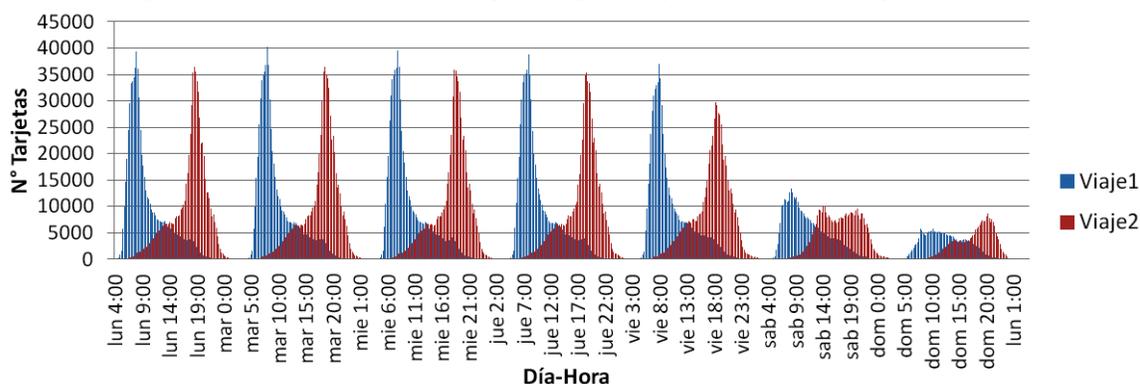
La Figura ?? muestra la distribución horaria de los viajes realizados por tarjetas que efectúan solo un viaje al día. Comparando con la distribución del total de viajes, se puede ver que la punta mañana es mucho más pronunciada que la punta tarde. Asumiendo que cada uno de estos viajes tiene su par en el día, el cual no se observa porque se utiliza otro modo (como autoacompañante o taxi), se evade el pago del pasaje o cualquier otra razón, esto ocurre más frecuentemente durante la tarde/noche, presumiblemente (por horario) para el viaje de retorno. Los días laborales y el domingo tienen los mismos horarios punta observados para el total de viajes, mientras que el día sábado la punta se traslada a la mañana, entre las 8:15 y 8:30 hrs.

Figura 5.2: Distribución de viajes, tarjetas tipo adulto, un viaje diario.



El caso de las tarjetas que realizan dos viajes diarios se muestra en la Figura ???. La distribución del primer viaje del día se muestra en azul y la del segundo en rojo. Esta distribución es casi idéntica a la del total de viajes si se considera la suma de ambas campanas. Siendo éste el comportamiento más común de los cinco observados, es lógico que la distribución del total de viajes muestre con más fuerza el comportamiento que aquí se observa. La punta medio día no aparece porque se forma de la suma de quienes realizan el primer y segundo viaje en este horario. La principal diferencia es que el efecto de dispersión de la campana de punta tarde del día viernes es más marcado.

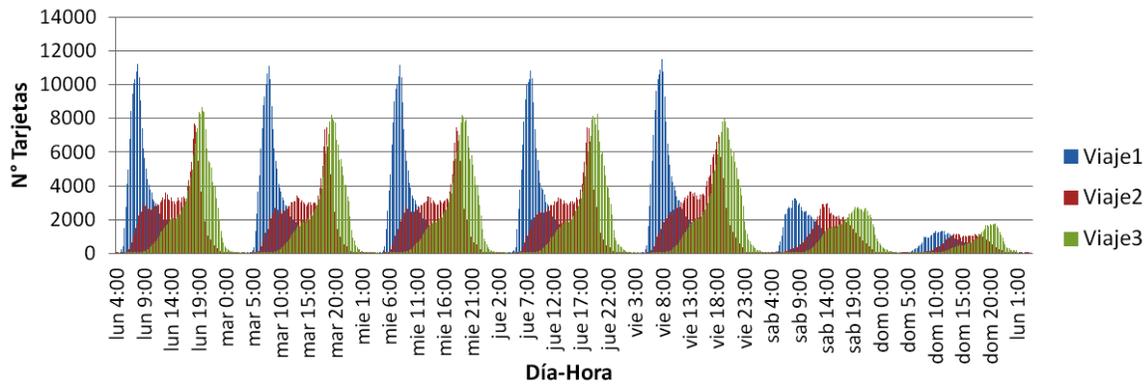
Figura 5.3: Distribución de viajes, tarjetas tipo adulto, dos viajes diarios.



Con la visualización anterior también es posible apreciar las diferencias en la distribución de los viajes entre día laboral, sábado y domingo. Los días de fin de semana las campanas de viaje tienen proporcionalmente mucha mayor dispersión respecto a los días hábiles y sus formas son completamente diferentes entre sí.

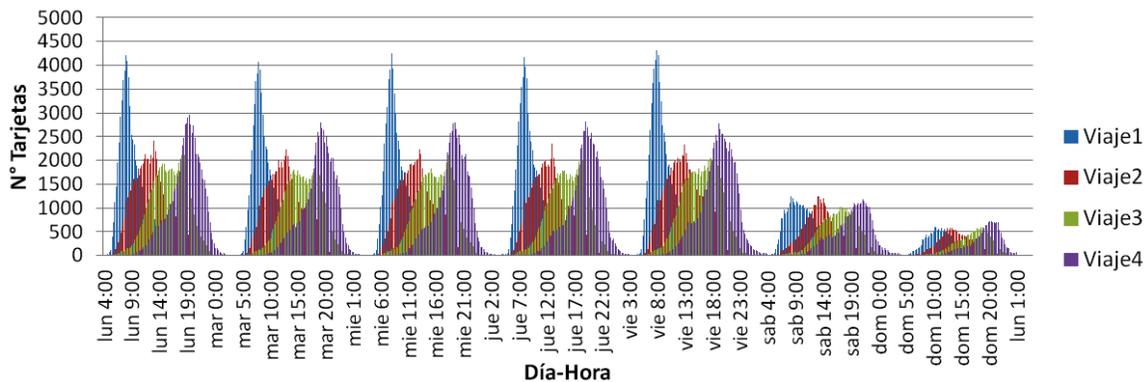
Para las tarjetas que realizan tres viajes diarios, la distribución horaria de éstos se muestra en la Figura ???. Se puede ver cambios sustanciales. La distribución del segundo viaje empieza poco después de la punta mañana y se mantiene casi constante hasta el horario punta tarde, que ocurre a la misma hora señalada en los casos anteriores. Desde ahí se traspone en gran medida con el tercer viaje del día, campana que tiene su punto máximo aproximadamente una hora después del segundo viaje, entre 19:15 y 19:30 hrs. Este comportamiento se asocia a quienes realizan una actividad como *comprar* o *recoger a alguien* después de abandonar el lugar de trabajo. En este caso, las últimas dos campanas del día son mucho más dispersas que la primera y la forma de la campana central es completamente distinta de las otras. El día sábado se observan campanas de viaje similares distribuidas bastante uniformemente a lo largo del día, mientras que el domingo hay una leve concentración en el último viaje del día, que ocurre entre 19:00 y 21:00 hrs. Este es el segundo comportamiento más común e involucra al 19,5 % de los viajes.

Figura 5.4: Distribución de viajes, tarjetas tipo adulto, tres viajes diarios.



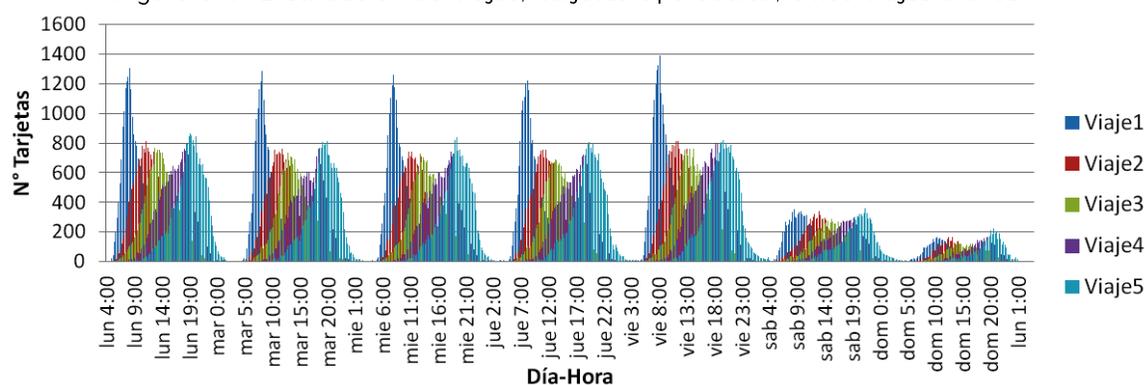
La Figura ?? exhibe el comportamiento de las tarjetas que realizan cuatro viajes al día. La primera campana del día no muestra diferencia respecto a los casos anteriores. El segundo viaje del día está centrado en las 13:00 hrs, que se asocia al horario de inicio de colación. El tercer viaje del día presenta una cantidad similar de viajes entre las 13:30 y 18:45 hrs, con dos pequeñas puntas a las 14:30 y 18:00 hrs. Es decir, en esta campana parecen juntarse quienes realizan un viaje posterior a la hora de almuerzo con quienes realizan una actividad al salir del trabajo, antes de volver a su hogar. El último viaje del día presenta una dispersión similar al caso anterior y mantiene su punto máximo entre 19:00 y 19:15 hrs, una hora después del tercer viaje. El comportamiento de fin de semana repite las características descritas en el caso antecedente.

Figura 5.5: Distribución de viajes, tarjetas tipo adulto, cuatro viajes diarios.



La Figura ?? muestra el comportamiento de las tarjetas que realizan cinco viajes diarios. Contrario a lo que podría creerse, las puntas del primer y último viaje siguen centradas en las 8:00-8:15 y 19:15-19:30 hrs, respectivamente. Aunque ciertamente aumenta la dispersión de los viajes, éstos se acumulan dentro del mismo espacio temporal. El segundo viaje del día adelanta su punta al lapso que va desde las 10:30 a las 11:45. El tercer viaje se concentra entre 13:00 y 14:00 hrs y los últimos dos viajes vuelven a traslaparse con cerca de una hora de diferencia entre sí. El fin de semana mantiene la tendencia mostrada en los dos casos anteriores.

Figura 5.6: Distribución de viajes, tarjetas tipo adulto, cinco viajes diarios.



Fuera de este análisis han quedado las distribuciones de viaje de las tarjetas que realizan seis o más viajes diarios. Sin embargo, estas corresponden a menos del 3 % del total de viajes. La tendencia muestra que para tres o más viajes diarios, el primer y último viaje no cambian sustancialmente de horario y que las campanas de viaje intermedias propenden a superponerse entre sí. También se observa un fuerte aumento en la dispersión de las campanas para tres o más viajes diarios, salvo en el primer viaje del día. Los fines de semana las campanas de viaje son similares en dispersión y altura; con excepción del día domingo donde la última campana del día es más pronunciada.

5.1.2. Patrón de viajes diarios de tarjetas tipo estudiante

La presente sección muestra un análisis similar al anterior, para las tarjetas tipo estudiante. El detalle según número de viajes diarios se expone en la Tabla ???. Se puede ver que para estas tarjetas el comportamiento más común es realizar dos viajes diarios, al igual que en el caso anterior, aunque las tarjetas tipo estudiante que realizan uno y dos viajes diarios son menos en proporción a las tipo adulto, siendo más aquellas que realizan tres o más viajes por día.

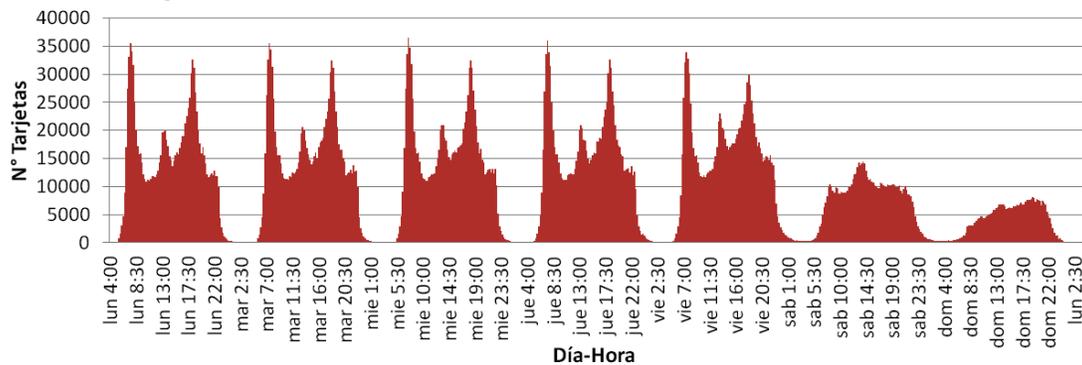
Tabla 5.2: Información del patrón de viajes diarios para tarjeta tipo estudiante.

N° de viajes diarios	N° de observaciones	N° de viajes semanales	% del total de viajes
1	649.833	649.833	9,1
2	1.543.011	3.086.022	43,2
3	573.762	1.721.286	24,1
4	253.274	1.013.096	14,2
5	80.540	402.700	5,6
Acumulado		6.872.937	96,1
Total de viajes		7.149.558	
Total de tarjetas		695.319	

La Figura ??? muestra la distribución temporal del total de viajes de las tarjetas tipo estudiante en intervalos de 15 minutos. En el eje vertical representa el número de tarjetas

que inicia su viaje en cada lapso temporal. De lunes a viernes esta distribución es bastante distinta a la de las tarjetas tipo adulto. La principal diferencia se observa los días hábiles, donde aparecen cuatro en vez de tres puntas de viaje, la última punta alrededor de las 21:30 hrs es atribuible a estudios en horario vespertino. El fin de semana el comportamiento es similar, la distribución de viajes del día sábado alcanza su punto mayor a las 14:00 hrs, mientras que el día domingo se observa la mayor demanda a las 19:00 hrs.

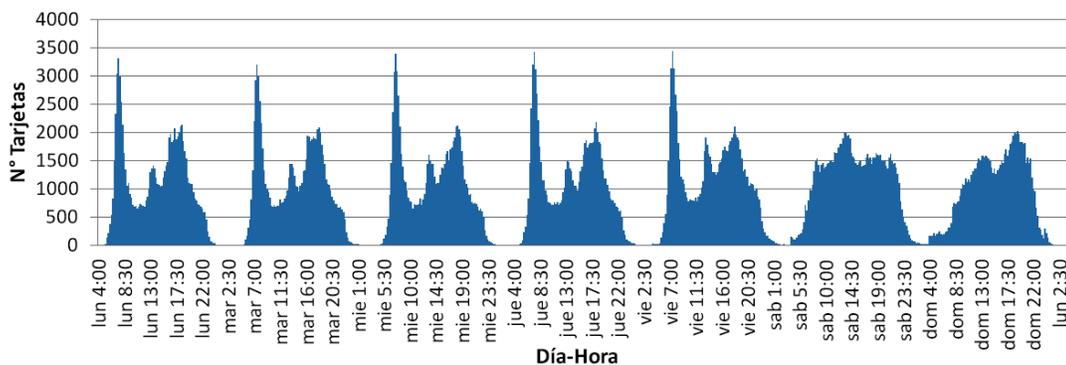
Figura 5.7: Distribución de viajes totales, tarjetas tipo estudiante.



Para entender su comportamiento, a continuación se presenta la distribución de viajes para tarjetas tipo estudiante que realizan entre uno y cinco viajes diarios.

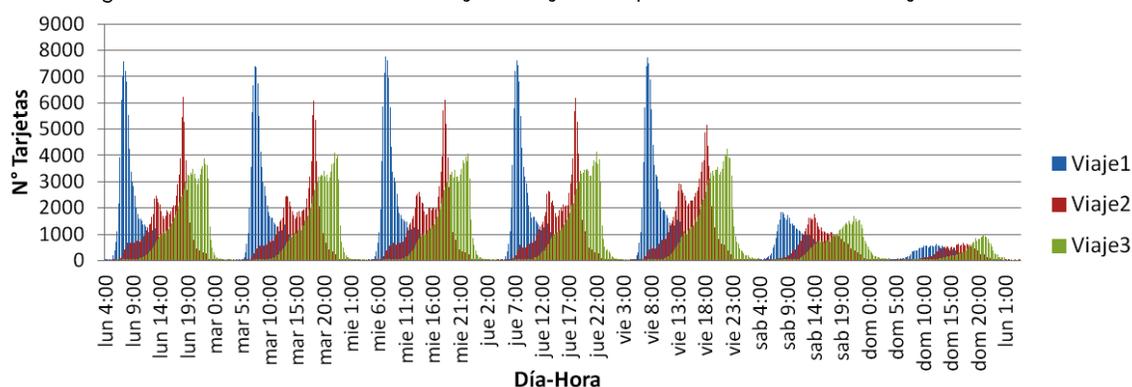
La Figura ?? muestra la distribución para las tarjetas que realizan un viaje al día. Este caso presenta solo tres puntas de viaje y, al igual que para la tarjeta adulto, se acentúa la diferencia entre la proporción de viajes de punta mañana y punta tarde respecto al comportamiento observado en el total de viajes. Este patrón de comportamiento se asocia al uso de modos no integrados y la evasión del pago del pasaje, entre otros. El día sábado el comportamiento es similar a lo observado en el total de viajes, en cambio el día domingo aparecen dos puntas entre 13:00 y 14:00 hrs y entre 18:45 y 19:30 hrs. Llama la atención que los estudiantes no reduzcan fuertemente sus viajes el día domingo, como ocurre con la tarjeta adulto.

Figura 5.8: Distribución de viajes, tarjetas tipo estudiante, un viaje diario.



La Figura ?? muestra la distribución de viajes de tarjetas que realizan dos viajes diarios. Se observa un singular comportamiento en la campana del primer viaje del día que tiene tres

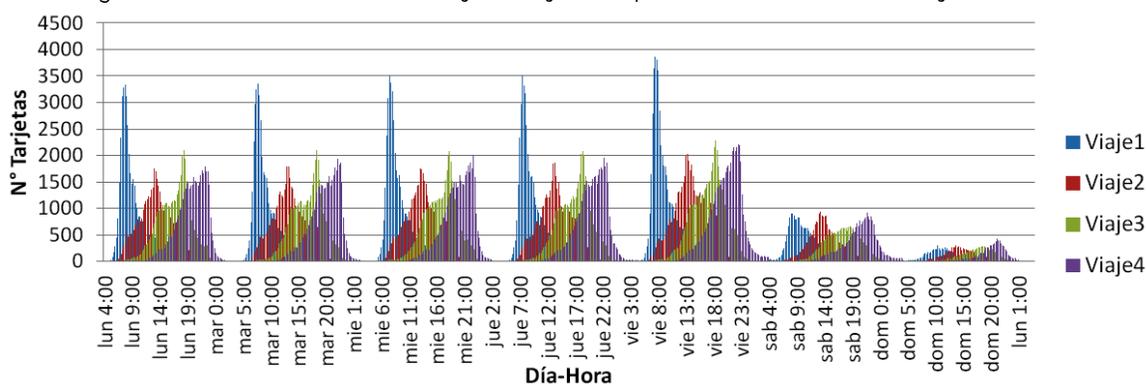
Figura 5.10: Distribución de viajes, tarjetas tipo estudiante, tres viajes diarios.



La Figura ?? muestra la distribución temporal de viajes para las tarjetas que realizan cuatro viajes diarios. Se reduce casi completamente el efecto de campanas con múltiples puntas. El primer viaje mantiene su máximo a las 7:30 hrs. El segundo viaje tiene una punta a las 13:00 hrs y una segunda punta mucho más débil a las 18:00 hrs que se observa solamente de lunes a jueves. El tercer y cuarto viaje del día tienen su máximo a las 18:15 y 22:00 hrs, respectivamente. El día sábado presenta máximos a las 8:00, 13:00 y 21:30 hrs para el primer, segundo y cuarto viaje del día. El tercer viaje presenta una distribución más suave con similar número de viajes entre las 17:00 y 18:30 hrs. El día domingo se comporta de manera análoga al caso de tres viajes diarios, presentando una punta solo en el último viaje del día.

Aunque la estructura de viajes se asemeja a la observada para la tarjeta adulto, en el caso de la tarjeta estudiante las puntas del primer y último viaje del día ocurren más temprano y más tarde, respectivamente.

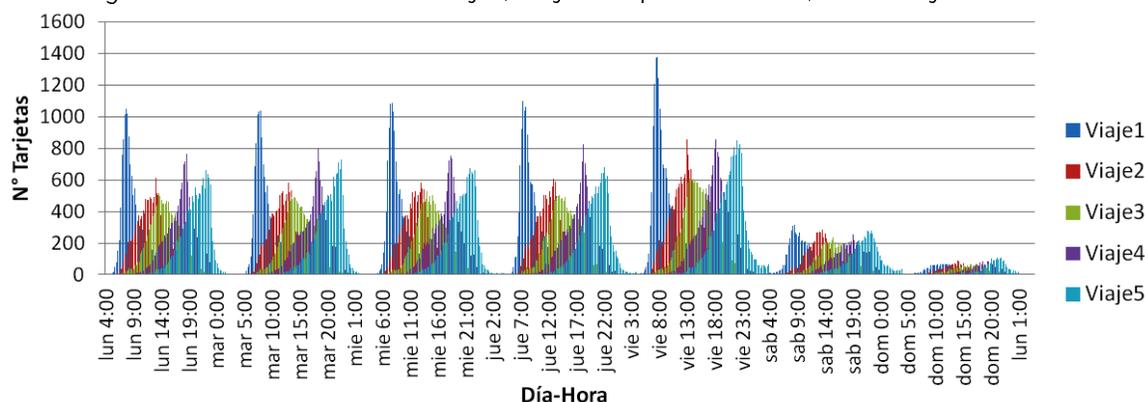
Figura 5.11: Distribución de viajes, tarjetas tipo estudiante, cuatro viajes diarios.



La Figura ?? muestra la distribución de viajes para las tarjetas que realizan cinco viajes diarios. Cada campana tiene un único horario punta. Los máximos se observan a las 7:30, 13:00, 13:30, 18:15 y 22:00 hrs. La distribución de viajes de punta mañana (primer viaje del día) es la campana con menor dispersión y el tercer viaje del día la con mayor. Los días de fin de semana se comportan de manera similar a lo expuesto en el caso previo. El primer y

último viaje del día ocurren en los mismos horarios ya descritos, por lo que las campanas de los viajes centrales tienen a juntarse.

Figura 5.12: Distribución de viajes, tarjetas tipo estudiante, cinco viajes diarios.



Las distribuciones de tarjetas que realizan seis o más viajes diarios no fueron estudiadas y corresponden a menos del 4 % del total de viajes. Se remarca la tendencia a la superposición de las campanas antes que la ampliación del horario de viaje, antes observada para la tarjeta adulto. Por otra parte, las tarjetas que realizan tres o más viajes muestran aumento de dispersión en todas las campanas de viaje del día, salvo la primera. Se intuye que el comportamiento de múltiples puntas en las campanas de viaje de estudiante se asocia a mezcla de tarjetas de estudiantes de distinto nivel. Adicionalmente, tarjetas tipo estudiante de personas que trabajan y estudian es una de las razones que explican que la tarjeta estudiante presente más viajes diarios por tarjeta en comparación a la tarjeta adulto.

Para ver el código (c++) aplicado a las bases de datos para obtener estos resultados, consultar Anexo A.

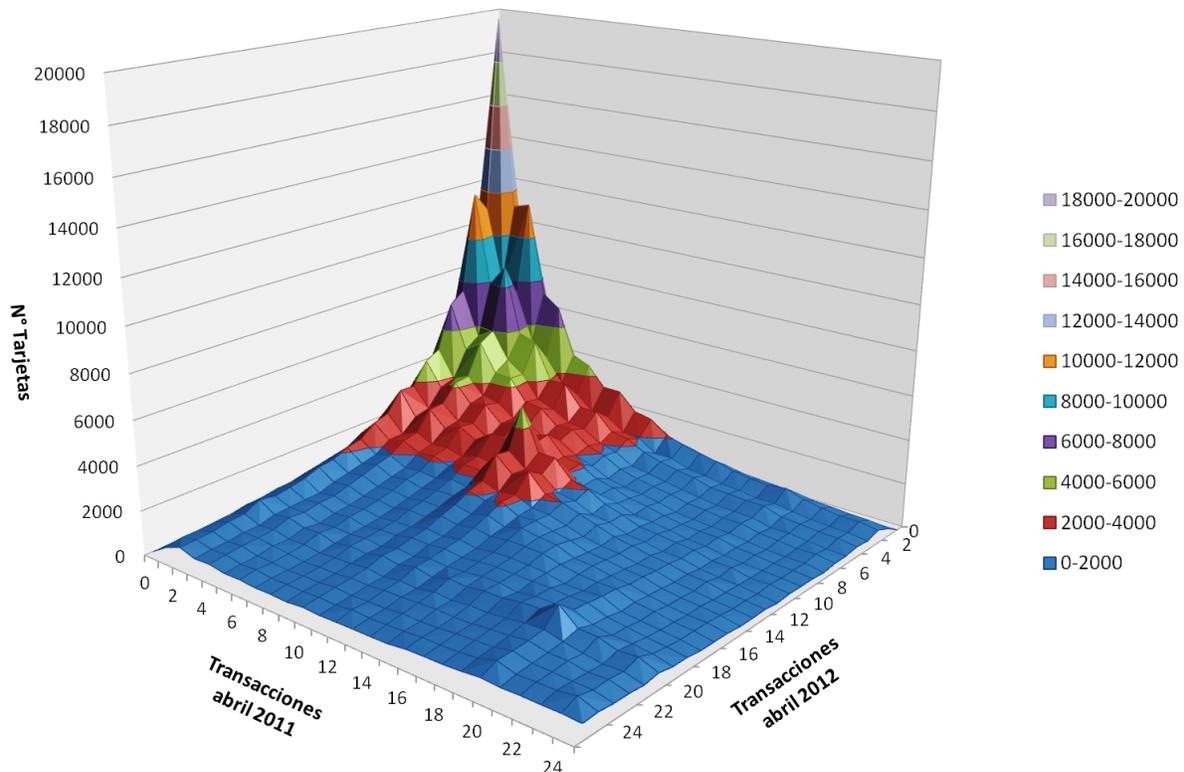
5.2. Patrón de transacciones entre dos cortes temporales

La permanencia de las tarjetas, mostrada en el capítulo tres, señala que existen usuarios que viajan de manera tan infrecuente que hay lapsos de al menos una semana en que no hay registro de transacciones. Se concluye que estudiar a los usuarios infrecuentes requiere de datos de periodos de tiempo mayores. Sin embargo, se abre la interrogante respecto del nivel de utilización de la tarjeta o la variación de la utilización de ésta entre cortes temporales.

Para ejemplificar, una tarjeta que aparece en dos cortes temporales realizando diez transacciones en cada uno no es igual a una tarjeta que realiza diez transacciones en el primer corte y solo dos en el segundo, o viceversa. En esta sección se analiza los diferentes casos: quienes mantienen el nivel de uso, quienes reducen su uso (posiblemente han optado por modos alternativos) y quienes aumentan el uso.

Para realizar este análisis se escogen las semanas de datos de abril-2011 y abril-2012, lo que además permite evitar cambios de comportamiento asociados a efectos estacionales. La Figura ?? muestra en los ejes horizontales el número de transacciones realizadas en una semana de viajes en cada corte temporal y en el eje vertical el número de tarjetas que realizan una o más transacciones en ambos cortes temporales. El total de tarjetas tipo adulto o estudiante con al menos una transacción en cada corte es de 1.228.499.

Figura 5.13: Patrón de transacciones semanales entre abril-2011 y abril-2012.



Este manto de viajes se ha truncado en 25 transacciones semanales por corte temporal. La tabla con el total de datos que completan la matriz de transacciones se encuentra en el Anexo B.

Se destaca el comportamiento común de realizar pocas transacciones en ambas semanas, correspondiente a usuarios que viajan infrecuentemente o que su principal modo de viaje no es transporte público. Notar que hasta las diez transacciones semanales, las transacciones pares son en su mayoría máximos locales atribuibles a que los viajes en pares (de ida y vuelta) son más comunes que aquellos impares (al menos un viaje sin su par).

También se destacan dos máximos locales particulares: (10, 10, 4.875) y (20, 20, 1.891) que corresponden al número de tarjetas que realizan 10 y 20 transacciones en cada corte temporal, respectivamente. Estos máximos podrían asociarse a viajes de ida y vuelta durante

los cinco días hábiles de la semana, cada viaje de una etapa en el caso del máximo (10,10) y de dos etapas en el caso del máximo (20,20). Estas tarjetas son buenas candidatas a estudios de regularidad de lugar y hora de viaje.

La siguiente tabla muestra el número de tarjetas que aumenta, mantiene y reduce el número de transacciones entre abril-2011 y abril-2012. $T_{2011} > T_{2012}$ significa que el número de transacciones realizadas en la semana de abril-2011 es estrictamente mayor que en abril-2012, es decir, las tarjetas que redujeron sus transacciones. Luego se muestra aquellas que mantienen ($T_{2011} = T_{2012}$) y aumentan el uso ($T_{2011} < T_{2012}$).

Tabla 5.3: Variación del número de transacciones por tarjeta entre abril-2011 y abril-2012.

Utilización	N° de tarjetas	%
$T_{2011} > T_{2012}$	596.573	48,6
$T_{2011} = T_{2012}$	95.669	7,8
$T_{2011} < T_{2012}$	536.257	43,7
Total	1.228.499	100

De las tarjetas que mantienen el número de transacciones (es decir, las que se ubican en la diagonal) la mayoría son aquellas que realizan entre una y cuatro transacciones en cada corte temporal. Es decir, usuarios poco frecuentes que probablemente combinan el uso de transporte público con otros modos.

En cuanto a la utilización, las tarjetas que reducen sus transacciones son cerca de un 5 % más que las que las aumentan. Por la separación de un año, se descartan efectos estacionales. Este resultado podría marcar una tendencia en el descenso del uso de transporte público y el reemplazo total o parcial por otros modos.

Para ver el código (c++) aplicado a las bases de datos para obtener estos resultados, consultar Anexo C.

5.3. Tiempo de viaje en vehículo y tiempo de trasbordo

El tiempo de viaje, trasbordo, acceso y egreso son algunas de las variables más estudiadas del transporte público por su relevancia en la calidad de servicio. En esta sección se analizan las variables tiempo de viaje en vehículo y tiempo de trasbordo. Ambos datos son obtenidos con la información de transacciones y las emisiones GPS, al aplicar el método de estimación de bajada descrito en el capítulo dos. Otras variables, como tiempo de acceso y egreso, no están disponibles en las bases de datos.

El tiempo de viaje en vehículo presentado se calcula como la suma de los tiempos de cada etapa del viaje, se utilice bus o metro. El tiempo de trasbordo es la suma de los tiempos transcurridos entre las etapas del viaje (tiempo de caminata más tiempo de espera). Es importante precisar que en los viajes de una etapa no hay trasbordo y que la obtención de tiempo

de trasbordo considera estos viajes para su cálculo, aunque el tiempo de trasbordo sea nulo. Los datos son agrupados en períodos de 15 minutos y en cada periodo se incluye el tiempo promedio de todos los pasajeros con la misma hora de inicio de viaje.

A continuación se muestran los resultados del análisis por tipo de tarjeta y zona de residencia. Se remarca que este análisis presenta los tiempos de viaje en vehículo y trasbordo observados para quienes habitan en cada zona, independientemente del origen y destino de sus viajes. Es decir, se muestran los promedios en tiempo dedicado a viaje y trasbordo para habitantes de cada zona de la ciudad.

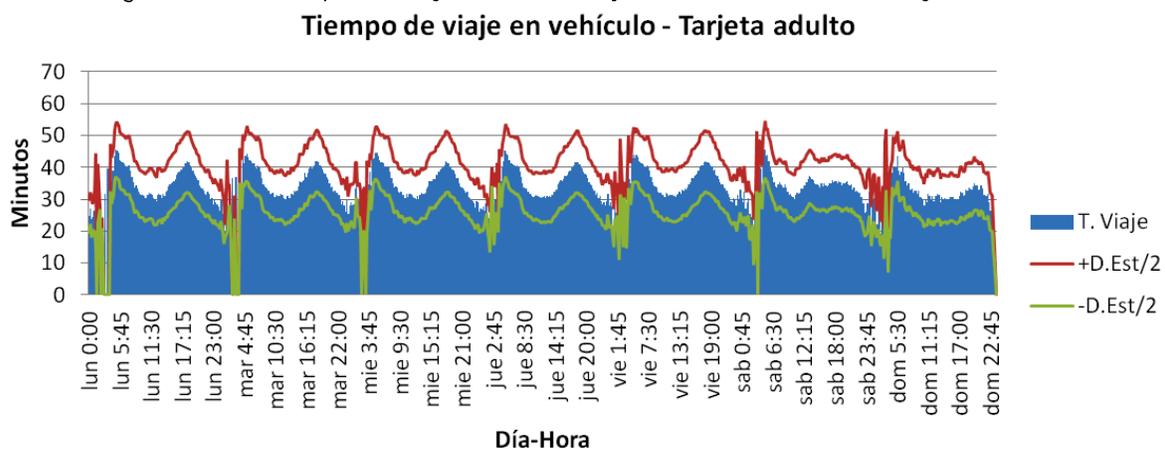
5.3.1. Tiempo de viaje en vehículo y de trasbordo de tarjetas tipo adulto

Esta sección expone los resultados de tiempo de viaje en vehículo y tiempo de trasbordo para la tarjeta tipo adulto. Es importante señalar que este análisis se realiza para todos los viajes, sin considerar si son de una, dos, tres o más etapas.

Las Figuras ?? y ?? muestran las distribuciones de tiempo de viaje en vehículo y trasbordo respectivamente. Además se presenta la desviación estándar en torno al promedio del tiempo. En el eje horizontal se indica el día y hora y en el vertical el tiempo en minutos.

Lo primero que se destaca es la similitud de la distribución de los tiempos entre días hábiles y las diferencias con los fines de semana. Los mayores tiempos de viaje se observan en las horas previas a la punta mañana y posiblemente corresponden a los viajes de largas distancias que comienzan antes precisamente por esta razón y no necesariamente debido a bajas velocidades. Los horarios de punta mañana y punta tarde tienen tiempos de viaje similares que bordean los 40 minutos. Los días de fin de semana no se aprecia el alza en punta tarde. Durante algunas horas de la madrugada la varianza es nula o inestable porque existen pocos o ningún viaje en ese horario.

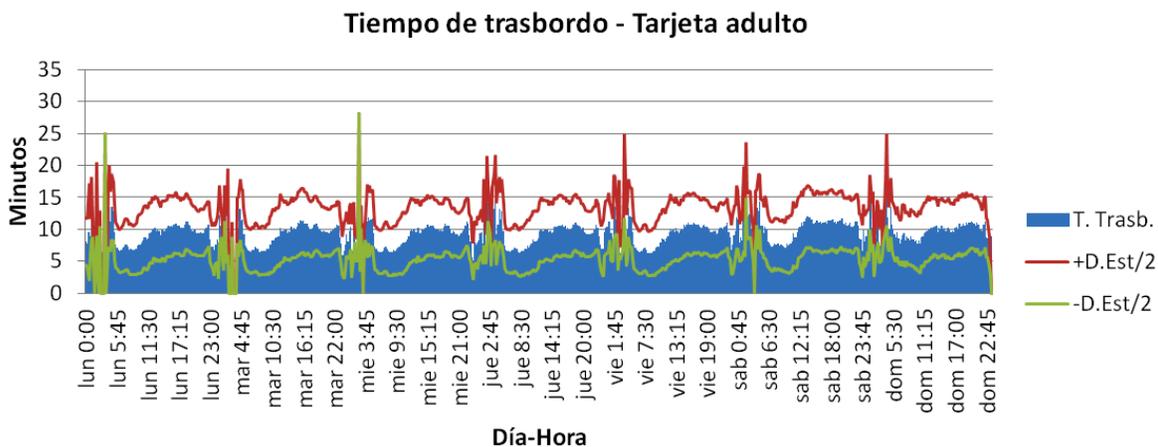
Figura 5.14: Tiempo de viaje en vehículo y desviación estándar, tarjeta adulto.



En cuanto al tiempo de trasbordo, este es menor en la punta mañana (cuando hay mayor

frecuencia de buses) y se mantiene relativamente estable a lo largo del día. Esta situación se repite el fin de semana, aunque con promedios de tiempo algo mayores respecto a los días hábiles. Este efecto se atribuye a la menor frecuencia de buses los fines de semana.

Figura 5.15: Tiempo de trasbordo y desviación estándar, tarjeta adulto.



La Tabla ?? contiene los promedios diarios de las variables por día en minutos. Se observan menores tiempos de viaje en vehículo los fines de semana, concordando con la menor congestión de estos días, además de diferencias en los pares origen-destino. Por otro lado, aumentan los tiempos de trasbordo, aunque no logran compensar el efecto de reducción del tiempo de viaje, especialmente el día domingo.

Tabla 5.4: Tiempo promedio de viaje en vehículo y trasbordo, tarjetas tipo adulto.

	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo
Tiempo de viaje [min]	37,8	37,8	37,9	37,8	38,0	34,6	32,0
Tiempo de trasbordo [min]	8,7	8,5	8,4	8,5	8,4	9,6	9,8
Suma [min]	46,6	46,3	46,2	46,3	46,4	44,3	41,8

Los gráficos y la tabla anteriores parecen indicar que no hay diferencias de estas variables entre días hábiles. Para corroborar la hipótesis de que las distribuciones de tiempo en los cinco días hábiles son las mismas, se propone promediar las distribuciones de tiempo de los cinco días hábiles y observar las diferencias de cada día respecto al promedio.

Al aplicar este procedimiento a las curvas de tiempos se obtienen los resultados mostrados en las Figuras ?? y ?? para tiempo de viaje y tiempo de trasbordo respectivamente. Se observa que en las horas de la madrugada existe mucha varianza, lo que se debe a los pocos datos de tiempos registrados en este horario y a la baja frecuencia de los servicios. Sin embargo, para el resto del día todas las curvas se mueven de manera similar, ningún día hábil presenta un comportamiento notoriamente distinto al de los otros.

Figura 5.16: Comparación entre días hábiles de variación del tiempo de viaje en vehículo(min), tarjeta adulto.

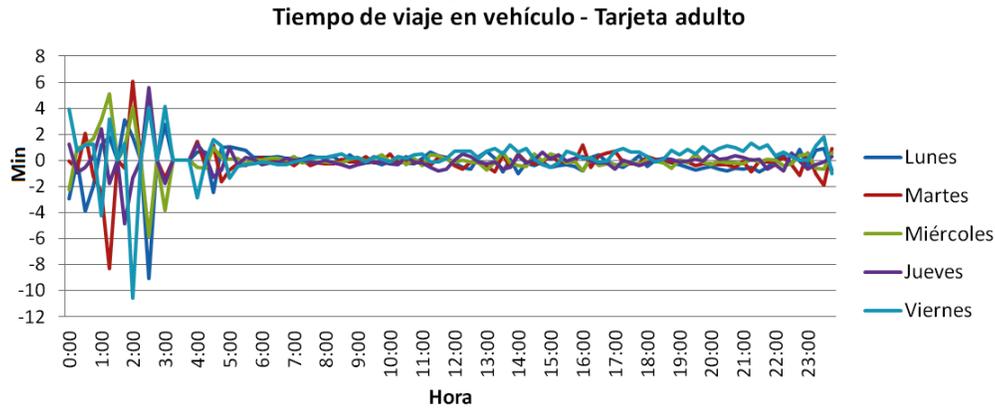
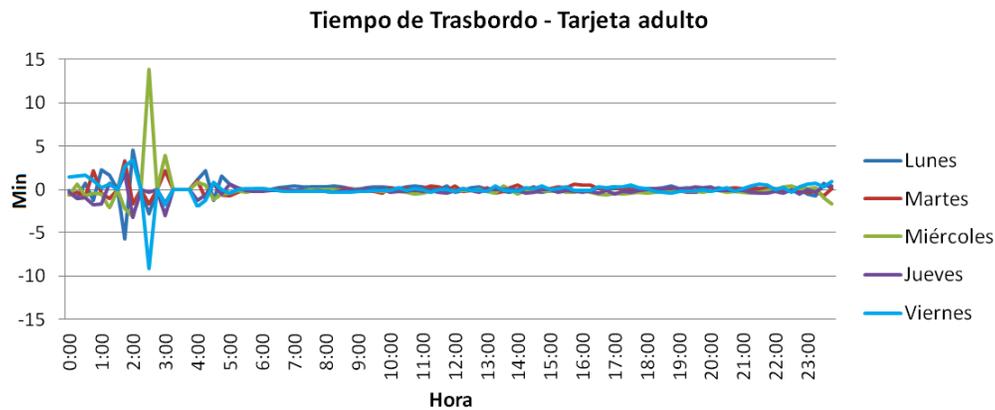


Figura 5.17: Comparación entre días hábiles de variación del tiempo de trasbordo(min), tarjeta adulto.



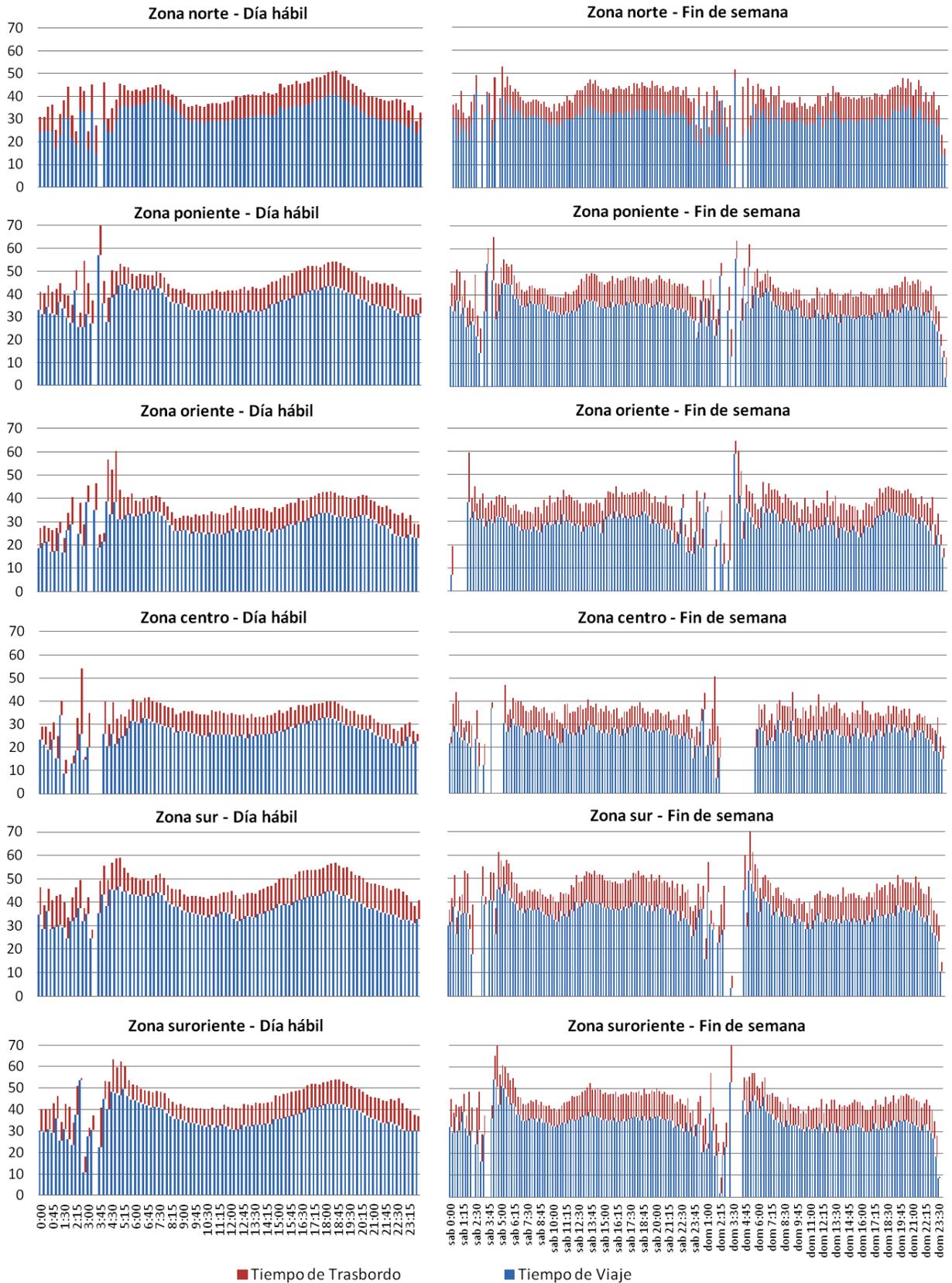
Considerando el resultado anterior, a continuación se presentan los tiempos de viaje en vehículo y traslado por zona de residencia, agrupados en día hábil, sábado y domingo.

La figura ?? muestra apiladas las variables tiempo de viaje en vehículo y traslado por zona, según hora de inicio del viaje. Como se mencionó, se mantiene el efecto de tiempos de viaje mayores en horarios punta los días hábiles, a diferencia del fin de semana en que el tiempo de viaje es más bien estable. El tiempo de traslado es menor en la punta mañana, coincidiendo con el periodo en que la frecuencia de transporte es mayor. La única excepción es la zona centro que presenta el comportamiento inverso, con menores tiempos de traslado en la punta tarde.

Las zonas donde habitan quienes destinan menos tiempo a viajar son oriente y centro, que en los horarios punta alcanzan los 40 minutos promedio, mientras que en los horarios valle oscilan entre los 30 y 35 minutos. Las zonas con peores tiempos son Suroriente, Sur y Poniente, con tiempos de entre 50 y 55 minutos de viaje promedio en hora punta. Los fines

de semana se observa diferencias similares en los tiempos.

Figura 5.18: Tiempo promedio de viaje en vehículo y trasbordo, tarjeta adulto, por zona.



La Tabla ?? muestra los promedios de tiempo de viaje en vehículo y tiempo de trasbordo por tipo de día y zona de residencia de los usuarios.

Se observan diferencias de más de 10 minutos en el tiempo de viaje promedio. Las zonas con menores tiempos de viaje en vehículo son centro y oriente, que bordean los 30 minutos en días hábiles. Los peores tiempos se presentan en las zonas sur, poniente y suroriente que borden los 40 minutos. El fin de semana los tiempos de viaje caen, especialmente el día domingo. La excepción es la zona oriente que presenta tiempos de viaje similares sin importar día de la semana.

El tiempo de trasbordo va de los 7,7 a los 9,2 minutos en día hábil. Los fines de semana este tiempo aumenta, alcanzando los 10,3 minutos en la zona suroriente.

Tabla 5.5: Tiempo de viaje en vehículo y de trasbordo por zona, tarjetas tipo adulto.

Variable	Zona	Día Hábil	Sábado	Domingo
Tiempo de Viaje [min]	Norte	35,4	31,8	30,5
	Poniente	39,6	35,6	32,5
	Oriente	30,4	29,7	30,0
	Centro	29,1	27,0	25,8
	Sur	41,0	37,5	34,1
	Suroriente	39,5	36,2	33,3
Tiempo de Traslado [min]	Norte	8,0	9,1	9,6
	Poniente	8,5	9,2	9,4
	Oriente	7,7	9,2	9,5
	Centro	7,9	8,9	9,2
	Sur	9,2	10,1	10,2
	Suroriente	8,8	10,3	10,2
Tiempos Viaje + Traslado [min]	Norte	43,4	40,9	40,1
	Poniente	48,1	44,8	41,9
	Oriente	38,2	38,8	39,6
	Centro	37,0	36,0	35,1
	Sur	50,1	47,6	44,3
	Suroriente	48,3	46,5	43,5

La zona sur presenta los mayores tiempos promedio de viaje y trasbordo en día laboral. Si se considerará la suma de ambos tiempos, esta misma zona presenta los peores índices en cualquier tipo de día. En contraposición, la zona centro presenta los mejores tiempos, seguida por la zona oriente. En general, las zonas con mejores y peores indicadores son casi las mismas para ambos tiempos, lo que genera que al sumarlos se acrecenten las diferencias de tiempos entre zonas, lo que en esta semana de datos supera los 13 minutos de diferencia entre la zona centro y sur en días hábiles.

Aunque es claro que existe relación entre lugar de residencia y tiempo de viaje, usualmente sólo se estudia la variación en los tiempos de viaje por origen-destino. La información expuesta

en esta sección cuantifica el impacto de la elección de lugar de residencia en los tiempos promedio de viaje en vehículo y trasbordo, por cada viaje realizado. Podría usarse como medida de accesibilidad de una zona o de calidad de servicio del transporte público. Además, este análisis podría ser relevante en la evaluación de políticas de vivienda social y planificación de la ciudad.

5.3.2. Tiempo de viaje en vehículo y de trasbordo de tarjetas tipo estudiante

Esta sección presenta los resultados de tiempo de viaje en vehículo y tiempo de trasbordo para la tarjeta tipo estudiante. Para estudiar estas variables, las siguientes figuras muestran las distribuciones de tiempo de viaje en vehículo y tiempo de trasbordo. Además se incluye la desviación estándar en torno al promedio de los tiempos. En la madrugada la varianza es nula o inestable debido a que existen pocos o ningún viaje en este periodo. En el eje horizontal se muestra el día y hora y en el eje vertical el tiempo en minutos.

Nuevamente se destaca la similitud en la distribución de los tiempos entre días hábiles y las diferencias con los fines de semana. Se repite el efecto de tiempo de viaje en vehículo mayor en la madrugada. A diferencia de la tarjeta adulto, los tiempos de viaje en punta tarde son bastante menores que en punta mañana. Los fines de semana no se aprecian alzas considerables en el tiempo por efecto de la punta tarde. El tiempo de trasbordo es menor en punta mañana (cuando hay mayor frecuencia) y se mantiene relativamente estable a lo largo del día.

Figura 5.19: Tiempo de viaje en vehículo y desviación estándar, tarjeta estudiante.

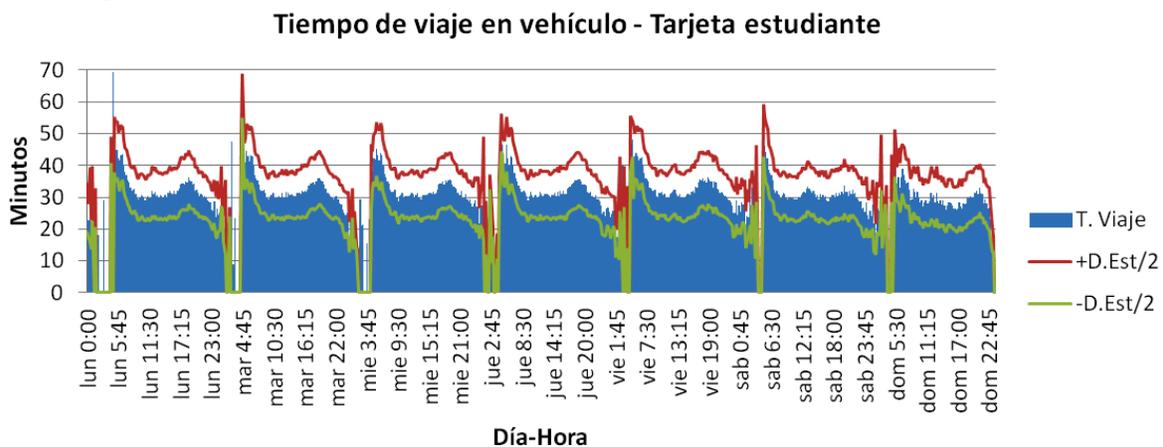
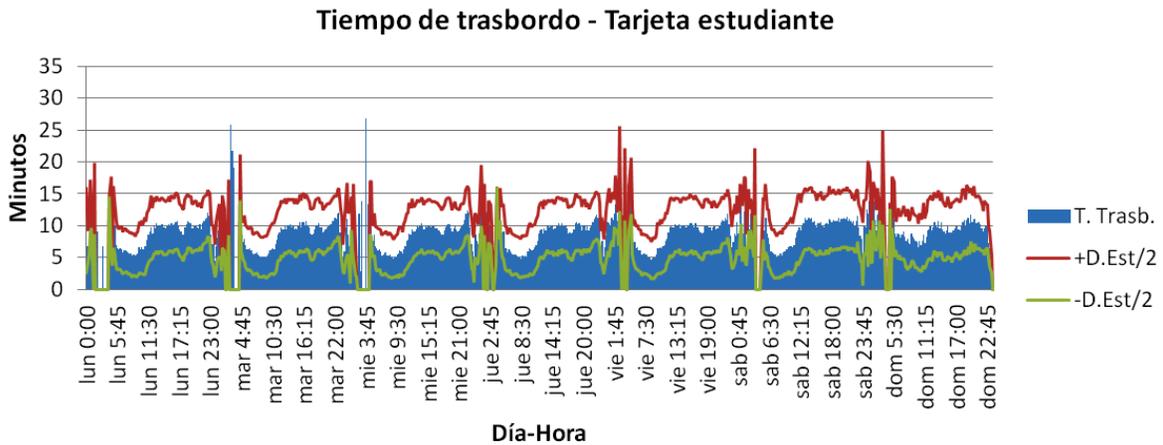


Figura 5.20: Tiempo de trasbordo y desviación estándar, tarjeta estudiante.



La Tabla ?? incluye los tiempos promedio diarios de las variables por día, en minutos. Se puede ver que, al igual que para las tarjetas adulto, se observan menores tiempos de viaje en vehículo los fines de semana, especialmente el domingo. El tiempo de trasbordo se comporta de manera inversa mostrando menores tiempos promedio los días hábiles.

Tabla 5.6: Tiempo promedio de viaje en vehículo y trasbordo, tarjetas tipo estudiante.

	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo
T. viaje [min]	33,4	33,3	33,4	33,3	33,6	31,0	29,2
T. trasbordo [min]	8,1	7,9	7,8	7,8	7,9	9,0	9,4
Suma [min]	41,5	41,3	41,2	41,2	41,5	39,9	38,6

Para corroborar la hipótesis de que las distribuciones de tiempo en los cinco días hábiles son las mismas, se propone promediar las distribuciones de tiempo de los cinco días hábiles y observar las diferencias de cada día respecto al promedio:

Figura 5.21: Comparación entre días hábiles de variación del tiempo de viaje en vehículo(min), tarjeta estudiante.

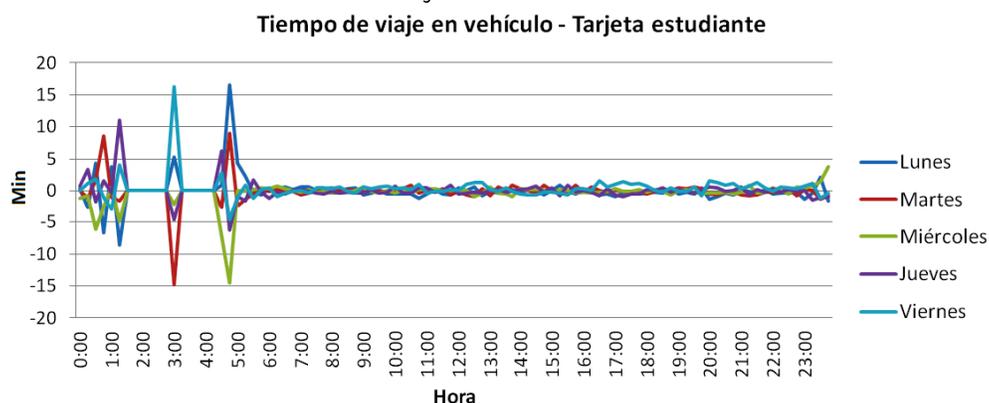
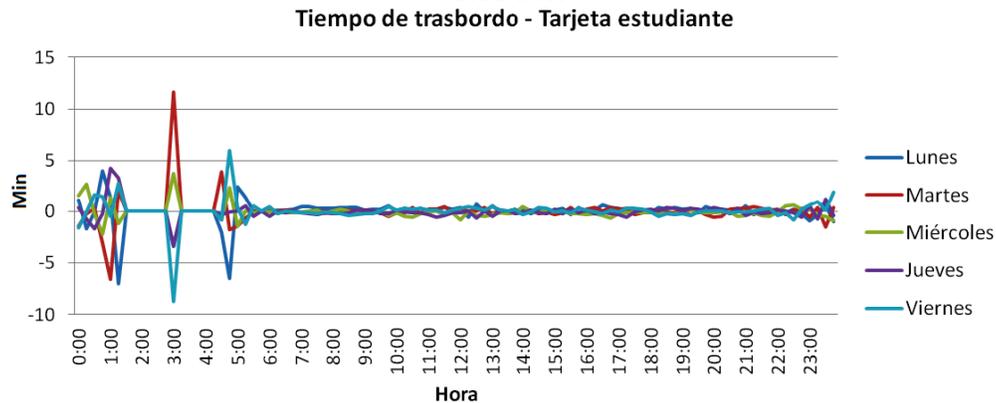


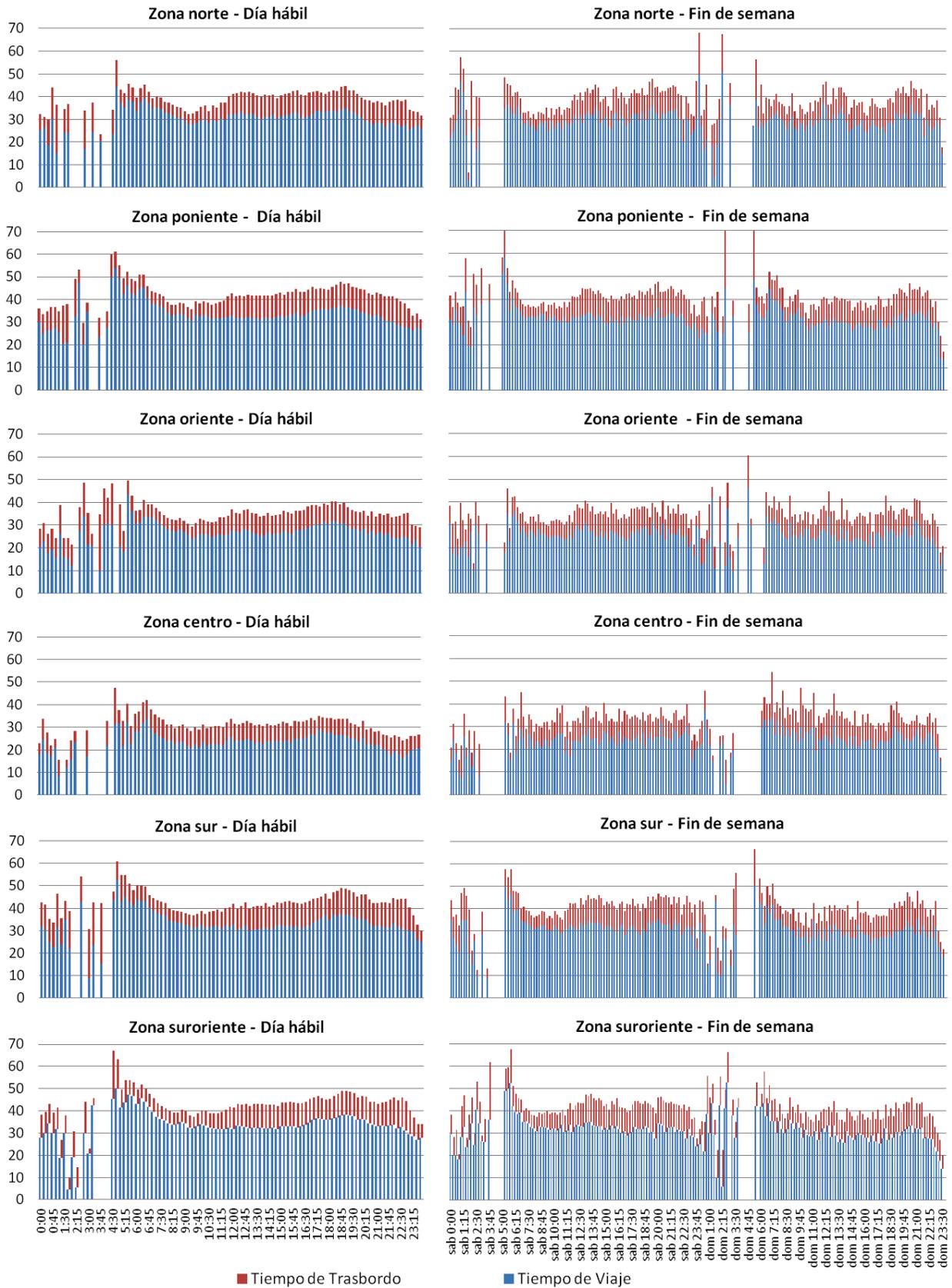
Figura 5.22: Comparación entre días hábiles de variación del tiempo de trasbordo (min), tarjeta estudiante.



Los cinco días hábiles se comportan de manera similar, salvo en las horas de la madrugada donde existe mucha varianza debido a los pocos datos de tiempo registrados en este horario y a la baja frecuencia de los servicios. Sin embargo, en otros horarios no se aprecia diferencias entre días hábiles. Por lo tanto, los tiempos de lunes a viernes se presentan en un único gráfico para día laboral promedio.

La figura ?? muestra los tiempos de viaje en vehículo y trasbordo por zona de la ciudad, agrupados en día hábil, sábado y domingo. Cabe recordar que se muestran las variables para quienes habitan en cada zona, independientemente del origen y destino de sus viajes. Los mayores tiempos de viaje se observan en la madrugada, previo a la punta mañana. Los menores se observan después de la hora punta mañana y desde las 21:30 hrs. Se repite el efecto de tiempos de viaje mayores en horarios punta los días hábiles, a diferencia del fin de semana en que el tiempo de viaje es más bien estable, salvo en la madrugada. El tiempo de trasbordo es menor en punta mañana, lo que típicamente corresponde al periodo en que la frecuencia del transporte es mayor.

Figura 5.23: Tiempo de viaje en vehículo y trasbordo, tarjeta estudiante, por zona.



Otro efecto que se repite es la diferencia en los tiempos por zona de residencia. La Tabla ?? muestra los promedios de tiempo de viaje en vehículo y trasbordo por día y zona de la ciudad.

Las zonas con menores tiempos de viaje en vehículo son centro y oriente, con 24,9 y 28,3 minutos en días hábiles, respectivamente. Los peores tiempos se presentan en las zonas suroriente (35,5), sur y poniente (34,8 en ambos casos). El fin de semana los tiempos de viaje son menores, especialmente el día domingo.

El tiempo de trasbordo va de los 7,2 a los 8,3 minutos en día hábil. Los fines de semana estos tiempos aumentan alcanzando los 9,6 minutos en las zonas sur y suroriente. Las zonas con mayor tiempo son las mismas que en el caso anterior, pero la zona con menor tiempo de trasbordo es la zona norte, desplazando a oriente y centro del primer lugar.

Tabla 5.7: Tiempo de viaje en vehículo y de trasbordo por zona, tarjetas tipo estudiante.

Variable	Zona	Día Hábil	Sábado	Domingo
Tiempo de Viaje [min]	Norte	32,2	29,9	29,3
	Poniente	34,8	32,1	30,8
	Oriente	28,3	26,3	25,7
	Centro	24,9	24,7	25,0
	Sur	34,8	32,1	29,6
	Suroriente	35,5	32,6	29,9
Tiempo de Traslado [min]	Norte	7,2	8,4	9,1
	Poniente	7,9	8,7	9,3
	Oriente	7,3	8,4	9,3
	Centro	7,4	9,0	9,1
	Sur	8,3	9,2	9,6
	Suroriente	8,3	9,6	9,6
Tiempos Viaje + Traslado [min]	Norte	39,5	38,4	38,4
	Poniente	42,7	40,7	40,1
	Oriente	35,6	34,8	35,0
	Centro	32,3	33,7	34,1
	Sur	43,1	41,3	39,3
	Suroriente	43,9	42,2	39,5

En general, los tiempos son menores para la tarjeta estudiante respecto a la tarjeta adulto. Sin embargo, se conservan diferencias de tiempo entre zonas que llegan a los 11 minutos entre centro y suroriente. Sería interesante analizar si estas diferencias son independientes de los distintos niveles de enseñanza.

Para ver el código (c++) aplicado a las bases de datos para obtener estos resultados, consultar Anexo D.

5.4. Perfil de uso de tiempo en Santiago

Utilizando el método de estimación de zona de residencia descrito en el capítulo 4 y el método de asignación de propósito de viaje descrito en el capítulo 2, la presente sección muestra los resultados de perfiles de uso de tiempo por propósito y zona en la ciudad para la semana de abril-2012. Este análisis se realiza con una muestra de 591.976 tarjetas a las que se pudo estimar residencia según lo descrito en el capítulo 4. Los datos permiten asignar el tiempo de los usuarios a las siguientes actividades: *Hogar*, *Trabajo*, *Estudio*, *Otros* y *Tiempo de Viaje*.

Antes de mostrar los perfiles, se analiza la forma de presentación de los datos. Es claro que los días hábiles, sábado y domingo tienen perfiles de uso de tiempo distintos entre sí, por lo que serán presentados de manera independiente. Algo menos obvio es si existen diferencias en el uso de tiempo de lunes a viernes. Para determinar si existe diferencias en la distribución de uso de tiempo entre días hábiles se propone normalizar las distribuciones, calcular una distribución promedio de los 5 días y observar las diferencias de cada día respecto al promedio. De esta forma, si un día de la semana se aleja fuertemente de las otras curvas en algún horario, significará que la proporción de usuarios realizando esa actividad ese día y en ese horario es diferente al promedio. Además, se incluye la franja que encierra +/- la desviación estándar de los datos (DE).

Las Figuras ?? a ?? muestran los resultados separados por actividad.

Para el propósito hogar, hasta las 14:30 hrs las tarjetas muestran un comportamiento similar de lunes a viernes. A partir de esta hora el día viernes escapa a la franja de +/- la desviación estándar y muestra diferencias de comportamiento, ya que la proporción de tarjetas que se encuentran en el hogar se incrementa respecto el promedio lunes-jueves. Desde las 19:45 hrs se invierten roles y se aprecia que el día viernes se encuentra bajo el promedio, indicando que a partir de esta hora el retorno al hogar es más lento (o que luego de esta hora se realizan otras actividades) respecto a lo observado de lunes a jueves.

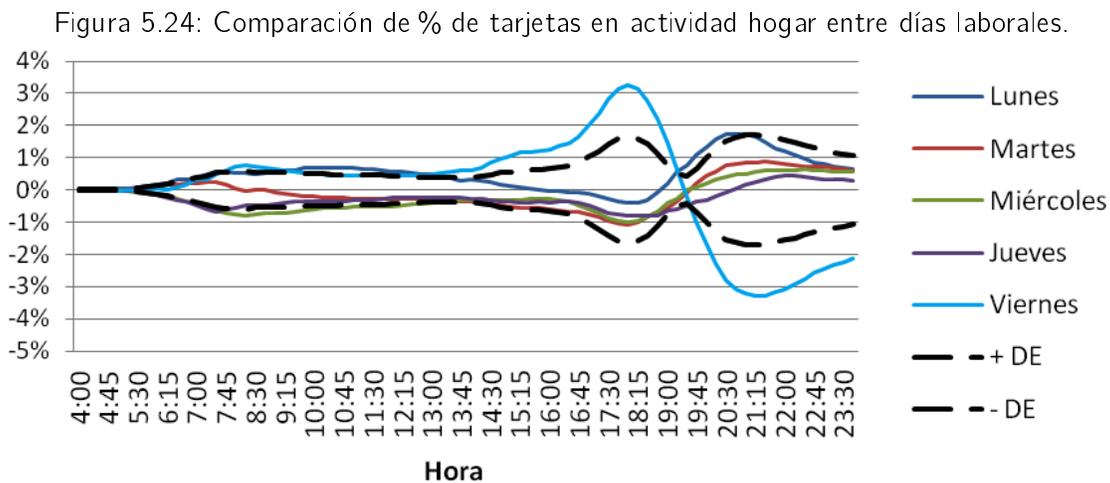
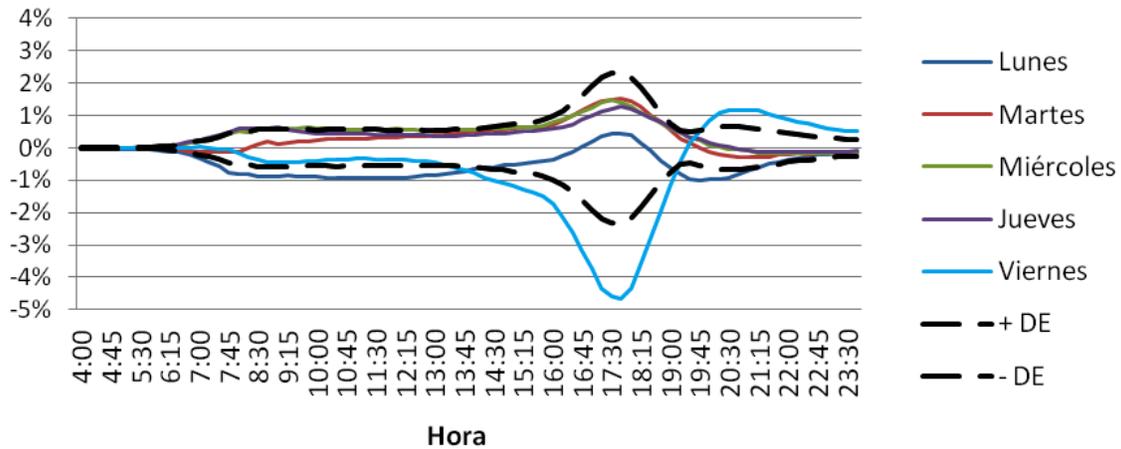


Figura 5.25: Comparación de % de tarjetas en actividad trabajo entre días laborales.

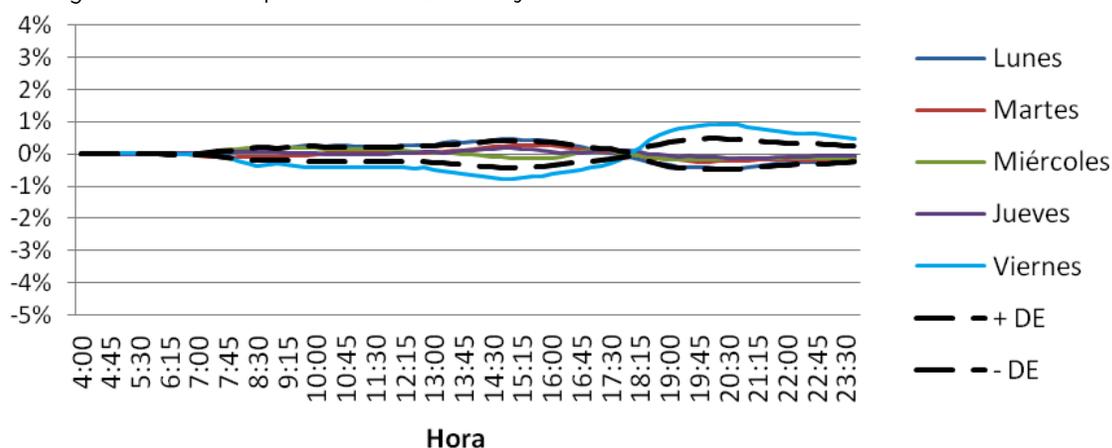


La comparación entre días hábiles para actividad trabajo también muestra diferencias el día viernes. La proporción de tarjetas que asignan su tiempo a la actividad trabajo es menor al promedio lunes-jueves a partir de las 14:00 y hasta las 19:15 hrs. Es decir, este día la gente abandona su sitio de trabajo más temprano.

En el caso de la comparación entre días hábiles para la actividad estudio, el día viernes es distinto casi en su totalidad (desde las 8:15 hrs), aunque las diferencias se incrementan a partir de las 14:00 hrs en que la curva abandona los límites de la franja +/- desviación estándar. Hasta las 18:00 hrs la actividad estudio el día viernes se realiza bajo el promedio lunes-jueves y luego de esta hora sobre el mismo. Parece extraño que a partir de las 18:00 hrs. del día viernes la actividad estudio se realice en mayor proporción que los días lunes a jueves.

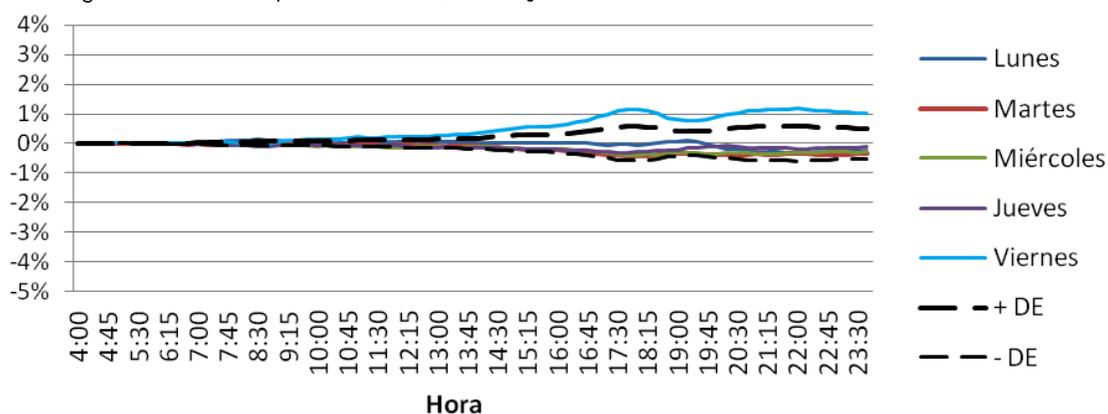
Es posible que las reglas para determinar propósito de viaje no representen adecuadamente la asignación de tiempo el día viernes. Parece más sensato pensar que este día a partir de las 18:00 hrs se asigna el tiempo a la actividad otros antes que estudio, incluso si la duración de la actividad supera las 2:00 hrs, como indica el método.

Figura 5.26: Comparación de % de tarjetas en actividad estudio entre días laborales.



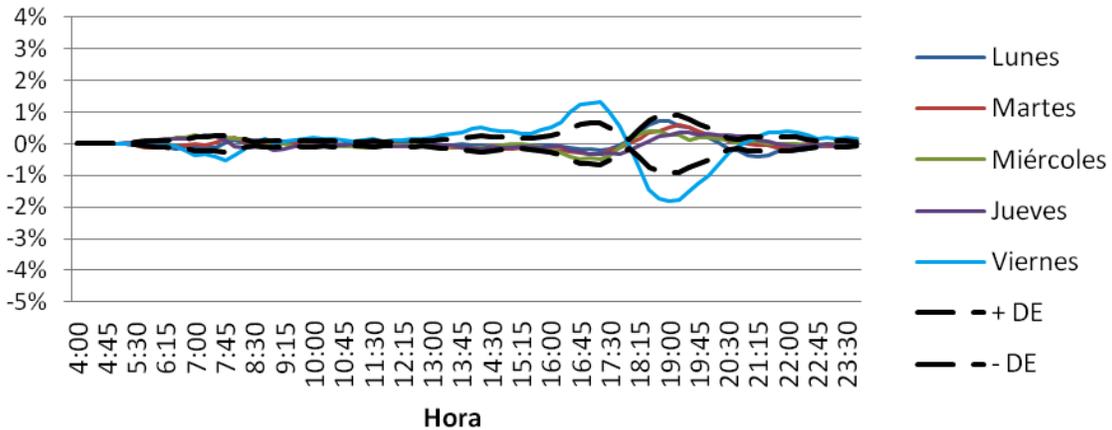
Aunque con menor variabilidad, la actividad otros también muestra diferencias sustanciales el día viernes a partir de las 13:00 hrs. La proporción de tarjetas en esta actividad es superior al promedio lunes-jueves, alcanzando la mayor diferencia a las 22:00 hrs.

Figura 5.27: Comparación de % de tarjetas en actividad otros entre días laborales.



La actividad viaje también se inicia más temprano el día viernes (14:00 a 14:15 hrs), siendo consistente con el retorno al hogar que inicia más temprano y adelantamiento de la hora punta a las 17:00 hrs, razón por la cual en el horario punta tarde tradicional (de salida del trabajo de lunes a jueves) la asignación de tiempo a la actividad viaje se encuentra por debajo del promedio.

Figura 5.28: Comparación de % de tarjetas en actividad viaje entre días laborales.

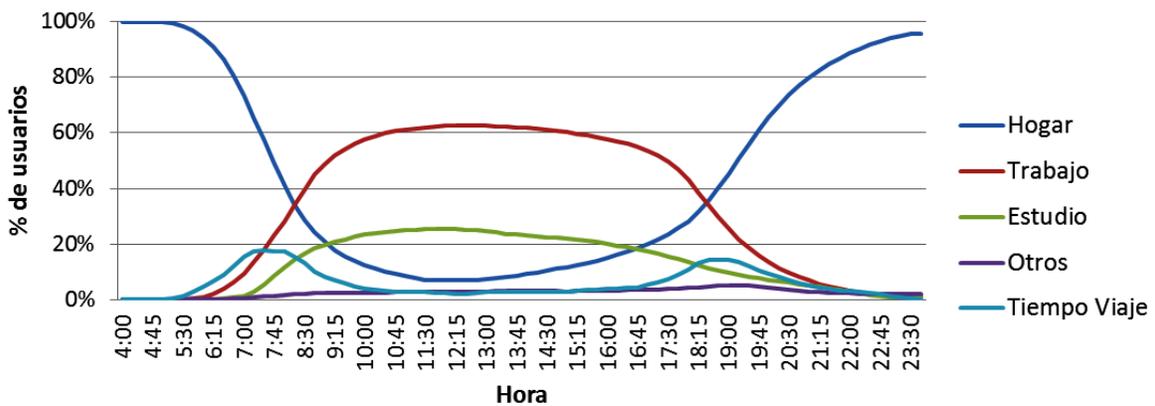


Es decir, para todas las actividades en estudio se observa similitud en el comportamiento lunes-jueves y diferencias el día viernes. Por tanto, el perfil de uso de tiempo se expondrá para el promedio de los días lunes-jueves, además de los días viernes, sábado y domingo.

Las Figuras ?? a ?? muestran el perfil de uso de tiempo del total de tarjetas a las que pudo estimárseles lugar de residencia, según tipo de día. En el eje horizontal se muestra la hora del día y en el eje vertical la proporción de tarjetas dedicadas a cada actividad. La primera figura promedia los días lunes a jueves (por su similitud). Luego se muestran viernes, sábado y domingo.

De lunes a jueves las tarjetas que realizan la actividad trabajo superan el 60% y las que realizan la actividad estudio se acercan al 30%. En el mismo periodo la actividad hogar muestra menos de un 10% de tarjetas en esta ocupación. La actividad otros bordea el 5% en la punta tarde.

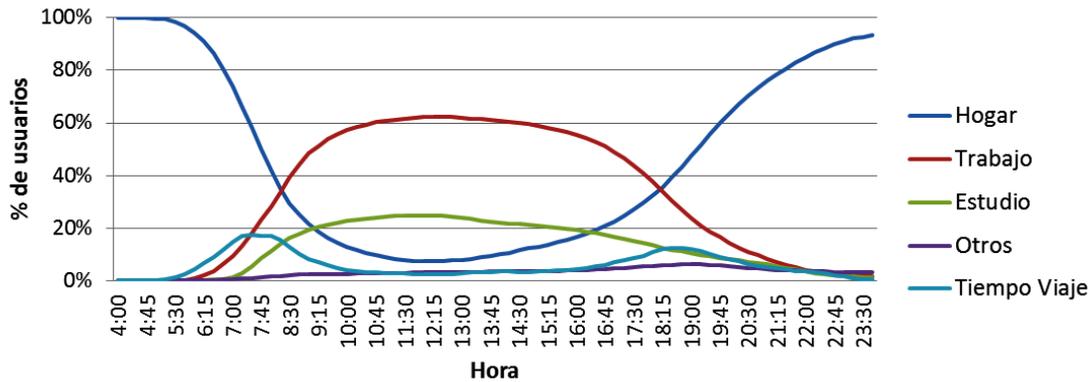
Figura 5.29: Perfil de uso de tiempo, promedio lunes a jueves.



Aunque el perfil del día viernes luce bastante similar al promedio lunes-jueves, tiene todas las diferencias descritas en en las Figuras ?? a ??. Al observar con detención se aprecia un

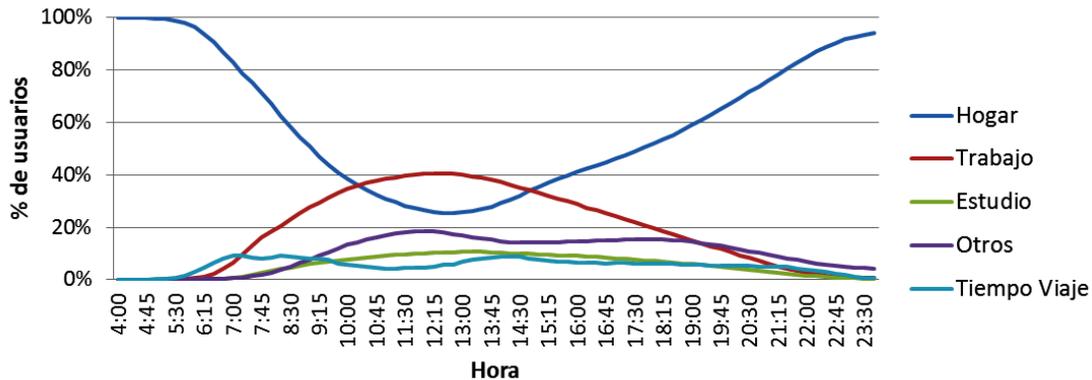
leve incremento en la actividad otros especialmente al final del día. La curva de tiempo de viaje es menos pronunciada en la punta tarde y las curvas de hogar y trabajo son más suaves. Todas estas diferencias son difíciles de distinguir a simple vista porque ocurren en variaciones porcentuales pequeñas, pero muy significativas respecto a la varianza de los datos.

Figura 5.30: Perfil de uso de tiempo, viernes.



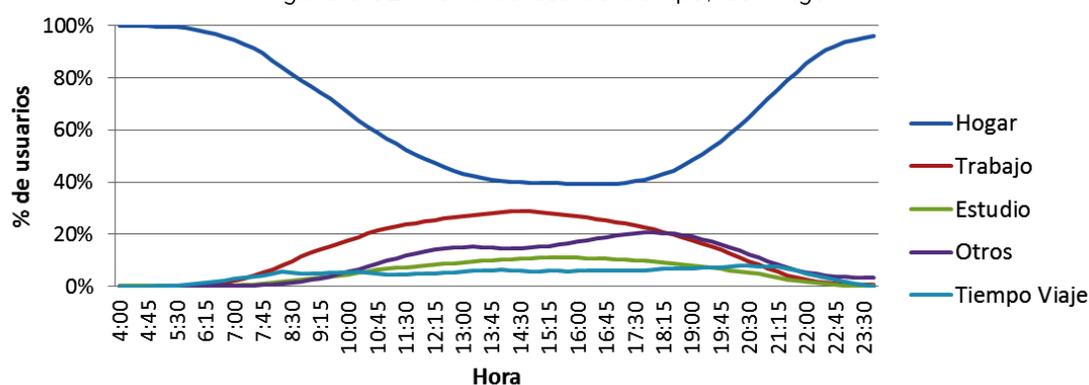
El día sábado las actividades trabajo y estudio reducen notoriamente su proporción, no solo en cuanto a número de tarjetas sino también a duración de la actividad. Aumenta la proporción de tarjetas que dedica su tiempo a los propósitos hogar y otros. La proporción de tarjetas en actividad viaje es más estable a lo largo del día, desapareciendo las marcadas punta mañana y punta tarde.

Figura 5.31: Perfil de uso de tiempo, sábado.



El día domingo se acentúan los efectos mostrados el sábado de reducción de tiempo asignado a trabajo y estudio. La actividad otros alcanza su máxima proporción semanal en torno a las 18:00 hrs, con el 20% de las tarjetas dedicadas a este propósito simultáneamente.

Figura 5.32: Perfil de uso de tiempo, domingo.



A continuación se muestra los perfiles de uso de tiempo por zona de residencia.

5.4.1. Comparación de perfiles de uso de tiempo entre zonas

En esta sección se presenta la comparación de perfiles de uso de tiempo entre zonas de la ciudad por propósito y tipo de día. Para ello se utiliza la zonificación y el método de predicción de zona de residencia expuestos en el capítulo cuatro. Las zonas de residencia son seis: norte, poniente, oriente, centro, sur y suroriente. Para facilitar la comparación, los resultados están normalizados por el número de tarjetas de la zona.

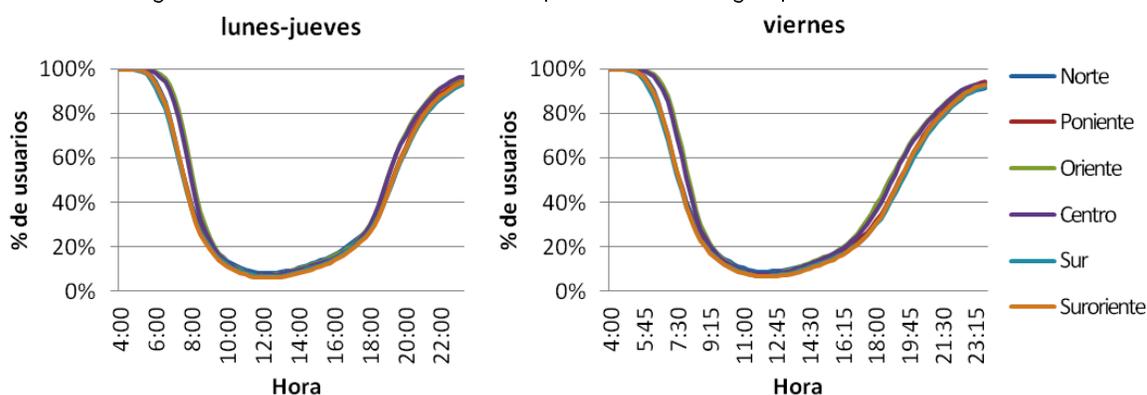
Antes de presentar los perfiles de uso de tiempo se muestra el número de observaciones por zona y tipo de día en la Tabla ???. Las zonas con más observaciones son poniente y suroriente, que coinciden con las de mayor población. Todas las zonas presentan más observaciones para el tipo de día lunes-jueves y menos para el día domingo.

Tabla 5.8: Número de observaciones por zona y tipo de día.

Tipo día	Norte	Poniente	Oriente	Centro	Sur	Suroriente	Observaciones
Lunes-Jueves	43.595	112.379	58.314	35.210	63.861	95.648	409.007
Viernes	40.838	105.372	54.760	33.448	60.405	90.472	385.295
Sábado	20.704	48.266	18.147	13.236	27.464	38.104	165.921
Domingo	8.503	20.220	8.679	6.686	10.865	16.071	71.024
Observaciones	34.918	89.053	44.977	27.744	50.597	75.320	322.609
% del total	10,8	27,6	13,9	8,6	15,7	23,3	100

En la Figura ?? se observa el perfil de uso de tiempo de la actividad hogar para las seis zonas de la ciudad. Aunque las curvas presenten formas casi idénticas, las zonas centro y oriente dejan el hogar después de las otras zonas y retornan al hogar antes que ellas. El fenómeno se repite el día viernes y en la punta tarde es más marcado en comparación al promedio lunes-jueves.

Figura 5.33: Perfil de uso de tiempo actividad hogar por zona, día laboral.

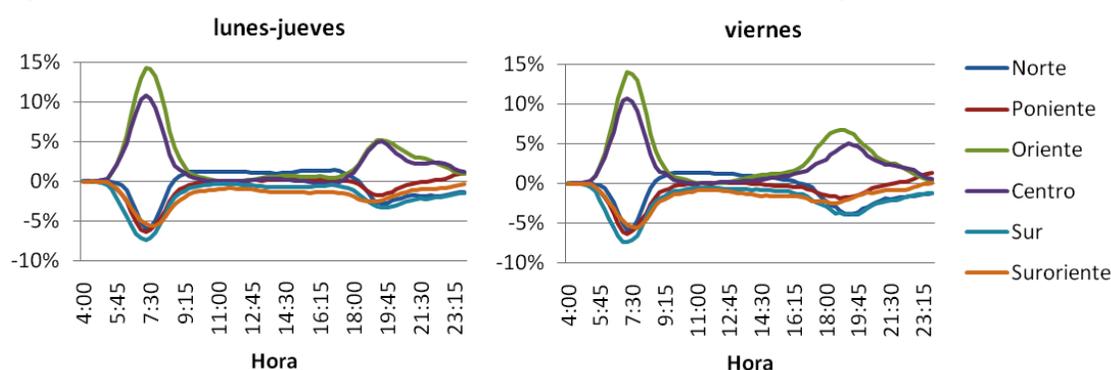


Para observar diferencias de comportamiento entre zonas, que es difícil de distinguir en este gráfico, se repite el proceso aplicado entre días hábiles, pero ahora para graficar las diferencias de cada zona respecto al promedio de las seis zonas.

La Figura ?? muestra la comparación entre zonas para el propósito hogar. Aquí se aprecia con mayor claridad que las tarjetas de las zonas oriente y centro asignan tiempo al hogar sobre el promedio de las otras zonas entre las 6:15 y 8:45 hrs. Esto se repite entre las 18:00 y 23:00 hrs.

El día viernes ocurre algo semejante, con la salvedad de que en la tarde se aprecian diferencias a partir de las 16:30. Es decir, de lunes a viernes las zonas centro y oriente dejan el hogar más tarde y llegan más temprano que las otras zonas. Notar que entre 9:00 y 18:00 hrs la misma proporción de tarjetas está dedicado a la actividad hogar en cada zona.

Figura 5.34: Comparación de perfiles de uso de tiempo actividad hogar por zona, día laboral.



Como se puede ver en la Figura ??, el día sábado es particularmente distinto porque la diferencia no es sólo en la hora de dejar el hogar, además entre las 5:45 y 15:00 hrs hay una proporción mayor de personas dedicadas a la actividad hogar en las zonas oriente y centro. Algo similar ocurre el día domingo, principalmente para la zona centro.

Figura 5.35: Perfil de uso de tiempo actividad hogar por zona, fin de semana.

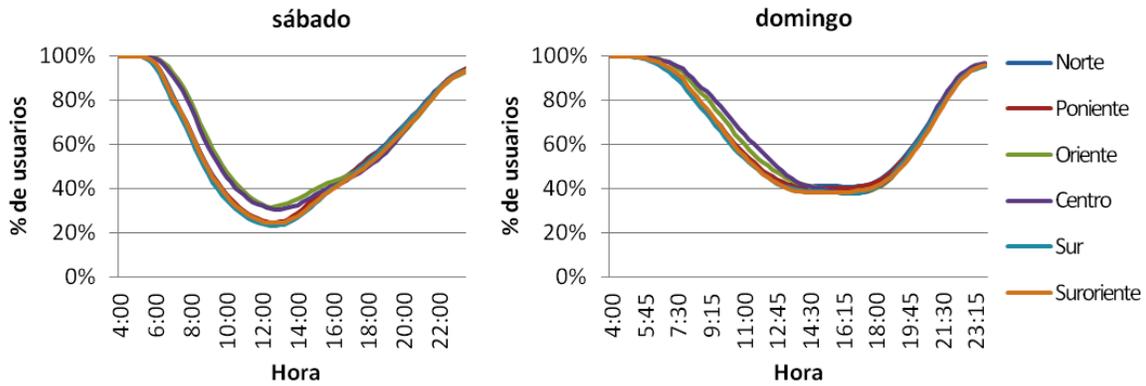
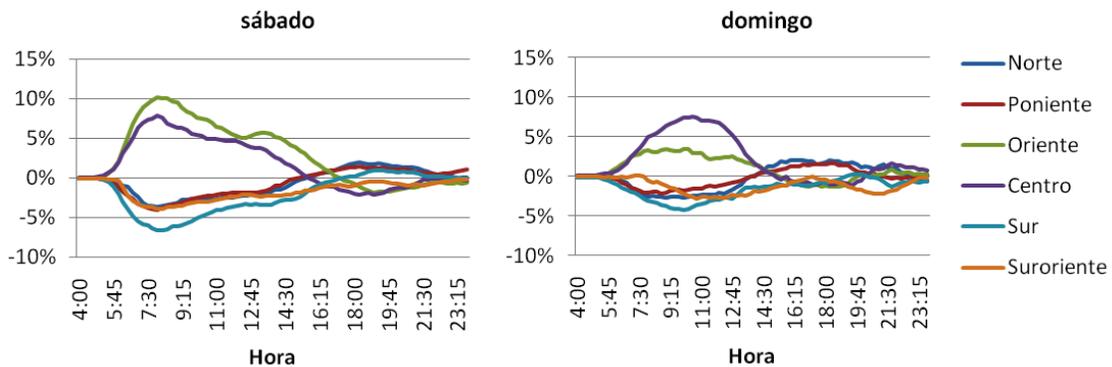
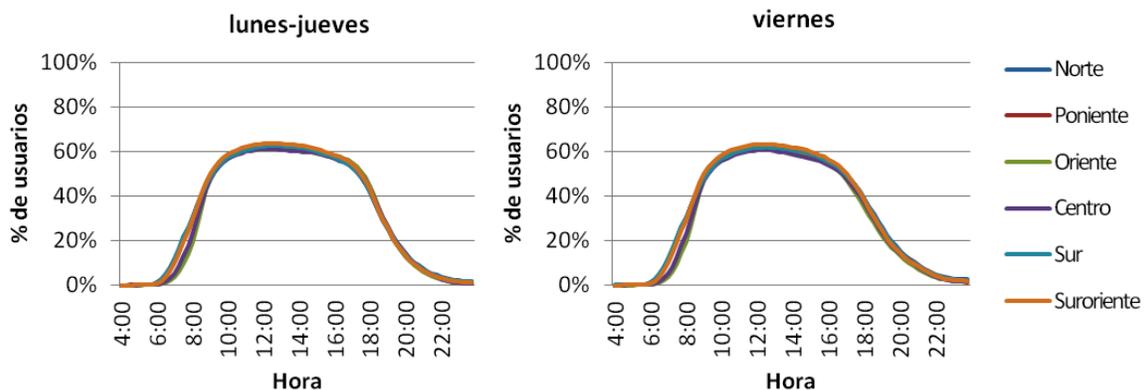


Figura 5.36: Comparación de perfiles de uso de tiempo actividad hogar por zona, fin de semana.



La Figura ?? presenta los resultados para el propósito trabajo. El comportamiento es similar entre zonas de lunes a viernes. En todas las zonas la asignación de tiempo a esta actividad alcanza el 60% de las tarjetas. Las diferencias se concentran en punta mañana, donde la proporción de personas de las zonas centro y oriente en actividad trabajo están bajo el promedio.

Figura 5.37: Perfil de uso de tiempo actividad trabajo por zona, día laboral.



Al comparar la asignación de tiempo a trabajo entre zonas (ver Figura ??), se aprecia la ventana horaria en que centro y oriente están bajo el promedio: de 6:00 a 9:00 hrs. El día viernes presenta el mismo comportamiento. Esto quiere decir que centro y oriente comienzan horario laboral más tarde, presentando la misma proporción de tarjetas en esta actividad a partir de las 9:00 hrs.

Figura 5.38: Comparación de perfiles de uso de tiempo actividad trabajo por zona, día laboral.

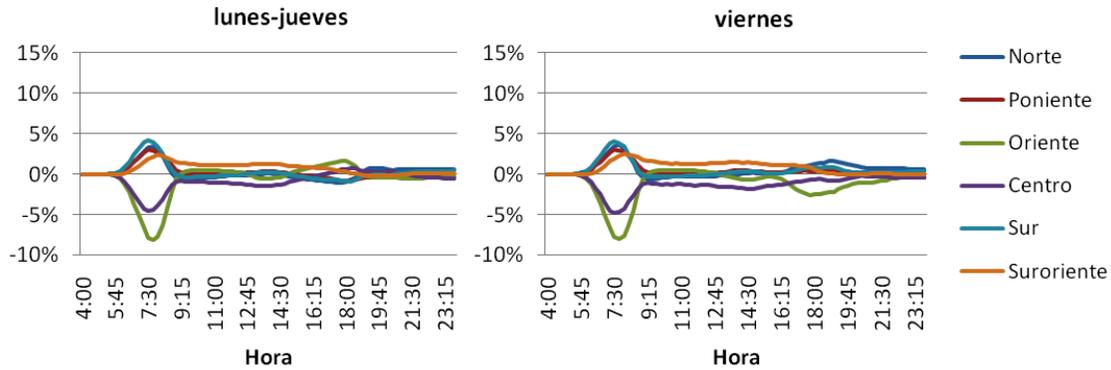
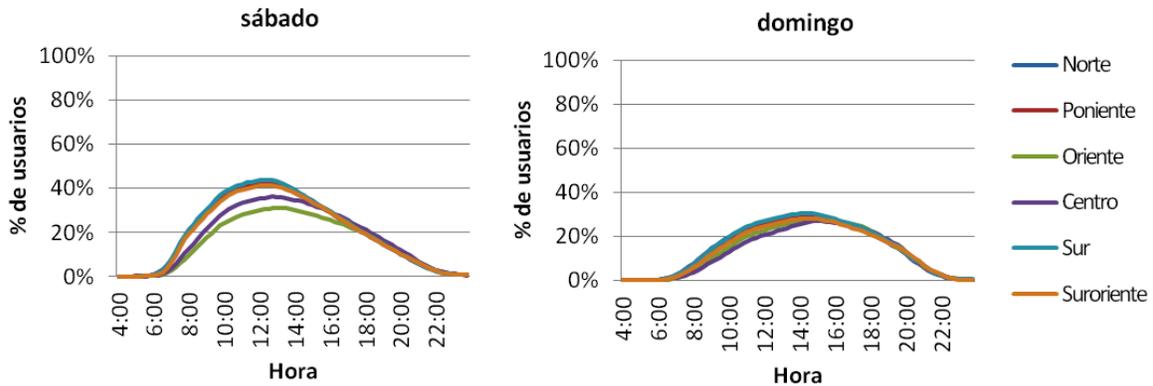
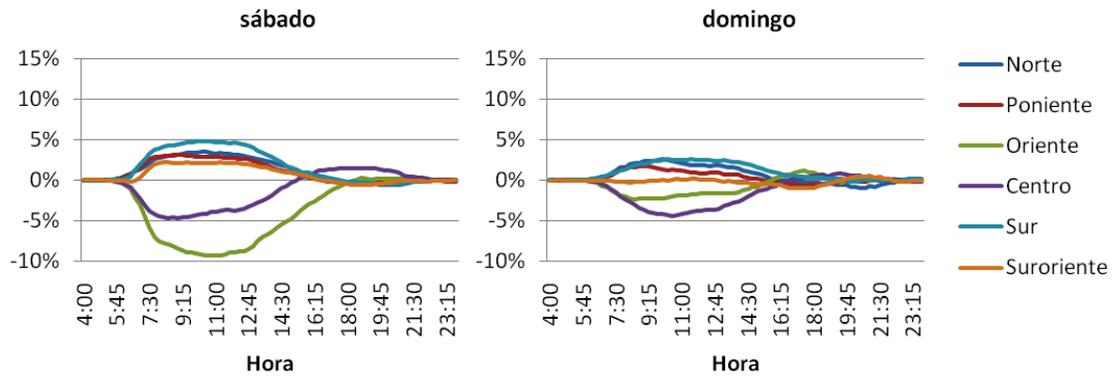


Figura 5.39: Perfil de uso de tiempo actividad trabajo por zona, fin de semana.



El día sábado no solo existe diferencias en el horario de inicio de la actividad, además es menor la proporción de personas de las zonas centro (-8%) y oriente (-12%) que dedican tiempo a la actividad trabajo. Se aprecia el mismo efecto el día domingo, aunque en menor proporción.

Figura 5.40: Comparación de perfiles de uso de tiempo actividad trabajo por zona, fin de semana.



Las Figuras ?? y ?? presentan los resultados para el propósito estudio en día hábil. De lunes a viernes la zona oriente comienza la actividad estudio después que el resto de las zonas y termina antes que éstas. Por otra parte, la proporción de tarjetas de la zona centro que asignan tiempo a la actividad estudio es superior a todas las otras zonas.

Figura 5.41: Perfil de uso de tiempo actividad estudio por zona, día laboral.

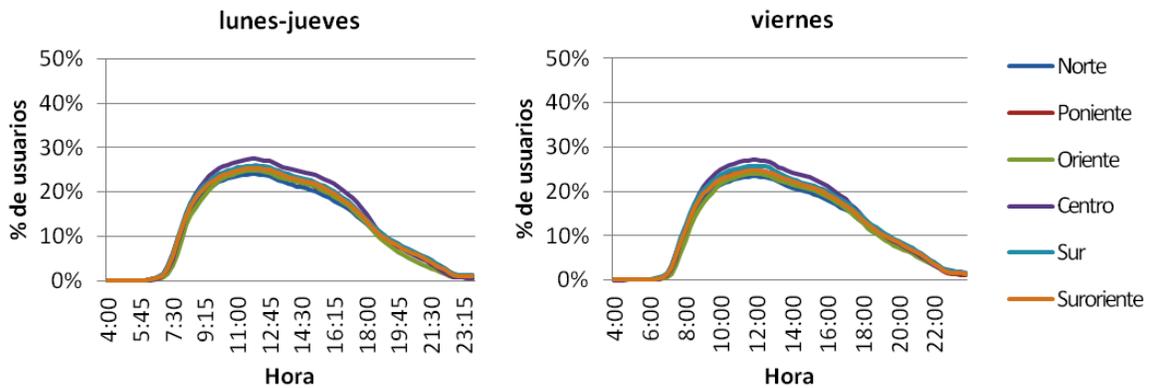
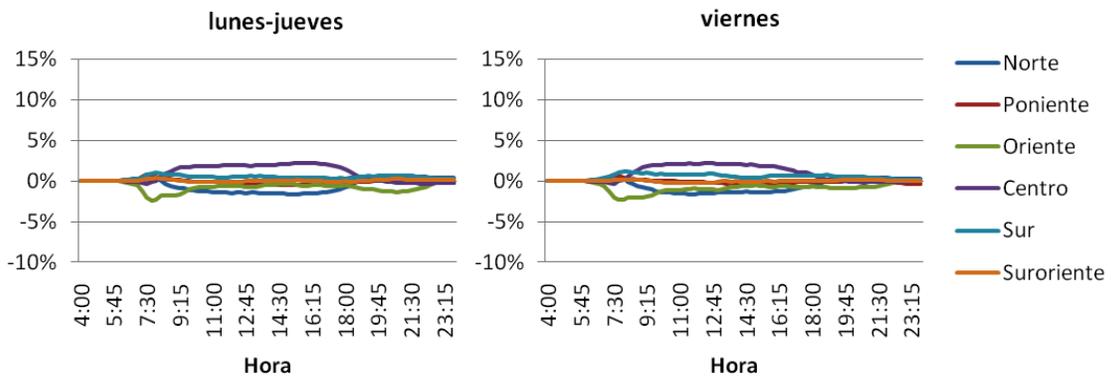


Figura 5.42: Comparación de perfiles de uso de tiempo actividad estudio por zona, día laboral.



El fin de semana la actividad estudio presenta importantes diferencias entre sábado y

domingo, aunque pequeñas diferencias entre zonas. El día sábado la zona oriente tiene una menor proporción de tarjetas que asignan su tiempo a la actividad estudio entre 7:30 y 14:00 hrs. El día domingo aparece la zona suroriente sobre el promedio de las otras zonas entre 11:00 y 21:00 hrs.

Este comportamiento es algo extraño y puede estar vinculado a una errónea asignación de propósito, particularmente porque el tiempo asignado a la actividad estudio (fuera del hogar) durante el día domingo supera al día sábado. Es probable que el domingo se observen actividades de más de cinco horas que estén dedicadas a la actividad otros, pero que el método asigna a la actividad estudio.

Figura 5.43: Perfil de uso de tiempo actividad estudio por zona, fin de semana.

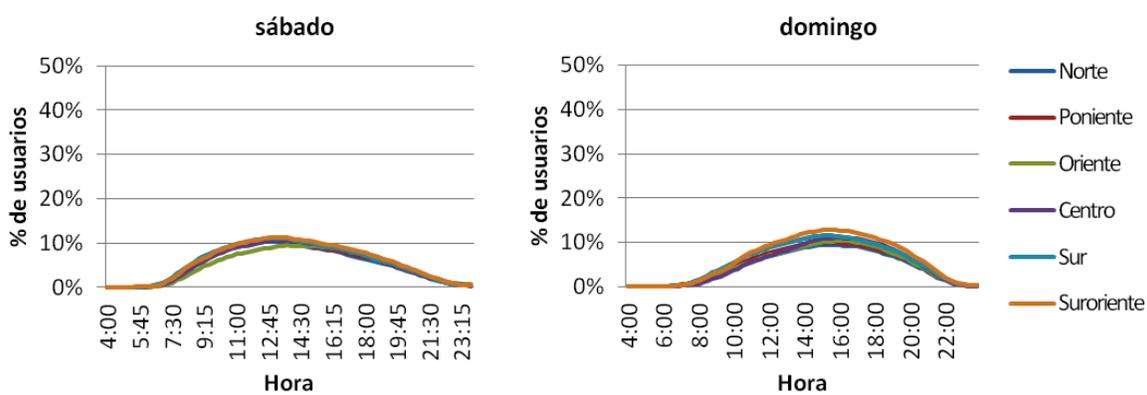
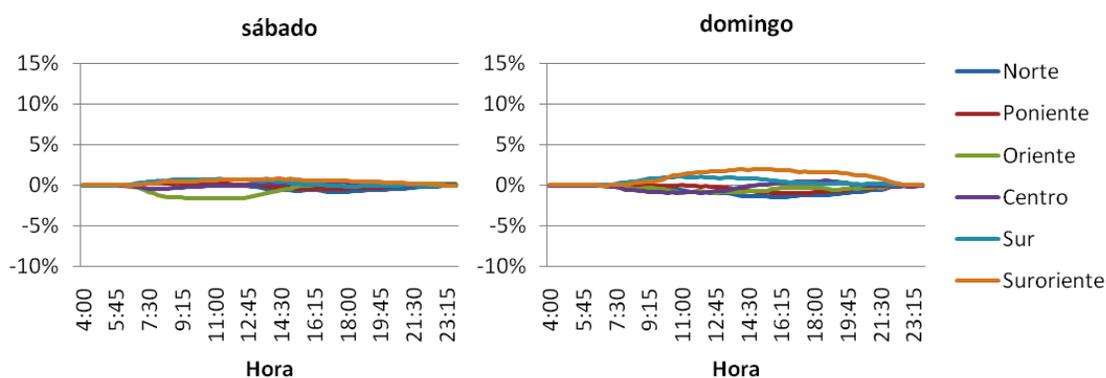


Figura 5.44: Comparación de perfiles de uso de tiempo actividad estudio por zona, fin de semana.



A continuación se presentan los resultados para el propósito otros. De lunes a viernes la actividad presenta curvas similares. Sin embargo, el día viernes esta actividad comienza más temprano, termina más tarde y alcanza una proporción mayor de tarjetas. Las zonas centro y oriente son las únicas que muestra un incremento de esta actividad en el horario de almuerzo. A las 19:00 hrs se observa la mayor proporción de tarjetas en esta ocupación.

Figura 5.45: Perfil de uso de tiempo actividad otros por zona, día laboral.

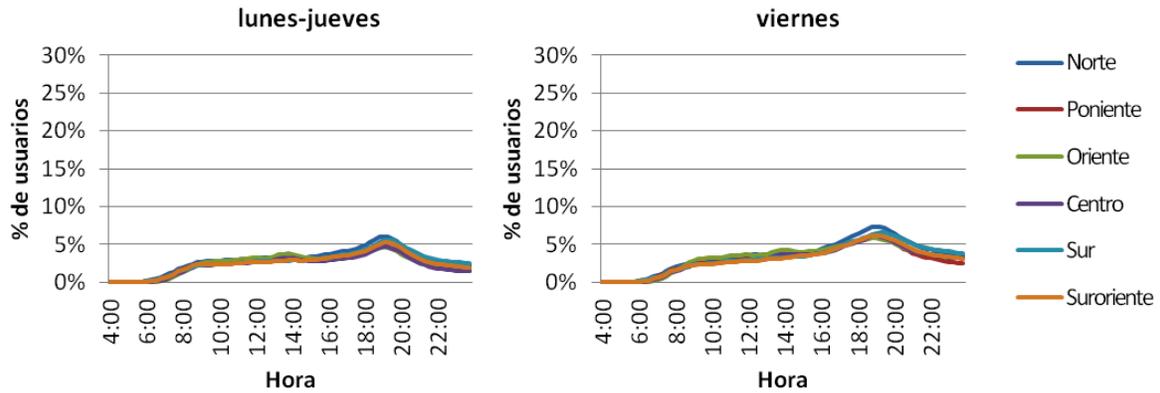
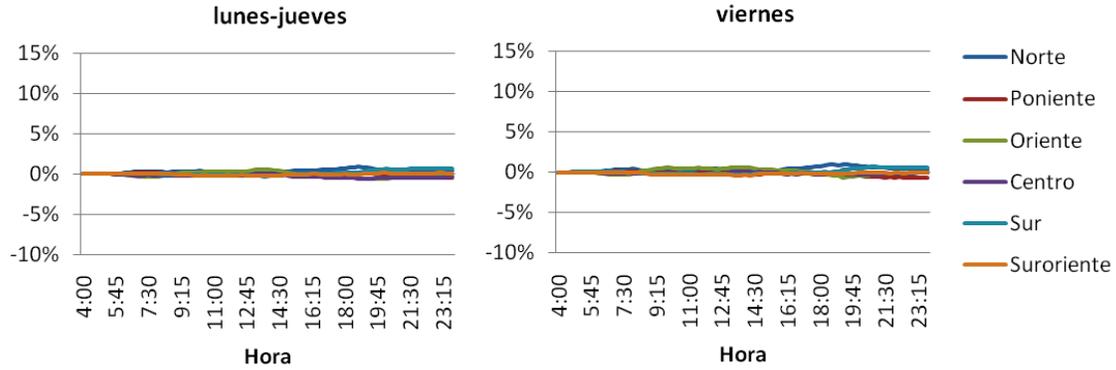


Figura 5.46: Comparación de perfiles de uso de tiempo actividad otros por zona, día laboral.



El día sábado la actividad otros en la zona oriente presenta una proporción de tarjetas mayor entre 9:00 y 18:00 hrs (ver figuras ?? y ??). Es probable que este comportamiento se asocie a una mayor disponibilidad de recursos para realizar actividades de carácter recreativo o deportivo. El día domingo el fenómeno se repite para las zonas centro y oriente entre 13:30 y 21:00 hrs.

Figura 5.47: Perfil de uso de tiempo actividad otros por zona, fin de semana.

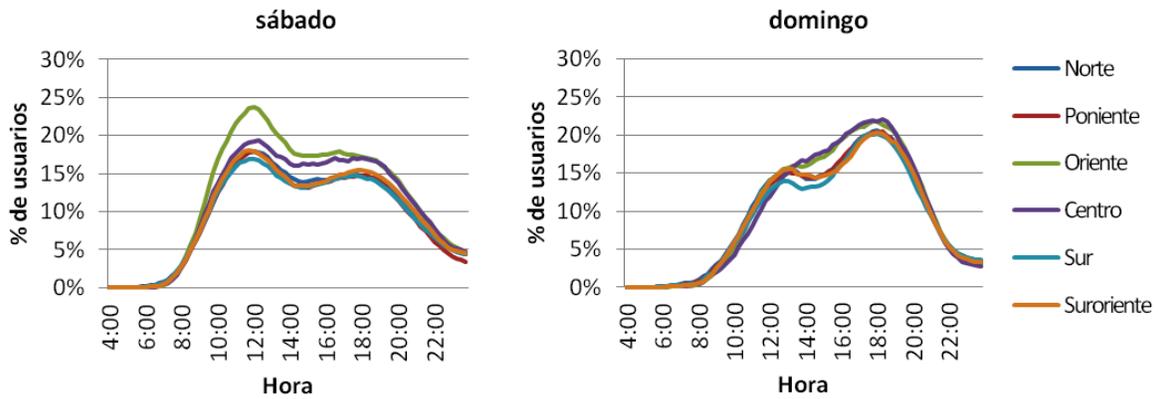
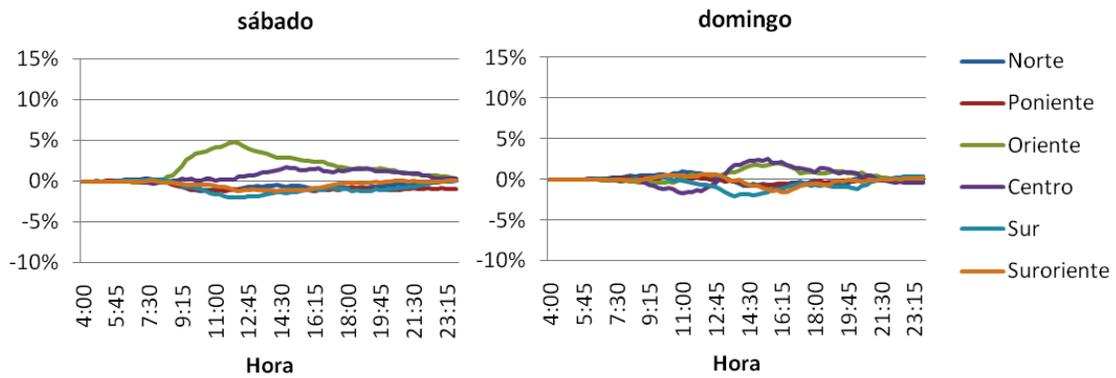


Figura 5.48: Comparación de perfiles de uso de tiempo actividad otros por zona, fin de semana.



Las figuras ?? y ?? presentan los resultados para actividad viaje en día hábil. El día viernes muestra una punta tarde más dispersa respecto al promedio lunes-jueves, para todas las zonas. La comparación entre zonas indica que quienes habitan en las zonas centro y oriente tienen tiempos de viaje menores, lo que se traduce en una menor proporción de personas en esta ocupación de manera simultanea y más tiempo disponible para asignar a otras actividades.

Figura 5.49: Perfil de uso de tiempo actividad viaje por zona, día laboral.

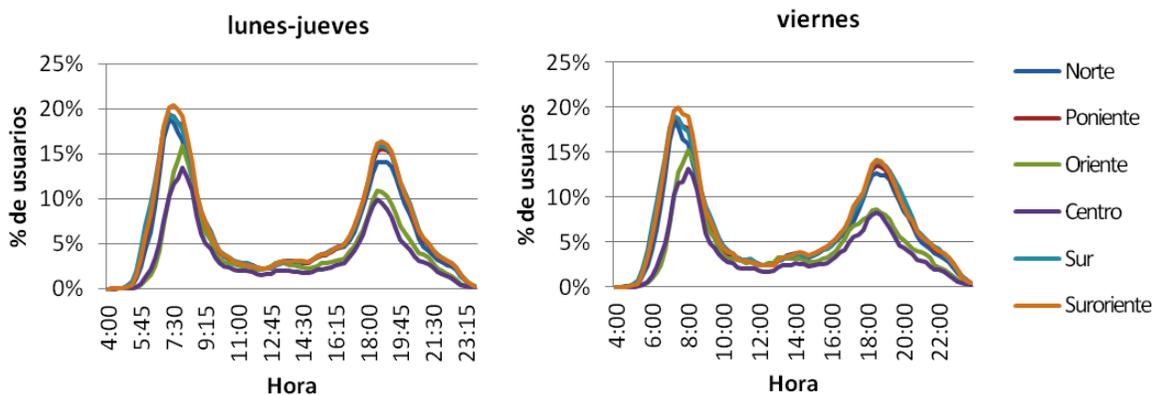
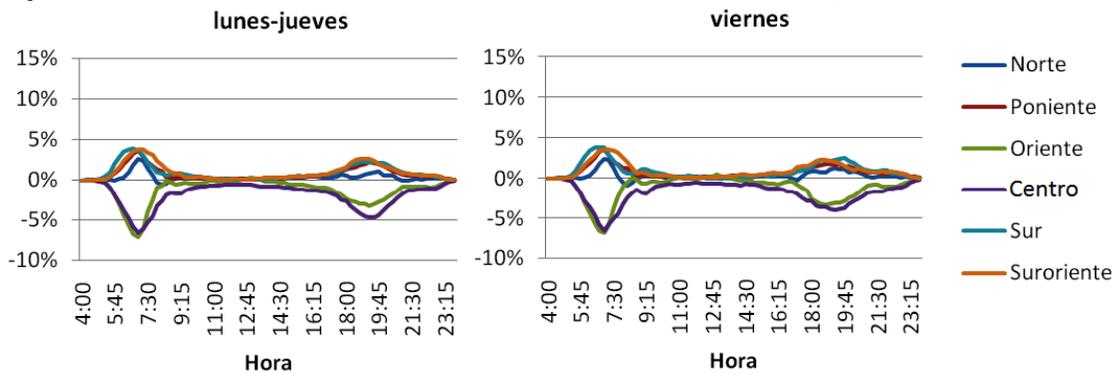


Figura 5.50: Comparación de perfiles de uso de tiempo actividad viaje por zona, día laboral.



El fin de semana se observa el mismo efecto en las zonas centro y oriente, que se encuentran bajo el promedio de asignación de tiempo respecto a las otras zonas a lo largo de ambos días.

Figura 5.51: Perfil de uso de tiempo actividad viaje por zona, fin de semana.

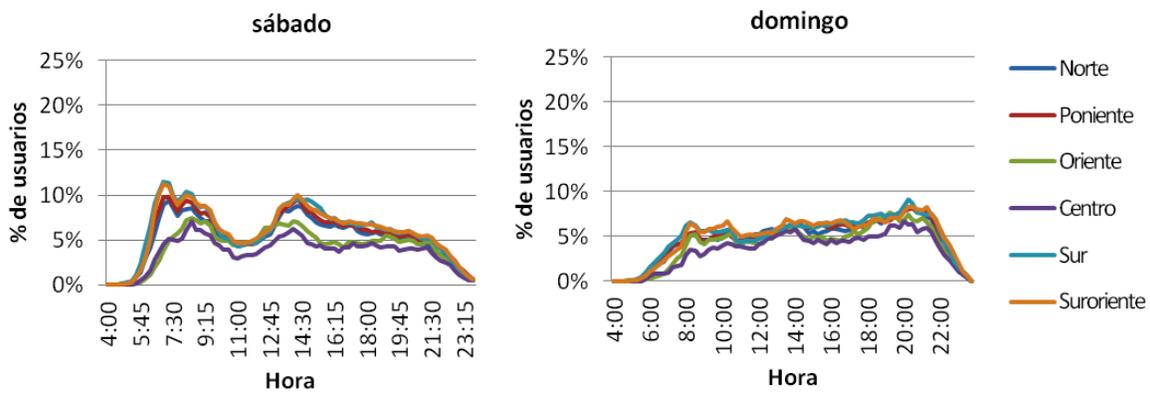
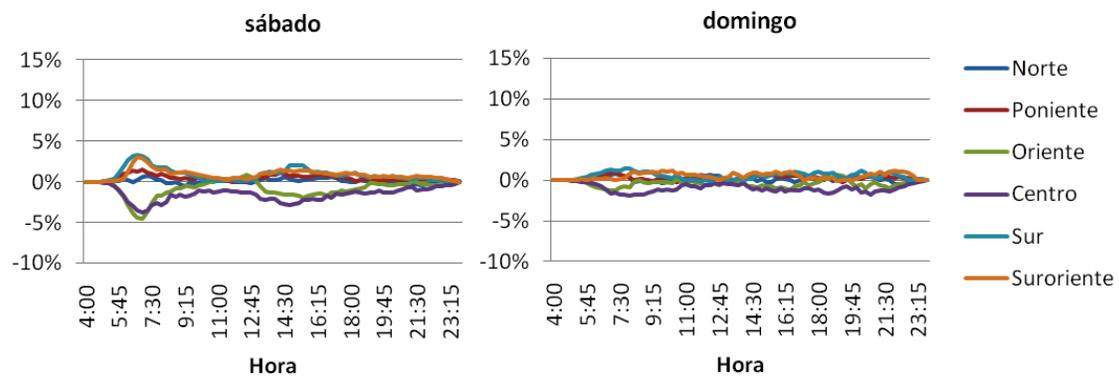


Figura 5.52: Comparación de perfiles de uso de tiempo actividad viaje por zona, fin de semana.



Los resultados de asignación de tiempo a viaje coinciden con lo mostrado en la sección de tiempos de viaje en vehículo y trasbordo.

Para ver el código (c++) aplicado a las bases de datos para obtener estos resultados, consultar Anexo E.

5.5. Síntesis y conclusiones

El capítulo de Resultados se divide en cuatro secciones. La primera clasifica a las tarjetas según el número de viajes diarios y muestra las distintas campanas de viaje para todas las tarjetas tipo adulto y estudiante. El patrón de viajes más común es realizar dos viajes diarios, típicamente para ejecutar la principal actividad, sea laboral o estudiantil. Como este patrón es el más numeroso (48,3 % del total de viajes para tarjeta adulto y 43,2 % para tarjeta estudiante), da forma a la visualización agregada del total de viajes que conocemos, ocultando otros comportamientos menos cuantiosos. Un 16 % del total de los viajes tarjeta adulto y un 9,1 % del total de viajes tarjeta estudiante son el único viaje del día. Esta conducta se explica por utilización de otros modos de transporte o evasión del pago y presenta una distribución de viajes muy distinta al caso recién descrito.

En el caso de la tarjeta adulto, para tres o más viajes se observan varias tendencias: el número de viajes diarios no afecta la dispersión de la primera campana de viajes del día, como si lo hace con las distribuciones del resto de los viajes; el primer y último viaje del día alcanzan su máximo a la misma hora independientemente de cuántos viajes se realicen, por lo tanto, las otras campanas se acumulan y superponen en un mismo espacio de tiempo; y los fines de semana las distribuciones de los viajes muestran similar altura y dispersión.

Por su parte, la tarjeta tipo estudiante ostenta una distribución de sus viajes completamente distinta y poco explorada en la literatura. Se destaca la presencia cuatro en vez de tres puntas de viaje, razón que justifica estudiar por separado a estos usuarios. Otra singularidad de este tipo de tarjeta es que una misma campana de viajes posee varias puntas que dan indicios de la agrupación de usuarios de distinto tipo y comportamiento en una misma visualización.

Como los patrones de comportamiento también tienen una componente temporal, la sección dos analiza la variación de la utilización de la tarjeta entre dos semanas de viaje separadas por un año de diferencia. Se cuantifica la proporción de tarjetas que aumentan, mantienen y reducen el número de transacciones entre uno y otro corte. La gran mayoría de los usuarios realiza cuatro o menos transacciones en cada semana, comportamiento correspondiente a quienes viajan infrecuentemente o que su principal modo de movilizarse no es transporte público.

Se destaca que dentro de los usuarios frecuentes aparecen dos máximos locales: quienes realizan 10 y 20 transacciones en los dos cortes de tiempo observados. Este patrón de comportamiento se atribuye a quienes realizan viajes de ida y vuelta durante los cinco días hábiles de la semana, cada viaje de una etapa en el caso del máximo (10,10) y de dos etapas en el

caso del máximo (20,20).

Los resultados muestran que quienes aumentan el número de transacciones (43,7 % de las tarjetas) son menos que quienes las reducen (48,6 %). Bajo este análisis, quienes aumentan su uso del transporte público son menos que quienes lo reducen.

La tercera sección expone las diferencias de tiempos de viaje en vehículo y trasbordo según tipo de tarjeta y zona de residencia. En términos generales, las tarjetas tipo estudiante presentan tiempos de viaje y trasbordo menores que las tarjetas tipo adulto; los tiempos de viaje en vehículo los fines de semana son menores que en días de semana; los siete días de la semana se aprecia un aumento del tiempo de viaje en punta mañana, pero sólo de lunes a viernes se observa el mismo efecto en punta tarde; y los tiempos de trasbordo son menores los días y horarios con mayor frecuencia de transporte (horarios punta respecto a valle y días hábiles respecto a fines de semana).

Al mirar las diferencias de tiempo por zona los resultados son contundentes: en la suma de tiempos de viaje en vehículo y trasbordo se observa que entre dos zonas hay diferencias de hasta 13 minutos por viaje para tarjetas tipo adulto y 11 minutos para tarjetas tipo estudiante.

Considerando los resultados de la sección uno, si una persona con tarjeta tipo adulto viaja de lunes a viernes y realiza dos viajes diarios (que es el patrón más común), quienes habitan en la zona sur gastan al mes 8 horas y 44 minutos más que quienes habitan en la zona centro, solo en tiempo de viaje en vehículo y trasbordo. Si los mismos cinco días hábiles se realizan tres viajes diarios (segundo comportamiento más común), la asignación de tiempo de viaje difiere en 13 horas y 6 minutos entre estas zonas. Para el caso de la tarjeta tipo estudiante las diferencias entre la zona centro y suroriente son de 7 horas y 44 minutos si se realizan dos viajes diarios y de 11 horas y 36 minutos si se realizan tres viajes diarios. Como la red de transporte en la zona centro es mucho más densa que en cualquier otra zona, es probable que estas diferencias crezcan al considerar tiempos de acceso y egreso.

Finalmente, la sección cuatro presenta los perfiles de uso de tiempo para la ciudad de Santiago por propósito y zona de residencia. La primera conclusión es que los perfiles de uso de tiempo del día viernes son distintos a lo observado de lunes a jueves. Las principales diferencias del día viernes son que la jornada laboral termina más temprano y que durante la tarde/noche aumenta el tiempo asignado a la actividad otros. Los fines de semana se reduce notoriamente el tiempo dedicado a trabajo y estudio y aumenta la actividad hogar y otros.

De lunes a viernes las tarjetas que realizan la actividad trabajo superan el 60 % y las que realizan la actividad estudio se acercan el 30 %. En el mismo periodo la actividad hogar muestra menos de un 10 % de tarjetas en esta ocupación. La actividad otros bordea el 5 % en la punta tarde. El día sábado las actividades trabajo y estudio reducen notoriamente su proporción, no sólo en cuanto a número de tarjetas sino también a duración de la actividad. Aumenta la proporción de tarjetas que dedica su tiempo a los propósitos hogar y

otros. El día domingo se acentúan los efectos mostrados el sábado de reducción de tiempo asignado a trabajo y estudio. La actividad otros alcanza su máxima proporción semanal en torno a las 18:00 hrs, con el 20 % de las tarjetas dedicadas a este propósito simultáneamente.

En cuanto a la comparación entre zonas, la sección tres ya revela importantes diferencias en la asignación de tiempo de viaje según zona de residencia en la ciudad. La sección cuatro amplía el análisis a otras actividades. Se destaca que quienes habitan en las zonas centro y oriente, además de pasar más tiempo en el hogar por tener menor tiempo de viaje (salen más tarde y retornan antes al hogar), dejan más tarde su hogar porque entran más tarde a trabajar. La proporción de tarjetas de las zonas centro y oriente en actividad trabajo es igual a las otras zonas a partir de las 9:00 de la mañana. En el caso de las diferencias en la proporción de personas que asignan su tiempo a la actividad estudio, estas diferencias parecen tener soporte en la composición de habitantes de cada zona. Además, los fines de semana estas dos zonas muestran una menor proporción de tarjetas en actividad trabajo.

Otra de las áreas de estudio que ha estado limitada por la escasez de datos es la valoración del tiempo. Para estudiar la valoración del tiempo y las razones para asignarlo a una determinada actividad se requiere datos de periodos de tiempo de al menos una semana, con el fin de observar un ciclo completo de trabajo y descanso. Por esta razón, las encuestas para estudiar valoración del tiempo tienen un alto costo y están limitadas a pequeñas muestras. La información de transacciones, las metodologías previamente desarrolladas y la estimación de información socioeconómica de la tarjeta terminan con estas limitantes y tienen el potencial de calibración de modelos de uso de tiempo para el Gran Santiago.

Capítulo 6

Conclusiones y futuras líneas de investigación

6.1. Síntesis y conclusiones

Los resultados de este trabajo se concentran en aumentar el conocimiento de los patrones de viaje con el objetivo de aportar a la planificación de la oferta de transporte público. Las primeras conclusiones se obtienen de los resultados del capítulo de análisis de datos. Se destaca que la mayoría de las tarjetas viajan de manera infrecuente y se desconoce su (i)regularidad de viaje y comportamiento. La importancia de este punto radica en que muy probablemente la variabilidad de la demanda está determinada por los patrones de viaje de los usuarios infrecuentes y por lo tanto, mejoras en la planificación pasan por conocer esa variabilidad de la demanda. Se detecta que existen tarjetas que si bien permanecen activas, no presentan registros de transacciones en lapsos de al menos siete días, razón por la cual se recomienda mirar periodos de tiempos mucho mayores.

Aproximadamente un 20 % del total de tarjetas que se observa en una semana corresponden a estudiantes. Si se considera que en promedio realizan más transacciones que la tarjeta adulto y que su distribución de viajes es completamente distinta, se tiene razones suficientes para argumentar la importancia de estudiar el comportamiento de esta tarjeta de manera independiente a la tarjeta adulto.

Los sistemas de transporte público con recolección de tarifa electrónica almacenan permanentemente información de los viajes. Una de las principales críticas a estos datos es que cuando la tarjeta de pago no es personalizada se desconoce toda información socioeconómica del pasajero. Con el trabajo realizado en esta tesis se rompe este paradigma. La metodología de estimación de zona de residencia entrega, a grandes rasgos, la ubicación geográfica del lugar donde habita el individuo. Si se refina este método con observaciones de periodos de tiempo mayores, podría conocerse la comuna e incluso un radio de manzanas en que vive. A esto se puede agregar información censal del nivel de ingreso de las zonas, enriqueciendo los datos de cada individuo en un ámbito que se descartaba para análisis.

La visualización agregada de la distribución de viajes por hora del día permiten conocer los periodos de alta y baja demanda de transporte. Sin embargo, esta visualización esconde valiosa información de la estructura de viajes que hay detrás de esta agregación. Al clasificar el comportamiento de las tarjetas según número de viajes diarios y observar la distribución de éstos se puede conocer mejor la estructura de la demanda.

Los resultados confirman que el comportamiento más común es realizar dos viajes diarios, como señala la intuición. Esta estructura de viaje es ampliamente conocida. Las principales diferencias se observan cuando las tarjetas realizan más viajes. La tendencia muestra que para tres o más viajes diarios, el primer y último viaje del día no cambian sustancialmente de horario y las campanas de viaje intermedias propenden a superponerse y acumularse en un mismo espacio de tiempo. También se observa un fuerte aumento en la dispersión de las distribuciones, con excepción de la primera del día.

La tarjeta tipo estudiante presenta una distribución de viajes completamente distinta y menos explorada en la literatura. Se destaca la presencia de cuatro en vez de tres puntas de viaje. Además, una misma campana de viajes posee varias puntas. Ambos hechos sugieren que la tarjeta tipo estudiante esta agrupando usuarios con patrones de viaje diferentes que no son recogidos adecuadamente al clasificar por número de viajes diarios. Muy probablemente, las múltiples puntas se deben a la coexistencia de estudiantes de distintos niveles de educación en una sola categoría. La agrupación de estos usuarios en una misma visualización esconde las diferencias entre sus estructuras de viaje.

La sección dos analiza la variación de la utilización de la tarjeta entre dos semanas de viaje separadas por un año de diferencia. Se cuantifica la proporción de tarjetas que aumentan, mantienen y reducen el número de transacciones entre uno y otro corte. Los resultados señalan que los usuarios que aumentan el uso del transporte público son menos que quienes lo reducen. Aunque hay factores involucrados en este análisis que no han sido considerados, como la incorporación de nuevos servicios al sistema que reducen el número de etapas por viaje, este resultado al menos marca tendencia y recuerda el círculo vicioso del transporte público y las consecuencias de deterioro del sistema en ausencia de medidas que lo fortalezcan.

Diferenciándose de los clásicos análisis, la sección tres muestra los tiempos de viaje en vehículo y trasbordo por zona de residencia en vez de origen-destino de viaje. Los resultados son contundentes: una persona con tarjeta tipo adulto que viaja de lunes a viernes y realiza dos viajes diarios (que es el patrón más común), utiliza 8 horas y 44 minutos más al mes si vive en la zona sur en vez de la zona centro. El mismo análisis para la tarjeta estudiante entrega una diferencia de 7 horas y 44 minutos entre las zonas centro y suroriente. Como la red de transporte en la zona centro es mucho más densa que en cualquier otra zona, es probable que estas diferencias crezcan al considerar tiempos de acceso y agreso. Este es el nivel de impacto que puede generar la elección de zona de residencia en la calidad de vida.

La sección cuatro presenta los perfiles de uso de tiempo para la ciudad de Santiago por

propósito y zona de residencia. Se concluye que los perfiles de uso de tiempo del día viernes son distintos a lo observado de lunes a jueves. Esto se debe a que la jornada laboral termina más temprano y que durante la tarde/noche aumenta el tiempo asignado a la actividad otros. Este análisis reconfirma las diferencias de asignación de tiempo entre usuarios con distintas zonas de residencia. Se concluye que quienes habitan en las zonas centro y oriente pasan más tiempo en el hogar porque tienen menores tiempos de viaje y porque entran más tarde a trabajar. Además, los fines de semana estas dos zonas muestran una menor proporción de tarjetas en actividad trabajo.

6.2. Líneas de investigación futura

Como se señaló en la sección anterior, un aspecto que se propone investigar es el estudio de patrones de viaje diarios entre los distintos tipos de tarjeta estudiante. Aquí se espera encontrar diferencias significativas de comportamiento según nivel de educación.

Por otra parte, se propone observar la distribución espacial de viajes de las tarjetas que realizan tres o más viajes diarios. Aunque la información indagada muestra que los viajes intermedios tienen origen y destino fuertemente ubicados en las zonas centro y oriente, este análisis abre otras ventanas al estudio de estructuras de viaje.

Sin duda, una posible línea de investigación es mejorar la metodología de estimación de zona de residencia que aquí se propone. Entre las falencias identificadas en la metodología está confundir zona de residencia con lugar de trabajo para quienes realizan turnos de noche y no poder estimar zona de residencia a los usuarios infrecuentes. Para refinar la estimación se requiere mirar periodos de tiempo mayores que permitan extraer información adicional de comportamiento. Además se propone incorporar información adicional para adjudicar nivel de ingreso a la tarjeta y de esta forma aumentar la información socioeconómica disponible.

El método de estimación de bajada no sólo puede mejorarse y re-aplicarse a los patrones estudiados, además puede utilizarse en otras variables de calidad de servicio que no fueron incluidas en esta tesis.

La línea de investigación que se considera prioritaria, por el impacto que puede causar en la planificación, es analizar bases de datos de periodos de tiempo mayores para estudiar patrones de viaje de usuarios que utilizan su tarjeta de manera esporádica. Se cree que esta rama puede aportar considerablemente a predecir la variabilidad de la demanda y mejorar la oferta de transporte público.

Finalmente, se estima que algunos de los resultados expuestos presentan sesgos por evasión de pago del pasaje. Aunque la evasión es una línea de investigación completamente diferente, se plantea incorporar resultados de la magnitud de este sesgo y determinar el impacto que puede causar en las conclusiones expuestas.

Capítulo 7

Referencias Bibliográficas

Agard, B., Morency, C. y Trépanier, M. (2009). Mining Smart Card Data from an urban Transit Network. IGI Global, pp. 1-11.

Bagchi, M. y White, P.R. (2004). What role for smart-card data from bus system? Municipal Engineer 157, pp. 39-46.

Bagchi, M. y White, P.R. (2005). The potencial of public transport smart card data. Transport Policy 12, pp. 464-474.

Bass, P., Donoso, P. y Munizaga, M. (2010). A model to assess public transport demand stability. Transportation Research Part A: Policy and Practice, 45 (2011), 755-764.

Beltrán, P., Gschwender, A., Munizaga, M., Palma, C., Zuñiga, M. (2012). Desarrollo de indicadores de nivel de servicio de Transantiago utilizando datos pasivos. XVII Congreso Panamericano de Ingeniería de Transito, Transporte y Logística - PANAM 2012, Santiago, Chile, 24-27 Septiembre.

Chriqui, C. y Robillard, P. (1975). "Common Bus Line", Transportation Science 9, 115-121.

Chu, K. y Chapleau, R. (2008). Enriching Archived Smart Card Transaction Data for Transit Demand Modeling. Transportation Research Record: Journal of the Transportation Research Board, No. 2063, Transportation Research Board of the National Academies, Washington, D.C., pp. 63-72.

Devillaine, F., (2012). Estimación de Viajes y Actividades en base a Sistemas Tecnológicos de transporte Público. Tesis presentada para la obtención del grado de Magíster en Ciencias de la Ingeniería, mención Transporte, Universidad de Chile.

Hofmann, M. y O' Mahony, M. (2005). Transfer Journey Identification and Analyses from Electronic Fare Collection Data. In: The 8th International IEEE Conference on Intelligent Transportation Systems – ITSC 2005, Viena, Austria, September 13-16.

Jang, W. (2010). Travel Time and Transfer Analysis Using Transit Smart Card Data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2144, Transportation Research Board of the National Academies, Washington, D.C., pp. 142-149.

Jara-Díaz, S., Munizaga, M. y Olgúin, J. (2010). The role of gender, age and location in the values of work behind time use patterns in Santiago, Chile. *Regional Science*.

Ma, X., Wu, Y., Wang, Y., Chen, F. y Liu, J. (2013). Mining smart card data for transit riders' travel patterns. *Transportation Research Part C: Emerging Technologies*, 36 (2013), 1-12.

Morency, C., Trépanier, M. y Agard, B. (2006). Analysing the variability of transit user behavior with smart card data. In: The 9th International IEEE Conference on Intelligent Transportation Systems – ITSC 2006, Toronto, Canada, September 17-20.

Morency, C., Trépanier, M. y Agard, B. (2007). Measuring transit use variability with smart-card data. *Transport Policy* 14 (3), pp. 193-203.

Munizaga, M. y Palma, C., (2012). Estimation of a disaggregate multimodal public transport origin - destination matrix from passive smart card data from Santiago, Chile. *Transportation Research Part C: Emerging Technologies*, 24, pp. 9-18.

Olgúin, J. (2009). Modelos de Uso de Tiempo para el gran Santiago. Tesis presentada para la obtención del grado de Magíster en Ciencias de la Ingeniería, mención Transporte, Universidad de Chile.

Ordoñez, S. y Erath, A. (2013). Estimating Dynamic Workplace Capacities using Public Transport Smart Card Data and a Household Travel Survey. *Transportation Research Board Annual Meeting 2013*, pp 13-3683.

Park, J.Y. y Kim, D.J. (2008). Use of Smart Card Data to Define Public Transit Use in Seoul, Korea. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2063, Transportation Research Board of The National Academies, Washington, D.C., pp. 3-9.

Pelletier, M., Trépanier, M. y Morency, C. (2011). Smart card data use in public transit: A literature review. *Transportation Research Part C: Emerging Technologies*, 19 (4), pp. 557-568.

Seaborn, C., Attanucci, J. y Wilson, N. H. M. (2009). Analyzing Multimodal Public Transport Journeys in London with Smart Card Fare Payment Data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2121, Transportation Research Board of The National Academies, Washington, D.C., pp. 55–62.

Sectra (2001). Encuestas de origen y destino de viajes del Gran Santiago 2001. http://www.sectra.gob.cl/Datos_e_Informacion_Espacial/Gran_Santiago/encuestas_movilidad.html.

Silva, D. (2012). Validación exógena de estimación de paradero de bajada y destino de actividades de usuarios de Transantiago usando encuestas origen/destino. Memoria presentada para optar al título de Ingeniero Civil, mención Transporte, Universidad de Chile.

Trépanier, M. y Morency, C. (2010). Assessing transit loyalty with Smart card data. 12th WCTR. Lisbon, Portugal.

Trépanier, M., Tranchant, N. y Chapleau, R. (2007). Individual Trip Destination Estimation in a Transit Smart Card Automated Fare Collection System. *Journal of Intelligent Transportation Systems* 11, 1-14.

Utsunomiya, M., Attanucci, J. y Wilson, N. (2006). Potential Uses of Transit Smart Card Registration and Transaction Data to Improve Transit Planning. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1971, Transportation Research Board of the National Academies, Washington, D.C., pp. 119-126.

Zhao, J. y Rahbee, A. (2007). Estimating a Rail Passenger Trip Origin-Destination Matrix Using Automatic Data Collection Systems, *Computer-Aided Civil and Infrastructure Engineering* 22, pp. 376-387.

Anexo A

Patrón de viajes diarios (código c++)

A.1. Tarjeta Adulto

El código se encuentra en el siguiente link:

<https://dl.dropboxusercontent.com/u/813345/PatronViajesAdulto.txt>

A.2. Tarjeta Estudiante

El código se encuentra en el siguiente link:

<https://dl.dropboxusercontent.com/u/813345/PatronViajesEscolar.txt>

Anexo B

Patrón de transacciones entre 2 cortes temporales (tabla datos)

La tabla con el total de datos que completan la matriz de transacciones de tarjetas en los cortes abril 2011 y abril 2012 se encuentra en el siguiente link:

<https://dl.dropboxusercontent.com/u/813345/Transacciones2CortesTemporales.txt>

Anexo C

Patrón de transacciones entre 2 cortes temporales (código c++)

El código se encuentra en el siguiente link:

<https://dl.dropboxusercontent.com/u/813345/TodosLosCortes.txt>

Anexo D

Tiempo de viaje en vehículo y tiempo de trasbordo (código c++)

D.1. Tarjeta Adulto

El código se encuentra en el siguiente link:

<https://dl.dropboxusercontent.com/u/813345/TiemposAdulto.txt>

D.2. Tarjeta Estudiante

El código se encuentra en el siguiente link:

<https://dl.dropboxusercontent.com/u/813345/TiemposEscolar.txt>

Anexo E

Comparación de perfiles de uso de tiempo entre zonas (código c++)

El código se encuentra en el siguiente link:

<https://dl.dropboxusercontent.com/u/813345/PerfilesXzonas.txt>