



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

MODELACIÓN Y ANÁLISIS ESPACIO-TEMPORAL DE LA DEMANDA DE VIAJES  
EN RIDESOURCING EN SANTIAGO DE CHILE

TESIS PARA OPTAR AL GRADO DE  
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN TRANSPORTE

MEMORIA PARA OPTAR AL TÍTULO DE  
INGENIERO CIVIL

BASTIÁN ALEJANDRO DOMÍNGUEZ HERNÁNDEZ

PROFESOR GUÍA:  
ALEJANDRO TIRACHINI HERNÁNDEZ

PROFESORA CO-GUÍA:  
MARCELA MUNIZAGA MUÑOZ

MIEMBROS DE LA COMISIÓN:  
CRISTIAN GUEVARA CUE  
GIOVANNI CIRCELLA

SANTIAGO DE CHILE  
2023

# Resumen

El enorme crecimiento de las plataformas ridesourcing – tales como Uber, Didi o Cabify – en la última década ha generado diversos impactos en varias ciudades alrededor del mundo. Debido a ello, es fundamental entender el funcionamiento de estos servicios, los factores que participan en su uso y el comportamiento de viajes, para así tomar decisiones que permitan potenciar los efectos positivos y contrarrestar los negativos. La liberación de datos de viajes por parte de las compañías de ridesourcing en el último tiempo ha permitido avanzar en este aspecto, sin embargo, los datos compartidos se han limitado principalmente a ciudades de Estados Unidos y China.

El presente trabajo de investigación busca extender el alcance de la literatura actual al contexto de una ciudad latinoamericana, incorporando mejoras en el método de modelación y el nivel de agregación temporal. Concretamente, en este estudio se examinan datos de viajes en Didi Chuxing realizados en Santiago de Chile entre el 13 y 26 de enero de 2020, con el objetivo de determinar los factores que participan en el uso de Didi e identificar el patrón espacio-temporal de viajes. Para ello, se construyen modelos de generación de viajes usando regresión lineal por mínimos cuadrados en dos etapas (2SLS), para diversas combinaciones día-período y a nivel de distrito censal. Se consideran dos variables dependientes: viajes por  $\text{km}^2$  y viajes por cada 1.000 hogares, cuyas variables explicativas engloban atributos de viaje, de sociodemografía, de entorno construido y de oferta de otros modos de transporte.

En general, los resultados muestran que una mayor demanda de ridesourcing está asociada a sectores con mayor poder adquisitivo y menor posesión de automóvil, mayor densidad poblacional, diversidad de actividades, menor frecuencia de buses y mayor presencia de paraderos y estaciones de metro, siendo consistentes con trabajos anteriores de este tipo. También se identifican puntos de interés que constituyen fuentes de generación de viajes, tales como centros comerciales, terminales de buses interurbanos y el Aeropuerto Internacional. Adicionalmente, al corregir por endogeneidad, los resultados también indican que la demanda de viajes se relacionan de manera inversa con el tiempo de espera. Por otro lado, la identificación del patrón de viajes permite inferir algunos aspectos del comportamiento de viajes en Santiago: la alta demanda durante los períodos nocturnos - especialmente durante viernes y sábado - sugiere que gran parte de los viajes en ridesourcing son de carácter ocasional, con propósitos de ocio y de retorno al hogar; mientras que el peak más pequeño observado durante la mañana en los días laborales sugiere que una parte considerable de los viajes en Didi son con motivo trabajo.

*A mi tía madrina, que siempre esperó este momento.*

# Agradecimientos

Siempre digo que durante mi carrera universitaria – y en mi vida en general – he tenido tres pilares, a quienes quiero agradecer a través estas palabras. En primer lugar a mi familia: mi mamá Olivia Hernández, mi papá Alejandro Domínguez, mi abuelita Mónica Hernández y mis hermanos Matías y Tomás. Gracias por todo el amor y apoyo incondicional que me han dado siempre, les debo todo lo que soy y esto también es de ustedes. Quiero mencionar especialmente a mi tía-madrina, Flor Hernández, quien siempre estuvo presente hasta donde pudo, espero que algún día pueda leer estas palabras y sepa lo mucho que significa para mí.

A mi pareja Mónica Martínez, mi compañera de vida y aventuras, quien me acompañó en este proceso incluso desde antes que iniciara. Sin ella, no habría podido llegar hasta esta instancia (en el sentido más literal). Gracias por todos los momentos buenos, malos y yetas que hemos vivido durante estos casi 10 años, por el inmenso amor que me das y por creer en mí siempre, incluso cuando yo mismo no lo hacía. Puedo decir con total seguridad que a tu lado me he convertido en una mejor versión de mí mismo.

A los amigos/as que han formado y forman parte de mi vida. A mis amigos/as de civil *Los Arrieros*: a Cyntia, la mejor amiga que me dio la etapa universitaria; a Cata, con quién entablé una amistad muy especial y que ha sabido estar en los momentos precisos; a Fran, Claudio y Ariel, por integrarme e invitarme a probar nuevas experiencias, cuidarme y retarme cuando fue necesario; a Diego, Octavio, Lulo, Karina, Michi, Benja, Indra, Clau, Nico, Mati y Rafa, con quienes compartí y seguimos compartiendo momentos valiosos tanto en la universidad como fuera de ella. A mis amigas de la mención Transporte, Sol y Monse, que con su energía positiva hicieron aún más agradable mi estadía aquí. A mis amigos/as de la época escolar: Emanuel, Marcela y Eve, que de algún modo estuvieron conmigo durante este proceso y de quienes he aprendido que no se necesita hablar a diario para tener una amistad duradera. Dicen que los amigos son la familia que uno elige, y creo que hay bastante de cierto en eso.

Además de estos tres pilares, agradezco a los profesores que marcaron mi estadía en el departamento de Transporte. A mis profesores guía Alejandro Tirachini y Marcela Munizaga, por confiarme este proyecto y ayudarme a enmendar el rumbo cuando no sabía cómo seguir; al profesor Francisco Martínez, por confiar en mis capacidades para ser auxiliar por primera vez; también al profesor Ángelo Guevara, por su gran disposición para ayudarme y resolver mis dudas en innumerables ocasiones. Finalmente, también agradezco a la comunidad del departamento de Transporte, por el cálido recibimiento y buena voluntad desde que llegué. Espero no haber dejado afuera a nadie y que puedan disculparme si lo hice.

# Tabla de Contenido

<b>1. Introducción</b>	<b>1</b>
<b>2. Estado del Arte</b>	<b>3</b>
2.1. Irrupción del ridesourcing e impactos en la movilidad	3
2.1.1. Ridesourcing vs. Taxis Tradicionales	4
2.1.2. Relación con Transporte Público	5
2.1.3. Impacto sobre la Posesión y Uso de Automóvil	6
2.1.4. Impacto en VKT	7
2.1.5. Accesibilidad y Equidad Social	8
2.2. Estudios basados en encuestas	10
2.3. Modelos de demanda basados en datos de viajes	12
<b>3. Metodología</b>	<b>16</b>
3.1. Datos	16
3.1.1. Recolección de Datos	16
3.1.2. Procesamiento de datos	20
3.1.3. Análisis Descriptivo	23
3.2. Modelación	40
3.2.1. Regresión Lineal OLS	40
3.2.2. Teorema de Gauss-Markov	42
3.2.3. Correcciones al modelo	42
3.2.4. Capacidad predictiva de los modelos	47

<b>4. Resultados y Discusión</b>	<b>49</b>
4.1. Resultados de los modelos	49
4.2. Interpretación de los modelos	54
4.2.1. Atributos de Ridesourcing	55
4.2.2. Sociodemografía	55
4.2.3. Entorno Construido	56
4.2.4. Oferta de transporte	57
4.3. Patrón de viajes	58
<b>5. Conclusiones</b>	<b>60</b>
<b>Bibliografía</b>	<b>67</b>
<b>Anexos</b>	<b>68</b>
<b>A. Bases de Datos</b>	<b>69</b>
A.1. Viajes en ridesourcing	69
A.2. Sociodemografía	70
A.3. Distritos Censales	71
A.4. Paraderos	71
A.5. Centros Comerciales	72
<b>B. Análisis Estadístico</b>	<b>73</b>
B.1. Histogramas de variables no temporales	73
B.2. Matriz de correlación	78
<b>C. Comparación de modelos con endogeneidad</b>	<b>79</b>
<b>D. Resultados viajes generados por cada 1.000 hogares</b>	<b>81</b>

# Índice de Tablas

2.1. Estudios sobre modelación de demanda en base a datos de viajes . . . . .	15
3.1. Atributos de ridesourcing . . . . .	17
3.2. Composición del área de estudio según uso de suelo. . . . .	19
3.3. Estadística descriptiva de atributos de viaje en ridesourcing . . . . .	23
3.4. Descripción y Estadística Descriptiva de las variables construidas . . . . .	26
3.5. Variables explicativas del modelo . . . . .	41
4.1. Resultados modelación viajes por km <sup>2</sup> período N2 (00:00 - 07:00) . . . . .	50
4.2. Resultados modelación viajes por km <sup>2</sup> período PM (07:00 - 10:00) . . . . .	51
4.3. Resultados modelación viajes por km <sup>2</sup> período FP (10:00 - 17:30) . . . . .	52
4.4. Resultados modelación viajes por km <sup>2</sup> período PT (17:30 - 21:00) . . . . .	53
4.5. Resultados modelación viajes por km <sup>2</sup> período N1 (21:00 - 00:00) . . . . .	54
A.1. Atributos de datos de solicitudes de viaje en Didi . . . . .	69
A.2. Atributos de base de datos de hogares EOD 2012 . . . . .	70
A.3. Atributos de base de datos de personas EOD 2012 . . . . .	70
A.4. Atributos de base de datos de distritos censales, Censo 2017 . . . . .	71
A.5. Atributos de base de datos de paraderos, PO Transantiago 2019-2 . . . . .	71
A.6. Centros Comerciales en el área de estudio . . . . .	72
B.1. Matriz de correlación de las variables no temporales . . . . .	78
C.1. Comparación de modelos para día laboral período PM, viajes/km <sup>2</sup> . . . . .	79

C.2. Comparación de modelos para día laboral período PT, viajes/km <sup>2</sup> . . . . .	80
D.1. Resultados modelación viajes por cada 1.000 hogares período N2 (00:00 - 07:00)	81
D.2. Resultados modelación viajes por cada 1.000 hogares período PM (07:00 - 10:00)	82
D.3. Resultados modelación viajes por cada 1.000 hogares período FP (10:00 - 17:30)	83
D.4. Resultados modelación viajes por cada 1.000 hogares período PT (17:30 - 21:00)	84
D.5. Resultados modelación viajes por cada 1.000 hogares período N1 (21:00 - 00:00)	85



# Índice de Ilustraciones

2.1. Demanda de taxis y ridesourcing en Nueva York y Chicago . . . . .	4
3.1. Número de observaciones descartadas según filtros aplicados . . . . .	21
3.2. Limpieza de base de datos de viajes en ridesourcing . . . . .	22
3.3. Histograma de tarifa del viaje . . . . .	23
3.4. Histograma de distancia recorrida . . . . .	24
3.5. Histograma de tiempo de viaje real . . . . .	24
3.6. Histograma de tiempo de espera . . . . .	24
3.7. Histograma de velocidad media . . . . .	25
3.8. Histograma de frecuencia de uso . . . . .	25
3.9. Perfil de Viajes . . . . .	27
3.10. Periodización de cada día de análisis . . . . .	28
3.11. Composición Temporal de Viajes . . . . .	28
3.12. Viajes promedio por hora según día-período . . . . .	29
3.13. Distribución espacial de viajes totales . . . . .	29
3.14. Distribución espacial de densidad de viajes . . . . .	30
3.15. Distribución espacial de viajes cada 1000 hogares . . . . .	30
3.16. Distribución espacial de tiempo de espera promedio . . . . .	31
3.17. Distribución espacial de velocidad media espacial . . . . .	31
3.18. Distribución espacial de % de población masculina . . . . .	32
3.19. Distribución espacial de % de población entre 18 y 39 años . . . . .	32

3.20. Distribución espacial de % de población con educación superior terminada	33
3.21. Distribución espacial de % de hogares de ingreso alto	33
3.22. Distribución espacial de % de hogares sin automóvil	34
3.23. Distribución espacial de % de superficie con uso urbano	34
3.24. Distribución espacial de número de hogares	35
3.25. Distribución espacial de densidad poblacional	35
3.26. Distribución espacial de mix de uso de suelo	36
3.27. Distribución espacial de porcentaje de suelo con uso habitacional	36
3.28. Distribución espacial de porcentaje de suelo con uso comercial	37
3.29. Distribución espacial de porcentaje de suelo con uso industrial	37
3.30. Distribución espacial de porcentaje de suelo con uso para oficinas	38
3.31. Distribución espacial de número de paraderos	38
3.32. Distribución espacial de densidad de paraderos	39
3.33. Distribución espacial de densidad de ciclovías	39
B.1. Histograma de proporción de población masculina	73
B.2. Histograma de proporción de población entre 18 y 39 años	73
B.3. Histograma de proporción de población con educación superior completada	74
B.4. Histograma de proporción de hogares de ingreso alto	74
B.5. Histograma de proporción de hogares sin auto	74
B.6. Histograma de densidad poblacional (log)	75
B.7. Histograma de mix de uso de suelo	75
B.8. Histograma de porcentaje de suelo con uso habitacional (H)	75
B.9. Histograma de porcentaje de suelo con uso comercial (C)	76
B.10. Histograma de porcentaje de suelo con uso industrial (I)	76
B.11. Histograma de porcentaje de suelo con uso de oficinas (O)	76
B.12. Histograma de densidad de paraderos	77
B.13. Histograma de densidad de ciclovías	77

# Capítulo 1

## Introducción

El concepto de *ridesourcing* o *ridehailing* se utiliza para referirse a servicios de transporte entregados por empresas como Uber, Didi o Cabify, en los que un usuario solicita un viaje mediante una aplicación de smartphone, en donde es puesto en contacto en tiempo real con un conductor que ofrece realizar dicho viaje. Estos servicios se caracterizan por ofrecer un servicio “puerta a puerta” a un costo que puede ser menor que el de los taxis tradicionales (Dias *et al.*, 2017; Lavieri *et al.*, 2018), con todas las comodidades que ofrece un automóvil sin necesidad de incurrir en los costos de poseer uno (Alemi *et al.*, 2018b), además de ser de fácil acceso, ya que sólo se requiere ingresar a la aplicación para solicitar el viaje y efectuar el pago. Adicionalmente, estas plataformas ofrecen la posibilidad de que varios usuarios compartan el mismo vehículo si su ruta es similar.

Debido a todas estas ventajas, las plataformas de ridesourcing han tenido enorme éxito desde su aparición hace una década, logrando un crecimiento exponencial desde entonces (Ghaffar *et al.*, 2020; Dey *et al.*, 2021). Por ejemplo, Uber – la compañía más grande del rubro – tardó seis años en alcanzar los mil millones de viajes a nivel mundial, pero solo tardó seis meses más en alcanzar los 2 mil millones de viajes, para luego alcanzar los 5 mil millones de viajes un año más tarde (Uber Blog, 2017). Este explosivo crecimiento ha causado que estos servicios produzcan múltiples impactos en la movilidad de varias ciudades alrededor del mundo. Dentro de los impactos positivos se encuentra la aparición de una nueva alternativa de transporte que otorga la cobertura y flexibilidad que no puede proporcionar el transporte público, sin tener que comprar o estacionar un automóvil (Rayle *et al.*, 2016; Ghaffar *et al.*, 2020). Por otra parte, los mismos autores señalan que la irrupción del ridesourcing puede impactar negativamente la ciudad a través de un aumento de la congestión y emisiones, debido a posibles sustituciones desde modos más eficientes – caminata, bicicleta y transporte público – e incluso debido a un eventual aumento de la cantidad de viajes.

Debido a la multiplicidad y magnitud de estos impactos, es fundamental entender cómo el uso de estas aplicaciones afecta los patrones de actividades de las personas, el uso de otros modos de transporte, la posesión de automóvil, la equidad social, la congestión y las emisiones de contaminantes, entre otros (Alemi *et al.*, 2018b). De esta manera, se espera que los planificadores y autoridades pertinentes puedan elaborar políticas eficientes que incentiven los impactos positivos e impidan los impactos negativos (Ghaffar *et al.*, 2020).

Varios autores coinciden en que uno de los mayores inconvenientes para investigar estos temas ha sido la falta de disponibilidad de datos de viajes en ridesourcing (Lavieri *et al.*, 2018; Alemi *et al.*, 2018b; Lavieri y Bhat, 2019; Ghaffar *et al.*, 2020; Marquet, 2020). Para superar este inconveniente, algunos investigadores han optado por utilizar encuestas como principal instrumento de adquisición de datos, permitiendo además obtener información a nivel de usuario (Ghaffar *et al.*, 2020). Afortunadamente, en los últimos años se han puesto a disposición algunas bases de datos espaciotemporales de viajes en ridesourcing, principalmente en ciudades de Estados Unidos y China. Cabe señalar que ambos tipos de datos – encuestas y datos espaciotemporales de viajes – tienen sus propias virtudes y debilidades, por lo que pueden ser utilizados de forma complementaria para corroborar resultados, especialmente con respecto a quienes utilizan estos servicios (Ghaffar *et al.*, 2020).

El objetivo de la presente investigación consiste en contribuir a la literatura existente sobre el entendimiento de este fenómeno, particularmente el comportamiento de viajes en este modo. Concretamente, se pretende responder las siguientes preguntas de investigación: (1) ¿Qué factores participan en la demanda de viajes en ridesourcing? y (2) ¿Cómo es el patrón espaciotemporal de viajes en ridesourcing? Para ello, se utilizan datos espaciotemporales de viajes realizados en Santiago de Chile mediante la plataforma Didi Chuxing para elaborar modelos de generación de viajes a nivel agregado, incluyendo como variables explicativas atributos de modo, características de sociodemografía, de entorno construido y oferta de otros modos de transporte en la ciudad.

Este estudio aporta varios elementos novedosos en relación a trabajos anteriores de este tipo. En primer lugar, de acuerdo a la investigación realizada por el autor, este es el primer trabajo realizado con datos de viajes en ridesourcing en una ciudad latinoamericana, expandiendo la exploración a contextos diferentes a los estudiados hasta la fecha. En segundo lugar, se elaboran modelos diferenciados por tipo de día y período, con el fin de capturar de manera más precisa los diferentes grados de influencia de los factores considerados en el análisis según el momento del día. Por último, se tiene la inclusión de los atributos de modo, en particular la velocidad y el tiempo de espera, siendo este último corregido por endogeneidad.

La estructura de este trabajo de investigación es la siguiente. Después de esta introducción, en el Capítulo 2 se presenta una recopilación y revisión de los trabajos anteriores que abordan el tema en cuestión. Luego, en el Capítulo 3 se describen los datos utilizados y la metodología para la elaboración de los modelos. Posteriormente, en el Capítulo 4 se presentan los resultados obtenidos. Para finalizar, en el Capítulo 5 se mencionan las principales conclusiones de este trabajo.

# Capítulo 2

## Estado del Arte

Varios autores afirman que el conocimiento acerca de los servicios ridesourcing aún es limitado (Hughes y MacKenzie, 2016; Alemi *et al.*, 2018a; Tirachini y del Río, 2019; Marquet, 2020). Sin embargo, la mayor disponibilidad de datos en el último tiempo ha permitido avanzar en diferentes ámbitos de este fenómeno. Según Alemi *et al.* (2018a) y Bekka *et al.* (2020), la literatura relacionada con el estudio del ridesourcing se divide en dos grandes grupos. El primer grupo – dentro del cual se encuentra este trabajo – investiga los factores que influyen en la elección y frecuencia de estos servicios, el perfil de los usuarios, propósitos de viaje y patrones espaciotemporales. El segundo grupo discute y evalúa los potenciales impactos del ridesourcing sobre la elección de modo, posesión de automóvil, equidad social, patrones de actividades, congestión y emisiones, entre otros. A los grupos antes mencionados se suman al menos dos grupos más: el tercer grupo estudia la oferta de ridesourcing y mecanismos de tarificación óptima (por ejemplo, Li *et al.*, 2021; Gómez-Lobo *et al.*, 2022); mientras que el cuarto y último grupo se dedica al estudio del trabajo y comportamiento de los conductores (por ejemplo, Hall y Krueger, 2018; Fielbaum y Tirachini, 2021).

En este capítulo se aborda y discute la literatura relacionada con el estudio de la aparición y uso de los servicios ridesourcing, considerando los primeros dos grupos. Para comenzar, en la Sección 2.1 se estudia el panorama general de la irrupción del ridesourcing y sus impactos en diferentes ámbitos de la movilidad, de forma similar a lo realizado en Marquet (2020). Luego, en la Sección 2.2 se revisan los trabajos que desarrollan modelos de comportamiento de viajes basados en encuestas a usuarios. Finalmente, en la Sección 2.3 se recopilan y analizan los trabajos que desarrollan modelos basados en datos masivos de viajes en ridesourcing.

### 2.1. Irrupción del ridesourcing e impactos en la movilidad

La irrupción de las plataformas ridesourcing y su rápido crecimiento ha causado varios impactos en la movilidad de varias ciudades alrededor del mundo, afectando su sistema de transporte y los patrones de actividades (Alemi *et al.*, 2018b; Ghaffar *et al.*, 2020). Además, este fenómeno se ha desarrollado en un período de tiempo relativamente corto, por lo cual estos impactos aún no han sido completamente comprendidos. Según Marquet (2020),

la literatura dedicada al ridesourcing ha avanzado considerablemente, con énfasis en los siguientes temas clave: (1) impacto sobre los taxis tradicionales, (2) impacto sobre la posesión y uso de automóvil, (3) impacto sobre VKT<sup>1</sup> y (4) equidad social. Además, varios autores han abordado la relación del ridesourcing con el transporte público (por ejemplo, Clewlow y Mishra, 2017; Hall *et al.*, 2018; Graehler *et al.*, 2019). A continuación, se revisa en detalle el panorama actual en cada uno de estos temas.

### 2.1.1. Ridesourcing vs. Taxis Tradicionales

Antes de la aparición de las plataformas ridesourcing, los principales proveedores de servicios de transporte tipo “puerta a puerta” solían ser los taxis, entregando mayor alcance y flexibilidad horaria que el transporte público, sin la necesidad de contar con un lugar para estacionarse (Aguilera-García *et al.*, 2022). Sin embargo, la llegada de esta nueva alternativa de transporte a las ciudades cambió drásticamente dicho panorama, ya que al proveer el mismo servicio compite directamente con la industria de los taxis. Por ejemplo, en Nueva York, la demanda diaria de viajes en taxi oscila entre 400.000 y 500.000 viajes entre 2010 y 2014, pero con la aparición de Uber y Lyft a partir de 2015 la demanda de viajes en taxi disminuye de forma sostenida hasta los 200.000 viajes diarios antes de la pandemia de covid-19, mientras que durante el mismo período los viajes diarios en ridesourcing aumentaron de 70.000 a 750.000 (NYC Taxi and Limousine Commission, 2022). En Chicago se tiene un escenario similar: las plataformas ridesourcing entraron con fuerza en la ciudad, a tal punto que desde su ingreso concentran aproximadamente el 80 % de la demanda total en servicios “puerta a puerta” (Chicago Data Portal, 2022). La Figura 2.1 muestra en detalle la evolución temporal de la demanda de viajes tanto en taxis como en ridesourcing.

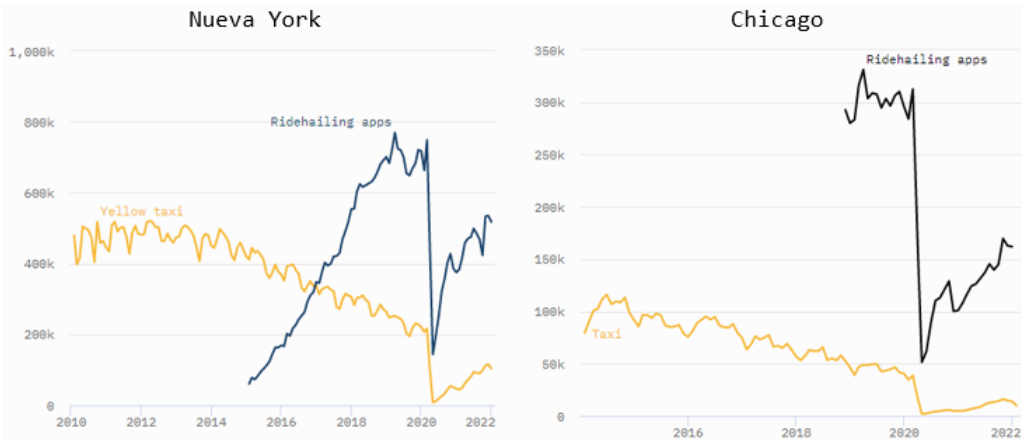


Figura 2.1: Demanda de taxis y ridesourcing en Nueva York y Chicago

La pregunta que surge entonces es ¿a qué se debe esta diferencia de demanda si ambas industrias ofrecen el mismo servicio? La respuesta es que, a pesar de ofrecer el mismo servicio de transporte, existen diferencias importantes que inciden en las preferencias de los usuarios por uno u otro modo. La primera gran diferencia es que el ridesourcing funciona en base a

<sup>1</sup>VKT (*vehicle kilometers traveled*) es un indicador utilizado en transporte para medir la cantidad de viajes realizados por todos los vehículos dentro de un área geográfica durante determinado período de tiempo.

aplicaciones de smartphone, mientras que los taxis que deben ser solicitados en plena calle o a través de una llamada telefónica (aunque en el último tiempo algunas empresas de taxis han habilitado aplicaciones para solicitar viajes, como es el caso de EasyTaxi en Brasil, Chile y México, entre otros). Esto permite que los servicios ridesourcing sean de fácil acceso para los usuarios, más eficientes en términos de tiempo de utilización (Rayle *et al.*, 2016; Cramer y Krueger, 2016) e incluso con menores tiempos de espera (Rayle *et al.*, 2016; Brown y LaValle, 2021). Además de las diferencias mencionadas, Brown y LaValle (2021) identifican una serie de elementos que hacen que los usuarios se inclinen por el ridesourcing en desmedro de los taxis, tales como la forma de pago y capacidad de retroalimentación, ya que hacen que este modo sea más asequible y confiable.

Por otra parte, debido a su relativamente reciente aparición en el mercado, actualmente el ridesourcing posee menos restricciones regulatorias que los taxis en términos de esquema tarifario, oferta disponible, estándares de seguridad de los vehículos y nivel de expertiz de los conductores, entre otros (Rayle *et al.*, 2016; Brown y LaValle, 2021; Aguilera-García *et al.*, 2022). Esto ha sido un motivo frecuente de quejas por parte de la industria de taxis en varias ciudades alrededor del mundo, argumentando que no pueden competir en igualdad de condiciones (Rayle *et al.*, 2016; Aguilera-García *et al.*, 2022).

### 2.1.2. Relación con Transporte Público

Hasta el momento, el efecto neto del ridesourcing sobre el transporte público no es del todo claro, ya que hay argumentos para afirmar que son tanto complementarios como sustitutos. De hecho, existe una vasta literatura que apoya ambas posturas (Jin *et al.*, 2018), cuyos resultados dependen fuertemente del contexto de cada ciudad (Wang y Yang, 2019). De acuerdo a Hall *et al.* (2018) y Jin *et al.* (2018), el efecto complementario aparece debido a que el ridesourcing puede cubrir tramos que el transporte público no puede debido a sus recorridos fijos (problema de la última milla), mientras que el efecto sustitutivo surge naturalmente debido a la inclusión de un nuevo modo de transporte con características más favorables para el usuario, especialmente en períodos y sitios donde la disponibilidad de transporte público es escasa. En Kong *et al.* (2020) se reafirma este punto a través de un análisis espaciotemporal a nivel desagregado de los viajes realizados mediante la aplicación Didi Chuxing en Chengdu, China: sus resultados muestran que existe mayor grado de sustitución en los sectores con menor cobertura de transporte público y en horarios en donde el funcionamiento del mismo es menor.

Dentro de los estudios que sugieren que el ridesourcing y el transporte público son complementarios, se encuentra el de Zhang y Zhang (2018). En este trabajo se estima una regresión binomial negativa inflada en ceros a partir de datos de la Encuesta Nacional de Viajes de Hogares (NHTS) de Estados Unidos, obteniendo que un aumento en un 1% del uso de transporte público está asociado a un aumento del 1,2% de la frecuencia del uso de ridesourcing y a un aumento del 5,7% de la probabilidad de uso de estos servicios, principalmente en sectores con mayor densidad poblacional y menor posesión de automóvil. Hall *et al.* (2018) recolecta datos de viajes en Uber y transporte público entre 2004 y 2015, encontrando que en general Uber complementa el transporte público en Estados Unidos, aumentando la demanda de transporte público en ciudades grandes.



Sin embargo, existen varios estudios empíricos recientes apuntan a que el efecto de la sustitución es más fuerte que la complementariedad. Por ejemplo, a partir de una encuesta de interceptación realizada a usuarios de ridesourcing en San Francisco, en [Rayle et al. \(2016\)](#) se obtuvo que un 33 % de los usuarios hubieran realizado el viaje en cuestión en bus o metro. Por su parte, [Clewlow y Mishra \(2017\)](#) realizaron una encuesta en siete áreas metropolitanas de los Estados Unidos, obteniendo que el uso de ridesourcing reduce en un 6 % el uso de transporte público. Más tarde, [Tirachini y del Río \(2019\)](#) realizaron un análisis similar al de [Rayle et al. \(2016\)](#) en Santiago de Chile, obteniendo que un 37,6 % de los viajes realizados en ridesourcing sustituyen viajes que anteriormente se hubieran realizado en transporte público, mientras que sólo un 3,4 % de los viajes en ridesourcing son combinados con bus o metro. [Graehler et al. \(2019\)](#) analiza los factores que inciden en la disminución de la demanda en transporte público en Estados Unidos, obteniendo que el ridesourcing reduce en promedio un 1.5 % la demanda de viajes en transporte público cada año. Finalmente, en [Kong et al. \(2020\)](#) se obtiene que el 33.1 % de los viajes en ridesourcing son potencialmente sustituibles por viajes en transporte público.

Según [Tirachini \(2020\)](#), este tema es particularmente relevante para la toma de decisiones por parte de las autoridades. Si estos modos resultan ser complementarios, se pueden adoptar medidas para potenciar esta relación y así incentivar el uso de transporte público, por ejemplo, a través de convenios que permita a los usuarios acceder a descuentos en la tarifa si se combinan ambos modos. Por ejemplo, en 2016 en el condado de Pinellas, Florida, la Autoridad de Tránsito de Pinellas Suncoast (PSTA) comenzó a subsidiar viajes en Uber para cualquier usuario que viaje desde o hacia paraderos de buses designados, ubicados en sectores con baja cobertura de transporte público ([Bloomberg CityLab, 2016](#)). Por el contrario, si ambos modos resultan ser sustitutos, puede ser perjudicial para la movilidad de la ciudad y especialmente para los usuarios (cautivos) de transporte público, ya que además de aumentar la congestión, la disminución de la demanda de este modo puede causar que los operadores de transporte público decidan reducir la frecuencia de sus vehículos, provocando un aumento de los tiempos de espera.

### 2.1.3. Impacto sobre la Posesión y Uso de Automóvil

Según la industria de las plataformas ridesourcing, esta nueva tecnología – especialmente a través del uso compartido de un automóvil privado – permitiría reducir el número de vehículos en circulación y estacionados en espacios públicos ([Jin et al., 2018](#); [Bekka et al., 2020](#)). Sin embargo, este efecto no se ha manifestado en gran medida hasta el momento, pues la evidencia empírica muestra que los usuarios de ridesourcing se caracterizan por tener menor acceso al automóvil que la población en general ([Rayle et al., 2016](#); [Bekka et al., 2020](#)), lo que sugiere que los usuarios de ridesourcing provienen de otros modos de transporte.

Si bien existen usuarios que han reducido su uso de automóvil a raíz de la aparición del ridesourcing, esta proporción resulta ser una minoría. Por ejemplo, según [Rayle et al. \(2016\)](#), el 90 % de los usuarios encuestados no modifica su posesión de automóvil desde que comenzaron a utilizar el ridesourcing, mientras que en [Clewlow y Mishra \(2017\)](#) esta cifra alcanza el 91 %. Por otro lado, los resultados de [Henao y Marshall \(2019\)](#) y [Tirachini y del Río \(2019\)](#) indican que, de no haber contado con el ridesourcing como opción, la cantidad



de usuarios que hubieran realizado el viaje en automóvil no superan el 20%, de hecho, la mayor parte sustituyen modos como el transporte público y taxis. A pesar de lo que indican estos resultados, es importante tener en cuenta sus limitaciones. Por ejemplo, un usuario puede declarar que está sustituyendo viajes que anteriormente hubiera realizado en transporte público, pero existe la posibilidad de que el uso de ridesourcing sea la principal razón por la cual no adquiere un automóvil privado (Tirachini, 2020).

Anderson (2014), por otro lado, afirma que el ridesourcing incluso puede potenciar la posesión y el uso de automóvil si se considera la perspectiva de los conductores, ya que los ingresos obtenidos permiten subsidiar el uso y en algunos casos incentivar la compra de un vehículo propio para entrar a este mercado.

#### 2.1.4. Impacto en VKT

El VKT es un indicador clave para representar la demanda de viajes, debido a los efectos que tiene sobre la congestión y el medio ambiente (Wu y MacKenzie, 2021). La influencia del ridesourcing sobre el VKT es el resultado de los impactos que genera esta nueva opción en el uso de los otros modos de transporte (Tirachini y Gómez-Lobo, 2020; Marquet, 2020), además de la demanda inducida (Wu y MacKenzie, 2021). Dado que la magnitud de los impactos sobre cada modo varía en cada ciudad, aún no existe consenso sobre el efecto neto del ridesourcing sobre el VKT.

Al comparar el uso de ridesourcing con el uso de automóvil propio, existen argumentos que apuntan a una reducción del VKT. En primer lugar, los servicios de ridesourcing no requieren estacionamiento, lo que permite ahorrar tiempo y distancia en búsqueda de uno. En segundo lugar, si bien la adquisición de un automóvil requiere una fuerte inversión económica, el costo marginal de realizar un viaje en auto propio es significativamente menor que realizar un viaje en ridesourcing, lo cual incentiva a las personas a realizar viajes más largos y en mayor cantidad (Tirachini y Gómez-Lobo, 2020). Sin embargo, también se debe tener en cuenta que el uso del ridesourcing implica un recorrido mayor en relación a realizar el mismo viaje en automóvil, debido a la circulación del vehículo mientras no lleva pasajeros (*deadheading*), aumentando el VKT (Clewlow y Mishra, 2017; Tirachini y Gómez-Lobo, 2020; Young y Farber, 2019). De hecho, Cramer y Krueger (2016) obtuvieron que el 36% de los VKT de Uber en Los Ángeles se recorren sin pasajeros, mientras que en Seattle alcanza el 45%. Wenzel *et al.* (2019) analiza 1,5 millones de viajes realizados mediante la plataforma RideAustin, obteniendo que el 26% del VKT corresponde a viajes sin pasajeros.

Por otra parte, la sustitución del transporte público y modos activos como la bicicleta y la caminata generan un aumento en el VKT y también en la congestión, ya que se están sustituyendo modos que son más eficientes en términos de utilización del espacio vial. Este hecho es una de las principales críticas que reciben las compañías de ridesourcing, ya que va en contra de los objetivos del transporte sustentable (Young y Farber, 2019).

Hasta el momento, existen varios estudios que han intentado estimar el efecto del ridesourcing sobre el VKT. Por ejemplo, Tirachini y Gómez-Lobo (2020) realizan una simulación de Montecarlo para estimar el efecto sobre el VKT de los viajes realizados mediante Uber en Chile. Sus resultados indican que en la situación base el uso de ridesourcing aumenta el

VKT, más aún, los autores afirman que es necesario incrementar significativamente la tasa de ocupación y la cantidad de viajes compartidos para lograr revertir este efecto. Por su parte, [Bekka \*et al.\* \(2020\)](#) analiza el efecto de Uber sobre el VKT a través de la reducción de la posesión de automóvil en París, obteniendo que hasta el momento el efecto es mínimo. De manera similar, [Wu y MacKenzie \(2021\)](#) utilizan los datos de la Encuesta Nacional de Viajes de Hogares (NHTS) en Estados Unidos para estimar el efecto del uso de ridesourcing sobre el VKT para distintos grupos de población según su frecuencia de uso y disponibilidad de automóvil propio, obteniendo que el uso de ridesourcing produce un aumento neto de 12,6 millones VKT diarios en comparación al escenario hipotético en que ninguno de los encuestados usa estos servicios. Además, sus resultados indican que únicamente los individuos que poseen mayor acceso al automóvil reducen su VKT, aunque sigue siendo más alto que el resto de la población.

### 2.1.5. Accesibilidad y Equidad Social

La accesibilidad al transporte es uno de los aspectos centrales en la economía y especialmente de la vida social de las personas, ya que permite acceder a bienes, servicios, educación, empleos, salud y actividades de interacción social, entre otros ([National Academies of Sciences, Engineering, and Medicine, 2016](#)). Por lo tanto, es fundamental que las autoridades sean capaces de garantizar este derecho a toda la población. Sin embargo, existen varios segmentos de la población que ven limitada esta accesibilidad debido a factores como bajos ingresos, situaciones de discapacidad o discriminación ([National Academies of Sciences, Engineering, and Medicine, 2016](#)). En este contexto, la aparición del ridesourcing podría ser una contribución al mejoramiento de la accesibilidad o bien podría acentuar este problema, por lo cual este aspecto es uno de los temas claves que se aborda en la literatura.

#### Población de bajos ingresos

La posesión de automóvil otorga a las personas mayor accesibilidad, ya que les permite acudir a diversas actividades – tales como trabajo, educación y ocio – a mayores distancias. Sin embargo, sus costos de adquisición y mantenimiento suelen ser bastante altos, especialmente para las hogares de ingresos bajos, que suelen ser justamente los que viven más lejos de los centros de actividades. En este contexto, la cobertura de sectores de bajos ingresos es una de las oportunidades de mejora que puede ofrecer el ridesourcing, ya que como se mencionó anteriormente, permite acceder a los beneficios que proporciona un automóvil sin tener que adquirir comprar uno ([National Academies of Sciences, Engineering, and Medicine, 2016](#)).

No obstante lo anterior, [Kodrinsky y Lewenstein \(2014\)](#) identifican una serie de barreras que limitan particularmente el uso de ridesourcing por parte de la población de menores ingresos. En primer lugar está el ámbito logístico y financiero: para poder solicitar un viaje en las plataformas ridesourcing se requiere tener un smartphone con acceso a internet y una tarjeta de crédito o débito, lo cual naturalmente es menos frecuente en este segmento de la población. Por otra parte, los conductores no tienen incentivos para servir en los sectores de bajos ingresos debido a que no son rentables y son percibidos como más inseguros.

## Población Rural

Otro grupo vulnerable en términos de movilidad es la población rural. Debido a las grandes distancias que deben recorrer para acceder a los centros de actividades y a la baja cobertura de transporte público que suelen tener estos sectores, este segmento de la población depende más de los vehículos privados que las personas que viven en sectores urbanos (National Academies of Sciences, Engineering, and Medicine, 2016). Por lo tanto, la movilidad de las personas que no pueden acceder a un automóvil propio en estos sectores – por ejemplo, familias de escasos recursos o gente de tercera edad – se encuentra tremendamente restringida. En este aspecto, el ridesourcing no ha logrado ser una solución efectiva hasta el momento, ya que la evidencia muestra que su uso se da principalmente en áreas urbanas (Clewlow y Mishra, 2017; Alemi *et al.*, 2018b; Aguilera-García *et al.*, 2022). Esto se debe en parte al perfil del usuario de ridesourcing – personas jóvenes con gran acceso a tecnologías de información – y, dado que la cobertura del ridesourcing depende principalmente de los conductores, quienes tienen más incentivos a operar en zonas con mayor densidad poblacional y así obtener mayores ganancias. Para solucionar este problema, se requiere implementar medidas que incentiven a los conductores a operar en estos sectores, por ejemplo, a través de subsidios (National Academies of Sciences, Engineering, and Medicine, 2016).

## Discriminación Racial

La tecnología utilizada por las plataformas ridesourcing podrían combatir la discriminación racial que ocurre en otros modos de transporte, ya que los conductores deben aceptar la solicitud del viaje sin conocer la raza o etnia del usuario, aunque existe la posibilidad de que algunos conductores rechacen las solicitudes por parte de usuarios cuyos nombres “suenen” de origen afroamericano (National Academies of Sciences, Engineering, and Medicine, 2016). De hecho, la investigación de Ge *et al.* (2016) sugiere que efectivamente existe discriminación hacia estos usuarios, ya que los resultados muestran que sus solicitudes tardan entre un 16% a un 28% más en ser aceptadas, además de que su tasa de cancelación es hasta tres veces mayor que la de los varones con nombres que suenen de raza blanca. Sin embargo, los autores son enfáticos en señalar que la discriminación racial no se da exclusivamente en las plataformas de ridesourcing, ya que existen estudios previos que dan cuenta de la existencia de discriminación en otros modos de transporte. Por ejemplo, en Brown (2019) se señala que los taxistas suelen evitar los barrios “negros”, mientras que Lyft incluso efectúa una mayor cantidad de viajes per cápita en estos barrios que en aquellos sin mayoría racial.

## Perspectiva de Género

En términos de movilidad y transporte, la población femenina se enfrenta sistemáticamente a diversas situaciones desfavorables. En primer lugar, su capacidad para adquirir y conducir vehículos se ve frenada por barreras sociales, financieras y legales, impidiéndoles viajar de forma independiente o trabajar como conductoras en servicios de transporte (IFC *et al.*, 2018). El panorama tampoco es alentador si se analiza el uso de otros modos – tales como el transporte público, caminata o bicicleta – ya que muchas mujeres se ven expuestas

a numerosas situaciones de violencia y acoso, afectando su percepción de seguridad e incluso sus patrones de movilidad. Por ejemplo, en Chile 6 de cada 10 mujeres ha sufrido algún tipo de acoso sexual en espacios públicos, más aún, entre un 40 y un 50 % las agresiones a mujeres ocurre en los buses, entre un 20 y 30 % en la calle y entre un 15 y 20 % en el metro (Ministerio de Transporte y Telecomunicaciones, 2018).

La aparición del ridesourcing como nueva alternativa de transporte podría significar una oportunidad para mejorar este panorama. Si bien la literatura sobre los impactos del ridesourcing ha avanzado bastante en los últimos años, la literatura dedicada al análisis desde una perspectiva de género aún es bastante escasa (Rizk *et al.*, 2018). En IFC *et al.* (2018) se realiza un análisis exhaustivo de los impactos del ridesourcing desde una perspectiva de género en seis países (Reino Unido, México, Egipto, Sudáfrica, India e Indonesia), destacando los beneficios y barreras tanto para conductoras como pasajeras. De acuerdo a los autores, la participación de mujeres como conductoras permite reducir la brecha de género en una industria claramente dominada por los hombres, además de potenciar sus ingresos; por otro lado, la principal barrera que impide una mayor participación es la percepción de inseguridad. En el caso de las pasajeras, los autores plantean que las mujeres se sienten más seguras en este modo de transporte gracias a que pueden conocer la información del conductor de antemano y a que el viaje está siendo rastreado. Sin embargo, las plataformas ridesourcing no han estado exentas de episodios de violencia y abuso sexual<sup>2</sup>, por lo cual se han impulsado iniciativas tales como plataformas ridesourcing conformadas sólo por conductoras, o bien solicitar específicamente una conductora a las plataformas tradicionales (Rizk *et al.*, 2018; Tirachini, 2020).

## 2.2. Estudios basados en encuestas

Uno de los principales obstáculos en la investigación relacionada con el uso de ridesourcing ha sido la falta de datos de viajes. A pesar de contar con información de los usuarios y de las características de los viajes realizados, las compañías de ridesourcing han sido reacias a liberar dichos datos con el fin de resguardar la privacidad de sus usuarios (Lavieri y Bhat, 2019) y sus propios intereses. En consecuencia, la mayoría de la literatura está basada en datos recolectados a través de encuestas. Este método de recolección de datos tiene la ventaja de ofrecer información más detallada a nivel de usuario, aunque su nivel de cobertura es menor y suele estar más expuesto a sesgos. Debido a estas características, los estudios basados en encuestas siguen tres direcciones: (1) caracterización de usuarios de ridesourcing y sus propósitos de viaje, (2) comportamiento de viajes; y (3) impactos del ridesourcing (Li *et al.*, 2019). Las última fue abordada en la sección anterior, por lo que en esta sección se hará énfasis en el primer y segundo tópico.

Según Lavieri y Bhat (2019), gran parte de los estudios basados en encuestas son de naturaleza exploratoria, por lo que sus resultados se apoyan en estadística descriptiva (por ejemplo, Rayle *et al.*, 2016; Clewlow y Mishra, 2017; Young y Farber, 2019); mientras que

---

<sup>2</sup>Entre 2017 y 2018, se reportaron cerca de 6.000 casos de abuso sexual y casi 500 casos de violación en viajes en Uber realizados en Estados Unidos, ver [https://www.cnnchile.com/mundo/uber-informe-agresion-sexual-estados-unidos\\_20191206/](https://www.cnnchile.com/mundo/uber-informe-agresion-sexual-estados-unidos_20191206/), fecha de acceso 17-03-2022

un grupo más pequeño ha logrado desarrollar modelos econométricos a nivel individual (por ejemplo, [Dias et al., 2017](#); [Alemi et al., 2018a](#); [Aguilera-García et al., 2022](#)). A pesar de las diferencias de metodología, las conclusiones obtenidas son similares en términos de caracterización de usuarios y comportamiento de viajes, permitiendo definir un perfil de demanda en este modo.

En relación a las características de los usuarios, los resultados de [Rayle et al. \(2016\)](#) sugieren que los usuarios de ridesourcing en San Francisco son más jóvenes y poseen mejor educación y mayor poder adquisitivo que el promedio de la población. [Clewlow y Mishra \(2017\)](#) obtienen que sólo el 4% de las personas mayores de 65 años han utilizado estos servicios, versus un 36% por parte de la población entre 18 y 29 años; además, la población de mayores ingresos y nivel educativo utilizan estos servicios hasta con dos veces más frecuencia que la población de menores ingresos y nivel educativo. [Dias et al. \(2017\)](#), por su parte, a través de un modelo probit bivariado, afirman que los usuarios de ridesourcing tienden a ser jóvenes, de ingresos y nivel educativo alto, además de vivir en sectores con gran densidad poblacional. Por su parte, [Alemi et al. \(2018a\)](#) estudia los factores que influyen en el uso de Uber y Lyft a través de un modelo logit binario, cuyos resultados indican que los *millennials* de alto nivel educativo y poder adquisitivo son más propensos a utilizar estos servicios.

La mayoría de los estudios se han desarrollado en ciudades estadounidenses, aunque también se han encontrado investigaciones en Chile, Canadá y España, entre otros. Si bien el contexto es diferente en cada país, los resultados no difieren demasiado entre sí. [Young y Farber \(2019\)](#) obtienen que los usuarios de ridesourcing en Toronto suelen ser adultos jóvenes (entre 20 y 39 años), de ingreso alto y además sin posesión de automóvil en el hogar. En Santiago de Chile, [Tirachini y del Río \(2019\)](#) obtienen que los usuarios más jóvenes y con mayor poder adquisitivo utilizan esta alternativa con mayor frecuencia; más aún, los usuarios de Cabify poseen mayores ingresos que los usuarios de Uber, atribuyendo este hecho a que Uber admite el pago mediante efectivo, mientras que Cabify sólo acepta tarjetas de débito o crédito. Por último, [Aguilera-García et al. \(2022\)](#) sugieren que en España hay mayor tendencia a utilizar plataformas ridesourcing por parte de mujeres, adultos jóvenes y personas de ingreso alto cuando el viaje a realizar no es de carácter urgente; mientras que las personas mayores de 49 años utilizan estos servicios cuando su viaje requiere algo más de urgencia.

Con respecto al comportamiento de viajes, diversos autores coinciden en que los viajes en ridesourcing son de carácter mayoritariamente ocasional y que sus principales propósitos son el ocio y el retorno al hogar ([Rayle et al., 2016](#); [de Souza Silva et al., 2018](#); [Tirachini y del Río, 2019](#); [Aguilera-García et al., 2022](#)). Los viajes de retorno al hogar se pueden asociar a actividades laborales, pero principalmente a actividades sociales, de hecho, una de las principales razones por la cual los usuarios optan por utilizar el ridesourcing en lugar de conducir sus propios vehículos es el consumo de alcohol ([Clewlow y Mishra, 2017](#); [Tirachini y del Río, 2019](#)). Este comportamiento se refleja en el patrón de viajes, ya que los mayores niveles de demanda se registran durante los períodos nocturnos y fines de semana ([Tirachini y del Río, 2019](#); [Bi y Ye, 2021](#)).

## 2.3. Modelos de demanda basados en datos de viajes

Durante el último tiempo, algunas compañías de ridesourcing han liberado datos espacio-temporales de viajes. Hasta el momento, esto se ha dado únicamente con datos de viajes en ciudades de Estados Unidos (Uber en Nueva York; RideAustin en Austin, Texas; Lyft en Los Ángeles, California; múltiples compañías en Chicago) y China (Didi Chuxing en Chendu). Estos datos se caracterizan por su gran cobertura en términos de espacio y tiempo, ya que típicamente contienen información de todos los viajes realizados en la ciudad correspondiente durante un período determinado, lo que ha permitido a los investigadores desarrollar modelos econométricos a nivel agregado. Sin embargo, el gran inconveniente que poseen estos datos es que la información de los usuarios se encuentra anonimizada.

Si bien en la literatura previa existen varios trabajos en que se analiza la demanda de viajes en taxis a través de datos espacio-temporales (por ejemplo, [Correa et al., 2017](#); [Chen et al., 2021](#)), el primer trabajo de este tipo aplicado a viajes en ridesourcing del cual se tiene registro es el de [Correa et al. \(2017\)](#). En este estudio, se usan datos de viajes en Uber y taxis en Nueva York entre 2014 y 2015 para estimar la generación de viajes a través de modelos de regresión lineal y de correlación espacial. Los resultados indican que el uso de ridesourcing está ligado a sectores de mayor ingreso, nivel educativo y con menor posesión de automóvil, así como también los sectores con más puestos de empleos, oferta de vías y paraderos de buses.

[Gerte et al. \(2018\)](#) utiliza la misma base de datos para elaborar un modelo lineal con efectos aleatorios mediante mínimos cuadrados generalizados, aunque considera variables distintas a las utilizadas en [Correa et al. \(2017\)](#), dando mayor relevancia a factores temporales como el clima. Sus resultados muestran que una mayor demanda está relacionada con los sectores con mayor población joven, con mayor presencia masculina y nivel educativo, además de los sectores con mayor uso de suelo residencial y de retail. En cuanto a las variables temporales, se observa un mayor uso de ridesourcing cuando las condiciones climáticas no son favorables, como en invierno o en días de lluvia.

Posteriormente, la liberación de datos de viajes mediante el servicio RideAustin en Austin, Texas entre 2016 y 2017, permitió el desarrollo de nuevas investigaciones en este ámbito. [Lavieri et al. \(2018\)](#) describen la generación de viajes a través de un modelo de conteo con retardo espacial, tanto para día laboral como fin de semana. En este estudio también se analiza la distribución de viajes mediante un modelo de partición fraccionaria. Los resultados apuntan a que la demanda de ridesourcing es mayor en las zonas con población más joven y menor disponibilidad de automóvil propio; además de las zonas con mayor densidad poblacional, densidad de empleos (de retail) y con presencia de parques. Un resultado llamativo de este trabajo está relacionado con el ingreso: durante los días laborales la demanda de ridesourcing aumenta con el ingreso, mientras que durante el fin de semana se tiene el efecto contrario. Se observa un comportamiento similar en la composición por género: durante los días laborales se tiene una relación positiva entre la demanda y la proporción de varones, mientras que durante el fin de semana esta relación es negativa, lo que indica una mayor influencia de la población femenina en los viajes durante este período.

En [Yu y Peng \(2019\)](#) se usa la misma base de datos, aunque considerando un enfoque de modelación diferente. En este trabajo se utiliza una regresión Poisson geográficamente



ponderada para dar cuenta de una eventual correlación espacial de las variables, además se incorporan variables de entorno construido, tales como densidades de población y empleo, uso de suelo e infraestructura. A nivel de sociodemografía, los resultados indican que existe mayor demanda de viajes en los sectores de población más joven y con mayor nivel educativo, mientras que para el ingreso se observan dos efectos: durante los días laborales existe una relación negativa y durante el fin de semana la relación es positiva, contrario a lo obtenido en [Lavieri \*et al.\* \(2018\)](#). En cuanto a entorno construido, existe mayor demanda en zonas con mayor densidad de población, diversidad de actividades y accesibilidad al transporte público.

[Brown \(2019\)](#) utiliza una regresión binomial negativa para estimar la generación de viajes en Los Ángeles, California, usando datos de viajes en Lyft realizados entre septiembre y noviembre de 2016. Sus resultados indican que los vecindarios con mayor presencia de población joven, mayor densidad de población y de empleos, con mayor presencia de paraderos y con menor disponibilidad de estacionamientos están asociadas a una mayor demanda de viajes.

Más tarde, [Ghaffar \*et al.\* \(2020\)](#) estudian los viajes realizados en Chicago entre noviembre de 2018 y marzo de 2019, para desarrollar una regresión binomial negativa con efectos aleatorios y así modelar el uso de ridesourcing en dicha ciudad, incorporando elementos como el clima y la delincuencia. Reafirmando los resultados de investigaciones anteriores, los resultados de su trabajo sugieren que existe mayor demanda de viajes en los sectores con mayor ingreso, mayor presencia de hogares sin automóvil, mayor densidad de población y de empleos, mayor diversidad de actividades, mayor presencia de restaurantes y menor disponibilidad de estacionamientos, además de las zonas con mayor cantidad de homicidios. Adicionalmente, se observa mayor demanda de viajes durante los fines de semana y durante los días con temperaturas más bajas y menor precipitación.

Con la misma base de datos, [Marquet \(2020\)](#) opta por elaborar un modelo de distribución de viajes a través de una regresión poisson truncada, considerando elementos relacionados con la sociodemografía y el entorno construido, con especial énfasis en las características de los vecindarios en donde comienza y finaliza el viaje, además de la “caminabilidad” del viaje, definida como un indicador de la accesibilidad. Sus resultados sugieren una asociación positiva entre la demanda de viajes en un par origen-destino con el ingreso, la densidad de población, la densidad de empleo y la caminabilidad. A diferencia de los trabajos anteriores, se obtiene una relación negativa entre la demanda y la ausencia de automóvil en el hogar.

[Sabouri \*et al.\* \(2020\)](#) utilizan datos de viajes de Uber en 24 ciudades de Estados Unidos entre enero y marzo de 2017 para construir un modelo lineal jerárquico, incluyendo variables de sociodemografía y entorno construido. Además, incluye el tiempo de viaje como atributo de modo. Los resultados relacionan una mayor demanda de viajes con sectores con mayor población y empleo, ingreso promedio, mix de uso de suelo y densidad de paraderos; mientras que se relaciona negativamente con la edad promedio y la posesión de automóvil. Naturalmente, también se obtuvo una relación negativa entre la demanda y el tiempo de viaje.

Los estudios mencionados en esta sección utilizan bases de datos de ciudades estadounidenses. Hasta la fecha, sólo se tiene conocimiento de un estudio de este tipo realizado con datos de viajes en otro país, realizado por [Wang y Noland \(2021\)](#). Los autores utilizan datos de viajes en Didi Chuxing realizados en Chengdu, China, durante noviembre de 2016 para estimar un modelo de regresión lineal por mínimos cuadrados ordinarios y una regresión

geográficamente ponderada. Sorprendentemente, no se consideran variables de tipo sociodemográfica, a excepción del valor de las viviendas como proxy de la condición socioeconómica de cada unidad espacial. Se obtiene que la demanda se relaciona positivamente con la densidad de población, el valor de las viviendas, la cantidad de restaurantes y la proximidad a estaciones de metro. Sin embargo, un resultado llamativo es que no se observa correlación entre la demanda y el mix de uso de suelo.

Los modelos utilizados y las variables consideradas en cada uno de estos trabajos se reportan en la Tabla [2.1](#).



Tabla 2.1: Estudios sobre modelación de demanda en base a datos de viajes

Autores	Área de Estudio	Período	Variable Explicada	Modelo(s)	Sociodemografía	Entorno Construido	Oferta de Transporte	Otros
Correa et al (2017)	Nueva York	Abr 2014	Viajes totales	Modelo de retardo espacial	Ingreso (+)	Total Empleos (+)	Num. paraderos (+)	Tiempo de acceso a TP (+)
		Sep 2014			Nivel Educativo (+)	Oferta vías (+)		
Gerte et al (2018)	Nueva York	Ene 2015	Viajes semanales	Regresión Lineal GLS con efectos aleatorios	Pos. Automóvil (-)	Superficie Residencial (+)	Superficie Retail (+)	Temperatura (-)
		Jul 2015			Tamaño de Hogar (+)	Superficie Retail (+)		
Lavrieri et al (2018)	Austin	Ago 2016	Viajes diarios (lab y fds)	Modelo de conteo con retardo espacial	Edad (-)	log  Densidad poblacional (+)	Frecuencia buses (-)	Lluvia (+)
		Ene 2017			% Hombres (+/-)	log  Densidad empleos retail (+)		
Yu y Peng (2019)	Austin	Sep 2016	Viajes diarios	Regresión Poisson geográficamente ponderada	Ingreso (+)	Universidad de Texas (+)	Densidad paraderos	log  Densidad paraderos
		Mar 2017			Pos. Automóvil (-)	Densidad poblacional (+)		
Brown (2019)	Los Ángeles	Sep 2016	Viajes per cápita	Regresión Binomial Negativa	Mix población/empleos (+)	Mix población/empleos (+)	Distancia de acceso (-)	Temperatura (-)
		Nov 2016			Edad (-)	log  Densidad poblacional (+)		
Marquet (2020)	Chicago	Nov 2018	Viajes diarios	Regresión Poisson Truncada	Ingreso (+)	Densidad retail (+)	Densidad de acceso (-)	Lluvia (-)
		Jun 2019			Pos. Automóvil (+)	Densidad oficinas (-)		
Ghaffar (2020)	Chicago	Nov 2018	Viajes diarios	Regresión Binomial Negativa con efectos aleatorios	Pos. Automóvil (-)	Densidad industrias (+)	Num. paraderos (+)	Estacionamientos (-)
		Mar 2019			Tamaño de Hogar (-)	Caminabilidad (+)		
Sabouni et al (2020)	EE.UU.	Ene 2017	log  Viajes mensuales	Modelo Lineal Jerárquico	Ingreso (+)	log  Densidad poblacional (+)	Densidad paraderos (+)	Tiempo de Viaje (-)
		Mar 2017			Edad (-)	log  Densidad empleos (+)		
Dey (2021)	Nueva York	Ene 2015	Viajes totales	Regresión Binomial Negativa	Pos. Automóvil (-)	Mix de Uso de Suelo (+)	Num. estaciones de metro (-)	Nieve (-)
		Dic 2018			Ingreso (+)	Restaurantes (+)		
Wang y Noland (2021)	Chengdtú	Nov 2016	log  Viajes totales (por periodo)	Regresión lineal geográficamente ponderada	Pos. Automóvil (-)	Restaurantes (+)	Cobertura bus (+)	Cobertura metro (+)
		Nov 2016			Ingreso (+)	Aeropuerto (+)		
					Densidad poblacional (+)	Deportes y Entrenamiento (+)		

# Capítulo 3

## Metodología

El tercer capítulo de esta tesis tiene por objetivo describir detalladamente las actividades realizadas durante la presente investigación. Para comenzar, en la Sección 3.1 se detallan las actividades relacionadas con los procesos de recolección, exploración, procesamiento y descripción de los datos. Posteriormente, en la Sección 3.2 se abordan las labores relacionadas con la modelación de la demanda, específicamente la formulación y estimación de los modelos, además de evaluar la capacidad predictiva de dichos modelos.

### 3.1. Datos

#### 3.1.1. Recolección de Datos

##### **Viajes en ridesourcing**

La base de datos principal contiene datos de viajes en ridesourcing, los cuales fueron facilitados por Didi Chuxing a través de un convenio con la Universidad de Chile. Concretamente, esta base de datos contiene información de las solicitudes de viaje realizadas a través de la aplicación en la Región Metropolitana de Santiago, entre los días 13 y 26 de enero de 2020. En total se tienen 1.369.643 solicitudes, donde cada una cuenta con 67 atributos. Sin embargo, debido a que la base de datos no contaba con un diccionario, no fue posible identificar todos los atributos.

Dentro de los atributos más relevantes se encuentran identificadores (ID's) de la solicitud, del usuario y del conductor; la fecha en que se realiza la solicitud, dirección del origen, distancia recorrida entre origen y destino, duración del viaje predicha por la aplicación y tarifa total. También se cuenta con los instantes de diferentes etapas del viaje, tales como el de solicitud del servicio, llegada del vehículo, cancelación del viaje (si corresponde) y finalización del viaje, entre otros. Sin embargo, para resguardar la privacidad de los usuarios, su información se encuentra anonimizada y tampoco se cuenta con información del destino del viaje.

Además, a partir de los atributos entregados en la base de datos se construyen nuevas variables, tales como el tiempo de espera, el tiempo real de viaje y la velocidad media. En efecto, el tiempo de espera se calcula como el tiempo transcurrido entre la solicitud del servicio y la llegada del vehículo; de manera similar, el tiempo real de viaje se calcula como el tiempo transcurrido entre la llegada del vehículo y la finalización del viaje, mientras que la velocidad media se calcula como la distancia recorrida dividida en el tiempo real de viaje. Estas nuevas variables permiten caracterizar los viajes de forma más precisa, establecer criterios para la limpieza de datos e incluso incorporar información del propio modo en el modelo.

La Tabla 3.1 muestra una descripción de los atributos identificados en esta base de datos, además de las variables construidas, junto con un ejemplo de una observación.

Tabla 3.1: Atributos de ridesourcing

Atributo	Tipo	Definición	Ejemplo
order_id	int	ID de solicitud	87967099937070
passenger_id	int	ID de usuario	87961068791715
driver_id	int	ID de conductor	650910951289411
stat_start_hour	datetime	Fecha del viaje	2020-01-16
starting_name	string	Dirección de origen	Portales 5225_Lo Prado_ Chile
county	string	Comuna de origen	LO PRADO
start_dest_dis	float	Distancia OD [km]	1,115
forecast_dur	int	Duración predicha [min]	2
total_fee	int	Tarifa	1100
call_time	datetime	Instante en que se solicita el servicio	2020-01-16 11:18:01
arrive_time	datetime	Instante en que llega el vehículo	2020-01-16 11:23:25
cancel_time	datetime	Instante en que se cancela el viaje	1971-01-01 00:00:00
finish_time	datetime	Instante de finalización del servicio	2020-01-16 11:29:36
wait_time	float	Tiempo de espera [min]	5,2
travel_time	float	Tiempo de viaje real [min]	6,2
speed	float	Velocidad media [km/hr]	10,8

Los datos de viajes en ridesourcing se complementan con otras fuentes de datos que incluyen características de: (1) sociodemografía, (2) entorno construido y (3) oferta de transporte. Estos datos se imputan a las observaciones a través del origen del viaje: si bien esto es coherente para las variables de entorno construido y de oferta de transporte, es importante tener en cuenta que para los datos de sociodemografía lo sería sólo para los viajes basados en el hogar, los cuales ocurren principalmente durante la punta mañana de la ciudad. A continuación, se detalla el proceso de recolección de las variables de cada una de estas dimensiones.

## Sociodemografía

Este conjunto de características incluye aspectos relacionados con las personas y hogares que conforman el área de estudio, tales como el ingreso, edad, nivel educativo y posesión de automóvil, entre otros. De acuerdo a la literatura (Rayle *et al.*, 2016; Yu y Peng, 2019), estas variables suelen ser buenos predictores para explicar la demanda de viajes en ridesourcing.

Los datos de sociodemografía se obtienen de la Encuesta Origen-Destino de viajes del 2012 para Santiago (SECTRA, 2013), específicamente las bases de datos de hogares y de personas. Si bien el objetivo principal de la Encuesta Origen-Destino es caracterizar el comportamiento y patrón de viajes de la población de la Región Metropolitana, proporciona información socioeconómica valiosa que puede ser georreferenciada.

La base de datos de hogares contiene 18.264 observaciones, donde cada una corresponde a un hogar encuestado. Cada observación contiene 21 atributos, dentro de los cuales destacan el ID del hogar, coordenadas geográficas, número de personas, número de automóviles e ingreso per cápita. En base a esta información, se definen las siguientes categorías de hogares:

- Ingreso: Se considera la clasificación utilizada en la EOD 2012, dividiendo los umbrales por el tamaño de hogar promedio (3,23 habitantes por hogar), obteniendo las siguientes categorías de ingreso:
  - Bajo: Ingreso familiar bajo 400.000 CLP, equivalente a un ingreso per cápita bajo 123.839 CLP
  - Medio: Ingreso familiar entre 400.001 CLP y 1.600.000 CLP, equivalente a un ingreso per cápita entre 123.840 CLP y 495.356 CLP.
  - Alto: ingreso familiar sobre 1.600.000 CLP, equivalente a un ingreso per cápita sobre 495.356 CLP.
- Posesión de automóvil: Sin auto, 1 auto, 2 o más autos.

Por otra parte, la base de datos de personas contiene 60.054 observaciones, donde cada una corresponde a una persona encuestada. Cada observación posee 39 atributos, siendo los más relevantes para este trabajo el ID de la persona, el ID del hogar en el que vive, edad, sexo y nivel educativo. En este caso, se definieron las siguientes categorías de personas:

- Sexo: Masculino y Femenino
- Edad: 18–39 años, 40–59 años, 60+ años
- Nivel Educativo: Educación superior terminada, educación superior no terminada.

## Entorno Construido

El entorno construido es un concepto multidimensional que captura elementos del entorno en donde las personas viven y realizan sus actividades, tales como el uso de suelo, edificaciones y calles. La inclusión de estas variables es relevante, ya que el entendimiento del impacto del entorno construido sobre los viajes permite modelar de mejor manera la demanda de viajes y elaborar estrategias de planificación más eficientes (Yu y Peng 2019). Para efectos de este trabajo, se recolecta información relacionada con la densidad poblacional, el uso de suelo y algunos sitios de interés.

La información de población se obtiene del Censo 2017 (Instituto Nacional de Estadísticas 2018). En el sitio web de esta institución, se encuentra disponible un archivo en formato shp que contiene un mapa de la Región Metropolitana zonificada por distritos censales (en

adelante DC). Este archivo cuenta además con una base de datos que reporta información de la población y superficie de cada DC, con lo cual es posible calcular la densidad poblacional de cada una de estas unidades de análisis.

Los datos de uso de suelo se obtienen a partir de un archivo en formato shp disponible para todo público en el sitio web del [Centro de Desarrollo Urbano Sustentable \(2016\)](#). Este archivo contiene información del uso de suelo de Santiago para el año 2014 y está construida a partir de la Base de Datos de Bienes Raíces del Servicio de Impuestos Internos. Las observaciones se reportan a nivel de predio, considerando variables como la superficie total, el avalúo fiscal y la superficie destinada a cada actividad, entre otros. En la Tabla 3.2 se reportan los porcentajes de cada actividad realizada en el área de estudio. Cabe señalar que las actividades reportadas corresponden a las activamente desarrolladas, es decir, aquellas que son relevantes en la generación y/o atracción de viajes en ridesourcing, mientras que dentro de la categoría Otros se han incluido las actividades de Bodega y Almacenaje (L), Estacionamientos (Z) y Transporte y Telecomunicaciones (T), Sitios Eriazos (W) y otras actividades no especificadas en la base de datos.

Tabla 3.2: Composición del área de estudio según uso de suelo.

Actividad	Símbolo	% Superficie
Habitacional	H	56,53 %
Comercio	C	10,86 %
Industria	I	9,24 %
Oficina	O	4,09 %
Educación y Cultura	E	3,65 %
Salud	S	0,86 %
Hotel	G	0,56 %
Culto	Q	0,53 %
Administración Pública y Defensa	P	0,52 %
Transporte y Telecomunicaciones	T	0,44 %
Deporte y Recreación	D	0,38 %
Otros		12,34 %
Total		100,00 %

Adicionalmente, se construye de manera manual un catastro de localizaciones específicas que podrían constituir una fuente importante de generación de viajes en ridesourcing, conocidas en la literatura como Puntos de Interés o POI's, por sus siglas en Inglés. En este caso, se consideraron aeropuertos, terminales de buses y centros comerciales. Esta base de datos contiene como mínimo las coordenadas geográficas de cada uno de estos lugares.

En el caso de los aeropuertos, sólo se tiene en consideración el Aeropuerto Internacional Arturo Merino Benitez, el cual se encuentra en la comuna de Pudahuel. En cuanto a los terminales de buses, se identifican cuatro dentro del área de estudio: (1) Terminal San Borja, ubicado en la comuna de Estación Central; (2) Terminal Alameda, ubicado en la misma comuna; (3) Terminal Pajaritos, ubicado en la comuna de Lo Prado; y (4) Terminal Los Héroes, el cual se encuentra en la comuna de Santiago. En ambos casos, se registran sus coordenadas geográficas desde Google Maps.

Por otro lado, para los centros comerciales se registran las coordenadas geográficas y su GLA<sup>1</sup>. Luego, usando como base la clasificación definida por el Consejo Internacional de Centros Comerciales (2017), a cada centro comercial se le asigna una de las siguientes categorías de acuerdo a su GLA:

1. Centro Local: GLA menor a 40.000 m<sup>2</sup>.
2. Regional: GLA entre 40.000 y 80.000 m<sup>2</sup>.
3. Superregional: GLA mayor a 80.000 m<sup>2</sup>.

## Oferta de transporte

Finalmente, se incorporan variables de oferta de otros modos de transporte, con el objetivo de capturar su relación con el uso de ridesourcing, es decir, si son sustitutos o complementarios. Los modos de transporte considerados en este trabajo son transporte público y bicicleta.

Con respecto al transporte público, se recopilan datos de la Red Metropolitana de Movilidad – en ese entonces Transantiago – y del Metro de Santiago. En el caso de los buses, se consideran dos variables: infraestructura (paraderos) y disponibilidad (frecuencia). Los datos de paraderos se obtienen del programa de operación del segundo semestre del 2019, el cual está disponible en el sitio web del Directorio de Transporte Público Metropolitano (2019). Esta base de datos está conformada por 11.274 observaciones y contiene información de la ubicación geográfica y los servicios que se detienen en cada uno de ellos. Por otra parte, desde el mismo sitio web se obtienen datos de frecuencias de cada recorrido durante enero del 2020.

En cuanto a la disponibilidad de metro, se obtienen datos de las estaciones del Metro de Santiago a partir de un archivo en formato shp disponible en el sitio web del Observatorio de Ciudades UC (2019). Este archivo cuenta con una base de datos que contiene el nombre, las coordenadas geográficas y la línea a la que pertenece cada estación. A la fecha, la red del Metro de Santiago cuenta con 136 estaciones, repartidas en siete líneas.

Por último, la factibilidad de viajar en bicicleta se captura mediante la oferta de ciclovías en la ciudad. Dicha información se obtiene del Visor de Ciclovías (Ministerio de Transporte y Telecomunicaciones, 2021), el cual contiene el catastro de ciclovías de todo el país.

### 3.1.2. Procesamiento de datos

#### Definición del área de estudio y zonificación

Inicialmente, se considera como área de estudio la Región Metropolitana de Santiago, debido a la cobertura de los datos de viajes en ridesourcing. Sin embargo, se decide acotar el análisis a las 32 comunas de la provincia de Santiago y las comunas de San Bernardo y Puente Alto, ya que su composición es mayoritariamente urbana y existe una mayor calidad y

---

<sup>1</sup>La superficie bruta alquilable (SBA o GLA por sus siglas en inglés), corresponde a la superficie disponible para la venta de productos y servicios. No incluye la superficie destinada a estacionamientos.

disponibilidad de información en las bases de datos complementarias. La zonificación utilizada para esta área de estudio es por distritos censales definida por el INE para el Censo 2017, contando con 339 zonas entre las 34 comunas. Finalmente, los datos de viajes, así como también los datos de sociodemografía, uso de suelo e infraestructura de transporte público son mapeados a cada distrito censal mediante el software QGIS.

## Geocodificación y Limpieza de datos

Dado que la base de datos de viajes en ridesourcing cuenta con la dirección de la solicitud, se realiza un proceso de geocodificación de viajes mediante el complemento “Geocode by Awesome Table” de Google Sheets. De esta manera, se tienen las coordenadas geográficas del origen de cada viaje realizado, permitiendo georreferenciar cada observación.

Posteriormente, se realiza una limpieza de los datos considerando una serie de filtros, con el fin de conservar los viajes cuyas características sean consistentes para efectos de modelación. Como resultado de la geocodificación, se eliminaron 10.681 observaciones por errores del proceso (0.8 % de las observaciones). Luego, se dejaron de lado 352.328 observaciones correspondientes a viajes cancelados (25.7 % del total de observaciones). Como tercera etapa de filtros, se eliminaron viajes cuyas características son consideradas como inverosímiles:

- Viajes con distancia recorrida inferior a 200 [m] o superior a 50 [km]
- Viajes con tiempo de viaje superior a 120 [min]
- Viajes con velocidad media inferior a 3 [km/h] o superior a 100 [km/h]
- Viajes con tiempo de espera superior a 30 [min]

La aplicación de los cuatro filtros de atributos descartan 55.470 observaciones (5.5 % de los viajes realizados), sin embargo, es necesario tener en cuenta que gran parte de estas observaciones fueron descartadas por más de un criterio. La Figura 3.1 muestra un diagrama de Venn que contiene la cantidad de observaciones descartadas de acuerdo a la cantidad de filtros considerados. De la figura se puede observar que de las 55.470 observaciones descartadas, 55.046 fueron descartadas únicamente por el filtro de velocidad, lo que equivale al 99.2 % de las observaciones eliminadas.

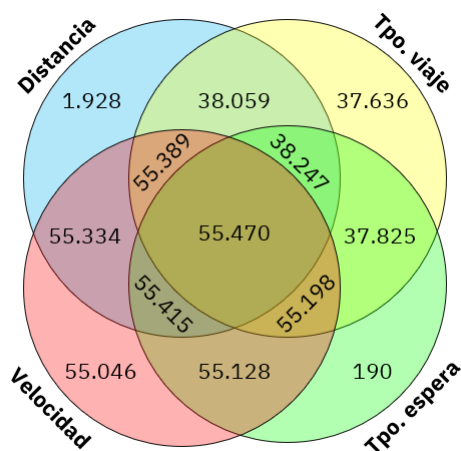


Figura 3.1: Número de observaciones descartadas según filtros aplicados

Finalmente, se descartan los viajes cuyo origen se encuentra fuera del área de estudio, correspondientes a 6.685 observaciones (0.7% de los viajes realizados). De esta manera, la base de datos se reduce a 944.479 observaciones. El gráfico de la Figura 3.2 resume el proceso de limpieza de la base de datos de ridesourcing.

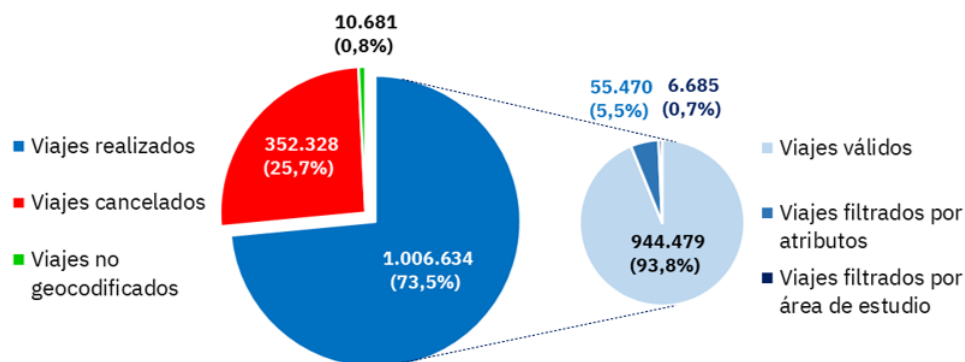


Figura 3.2: Limpieza de base de datos de viajes en ridesourcing

## Construcción de variables a nivel de DC

Una vez que los datos han sido georreferenciados, se construyen variables a nivel de DC, las que finalmente son incluidas en los modelos de demanda. A partir de los datos de viajes en ridesourcing, se construyen las variables de número de viajes, número de usuarios, número de conductores, distancia recorrida promedio, tiempo de viaje promedio, velocidad media espacial y tiempo de espera promedio.

Por otra parte, a partir de los datos de sociodemografía se calculan los valores promedio del ingreso y la tasa de motorización, así como también las composiciones porcentuales de categorías de ingreso, de tasa de motorización, de género, de edad y de nivel educativo.

En cuanto a las variables de entorno construido, se obtienen el número de aeropuertos, terminales de buses y centros comerciales en cada DC, además de la composición según uso de suelo. Más aún, se calcula el mix de uso de suelo, un indicador desarrollado por [Cervero y Kockelman \(1997\)](#) en base a la entropía de Shannon, que mide la diversidad de actividades del DC. Este indicador está dado por la expresión de la Ecuación 3.1:

$$land\ use\ mix = \sum_{j=1}^J \frac{P_j \cdot \ln P_j}{\ln J} \quad (3.1)$$

donde  $P_j$  es la propoción del  $j$ -ésimo uso de suelo.

Además, para corregir parte del efecto tamaño, se calcula para cada DC el porcentaje de superficie destinado a uso urbano, como la razón entre la suma de las áreas de los polígonos<sup>2</sup> de predio de un DC y el área del polígono de dicho DC. Luego, se calcula la superficie urbana como la superficie real del DC ponderada por el porcentaje urbano.

<sup>2</sup>Estructuras vectoriales de QGIS que consisten en secuencias cerradas de vértices.



Por último, como variables de oferta de transporte se calcula el número de paraderos y de estaciones de metro, además de la frecuencia promedio de buses en cada DC. También se calcula la densidad de paraderos, definida como el número de paraderos por unidad de superficie urbana. Del mismo modo, se calcula la densidad de ciclovías.

### 3.1.3. Análisis Descriptivo

#### Estadística Descriptiva

En primer lugar, en la Tabla 3.3 se reporta la estadística descriptiva de los atributos de viajes en ridesourcing. Un resultado interesante de esta tabla es que los indicadores estadísticos del tiempo de viaje real son sistemáticamente mayores que el tiempo de viaje predicho por la aplicación, lo que da indicios de la influencia de la congestión.

Tabla 3.3: Estadística descriptiva de atributos de viaje en ridesourcing

Variable	Prom.	Std.	Min.	25 %	50 %	75 %	Max.
Tarifa [CLP]	2.697	1.945	1.056	1.394	2.088	3.267	55.718
Distancia recorrida [km]	6,4	5,3	0,2	2,9	4,7	8,0	50,0
Tiempo de viaje predicho [min]	12,7	7,3	1,0	7,0	11,0	16,0	109,0
Tiempo de viaje real [min]	15,0	9,4	0,2	8,4	12,8	19,3	119,3
Tiempo de espera [min]	4,5	2,6	0,1	2,7	4,0	5,7	30,0
Velocidad media [km/hr]	24,8	9,9	3,0	18,4	23,2	29,1	100,0

Para complementar esta información, en las Figuras 3.3, 3.4, 3.5, 3.6 y 3.7 se reportan los histogramas de distancia, tiempo de viaje real, tiempo de espera y velocidad media, respectivamente. En cada uno de ellos se señala la posición del promedio (negro), del 25 % (azul), del 50 % (verde) y del 75 % (rojo). Se observa que todos los histogramas presentan una forma log-normal, como consecuencia de la acumulación de observaciones en los valores más bajos, pero con valores máximos bastante altos. De hecho, en todos los histogramas se aprecia que la moda es cercano al 25 % y el promedio cubre más del 50 % de las observaciones.

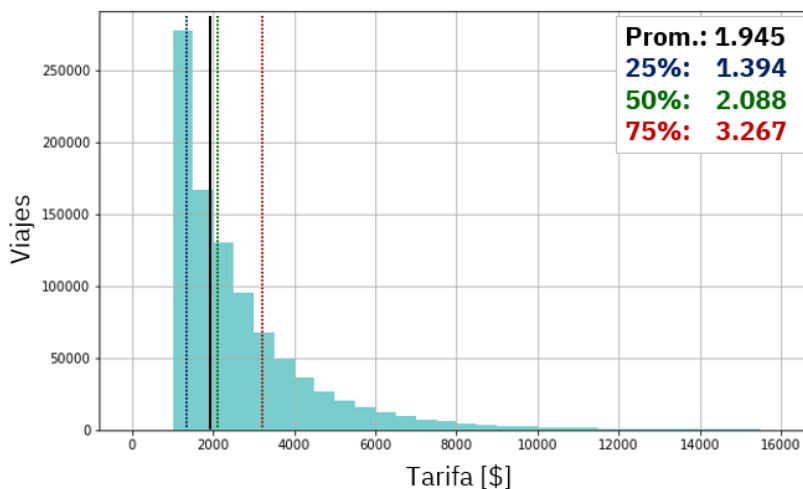


Figura 3.3: Histograma de tarifa del viaje

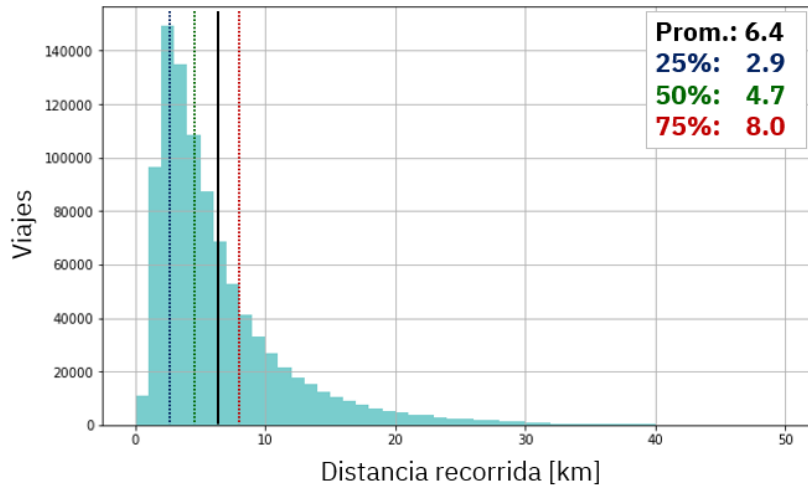


Figura 3.4: Histograma de distancia recorrida

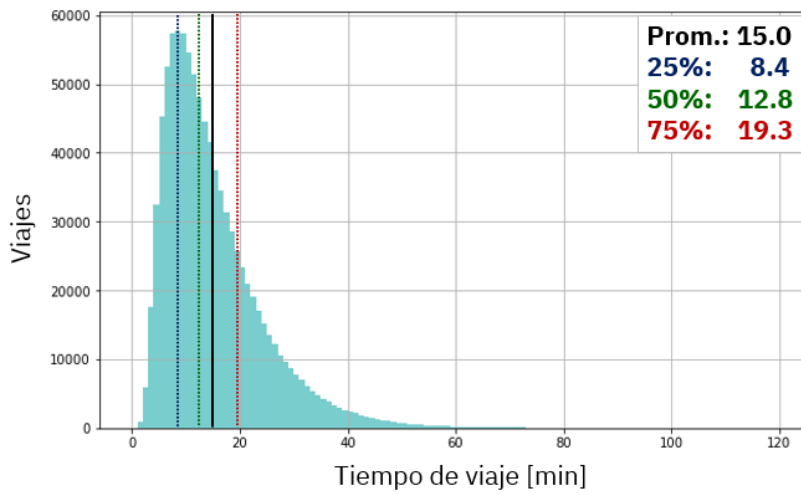


Figura 3.5: Histograma de tiempo de viaje real

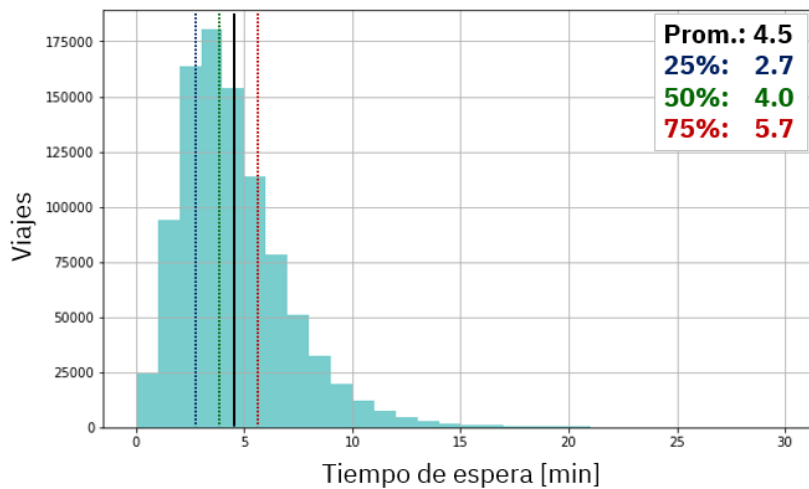


Figura 3.6: Histograma de tiempo de espera

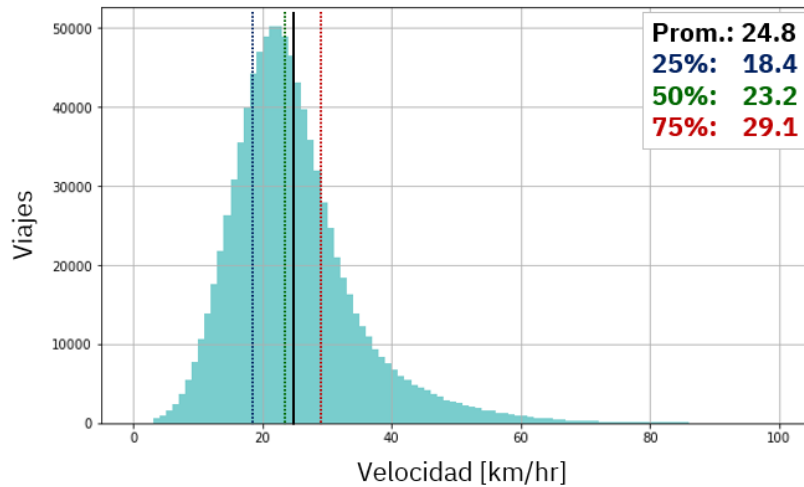


Figura 3.7: Histograma de velocidad media

Además, en la Figura 3.8 se reporta un histograma de frecuencia de uso, donde se observa una forma exponencial, lo que permite afirmar que el uso de ridesourcing es preferentemente ocasional. De hecho, de los 303.471 usuarios que utilizaron el servicio durante estas dos semanas, el 42 % utiliza el servicio solo una vez, mientras que el 62.3 % no lo utiliza más de dos veces dentro de este período de análisis.

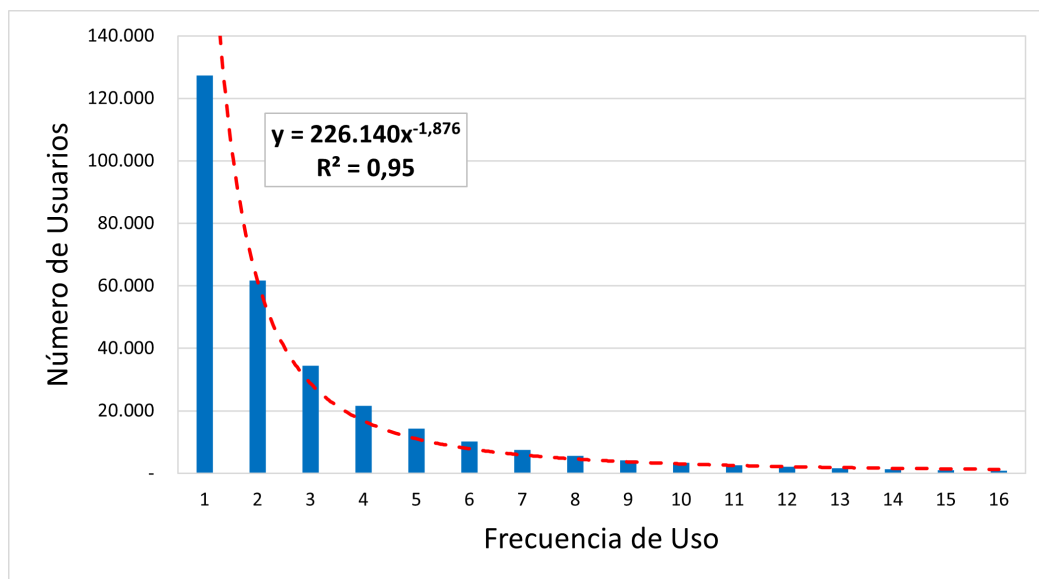


Figura 3.8: Histograma de frecuencia de uso

Finalmente, en la Tabla 3.4 se reporta la estadística descriptiva de las variables consideradas a nivel de DC en la sección 3.1.2.

Tabla 3.4: Descripción y Estadística Descriptiva de las variables construidas

Variable	Descripción	Prom.	Std.	Min.	25 %	50 %	75 %	Max.
<b>Viajes en ridesourcing</b>								
trips	Viajes totales	2.778	2.253	200	1.414	2.155	3.331	14.011
daily_trips	Viajes diarios	198	161	14	101	154	238	1.001
daily_trips_per_km2	Viajes diarios por km <sup>2</sup>	180	200	2	79	127	193	1.663
mean_dist	Distancia recorrida promedio	6,43	1,11	5,00	5,80	6,30	6,70	20,60
mean_travel_time	Tiempo de viaje promedio	14,72	1,55	11,80	13,60	14,70	15,70	25,80
spatial_mean_speed	Velocidad media espacial	26,20	2,73	21,00	24,40	26,10	27,70	47,90
wait_time	tiempo de espera promedio	4,58	0,60	3,34	4,17	4,54	4,88	7,44
<b>Sociodemografía</b>								
mean_income	Ingreso promedio en miles de pesos	678	354	301	476	562	734	3.099
mean_car_own	Número de autos promedio por hogar	0,54	0,41	-	0,30	0,40	0,60	2,80
prop_low_income	% hogares de ingreso bajo	0,32	0,18	-	0,18	0,31	0,44	1,00
prop_med_income	% hogares de ingreso medio	0,59	0,13	-	0,51	0,60	0,69	1,00
prop_high_income	% hogares de ingreso alto	0,09	0,15	-	-	0,02	0,10	0,64
prop_male	% hombres	0,48	0,02	0,44	0,48	0,49	0,49	0,54
prop_female	% mujeres	0,52	0,02	0,46	0,51	0,51	0,52	0,56
prop_age_18_39	% personas entre 18 y 39 años	0,34	0,08	0,10	0,29	0,34	0,38	0,60
prop_age_40_59	% personas entre 40 y 59 años	0,27	0,08	-	0,22	0,27	0,31	0,54
prop_age_60+	% personas mayores de 60 años	0,18	0,09	-	0,12	0,17	0,23	0,83
prop_higher_educ	% personas con educ. superior terminada	0,39	0,20	-	0,24	0,37	0,53	0,92
prop_car_0	% hogares sin auto	0,61	0,19	-	0,52	0,63	0,74	1,00
prop_car_1	% hogares con 1 auto	0,26	0,11	-	0,19	0,26	0,33	0,67
prop_car_2+	% hogares con 2 o más autos	0,13	0,18	-	-	0,07	0,17	0,96
<b>Entorno Construido</b>								
population	Población	17.895	12.531	440	11.482	14.926	20.431	98.596
num_hh	Número de hogares	5.694	3.819	161	3.627	4.747	6.453	28.017
hh_size	Número de habitantes promedio por hogar	3,21	0,89	1,80	2,90	3,30	3,50	15,70
area	Superficie	4,14	16,61	0,38	0,96	1,34	2,08	262,09
%_urb	% de superficie urbana	79,78	15,47	-	80,04	84,12	86,96	97,00
area_eff	Superficie urbana	2,11	3,77	0,15	0,81	1,12	1,69	45,93
pop_density	Número de habitantes por km <sup>2</sup>	13.813	8.178	63	9.337	12.950	16.986	69.713
%H	% de predios con uso habitacional	0,73	0,22	-	0,62	0,82	0,89	0,99
%C	% de predios con uso comercial	0,11	0,11	-	0,04	0,08	0,15	0,74
%O	% de predios con uso de oficinas	0,02	0,06	-	-	0,01	0,02	0,55
%I	% de predios con uso industrial	0,08	0,14	-	-	0,01	0,10	0,89
%E	% de predios con uso educacional	0,05	0,05	-	0,02	0,04	0,06	0,53
%S	% de predios con uso salud	0,01	0,02	-	-	-	0,01	0,17
%G	% de predios con uso hotel	-	0,01	-	-	-	-	0,11
%Q	% de predios con uso culto	0,01	0,01	-	-	0,01	0,01	0,05
%P	% de predios con uso adm. pública y defensa	0,01	0,03	-	-	-	-	0,56
%D	% de predios con uso deporte y recreación	-	0,01	-	-	-	-	0,19
land_use_mix	Mix de uso de suelo	0,34	0,17	-	0,21	0,31	0,47	0,78
num_mall1	Num. malls con GLA <40.000 m <sup>2</sup>	0,06	0,31	-	-	-	-	4,00
num_mall2	Num. malls con 40.000 <GLA <80.000 m <sup>2</sup>	0,02	0,15	-	-	-	-	1,00
num_mall3	Num. malls con GLA >80.000 m <sup>2</sup>	0,03	0,18	-	-	-	-	2,00
dummy_airport	1 si DC posee aeropuerto, 0 si no	-	0,05	-	-	-	-	1,00
dummy_bus_station	1 si DC posee terminal de buses, 0 si no	0,01	0,11	-	-	-	-	1,00
dummy_CBD	1 si DC pertenece a comuna CBD, 0 si no	0,17	0,37	-	-	-	-	1,00
<b>Oferta de otros modos</b>								
num_stops	Número de paraderos	32,82	18,99	4,00	20,00	28,00	40,00	135,00
stops_density	Número de paraderos por km <sup>2</sup>	25,19	12,45	1,28	17,43	25,32	32,02	113,33
f_lab_n2	Frec. promedio buses lun-vie 00:00 - 07:00	1,06	0,27	0,32	0,87	1,02	1,19	2,03
f_lab_pm	Frec. promedio buses lun-vie 07:00 - 10:00	5,58	0,96	2,61	5,02	5,46	6,04	9,21
f_lab_fp	Frec. promedio buses lun-vie 10:00 - 17:30	4,72	0,75	2,52	4,31	4,70	5,09	6,96
f_lab_pt	Frec. promedio buses lun-vie 17:30 - 21:00	5,43	0,89	2,75	4,92	5,33	5,74	9,61
f_lab_n1	Frec. promedio buses lun-vie 21:00 - 00:00	3,57	0,61	2,03	3,25	3,50	3,86	5,70
f_sab_n2	Frec. promedio buses sáb. 00:00 - 07:00	0,76	0,22	0,23	0,61	0,75	0,88	1,45
f_sab_pm	Frec. promedio buses sáb. 07:00 - 10:00	4,10	0,67	2,13	3,79	4,11	4,46	6,25
f_sab_fp	Frec. promedio buses sáb. 10:00 - 17:30	4,02	0,64	2,10	3,72	4,02	4,35	6,16
f_sab_pt	Frec. promedio buses sáb. 17:30 - 21:00	4,00	0,64	2,03	3,69	4,01	4,32	6,32
f_sab_n1	Frec. promedio buses sáb. 21:00 - 00:00	2,97	0,50	1,29	2,68	2,96	3,22	4,85
f_dom_n2	Frec. promedio buses dom. 00:00 - 07:00	0,37	0,20	0,00	0,22	0,35	0,50	0,97
f_dom_pm	Frec. promedio buses dom. 07:00 - 10:00	2,84	0,50	1,29	2,55	2,86	3,14	4,37
f_dom_fp	Frec. promedio buses dom. 10:00 - 17:30	3,51	0,56	1,80	3,23	3,49	3,81	5,22
f_dom_pt	Frec. promedio buses dom. 17:30 - 21:00	3,61	0,58	1,72	3,31	3,61	3,96	5,71
f_dom_n1	Frec. promedio buses dom. 21:00 - 00:00	2,66	0,47	1,08	2,34	2,65	2,92	4,07
dummy_metro	1 si DC tiene estación de metro, 0 si no	0,46	0,50	-	-	-	1,00	1,00
cyclepath_density	km. de ciclovías disponibles por km <sup>2</sup>	3,54	4,09	0,00	0,00	2,22	5,36	21,08

## Análisis Temporal y Patrón de Viajes

Debido a que el período de análisis es de dos semanas, no se tienen cambios significativos en las variables de sociodemografía, entorno construido y oferta de transporte público, por lo cual en esta subsección se analizará solamente el comportamiento temporal de los viajes. La Figura 3.9 muestra el perfil de viajes para las dos semanas de estudio. En ella, se observa un comportamiento similar en ambas semanas, lo que indica una regularidad en el patrón de viajes. Durante los días laborales se observa un patrón claro: existe un período “punta mañana” entre las 07:00 y las 10:00 horas, además de un período “punta tarde” que se amplía a medida que transcurren los días de la semana. Es importante señalar que el comportamiento en el día viernes es ligeramente diferente al resto de los días laborales, ya que el peak de la tarde comienza más temprano y es mucho más pronunciado. Luego, durante el sábado, desaparece el peak de la mañana y además el peak de la tarde aumenta a tal punto que se observa la demanda máxima semanal. Finalmente, durante el domingo la demanda de viajes disminuye considerablemente en comparación a los demás días y además esta demanda presenta una forma más plana.

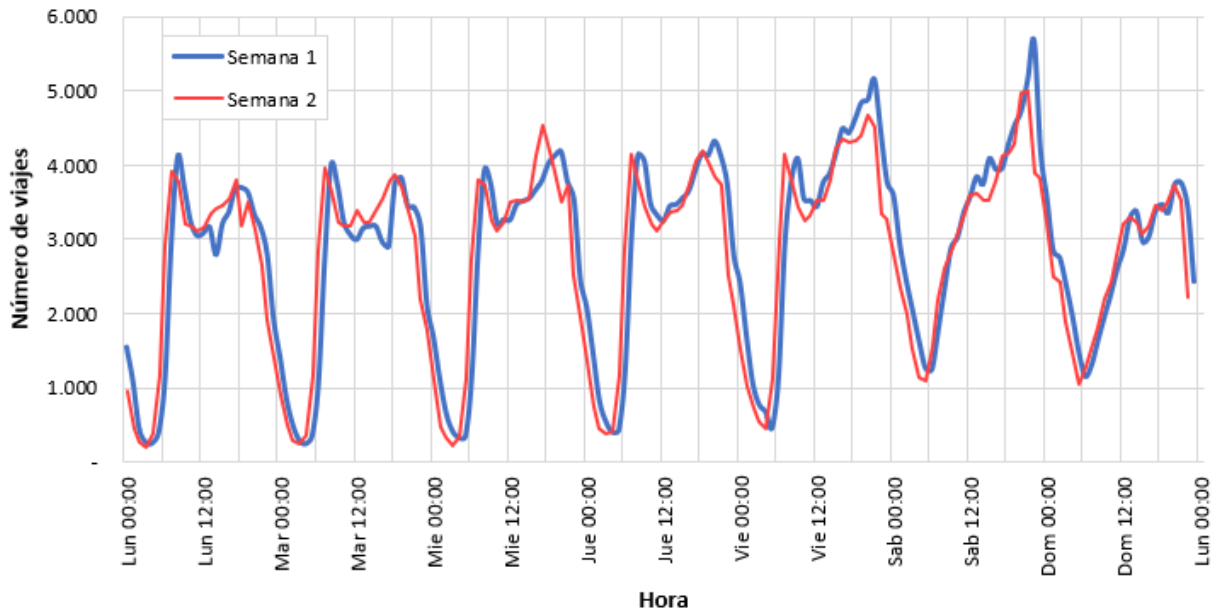


Figura 3.9: Perfil de Viajes

Con el objetivo de realizar un análisis más detallado y capturar patrones de viajes específicos, se definen cuatro tipos de día: lunes a jueves, viernes, sábado y domingo. Además, se descompone cada día en cinco períodos, de acuerdo a lo descrito en la Figura 3.10: nocturno 2 (N2: 00:00 - 06:59 hrs.), punta mañana (PM: 07:00 - 09:59 hrs.), fuera de punta (FP: 09:00 - 17:29 hrs.), punta tarde (PT: 17:30 - 20:59 hrs.) y nocturno 1 (N1: 21:00 - 23:59 hrs.).

Para la punta mañana, durante los días laborales se observa una proporción del orden del 15 % del total de viajes diario, mientras que durante el fin de semana se reduce a menos del 10 %. En el período fuera de punta se observa una proporción cercana al 40 % durante los días laborales y disminuye a un 34 % durante el fin de semana. Luego, en punta tarde se tiene una proporción de aproximadamente 20 % durante toda la semana. Por otra parte,

para el período nocturno 1 se observa un ligero aumento durante viernes y sábado: pasa de aproximadamente un 15 % entre domingo y jueves a poco menos del 20 % durante viernes y sábado. Por último, se aprecia un aumento significativo en la proporción de viajes para el período nocturno 2, pasando de un 6 y 7 % durante los días laborales a un 14 y 20 % durante el fin de semana. Lo expuesto anteriormente se ilustra en la Figura 3.11.



Figura 3.10: Periodización de cada día de análisis

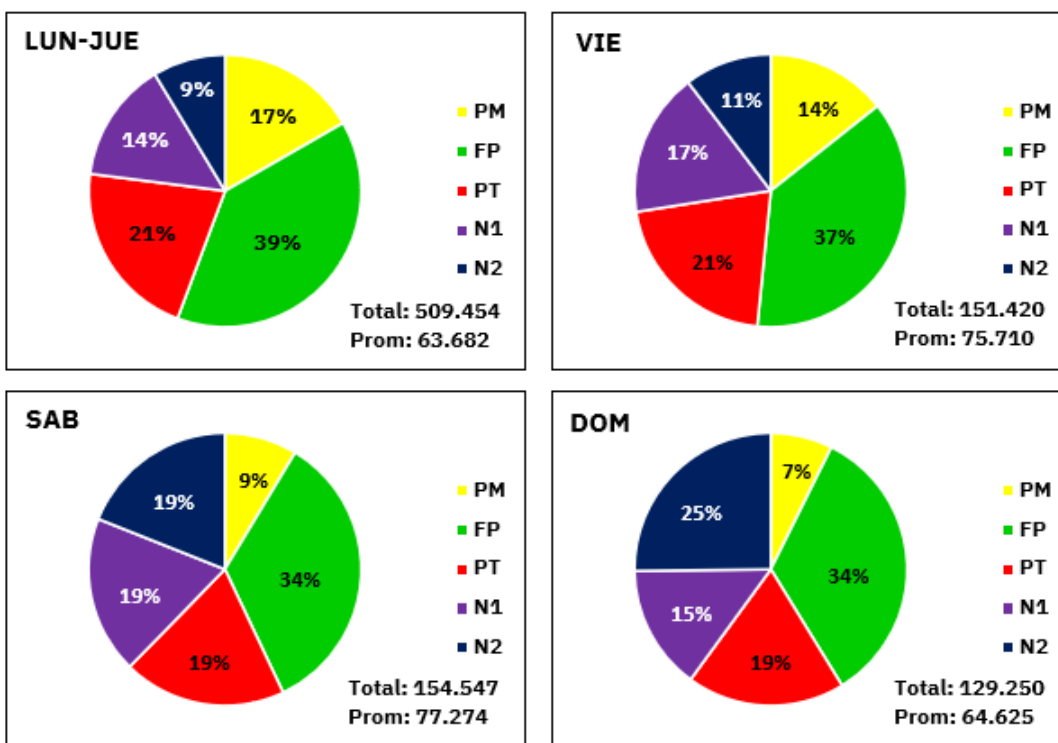


Figura 3.11: Composición Temporal de Viajes

Es importante notar que los porcentajes reportados en la Figura 3.11 están influenciados por el tamaño de cada período, por lo que no son suficientes para dar una idea precisa de la demanda específica durante cada uno. Para eliminar el efecto tamaño, en la Figura 3.12 se reporta el número de viajes por hora promedio de cada período y cada tipo de día.

En la figura, se observa que durante los días laborales, la demanda durante la madrugada es del orden de los 1.000 viajes por hora, pero durante el fin de semana aumenta al doble o incluso más. Por otro lado, para el período PM ocurre lo contrario: durante los días laborales se tiene una demanda del orden de los 4.000 viajes por hora y durante el fin de semana disminuye a cerca de la mitad. En cuanto al período FP, la demanda es relativamente estable

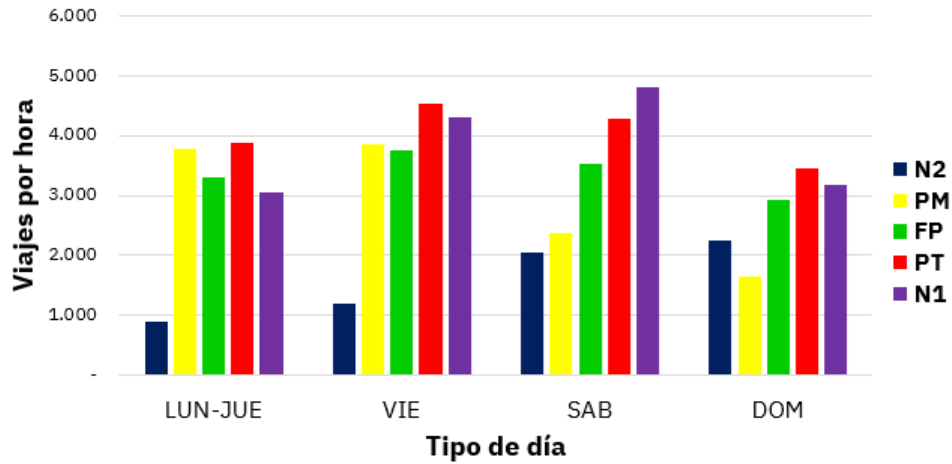


Figura 3.12: Viajes promedio por hora según día-período

durante toda la semana, al igual que el período PT donde se tiene un aumento pequeño (entre 10 y 15 %) durante el viernes y sábado. Por último, en el período N1, se observa un aumento más notorio durante viernes y sábado.

## Análisis Espacial

### *Viajes en ridesourcing*

En la Figura 3.13 se muestra la distribución espacial de los viajes. Este primer mapa corresponde a la distribución de los viajes totales por DC, donde se observa que los viajes se concentran principalmente en los sectores centro y oriente de Santiago, en zonas pertenecientes a comunas con gran población y en algunos sectores particulares.

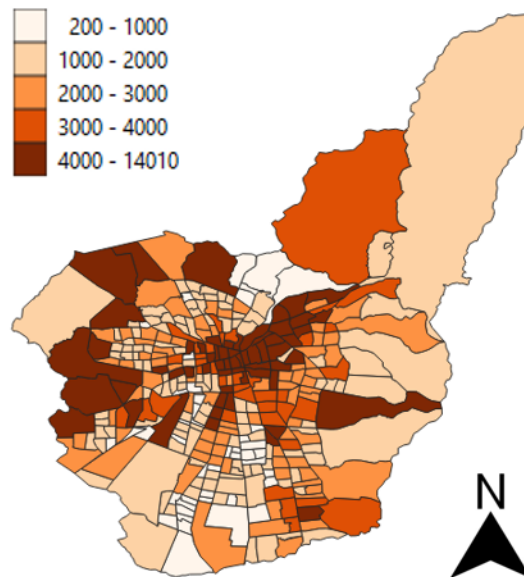


Figura 3.13: Distribución espacial de viajes totales

Sin embargo, esta distribución está fuertemente influenciada por el tamaño de los DC. Para eliminar este efecto, se presentan los mapas de dos variables de viaje normalizadas: en la Figura 3.14 se muestra la distribución espacial de la densidad de viajes, definida como la cantidad de viajes por unidad de superficie urbana, mientras que en la Figura 3.15 se muestra el total de viajes por cada 1.000 hogares. Si bien en ambos mapas se acentúa el comportamiento descrito anteriormente, se observan algunas diferencias. En el caso de la densidad de viajes, la concentración de viajes es más notoria en el centro, ya que las zonas de dicho sector son más pequeñas. Por otro lado, al normalizar por hogares, cobran relevancia algunas zonas fuera del centro, debido a que son zonas con pocos hogares (por ejemplo, sectores de uso industrial).

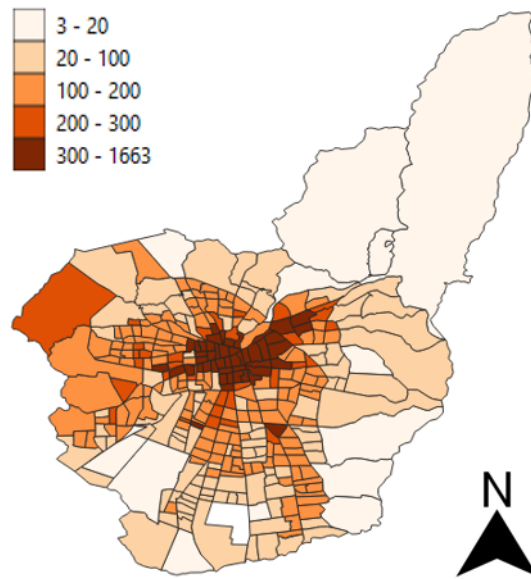


Figura 3.14: Distribución espacial de densidad de viajes

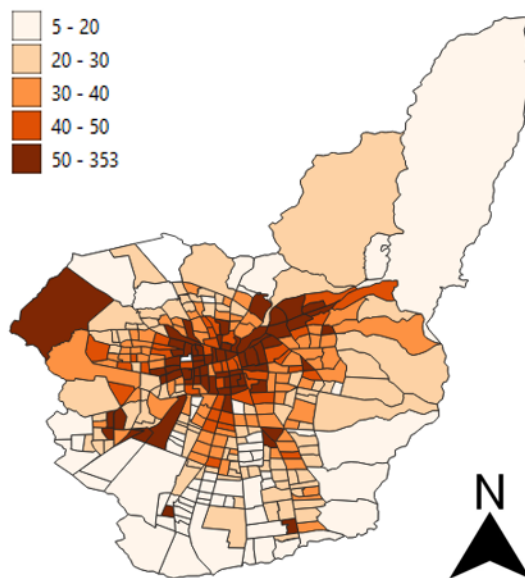


Figura 3.15: Distribución espacial de viajes cada 1000 hogares



En cuanto a los atributos de viaje, en las Figuras 3.16 y 3.17 se muestran las distribuciones espaciales del tiempo de espera promedio y de la velocidad media espacial. A partir de ambos mapas, se logra apreciar que en el interior del anillo Vespucio se tienen menores tiempos de espera, pero también se tiene una menor velocidad; mientras que en las zonas de la periferia ocurre lo contrario: si bien los tiempos de espera son mayores, la velocidad de circulación promedio también es mayor. Ambas distribuciones pueden explicarse en parte debido a la influencia de la congestión.

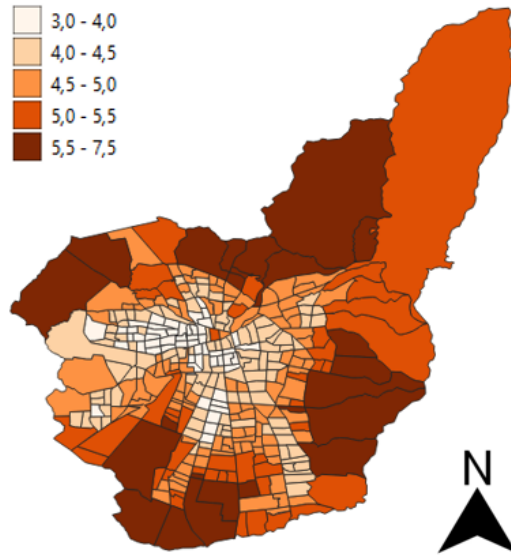


Figura 3.16: Distribución espacial de tiempo de espera promedio

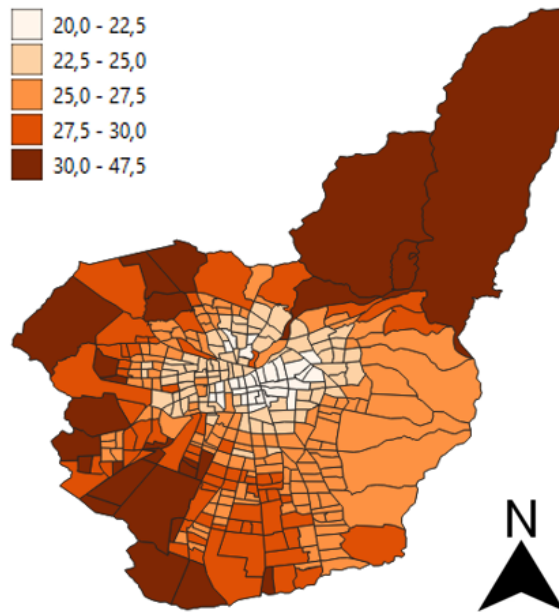


Figura 3.17: Distribución espacial de velocidad media espacial

## Sociodemografía

En primer lugar, se presentan los mapas de las variables calculadas a nivel de persona. En la Figura 3.18 se muestra la distribución espacial del porcentaje de población masculina de cada DC. A pesar de que se observa una menor población masculina en parte del sector oriente, se debe tener en cuenta que el promedio zonal está en torno al 48% y su dispersión es bastante baja. Por lo tanto, se tiene una composición por género es bastante equilibrada.

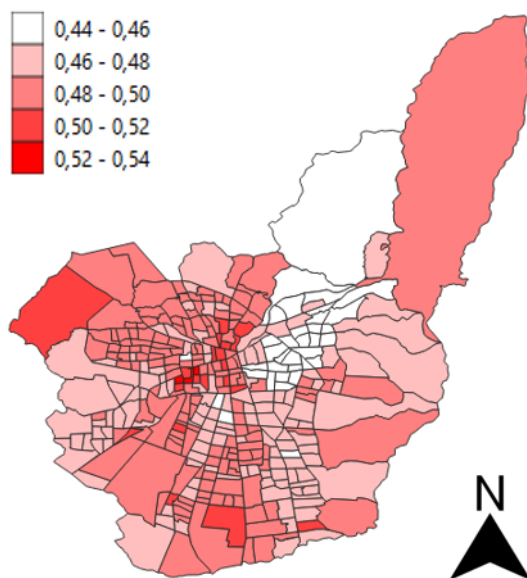


Figura 3.18: Distribución espacial de % de población masculina

Por otro lado, el mapa de la Figura 3.19 muestra la distribución espacial del porcentaje de población entre 18 y 39 años en cada DC. En este caso tampoco se observa segregación, pero la dispersión es bastante mayor que la de distribución por género.

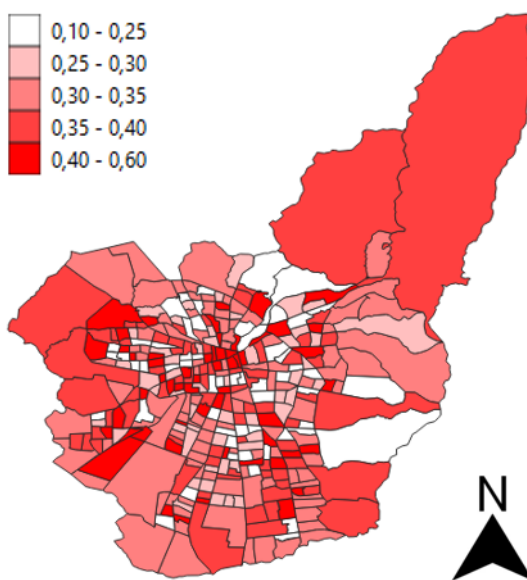


Figura 3.19: Distribución espacial de % de población entre 18 y 39 años

En la Figura 3.20 se muestra la distribución espacial del porcentaje de población con educación superior terminada. A diferencia de los dos mapas anteriores, se observa una clara segregación de los valores de esta variable: en el sector oriente existe una alta concentración de personas con educación superior terminada, llegando incluso al 65 %; la zona centro también posee una concentración alta (sobre el 50 %), aunque menor que la de la zona oriente; mientras que en el resto de la ciudad este porcentaje no supera el 20 %.

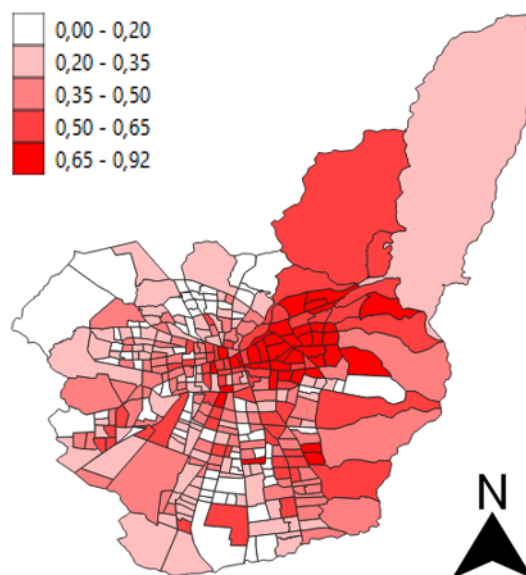


Figura 3.20: Distribución espacial de % de población con educación superior terminada

La Figura 3.21 muestra la distribución espacial del porcentaje de hogares de ingreso alto. En este caso también se observa una distribución altamente segregada: los hogares de mayores ingresos se concentran principalmente en la zona oriente y en menor medida en el centro de Santiago. En el resto de la ciudad, los hogares de ingreso alto no superan el 5 %.

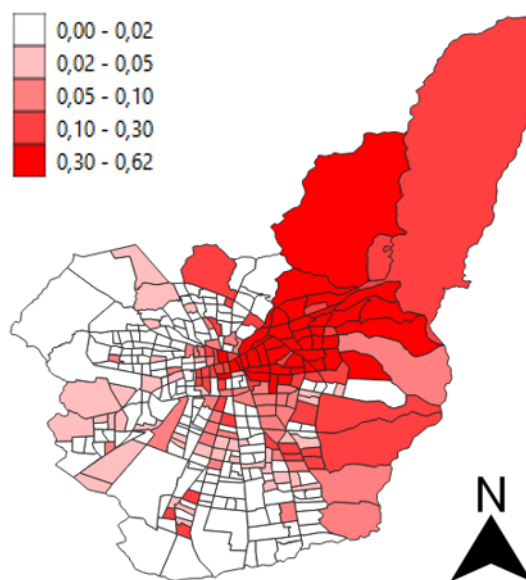


Figura 3.21: Distribución espacial de % de hogares de ingreso alto

Finalmente, en la Figura 3.22 se muestra la distribución espacial del porcentaje de hogares sin automóvil. Como era de esperar, la mayor proporción de hogares sin automóvil se observa justamente en las zonas en donde hay una menor concentración de hogares de ingreso alto.

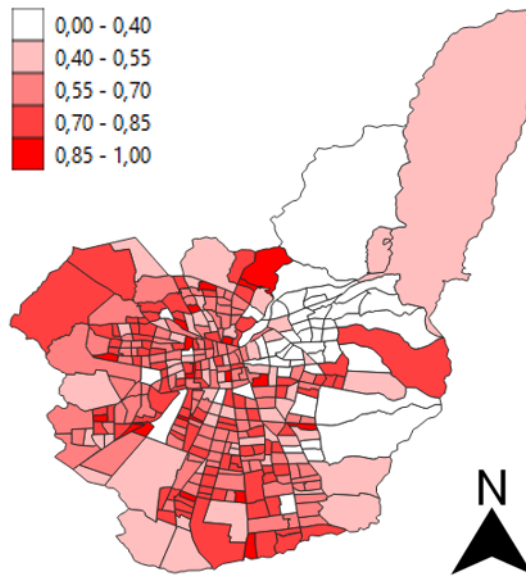


Figura 3.22: Distribución espacial de % de hogares sin automóvil

### *Entorno Construido*

La Figura 3.23 muestra la distribución espacial del porcentaje de superficie destinada a uso urbano en cada DC. En concordancia con la intuición, las zonas ubicadas al interior del anillo Vespucio presentan un mayor uso urbano (sobre el 80 %) en comparación con las zonas del exterior, a excepción de algunos sectores que podrían corresponder a subcentros.

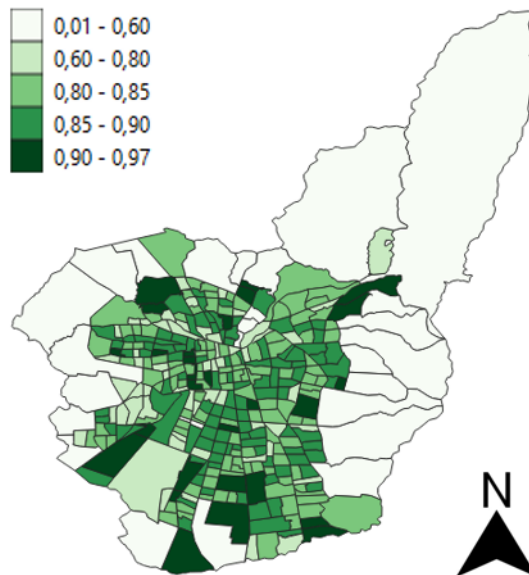


Figura 3.23: Distribución espacial de % de superficie con uso urbano

Luego, en la Figura 3.24 se presenta la distribución espacial del número de hogares por DC. En este caso, se observa que la mayor cantidad de hogares se concentra en el centro y parte del sector oriente de Santiago. A pesar de que las zonas son más pequeñas, estos sectores se caracterizan por tener una gran cantidad de edificios residenciales. También se observa un gran número de hogares en la periferia, aunque la situación es diferente: hay mayor presencia de casas y el tamaño de las zonas es mayor.

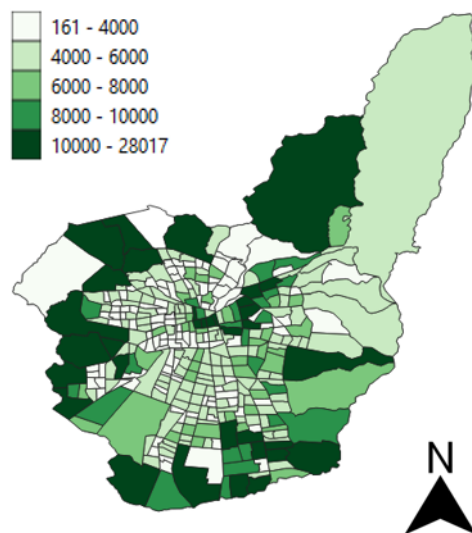


Figura 3.24: Distribución espacial de número de hogares

En la Figura 3.25 se muestra la distribución espacial de la densidad poblacional. En la figura se observa que las zonas con mayor densidad poblacional se ubican en el centro de Santiago, y comienza a disminuir a medida que las zonas están más alejadas del centro, principalmente debido a que las zonas alejadas son más grandes.

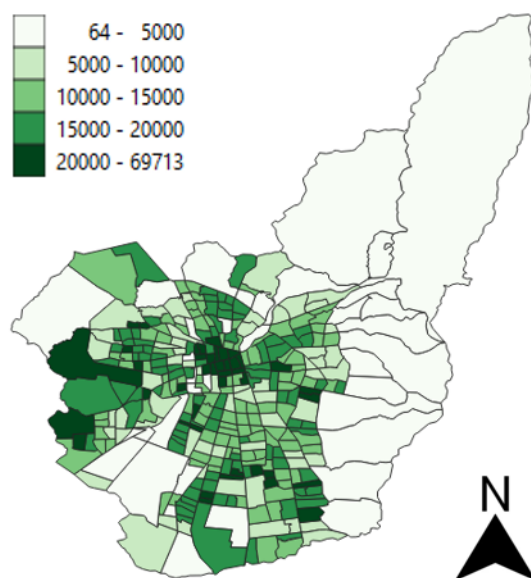


Figura 3.25: Distribución espacial de densidad poblacional

Por otro lado, en el mapa de la Figura 3.26 muestra la distribución espacial del mix de uso de suelo. Se observa que las zonas con un mix de uso de suelo más alto son principalmente aquellas que se ubican en el centro, salvo algunas excepciones. Como se muestra más adelante, esto se debe a que en el centro hay mayor diversidad de actividades, mientras que en la periferia el suelo está destinado principalmente a uso residencial e industrial.

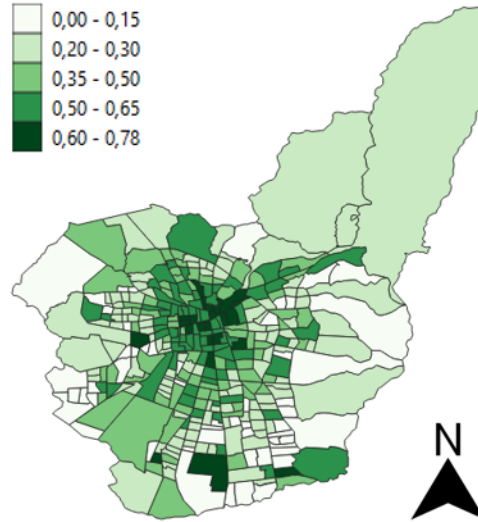


Figura 3.26: Distribución espacial de mix de uso de suelo

A continuación, se presentan los mapas del porcentaje de suelo destinado a las actividades más relevantes. La Figura 3.27 describe la distribución espacial del porcentaje de suelo destinado al uso habitacional, donde se observa que las zonas ubicadas en las afueras de Santiago son las que poseen un mayor uso habitacional, a excepción del sector norponiente.

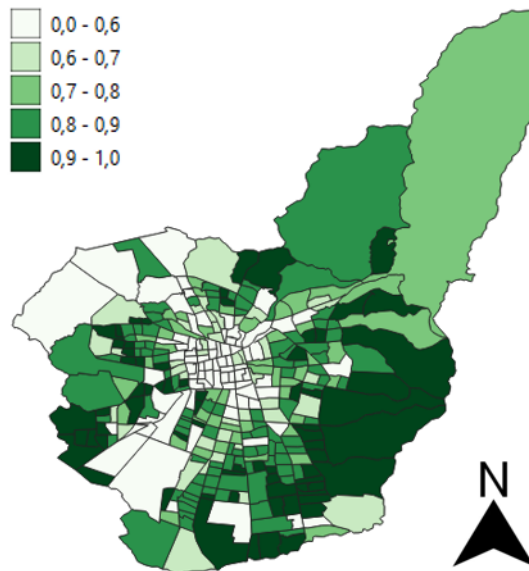


Figura 3.27: Distribución espacial de porcentaje de suelo con uso habitacional

En la Figura 3.28 se muestra la distribución espacial del porcentaje de uso de suelo con uso comercial. Contrariamente al mapa anterior, se tiene que las zonas con mayor composición comercial son las ubicadas en el centro de Santiago, además de algunas localizaciones específicas donde se ubican centros comerciales.

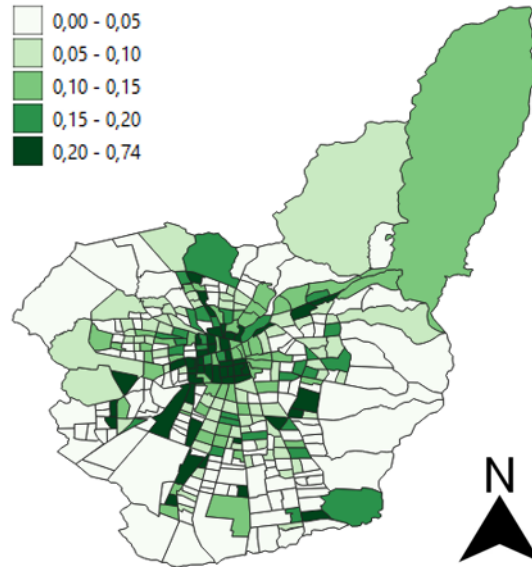


Figura 3.28: Distribución espacial de porcentaje de suelo con uso comercial

Por su parte, en la Figura 3.29 se describe la distribución espacial del porcentaje de suelo utilizado para industrias. En la figura se observa que las industrias se encuentran muy sectorizadas: en general no supera el 5% de la superficie, pero en algunos sectores de la periferia sobrepasa el 30% de participación.

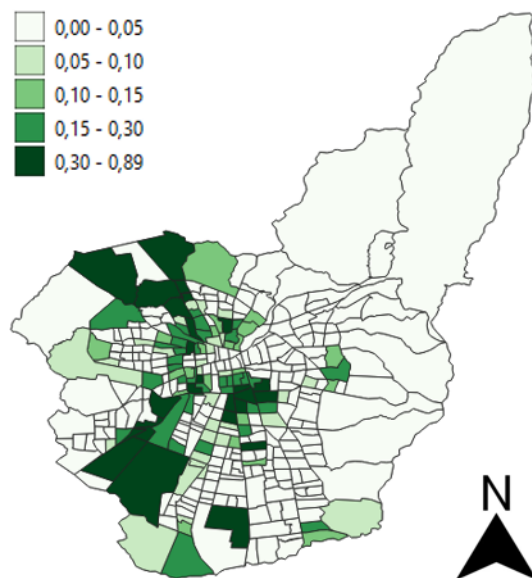


Figura 3.29: Distribución espacial de porcentaje de suelo con uso industrial



Finalmente, la Figura 3.30 muestra la distribución espacial del porcentaje de suelo destinado a oficinas. En gran parte del área de estudio el porcentaje de oficinas no supera el 5%, pero existe una gran concentración en el centro y parte del sector oriente de Santiago, alcanzando incluso el 50% en algunos casos.

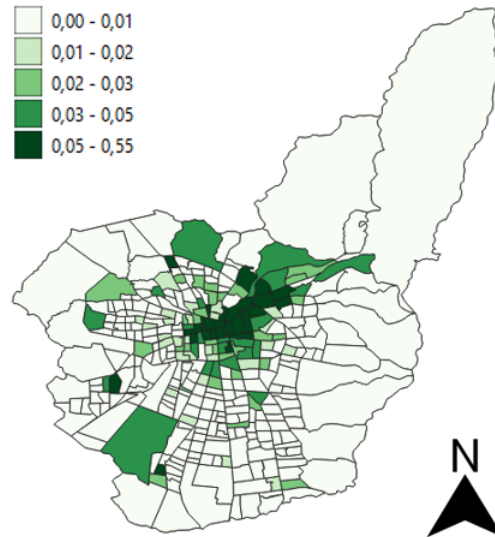


Figura 3.30: Distribución espacial de porcentaje de suelo con uso para oficinas

### *Oferta de Transporte Público*

La Figura 3.31 muestra el número de paraderos por DC, donde se observa una mayor cantidad en la periferia de Santiago. Sin embargo, al igual que en el caso de los viajes, esta variable puede estar influenciada por el tamaño de la zona, por lo cual en la Figura 3.32 se reporta la distribución espacial de la densidad de paraderos. En este segundo mapa se puede ver que la situación cambia diametralmente: debido a su gran tamaño, las zonas periféricas presentan una densidad de paraderos muy baja en comparación a las zonas del centro.

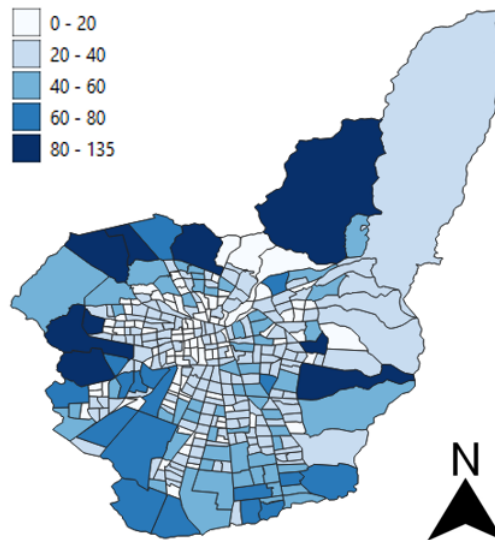


Figura 3.31: Distribución espacial de número de paraderos



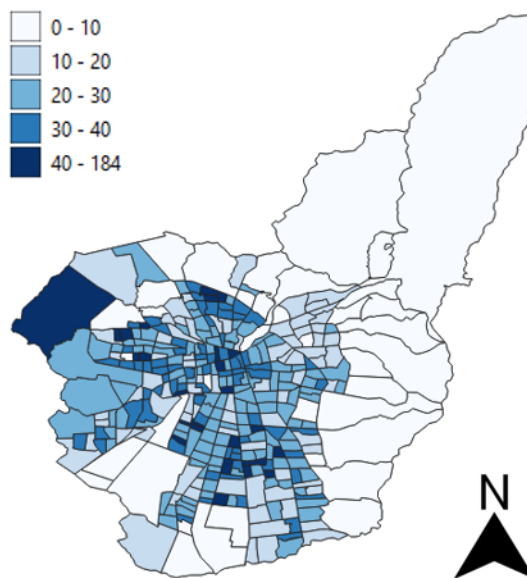


Figura 3.32: Distribución espacial de densidad de paraderos

Finalmente, en la Figura [3.33](#) se reporta la distribución espacial de la densidad de ciclo-  
vías. Se puede observar que la oferta de ciclo-  
vías está muy concentrada en algunos sectores  
específicos de la capital, específicamente en los sectores oriente, poniente y sur, donde la  
densidad es al menos dos veces mayor que en el resto de la ciudad. Es importante señalar  
que debido a la estructura de los datos, los kilómetros de ciclo-  
vías utilizados para calcular la  
densidad no corresponden a los kilómetros contenidos dentro de cada distrito censal, sino que  
corresponde a los kilómetros totales de las ciclo-  
vías que pasan por la zona, como aproximación  
de la accesibilidad que otorga el modo bicicleta.

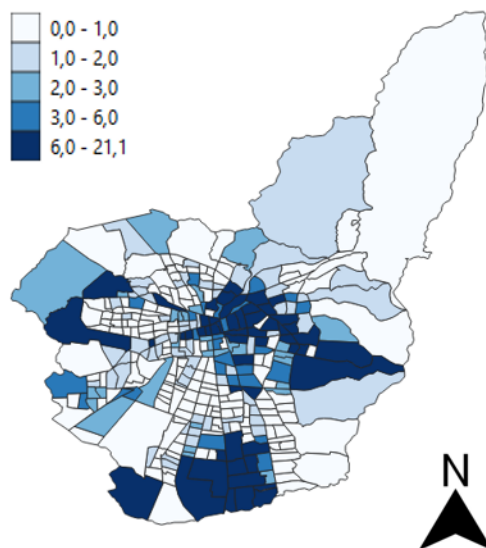


Figura 3.33: Distribución espacial de densidad de ciclo-  
vías

## 3.2. Modelación

### 3.2.1. Regresión Lineal OLS

En primera instancia, se considera la aplicación de un modelo de regresión lineal múltiple. Este modelo se caracteriza por establecer una relación lineal entre la variable que se desea explicar (variable dependiente) y un conjunto de variables explicativas. La forma funcional de este modelo se muestra en la ecuación (3.2):

$$y_i = \beta_0 + \sum_{k=1}^K \beta_k \cdot x_{ik} + \varepsilon_i \quad i = 1, \dots, n \quad (3.2)$$

donde  $y_i$  es la observación  $i$  de la variable dependiente;  $x_{ik}$  es la observación  $i$  de la  $k$ -ésima variable explicativa;  $\beta_0$  es el intercepto, que indica el valor promedio de la variable explicada cuando las variables explicativas valen cero;  $\beta_k$  es el coeficiente de regresión, que mide la influencia individual de  $x_{ik}$ ; mientras que  $\varepsilon_i$  es el error o residuo, el cual captura tanto el efecto de variables no incluidas en el modelo y también la propia variabilidad de los datos.

El modelo también puede ser representado de forma matricial:

$$Y = X\beta + \varepsilon \quad (3.3)$$

donde  $Y = \{y_i\}$ ,  $X = \{x_{ik}\}$ ,  $\beta = \{\beta_k\}$  y  $\varepsilon = \{\varepsilon_i\}$ . Cabe señalar que para que esta forma matricial sea consistente con la ecuación (3.2), la matriz  $X$  incluye en su primera fila un vector de 1's, mientras que el vector  $\beta$  incluye como primer término el intercepto  $\beta_0$ .

La estimación de los coeficientes del modelo se realiza mediante el criterio de mínimos cuadrados ordinarios (OLS: *ordinary least squares*). Este método de estimación permite que la recta de regresión sea la que mejor se ajusta a los datos disponibles, a través de un problema de optimización que minimiza la suma de los cuadrados de los residuos.

En este trabajo se consideran dos variables dependientes: (1) número de viajes por km<sup>2</sup> y (2) número de viajes por cada 1.000 hogares, normalizadas en la cantidad de días (lun–jue: 4, vie: 1, sab: 1, dom: 1). Ambas variables dependientes permiten eliminar el efecto del tamaño de las zonas, aunque presentan algunas diferencias fundamentales: la variable dependiente (1) representa una densidad de viajes generados independientemente del tipo de actividades que se realicen en cada DC, mientras que la variable dependiente (2) se asimila de mejor manera a modelos basados en el hogar, los cuales son habituales en la literatura de transporte.

Además, debido a la naturaleza no negativa de la cantidad de viajes al ser una variable de conteo – independientemente de que se encuentre normalizada – no es adecuado utilizar directamente una regresión lineal, ya que se podrían generar valores ajustados negativos. Para garantizar que la estimación sea consistente con esta naturaleza no negativa de las variables explicadas, se establece una relación exponencial con las variables explicativas, dando origen a un modelo log-lineal, dado por la expresión de la ecuación (3.4):

$$Y = e^{X\beta + \varepsilon} \implies Y' = \ln(Y) = X\beta + \varepsilon \quad (3.4)$$

Esta forma funcional implica que un aumento de una unidad de una variable explicativa  $X_i$  produce un aumento de  $(100 \cdot \beta_i) \%$  de  $Y$ .

Un inconveniente del uso de logaritmos es la aparición de ceros en la variable explicada. Para tratar este problema en regresión lineal existen diversos métodos, siendo los más utilizados el uso de  $\log(Y + \Delta)$  (por ejemplo, [Dey et al., 2021](#)) y la eliminación de observaciones cero. Según [Bellego et al. \(2022\)](#), el sumar un término arbitrario dentro del logaritmo resulta problemático debido a que se produce endogeneidad y además reduce artificialmente los errores estándar, mientras que al descartar observaciones con valor cero induce un sesgo de selección. No obstante, dado que en este caso sólo se tienen 43 combinaciones DC-período en donde no hay viajes (un 0,63 % de las 6.780 combinaciones DC-período) y 42 de ellas se dan en períodos nocturnos, se opta por eliminar dichas observaciones en los casos correspondientes.

Se estima un modelo para cada combinación día–período, teniendo así un total de 20 combinaciones, con el objetivo de identificar los diferentes grados de participación de cada variable explicativa en cada una de estas combinaciones. Dado que se tienen dos variables explicadas, en total se estiman 40 modelos. Las variables explicativas consideradas en los modelos se detallan en la Tabla [3.5](#).

Tabla 3.5: Variables explicativas del modelo

Variable	Descripción
<b><i>Atributos de Ridesourcing</i></b>	
wait_time	Tiempo de espera promedio
spatial_mean_speed	Velocidad media espacial
<b><i>Sociodemografía</i></b>	
prop_male	% de población masculina
prop_18_39	% de población entre 18 y 39 años
prop_med_income	% de hogares de ingreso medio
prop_high_income	% de hogares de ingreso alto
prop_higher_educ	% de población con educación superior terminada
prop_car_1	% de hogares con 1 automóvil
prop_car_2+	% de hogares con 2 o más automóviles
<b><i>Entorno Construido</i></b>	
log_pop_density	Densidad Poblacional (log)
land_use_mix	Mix de uso de suelo
num_mall_1	Núm. de centros comerciales con GLA <40.000 m <sup>2</sup>
num_mall_2	Núm. de centros comerciales con GLA entre 40.000 y 80.000 m <sup>2</sup>
num_mall_3	Núm. de centros comerciales con GLA >80.000 m <sup>2</sup>
dummy_airport	1 si el DC posee aeropuerto, 0 si no
dummy_bus_station	1 si el DC posee terminal de buses, 0 si no
dummy_CBD	1 si el DC pertenece a una comuna definida como CBD, 0 si no
<b><i>Oferta de otros modos</i></b>	
stops_density	Número de paraderos por km <sup>2</sup>
frec_d_p	Frecuencia de buses en el tipo de día d y período p
dummy_metro	1 si el DC posee estación de metro, 0 si no
cycleways_density	Densidad de ciclovías [km/km <sup>2</sup> ]

### 3.2.2. Teorema de Gauss-Markov

De acuerdo a [Stock y Watson \(2011\)](#), el Teorema de Gauss-Markov establece que bajo una serie de condiciones, el estimador OLS – denotado por  $\hat{\beta}$  – es el mejor estimador lineal insesgado (MEI o BLUE: *best linear unbiased estimator*), en el sentido de que son los estimadores de mínima varianza entre los estimadores lineales e insesgados. Estas condiciones se conocen como condiciones de Gauss-Markov, las cuales se detallan a continuación:

1. **Exogeneidad:** La esperanza condicional de los errores – dado  $X$  – es cero, lo que permite garantizar que el estimador sea insesgado, es decir, que el valor esperado de dicho estimador sea el valor del parámetro real. Matemáticamente:

$$E(\varepsilon|X) = 0 \implies E(\hat{\beta}|X) = 0 \quad (3.5)$$

2. **Homocedasticidad:** La varianza de los errores es constante en las observaciones. Junto con la condición de no autocorrelación (estas condiciones en conjunto se conocen como esfericidad), permite asegurar que el estimador es eficiente, es decir, que su varianza es menor que la de otros estimadores lineales insesgados.

$$E(\varepsilon^2|X) = Var(\varepsilon|X) = \sigma^2 \quad (3.6)$$

3. **No autocorrelación:** Los errores o residuos no están correlacionados entre diferentes observaciones. Como se señaló en la condición 2, junto con la condición de homocedasticidad permite asegurar que el estimador es eficiente.

$$E(\varepsilon_i \varepsilon_j | X) = Cov(\varepsilon_i, \varepsilon_j | X) = 0 \quad (3.7)$$

Si además de las condiciones mencionadas anteriormente, se satisface el supuesto de distribución normal de los errores, es decir:

$$\varepsilon_i | X \sim N(0, \sigma^2) \quad (3.8)$$

entonces el estimador OLS es el estimador máximo verosímil, y por lo tanto, el mejor estimador insesgado (MEI o BUE: *best unbiased estimator*), es decir, es el estimador insesgado con menor varianza, incluyendo estimadores tanto lineales como no lineales.

### 3.2.3. Correcciones al modelo

A pesar de que gran parte de las condiciones descritas en la subsección anterior son consideradas como supuestos de los estimadores OLS ([Stock y Watson 2011](#)), es esencial corroborar que estas condiciones efectivamente se están cumpliendo. En caso de que aquello no esté ocurriendo, es necesario realizar las correcciones correspondientes para que el modelo funcione correctamente.

## 1. Endogeneidad

La endogeneidad se produce cuando no se cumple la condición de exogeneidad, es decir, cuando la esperanza condicional de los errores es distinta de cero, causando que el estimador OLS sea sesgado e inconsistente. Puede ocurrir por varias razones, tales como omisión de variables explicativas, errores de especificación o determinación simultánea, entre otros. Así, se dice que una variable explicativa  $x_k$  es endógena cuando está correlacionada con el error  $\varepsilon$  (Wooldridge 2002). Para ilustrar esto, en la ecuación (3.9) se tiene nuevamente la forma funcional del modelo<sup>3</sup>:

$$y = \beta_0 + \sum_{k=1}^K \beta_k \cdot x_k + \varepsilon \quad (3.9)$$

donde, sin pérdida de generalidad, se asumirá que la variable  $x_K$  es potencialmente endógena.

En la modelación de viajes en ridesourcing, se sospecha la existencia de endogeneidad en los atributos de viaje debido a un problema de determinación simultánea. En el caso del tiempo de espera, se debe a que esta variable depende del número de conductores que estén circulando en el sector, que a su vez depende de la demanda de viajes en ridesourcing. También es razonable pensar que existe endogeneidad en la velocidad: por una parte, una menor velocidad desincentiva el uso de ridesourcing, pero por otro lado esta reducción de velocidad podría estar asociada a una mayor congestión, donde la demanda suele ser mayor.

Una manera de corregir el efecto de la endogeneidad de una o más variables explicativas en el modelo es mediante el uso de variables instrumentales (VI). De acuerdo a (Wooldridge 2002), las variables instrumentales – denotadas por  $z$  – son variables observables que no se encuentran en la regresión principal (ecuación (3.2)) y deben cumplir dos condiciones:

- Validez: la VI no debe estar correlacionada con el error, es decir,  $Cov(z_1, \varepsilon) = 0$ . En otras palabras, la variable instrumental es exógena en la ecuación (3.9).
- Relevancia: la VI está parcialmente correlacionada con la variable endógena  $x_K$ . En estricto rigor, se requiere que en la regresión de la variable endógena sobre todas las variables exógenas:

$$x_K = \gamma_0 + \sum_{k=1}^{K-1} \gamma_k \cdot x_k + \theta_1 \cdot z_1 + \delta_K \quad (3.10)$$

el coeficiente  $\theta_1$  es distinto de cero.

Las variables instrumentales se incorporan al modelo a través de una regresión lineal por mínimos cuadrados en 2 etapas (2SLS: *two stage least squares*). La primera etapa consiste en una regresión auxiliar para cada variable endógena en función de las variables exógenas del modelo principal y de las variables instrumentales. Para un conjunto de  $M$  variables instrumentales, se tiene la siguiente regresión auxiliar:

$$x_K = \gamma_0 + \sum_{k=1}^{K-1} \gamma_k \cdot x_k + \sum_{m=1}^M \theta_m \cdot z_m + \delta_K \quad (3.11)$$

---

<sup>3</sup>Esta expresión es equivalente a la de la ecuación (3.2), se omite el subíndice  $i$  para simplificar la notación.

El resultado de esta primera etapa es un conjunto de estimadores  $\hat{\gamma} = \{\hat{\gamma}_k\}$  y  $\hat{\theta} = \{\hat{\theta}_m\}$ . Usando estos estimadores, se calcula el valor ajustado  $\hat{x}_K$  como:

$$\hat{x}_K = \hat{\gamma}_0 + \sum_{k=1}^{K-1} \hat{\gamma}_k \cdot x_k + \sum_{m=1}^M \hat{\theta}_m \cdot z_m \quad (3.12)$$

Luego, la segunda etapa consiste en realizar la regresión principal de la ecuación (3.9), reemplazando las variables endógenas por el valor ajustado de la ecuación (3.12):

$$y = \beta_0 + \sum_{k=1}^{K-1} \beta_k \cdot x_k + \beta_K \cdot \hat{x}_K + \varepsilon \quad (3.13)$$

Para corregir la endogeneidad en el tiempo de espera promedio y la velocidad media espacial, se proponen las siguientes variables instrumentales:

- Variable en el período anterior (*variable\_name\_1*)
- Promedio de la variable a nivel de comuna a la que pertenece el DC, excepto el mismo DC en cuestión (*variable\_name\_2*).

El objetivo de considerar estas variables instrumentales es poder incluir en los instrumentos una componente temporal y una espacial. El uso de la componente temporal se sustenta en el hecho de que cada zona posee la misma infraestructura independientemente del período, por lo que se espera que exista correlación entre el tiempo de espera o la velocidad de una misma zona en distintos períodos. De manera similar, la inclusión de la componente espacial está relacionada con que las zonas vecinas en general presentan una infraestructura similar, lo que explicaría una eventual correlación en el tiempo de espera y la velocidad en zonas aledañas, pero no necesariamente responden a los mismos shocks de demanda.

La estimación de 2SLS en dos etapas como se describió requiere hacer una corrección de los errores estándar, porque la segunda etapa ignora que se usaron regresores estimados. La estimación de los modelos se realiza con el software RStudio, mediante el paquete *AER* (*Applied Econometrics in R*), específicamente el método *ivreg*, que considera la estimación 2SLS con la corrección de los errores estándar correspondiente. Este método recibe como argumentos la especificación de la regresión principal – incluyendo la variable potencialmente endógena – y los candidatos a instrumentos, entregando como resultado los estimadores 2SLS.

Adicionalmente, el paquete *AER* de RStudio cuenta con una serie de tests de hipótesis, los cuales se detallan a continuación:

- **Test de Instrumentos Débiles:** Se dice que un instrumento es débil cuando su correlación con la variable endógena es cercana a cero, causando que la regresión por 2SLS sea poco confiable (Andrews *et al.* 2019). El test de instrumentos débiles consiste en un test-F sobre la primera regresión bajo la hipótesis nula de que los instrumentos son débiles. Por lo tanto, el resultado deseable es que la hipótesis nula sea rechazada.

- **Test de Wu-Hausman:** Permite detectar la existencia de endogeneidad a través de la comparación de los estimadores 2SLS y OLS. La hipótesis nula sugiere la existencia de endogeneidad en la variable en cuestión.
- **Test de Sargan:** Verifica la validez de los instrumentos a través de una prueba de sobreidentificación, que requiere tener más instrumentos que variables endógenas, como ocurre en esta aplicación. La hipótesis nula corresponde a que los instrumentos son exógenos (no están correlacionados con el error), por lo cual el resultado deseable es que dicha hipótesis no sea rechazada.

Debido a la estructura de la función *ivreg*, sólo se puede incorporar una variable endógena a la vez. Para ilustrar el efecto de la endogeneidad en cada una de las variables, se estimaron modelos para los períodos PM y PT usando como variable dependiente la generación de viajes por km<sup>2</sup>, considerando tres métodos de estimación: OLS, 2SLS usando el tiempo de espera como variable endógena y 2SLS usando la velocidad como variable endógena. El detalle de los resultados se encuentra en el Anexo C.

## 2. Heterocedasticidad:

La heterocedasticidad se produce cuando la varianza de los errores no es constante en la muestra, es decir, cuando no se cumple la condición de homocedasticidad. Se puede producir por distintas causas, tales como agregación, existencia de multicolinealidad o falta de normalidad de las variables, entre otros. Si bien los estimadores que sufren de heterocedasticidad siguen siendo insesgados, dejan de ser estimadores eficientes debido a que no presentan varianza mínima y, además, los errores estándar no pueden obtenerse de manera directa, sino que requieren una corrección.

Existen diversos métodos que permiten detectar la existencia de heterocedasticidad. Para efectos de este trabajo se utiliza el test de Breusch-Pagan, el cual consiste en realizar una regresión de los residuos al cuadrado sobre las mismas variables explicativas consideradas en el modelo principal, tal como se muestra en la ecuación (3.14)

$$\hat{\varepsilon}^2 = \alpha_0 + \sum_{k=1}^K \alpha_k \cdot x_k + u \quad (3.14)$$

Si se concluye que  $\alpha_k = 0$  para  $k = 1, \dots, K$ , se tiene que la varianza de los errores no depende de las variables explicativas, lo cual sugiere que el modelo es homocedástico. En estricto rigor, se tiene que el estadístico de este test está dado por  $BP = n \cdot R^2$ , siendo la hipótesis nula que dicho estimador distribuye  $\chi_K^2$ . La implementación de este test en RStudio se lleva a cabo a través del método *bptest*, perteneciente al paquete *lmtest*. Este método requiere como argumento únicamente el modelo original.

Una forma de corregir la heterocedasticidad cuando no se conocen las varianzas es a través de la estimación por mínimos cuadrados generalizados factibles (FGLS: *feasible generalized least squares*). Una versión del modelo FGLS corresponde al método de mínimos cuadrados ponderados, el cual consiste en construir un modelo auxiliar a través de la ponderación del



modelo lineal original por el inverso de la varianza (o un estimador de ella, dado que es desconocida). El resultado de esto si el estimador de la varianza por observación es adecuado, es que las varianzas de los residuos se hacen constantes, logrando que el modelo sea homocedástico.

En este trabajo, se obtiene que prácticamente la totalidad de los modelos presenta heterocedasticidad. Sin embargo, no fue posible dar con un estimador adecuado para corregir los modelos por heterocedasticidad, en combinación con el problema de endogeneidad en el contexto en estudio. En consecuencia, en vez de tratar de recuperar la eficiencia de los estimadores, se optó por corregir los errores estándar a través de métodos no paramétricos. En particular, se utilizó *bootstrapping*, un proceso que consiste en realizar  $R$  muestreos aleatorios con reemplazo y realizar una estimación en base a dichas muestras. Esto permite recolectar  $R$  estimadores para cada variable explicativa y posteriormente calcular su valor promedio, error estándar, valor del test- $t$  y valor- $p$ . En este caso, se consideraron 500 repeticiones.

### 3. Autocorrelación Espacial:

Un aspecto que puede estar influyendo en la correlación entre residuos, particularmente en los modelos espaciales, es la ubicación geográfica de las observaciones. Se dice que en un área geográfica existe autocorrelación espacial cuando las observaciones (zonas) que están más cerca entre sí tienden a ser más parecidas, siguiendo lo que se conoce como la Ley de Geografía de Tobler: “*todo esta relacionado con todo lo demás, pero las cosas cercanas están más relacionadas que las cosas distantes*” (Tobler, 1970). Resulta natural pensar que en este caso puede existir autocorrelación espacial, ya que los límites de los distritos censales son de carácter administrativo y no debieran haber grandes diferencias en el comportamientos de viaje entre zonas vecinas.

Para detectar la presencia de autocorrelación espacial, existen algunos indicadores que miden el nivel de autocorrelación espacial a nivel global para una variable particular, siendo el más destacado el estadístico  $I$  de Moran (Borrego Sánchez, 2018). Este estadístico está dado por la siguiente expresión:

$$I = \frac{N}{\sum_{i=1}^N \sum_{j=1}^N W_{ij}} \cdot \frac{\sum_{i=1}^N \sum_{j=1}^N W_{ij} (z_i - \bar{z})(z_j - \bar{z})}{\sum_{i=1}^N \sum_{j=1}^N (z_i - \bar{z})^2} \quad (3.15)$$

donde  $N$  es el número de observaciones,  $z_i$  y  $z_j$  los valores de la variable  $Z$  para la  $i$ -ésima y  $j$ -ésima observación, respectivamente. Por su parte,  $\{W_{ij}\}$  es una matriz de pesos que, en base a algún criterio, define el vecindario de cada zona. Este estadístico toma valores entre -1 y 1: los valores cercanos a los extremos indican una alta correlación (positiva o negativa según el signo), mientras que los valores cercanos a cero sugieren que no hay autocorrelación espacial.

El estadístico  $I$  de Moran puede utilizarse como un test de autocorrelación espacial utilizando los residuos  $\hat{\varepsilon} = Y' - X\hat{\beta}$  de la regresión OLS para su cálculo. El índice de Moran aplicado a los residuos, adopta la siguiente forma:



$$I = \frac{N}{W_0} \cdot \frac{\hat{\varepsilon}^T W \hat{\varepsilon}}{\hat{\varepsilon}^T \hat{\varepsilon}} \quad (3.16)$$

donde  $W_0 = \sum_{i=1}^N \sum_{j=1}^N W_{ij}$  y  $W = W_{ij}$ . El test de Moran realiza una aproximación normal sobre el valor estandarizado de  $I$ : la hipótesis nula indica que no existe correlación espacial, mientras que la hipótesis alternativa indica que si la hay. El test estadístico queda del siguiente modo:

$$Z(I) = \frac{I - E(I)}{\text{Var}(I)} \sim N(0, 1) \quad (3.17)$$

Para implementar este test en RStudio, se requiere instalar previamente los siguientes paquetes:

- *rgdal*: de acuerdo a la descripción del propio paquete, proporciona enlaces a la librería de abstracción de datos espaciales ('GDAL') y otorga acceso a las operaciones de proyección y transformación de la librería 'PROJ', por lo que también es necesaria la instalación de dichas librerías. En la práctica, el paquete *rgdal* permite leer datos (por ejemplo, en formato .shp) y entregar un objeto espacial, conservando sus características geométricas.
- *spdep*: este paquete posee un conjunto de funciones que permite construir las matrices de pesos espaciales a partir de la cercanía de los polígonos y también permite llevar a cabo los test de autocorrelación espacial.

La implementación propiamente tal se realiza mediante el comando *lm.morantest*, el cual requiere como argumentos el vector de residuos y la matriz de pesos en formato de lista. Entre los resultados se reporta el valor del índice  $I$  de Moran, su valor esperado y varianza, además del  $p$ -valor. Este proceso se realiza para los 40 modelos, cuyos resultados se reportan en el siguiente capítulo.

Si bien existen modelos específicos que abordan la corrección de la autocorrelación espacial para recuperar la eficiencia de los parámetros, en este trabajo no son utilizados. En lugar de posiblemente comprometer la consistencia de los estimadores por autocorrelación que pueda basarse en supuestos arbitrarios, el esfuerzo se centra en corregir los errores estándar para hacer inferencias. A este respecto, cabe señalar que los errores estándar obtenidos a través de las regresiones con bootstrapping para la corrección por heterocedasticidad también contemplan la corrección por autocorrelación espacial.

### 3.2.4. Capacidad predictiva de los modelos

Finalmente, una vez que se han estimado los modelos, se evalúa la capacidad predictiva de cada uno de ellos. Para dicha labor, se utilizan los datos de la primera semana para estimar los modelos y posteriormente se evalúa cada modelo con los datos de la segunda semana. La capacidad predictiva de cada modelo se mide a través del error porcentual medio absoluto

(MAPE, por sus siglas en inglés), dado por la Ecuación (3.18):

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{T_i}{t_i} - 1 \right| \quad (3.18)$$

donde  $T_i$  es el valor de la variable explicada estimada mediante los modelos,  $t_i$  es el valor de la variable explicada observada en los datos y  $N$  es la cantidad de observaciones.

La principal ventaja de este indicador es que permite medir la capacidad de predicción de manera relativa, facilitando su interpretación. No obstante lo anterior, es importante señalar que se pueden obtener valores que superan el 100 % si la diferencia entre el valor estimado y el valor observado es demasiado grande. Por otra parte, la principal desventaja de este indicador es que no funciona cuando el valor de la variable explicada observada ( $t_i$ ) es cero, ya que la expresión se indetermina.

# Capítulo 4

## Resultados y Discusión

### 4.1. Resultados de los modelos

A continuación, se presentan los resultados de los modelos descritos en la sección 3.2 para cada una de las combinaciones día-período. Cada tabla contiene los resultados de cuatro modelos, correspondientes a cada tipo de día para un período determinado. Las tablas reportadas en esta sección contienen los valores de los coeficientes, errores estándar (números entre paréntesis) y el nivel de significancia de las variables en cada modelo, este último siguiendo la notación utilizada en el software RStudio<sup>1</sup>. Los errores estándar corresponden a los obtenidos a través del proceso de *bootstrapping* descrito en el capítulo anterior. Además, para cada modelo se reporta la bondad de ajuste, los resultados de los test de hipótesis de endogeneidad (Instrumentos Débiles, Wu-Hausman y Sargan), el resultado del test de correlación espacial de Moran y la capacidad de predicción, a través del indicador MAPE<sup>2</sup>.

Cabe señalar que las variables que no fueron consideradas en los modelos resultaron ser sistemáticamente no significativas, como es el caso de las proporciones de población por género (*prop\_male*; *prop\_female*), las proporciones de población por rango de edad (*prop\_age\_18\_39*; *prop\_age\_40\_59*; *prop\_age\_60\_*) y la proporción de población con educación superior terminada (*prop\_higher\_educ*), entre otros. También existen variables que fueron incluidas en los modelos a pesar de que no son estadísticamente significativas (al 10%) en algunos de ellos, ya que permite mostrar los diferentes grados de influencia dependiendo del período o día que se está analizando. Dichos casos se explican más adelante.

Dado que ambas variables dependientes presentan resultados similares, se presentan únicamente los resultados para los viajes generados por km<sup>2</sup> los resultados para los viajes generados por cada 1.000 hogares se muestran en detalle en el Anexo D. Además, los resultados reportados en las siguientes tablas consideran como variable endógena el tiempo de espera, ya que este tipo de modelo presenta el comportamiento más plausible, de acuerdo a lo presentado en el Anexo C.

---

<sup>1</sup>Códigos de significancia de acuerdo a Prob(>|t|): 0 (\*\*\*) 0.001 (\*\*) 0.01 ( \* ) 0.05 ( . ) 0.10 ( ) 1

<sup>2</sup>Las observaciones de la primera semana con valor cero no fueron consideradas para el cálculo.

Tabla 4.1: Resultados modelación viajes por km<sup>2</sup> período N2 (00:00 - 07:00)

Variable	LUN-JUE			VIE			SAB			DOM		
intercepto	-5,855	(2,699)	***	-0,199	(2,705)		-1,688	(1,821)		1,578	(7,057)	
<i>Atributos de Ridesourcing</i>												
wait_time	-0,317	(0,290)	**	-0,847	(0,263)	***	-0,849	(0,258)	***	-0,924	(0,686)	***
spatial_mean_speed	-0,028	(0,019)	*	-0,002	(0,016)		0,003	(0,021)		-0,004	(0,030)	
<i>Sociodemografía</i>												
prop_med_income	1,275	(0,373)	***	0,520	(0,488)		-0,039	(0,464)		-0,047	(0,905)	
prop_high_income	1,781	(0,478)	**	1,966	(0,648)	**	1,467	(0,647)	*	-0,011	(1,066)	
prop_car_1	0,036	(0,506)		-0,689	(0,592)		0,791	(0,452)	.	-0,062	(0,593)	
prop_car_2+	-0,380	(0,406)		-0,168	(0,576)		-0,224	(0,567)		-0,328	(0,658)	
<i>Entorno Construido</i>												
log_pop_density	1,008	(0,135)	***	0,687	(0,148)	***	0,891	(0,121)	***	0,735	(0,286)	***
land_use_mix	1,413	(0,599)	***	0,851	(0,493)	.	0,866	(0,407)	*	0,251	(1,227)	
num_mall_1	0,007	(0,091)		0,054	(0,121)		0,023	(0,114)		0,017	(0,131)	
num_mall_2	-0,117	(0,443)		-0,585	(0,426)		-0,827	(0,633)	*	-0,225	(0,525)	
num_mall_3	0,412	(0,228)	.	0,274	(0,222)		0,257	(0,184)		0,307	(0,171)	
dummy_airport	3,484	(0,917)	***	4,171	(1,196)	**	3,502	(1,213)	**	1,151	(1,241)	
dummy_bus_station	1,204	(0,624)	**	0,991	(0,487)	.	0,977	(0,523)	*	0,841	(0,516)	.
dummy_CBD	0,109	(0,148)		0,250	(0,178)		-0,147	(0,178)		-0,057	(0,230)	
<i>Oferta de otros modos</i>												
stops_density	0,005	(0,005)		0,004	(0,005)		-0,003	(0,007)		0,003	(0,007)	
frec	-0,557	(0,197)	***	-0,239	(0,228)		-0,409	(0,243)	.	-0,079	(0,317)	
dummy_metro	0,190	(0,133)	*	-0,045	(0,158)		0,049	(0,130)		0,219	(0,133)	.
cycleways_density	0,028	(0,015)	.	0,050	(0,021)	*	0,059	(0,023)	**	0,026	(0,022)	
<i>Bondad de Ajuste</i>												
R <sup>2</sup>	0,697			0,421			0,572			0,470		
<i>Tests de Endogeneidad</i>												
Instrumentos Débiles	37,518			15,490			26,320			9,887		
Wu-Haussman	0,419			26,308			29,860			13,731		
Sargan	22,423			8,573			17,890			49,932		
<i>Correlación Espacial</i>												
Moran	0,339			0,129			0,087			0,187		
<i>Capacidad Predictiva</i>												
MAPE	86,5 %			126,3 %			36,9 %			38,7 %		
<i>Num. Observaciones</i>												
N	336			332			331			330		

Tabla 4.2: Resultados modelación viajes por km<sup>2</sup> período PM (07:00 - 10:00)

Variable	LUN-JUE			VIE			SAB			DOM		
intercepto	-2,205	(0,775)	**	-3,689	(0,819)	***	-2,981	(1,233)	**	-4,163	(1,557)	**
<i>Atributos de Ridesourcing</i>												
wait_time	-0,317	(0,045)	***	-0,253	(0,064)	***	-0,593	(0,124)	***	-0,540	(0,168)	***
spatial_mean_speed	-0,065	(0,009)	***	-0,031	(0,009)	***	-0,008	(0,007)		-0,002	(0,010)	
<i>Sociodemografía</i>												
prop_med_income	0,578	(0,204)	**	0,548	(0,234)	*	0,666	(0,301)	*	1,112	(0,357)	***
prop_high_income	1,819	(0,375)	***	2,113	(0,365)	***	0,068	(0,513)		0,073	(0,560)	
prop_car_1	-0,179	(0,213)		-0,180	(0,248)		0,155	(0,335)		-0,033	(0,418)	
prop_car_2+	-0,368	(0,303)		-0,513	(0,266)	.	-1,089	(0,388)	**	-0,875	(0,454)	*
<i>Entorno Construido</i>												
log_pop_density	0,832	(0,060)	***	0,866	(0,061)	***	0,929	(0,087)	***	0,932	(0,097)	***
land_use_mix	0,457	(0,190)	*	0,719	(0,220)	**	0,484	(0,321)		0,649	(0,368)	.
num_mall_1	-0,141	(0,066)	.	-0,228	(0,069)	**	-0,006	(0,091)		-0,028	(0,131)	
num_mall_2	0,207	(0,189)		0,023	(0,193)		-0,170	(0,261)		0,403	(0,275)	
num_mall_3	0,213	(0,111)		0,206	(0,088)	.	0,225	(0,152)		-0,038	(0,160)	
dummy_airport	1,591	(0,451)	*	2,782	(0,866)	.	3,542	(0,924)	***	2,320	(0,835)	*
dummy_bus_station	0,591	(0,373)	*	0,821	(0,489)	**	0,763	(0,924)	*	0,526	(0,375)	
dummy_CBD	0,017	(0,093)		0,099	(0,094)		-0,112	(0,135)		-0,373	(0,171)	*
<i>Oferta de otros modos</i>												
stops_density	0,005	(0,002)	.	0,005	(0,003)	.	-0,004	(0,004)		0,001	(0,003)	
frec	-0,061	(0,030)	*	-0,057	(0,033)	.	-0,098	(0,052)	.	-0,044	(0,097)	
dummy_metro	0,103	(0,044)	.	0,165	(0,056)	**	0,092	(0,066)		0,031	(0,093)	
cycleways_density	-0,008	(0,009)		-0,003	(0,011)		0,004	(0,014)		0,030	(0,016)	.
<i>Bondad de Ajuste</i>												
R <sup>2</sup>	0,812			0,751			0,697			0,552		
<i>Tests de Endogeneidad</i>												
Instrumentos Débiles	101,562		***	41,645		***	21,968		***	14,460		***
Wu-Haussman	13,965		***	13,642		***	24,680		***	25,760		***
Sargan	1,913			1,144			3,064		.	7,660		**
<i>Correlación Espacial</i>												
Moran	0,328		***	0,282		***	0,107		***	0,002		***
<i>Capacidad Predictiva</i>												
MAPE	14,6 %			23,8 %			35,6 %			171,6 %		
<i>Num. Observaciones</i>												
N	339			339			338			339		

Tabla 4.3: Resultados modelación viajes por km<sup>2</sup> período FP (10:00 - 17:30)

Variable	LUN-JUE		VIE		SAB		DOM					
intercepto	1,032	(0,730)	-0,198	(0,820)	-0,975	(0,713)	-0,881	(0,967)				
<i>Atributos de Ridesourcing</i>												
wait_time	-0,487	(0,045)	***	-0,375	(0,060)	***	-0,457	(0,076)	***	-0,559	(0,074)	***
spatial_mean_speed	-0,043	(0,010)	***	-0,034	(0,008)	***	-0,025	(0,010)	**	-0,026	(0,009)	**
<i>Sociodemografía</i>												
prop_med_income	0,332	(0,166)	.	0,563	(0,180)	**	0,346	(0,169)	.	0,349	(0,221)	.
prop_high_income	0,779	(0,265)	**	1,387	(0,306)	***	0,651	(0,263)	*	0,263	(0,298)	.
prop_car_1	-0,379	(0,253)	.	-0,377	(0,264)	.	-0,249	(0,263)	.	0,069	(0,240)	.
prop_car_2+	-0,235	(0,248)	.	-0,527	(0,254)	*	-0,926	(0,258)	***	-0,696	(0,260)	**
<i>Entorno Construido</i>												
log_pop_density	0,567	(0,056)	***	0,616	(0,062)	***	0,791	(0,052)	***	0,806	(0,065)	***
land_use_mix	0,830	(0,176)	***	1,166	(0,186)	***	0,992	(0,188)	***	0,786	(0,211)	***
num_mall_1	0,190	(0,123)	**	0,145	(0,102)	.	0,109	(0,118)	.	-0,010	(0,090)	.
num_mall_2	0,263	(0,167)	.	0,228	(0,173)	.	0,337	(0,175)	*	0,207	(0,148)	.
num_mall_3	0,611	(0,104)	***	0,585	(0,112)	***	0,400	(0,108)	**	0,322	(0,130)	*
dummy_airport	2,676	(0,390)	***	1,876	(0,438)	**	2,719	(0,531)	***	2,143	(0,486)	***
dummy_bus_station	0,839	(0,204)	***	0,769	(0,274)	***	0,582	(0,281)	*	0,629	(0,358)	**
dummy_CBD	0,208	(0,090)	**	0,290	(0,101)	***	0,172	(0,093)	*	-0,069	(0,082)	.
<i>Oferta de otros modos</i>												
stops_density	0,006	(0,002)	**	0,006	(0,002)	**	0,003	(0,002)	.	0,006	(0,003)	*
frec	-0,037	(0,036)	.	-0,017	(0,037)	.	-0,087	(0,042)	*	-0,028	(0,049)	.
dummy_metro	0,133	(0,048)	**	0,115	(0,051)	*	0,191	(0,052)	***	0,052	(0,054)	.
cycleways_density	0,035	(0,012)	***	0,024	(0,011)	*	0,020		.	0,006	(0,009)	.
<i>Bondad de Ajuste</i>												
R <sup>2</sup>	0,841		0,812		0,809		0,825					
<i>Tests de Endogeneidad</i>												
Instrumentos Débiles	153,019		***	69,263		***	65,255		***	47,098		***
Wu-Haussman	20,025		***	11,353		***	34,847		***	28,618		***
Sargan	0,053			0,027			8,142		**	0,014		
<i>Correlación Espacial</i>												
Moran	0,195		***	0,182		***	0,149		***	0,132		***
<i>Capacidad Predictiva</i>												
MAPE	8,2 %		9,3 %		9,7 %		10,2 %					
<i>Num. Observaciones</i>												
N	339		339		339		339					

Tabla 4.4: Resultados modelación viajes por km<sup>2</sup> período PT (17:30 - 21:00)

Variable	LUN-JUE		VIE		SAB		DOM					
intercepto	0,254	(0,676)	-0,880	(0,853)	-1,010	(0,887)	-1,393	(0,985)	.			
<i>Atributos de Ridesourcing</i>												
wait_time	-0,375	(0,048)	***	-0,311	(0,061)	***	-0,433	(0,068)	***	-0,492	(0,069)	***
spatial_mean_speed	-0,052	(0,012)	***	-0,031	(0,009)	**	-0,017	(0,008)	*	-0,004	(0,009)	
<i>Sociodemografía</i>												
prop_med_income	0,406	(0,192)	*	0,503	(0,209)	*	-0,009	(0,225)		0,488	(0,200)	*
prop_high_income	1,638	(0,338)	***	1,720	(0,355)	***	0,531	(0,316)		0,436	(0,355)	
prop_car_1	-0,283	(0,261)		-0,107	(0,255)		0,173	(0,235)		-0,288	(0,229)	
prop_car_2+	-0,391	(0,294)		-0,663	(0,275)	*	-0,617	(0,298)	*	-0,497	(0,327)	.
<i>Entorno Construido</i>												
log_pop_density	0,506	(0,056)	***	0,565	(0,068)	***	0,682	(0,068)	***	0,656	(0,078)	***
land_use_mix	1,098	(0,187)	***	1,291	(0,204)	***	0,787	(0,220)	***	0,566	(0,239)	**
num_mall_1	0,096	(0,125)		0,086	(0,111)		0,116	(0,104)		0,086	(0,093)	
num_mall_2	0,553	(0,157)	***	0,461	(0,174)	**	0,523	(0,197)	**	0,639	(0,238)	***
num_mall_3	0,789	(0,103)	***	0,653	(0,120)	***	0,732	(0,092)	***	0,951	(0,131)	***
dummy_airport	2,407	(0,475)	***	1,635	(0,518)	*	2,277	(0,620)	***	2,350	(0,508)	***
dummy_bus_station	1,196	(0,281)	***	1,066	(0,400)	***	0,795	(0,361)	***	1,097	(0,436)	***
dummy_CBD	0,427	(0,123)	***	0,408	(0,105)	***	0,237	(0,122)	**	0,012	(0,136)	
<i>Oferta de otros modos</i>												
stops_density	0,007	(0,003)	**	0,011	(0,003)	***	0,003	(0,003)		0,003	(0,002)	
frec	-0,034	(0,027)		-0,040	(0,029)		-0,070	(0,039)	.	0,003	(0,045)	
dummy_metro	0,275	(0,055)	***	0,289	(0,052)	***	0,168	(0,056)	**	0,076	(0,056)	
cycleways_density	0,037	(0,015)	***	0,019	(0,012)	.	0,025	(0,013)	*	0,022	(0,011)	.
<i>Bondad de Ajuste</i>												
R <sup>2</sup>	0,807		0,775		0,796		0,796					
<i>Tests de Endogeneidad</i>												
Instrumentos Débiles	209,936		***	92,538		***	85,300		***	59,450		***
Wu-Haussman	57,893		***	42,321		***	28,220		***	32,060		***
Sargan	1,458			0,034			0,290			6,710		**
<i>Correlación Espacial</i>												
Moran	0,182		***	0,106		***	0,188		***	0,146		***
<i>Capacidad Predictiva</i>												
MAPE	15,2 %			63,9 %			17,5 %			26,5 %		
<i>Num. Observaciones</i>												
N	339			339			339			339		

Tabla 4.5: Resultados modelación viajes por km<sup>2</sup> período N1 (21:00 - 00:00)

Variable	LUN-JUE			VIE			SAB			DOM		
intercepto	-2,510	(1,465)	*	-1,156	(1,389)		0,658	(1,596)		-0,717	(1,610)	
<i>Atributos de Ridesourcing</i>												
wait_time	-0,754	(0,129)	***	-0,461	(0,111)	***	-0,674	(0,216)	***	-0,510	(0,102)	***
spatial_mean_speed	0,032	(0,033)	.	-0,019	(0,015)	.	-0,015	(0,024)		-0,031	(0,018)	**
<i>Sociodemografía</i>												
prop_med_income	0,664	(0,363)	*	1,130	(0,340)	***	0,474	(0,461)		0,097	(0,387)	
prop_high_income	1,177	(0,444)	*	1,972	(0,454)	***	-0,005	(0,607)		-0,888	(0,635)	
prop_car_1	0,426	(0,344)		-0,192	(0,312)		0,073	(0,382)		0,307	(0,367)	
prop_car_2+	0,057	(0,426)		-0,708	(0,361)	.	-0,762	(0,435)	.	-0,144	(0,544)	
<i>Entorno Construido</i>												
log_pop_density	0,842	(0,103)	***	0,676	(0,101)	***	0,719	(0,105)	***	0,801	(0,124)	***
land_use_mix	0,744	(0,298)	*	0,854	(0,288)	**	0,486	(0,379)		0,592	(0,362)	.
num_mall_1	0,113	(0,127)		0,096	(0,096)		0,216	(0,123)		0,121	(0,133)	
num_mall_2	0,147	(0,461)		-0,119	(0,342)		0,315	(0,293)		-0,077	(0,468)	
num_mall_3	0,916	(0,188)	***	0,698	(0,117)	***	0,662	(0,145)	**	0,892	(0,130)	***
dummy_airport	2,111	(0,953)	*	0,174	(0,671)		-2,190	(1,225)	.	1,966	(0,873)	.
dummy_bus_station	1,168	(0,623)	**	1,490	(0,582)	***	1,176	(0,402)	**	1,557	(0,665)	***
dummy_CBD	0,167	(0,097)		0,259	(0,100)	.	-0,086	(0,149)		-0,161	(0,132)	
<i>Oferta de otros modos</i>												
stops_density	0,002	(0,004)		0,011	(0,004)	**	0,015	(0,005)	**	0,008	(0,005)	.
frec	-0,431	(0,078)	***	-0,310	(0,097)	***	-0,306	(0,137)	***	-0,346	(0,120)	***
dummy_metro	0,222	(0,085)	*	0,253	(0,075)	**	0,048	(0,085)		0,182	(0,087)	.
cycleways_density	0,047	(0,014)	**	0,028	(0,015)	.	0,012	(0,019)		0,016	(0,018)	
<i>Bondad de Ajuste</i>												
R <sup>2</sup>	0,676			0,692			0,544			0,607		
<i>Tests de Endogeneidad</i>												
Instrumentos Débiles	109,41		***	58,733		***	25,540		***	53,736		
Wu-Haussman	18,09		***	14,711		***	45,620		***	39,244		
Sargan	16,50		***	5,108		*	14,120		***	0,518		
<i>Correlación Espacial</i>												
Moran	0,353		***	0,184		***	0,138		***	0,187		
<i>Capacidad Predictiva</i>												
MAPE	57,4 %			59,4 %			32,4 %			156,8 %		
<i>Num. Observaciones</i>												
N	336			333			335			337		

## 4.2. Interpretación de los modelos

En general, se obtiene un mejor ajuste y capacidad de predicción para los modelos de día laboral en los períodos PM, FP y PT, lo cual sugiere que las variables consideradas para estos modelos explican en gran parte la demanda de viajes en ridesourcing. Sin embargo, para el fin de semana y especialmente en los períodos nocturnos (N1 y N2) no se logra el mismo rendimiento, por lo que se sospecha que existen otras variables no incluidas en el modelo que explican la demanda en estos períodos. Los resultados obtenidos son similares para ambas variables dependientes consideradas, a excepción de que el coeficiente de densidad poblacional presenta signos opuestos. Además, los modelos de viajes diarios normalizados por superficie presentan mejor bondad de ajuste, posiblemente debido a que parte de las variables explicativas están naturalmente normalizadas por unidad de superficie, como por ejemplo la densidad poblacional, la densidad de paraderos y la densidad de ciclovías.



### 4.2.1. Atributos de Ridesourcing

En todos los modelos se observa una relación negativa entre la demanda y el tiempo de espera, es decir, las zonas con menor tiempo de espera generan una mayor demanda. Esta relación es razonable debido a que un menor tiempo de espera es un incentivo a elegir el ridesourcing como modo de transporte, e incluso a preferir Didi por sobre otras plataformas. Además, los test de hipótesis indican que efectivamente existe endogeneidad en el tiempo de espera y que las variables instrumentales propuestas son apropiadas.

El caso de la velocidad es más complejo. En los períodos laborales—diurnos existe una relación negativa con la demanda, lo cual indica que existe mayor demanda en las zonas con menor velocidad. Es bastante improbable que los usuarios prefieran deliberadamente este modo cuando es más lento, más bien lo que debe estar ocurriendo es que los DC con menor velocidad presentan mayor congestión debido a que son justamente los DC con mayor demanda. De hecho, en la estimación de los modelos se estudió una posible endogeneidad en esta variable, bajo la premisa de que el alto flujo de vehículos de Didi genera congestión, pero los resultados indicaron que la endogeneidad es más débil que la detectada en el tiempo de espera e incluso en algunos períodos no se detectó (ver Tabla C.2). Esto podría explicarse debido a que la congestión no depende solo del ridesourcing — y menos aún de Didi por sí solo — sino que es resultado de los flujos de los distintos modos que circulan en el sector.

### 4.2.2. Sociodemografía

En concordancia con la literatura (Correa *et al.*, 2017, Lavieri *et al.*, 2018, Ghaffar *et al.*, 2020, Marquet, 2020, Dey *et al.*, 2021), los modelos sugieren que hay mayor demanda en las zonas con mayor presencia de hogares de ingreso alto. Sin embargo, se debe señalar que esta relación positiva entre la demanda y el ingreso sólo se da en los días laborales. Aún así esto es razonable, pues los usuarios de ingreso alto tienen más facilidades para pagar de forma recurrente (por ejemplo, para viajar al trabajo) las tarifas de un viaje en ridesourcing, usualmente más altas que las tarifas del transporte público. Por otro lado, los resultados sugieren que durante el fin de semana el uso de ridesourcing es más transversal en términos socioeconómicos. Este resultado es similar al obtenido en Lavieri *et al.* (2018).

En cuanto a la posesión de automóvil, si bien en los modelos se obtuvo que los coeficientes no son estadísticamente significativos (posiblemente debido a su alta correlación con el ingreso), también indican una relación inversa entre la demanda y la proporción de hogares con auto, lo que sugiere que los usuarios sin automóvil son los más propensos a usar este modo, siendo también consistente con todos los estudios mencionados en la sección 2.3 que consideran este factor (ver Tabla 2.1). Es natural pensar que una persona que no tiene automóvil propio tiene mayores incentivos a utilizar estos servicios que una persona que sí lo tiene, ya que estos últimos sólo suelen dejar de lado su propio vehículo en situaciones puntuales, como por ejemplo ante la falta de estacionamientos o el consumo de alcohol. De hecho, de acuerdo a una encuesta sobre uso de Uber realizada en Tirachini y Gómez-Lobo (2020) indica que más del 20% de las personas encuestadas utilizan Uber debido — en parte — a la falta de disponibilidad o al alto costo de los estacionamientos, mientras que más del 40% lo utiliza — también en parte — debido a que así no deben conducir luego del consumo de alcohol.

Por otra parte, los resultados del modelo no arrojan una relación significativa de la demanda con la composición etaria y de género. Esto se podría explicar por la distribución espacial de ambas variables: como se mostró en la sección de análisis espacial, ambas variables presentan una distribución homogénea (ver Figuras 3.18 y 3.19), a diferencia de las variables de ingreso y posesión de automóvil que están mucho más segregadas. Además, es importante recordar que a cada viaje se están imputando las características correspondientes al origen del viaje, lo cual es más apropiado para los viajes generados desde el hogar. Para obtener una relación más definida de la demanda con la edad y género, sería más conveniente realizar una encuesta a nivel de usuario.

Por último, en el caso del nivel educativo de la población, tampoco se obtuvo una relación estadísticamente significativa con la demanda de viajes, a pesar de que esta variable presenta una distribución espacial segregada (ver Figura 3.20). El motivo por el cual puede estar sucediendo esto es que el nivel educativo de la población está fuertemente correlacionado con sus niveles de ingreso, es decir, las personas con un alto nivel educativo suelen pertenecer a estratos socioeconómicos altos. De hecho, el coeficiente de correlación de Pearson entre estas dos variables es de 0,7 (ver Anexo B).

### 4.2.3. Entorno Construido

En primer lugar, se observan efectos opuestos en relación de la demanda con la densidad poblacional, dependiendo de la variable dependiente modelada. En el caso del número de viajes normalizados por superficie, se observa una relación positiva y estadísticamente significativa entre la demanda y la densidad poblacional en todos los días y períodos, siendo consistente con los resultados hallados en la literatura (Ghaffar *et al.*, 2020; Lavieri *et al.*, 2018; Yu y Peng, 2019, entre otros). Este resultado sugiere que hay mayor densidad de viajes en las zonas con mayor densidad poblacional, las cuales se ubican principalmente en el centro de Santiago. Por otro lado, al modelar los viajes diarios normalizados por hogares, se observa una relación negativa con la densidad poblacional, es decir, cuando la densidad poblacional crece, los viajes por hogar disminuyen. Esto se puede explicar por las diferencias de distribución espacial de la densidad poblacional y el número de hogares: como se mencionó en la sección de análisis espacial, los sectores centro y oriente presentan una alta densidad de población y número de hogares debido a la gran cantidad de edificios residenciales que existen en dichos sectores; por el contrario, los sectores periféricos presentan un gran número de hogares, pero baja densidad de población, debido a la mayor extensión de las zonas y a que las viviendas son mayoritariamente casas.

En segundo lugar, los modelos indican una relación positiva entre la demanda y el mix de uso de suelo en todas las combinaciones día–período, incluyendo los fines de semana y períodos nocturnos, es decir, hay mayor demanda en las zonas con mayor diversidad de actividades. Considerando que los modelos son de generación de viajes, una posible explicación para esto es que los viajes realizados en ridesourcing son desde los grandes centros de actividades hacia el hogar, en otras palabras, viajes de retorno. Más aún, se observa que durante el período punta mañana la relación es más débil, lo cual podría explicarse debido a que durante este período debieran haber más viajes al trabajo, los cuales son realizados desde el hogar.

Además, los modelos reafirman la importancia de los centros comerciales en la demanda de viajes, pues se observa una relación positiva de la demanda con la presencia de centros comerciales durante los períodos en que están en funcionamiento (FP y PT), especialmente con centros comerciales de gran envergadura. Del mismo modo, se verifica la gran influencia del aeropuerto y los terminales de buses. En el caso de la variable dummy de aeropuerto, se obtuvo una relación positiva y estadísticamente significativa en todas las combinaciones día-período, inclusive en los períodos nocturnos y fines de semana. Por su parte, la variable dummy terminal presenta una relación algo más débil, pero aún así constituyen una fuente importante de generación de viajes en ridesourcing.

Por último, se tiene que la variable dummy CBD cobra mayor relevancia durante los períodos FP y PT, donde se asocia positivamente con la demanda. Este resultado indica que durante estos períodos se generan más viajes desde los centros de actividades, posiblemente de regreso al hogar desde actividades como el trabajo o compras.

#### 4.2.4. Oferta de transporte

Los modelos revelan en general una relación positiva de la demanda con la presencia de paraderos y estaciones de metro, coincidiendo con los resultados de trabajos anteriores que incluyeron estas variables (Correa *et al.*, 2017; Brown, 2019; Ghaffar *et al.*, 2020; Sabouri *et al.*, 2020; Dey *et al.*, 2021). Este resultado indica que se tiene una mayor demanda en los sectores con mayor presencia de paraderos y estaciones de metro, principalmente durante los períodos FP, PT y N1. Una posible hipótesis para explicar este resultado es que los usuarios están usando los servicios ridesourcing para el viaje de última milla, lo que sugiere que este modo se complementa con el transporte público. Sin embargo, esto no concuerda con los resultados de Tirachini y del Río (2019), quienes obtuvieron que el 96,1% de los usuarios utiliza ridesourcing sin combinar con otros modos. Una segunda opción es que los usuarios están viajando en este modo para retornar al hogar desde centros de actividades, que son justamente las zonas con mayor densidad de paraderos y de estaciones de metro. Una tercera posibilidad es que los usuarios de ridesourcing suelen ser multimodales (Tirachini, 2020): si bien la evidencia sugiere que los usuarios no suelen combinar con otros modos dentro de un mismo viaje, siguen siendo usuarios habituales de otros modos y particularmente de transporte público, por lo que pueden combinar modos en el ida y vuelta. Dicho comportamiento sólo se puede dar en zonas que son servidas por el transporte público. Resulta importante señalar que en la realidad, lo más probable es que la relación observada entre el uso del ridesourcing y la infraestructura de transporte público sea una combinación entre estos y otros factores.

En cuanto a la frecuencia de los buses, se obtuvo una relación negativa con la demanda durante los períodos PM, N1 y N2, es decir, hay mayor uso de ridesourcing en los sectores y períodos en donde la frecuencia de buses es más escasa. Esto es consistente con lo obtenido por Lavieri *et al.* (2018), quienes también utilizan esta variable. Este resultado sugiere que parte del uso del ridesourcing se debe a la falta de disponibilidad de transporte público: si bien en el período PM la frecuencia de buses es mayor, las personas suelen contar con menos holgura de tiempo y una baja frecuencia es más preocupante; por otro lado, durante los períodos N1 y N2 la frecuencia disminuye considerablemente — en algunos casos a tal punto que ya no circulan buses — lo que obliga a las personas a optar por otras alternativas. Este

efecto no se observa en los períodos FP y PT, donde la disponibilidad de buses es mayor y además en dichos períodos las personas cuentan con mayor holgura de tiempo, a diferencia de lo que ocurre en la punta mañana.

Finalmente, se tiene que en general la oferta de ciclovías se relaciona de forma positiva con el uso de ridesourcing, lo que sugiere que estos modos podrían ser complementarios. No obstante, debido al bajo uso combinado del ridesourcing con otros modos, este resultado seguramente se explica debido a que la oferta de ciclovías está parcialmente correlacionada con los sectores definidos como CBD, que son zonas en donde se tiene una mayor demanda.

### 4.3. Patrón de viajes

Los análisis realizados en el patrón de viajes revelan ciertas semejanzas con algunos de los resultados obtenidos por Tirachini y del Río (2019). En primer lugar, se verifica que los días con mayor demanda de viajes son viernes y sábado: en Tirachini y del Río (2019) el 28,3 % de los usuarios encuestados responde que el viernes es el día que utilizan más ridesourcing y un 35,9 % responde que el sábado; mientras que en este trabajo se obtiene que el 16,0 % de los viajes se realizaron durante los viernes y un 16,4 % los días sábado. En segundo lugar, los horarios de mayor demanda también coinciden: en Tirachini y del Río (2019) el 35,6 % de los usuarios responde que el horario en que más utilizan ridesourcing es entre 20:00 y 23:59, mientras que un 22,7 % responde entre 17:00 y 20:00; por otro lado, en este trabajo se obtuvo que el 21,5 % y un 17,4 % de los viajes se realizan respectivamente en dichos horarios. Por último, se observan semejanzas en la frecuencia de uso: en Tirachini y del Río (2019) un 58 % de los usuarios usaron Uber entre 2 y 4 veces en un mes, un 19 % entre 5 y 8 veces y un 9 % lo utilizaron más de 8 veces. Si se amplifican por dos las frecuencias obtenidas en la Figura 3.8, se tiene de manera aproximada que en un mes el 62,4 % utilizaron Didi entre 2 y 4 veces, un 11,8 % entre 5 y 8 veces y el 25,8 % restante más de 8 veces.

A pesar de que el período de análisis es de solo dos semanas, se logra apreciar una regularidad en la demanda a escala semanal. Sin embargo, no existe garantía de que éste sea el patrón más recurrente durante el año, ya que este período está ubicado dentro de la temporada de vacaciones, por lo que no se están considerando, por ejemplo, los viajes con motivo estudio; aunque por otro lado pueden haber más viajes debido a una mayor presencia de turistas. Además, se debe tener en consideración que el período de análisis es muy cercano al período en que se produjo el Estallido Social<sup>3</sup> (18 de octubre de 2019), lo cual podría haber afectado el patrón de viajes.

Además de la regularidad semanal, también se observa una regularidad dentro de la semana, lo que permite inferir algunos propósitos de viaje. El peak en el período PM de los días laborales se puede atribuir principalmente a viajes realizados con propósito trabajo. De hecho, en Tirachini y del Río (2019) los viajes con propósito trabajo conforman el 17,4 % del total de viajes realizado por los encuestados, mientras que en este trabajo se tiene que los viajes realizados durante el período PM en día laboral corresponden al 11,3 % del total de viajes. Esta hipótesis se puede reafirmar al comparar con el patrón de viajes observado

<sup>3</sup><https://www.dw.com/es/la-cronolog%C3%ADa-del-estallido-social-de-chile/a-51407726>

durante la mañana del sábado y domingo, donde no se observa este peak. Más aún, dado que los viajes con propósito trabajo son recurrentes, es razonable pensar que son realizados por usuarios con un nivel de ingreso alto, lo que es consistente con los resultados de los modelos.

El período fuera de punta no es tan claro, aunque la fuerte influencia de la presencia de centros comerciales en la demanda sugiere que una parte importante de los viajes podría ser de retorno al hogar luego de actividades de compras. De acuerdo a los resultados de [Tirachini y del Río \(2019\)](#), el 17,4 % de las personas encuestadas reporta usar ridesourcing con propósito de compras, mientras que los resultados obtenidos en el presente trabajo indican viajes realizados en el período fuera de punta corresponden al 37,3 %. Considerando que la naturaleza de este propósito no es tan frecuente en comparación a otros como trabajo o estudio, no explicaría completamente los viajes realizados en este horario. Otros posibles propósitos son ocio, salud o visitas.

El peak que surge en la punta tarde de los días laborales se puede explicar por los viajes de retorno al hogar desde el trabajo. Este comportamiento se acentúa durante el viernes, ya que el crecimiento de la demanda comienza más temprano y el peak es mayor, lo que podría explicarse debido a que la jornada laboral suele terminar más temprano en este día, dando paso a viajes de ocio y entretenimiento (55,4 % de las personas usan ridesourcing con este propósito, según [Tirachini y del Río, 2019](#)), que son de carácter ocasional. Esto es aún más notorio en el sábado, donde se tiene la máxima demanda de la semana. Estas actividades de ocio – generalmente en bares y restaurantes – se extienden hasta la madrugada, por lo cual el aumento de la demanda en las noches de viernes y sábado se debe a viajes de retorno al hogar desde estas actividades.

# Capítulo 5

## Conclusiones

En este trabajo se investigaron los factores que participan en el uso de ridesourcing en Santiago de Chile. En particular, se utilizaron datos de viajes realizados mediante la plataforma Didi Chuxing, junto con datos de sociodemografía, entorno construido y oferta de transporte para desarrollar modelos de generación de viajes por día-período a nivel de distrito censal. Entre las contribuciones de este estudio a la literatura se encuentra la extensión del alcance de esta línea de investigación al contexto latinoamericano, la corrección de los modelos para el cumplimiento de las condiciones de Gauss-Markov y la detección de diferentes grados de influencia de los factores de acuerdo al día y período en que se realizan los viajes.

A pesar de que los contextos varían entre ciudades y varían aún más entre países, los resultados de este estudio presentan varias semejanzas con los trabajos previos de este tipo, realizados en su mayoría en ciudades de Estados Unidos. Los resultados obtenidos sugieren que la demanda de viajes es mayor en sectores con (1) mayor poder adquisitivo, (2) menor posesión de automóvil, (3) mayor densidad poblacional, (4) mayor diversidad de actividades, (5) mayor presencia de centros comerciales, (6) presencia de aeropuertos y terminales de buses, (7) mayor cantidad de infraestructura de transporte público, (8) menor frecuencia de buses y (9) menor tiempo de espera. En este aspecto, se tienen tres resultados interesantes: en primer lugar, se obtuvo que existe una mayor participación de los hogares de ingreso alto durante los días laborales, mientras que durante el fin de semana su uso es transversal; en segundo lugar, se obtuvo que el tiempo de espera se relaciona de manera negativa con el tiempo de espera, que si bien es un resultado esperable debido a que un mayor tiempo de espera naturalmente desincentiva el uso del modo en cuestión, es un aspecto que no se había medido en estudios anteriores; en tercer lugar, se obtuvo que la demanda se relaciona de forma negativa con la velocidad en períodos con mayor actividad, dando cuenta de la existencia de congestión en los sectores con alta demanda.

De manera complementaria, se examinó el patrón espaciotemporal de viajes en estos servicios, lo que permitió inferir algunos de los propósitos de viaje para los cuales se usan estos servicios. Se observaron peaks en la mañana y en la tarde, además de un aumento considerable de la demanda en la noche de viernes y sábado, por lo que se concluye que los principales propósitos de viaje son ocio, retorno al hogar y trabajo. Los resultados de este análisis son consistentes con estudios previos realizados en Santiago.

Los resultados y métodos de este trabajo pueden cobrar gran relevancia para diversos actores. Por una parte, son relevantes para los operadores de Didi Chuxing (y eventualmente otras compañías de ridesourcing), ya que el conocimiento del comportamiento de los viajes realizados en este modo les permite elaborar estrategias para servir esta demanda de manera más eficiente, por ejemplo a través de la reasignación de conductores o la estimación de la flota óptima (Lavieri *et al.*, 2018). Por otra parte, estos resultados también pueden resultar de gran importancia para los entes reguladores y las autoridades que elaboran políticas públicas, ya que el conocer los factores que participan en el uso de ridesourcing y los impactos que éste provoca permite enfocar correctamente los esfuerzos para incentivar o desincentivar su uso según la situación lo requiera.

## Limitaciones del estudio

En primer lugar, se debe tener en cuenta que el período de análisis es de sólo dos semanas, a diferencia de estudios previos que contaron con varios meses de observaciones. Como se mencionó anteriormente, a pesar de que se logró observar cierta regularidad en el patrón de viajes, no se puede garantizar que éste sea el patrón de viajes que se observaría durante todo el año. Esta situación es especialmente problemática para los días viernes, sábado y domingo, ya que sólo se cuenta con un día para estimar y uno para predecir. De hecho, en general se obtuvo que la capacidad de predicción para estos días es bastante menor que para el resto de los días.

En segundo lugar está la falta de información del usuario. Con el fin de resguardar la privacidad de los usuarios de Didi, los datos de viajes se encuentran anonimizados. Este problema no es exclusivo de este estudio, ya que las bases de datos utilizadas en trabajos anteriores (por ejemplo, las de Chicago y Austin) han utilizado ésta y otras estrategias con el mismo fin. Para superar esta dificultad, se consideraron las características sociodemográficas de los DC en que se realizaron los viajes. Esta acción en sí misma también es una limitación, ya que se están imputando a los usuarios las características de la zona en la cual solicitaron el viaje, lo cual es válido sólo en los viajes basados en el hogar. Si bien los resultados obtenidos responden correctamente a este comportamiento durante la punta mañana laboral, es posible que de contar con la información real de usuario se detecten otros comportamientos en el resto de los períodos.

Por último, se tiene la inconsistencia en las fechas de recolección de datos. Los datos de viajes en ridesourcing son de enero de 2020; los datos de sociodemografía fueron recolectados en el año 2012, año en que se realizó la última encuesta origen-destino; los datos de población corresponden al último Censo, realizado en 2017; los datos de uso de suelo son del año 2014; mientras que los datos de paraderos corresponden al programa de operación del 2019 y 2020. Los datos de estaciones de metro, centros comerciales, aeropuerto y terminales de buses son actuales debido a que fueron recolectados de manera manual. Si bien esto podría producir algunas imprecisiones en los resultados, la naturaleza de variables como la distribución de ingresos (por categorías) y el uso de suelo permite asumir que los cambios en el tiempo no son significativos.

## Desafíos Futuros

A pesar de los grandes avances en la literatura del ridesourcing, la investigación del uso de este modo de transporte mediante datos de viajes aún está dando sus primeros pasos, especialmente en Latinoamérica. Este método de recolección de datos abre una puerta para profundizar ésta y otras líneas de investigación. En particular, la base de datos de viajes en Didi tiene el potencial suficiente para refinar los resultados obtenidos en este trabajo e incluso abordar otros temas dentro de la literatura.

En el ámbito de la modelación se puede avanzar tanto en la especificación como en las variables consideradas. Como primera aproximación, en este trabajo se utilizó la regresión lineal por 2SLS para capturar la relación del uso de ridesourcing con las distintas variables, garantizando que se cumplan las condiciones fundamentales en la estimación; sin embargo, se puede avanzar en este aspecto a través de otros métodos de estimación o mediante la implementación de modelos que consideren autocorrelación espacial. En cuanto a las variables de los modelos, se sugiere incorporar la presencia de bares y restaurantes, ya que se sospecha que son un determinante importante en la generación de viajes en ridesourcing, especialmente en los períodos nocturnos, donde se obtuvo un menor rendimiento. Para ello, se requiere un catastro con la ubicación de estos lugares.

Por otra parte, a partir de estos datos se pueden estimar algunos de los impactos abordados en la sección 2.1, particularmente el impacto sobre otros modos de transporte como taxi o transporte público. No obstante, para ello se necesita complementar esta base de datos con información de uso de otros modos, por ejemplo, a través de las transacciones de tarjetas “Bip!”. De esta manera, se podría estimar — al menos de forma parcial — el efecto del ridesourcing sobre el VKT. Dado que se cuenta con el ID de los conductores, se podría potenciar esto último a través de la estimación del deadheading. Otras posibles líneas de investigación utilizando estos datos son el análisis espacio-temporal de los tiempos de espera, comportamiento de conductores y exploración de factores que influyen en la cancelación de los viajes.



# Bibliografía

- Aguilera-García, A., Gomez, J., Velázquez, G., y Vassallo, J. M. (2022). Ridesourcing vs. traditional taxi services: Understanding users' choices and preferences in Spain. *Transportation Research Part A: Policy and Practice*, 155:161–178.
- Alemi, F., Circella, G., Handy, S., y Mokhtarian, P. (2018a). What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California. *Travel Behaviour and Society*, 13:88–104.
- Alemi, F., Circella, G., Mokhtarian, P., y Handy, S. (2018b). Exploring the latent constructs behind the use of ridehailing in California. *Journal of Choice Modelling*, 29:47–62.
- Anderson, D. (2014). Not just a taxi? For-profit ridesharing, driver strategies, and VMT. *Transportation*, 41.
- Andrews, I., Stock, J. H., y Sun, L. (2019). Weak Instruments in Instrumental Variables Regression: Theory and Practice. *Annual Review of Economics*, 11(1):727–753.
- Bekka, A., Louvet, N., y Adoue, F. (2020). Impact of a ridesourcing service on car ownership and resulting effects on vehicle kilometers travelled in the Paris Region. *Case Studies on Transport Policy*, 8(3):1010–1018.
- Bellego, C., Benatia, D., y Pape, L. (2022). Dealing with logs and zeros in regression models.
- Bi, H. y Ye, Z. (2021). Exploring ridesourcing trip patterns by fusing multi-source data: A big data approach. *Sustainable Cities and Society*, 64:102499.
- Bloomberg CityLab (2016). A Florida transit agency takes on the digital divide in a partnership with Uber. <https://www.bloomberg.com/news/articles/2016-06-20/pinellas-county-partners-with-uber-to-offer-free-dial-a-ride-services>. Fecha de acceso 08-03-2022.
- Borrego Sánchez, J. (2018). Modelos de regresión para datos espaciales. Universidad de Sevilla.
- Brown, A. (2019). Redefining car access. *Journal of the American Planning Association*, 85(2):83–95.
- Brown, A. y LaValle, W. (2021). Hailing a change: comparing taxi and ridehail service quality in Los Angeles. *Transportation*, 48.

- Centro de Desarrollo Urbano Sustentable (2016). Uso de suelo - Chile. [http://datos.cedeus.cl/layers/geonode:cl\\_uso\\_suelo\\_geo](http://datos.cedeus.cl/layers/geonode:cl_uso_suelo_geo). Fecha de acceso 18-07-2021.
- Cervero, R. y Kockelman, K. (1997). Travel demand and the 3ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, 2(3):199–219.
- Chen, C., Feng, T., Shao, M., y Yao, B. (2021). Understanding the determinants of spatial-temporal mobility patterns based on multi-source heterogeneous data. *Transportation Research Procedia*, 52:477–484.
- Chicago Data Portal (2022). Taxi and Ridehailing Usage in Chicago. <https://toddschneider.com/dashboards/chicago-taxi-ridehailing-data/>. Fecha de acceso 18-02-2022.
- Clewlou, R. R. y Mishra, G. S. (2017). Disruptive transportation: The adoption, utilization, and impacts of ride-hailing in the United States.
- Consejo Internacional de Centros Comerciales (2017). U.S. Shopping-Center Classification and Characteristics. [https://www.icsc.com/uploads/research/general/US\\_CENTER\\_CLASSIFICATION.pdf](https://www.icsc.com/uploads/research/general/US_CENTER_CLASSIFICATION.pdf). Fecha de acceso 12-11-2020.
- Correa, D., Xie, K., y Ozbay, K. (2017). Exploring the taxi and uber demand in new york city: An empirical analysis and spatial modeling.
- Cramer, J. y Krueger, A. B. (2016). Disruptive change in the taxi business: The case of Uber. *The American Economic Review*, 106(5):177–182.
- de Souza Silva, L. A., de Andrade, M. O., y Alves Maia, M. L. (2018). How does the ride-hailing systems demand affect individual transport regulation? *Research in Transportation Economics*, 69:600–606. Competition and Ownership in Land Passenger Transport (selected papers from the Thredbo 15 conference).
- Dey, B. K., Tirtha, S. D., Eluru, N., y Konduri, K. C. (2021). Transformation of ridehailing in New York City: A quantitative assessment. *Transportation Research Part C: Emerging Technologies*, 129:103235.
- Dias, F., Lavieri, P., Garikapati, V., Astroza, S., Pendyala, R., y Bhat, C. (2017). A behavioral choice model of the use of car-sharing and ride-sourcing services. *Transportation*, 44.
- Directorio de Transporte Público Metropolitano (2019). Planes Operacionales Históricas. <https://www.dtpm.cl/index.php/noticias/plan-operacional-historico>. Fecha de acceso 14-06-2021.
- Fielbaum, A. y Tirachini, A. (2021). The sharing economy and the job market: the case of ride-hailing drivers in Chile. *Transportation*, 48.
- Ge, Y., Knittel, C. R., MacKenzie, D., y Zoepf, S. (2016). Racial and gender discrimination in transportation network companies. Working Paper 22776, National Bureau of Economic Research.

- Gerte, R., Konduri, K., y Eluru, N. (2018). Is there a limit to adoption of dynamic ridesharing systems? evidence from analysis of uber demand data from new york city. *Transportation Research Record: Journal of the Transportation Research Board*, 2672:036119811878846.
- Ghaffar, A., Mitra, S., y Hyland, M. (2020). Modeling determinants of ridesourcing usage: A census tract-level analysis of chicago. *Transportation Research Part C: Emerging Technologies*, 119:102769.
- Graehler, M., Mucci, R. A., y Erhardt, G. D. (2019). Understanding the recent transit ridership decline in major US cities: Service cuts or emerging modes. En *98th Annual Meeting of the Transportation Research Board, Washington, DC*.
- Gómez-Lobo, A., Tirachini, A., y Gutierrez, I. (2022). Optimal prices for ridesourcing in the presence of taxi, public transport and car competition. *Transportation Research Part C: Emerging Technologies*, 137:103591.
- Hall, J. D., Palsson, C., y Price, J. (2018). Is Uber a substitute or complement for public transit? *Journal of Urban Economics*, 108:36–50.
- Hall, J. V. y Krueger, A. B. (2018). An Analysis of the labor market for Uber’s driver-partners in the United States. *ILR Review*, 71(3):705–732.
- Henao, A. y Marshall, W. (2019). The impact of ride-hailing on vehicle miles traveled. *Transportation*, 46.
- Hughes, R. y MacKenzie, D. (2016). Transportation network company wait times in greater seattle, and relationship to socioeconomic indicators. *Journal of Transport Geography*, 56:36–44.
- IFC, Accenture, y Uber (2018). Driving toward equality: Women, ride-hailing, and the sharing economy.
- Instituto Nacional de Estadísticas (2018). Censo 2017 Distrito Censal: Población, viviendas por área y densidad. [https://geoine-ine-chile.opendata.arcgis.com/datasets/c6e62e8bd9f04a6a9df46fbeaf1f22fd\\_0/about](https://geoine-ine-chile.opendata.arcgis.com/datasets/c6e62e8bd9f04a6a9df46fbeaf1f22fd_0/about). Fecha de acceso 31-05-2021.
- Jin, S. T., Kong, H., Wu, R., y Sui, D. Z. (2018). Ridesourcing, the sharing economy, and the future of cities. *Cities*, 76:96–104.
- Kodransky, M. y Lewenstein, G. (2014). *Connecting Low-Income People to Opportunity with Shared Mobility*. Institute for Transportation and Development Policy and Living Cities.
- Kong, H., Zhang, X., y Zhao, J. (2020). How does ridesourcing substitute for public transit? a geospatial perspective in Chengdu, China. *Journal of Transport Geography*, 86:102769.
- Lavieri, P., F. Dias, F., Ruíz-Juri, N., Kuhr, J., y Bhat, C. (2018). A model of ridesourcing demand generation and distribution. *Transportation Research Record Journal of the Transportation Research Board*, 2672.
- Lavieri, P. S. y Bhat, C. R. (2019). Investigating objective and subjective factors influencing the adoption, frequency, and characteristics of ride-hailing trips. *Transportation Research Part C: Emerging Technologies*, 105:100–125.

- Li, S., Yang, H., Poolla, K., y Varaiya, P. (2021). Spatial pricing in ride-sourcing markets under a congestion charge. *Transportation Research Part B: Methodological*, 152:18–45.
- Li, W., Pu, Z., Li, Y., y (Jeff) Ban, X. (2019). Characterization of ridesplitting based on observed data: A case study of Chengdu, China. *Transportation Research Part C: Emerging Technologies*, 100:330–353.
- Marquet, O. (2020). Spatial distribution of ride-hailing trip demand and its association with walkability and neighborhood characteristics. *Cities*, 106:102926.
- Ministerio de Transporte y Telecomunicaciones (2018). Política de equidad de género en transportes.
- Ministerio de Transporte y Telecomunicaciones (2021). Visor de Ciclovías. <https://ciclovias.visorterritorial.cl/>. Fecha de acceso 14-06-2021.
- National Academies of Sciences, Engineering, and Medicine (2016). *Between Public and Private Mobility: Examining the Rise of Technology-Enabled Transportation Services*. The National Academies Press, Washington, DC.
- NYC Taxi and Limousine Commission (2022). Taxi and Ridehailing Usage in New York City. <https://toddschneider.com/dashboards/nyc-taxi-ridehailing-uber-lyft-data/>. Fecha de acceso 18-02-2022.
- Observatorio de Ciudades UC (2019). Estaciones Actuales Metro de Santiago. [https://ideocuc-ocuc.hub.arcgis.com/datasets/cc2e1d8419e64c7cb9502c58c00ba33b\\_0/about](https://ideocuc-ocuc.hub.arcgis.com/datasets/cc2e1d8419e64c7cb9502c58c00ba33b_0/about). Fecha de acceso 16-04-2021.
- Rayle, L., Dai, D., Chan, N., Cervero, R., y Shaheen, S. (2016). Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. *Transport Policy*, 45:168–178.
- Rizk, N., Salem, N., y Weheba, N. (2018). A gendered analysis of ridesharing: Perspectives from Cairo, Egypt. *Urban Transport in the Sharing Economy Era*, pp. 32–60.
- Sabouri, S., Park, K., Smith, A., Tian, G., y Ewing, R. (2020). Exploring the influence of built environment on Uber demand. *Transportation Research Part D: Transport and Environment*, 81:102296.
- SECTRA (2013). Encuestas de Movilidad. [http://www.sectra.gob.cl/encuestas\\_movilidad/encuestas\\_movilidad.htm](http://www.sectra.gob.cl/encuestas_movilidad/encuestas_movilidad.htm). Fecha de acceso 20-06-2021.
- Stock, J. H. y Watson, M. W. (2011). *Introduction to Econometrics*. Pearson/Addison Wesley.
- Tirachini, A. (2020). Ride-hailing, travel behaviour and sustainable mobility: an international review. *Transportation*, 47.
- Tirachini, A. y del Río, M. (2019). Ride-hailing in Santiago de Chile: Users' characterisation and effects on travel behaviour. *Transport Policy*, 82:46–57.

- Tirachini, A. y Gómez-Lobo, A. (2020). Does ride-hailing increase or decrease vehicle kilometers traveled (VKT)? A simulation approach for Santiago de Chile. *International Journal of Sustainable Transportation*, 14:1–18.
- Tobler, W. R. (1970). A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46:234–240.
- Uber Blog (2017). Uber hits 5 billion rides milestone. <https://www.uber.com/en-SG/blog/uber-hits-5-billion-rides-milestone/>. Fecha de acceso 28-12-2020.
- Wang, H. y Yang, H. (2019). Ridesourcing systems: A framework and review. *Transportation Research Part B: Methodological*, 129:122–155.
- Wang, S. y Noland, R. B. (2021). Variation in ride-hailing trips in Chengdu, China. *Transportation Research Part D: Transport and Environment*, 90:102596.
- Wenzel, T., Rames, C., Kontou, E., y Henao, A. (2019). Travel and energy implications of ridesourcing service in Austin, Texas. *Transportation Research Part D: Transport and Environment*, 70:18–34.
- Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.
- Wu, X. y MacKenzie, D. (2021). Assessing the VMT effect of ridesourcing services in the US. *Transportation Research Part D: Transport and Environment*, 94:102816.
- Young, M. y Farber, S. (2019). The who, why, and when of Uber and other ride-hailing trips: An examination of a large sample household travel survey. *Transportation Research Part A: Policy and Practice*, 119:383–392.
- Yu, H. y Peng, Z.-R. (2019). Exploring the spatial variation of ridesourcing demand and its relationship to built environment and socioeconomic factors with the geographically weighted poisson regression. *Journal of Transport Geography*, 75:147–163.
- Zhang, Y. y Zhang, Y. (2018). Exploring the relationship between ridesharing and public transit use in the United States. *International Journal of Environmental Research and Public Health*, 15:1763.

# Anexos

# Anexo A

## Bases de Datos

### A.1. Viajes en ridesourcing

Tabla A.1: Atributos de datos de solicitudes de viaje en Didi

Nombre	Tipo	Nombre	Tipo	Nombre	Tipo
stat_start_hour	otro	finish_time	instante	travel_id	ID
stat_end_hour	otro	arrive_time	instante	is_carpool_show	dummy
order_id	ID	charge_dur	otro	history_dest_change_cnt	otro
product_id	ID	charge_dur_fee	otro	is_history_dest_change	dummy
city_id	ID	will_pay	otro	is_serial_ord	dummy
county	ID	total_fee	otro	is_td_grab_before_cancel	dummy
passenger_id	ID	dynamic_amt	otro	is_td_grab_after_pas_cancel	dummy
driver_id	ID	combo_type	otro	is_td_grab_after_dri_cancel	dummy
order_status	categ.	close_reason	categ.	is_td_grab_after_serv_cancel	dummy
complete_type	categ.	call_pay_type	categ.	is_td_call	dummy
source_type	otro	require_level	categ.	is_td_finish	dummy
start_dest_dis	otro	grap_car_level	categ.	is_td_grab	dummy
call_time	instante	car_id	ID	td_grab_dur	otro
modify_time	instante	charge_dis	otro	td_real_time_actual_success_grap_dur	otro
grab_time	instante	is_broad	dummy	is_td_real_time_driver_arrive	dummy
forecast_dur	otro	compound_type	categ.	is_td_reassign_grab	dummy
starting_name	otro	dynamic_times	otro	is_td_dynamic_call_order	dummy
is_dynamic	dummy	count_price_type	categ.	is_td_reassign_grab_after_dri_cancel	dummy
est_pick_dur	otro	is_fixed_price_call	dummy	is_td_last_call	dummy
grap_order_type	categ.	pax_app_version	otro	country_code	otro
departure_time	instante	initial_order_ID	ID	pt	otro
cancel_time	instante	nocarpool_cap_amt	otro		
begin_charge_time	instante	broad_time	instante		

## A.2. Sociodemografía

Tabla A.2: Atributos de base de datos de hogares EOD 2012

Nombre	Tipo	Descripción	Ejemplo
Hogar	int (id)	ID de Hogar	102791
Sector	int (id)	ID de Macrozona	2
Zona	int (id)	ID de Zona	418
Comuna	string	Comuna del hogar	MAIPU
DirCoordX	float	Longitud (UTM)	335966,573
DirCoordY	float	Latitud (UTM)	6284758,905
Fecha	datetime	Fecha de encuesta	07-02-2013
DiaAsig	string	Día de encuesta	jueves
TipoDia	int (cat)	Tipo de día (1: laboral, 2: fin de semana)	1
Temporada	int (cat)	Temporada (1: normal, 2: estival)	2
NumPer	int	Número de personas que habita el hogar	2
NumVeh	int	Número de vehículos en el hogar	1
NumBicAdu	int	Número de bicicletas (adulto)	1
NumBicNino	int	Número de bicicletas (nino)	0
Propiedad	int (cat)	Tipo de propiedad de la vivienda	1
MontoDiv	int	Monto de dividendo	100899
MontoArr	int	Monto de arriendo	220000
IngresoHogar	int	Ingreso total de los habitantes del hogar	1786826
Factor	float	Factor de expansión	313,02

Tabla A.3: Atributos de base de datos de personas EOD 2012

Nombre	Tipo	Descripción	Ejemplo
Hogar	int (id)	ID de hogar	100010
Persona	int (id)	ID de persona	10001001
AnoNac	int	Año de nacimiento	1958
Sexo	int (cat)	Sexo de la persona (1: hombre, 2: mujer)	1
Relacion	int (cat)	Relación con el jefe de hogar	1
Viajes	int	Viajes realizados	2
LicenciaConducir	int (cat)	Tipo de Licencia de Conducir	3
PaseEscolar	int (cat)	Posee Pase Escolar (1: si, 2: no)	2
Adulto Mayor	int (cat)	Posee Tarjeta Adulto Mayor (1: si, 2: no)	2
Estudios	int (cat)	Nivel de Estudios	4
Curso	int (cat)	Curso alcanzado	8
Actividad	string (cat)	Estado laboral	A
Ocupación	int (cat)	Ocupación laboral	5
ActividadEmpresa	int (cat)	Tipo de actividad de la empresa	5
JornadaTrabajo	int (cat)	Tipo de jornada laboral	1
DondeEstudia	int (cat)	Lugar de estudio actual	1
DirActividadCoordX	float	Longitud Lugar de Trabajo (UTM)	323492,978
DirActividadCoordY	float	Latitud Lugar de Trabajo (UTM)	6264121,901
DirEstudiosCoordX	float	Longitud Lugar de Estudio (UTM)	338936,332
DirEstudiosCoordY	float	Latitud Lugar de Estudio (UTM)	6266321,283
NoUsaTransantiago	string (cat)	Motivo por el cual no usa Transantiago	G
Discapacidad	int (cat)	Tipo de discapacidad	9
Ingreso	int	Nivel de ingresos	240000
TramoIngreso	int (cat)	Categoría de Ingreso	2
Factor	float	Factor de Expansión	85,585



## A.3. Distritos Censales

Tabla A.4: Atributos de base de datos de distritos censales, Censo 2017

Nombre	Tipo	Descripción	Ejemplo
OBJECTID	int (id)	ID de distrito censal	94
REGION	int (id)	Código Región	13
NOM_REGION	string	Nombre de Región	REGIÓN METROPOLITANA DE SANTIAGO
PROVINCIA	int (id)	Código Provincia	131
NOM_PROVINCIA	string	Nombre de Provincia	SANTIAGO
COMUNA	int (id)	Código Comuna	13119
NOM_COMUNA	string	Nombre de Comuna	MAIPÚ
COD_DISTRITO	int (id)	Código Distrito	4
TIPO_DISTRITO	string	Tipo de Distrito (urbano/rural)	URBANO
CODIGO <sub>C</sub> 17	int (id)	Código Censo 2017	1311904
T_HOM_R	int	Población Masculina Rural	0
T_MUJ_R	int	Población Femenina Rural	0
T_POB_R	int	Población Rural	0
T_HOM_U	int	Población Masculina Urbana	227
T_MUJ_U	int	Población Femenina Urbana	213
T_POB_U	int	Población Urbana	440
T_HOM	int	Población Masculina	227
T_MUJ	int	Población Femenina	213
T_POB	int	Población Total	440
T_VIV_U	int	Número de Viviendas Urbanas	157
T_VIV_R	int	Número de Viviendas Rurales	0
T_VIV	int	Número de Viviendas	157
SUPERFICIE	float	Superficie en km <sup>2</sup>	1,03855813
DENSIDAD	float	Densidad poblacional	423,66430664

## A.4. Paraderos

Tabla A.5: Atributos de base de datos de paraderos, PO Transantiago 2019-2

Nombre	Tipo	Descripción	Ejemplo
ID	int (id)	ID de paradero	3131
CODINFRA	string	Código paradero (operadores)	L-13-14-105-OP
SIMT	string	Código paradero (usuario)	PI1349
COMUNA	string	Comuna donde se ubica el paradero	MAIPU
DESDECRUCE	string	Vía transversal anterior al paradero	CERRO SOMBRERO
HACIACRUCE	string	Vía transversal posterior al paradero	AVENIDA LOS PAJARITOS
NOMBRE_PAR	string	Nombre de paradero	Camino a Melipilla / esq. Av. Los Pajaritos
X	float	longitud (UTM)	336960,11
Y	float	latitud (UTM)	6288955,79
NSERVICIOS	int	Número de servicios que se detienen en el paradero	4
SERVICIOS	string	Lista de servicios que se detienen en el paradero	113R; 113cR; 115R; 115cI

## A.5. Centros Comerciales

Tabla A.6: Centros Comerciales en el área de estudio

Centro Comercial	Empresa	Comuna	Superficie [m <sup>2</sup> ]	Categoría
Alto Las Condes	Cencosud	Las Condes	121.215	3
Apumanque	Otros	Las Condes	35.000	1
Arauco El Bosque	Arauco	El Bosque	30.000	1
Arauco Estación	Arauco	Estación Central	67.000	2
Arauco Maipú	Arauco	Maipú	72.500	2
Arauco Quilicura	Arauco	Quilicura	32.000	1
Costanera Center	Cencosud	Providencia	129.829	3
Easton Outlet Mall	Otros	Quilicura	21.000	1
Espacio Urbano Gran Avenida	Espacio Urbano	San Miguel	-	1
Espacio Urbano La Dehesa	Espacio Urbano	Lo Barnechea	-	1
Espacio Urbano Las Rejas	Espacio Urbano	Estación Central	-	1
Espacio Urbano Plaza Maipú	Espacio Urbano	Maipú	-	1
Espacio Urbano Puente Alto	Espacio Urbano	Puente Alto	-	1
Mall Espacio M	Otros	Santiago	34.000	1
Mall Florida Center	Cencosud	La Florida	123.188	3
Mall Paseo Quilín	Otros	Peñalolén	-	1
Mall Paseo San Bernardo	Otros	San Bernardo	45.000	2
Mall Plaza Alameda	Mall Plaza	Estación Central	60.000	2
Mall Plaza Egaña	Mall Plaza	La Reina	92.000	3
Mall Plaza Los Domínicos	Mall Plaza	Las Condes	92.000	3
Mall Plaza Norte	Mall Plaza	Huechuraba	119.000	3
Mall Plaza Oeste	Mall Plaza	Cerrillos	155.000	3
Mall Plaza Sur	Mall Plaza	San Bernardo	77.000	2
Mall Plaza Tobalaba	Mall Plaza	Puente Alto	77.000	2
Mall Plaza Vespucio	Mall Plaza	La Florida	169.000	3
Mall Sport	Cencosud	Las Condes	45.500	2
Mall VIVO El Centro	Mall VIVO	Santiago	19.692	1
Mall VIVO Imperio	Mall VIVO	Santiago	30.348	1
Mall VIVO Los Trapenses	Mall VIVO	Lo Barnechea	26.679	1
Mall VIVO Panorámico	Mall VIVO	Providencia	6.953	1
Parque Arauco	Arauco	Las Condes	112.000	3
Patio Centro	Mall VIVO	Santiago	19.692	1
Portal La Dehesa	Cencosud	Lo Barnechea	66.734	2
Portal La Reina	Cencosud	La Reina	38.198	1
Portal Ñuñoa	Cencosud	Ñuñoa	32.396	1
VIVO Outlet La Florida	Mall VIVO	La Florida	19.295	1
VIVO Outlet Maipú	Mall VIVO	Maipú	9.792	1

# Anexo B

## Análisis Estadístico

### B.1. Histogramas de variables no temporales

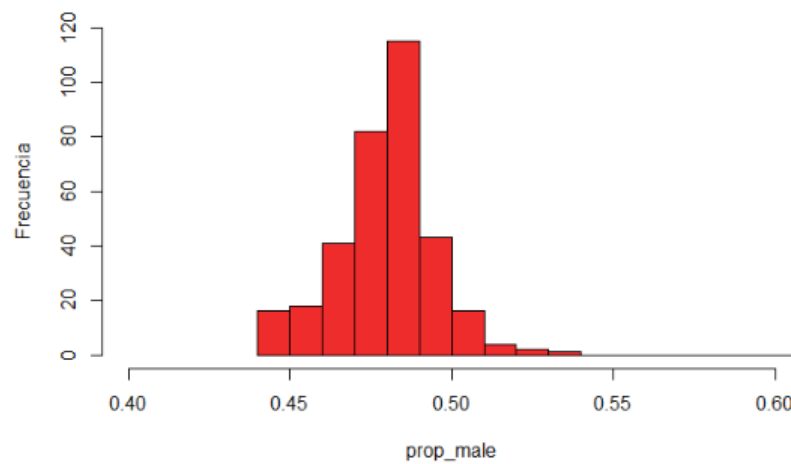


Figura B.1: Histograma de proporción de población masculina

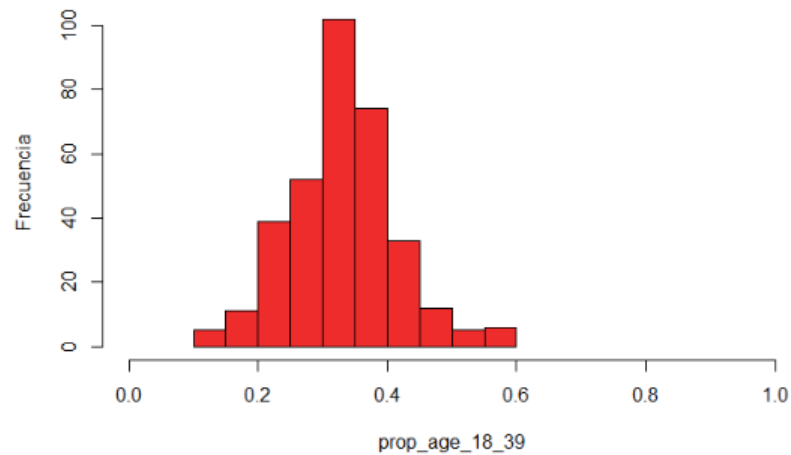


Figura B.2: Histograma de proporción de población entre 18 y 39 años

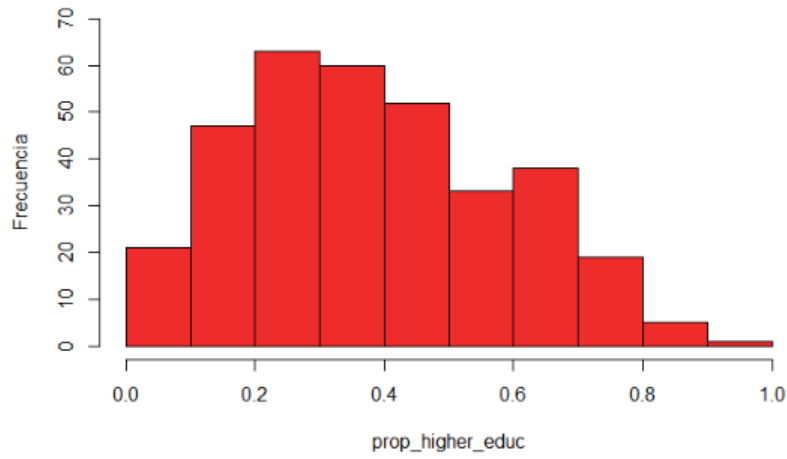


Figura B.3: Histograma de proporción de población con educación superior completada

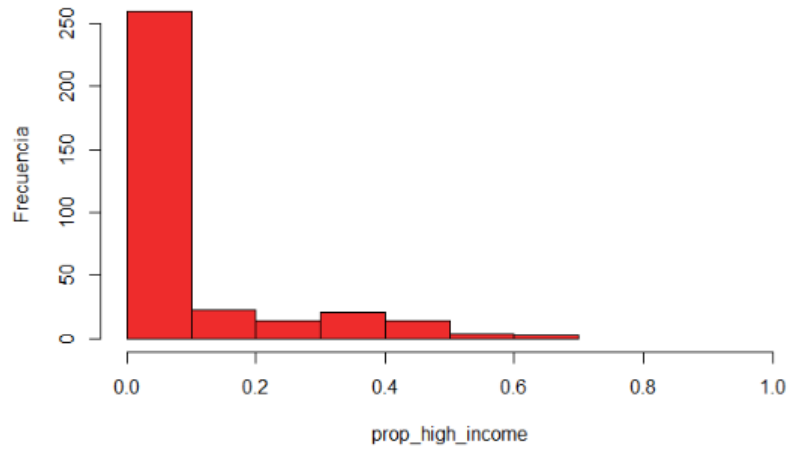


Figura B.4: Histograma de proporción de hogares de ingreso alto

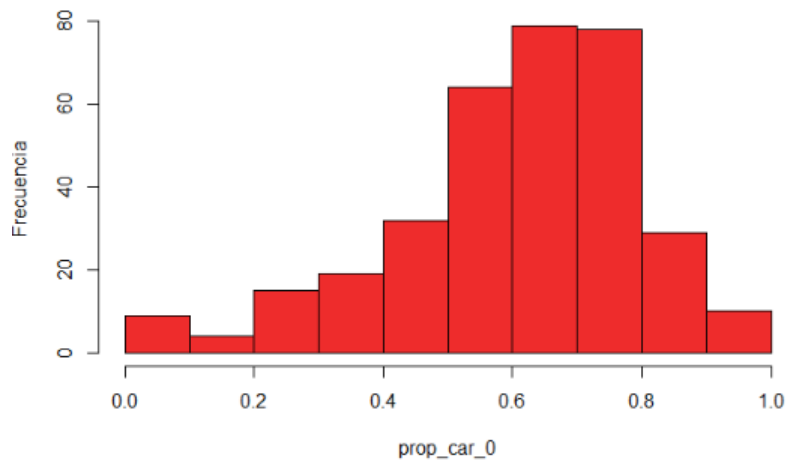


Figura B.5: Histograma de proporción de hogares sin auto

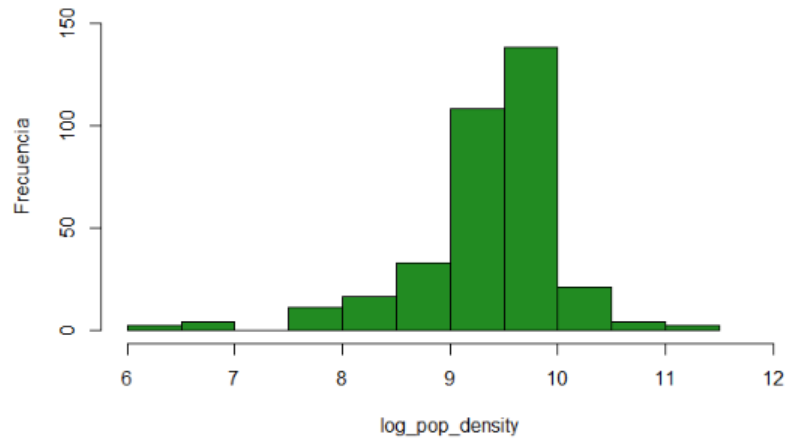


Figura B.6: Histograma de densidad poblacional (log)

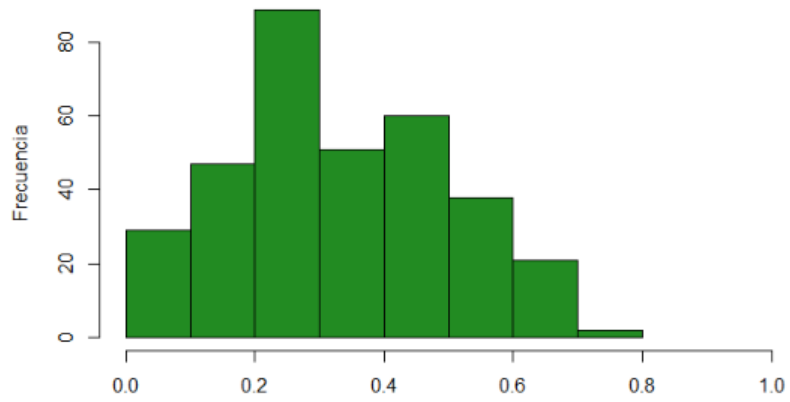


Figura B.7: Histograma de mix de uso de suelo

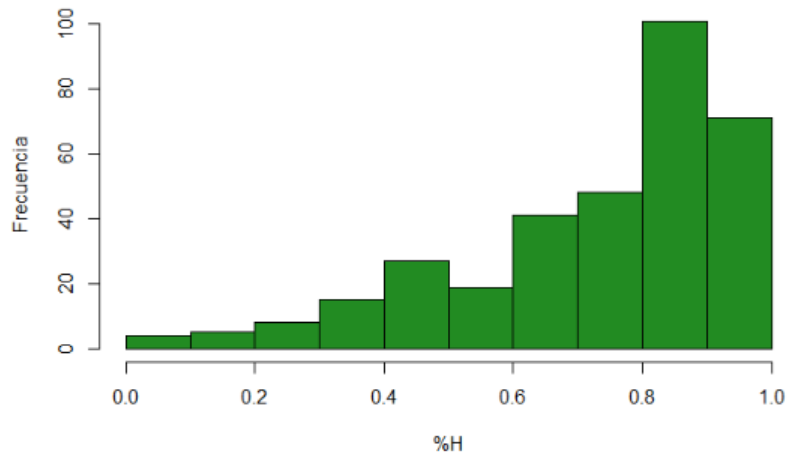


Figura B.8: Histograma de porcentaje de suelo con uso habitacional (H)

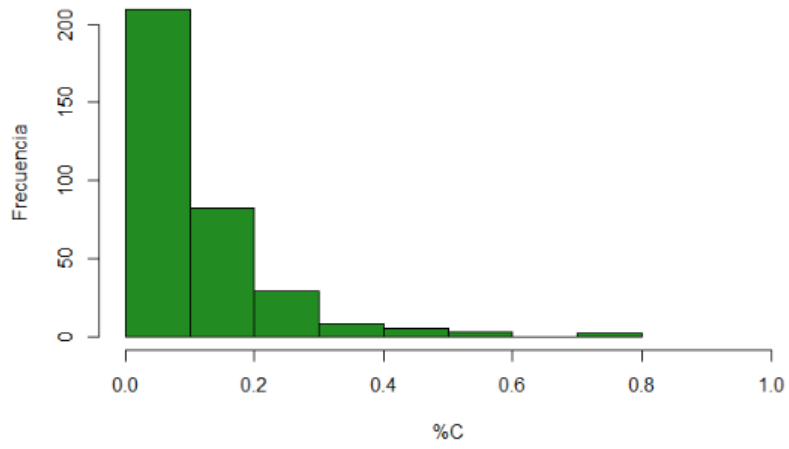


Figura B.9: Histograma de porcentaje de suelo con uso comercial (C)

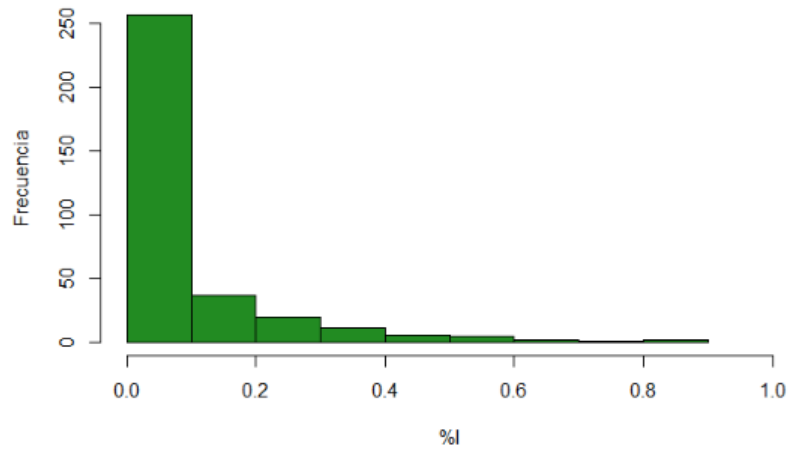


Figura B.10: Histograma de porcentaje de suelo con uso industrial (I)

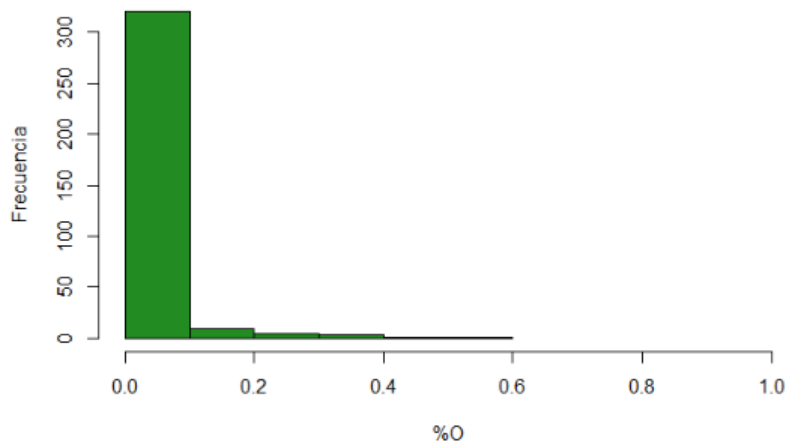


Figura B.11: Histograma de porcentaje de suelo con uso de oficinas (O)

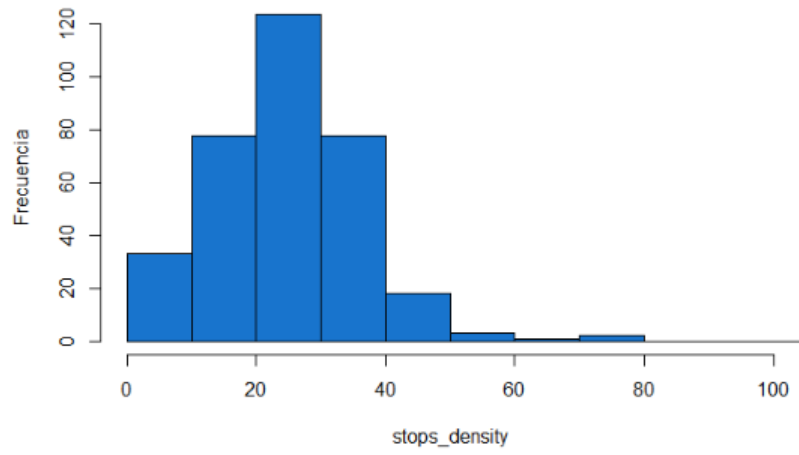


Figura B.12: Histograma de densidad de paraderos

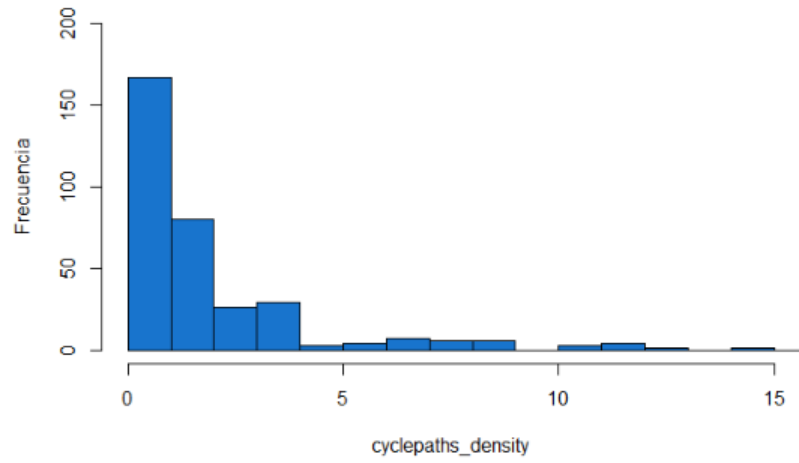


Figura B.13: Histograma de densidad de ciclovías





# Anexo C

## Comparación de modelos con endogeneidad

Tabla C.1: Comparación de modelos para día laboral período PM, viajes/km<sup>2</sup>

Variable	OLS			2SLS (tiempo espera)			2SLS (velocidad)		
intercepto	-3,285	(0,608)	***	-2,205	(0,775)	**	-5,603	(1,808)	***
<i>Atributos de Ridesourcing</i>									
wait_time	-0,172	(0,032)		-0,317	(0,045)	***	0,139	(0,185)	
spatial_mean_speed	-0,064	(0,009)	***	-0,065	(0,009)	***	-0,061	(0,011)	***
<i>Sociodemografía</i>									
prop_med_income	0,602	(0,190)	**	0,578	(0,204)	**	0,655	(0,293)	**
prop_high_income	1,572	(0,323)	*	1,819	(0,375)	***	1,042	(0,599)	*
prop_car_1	-0,180	(0,216)		-0,179	(0,213)		-0,182	(0,287)	
prop_car_2+	-0,408	(0,242)	.	-0,368	(0,303)		-0,493	(0,388)	.
<i>Entorno Construido</i>									
log_pop_density	0,870	(0,050)	***	0,832	(0,060)	***	0,953	(0,111)	***
land_use_mix	0,682	(0,176)	***	0,457	(0,190)	*	1,167	(0,367)	***
num_mall_1	-0,145	(0,073)	.	-0,141	(0,066)	.	-0,153	(0,087)	.
num_mall_2	0,251	(0,159)	.	0,207	(0,189)		0,346	(0,251)	.
num_mall_3	0,217	(0,126)		0,213	(0,111)		0,225	(0,156)	
dummy_airport	1,065	(0,596)		1,590	(0,451)	*	-0,063	(0,879)	
dummy_bus_station	0,646	(0,223)	**	0,591	(0,373)	*	0,765	(0,262)	**
dummy_CBD	0,125	(0,087)	*	0,017	(0,093)		0,358	(0,177)	*
<i>Oferta de otros modos</i>									
stops_density	0,005	(0,002)		0,005	(0,002)	.	0,007	(0,004)	*
frec	-0,076	(0,024)	*	-0,061	(0,030)	*	-0,110	(0,036)	***
dummy_metro	0,132	(0,051)	***	0,103	(0,044)	.	0,195	(0,072)	**
cycleways_density	-0,012	(0,010)	**	-0,008	(0,009)		-0,019	(0,012)	
<i>Bondad de Ajuste</i>									
R <sup>2</sup>	0,833			0,812			0,770		
<i>Tests de Endogeneidad</i>									
Instrumentos Débiles	N/A			101,562			***		
Wu-Haussman	N/A			13,965			***		
Sargan	N/A			1,913			1,939		
<i>Correlación Espacial</i>									
Moran	0,342			***			0,328		
<i>Capacidad Predictiva</i>									
MAPE	14,8%			14,6%			18,5%		

Tabla C.2: Comparación de modelos para día laboral período PT, viajes/km<sup>2</sup>

Variable	OLS		2SLS (tiempo espera)		2SLS (velocidad)				
intercepto	-1,136	(0,604)	.	0,254	(0,676)	-2,842	(3,413)		
<i>Atributos de Ridesourcing</i>									
wait_time	-0,175	(0,033)	***	-0,375	(0,048)	***	0,070	(0,487)	
spatial_mean_speed	-0,059	(0,010)	***	-0,052	(0,012)	**	-0,068	(0,023)	***
<i>Sociodemografía</i>									
prop_med_income	0,489	(0,182)	**	0,406	(0,192)	*	0,591	(0,344)	*
prop_high_income	1,379	(0,306)	***	1,638	(0,338)	***	1,062	(0,770)	.
prop_car_1	-0,235	(0,209)	.	-0,283	(0,261)	.	-0,177	(0,342)	.
prop_car_2+	-0,452	(0,234)	.	-0,391	(0,294)	.	-0,526	(0,382)	.
<i>Entorno Construido</i>									
log_pop_density	0,578	(0,049)	***	0,506	(0,056)	***	0,667	(0,1822)	***
land_use_mix	1,140	(0,165)	***	1,098	(0,187)	***	1,191	(0,238)	***
num_mall_1	0,054	(0,071)	.	0,096	(0,125)	.	0,003	(0,140)	.
num_mall_2	0,558	(0,153)	***	0,553	(0,157)	***	0,565	(0,269)	***
num_mall_3	0,718	(0,122)	***	0,789	(0,103)	***	0,631	(0,243)	***
dummy_airport	1,779	(0,565)	.	2,407	(0,475)	.	1,009	(1,887)	.
dummy_bus_station	1,072	(0,216)	***	1,196	(0,281)	***	0,919	(0,528)	**
dummy_CBD	0,354	(0,082)	***	0,427	(0,123)	***	0,263	(0,194)	.
<i>Oferta de otros modos</i>									
stops_density	0,008	(0,002)	***	0,007	(0,003)	**	0,010	(0,007)	**
frec	-0,042	(0,025)	.	-0,034	(0,027)	.	-0,050	(0,041)	.
dummy_metro	0,295	(0,050)	***	0,275	(0,055)	***	0,320	(0,078)	***
cycleways_density	0,019	(0,010)	.	0,037	(0,015)	***	-0,003	(0,045)	.
<i>Bondad de Ajuste</i>									
R <sup>2</sup>	0,837		0,807		0,797				
<i>Tests de Endogeneidad</i>									
Instrumentos Débiles	N/A		209,936		***		1,699		
Wu-Haussman	N/A		57,893		***		0,440		
Sargan	N/A		1,458				0,584		
<i>Correlación Espacial</i>									
Moran	0,188		***		0,182		***		
<i>Capacidad Predictiva</i>									
MAPE	19,1 %		15,2 %		25,1 %				

# Anexo D

## Resultados viajes generados por cada 1.000 hogares

Tabla D.1: Resultados modelación viajes por cada 1.000 hogares período N2 (00:00 - 07:00)

Variable	LUN-JUE			VIE			SAB			DOM		
intercepto	2,359	(2,485)	.	7,632	(2,638)	***	6,501	(1,830)	***	9,829	(7,555)	***
<i>Atributos de Ridesourcing</i>												
wait_time	-0,267	(0,275)	*	-0,754	(0,256)	***	-0,777	(0,245)	***	-0,873	(0,715)	***
spatial_mean_speed	-0,026	(0,018)	*	-0,003	(0,016)		-0,001	(0,020)		-0,005	(0,029)	
<i>Sociodemografía</i>												
prop_med_income	1,082	(0,367)	***	0,389	(0,488)		-0,155	(0,439)		-0,153	(0,997)	
prop_high_income	1,223	(0,491)	*	1,409	(0,699)	*	0,933	(0,573)		-0,469	(1,185)	
prop_car_1	0,071	(0,502)		-0,616	(0,504)		0,768	(0,428)	.	-0,046	(0,660)	
prop_car_2+	-0,236	(0,408)		-0,018	(0,556)		-0,062	(0,521)		-0,186	(0,717)	
<i>Entorno Construido</i>												
log_pop_density	-0,016	(0,123)		-0,315	(0,145)	*	-0,128	(0,120)		-0,295	(0,302)	*
land_use_mix	1,265	(0,564)	***	0,742	(0,457)	.	0,720	(0,396)	*	0,100	(1,318)	
num_mall_1	-0,004	(0,099)		0,040	(0,117)		0,016	(0,103)		0,010	(0,124)	
num_mall_2	-0,037	(0,382)		-0,512	(0,433)		-0,711	(0,586)	*	-0,158	(0,508)	
num_mall_3	0,368	(0,220)	.	0,248	(0,217)		0,250	(0,192)		0,282	(0,197)	
dummy_airport	4,414	(0,894)	***	4,986	(1,188)	***	4,477	(1,025)	***	2,186	(1,169)	.
dummy_bus_station	1,144	(0,678)	**	0,951	(0,444)	.	0,911	(0,546)	*	0,776	(0,458)	
dummy_CBD	0,134	(0,152)		0,273	(0,192)		-0,114	(0,170)		-0,033	(0,223)	
<i>Oferta de otros modos</i>												
stops_density	0,005	(0,005)		0,004	(0,005)		-0,003	(0,006)		0,003	(0,007)	
frec	-0,606	(0,199)	***	-0,309	(0,220)		-0,466	(0,227)	*	-0,162	(0,327)	
dummy_metro	0,182	(0,128)	*	-0,025	(0,153)		0,048	(0,127)		0,202	(0,123)	.
cycleways_density	0,017	(0,016)		0,037	(0,020)	.	0,046	(0,019)	*	0,015	(0,022)	
<i>Bondad de Ajuste</i>												
R <sup>2</sup>	0,459			0,135			0,197			0,017		
<i>Tests de Endogeneidad</i>												
Instrumentos Débiles	37,518		***	15,490		***	26,320		***	9,887		***
Wu-Haussman	0,399			21,350		***	28,330		***	13,918		***
Sargan	22,846		***	10,350		**	18,800		***	51,058		***
<i>Correlación Espacial</i>												
Moran	0,346		***	0,149		***	0,095		***	0,207		***
<i>Capacidad Predictiva</i>												
MAPE	66,4%			249,1%			104,1%			347,4%		
<i>Num. Observaciones</i>												
N	336			332			331			330		

Tabla D.2: Resultados modelación viajes por cada 1.000 hogares período PM (07:00 - 10:00)

Variable	LUN-JUE			VIE			SAB			DOM		
intercepto	5,705	(0,777)	***	4,547	(0,851)	***	5,445	(1,253)	***	3,981	(1,508)	***
<i>Atributos de Ridesourcing</i>												
wait_time	-0,260	(0,047)	***	-0,217	(0,064)	***	-0,552	(0,111)	***	-0,478	(0,161)	***
spatial_mean_speed	-0,054	(0,009)	***	-0,025	(0,009)	***	-0,006	(0,007)		-0,001	(0,010)	
<i>Sociodemografía</i>												
prop_med_income	0,426	(0,224)	*	0,381	(0,238)	.	0,485	(0,300)	.	0,893	(0,356)	**
prop_high_income	1,247	(0,374)	***	1,517	(0,406)	***	-0,401	(0,448)		-0,422	(0,539)	
prop_car_1	-0,157	(0,206)		-0,179	(0,246)		0,127	(0,333)		-0,060	(0,430)	
prop_car_2+	-0,242	(0,294)		-0,385	(0,297)		-0,974	(0,377)	**	-0,718	(0,453)	.
<i>Entorno Construido</i>												
log_pop_density	-0,184	(0,060)	***	-0,159	(0,065)	**	-0,118	(0,092)	.	-0,100	(0,094)	
land_use_mix	0,370	(0,187)	*	0,567	(0,213)	**	0,302	(0,310)		0,493	(0,362)	
num_mall_1	-0,159	(0,068)	*	-0,240	(0,071)	**	-0,020	(0,095)		-0,037	(0,126)	
num_mall_2	0,268	(0,194)	.	0,106	(0,208)		-0,081	(0,259)		0,416	(0,278)	
num_mall_3	0,181	(0,116)		0,172	(0,095)		0,203	(0,147)		-0,048	(0,154)	
dummy_airport	2,271	(0,446)	***	3,285	(0,875)	***	4,316	(0,832)	***	3,116	(0,847)	**
dummy_bus_station	0,540	(0,401)	*	0,708	(0,355)	**	0,696	(0,437)	*	0,525	(0,346)	
dummy_CBD	0,074	(0,091)		0,130	(0,094)		-0,080	(0,132)		-0,316	(0,175)	.
<i>Oferta de otros modos</i>												
stops_density	0,005	(0,002)	.	0,005	(0,003)	.	-0,003	(0,004)		0,001	(0,004)	
frec	-0,073	(0,028)	**	-0,067	(0,036)	*	-0,090	(0,049)	.	-0,033	(0,095)	
dummy_metro	0,098	(0,047)	.	0,153	(0,058)	**	0,078	(0,063)		0,033	(0,093)	
cycleways_density	-0,021	(0,009)	*	-0,016	(0,011)		-0,009	(0,013)		0,016	(0,015)	
<i>Bondad de Ajuste</i>												
R <sup>2</sup>	0,518			0,372			0,199			0,112		
<i>Tests de Endogeneidad</i>												
Instrumentos Débiles	101,562		***	41,645		***	21,968		***	14,460		***
Wu-Haussman	8,812		**	10,235		**	25,093		***	21,990		***
Sargan	0,065			1,609			5,678		*	10,470		**
<i>Correlación Espacial</i>												
Moran	0,322		***	0,286		***	0,120		***	0,012		***
<i>Capacidad Predictiva</i>												
MAPE	37,1%			46,2%			174,3%			136,3%		
<i>Num. Observaciones</i>												
N	339			339			338			339		

Tabla D.3: Resultados modelación viajes por cada 1.000 hogares período FP (10:00 - 17:30)

Variable	LUN-JUE			VIE			SAB			DOM		
intercepto	8,913	(0,706)	***	7,871	(0,814)	***	7,052	(0,644)	***	7,084	(0,980)	***
<i>Atributos de Ridesourcing</i>												
wait_time	-0,431	(0,044)	***	-0,331	(0,062)	***	-0,401	(0,068)	***	-0,511	(0,075)	***
spatial_mean_speed	-0,034	(0,009)	***	-0,027	(0,009)	***	-0,018	(0,009)	*	-0,016	(0,009)	.
<i>Sociodemografía</i>												
prop_med_income	0,193	(0,166)		0,389	(0,198)	*	0,186	(0,171)		0,196	(0,223)	
prop_high_income	0,310	(0,277)		0,835	(0,326)	**	0,120	(0,274)		-0,225	(0,301)	
prop_car_1	-0,356	(0,249)	.	-0,345	(0,247)		-0,207	(0,275)		0,093	(0,248)	
prop_car_2+	-0,114	(0,249)		-0,366	(0,266)		-0,747	(0,263)	**	-0,582	(0,278)	*
<i>Entorno Construido</i>												
log_pop_density	-0,443	(0,055)	***	-0,402	(0,064)	***	-0,236	(0,049)	***	-0,219	(0,067)	***
land_use_mix	0,690	(0,171)	***	1,001	(0,169)	***	0,828	(0,196)	***	0,619	(0,208)	***
num_mall_1	0,155	(0,125)	*	0,115	(0,097)		0,087	(0,118)		-0,023	(0,095)	
num_mall_2	0,333	(0,174)	*	0,305	(0,163)	*	0,395	(0,192)	*	0,278	(0,148)	.
num_mall_3	0,571	(0,103)	***	0,552	(0,118)	***	0,377	(0,103)	**	0,291	(0,136)	*
dummy_airport	3,407	(0,376)	***	2,686	(0,446)	***	3,398	(0,514)	***	2,924	(0,514)	***
dummy_bus_station	0,741	(0,239)	***	0,674	(0,250)	**	0,518	(0,293)	*	0,552	(0,337)	*
dummy_CBD	0,235	(0,096)	**	0,309	(0,105)	**	0,190	(0,088)	*	-0,044	(0,079)	
<i>Oferta de otros modos</i>												
stops_density	0,007	(0,002)	**	0,007	(0,002)	**	0,004	(0,003)		0,006	(0,003)	*
frec	-0,042	(0,036)		-0,024	(0,037)		-0,085	(0,040)	*	-0,031	(0,048)	
dummy_metro	0,137	(0,046)	**	0,115	(0,054)	*	0,188	(0,051)	***	0,050	(0,049)	
cycleways_density	0,020	(0,011)	*	0,009	(0,011)		0,006	(0,010)		-0,006	(0,009)	
<i>Bondad de Ajuste</i>												
R <sup>2</sup>	0,663			0,607			0,523			0,513		
<i>Tests de Endogeneidad</i>												
Instrumentos Débiles	153,019			69,263			65,255			47,098		
Wu-Haussman	14,706			9,382			31,468			25,844		
Sargan	0,159			0,506			5,411			0,151		
<i>Correlación Espacial</i>												
Moran	0,173			0,166			0,150			0,130		
<i>Capacidad Predictiva</i>												
MAPE	10,9%			12,5%			13,4%			14,1%		
<i>Num. Observaciones</i>												
N	339			339			339			339		

Tabla D.4: Resultados modelación viajes por cada 1.000 hogares período PT (17:30 - 21:00)

Variable	LUN-JUE			VIE			SAB			DOM		
intercepto	8,261	(0,668)	***	7,248	(0,809)	***	7,109	(0,853)	***	6,869	(1,061)	***
<i>Atributos de Ridesourcing</i>												
wait_time	-0,342	(0,050)	***	-0,277	(0,061)	***	-0,396	(0,065)	***	-0,456	(0,065)	***
spatial_mean_speed	-0,039	(0,012)	***	-0,023	(0,009)	*	-0,012	(0,008)		-0,002	(0,010)	
<i>Sociodemografía</i>												
prop_med_income	0,248	(0,190)		0,288	(0,205)		-0,166	(0,220)		0,284	(0,205)	
prop_high_income	1,105	(0,334)	***	1,116	(0,333)	***	0,022	(0,310)		-0,129	(0,342)	
prop_car_1	-0,260	(0,258)		-0,083	(0,245)		0,190	(0,248)		-0,262	(0,245)	
prop_car_2+	-0,264	(0,279)		-0,486	(0,276)	.	-0,496	(0,314)	.	-0,344	(0,286)	
<i>Entorno Construido</i>												
log_pop_density	-0,508	(0,057)	***	-0,450	(0,066)	***	-0,342	(0,069)	***	-0,364	(0,081)	***
land_use_mix	0,942	(0,175)	***	1,119	(0,214)	***	0,613	(0,216)	**	0,400	(0,242)	.
num_mall_1	0,074	(0,14)		0,067	(0,107)		0,098	(0,101)		0,074	(0,092)	
num_mall_2	0,599	(0,155)	***	0,513	(0,186)	**	0,565	(0,192)	***	0,700	(0,239)	***
num_mall_3	0,748	(0,117)	***	0,617	(0,128)	***	0,695	(0,093)	***	0,915	(0,123)	***
dummy_airport	3,245	(0,481)	***	2,460	(0,594)	***	3,044	(0,630)	***	3,255	(0,522)	***
dummy_bus_station	1,089	(0,271)	***	0,962	(0,396)	***	0,737	(0,338)	**	1,008	(0,506)	***
dummy_CBD	0,438	(0,115)	***	0,403	(0,106)	***	0,244	(0,122)	**	0,027	(0,139)	
<i>Oferta de otros modos</i>												
stops_density	0,007	(0,003)	**	0,011	(0,003)	***	0,003	(0,003)		0,003	(0,002)	
frec	-0,043	(0,026)	.	-0,044	(0,026)	.	-0,067	(0,040)	.	0,002	(0,041)	
dummy_metro	0,267	(0,055)	***	0,274	(0,052)	***	0,159	(0,057)	**	0,065	(0,051)	
cycleways_density	0,023	(0,015)	*	0,005	(0,012)		0,011	(0,013)		0,010	(0,012)	
<i>Bondad de Ajuste</i>												
R <sup>2</sup>	0,657			0,573			0,494			0,4634		
<i>Tests de Endogeneidad</i>												
Instrumentos Débiles	209,936		***	92,538		***	85,297		***	59,447		***
Wu-Haussman	54,645		***	38,098		***	29,494		***	30,791		***
Sargan	1,635			0,161			0,938			4,658		*
<i>Correlación Espacial</i>												
Moran	0,159		***	0,096		***	0,210		***	0,142		***
<i>Capacidad Predictiva</i>												
MAPE	18,0 %			22,6 %			19,3 %			29,7 %		
<i>Num. Observaciones</i>												
N	339			339			339			339		

Tabla D.5: Resultados modelación viajes por cada 1.000 hogares período N1 (21:00 - 00:00)

Variable	LUN-JUE			VIE			SAB			DOM		
intercepto	5,754	(1,321)	***	7,092	(1,375)	***	9,001	(1,489)	***	7,516	(1,690)	***
<i>Atributos de Ridesourcing</i>												
wait_time	-0,709	(0,124)	***	-0,425	(0,112)	***	-0,643	(0,215)	***	-0,480	(0,104)	***
spatial_mean_speed	0,035	(0,034)	*	-0,017	(0,016)		-0,013	(0,026)		-0,027	(0,018)	**
<i>Sociodemografía</i>												
prop_med_income	0,477	(0,388)		0,926	(0,345)	**	0,291	(0,464)		-0,069	(0,376)	
prop_high_income	0,631	(0,455)		1,382	(0,433)	**	-0,517	(0,580)		-1,375	(0,578)	*
prop_car_1	0,443	(0,348)		-0,166	(0,302)		0,107	(0,369)		0,311	(0,366)	
prop_car_2+	0,197	(0,433)		-0,533	(0,351)		-0,594	(0,417)		-0,019	(0,544)	
<i>Entorno Construido</i>												
log_pop_density	-0,186	(0,100)	*	-0,339	(0,100)	***	-0,311	(0,103)	***	-0,227	(0,126)	*
land_use_mix	0,569	(0,294)	.	0,680	(0,281)	*	0,287	(0,372)		0,407	(0,376)	
num_mall_1	0,099	(0,134)		0,083	(0,096)		0,200	(0,130)		0,111	(0,131)	
num_mall_2	0,200	(0,448)		-0,049	(0,341)		0,364	(0,251)		-0,014	(0,438)	
num_mall_3	0,861	(0,170)	***	0,657	(0,109)	***	0,632	(0,141)	**	0,849	(0,131)	***
dummy_airport	2,959	(0,902)	**	1,198	(0,680)		-1,099	(1,242)		2,900	(0,901)	**
dummy_bus_station	1,082	(0,599)	**	1,391	(0,585)	***	1,091	(0,390)	**	1,476	(0,694)	***
dummy_CBD	0,172	(0,096)		0,258	(0,096)	*	-0,067	(0,152)		-0,141	(0,138)	
<i>Oferta de otros modos</i>												
stops_density	0,002	(0,004)		0,011	(0,004)	**	0,015	(0,005)	**	0,008	(0,005)	.
frec	-0,435	(0,080)	***	-0,319	(0,095)	***	-0,309	(0,140)	***	-0,333	(0,128)	***
dummy_metro	0,216	(0,083)	*	0,237	(0,073)	**	0,039	(0,090)		0,167	(0,087)	
cycleways_density	0,033	(0,014)	*	0,014	(0,014)		-0,001	(0,020)		0,005	(0,018)	.
<i>Bondad de Ajuste</i>												
R <sup>2</sup>	0,439			0,416			0,087			0,195		
<i>Tests de Endogeneidad</i>												
Instrumentos Débiles	109,410		***	58,730		***	25,542		***	53,736		***
Wu-Haussman	19,320		***	14,190		***	43,386		***	35,663		***
Sargan	17,280		***	5,860		*	9,966		**	1,446		
<i>Correlación Espacial</i>												
Moran	0,344		***	0,181		***	0,151		***	0,190		***
<i>Capacidad Predictiva</i>												
MAPE	62,8%			43,8%			58,1%			65,5%		
<i>Num. Observaciones</i>												
N	336			333			335			337		