



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

“FORMULACIÓN Y ANÁLISIS DE MODELOS DE DEMANDA AGREGADA
DE VALIDACIONES Y VIAJES DE TRANSANTIAGO”

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL INDUSTRIAL

JUAN ANDRÉS CAVADA HERRERA

PROFESOR GUÍA:
ALEJANDRO TIRACHINI HERNÁNDEZ

MIEMBROS DE LA COMISIÓN:
RICARDO SAN MARTÍN ZURITA
RODOLFO URRUTIA URIBE

SANTIAGO DE CHILE
2014

RESUMEN DE LA MEMORIA PARA OPTAR AL
TITULO DE: Ingeniero Civil Industrial
POR: Juan Andrés Cavada Herrera
FECHA: Marzo 2014
PROFESOR GUÍA: Alejandro Tirachini Hernández

Formulación y análisis de modelos de demanda agregada de Validaciones y Viajes de Transantiago

El presente trabajo muestra la formulación y análisis de los primeros modelos de demanda agregada tanto para validaciones como para viajes realizados en el sistema de transporte público de la ciudad de Santiago. Puesto en marcha en el año 2007, Transantiago es un sistema integrado de transporte público que involucra tanto a buses como a Metro, bajo un sistema tarifario único a través de la tarjeta Bip!

Se formularon diversos modelos tanto para validaciones como para viajes. La importancia de estos modelos radica en que son herramientas necesarias tanto para entender los impactos que cambios en las variables o distintas políticas puedan tener sobre la demanda, así como poder realizar proyecciones financieras con un menor nivel de incertidumbre, fundamental para poder tomar decisiones.

Para la formulación de los modelos, se utilizaron regresiones lineales multivariadas, probando y comparando distintas formas funcionales con el fin de obtener modelos correctamente especificados. Las formas funcionales estimadas son lineal, lineal-log, log-lineal y log-log. Los datos utilizados son agregados y de carácter mensual, desde Junio 2007 hasta Diciembre 2012 para los modelos de validaciones, y desde enero 2009 hasta diciembre 2012 para el modelo de viajes.

Luego de comprobar que se cumplieran todos los supuestos necesarios para un correcto modelo de regresión lineal multivariada, se procedió a utilizar los datos del año 2013 con el fin de validar los modelos. Para el modelo de validaciones y para el modelo de viajes el pronóstico del año 2013 presentó errores porcentuales absolutos de 2.38% y 2.95% respectivamente.

Importantes relaciones fueron estimadas durante este estudio, como por ejemplo el la mayor elasticidad precio que presenta el bus con respecto al metro, además del impacto de otras variables relevantes como los coche kilómetros ofertados por metro, la frecuencia promedio de los buses y la longitud de la red de buses.

Tanto el modelo de validaciones como el de viajes resultaron entregar buenos resultados en el pronóstico a corto y mediano plazo, por lo que se cree que pueden ser herramientas fundamentales para el análisis y para realizar proyecciones sobre el sistema de transporte público de Santiago.

Agradecimientos

Se agradece el apoyo y colaboración de Antonio Gschwender y Transantiago, así como a Daniel Schwarz y Metro por su constante ayuda y colaboración con la memoria.

Agradecimientos también a Marcela Munizaga, Ricardo San Martín y Rodolfo Urrutia por sus comentarios y revisiones.

Especial agradecimiento a Alejandro Tirachini, quien logró que me interesara por el mundo del transporte, por su gran dedicación y preocupación durante todo el desarrollo de la memoria, así como también por su colaboración, comentarios y guía en este proceso.

Esta memoria fue financiada por Fondecyt (proyecto 11130227).

Tabla de Contenido

1.	Introducción	1
1.1	Antecedentes Generales	1
1.2	Justificación	3
1.3	Objetivos	5
1.3.1	Objetivo General	5
1.3.2	Objetivos Específicos	5
1.4	Estructura	5
2.	Revisión Bibliográfica	6
2.1	Modelos de demanda en Santiago.....	6
2.2	Modelos de Viajes de otros países y comparación de elasticidades.	7
3.	Descripción de los datos.....	11
4.	Estimación Modelos	24
4.1	Supuestos del modelo de regresión lineal	24
4.1.1	Linealidad	24
4.1.2	Independencia.....	24
4.1.3	Homocedasticidad.....	24
4.1.4	Normalidad.....	25
4.1.5	No Colinealidad	25
4.2	Formas funcionales	25
4.3	Estadísticos descriptivos	28
4.4	Modelos de Validaciones	28
4.4.1	Matriz de Covarianzas	29
4.4.2	Resultados de los Modelos.....	29
4.4.3	Modelo Agregado de Validaciones.....	30
4.4.4	Modelo de Validaciones de Buses.....	35
4.4.5	Modelo de Validaciones de Metro.....	39
4.4.6	Comparación modelos Lineales.....	43
4.4.7	Comparación modelos normalizados y sin normalizar	44
4.4.8	Comparación de formas funcionales para modelo de validaciones	46
4.4.9	Modelo Agregado Log-Log de Validaciones	48

4.4.10	Modelo log-log de Validaciones de Buses.....	53
4.4.11	Modelo log-log de validaciones de Metro	54
4.4.12	Elasticidades	55
4.4.13	Resultados	56
4.5	Modelo de Viajes.....	57
4.5.1	Matriz de Covarianzas	57
4.5.2	Resultado del Modelo de Viajes.....	58
5.	Pronóstico y modelación distintos escenarios.....	62
5.1	Predicción modelo de Transacciones.....	62
5.1.1	Predicción con datos utilizados para estimación: 2007-2012.....	62
5.1.2	Predicción con datos no utilizados para estimación: 2013	63
5.1.3	Escenarios y pronóstico a 3 años	64
5.1.4	Aumento de la red de Metro.....	67
5.2	Predicción Modelo Viajes.....	69
5.2.1	Predicción con datos utilizados para estimación: 2009-2012.....	69
5.2.2	Predicción con datos no utilizados para estimación: 2013	70
6.	Conclusiones.....	73
	Bibliografía	76

Índice de Tablas

Tabla 1 – Comparación autores modelos de demanda agregada.....	7
Tabla 2 – Estadísticos Descriptivos modelo Validaciones	28
Tabla 3 – Matriz de Covarianzas.....	29
Tabla 4 . Modelo Agregado de Validaciones.....	31
Tabla 5 – Modelo de Validaciones de Buses	36
Tabla 6 – Modelo Validaciones Metro	40
Tabla 7 – Comparación modelos lineales de validaciones	43
Tabla 8 – Comparación modelos agregado de validaciones Datos normalizado y sin normalizar	44
Tabla 9 – Comparación modelo validaciones de buses datos no normalizado y sin normalizar	45
Tabla 10 – Comparación modelo validaciones de Metro datos normalizados y sin normalizar.....	46
Tabla 11 – Comparación Formas Funcionales modelo agregado de Validaciones	47
Tabla 12 – Modelo agregado (Bus + metro) Log-Log de Validaciones.....	49
Tabla 13 – Modelo log-log validaciones (Km Comerciales).....	52
Tabla 14 – Modelo log-log de validaciones de buses.....	53
Tabla 15 – Modelo log-log de validaciones de buses (Km Comerciales).....	54
Tabla 16 – Modelo log log de validaciones de Metro	54

Tabla 17 – Matriz de Covarianza Modelo de Viajes	57
Tabla 18 – Modelos lineales anidados de Viajes Transantiago	59
Tabla 19 – Cuadro comparativo resultados de modelos anidados	59
Tabla 20 – Modelos log-log anidados de Viajes Transantiago	61
Tabla 21 – Resultados Validación modelo datos 2013.....	64
Tabla 22 – Resumen Resultados distintos escenarios Tarifa	65
Tabla 23 – Resumen resultados distintos escenarios Longitud Red Buses.....	66
Tabla 24 – Impacto Coche Km Metro en modelo de transacciones de buses	67
Tabla 25 - Impacto Coche Km Metro en modelo de transacciones de buses (sin tarifa).....	68
Tabla 26 –Impacto Coche Km Metro en modelo agregado de transacciones (sin tarifa).....	69
Tabla 27 - MAE y MAPE modelos de Viajes.....	70
Tabla 28 – Validación modelo de Viajes datos 2013.....	71
Tabla 29- MAE y MAPE modelos de viajes (2013).....	72

Índice de Figuras

Figura 1 – Ingreso, Costos y resultado operacional Dic.2012	3
Figura 2 – Proyección Ingresos y subsidios Transantiago	3
Figura 3 – Viajes	11
Figura 4 – Validaciones	11
Figura 5 – Tasa de Desempleo Santiago.....	12
Figura 6 – Índice General de Remuneraciones	13
Figura 7 – Precipitaciones	13
Figura 8 – Tarifas	14
Figura 9 – Precio Gasolina	15
Figura 10 – Longitud Red Buses	16
Figura 11 – Longitud Red Metro	16
Figura 12 – Kilómetros Comerciales	17
Figura 13 – Coche Kilómetro Metro	18
Figura 14 – ICR e ICF	18
Figura 15 – Frecuencia Programada Buses	19
Figura 16 – Frecuencia Programada Servicio Sentido(SS)	19
Figura 17 – Frecuencia real vs Frecuencia programada en horario Punta Mañana	20
Figura 18 – Frecuencia promedio SS real vs Frecuencia promedio programada SS	20
Figura 19 – Índice Automotor	21
Figura 20 – Evasión	22
Figura 21 – Gráficos de Regresiones Parciales modelo agregado de validaciones.....	33
Figura 22 – Histograma y gráfico P-P normal modelo agregado.....	33
Figura 23 - Gráfico de dispersión modelo agregado.....	34
Figura 24 – Gráficos de Regresiones Parciales modelo de Validaciones de Buses	37
Figura 25 – Histograma y Gráfico P-P normal modelo de buses.....	38
Figura 26 – Gráfico de Dispersión modelo de validaciones de Buses.....	38

Figura 27 – Gráficos de regresiones parciales modelo de validaciones Metro.....	41
Figura 28 – Histograma y Gráfico P-P normal modelo validaciones Metro	42
Figura 29 – Gráfico de Dispersión modelo de Validaciones Metro.....	42
Figura 30 – Gráficos de regresiones parciales modelo logarítmico de Validaciones	50
Figura 31 – Histograma y gráfico P-P normal modelo logarítmico de Validaciones	51
Figura 32 – Gráfico de Dispersión modelo logarítmico de Validaciones.....	51
Figura 33 – Histograma y Gráfico P-P normal modelo de Viajes	60
Figura 34 – Gráfico de dispersión modelo de Viajes	60
Figura 35 – Validaciones Real vs Validaciones Estimada (2007-2009).....	62
Figura 36 - Validaciones Real vs Validaciones Estimada (2010-2012).....	62
Figura 37 – Validación modelo validaciones con datos 2013	63
Figura 38 – Demanda Real vs Demanda estimada modelo de Viajes	70
Figura 39 – Validación modelo de viajes con datos 2013	71

1. Introducción

Esta memoria estudia la demanda de transporte público de la ciudad de Santiago incorporando una serie de variables de calidad de servicio, tarifas, estacionalidades, alcance espacial de la red entre otras, con el fin de poder entender el impacto que estas tienen tanto sobre la cantidad de validaciones como sobre la cantidad de viajes realizadas en Transantiago.

En la sección 1.1 se describen los antecedentes generales del sistema de transporte de Santiago. Luego en 1.2 se explica la justificación de realizar este estudio, denotando algunas de las aplicaciones prácticas que se podrán llevar a cabo con los resultados de la memoria. En la sección 1.3 se muestran los objetivos de la memoria, finalizando el capítulo con la sección 1.4 en donde se introduce la estructura del trabajo realizado.

1.1 Antecedentes Generales

La evolución del transporte público en Santiago ha sido importante y necesaria. Como indican Muñoz y Gschwender(2008), el sistema de buses fue votado como el peor servicio público de la ciudad en el año 2003. Buses en precarias condiciones, operadores no profesionales, competición peligrosa entre buses por llevar pasajeros, integración inexistente con servicio de Metro, rutas ineficientes y altas tasas de accidentes eran la tónica del sistema.

Es por este motivo que se decidió crear un sistema de transporte público capaz de revertir dicha situación, idea anunciada en febrero del año 2000. Después de seis años de planificación el proyecto fue implementado.

A grandes rasgos con la implementación de este nuevo sistema, se buscaba crear una industria formal, en donde los operadores tengan ingresos garantizados en un alto porcentaje y dejen de ser variables (con el fin de incentivar la competencia entre operadores al momento de licitar). Se buscaba además una fuerte regulación por parte del estado (antes existía únicamente una regulación tarifaria), y que exista una fiscalización en la oferta. Además de esto se buscaron ciertos estándares de calidad y por sobre todo, una integración tarifaria. (Simonetti, 2012),

Evidencias sobre el impacto de la implementación de un modelo con integración tarifaria en otros lugares, ha mostrado cambios importantes y ha sido el sistema favorecido en distintos países, como Alemania, Suiza y Austria (Pucher y Kurth, 1995), y posteriormente España (Matas, 2004).

Este nuevo sistema, compuesto de una red de buses alimentadores y troncales, en donde el Metro juega un papel principal como columna vertebral de los recorridos, fue catalogado por Hidalgo y Graftieaux (2007), como el plan de modernización de transporte más ambicioso de la última década en una ciudad grande. La elección de un sistema troncales-alimentadores, entre otras ventajas, permite una reducción en la cantidad de buses-km y tamaño de flota, conllevando a su vez a una reducción en los costos operacionales y una serie de externalidades positivas, como lo son la reducción en la polución (junto a estándares mínimos de los nuevos buses pedido por el Gobierno), disminución en la cantidad de ruido y de accidentes (Gschwender, 2005).

Sin embargo Transantiago ha vivido una serie de problemas, comenzando con su puesta en marcha, en donde se dieron de baja los buses anteriores que significó una disminución importante en el número de buses que transitaba por la ciudad, siendo estos buses insuficientes para satisfacer la demanda. El diseño de los contratos de los operadores de Transantiago presentaron problemas serios también, ya que permitían irregularidades en el servicio sin que estas fuesen sancionadas. Ambos problemas desencadenaron en grandes tiempos de espera por parte de los usuarios, además de elevados niveles de hacinamiento tanto en los buses como en el metro.

Hoy en día el sistema de transporte público de la ciudad de Santiago está compuesto por los buses de transporte público urbano de la ciudad, operado por 7 empresas privadas; y el Metro de Santiago. Tanto buses como Metro está integrados tarifariamente mediante un único medio de acceso electrónico, la tarjeta Bip!

Mensualmente se realizan sobre 150 millones de validaciones, lo que lleva a un promedio sobre las 3.200 validaciones por minuto. Desde la puesta en marcha en Febrero 2007, este sistema de transporte público ha tenido problemas y detractores, así como también una serie de medidas que han aportado un cambio positivo, buscando entre otras cosas mejorar la calidad de servicio entregada a los usuarios, estabilizar las condiciones financieras del sistema, contar con empresas sustentables en el largo plazo y potenciar el transporte público de forma de ayudar a la descongestión de la ciudad de Santiago.

Dentro de estas medidas, se encuentran el proceso de rediseño en el esquema regulatorio del Sistema, otorgándole un mayor rol regulador al Ministerio de Transportes y Telecomunicaciones (MTT), en desmedro de carácter supra-integrador que tenía el Administrador Financiero de Transantiago (AFT), de quien dependían los servicios de gestión de pago y gestión de flota, y la operación de sistemas centrales que procesen la información y gestionen la seguridad del sistema. (Simonetti, 2012)

La creación de nuevos contratos con los operadores de buses, con el fin de incentivar una mejora en la calidad de servicio por parte de los operadores, dar mayor protagonismo a las empresas en el desarrollo de recorridos y tener un mayor control de la evasión. Además hoy se cuenta con información al usuario en tiempo de llegadas de buses a paraderos en tiempo real.

Importante es mencionar que en estos nuevos contratos, se estipula que el ingreso del operador pasa de ser garantizado en una alta proporción a estar definido en base a los servicios efectivamente prestados, función de los pasajeros transportados y kilómetros recorridos. (Simonetti, 2013)

Estas medidas han permitido una mejora en algunos de los índices de calidad de servicio y en los resultados financieros. (Figura 1).

Ingresos, costos y resultado operacional (MM\$ Dic. 2012)

Mes	Ingresos	Egresos	Resultado antes de subsidio
2009	396.410	778.814	-382.403
2010	466.935	840.325	-373.390
2011	513.688	882.070	-368.382
2012	529.823	864.540	-334.717

Figura 1 – Ingreso, Costos y resultado operacional Dic.2012 Fuente: Simotti 2012

En términos financieros, Transantiago ha contado con un subsidio permanente del Estado, y uno transitorio que decae en el tiempo. Proyecciones de comienzos del 2011 (Figura 2) muestran la evolución tanto de los ingresos esperados, los subsidios y la brecha que se produce debido a la disminución del subsidio transitorio, la que debería ser cubierta con un alza en las tarifas. En este contexto en el mes de Marzo de 2010, se creó el Panel de Expertos Tarifarios, encargados de revisar mes a mes las tarifas del Sistema. La contraparte de ellos es la Gerencia de Finanzas Corporativas del Transantiago.

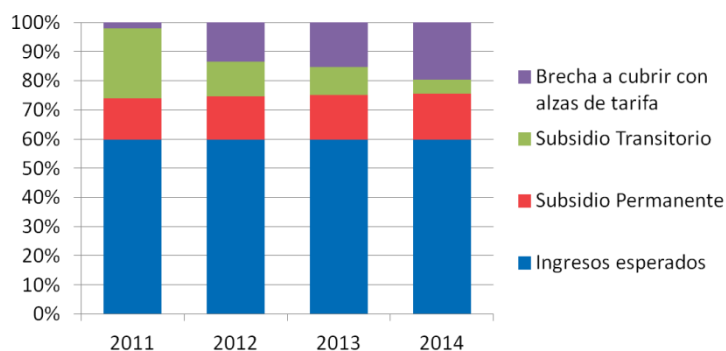


Figura 2 – Proyección Ingresos y subsidios Transantiago Fuente: Simonetti 2012

Un aspecto fundamental para proyectar el sistema hacia el futuro es contar con herramientas de estimación de demanda. Sin embargo actualmente no existe ningún modelo de estimación de demanda de viajes, por lo que para la preparación de programas de operación en Transantiago se realiza el supuesto que la demanda se mantiene constante.

1.2 Justificación

La justificación para la realización de estos modelos se puede separar en varias áreas distintas.

La razón de utilizar un método cuantitativo como los modelos estadísticos que se muestran en este trabajo, es respuesta a tanto a la importancia que tiene el transporte público en una ciudad como Santiago, como a la naturaleza estocástica de los problemas de transporte. Esta naturaleza estocástica del transporte es a grandes rasgos resultado del rol que la gente juega en el transporte público, ya que día a día deben tomar un gran número de decisiones que dependerán de un gran número de posibilidades. Es por esto que se necesita de un modelo econométrico capaz de estimar y predecir en base a distintos factores o variables que tienen un impacto en la decisión de cada individuo a la hora de utilizar el transporte público.

Ahora bien, dentro de los involucrados principales y las ventajas de poder estimar estos modelos, se pueden mencionar la utilidad del modelo de demanda agregada de viajes para la Gerencia de Finanzas Corporativas. En conjunto con los nuevos contratos firmados con los operados el año 2012, en donde la cantidad de validaciones son utilizadas para cuantificar los pagos a los operadores, una mejora en las herramientas de proyección permitiría poder realizar proyecciones financieras del sistema de transporte público, así como de cada unidad de negocio.

Dicho modelo resulta además de vital importancia para la toma de decisiones de la Autoridad, en materias de tarifa y eficiencia del uso de los recursos de Transantiago.

Actualmente Transantiago no cuenta con un modelo de demanda agregada. En este escenario, se utiliza el supuesto de que la demanda se mantiene constante¹, por lo que muchas veces no son consideradas ciertas interacciones o impacto que pueden causar una variación en ciertas variables, situación que puede llevar a una toma de decisiones incorrecta o con información faltante. Por ejemplo no se ha estimado el impacto neto que tiene las extensiones de Metro en el número de validaciones tanto en Metro como buses, análisis que se pueden hacer con los modelos de esta memoria.

Un segundo beneficiado de los modelos es la Gerencia de Planificación de Servicios. Ellos son los encargados de la gestión de los cambios en el Sistema en relación con los servicios de buses, en particular están a cargo del proceso de cambio en los Programas de Operación. Si bien ellos utilizan como supuesto que las matrices origen – destino de viajes (O/D) se mantendrán constantes, el saber cómo variará la demanda al cambiar aspectos como el largo de la red, el número de servicios o las frecuencias puede constituir una herramienta de gran utilidad tomando en cuenta una planificación anual a nivel macro.

Además el trabajo realizado dejará una importante cantidad de información recolectada y analizada, que junto con los resultados y análisis de elasticidades pueden colaborar para la creación o cambios en políticas de transporte (Dergay y Hanly, 2002) y estas mismas políticas, en conjunto con proyecciones financieras estimar el impacto que podrían causar subsidios u otras decisiones a tomar en pos de aumentar el uso de transporte público.

¹ Gerencia de Finanzas Corporativas de Transantiago, comunicación personal.

1.3 Objetivos

1.3.1 Objetivo General

El objetivo general es la estimación de modelos de demanda, tanto de viajes como de validaciones, para entender cómo se comportan variables relevantes y el impacto que tienen, de manera de que puedan ser traducidas en decisiones y proyecciones que ayuden a una correcta planificación financiera y operacional de Transantiago

1.3.2 Objetivos Específicos

- La formulación de un modelo de demanda agregada de viajes en el sistema de transporte público de Santiago.
- Análisis e interpretación de elasticidades e interacciones entre las variables explicativas del modelo de demanda agregada de viajes.
- La formulación de un modelo de agregado de validaciones.
- La formulación de un modelo de validaciones de buses.
- La formulación de un modelo de validaciones de Metro
- Análisis e interpretación de elasticidades e interacciones entre las variables explicativas de los modelos de validaciones.
- Un análisis predictivo en un plazo de 3 años para estudiar distintos escenarios, tanto en modelo de validaciones como en modelo de viajes.

1.4 Estructura

Este trabajo se encuentra estructurado en seis capítulos: Introducción; Revisión bibliográfica; Descripción de Datos; Estimación de modelos; Validación, Pronóstico y modelación distintos escenarios; y finalmente Conclusiones.

En el capítulo siguiente se revisa trabajos realizados por diversos autores en cuanto a modelación de demanda de transporte público en distintos escenarios, así como una revisión de los modelos de demanda realizados en Santiago.

El tercer capítulo muestra los datos recolectados y seleccionados a partir de los resultados del Capítulo 2, su fuente y una explicación de los mismos.

En el Capítulo 4, Estimación de modelos, se muestra los supuestos que se deben cumplir para una correcta formulación de un modelo de regresión lineal multivariada, y posterior a esto su aplicación para los modelos de validaciones y viajes. Se estudian además distintas transformaciones en cuanto a la forma funcional de los modelos, con el fin de seleccionar finalmente el modelo que mejor represente la variable explicada.

En el Capítulo 5 se validan los modelos encontrados en el Capítulo 4 con información del año 2013. Y posterior a eso se revisan distintos escenarios en un horizonte de 3

años. En este mismo capítulo se estudia el impacto que ha causado la extensión de las líneas 1 y 5 de Metro en el sistema, de manera de tener alguna luz de lo que se puede esperar con la apertura de las nuevas líneas de Metro.

Cierra este trabajo el Capítulo 6, en donde se hablará de los principales resultados y conclusiones obtenidas.

2. Revisión Bibliográfica

La revisión bibliográfica de la memoria se dividirá en dos secciones. La primera discutirá los trabajos que se han realizado en la ciudad de Santiago con respecto a la estimación de demanda en transporte público. La importancia de esto radica en poder observar distintas variables y metodologías que ya se han utilizado y han sido significativas, dado que si bien el sistema de transporte público en la ciudad ha cambiado de manera significativa, es interesante analizar la evolución no solo de variables del sistema de transporte, sino que también variables exógenas como la estacionalidad y cambios en la economía como los niveles de desempleo.

En la segunda sección se discutirán distintos trabajos de que han utilizados modelos econométricos para estudiar la demanda agregada y obtener elasticidades de distinto tipo. Se realizará una comparación de los modelos observando que variables han utilizado, el tipo de data utilizada y los principales resultados de los mismos.

2.1 Modelos de demanda en Santiago

Los modelos de demanda de transporte público se Santiago, encontrados en la literatura, son previos a la puesta en marcha de Transantiago. Jara-Díaz y Paredes, (1993) realizaron un modelo de estimación de demanda agregada para el Metro. Su modelo buscaba capturar los efectos de generación de viajes, a través del nivel de empleo y un efecto combinado con las precipitaciones mensuales debido al efecto de disminución de viajes de carácter más discrecional; alcance espacial de la red, medido como número de estaciones; y competencia modal de auto y bus, midiendo tanto el costo como la disponibilidad (Número de permisos de circulación, costo de la bencina, tarifa promedio metro y tarifa promedio bus). En este trabajo no se incluyeron variables frecuencia ofrecida, tiempos de viaje o calidad de servicio, pero sí incluye un rezago del mes anterior de manera de poder medir tendencia en la demanda.

Un nuevo trabajo (Jara-Díaz, 1995), busca mejorar el modelo propuesto anteriormente, eliminando la variable de rezago que explicaba un porcentaje muy grande de la demanda total. Los nuevos resultados confirmaron la utilidad del enfoque multivariado para entender el rol de cada variable bajo condiciones ceteris paribus.

En ambos modelos, se estudia únicamente la demanda de Metro, y no de Buses. Al momento de realizar estos trabajos, Metro y buses eran competencia, situación distinta a lo que se presenta hoy en día, en donde ambos pertenecen al mismo sistema de transporte y son complementarios y sustitutos al mismo tiempo.

Dentro de los resultados en dichos estudios, existen análisis desagregados por líneas de Metro (1 y 2), obteniéndose elasticidades precio de -0.15 y -0.14 respectivamente, y al analizar elasticidades precio cruzadas, concluye que los buses son alternativa a la

línea de Metro únicamente en la línea 2, mientras que los taxi-colectivo son la alternativa de transporte al Metro en la línea 1.

En este mismo trabajo se evidencian algunos principales problemas de utilizar variables de rezago para explicar la demanda, debido a que si bien puede llevar a mejores ajustes, explican gran parte de la demanda minimizando el impacto que tienen las distintas variables estudiadas.

2.2 Modelos de Viajes de otros países y comparación de elasticidades.

La literatura ofrece una gran cantidad de modelos de estimación de demanda de transporte público, para distintas ciudades, con distintos enfoques, horizontes de tiempo, variables y metodologías. En la Tabla 1 se muestra una comparación de los principales modelos encontrados, resaltando sus resultados y contribuciones.

Tabla 1 – Comparación autores modelos de demanda agregada

Autor	Modo	Área	Data	Variables Explicativas
Bresson et al (2003)	Buses	Inglaterra	Agregada anual	Tarifa Ingreso Vehículo-km
	Transporte Público	Francia		
Matas (2004)	Buses y Metro	España	Agregada anual	Tarifa Costo alternativas Longitud Red Población Nivel de suburbanización Empleo Ingreso
Dergay y Hanly (2002)	Buses	Reino Unido	Agregada anual	Tarifa Costo alternativas Ingreso Vehículo-km Población Densidad
García-Ferrer (2006)	Metro y Bus	España	Agregado Mensual	Tarifa Vehículo-km
Gaudry (1975)	Buses	Canadá	Agregado Mensual	Tarifa Parque Automotriz Índice de precio del consumidor Ingresos Clima Empleo
Velázquez (2009)	Buses	Colombia	Agregada anual	Tarifa Costo alternativas Estaciones Población

La Tabla 1 muestra el aporte de seis trabajos que realizaron modelos de demanda de transporte público. La mayoría de los casos son en Europa (5 de los 7 mostrados), siendo los restantes de América del Norte y América del Sur.

En Bresson et al (2003), se hace hincapié en que la mayoría de los modelos comúnmente utilizados usan la forma funcional de log-log. El supuesto subyacente de que las elasticidades son constantes para todos los niveles de variables dependientes y explicativas puede ser cuestionable. Acá se estudió el modelo en Inglaterra semi-log, en donde la elasticidad precio es proporcional al nivel de la tarifa obteniendo resultado ligeramente mejores que en el modelo log-log (en cuanto a ajuste), sugiriendo que la elasticidad precio aumenta con el nivel de tarifa. En cuanto para su modelo de Francia estos resultados no se replicaron, ya que los modelos log-log presentaron un mejor ajuste en comparación al modelo semi-log.

La elasticidad precio obtenidas en este estudio, fueron de -0.32 para el modelo de Francia, y -0.51 para Inglaterra. Dicha diferencia es posible explicarla por la composición de la muestra (Buses, tren y Metro versus sólo buses) junto con que en Francia sólo se obtuvo datos de zonas urbanas, mientras que Inglaterra fueron zonas urbanas y rurales.

La mayor conclusión del estudio, es que la demanda de transporte público es relativamente sensible a los cambios en las tarifas, por lo que políticas destinadas a la reducción de la misma (subsidios), pueden jugar un rol sustancial en incentivar el uso de transporte público en desmedro del uso de autos. Mismo efecto es posible observar a través de políticas públicas que incentiven el uso del transporte público bajo un sistema integrado de transporte y de alta calidad (Pucher y Kurth, 1995).

En Matas (2004), la autora estima ecuaciones para dos modelos diferentes, uno de Metro y otro de buses en la ciudad de Madrid, permitiendo el efecto de sustitución entre ambos modos. Las variables que definen los distintos modelos de transporte son principalmente el nivel de servicio y los costos. Sin embargo en el caso de los costos ocurre un fenómeno de que las tarifas de bus y Metro han crecido de forma muy similar desde 1981 hasta 1989, y desde 1990 en adelante han sido exactamente las mismas, por lo que presentan una correlación de 0.97 lo que impide la inclusión de ambas variables en el modelo.

Para medir calidad de servicio, tanto frecuencia como velocidad dieron no significativas para los modelos, motivo por el cual se utilizó el alcance de la red. Por otro lado las variables de Ingreso, Población, nivel de suburbanización y empleo presentan correlaciones en el rango de 0.895 y 0.984.

Las elasticidades cruzadas entre longitud de red Metro y viajes en bus fue de -0.21, por lo que el aumento en la red de Metro está correlacionado negativamente con la demanda de buses. En el caso de la longitud de red de bus, esta no tiene un impacto medible en la cantidad de viajes realizados en Metro.

También en relación a Madrid, García-Ferrer et al (2006), proponen un modelo en donde a diferencia de Matas, utiliza data mensual en vez de anual. Los autores dicen

que utilizar datos anuales o semestrales, puede conllevar a un gran número de problemas econométricos, tales como problemas de especificación, pequeños tamaños muestrales, gran número de regresores que a su vez se traduce en un número pequeño de grados de libertad y problemas de multicolinealidad, como lo sucedido con Matas, 2004.

Los mismos autores señalan sobre el pronóstico de la demanda o Forecast, que sus variables pueden ser divididas en 2 componentes principales. Aquellas relacionadas con cambios demográficos y económicos, así como otras variables externas, y aquellas relacionadas directamente con el sistema de transporte público.

Cuando se busca planificar a largo plazo, ambas componentes deben ser predichas ya que ambas involucran supuestos económicos sobre su comportamiento futuro. Sin embargo con la utilización de datos de carácter mensual, y cuando el pronóstico a mediano y corto plazo es el objetivo, se puede asumir que aquellas variables externas al transporte público, están incorporados en la tendencia estocástica. De este modo, son únicamente aquellas variables propias del sistema de transporte las responsables de los cambios importantes en la demanda de transporte público.

Al igual que García-Ferrer, Gaudry(1975) utilizó datos mensuales y agregados, y ambos a diferencia de aquellos que utilizaban data anual, ingresaron como variable explicativa la cantidad de días hábiles en el mes, y Gaudry además ingresó la cantidad de sábados, así como de domingos y festivos.

Ahora bien, con respectos a las elasticidades, existe una amplia literatura que abarca más que estos 6 autores elegidos por ser ejemplos prácticos de estudios realizados en distintas ciudades. Excelente material es posible ver en Balcombe et al (2004), en donde se realiza un resumen de diversas elasticidades, como lo son de tarifa para bus, tren, cómo varía esta misma según naturaleza del viaje, tipo de usuario entre otras. Junto con éste se muestra una compilación de estudios que señalan los impactos de calidad de servicio, tales como tiempo de viaje, intervalos o el ingreso, uso de auto, de suelo, precio de combustible entre otros.

Similar situación es posible observar en Litman (2013), en donde se estudian diversos factores que influyen en el comportamiento de viaje de los usuarios. Este autor también realiza un resumen de diversos estudios y describe cómo utilizar dicha información para predecir los impactos de cambios en factores como precio o calidad de servicio.

Otros autores que presentan recopilaciones de estudios son Goodwin (1992), Nijkamp y Peeping (1998) y Holmgren (2007). El motivo de la gran cantidad de estudios que hay al respecto y los distintos resultados obtenidos, está explicado en estos mismos textos, y corresponden principalmente a factores tales como el tipo de dato utilizado, el horizonte de tiempo, la forma funcional del modelo, el motivo del viaje y que estos valores pueden cambiar mucho dependiendo de la persona o grupo de personas, así como el momento que se tomen los datos (cambios en las preferencias por diferencias sociales, culturales, entre otros).

En lo que si hay consenso, es en las variables explicativas más utilizadas. Tsai, (2013) hace un resumen en donde se muestra que Tarifa, costo de alternativas, vehículos-km

e ingreso son las variables más comunes, lo que es posible de corroborar con los datos de la Tabla 1.

En relación a esto, Holmgren (2007) dice que un modelo de demanda ideal, debiese incluir como variables explicativas Tarifa, Ingreso, calidad de servicio, índice automotor y precio de combustible

3. Descripción de los datos

En esta sección se presentarán los datos obtenidos que fueron seleccionados como posibles candidatos a variables explicativas, su fuente y un análisis de los mismos, así como también se evaluarán las variables dependientes de los modelos.

Las variables dependientes en términos agregados, son tanto la demanda mensual de viajes, y la cantidad de validaciones en el sistema integrado Transantiago (Bus-Metro) (Figuras 3 y 4 respectivamente). De forma desagregada, se utilizará además la cantidad de validaciones separando por modo (Bus o Metro). En la Figura 3 se pueden observar los viajes desde el mes de enero 2009, hasta fines del 2012. Los viajes anteriores a dicha fecha no fueron posibles de conseguir debido a que no existen registros de ellos en Transantiago.

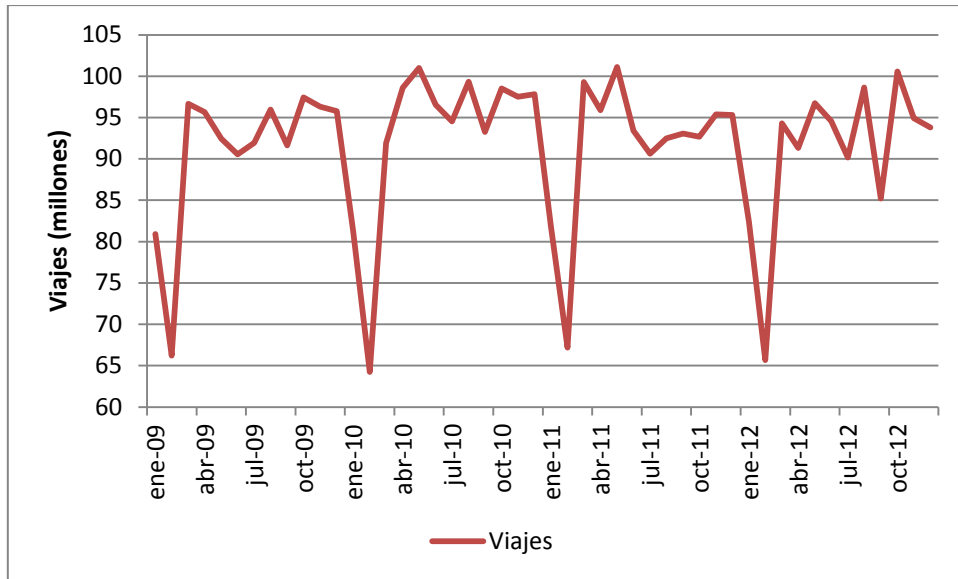


Figura 3 – Viajes Fuente: Transantiago

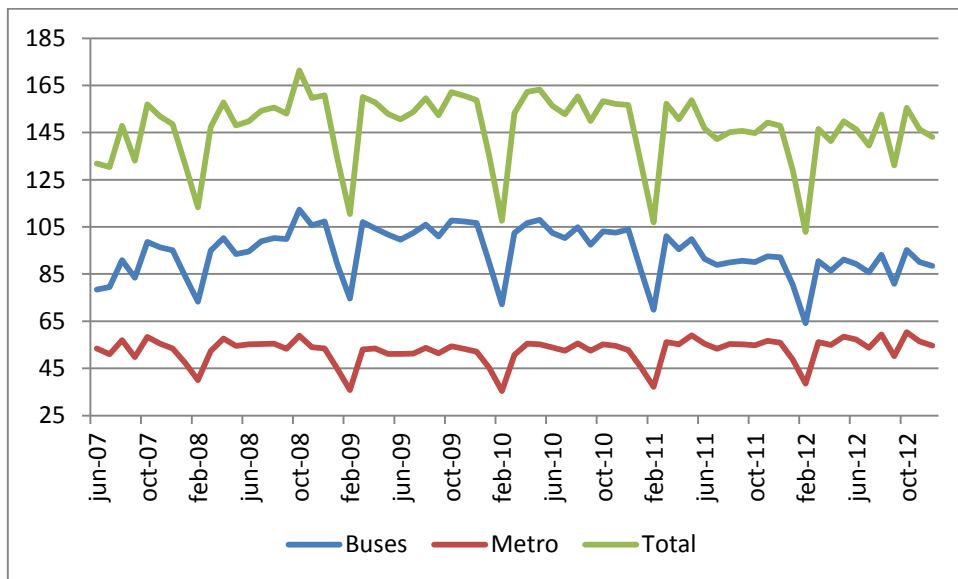


Figura 4 – Validaciones Fuente: Transantiago

En los 2 casos se puede observar de forma clara los ciclos de demanda anuales que se producen, con una disminución considerable en los meses de enero y febrero. Se puede observar además lo importante que es el Metro dentro del sistema de Transporte Público en Santiago, el cual ha tenido un aumento permanente en la demanda, mientras que para buses se observa una disminución en el total de validaciones, lo que no significa necesariamente una disminución en los viajes. Esto es posible corroborarlo con lo publicado por Simonetti, C (2013), en donde se muestra una disminución en la cantidad de etapas por viaje.

A su vez, para las variables explicativas es posibles separarlas en varios grupos, como aquellas variables generadoras de viajes, aquellas que representan otros modos de viajes, las variables correspondientes a cobertura del sistema de transporte público, y aquellas relacionadas con un nivel de servicio o confort para los usuarios.

En este primer grupo de variables, aquellas que influyen en la generación de viajes, se consideraron el nivel de Desempleo, medido por el INE con la metodología de trimestre móvil, las precipitaciones mensuales, obtenidas gracias a la Dirección Meteorológica de Chile y el Nivel de Ingreso, el cual esta medido como Remuneraciones hora, índice medido mensualmente por el Instituto Nacional de Estadística. La evolución de estas variables se observan en las figuras siguientes.

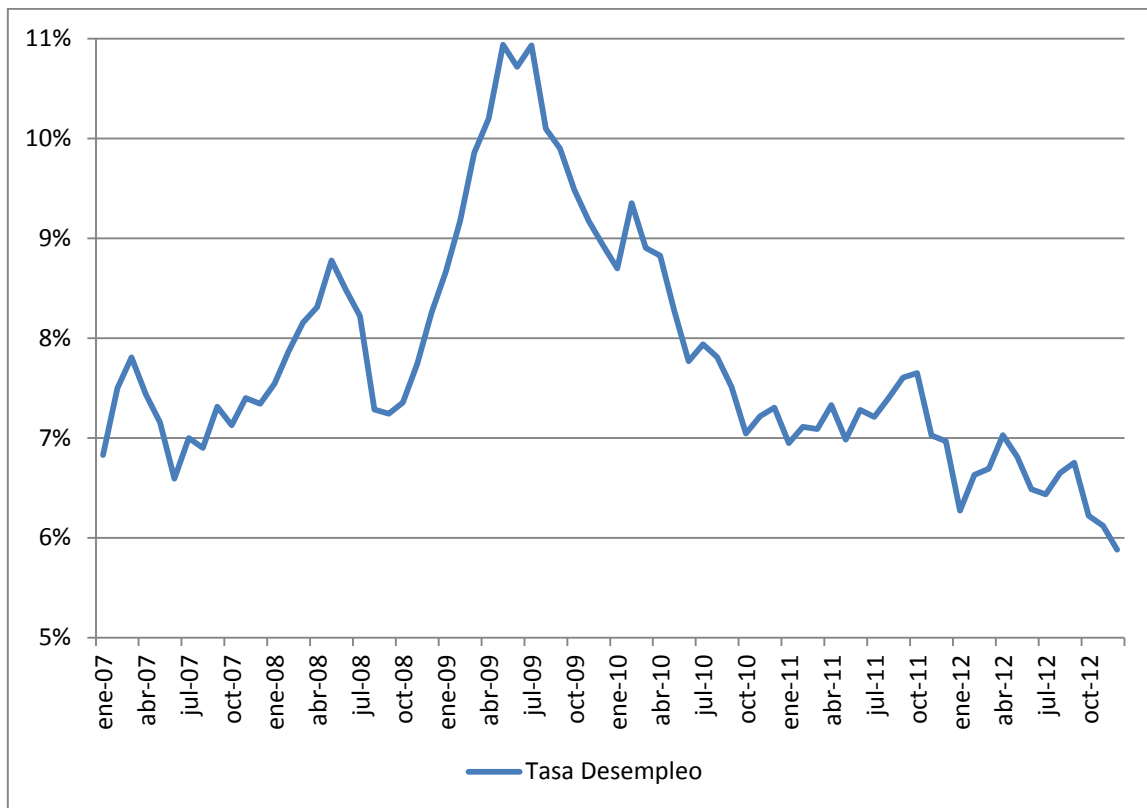


Figura 5 – Tasa de Desempleo Santiago Fuente: INE

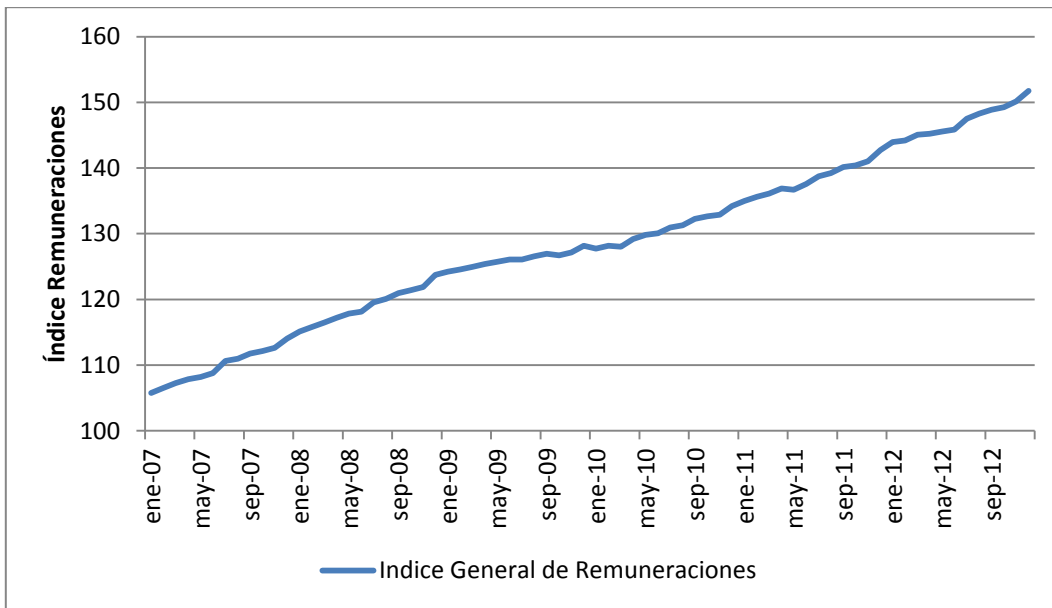


Figura 6 – Índice General de Remuneraciones Fuente: INE

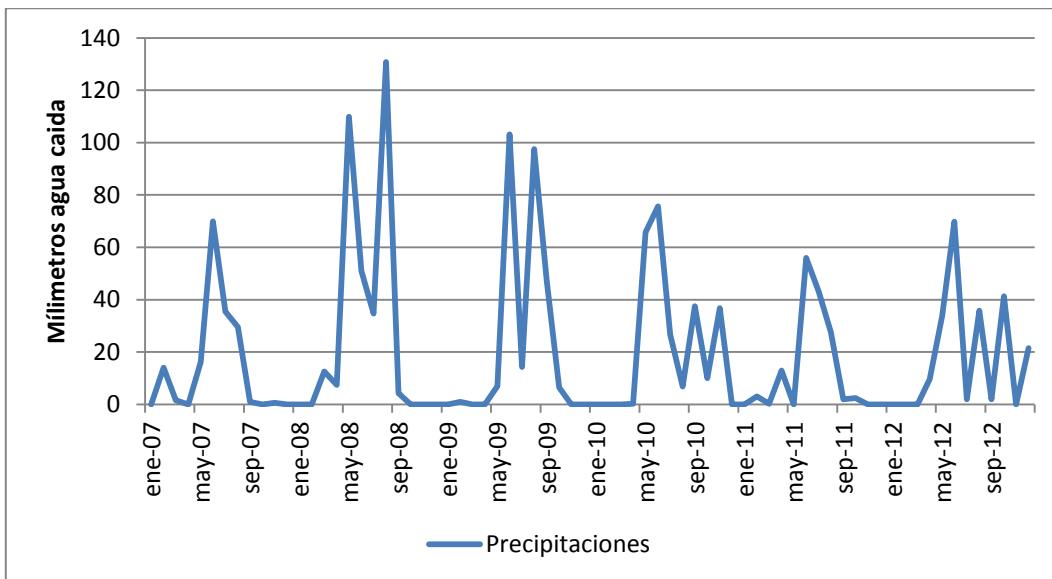


Figura 7 – Precipitaciones Fuente: Dirección Meteorológica de Chile

La tasa de desempleo (Figura 5) se encuentra en esta categoría, puesto a que el nivel de empleo es una fuente generadora de viajes obligatorios que se deben hacer de forma diaria. Es por esto que un bajo índice de desempleo debiese aumentar la cantidad de viajes. Es posible observar una fuerte alza en el desempleo producto de la crisis subprime en el 2008, que recién comienza a descender a mediados del 2009.

El índice general de remuneraciones (Figura 6) mide el ingreso promedio hora, y tiene la ventaja de ser medido de forma mensual. Este índice toma el valor de 100 en enero

del 2006, llegando a un valor sobre 155 en Julio 2013. Esta variable en particular, es entregada por el INE y es posible desagregarla por sectores productivos. Debido al uso que se le pretende dar, se prefirió utilizar el índice general, consistente en el promedio de los mismos.

Las precipitaciones (Figura 7) por su parte también son consideradas una variable generadora de viajes, debido a que son capaces de cambiar tanto su frecuencia, el modo y si se realiza el viaje o no. Ejemplo de esto es que pueden disminuir aquellos viajes no obligatorios, aumentar la congestión por lo que la gente puede preferir el transporte subterráneo, desincentivar el uso de la bicicleta o caminar, entre otros. Se logra ver un lógico aumento en las precipitaciones en los meses de invierno.

Pasando ahora al segundo grupo de variables, se mostrarán a continuación aquellas relacionadas con el costo de las alternativas u otros modos de transporte, como lo son la Tarifa Bus-Metro, y el precio del litro de bencina (95 s/p) como manera de calcular el costo de utilizar el auto.

Las Figuras 8 y 9 muestran los valores nominales de estas variables, en donde es posible observar que desde los inicios de Transantiago, las tarifas de Metro (hora valle y punta) y Buses han tenido alzas de manera simultánea, con una única excepción en el mes de Junio del 2010, en donde se diferenciaron las tarifas de buses y Metro valle. Por su parte la tarifa estudiante ha recibido alzas menores en términos absolutos, sin embargo en términos relativos las alzas han sido sólo menores a la tarifa Metro Punta.

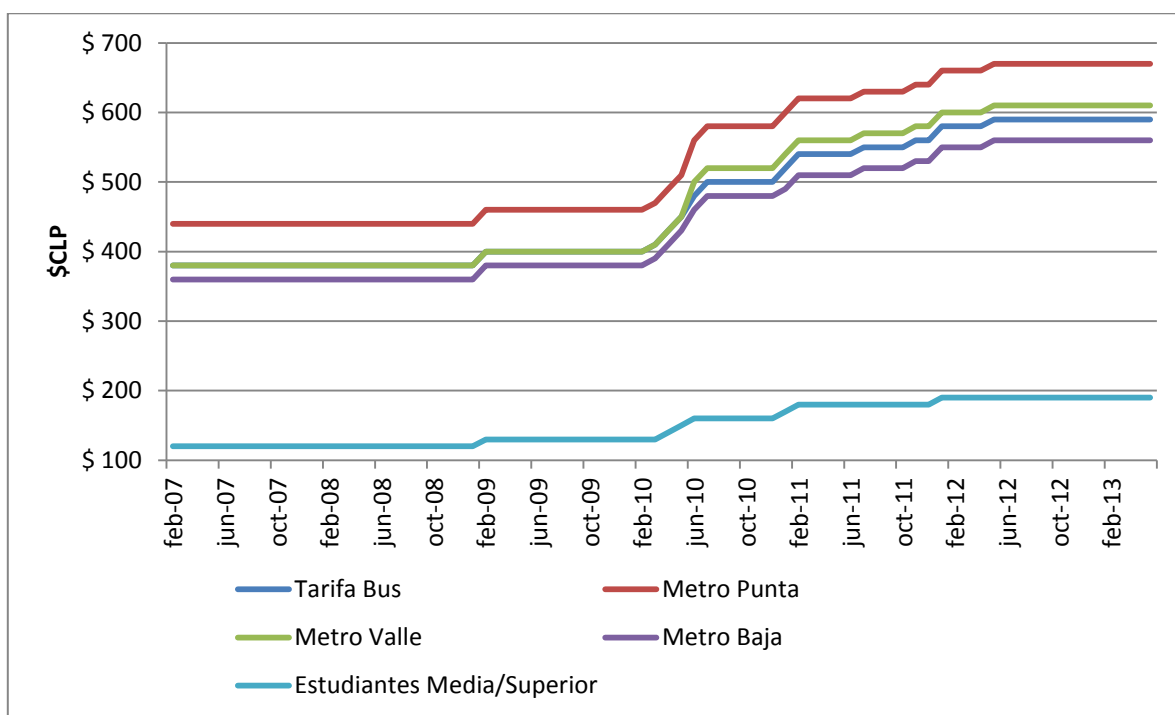


Figura 8 – Tarifas

Fuente: Transantiago

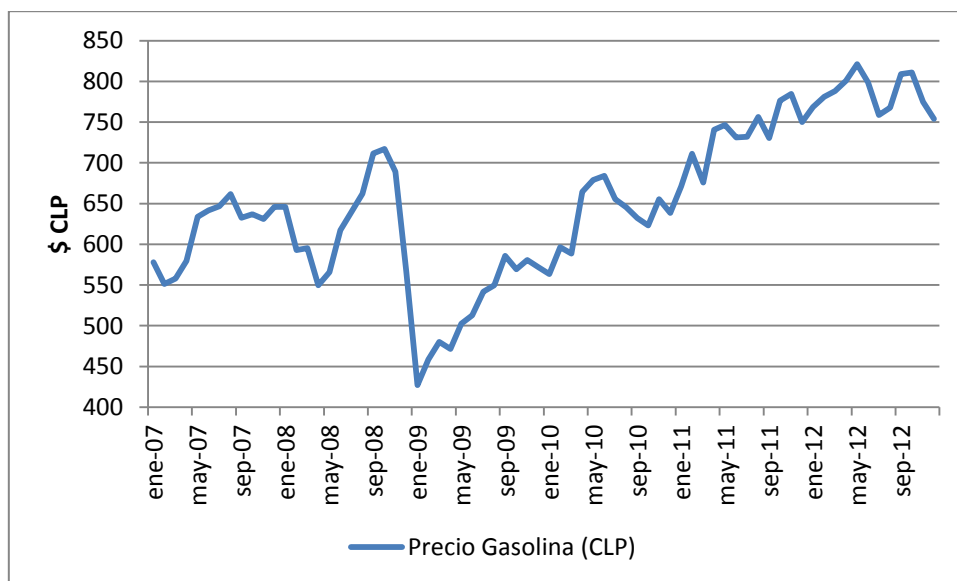


Figura 9 – Precio Gasolina Fuente: SERNAC y CNE

En cuanto a la gasolina, vale la pena mencionar el fuerte descenso que vivió el precio a fines del 2008, en dónde los precios del barril de petróleo bajaron sobre un 30%, debido a la desaceleración económica producto de la crisis, que conllevó a una disminución en la demanda del mismo.

Para los análisis ambas variables fueron corregidas por IPC, con fecha de diciembre de 2013. Es oportuno mencionar además que se estudió la posibilidad de utilizar los costos de Tarifas y Gasolina en función del índice de remuneración. La ventaja que tienen estos valores es que es posible estudiar las variaciones relativas al ingreso, y no únicamente su variación absoluta. Sin embargo estos valores de Tarifa / Índice remuneraciones y Precio Gasolina / Índice remuneraciones no resultaron significativos para ninguno de los modelos.

El tercer grupo de variables corresponde a las relacionadas con el nivel de cobertura espacial que ofrece los medios de transporte público, tales como la Longitud de la Red de buses, la Longitud de la Red de Metro. Junto a estas variables se mostrará la oferta de Buses y Metro.

La longitud de la red de Buses, Figura 10, se varía en base a los cambios en los Planes Operaciones, es decir, por fusión, extensión y creación de nuevos servicios tales como las variantes entre otros. Desde comienzos del 2007 hasta la fecha se ha aumentado la red de buses en cerca de 700km, con el efecto no tan solo de obtener una mayor cobertura, sino que con la creación de rutas más directas que han logrado una disminución en el número promedio de transbordos (Simonetti, 2013).

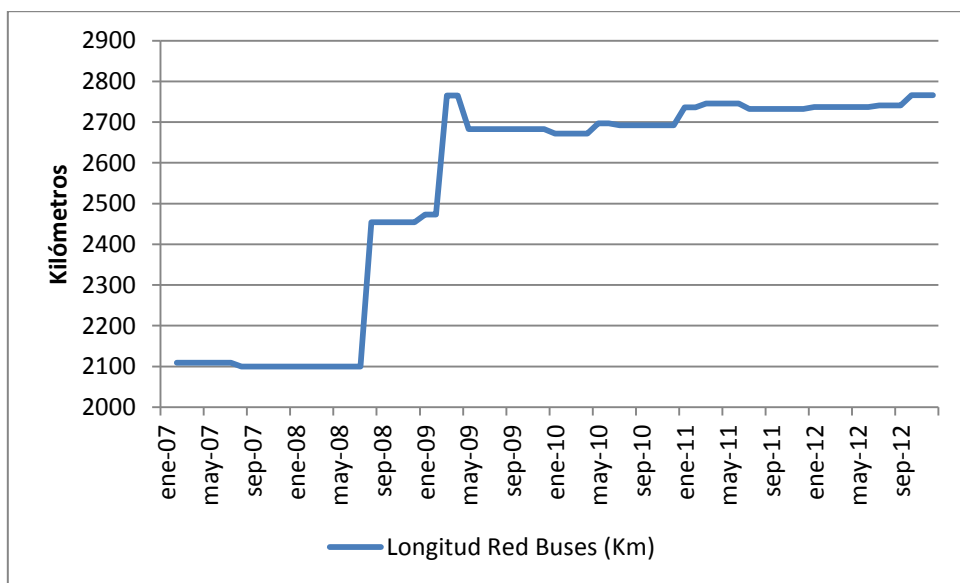


Figura 10 – Longitud Red Buses Fuente: Transantiago

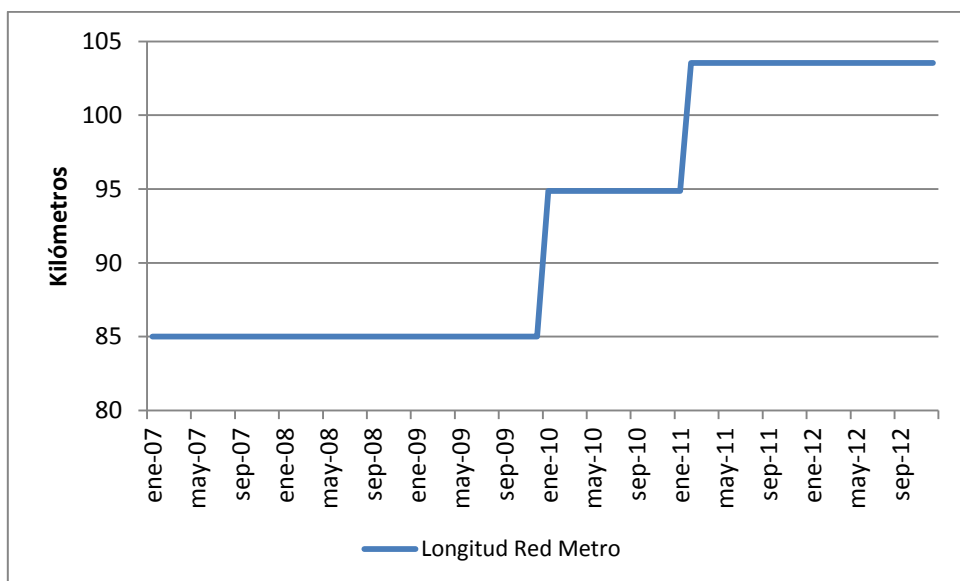


Figura 11 – Longitud Red Metro Fuente: Metro S.A

Por otra parte, han existido 3 adiciones a la red de Metro (Figura 11) desde el año 2007. La primera de ellas ocurrió en noviembre de 2009, cuando se abrió el Metro San José de la Estrella, una de las llamadas “estaciones fantasmas”. Esta apertura no está demostrada gráficamente en la Figura 11, debido a que no varió la longitud de la red. En enero del 2010 empezaron a operar 3 estaciones nuevas en la L1 y 5 nuevas estaciones en la L5, añadiéndose a la red 3,69 km y 5,87 km respectivamente. Finalmente en febrero del 2011 se abrieron 7 nuevas estaciones (Pudahuel – Plaza Maipú), sumando 8,11 km a la red, para contabilizar a la fecha una red de 103,5 km.

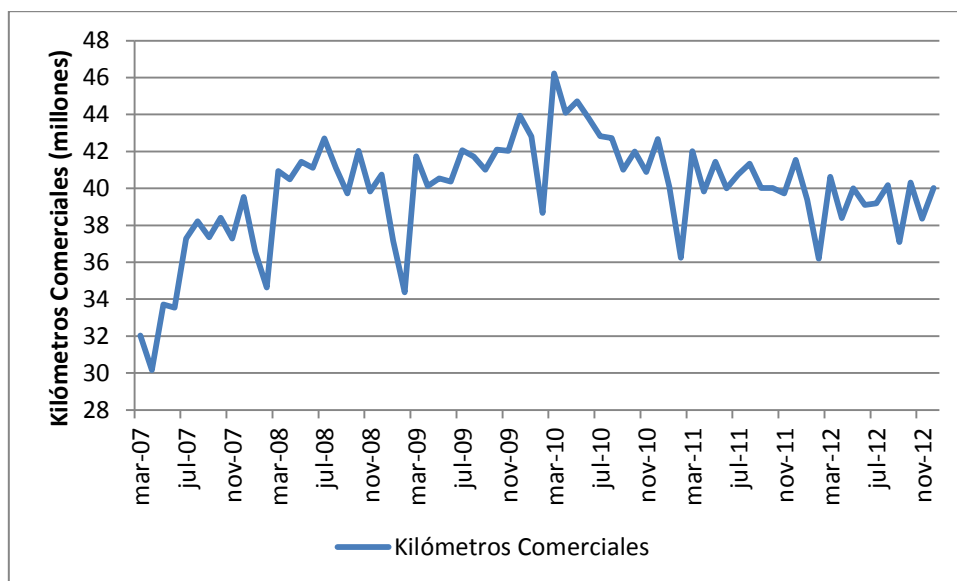


Figura 12 – Kilómetros Comerciales Fuente: Transantiago

Kilómetros Comerciales es una manera de medir la oferta de buses. Dicha oferta de buses mostrada (Figura 12) presenta un inconveniente, que es la correspondiente a los planes operacionales, por lo tanto es la oferta programada, la cual puede ser distinta a la oferta real. Para poder tener una estimación de la oferta real, se pueden utilizar índices de capacidad de transporte (ICT), pero estos índices son relativamente nuevos, dado que se comenzaron a medir junto con la implementación de los nuevos contratos a comienzos del 2012. Anterior a esto existieron indicadores tales como ICPH (índice de cumplimiento plazas-hora) y posteriormente ICPKH (índice de cumplimiento plazas-hora-km), por lo que no es posible utilizar algún indicador consiste en el tiempo de manera de observar la oferta real de buses de la forma de kilómetros comerciales. Una explicación y estudio más profundo de estos índices puede verse en Beltrán *et al* (2012).

Por otro lado, Metro si cuenta mediciones de oferta real, con agregación mensual, por hora de servicio, línea y vía de la misma. Esto entrega una gran gama de posibilidades para utilizar como variables, pero debido a que el estudio tiene un enfoque agregado, se decidió utilizar la oferta Coche Kilómetro entregada por Metro. (Figura 13)

En reuniones con la Gerencia de Planificación de Metro, adelantaron que en modelos de demanda que han realizado para análisis interno, esta variable resultaba significativa y estaba correlacionada con la demanda total de Metro, resultados que fueron replicados en los modelos de demanda de Metro obtenidos en este estudio.

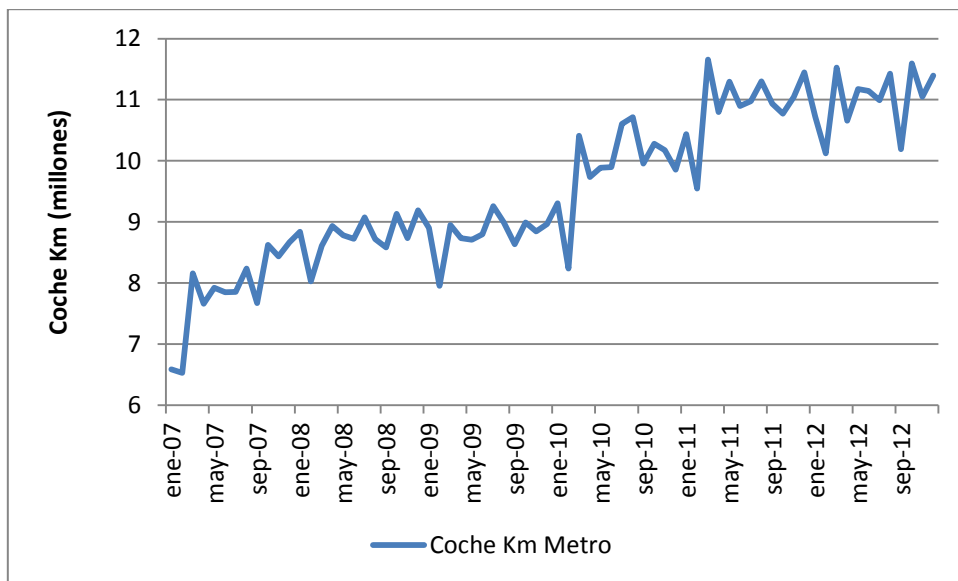


Figura 13 – Coche Kilómetro Metro Fuente: Metro S.A.

El cuarto grupo de variables, son aquellas capaces de medir o resultar proxy para el nivel de comodidad de transporte. Entre ellas se encuentran el ICR (índice cumplimiento de regularidad), la frecuencia tanto de buses como de metro, debido a que puede resultar una buena proxy a los tiempos de espera que dependen del intervalo entre buses, cuyo valor promedio es el inverso de la frecuencia; y el índice automotor, encargado de medir la cantidad de autos cada 1000 habitantes, el cuál puede utilizarse como una manera de ver cómo ha crecido la congestión vial en la ciudad de Santiago.

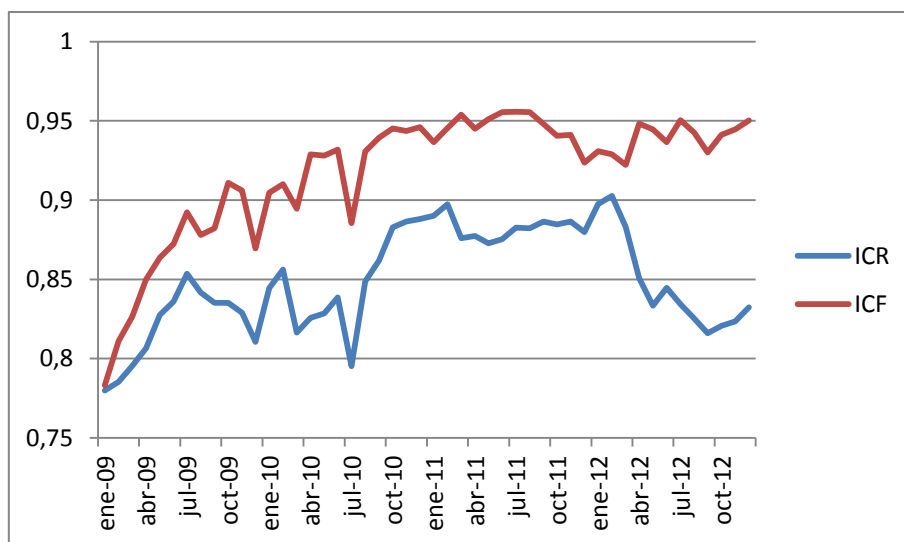


Figura 14 – ICR e ICF Fuente: Transantiago

La leyenda de las siguientes dos figuras corresponde al horario utilizado en Transantiago para dividir el día. Estos horarios son Punta Mañana (PM), Punta Tarde

(PT), Transición Punta Mañana (TPM), Fuera de Punta Mañana (FPM), Punta Mediodía (P Med), Fuera de Punta Tarde (FPT) y Transición Nocturna (T Noc). La sigla SS corresponde a Servicio Sentido.

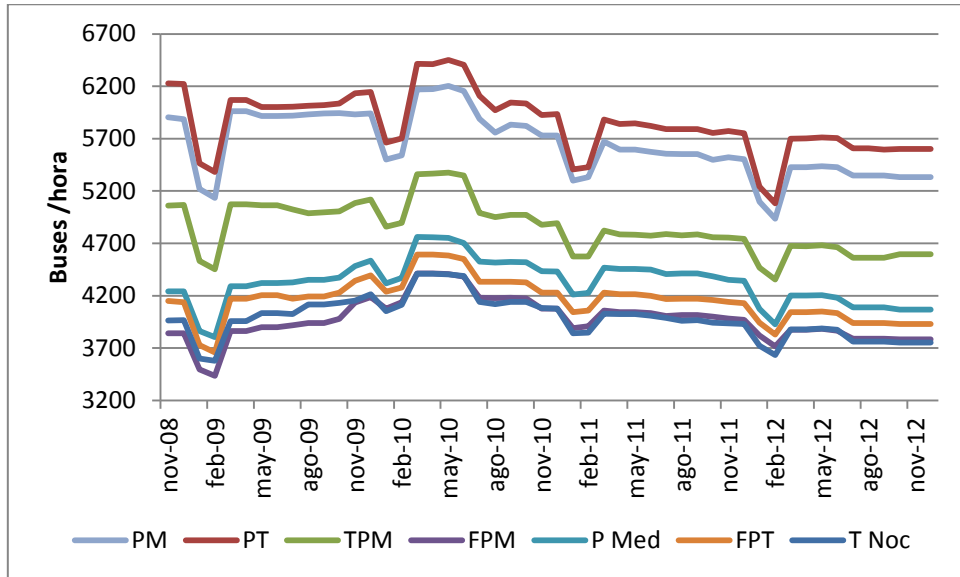


Figura 15 – Frecuencia Programada Buses Fuente: Transantiago y Elaboración Propia

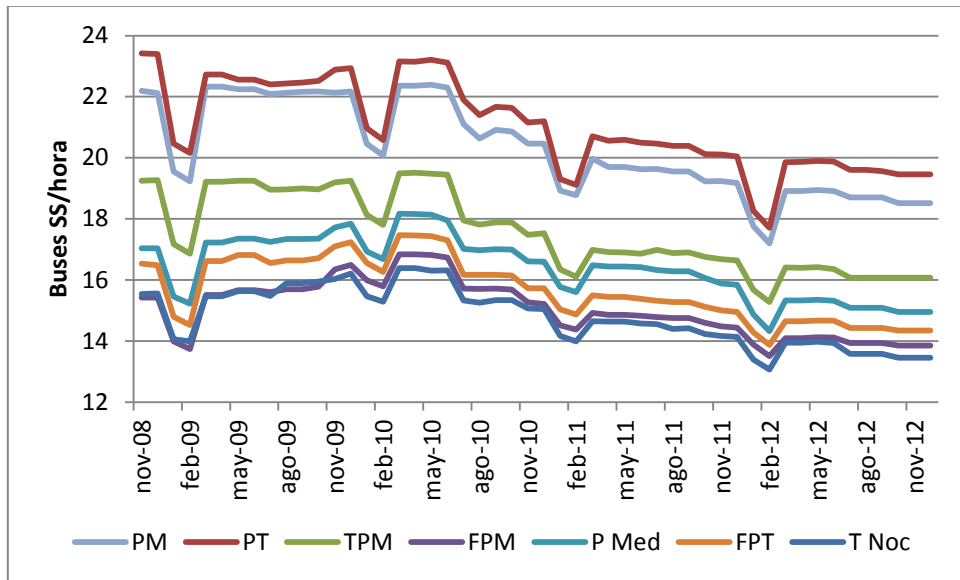


Figura 16 – Frecuencia Programada Servicio Sentido(SS) Fuente: Transantiago y Elaboración Propia

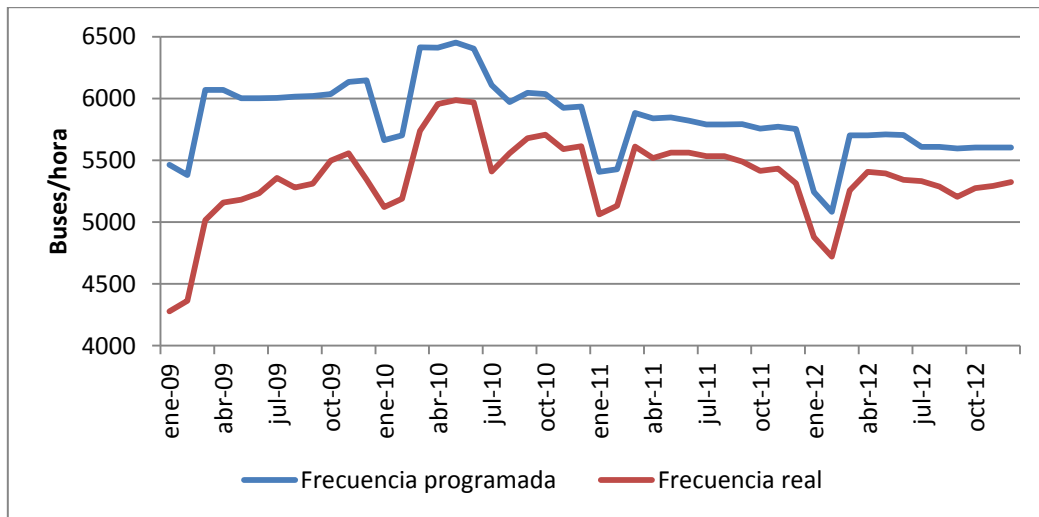


Figura 17 – Frecuencia real vs Frecuencia programada en horario Punta Mañana Fuente: Elaboración Propia

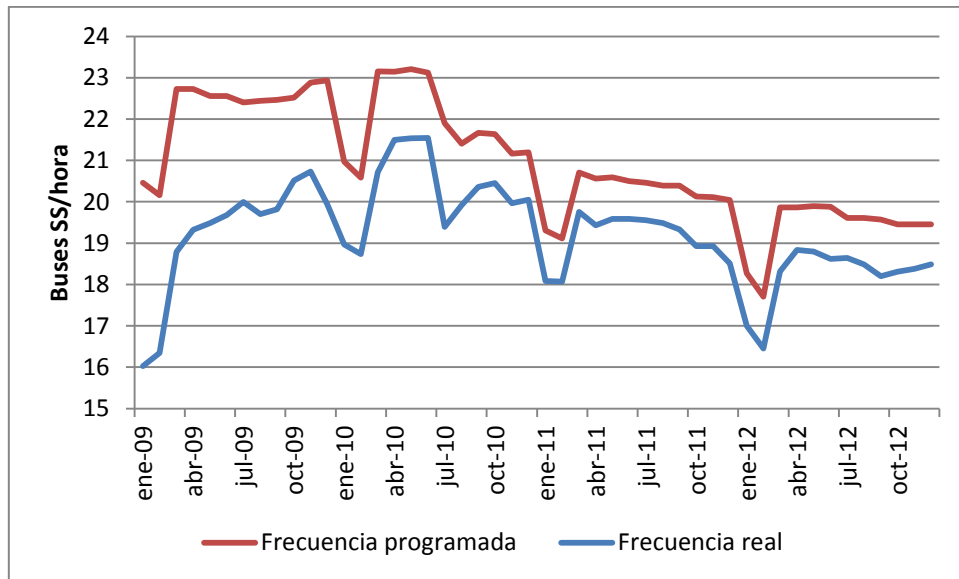


Figura 18 – Frecuencia promedio SS real vs Frecuencia promedio programada SS Fuente: Elaboración Propia

En las Figuras 15 y 16 se puede observar la frecuencia programada según los planes operacionales de Transantiago, así como la frecuencia promedio por servicio sentido.. Es importante considerar que para distintos horarios operan distintos servicios (Nocturnos, Expresos, Cortos, etc.), y tienen una frecuencia distinta, sin embargo de ambas figuras se puede desprender que todos los horarios se mueven de forma conjunta, por lo que no es errado el considerar algún horario en particular para demostrar el efecto de construir la frecuencia real (Figura 17) y la frecuencia promedio servicio sentido (SS) real (Figura 18). Para esto se multiplicó la frecuencia por el ICF, observando en algunos casos diferencias mayores a 1000 buses/hora entre lo programado y lo observado.

Es importante notar y motivo de una investigación más exhaustiva, que las frecuencias programadas han disminuido en los últimos años, lo que se puede deber a una mayor

cantidad de servicios para la misma flota, una adaptación de los planes operacionales, una respuesta a la demanda, u algún otro motivo.

Esta información de frecuencias es posible desagregar en servicios alimentadores y troncales, así como por Unidades de Negocio. Debido al enfoque del estudio, se decidió utilizar la frecuencia promedio servicio sentido, tomando en consideración tanto troncales como alimentadores.

Los valores de frecuencia utilizados corresponden a los estipulados en los Planes Operacionales de Transantiago, no a los valores obtenidos por GPS.

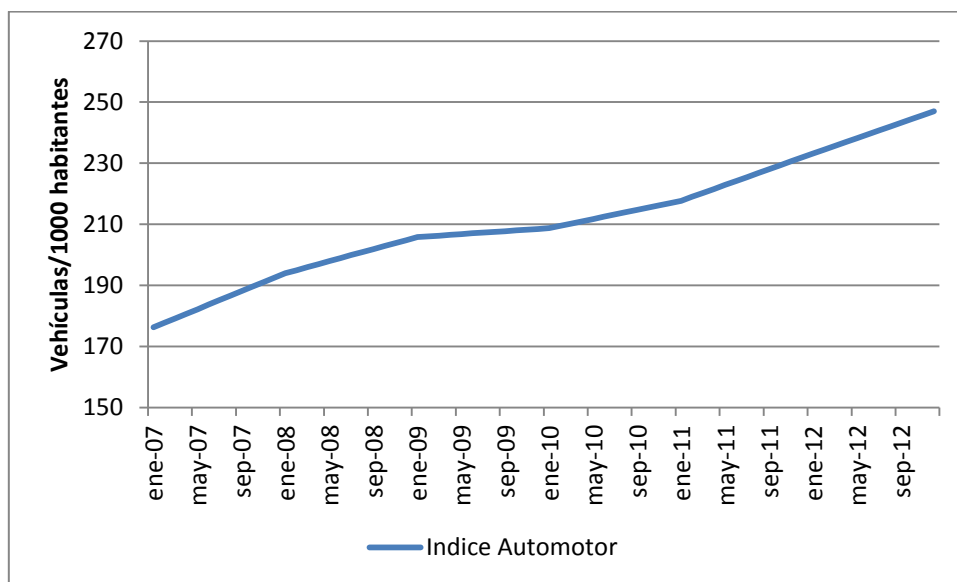


Figura 19 – Índice Automotor Fuente: Elaboración Propia en base a datos del INE

El índice automotor (Figura 19) fue construido en base a estimaciones demográficas para Gran Santiago del INE y el parque automotriz de Gran Santiago de la misma fuente. Los resultados son los esperados para países en vías del desarrollo y si bien el crecimiento en los últimos años ha sido notable, no difiere mucho del crecimiento de países similares.

Es estudiado en la literatura el utilizar el índice automotor como un equivalente al nivel de ingresos, por lo que no es extrañar que sea sólo un reflejo de la situación económica de Chile. Sin embargo, sea cual sea el motivo de este crecimiento, los resultados están a la vista en el día a día, aumentando la congestión vial y en consecuencia disminuyendo las velocidades de las vías, aumentando los tiempos de viaje tanto para automovilistas como para usuarios de Transantiago. Es por estos motivos la inclusión de este indicador como una variable que afecta el confort, y producto de la alta correlación teórica (y empírica) de este índice con los ingresos, es que se pretende mantener esta variable, y modificar los ingresos haciendo que tanto la gasolina, como las tarifas sean funciones del mismo para estudiar el efecto real que tienen las tarifas sobre la demanda de viajes y validaciones.

Finalmente se consideraran como variables explicativas la cantidad de días laborales que tiene un mes, los sábados e interfiados (se comportan de manera similar) y los domingos y festivos. Con esto se espera atrapar el efecto que los distintos tipos de días

tienen sobre el comportamiento de la demanda. Esta técnica ya ha sido utilizada anteriormente para el estudio de demanda de transporte público (Gaudry, 1975). Una segunda opción es normalizar los meses a 30 días, ponderando la demanda según corresponda (Jara-Díaz y Paredes; 1993)

Una variable particular que vale la pena mencionar, es la evasión. Una disminución en la misma aumenta la demanda observada a través de validaciones, pero el impacto en términos reales es menor considerando que los esta gente se está subiendo en los buses de todas formas. Se pueden cuantificar beneficios tanto por el conseguir que más usuarios paguen, o que los que no paguen no se suban, teniendo efectos distintos.

A esto se le debe sumar que las dificultades de medir evasión, a diferencia de la mayoría de los datos pasivos obtenidos a través de Transantiago, la medición de evasión requiere medidas en terreno, que por el puro hecho de estar midiéndolas interfieren en los niveles naturales de evasión. La medida oficial de evasión está a cargo del Programa Nacional de Fiscalización, del Ministerio de Transporte y Telecomunicaciones, pero estas medidas pueden diferir en valores de hasta un 100% en mediciones realizadas por otros organismos.

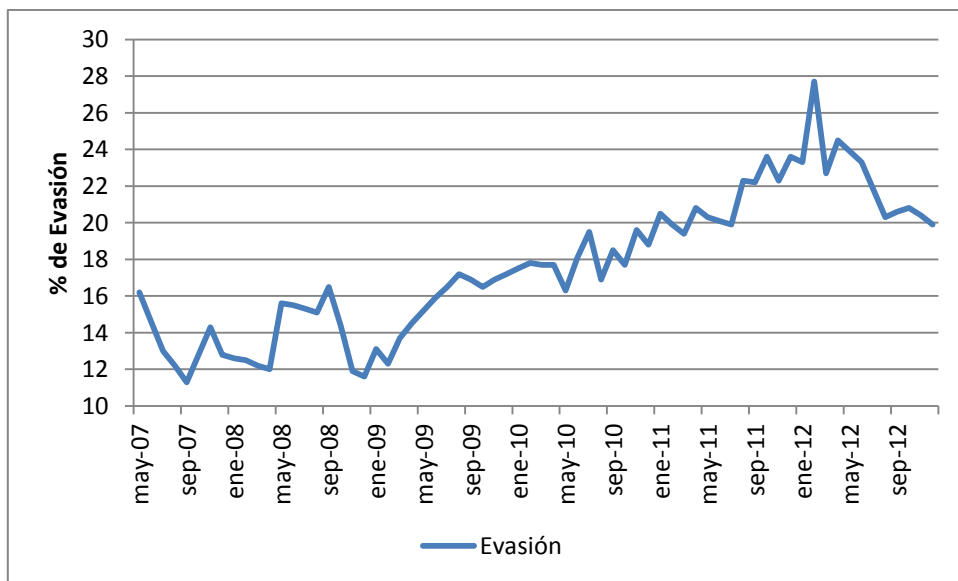


Figura 20 – Evasión Fuente: Fiscalización de Transportes

La evasión (Figura 20) no resultó ser significativa como variable explicativa ni en los modelos de validaciones ni en el modelo de viajes. Los nuevos y contratos firmados con los operadores de Transantiago en el año 2012 incentivan a que sean las mismas Unidades de Negocio las que combatan la evasión.

Una serie de indicadores y variables extras fueron creadas con el fin de explicar distintos fenómenos que pueden influir en la demanda del sistema de Transporte en Santiago. Entre ellas se pueden mencionar las marchas estudiantiles de carácter masivo, causantes de interrumpir o cambiar servicios por algunas horas. En distintos modelos se probaron la inclusión de estas marchas, siendo los resultados negativos para cada modelo, es decir, no resultó ser una variable significativa para modelar variaciones en la demanda.

Del mismo modo se crearon variables dummies para atrapar los efectos de paros, e irregularidades en servicios de buses. Si bien es de esperar que éstas se vean reflejadas en una variación en la cantidad de viajes o validaciones realizados, los resultados muestran que su impacto a nivel mensual de aquellas contingencias no son importantes o significativos para los modelos.

4. Estimación Modelos

Para estimar los modelos de demanda, tanto de validaciones como de viajes, se utilizaron modelos de regresión lineal múltiple. Los modelos de regresión lineal múltiple, al igual que cualquier otro modelo estadístico, se basan en una serie de supuestos que serán abordados a continuación. Durante el presente capítulo, los modelos presentados serán revisados con el objetivo de saber si es que cumplen con dichos supuestos, de manera de poder garantizar la validez de los mismos.

4.1 Supuestos del modelo de regresión lineal

4.1.1 Linealidad

Un incumplimiento en el supuesto de linealidad suele llamarse error de especificación del modelo. Suelen presentarse errores de especificación por motivos tales como la omisión de variables predictoras importantes, una relación entre variables dependientes e independientes no lineal o la inclusión de variables independientes que no sean significativas para el modelo.

La principal manera de revisar este supuesto es a través de un método gráfico, en donde se generan gráficos parciales entre la variable dependiente y cada una de las variables predictoras, pudiéndose observar de esta forma si es que existe alguna relación lineal, cuadrática, etc.

4.1.2 Independencia

En un modelo de regresión lineal bien construido, se espera que exista independencia entre los residuos. Cuando se trabaja con series temporales no es extraño encontrarse con residuos relacionados entre sí.

El estadístico utilizado para medir el grado de independencia de los residuos es el estadístico de Durbin-Watson (1951), cuyo valor oscila entre 0 y 4, tomando el valor de 2 cuando estos son independientes. Como regla de oro se asume independencia de los residuos cuando el estadístico se encuentra entre 1,5 y 2,5.

La forma en que es calculado este estadístico es:

$$DW = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}$$

En donde $e_i = Y_i - \hat{Y}_i$, Y_i es el valor de la variable explicada, mientras que \hat{Y}_i es valor estimado de esta variable según el modelo.

4.1.3 Homocedasticidad

El supuesto de homocedasticidad indica que para cada combinación de valores de las variables independientes, la varianza de los residuos debe ser constante. La falta de homocedasticidad es también llamada heterocedasticidad.

Entre las distintas formas de analizar la homocedasticidad de una regresión, se encuentra el método gráfico, en donde se grafican los residuos tipificados contra los pronósticos tipificados. Este gráfico debiese mostrar una nube de puntos sin un patrón

definido. Al verse un patrón (crecimiento o decrecimiento en el tiempo, aumento progresivo de la varianza, etc.) se puede decir que se está en presencia de heterocedastidad.

4.1.4 Normalidad

Otro supuesto que debe de seguir una correcta regresión lineal, es que los residuos de la misma se deben distribuir normalmente con media cero. Existen distintas técnicas de observar esto, como lo son los histogramas de residuos tipificados y los gráficos de probabilidad normal.

Un gráfico de probabilidad normal está representado por un eje con la probabilidad acumulada del residuo, contra la probabilidad teórica del mismo. De este modo se espera contar con una nube de puntos alineados sobre la diagonal del gráfico

4.1.5 No Colinealidad

No debe existir en un modelo bien construido un alto grado de relación lineal entre las variables independientes. El que exista un grado de relación lineal da paso a problemas de multicolinealidad. Para testear la no colinealidad se utilizaron dos métodos. En primer lugar se observaron las matrices de correlaciones, observándose de ese modo qué variables presentaban un nivel de colinealidad importante con otras.

El segundo método corresponde a utilizar los FIV o Factores de inflación de Varianza. Es un índice que indica cuánto es afectada la varianza de un coeficiente de la regresión debido a la multicolinealidad existente. Como regla de oro se dice que para un valor FIV bajo 5 no existen problemas de colinealidad, entre 5 y 10 existen problemas no severos, y mayor a 10 la varianza está muy inflada debido a la colinealidad. Sin embargo para algunos autores los problemas serios de multicolinealidad comienzan con $FIV > 5$. Mayor información se puede encontrar en Hair *et al*, 1995; Marquardt, 1970; Neter *et al* 1989). El Factor de Inflación de la Demanda para cada variable es calculado por el programa usado para la estimación de modelos, SPSS

4.2 Formas funcionales

Existen distintos modelos de regresiones utilizados en esta memoria. En esta sección se procederá a analizar la forma funcional de cada uno de ellos, de manera de ver las ventajas y limitaciones de los mismos, y las relaciones entre las variables explicativas y la variable dependiente.

La primera forma funcional es un modelo de regresión lineal, el cual considera una relación lineal entre las variables explicativas y la variable dependiente. El modelo lineal queda definido como se muestra en la ecuación (1).

$$Y = \beta_0 + \sum_{i=1}^n \beta_i X_i + \varepsilon \quad (1)$$

En este modelo, el coeficiente β_0 representa la constante, mientras que β_i el impacto que causa en Y la variación de una unidad en X_i . Las elasticidades en el modelo lineal

no son constantes, sino que varían según los valores de X_i e Y . Estas elasticidades son calculadas como la ecuación (2).

$$e_i = \frac{dY}{dX_i} \cdot \frac{X_i}{Y} \quad (2)$$

La siguiente forma funcional mostrada es la de doble logaritmo, conocida también como log-log. Esta forma funcional es derivada de la función exponencial mostrada en la ecuación (3).

$$Y = A \prod_i X_i^{\beta_i} \quad (3)$$

De este modo, la forma función log-log es obtenida transformando la ecuación (3), aplicando logaritmo en ambos lados de la ecuación:

$$\ln Y = \beta_0 + \sum_{i=1}^n \beta_i \ln X_i, \quad \text{en donde } \beta_0 = \ln(A) \quad (4)$$

La principal ventaja de los modelos log-log, es la facilidad de interpretar tanto sus parámetros, como las elasticidades del modelo. Las elasticidades en los modelos log-log son constantes (a diferencia del modelo lineal), y corresponde al término β_i . Esta elasticidad representa el impacto porcentual Y ante una variación de un 1% de la variable X . El cálculo de la elasticidad se puede observar en la ecuación (5)

$$e_i = \frac{dY}{dX_i} \cdot \frac{X_i}{Y} = A\beta_i X_i^{\beta_i-1} \cdot \frac{X_i}{AX_i^{\beta_i}} = \beta_i \quad (5)$$

Otra ventaja que es importante considerar, es que estos modelos tienden a reducir la dispersión original de una serie, lo que es importante pues disminuye el riesgo de que los modelos presenten heterocedasticidad.

Una tercera forma funcional nace de la función exponencial mostrada en la ecuación (6), y al igual que el caso anterior, al tomar logaritmo de ambos lados se obtiene la función log-lineal mostrada en ecuación (7).

$$Y = A \prod_i e^{X_i \beta_i} \quad (6)$$

$$\ln Y = \beta_0 + \beta_i X_i. \quad \text{en donde } \beta_0 = \ln(A) \quad (7)$$

En esta forma funcional (log-lineal), el término β_i representa el porcentaje de variación de Y al cambiar en una unidad el valor de X_i . La elasticidad en este caso no es constante, dado que como se observa en la ecuación (8), depende de X_i .

$$e_i = \frac{dY}{dX_i} \cdot \frac{X_i}{Y} = A\beta_i e^{X_i \beta_i} \cdot \frac{X_i}{Ae^{X_i \beta_i}} = X_i \beta_i \quad (8)$$

Finalmente se puede observar la forma funcional lineal-log, la cual nace de la forma exponencial mostrada en la ecuación (9), y quedando finalmente como se muestra en la ecuación (10).

$$e^Y = A \prod_i X_i \beta_i \quad (9)$$

$$Y = \ln(\beta_0) + \sum_i \ln(\beta_i X_i) \quad \text{en donde } \beta_0 = A \quad (10)$$

Dentro de los distintos criterios para poder elegir qué forma funcional elegir, está el comparar la bondad de ajuste de cada modelo. Este criterio sirve únicamente para aquellos modelos que tienen la misma variable dependiente. Para comparar modelos con variables dependientes distintas, como por ejemplo un modelo lineal contra un modelo cuya variable dependiente sea el logaritmo de la primera, entonces se utiliza el test de Ramsey (1969), llamado RESET por sus siglas en inglés (Regression Specification Error Test).

El test Ramsey es un test F sobre diferencias entre la bondad de ajuste R^2 entre un modelo lineal y uno no lineal. De este modo, sea la regresión

$$Y = \beta_0 + \beta_i X_i + \gamma_1 \widehat{Y}^2 + \gamma_2 \widehat{Y}^3 + \varepsilon_i$$

La hipótesis nula es $H_0: \gamma_1 = \gamma_2 = 0$, de manera de que si esta hipótesis es rechazada, entonces el modelo está mal especificado, y es posible que variables no lineales deben ser agregadas.

4.3 Estadísticos descriptivos

En la Tabla 2 se muestra los estadísticos descriptivos de las variables utilizadas para los modelos

Tabla 2 – Estadísticos Descriptivos modelo Validaciones

	Unidades	Mínimo	Máximo	Media	Desv. típ.
Transacciones	[millones pax/mes]	99,5	171,3	145,1	15,4
Viajes	[millones viajes/mes]	64,2	101,1	91,5	9,2
Tarifa	[\$ CLP]	404	596	495,1	68,1
Gasolina	[\$ CLP]	460	828	707,5	84,2
CocheKmMetro	[millones coche km]	6,52	11,66	9,55	1,28
KMComerciales	[millones km]	30,15	46,21	39,86	2,91
LongRedMetro	[km]	85,0	103,5	93,8	8,5
LongredBus	[km]	2100	2774	2549,05	266,8
Frecuencia	[buses servicio/hora]	17,7	26,6	21,6	2,01
DiasLaborales	[días]	17	23	20,84	1,28
Desempleo	[% desempleo]	,059	,109	,078	,012
IndiceAutomotriz	[autos / 1000 habitantes]	176,3	252,4	215,4	20,9
Precipitaciones	[mm / mes]	0,0	130,8	19,56	30,1
Evasión	[% evasión]	11,3	27,7	17,6	0,38
IngresoReal	[índice]	105,79	119,73	112,11	3,512

En la tabla se pueden observar distintas variables que fueron evaluadas en los modelos, siendo algunas de ellas consideradas en los modelos finales y otras como Índice Automotriz, Ingreso Real, etc., descartadas por diversos motivos. Dentro de estos motivos se encuentran problemas de correlaciones parciales, los cuales conllevan a colinealidad (ambos tratados durante el desarrollo de este capítulo) y algunas variables simplemente no resultaron ser significativas para los modelos de regresión múltiple probados.

4.4 Modelos de Validaciones

En esta sección se presentarán los modelos cuya variable dependiente corresponde a las validaciones o transacciones obtenidas a través de datos pasivos de la tarjeta Bip!. En primer lugar se mostrará la matriz de correlaciones parciales, para luego introducir los modelos de regresión lineal para las validaciones de buses, de Metro, y de ambos modos de manera agregada. Luego de esto se estudiarán las diferencias que existen al utilizar modelos normalizados y modelos sin normalizar y, para finalizar, se estudiarán diversas formas funcionales de estos mismos modelos.

4.4.1 Matriz de Covarianzas

La matriz de covarianza observada en la Tabla 3 corresponde a las variables independientes utilizando como variable de control las validaciones.

Tabla 3 – Matriz de Covarianzas

	Desemp leo	KMCo m	Longred Bus	LongRed Metro	IA	Gasoli na	DiasLabor ales	Precipitacio nes	PA	CocheKm Metro	Frecuen cia	Tari fa
Desempleo	1											
KMCom	,43	1										
LongredBus	,03	,26	1									
LongRedM etro	-,55	,00	,69	1								
IA	-,37	,00	,80	,89	1							
Gasolina	-,84	-,34	,01	,59	,39	1						
DiasLabor ales	-,05	,25	-,06	,01	-,03	-,01	1					
Precipitacio nes	,06	,20	-,09	-,09	-,08	,04	,01	1				
PA	-,37	,00	,81	,90	1	,39	-,03	-,08	1			
CocheKmM etro	-,52	,02	,69	,96	,90	,50	,13	-,10	,91	1		
Frecuencia	,32	-,20	-,53	-,63	-,63	-,25	,08	-,05	-,63	-,61	1	
Tarifa	-,63	-,15	,61	,94	,86	,63	,00	-,04	,87	,92	-,66	1

IA = Índice Automotor
PA= Parque Automotor

Es posible observar ciertos casos con altas correlaciones, ejemplo claro de esto es el índice automotriz, presentando altas correlaciones con la Longitud de la Red Metro, el Parque Automotriz, la oferta Coche-Km de Metro y la Tarifa. Esta alta correlación sugiere que puede haber problemas al ingresar dicha variable a los modelos, puesto que puede resultar en problemas de multicolinealidad.

Otras variables que presentan una alta correlación parcial es la Tarifa con Longitud de Red Metro (cada vez que se abrió una nueva estación de Metro subieron las tarifas). Además se pueden observar ciertas correlaciones como desempleo con el valor de Gasolina, en donde su alta correlación no necesariamente representa una correlación causal, sino que puede ser producto del tamaño muestral únicamente.

Tanto estos casos como aquellos no nombrados serán analizados a través del test de multicolinealidad realizado para cada modelo de regresión múltiple. Este test se basa en los factores de inflación de varianza FIV, en donde por regla de oro se acepta que factores menores a 10 no causan problemas serios de multicolinealidad

4.4.2 Resultados de los Modelos

A continuación se analizan los modelos base (regresiones lineales) para los modelos de validaciones. En primer lugar se analizará el modelo de validaciones agregado (Buses y Metro), para luego revisar de forma individual el modelo de Buses y el modelo de Metro.

Es importante mencionar que para todos los modelos se eliminó el mes de Marzo de 2010, debido al terremoto del 27 de Febrero del mismo año. Esto debido a que el mes de Marzo en particular no tuvo un desarrollo normal y agregarlo afecta el valor de los parámetros y la bondad de ajuste del modelo.

El motivo de utilizar como modelos base los modelos lineales, es que las distintas formas funcionales de los modelos corresponden a transformaciones realizadas sobre el modelo lineal.

4.4.3 Modelo Agregado de Validaciones

El modelo lineal siguiente (forma funcional en Ecuación 11) será considerado como el modelo base con el cual se compararán otras formas funcionales. Es importante notar que este modelo presenta un R cuadrado corregido de 0.922, lo que sugiere que un 92.2% de la variación de la demanda en cuanto a validaciones puede ser explicado a través de las variables explicativas seleccionadas. Además el F-Test indica que la relación entre la variable dependiente y sus predictores es significativa, ya que da un P-valor de 0,000.

$$Y = a_0 + a_e\delta_e + a_f\delta_f + a_j\delta_j + a_{mm}\delta_{mm} + a_{ml}\delta_{ml} + a_{lrb}LRB + a_{dl}DL + a_{frec}FREC + a_{tar}Tar + \varepsilon \quad (11)$$

En donde la variable dependiente Y corresponde a las validaciones, y las variables independientes y parámetros están definidos a continuación.

LRB: Longitud Red Bus (km)

DL: Días Laborales

FREC: Frecuencia (buses servicio sentido / hora)

Tar: Tarifa (Pesos)

Este modelo utiliza 5 variables dummy o mudas, correspondiente a Enero, Febrero, Julio, Mes Movilizado y Mes Largo. Las 3 primeras se utilizaron para modelos esos tres meses que presentan un comportamiento especial con respecto al resto de los meses (meses de vacaciones, donde existe una menor demanda). Las otras dos variables mudas utilizadas corresponden a Mes Largo (aquellos meses de 31 días), y mes movilizado, correspondiente a los meses de julio a octubre del año 2011, en donde el movimiento estudiantil tuvo un peak de adhesión y un gran número de colegios y universidades en Santiago se mantuvieron tomados o en paros.

$$\delta_e = \begin{cases} 1 & \text{si mes es enero} \\ 0 & \sim \end{cases}$$

$$\delta_f = \begin{cases} 1 & \text{si mes es febrero} \\ 0 & \sim \end{cases}$$

$$\delta_j = \begin{cases} 1 & \text{si mes es junio} \\ 0 & \sim \end{cases}$$

$$\delta_{mm} = \begin{cases} 1 & \text{si mes es mes movilizado} \\ 0 & \sim \end{cases}$$

$$\delta_{ml} = \begin{cases} 1 & \text{si mes es mes largo} \\ 0 & \sim \end{cases}$$

Tabla 4 . Modelo Agregado de Validaciones

Modelo Agregado de Validaciones	Coeficientes no estandarizados		Coeficientes tipificados			Estadísticos de colinealidad	
	B	Error típ.	Beta	t	Sig.	Tolerancia	FIV
(Constante) a_o	44,15	14,3		3,09	,003		
Enero a_e	-18,9	2,07	-,347	-9,14	,000	,830	1,205
Febrero a_f	-40,6	2,05	-,745	-19,84	,000	,847	1,180
Julio a_j	-6,73	1,86	-,134	-3,62	,001	,868	1,152
LongredBus a_{lrb}	,018	,003	,317	7,19	,000	,616	1,624
DiasLaborales a_{dl}	2,42	,41	,217	5,96	,000	,903	1,107
Frecuencia a_{frec}	1,53	,36	,213	4,24	,000	,475	2,106
Tarifa a_{tar}	-,047	,011	-,224	-4,11	,000	,403	2,484
MesMovilizado a_{mm}	-4,91	2,22	-,081	-2,21	,031	,885	1,130
MesLargo a_{ml}	3,29	1,17	,114	2,81	,007	,725	1,380

Observaciones 65

R cuadrado ,933

R cuadrado corregida ,922

Error típ. de la estimación 4,049

F(9 , 56) 86,829

P-Valor ,000

Analizando la significancia individual de cada variable explicativa en el modelo, se puede observar que todas las variables son significativas a un nivel de confianza de 95% y la mayoría son significativas a un 99% de confianza. Además se logra observar que el impacto de cada variable en la cantidad de validaciones se comporta de manera acorde a lo que la literatura sugiere. Las variables dummies de Enero, Febrero y Julio, logran explicar una baja en la demanda producida por efectos propios del comportamiento de esos meses (vacaciones escolares y universitarias, fecha preferida para salir de Santiago). Esta baja corresponde a 18.9 millones de validaciones para enero, 40.6 millones para Febrero y 6.7 millones para el mes de Julio.

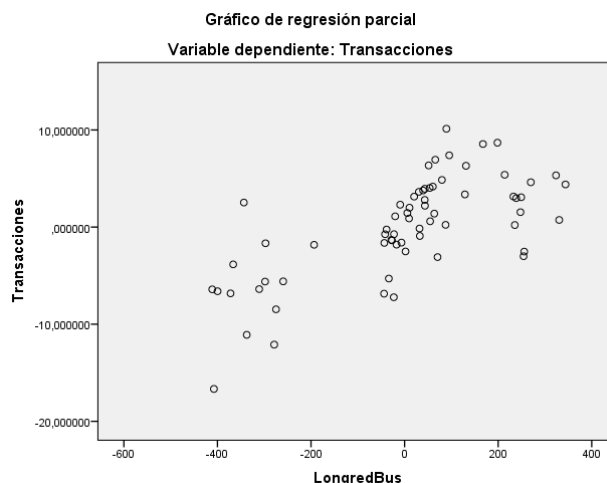
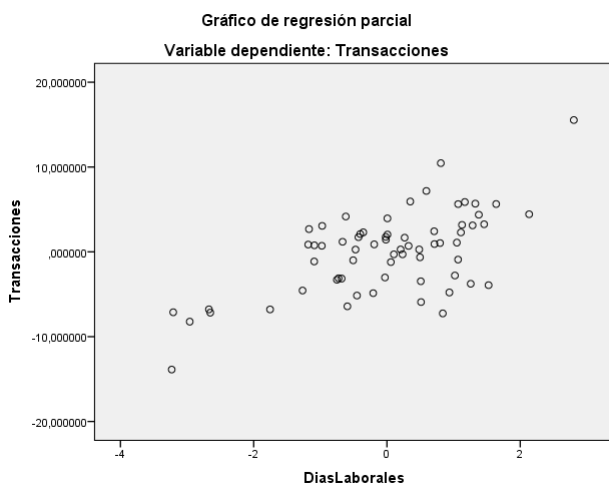
Se observa además una relación positiva entre la Longitud Red Bus y la cantidad de validaciones, lo que era de esperar debido a que una mayor longitud en la red es sinónimo de aumentar el alcance de la red. En cuanto a la relación Tarifas y demanda, se observa que es de signo negativo, como era de esperar. El valor obtenido indica que al aumentar en 1 peso el pasaje, se evidencia una disminución en la demanda de

47.000 validaciones al mes. La variable Mes Movilizado sugiere una relación negativa entre los meses en donde se intensificó el movimiento estudiantil y gran cantidad de establecimientos se sumaron a un paro general (desde Julio 2011 hasta Octubre 2011). Se estima que en tales meses hubo una disminución de 4.9 millones de validaciones en promedio.

Otras variables como Mes Largo, y Días Laborales también tienen un significado económico esperado, siendo ambas de valor positivo. Finalmente la Frecuencia entrega un signo positivo, es decir que al aumentar la frecuencia promedio en los servicios de Transantiago en 1 bus/hora para cada servicio, se ha observado una relación que sugiere que aumenta la cantidad de validaciones en 1.5 millones mensuales. Esto no necesariamente es una relación causa efecto, pues los planes operacionales se ajustan a variaciones estacionales de la demanda.

Se probaron distintas variables para este modelo que no quedaron finalmente. Entre ellas es importante mencionar las marchas que han ocurrido en importantes arterias de Santiago, condicionando el funcionamiento normal del transporte público. Sin embargo éstas no fueron significativas, lo que se puede explicar a que las marchas tan sólo interrumpen o cambian el funcionamiento normal por un corto periodo durante un día, cambio marginal si es que las evaluaciones son a un nivel mensual. Lo mismo ocurre para incidentes que afectaron el transporte público, como paro de conductores de alguna unidad de negocio, eventos, entre otros. Estas variables se desecharon en todos los modelos por los mismos motivos.

Se dejaron de lado variables como Ingreso Real, Parque Automotriz, el Valor de la Gasolina entre otros, debido a que no dieron significativos o presentaban una alta correlación, dando pie a problemas de multicolinealidad. Se pueden observar que los valores de FIV son todos bastante menores a 10, por lo que no se presentan problemas de colinealidad en este modelo.



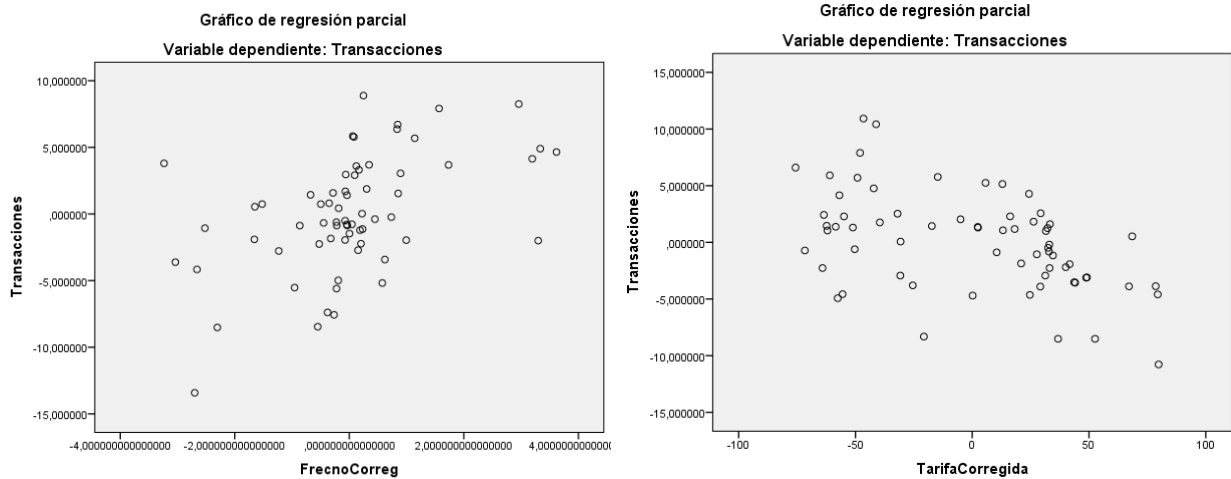


Figura 21 – Gráficos de Regresiones Parciales modelo agregado de validaciones

Se observa en los gráficos parciales que existe una clara relación lineal y positiva entre Transacciones y la cantidad de días laborales y entre transacciones y la longitud de red Bus. En cuanto a la frecuencia se observa una relación lineal positiva pero menos clara que los casos anteriores. Finalmente es posible notar que para el gráfico parcial de Transacciones con Tarifa, se observa una relación negativa. Todos los signos dan de acuerdo a lo encontrado por el modelo, y se logra observar además que esta regresión no viola el supuesto de linealidad.

Pasando a analizar ahora los residuos de los modelos, no se puede descartar la normalidad en los residuos, como es posible observar en la Figura 22.

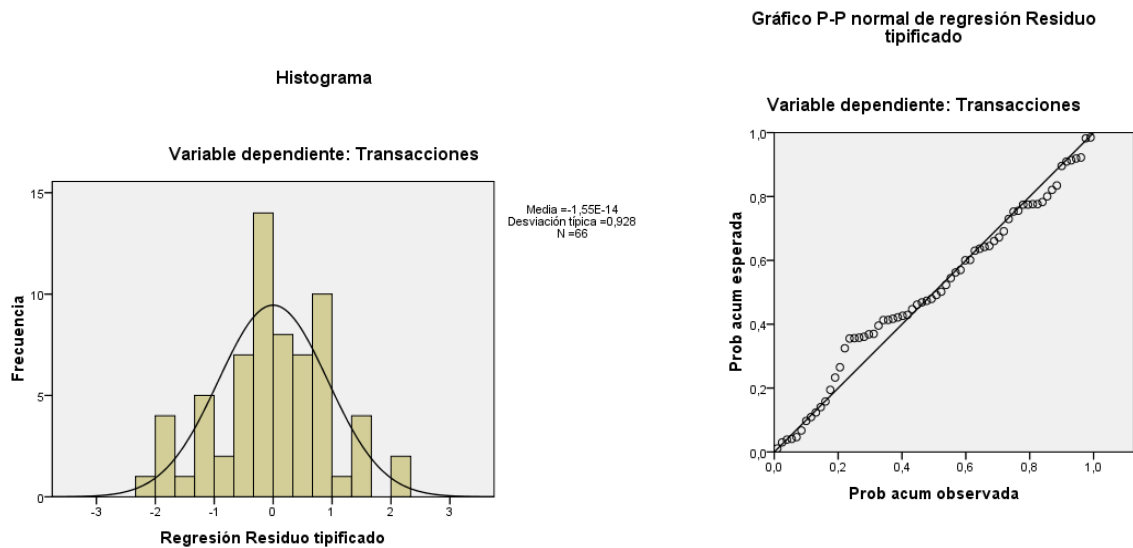


Figura 22 – Histograma y gráfico P-P normal modelo agregado

El gráfico de dispersión (Figura 23) permite observar un aumento en la varianza de los pronósticos menores a la media, sin embargo no es posible en base a este gráfico de

rechazar de manera concluyente el supuesto de homocedasticidad, ya que los datos anómalos corresponden a los meses de febrero.

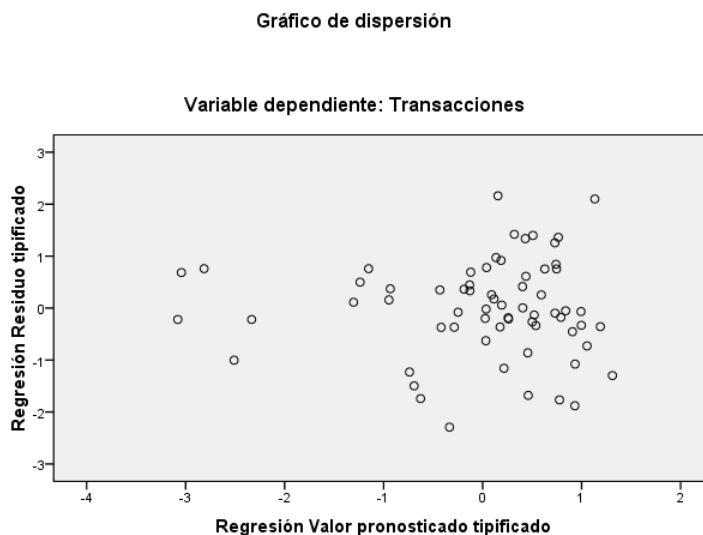


Figura 23 - Gráfico de dispersión modelo agregado

El test de independencia de los residuos señala que el estadístico de Durbin Watson tiene un valor de 1,436. El valor de dicho estadístico es muy cercano a 1,5, por lo que tampoco se puede rechazar de manera concluyente la independencia entre residuos.

4.4.4 Modelo de Validaciones de Buses

El siguiente modelo (Ecuación 12) también será considerado uno de los modelos base, y se trata de un modelo lineal de regresión múltiple. La variable dependiente corresponde a las validaciones realizadas únicamente en Buses, dejando de lado las validaciones en Metro en este modelo.

Este segundo modelo, presenta un R cuadrado de 0,889, es decir, un 88.9% de la variación de la demanda de validaciones en los buses puede ser explicado a través de las variables explicativas. Además el p-valor de 0,000 sugiere que existe una relación significativa entre la variable independiente y sus predictores.

$$Y = a_0 + a_e\delta_e + a_f\delta_f + a_j\delta_j + a_{mm}\delta_{mm} + a_{ml}\delta_{ml} + a_{lrb}LRB + a_{dl}DL + a_{frec}FREC + a_{tar}Tar + \varepsilon \quad (12)$$

En donde la variable dependiente Y corresponde a las validaciones o transacciones de buses, y las variables independientes y parámetros están definidos a continuación.

LRB: Longitud Red Bus (km)

DL: Días Laborales

FREC: Frecuencia (buses servicio sentido / hora)

Tar: Tarifa (Pesos)

Las variables mudas son las mismas utilizadas en el modelo agregado (buses + metro).

$$\delta_e = \begin{cases} 1 & \text{si mes es enero} \\ 0 & \sim \end{cases}$$

$$\delta_f = \begin{cases} 1 & \text{si mes es febrero} \\ 0 & \sim \end{cases}$$

$$\delta_j = \begin{cases} 1 & \text{si mes es junio} \\ 0 & \sim \end{cases}$$

$$\delta_{mm} = \begin{cases} 1 & \text{si mes es mes movilizado} \\ 0 & \sim \end{cases}$$

$$\delta_{ml} = \begin{cases} 1 & \text{si mes es mes largo} \\ 0 & \sim \end{cases}$$

Tabla 5 – Modelo de Validaciones de Buses

Modelo de Validaciones de Buses	Coeficientes no estandarizados		Coeficientes tipificados			Estadísticos de colinealidad	
	B	Error típ.	Beta	t	Sig.	Tolerancia	FIV
(Constante) a_o	17,51	12,21		1,43	0,157		
Enero a_e	-10,72	1,77	-0,28	-6,06	0	0,83	1,205
Febrero a_f	-23,94	1,75	-0,61	-13,68	0	0,847	1,18
Julio a_j	-3,99	1,59	-0,11	-2,50	0,015	0,868	1,152
LongredBus a_{lrb}	0,02	0,002	0,51	9,63	0	0,616	1,624
DiasLaborales a_{dl}	1,54	0,35	0,19	4,44	0	0,903	1,107
Frecuencia a_{frec}	1,32	0,31	0,26	4,27	0	0,475	2,106
Tarifa a_{tar}	-0,07	0,01	-0,47	-7,23	0	0,403	2,484
MesMovilizado a_{mm}	-4,22	1,9	-0,1	-2,22	0,03	0,885	1,13
MesLargo a_{ml}	2,46	1,00	0,12	2,46	0,017	0,725	1,38
Observaciones	65						
R cuadrado	,905						
R cuadrado corregida	,889						
Error típ. de la estimación	3,462						
F(9 , 56)	58,958						
P-Valor	,000						

Es posible observar que todas las variables son significativas con un nivel de confianza de 95%. Como es de esperar, las variables dummies para los meses de Enero, Febrero y Julio, así como el Mes Movilizado dan valores negativos. Del mismo modo se puede observar una relación de signo positivo entre la demanda y la Longitud de Red Bus, los días laborales dentro del mes y la Frecuencia.

El parámetro de la tarifa tiene un valor de -0,07, es decir que al aumentar el precio de la tarifa en 1 peso, disminuye la demanda mensual en 70.000 validaciones. Importante notar que este valor es mayor al caso de la demanda agregada, motivo de esto puede ser que los usuarios de Buses son más sensibles al aumento de tarifa, haciendo que los mismos busquen alternativas de transporte, o bien se vea reflejado en un aumento en la evasión (no captada por los datos pasivos de validaciones de las tarjetas Bip!).

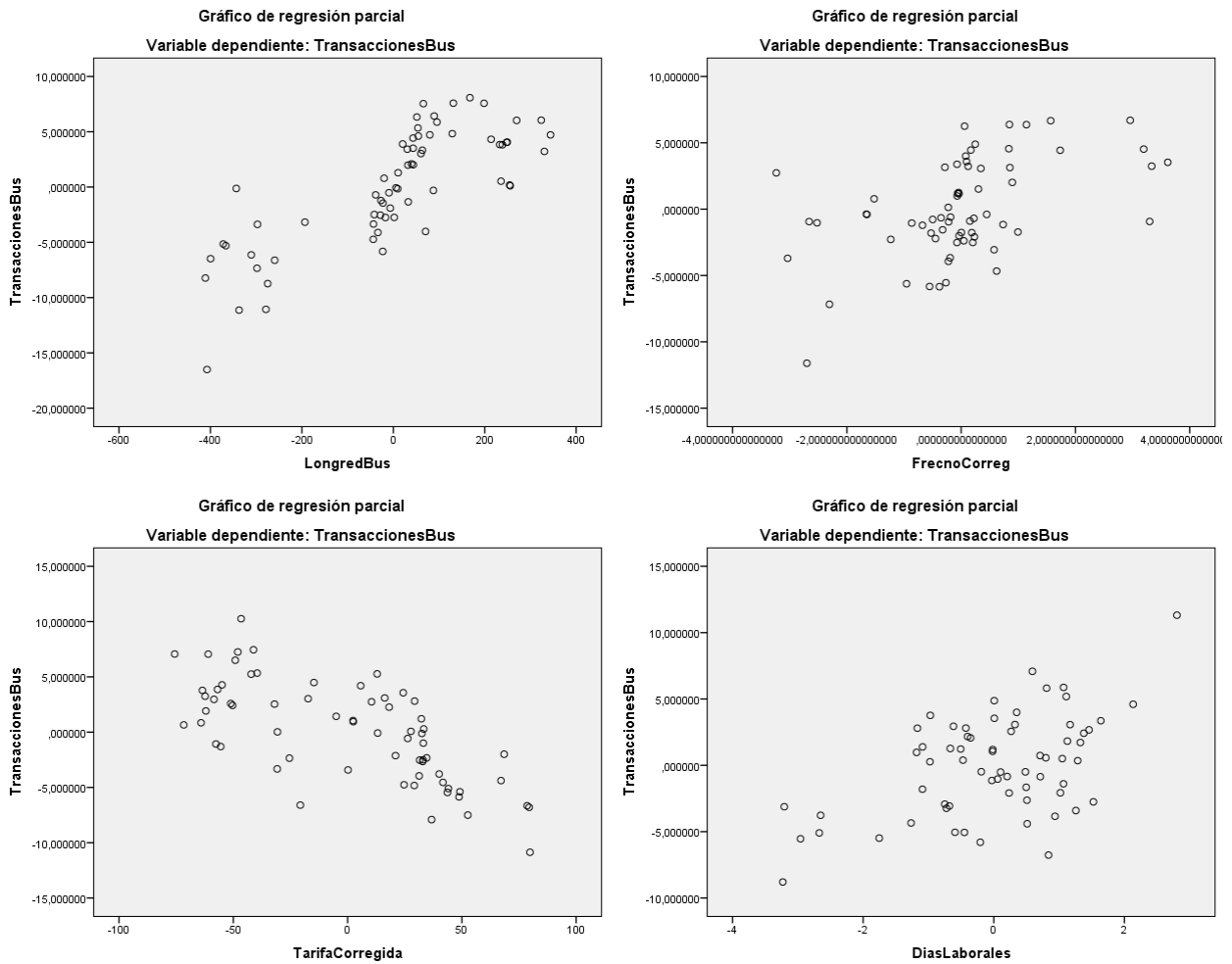


Figura 24 – Gráficos de Regresiones Parciales modelo de Validaciones de Buses

Al igual que en el modelo agregado (buses + metro), es posible ver una relación lineal y positiva para los gráficos de regresión parcial de longitud de red bus, días laborales y frecuencia. Para el gráfico de tarifa se obtiene una relación lineal negativa. Estos resultados se aprecian en los resultados del modelo, y permiten comprobar el supuesto de linealidad de la regresión, ya que ninguno de los gráficos muestra una evidente relación no lineal.

Los factores de inflación de varianza (FIV) son todos muy menores a 10, por lo que el modelo no presenta problemas de multicolinealidad

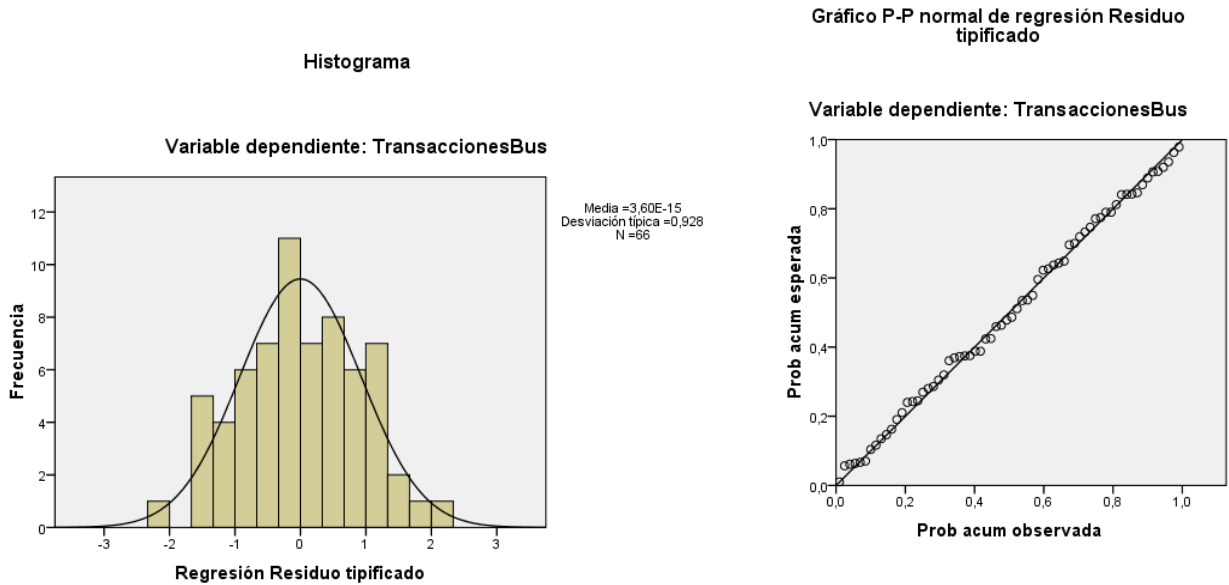


Figura 25 – Histograma y Gráfico P-P normal modelo de buses

Observando los gráficos de la figura 25, es válido decir que los residuos se distribuyen normalmente y con media cero, por lo que se respeta el supuesto de Normalidad de los Residuos.

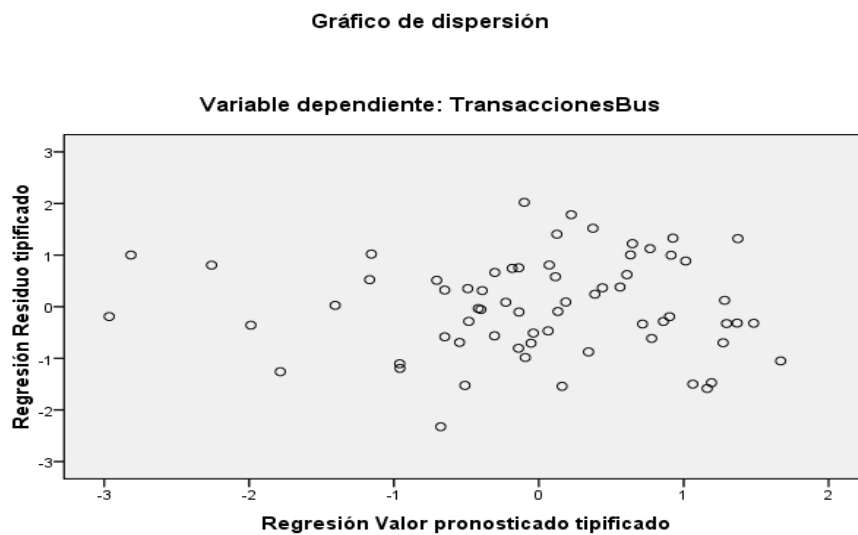


Figura 26 – Gráfico de Dispersión modelo de validaciones de Buses

El gráfico de dispersión no deja claro que las varianzas sean heterogéneas. Los pronósticos mayores que la media están más concentrados que los pronósticos

menores que la media, sin embargo no se ve una tendencia fuerte de manera de poder rechazar la homocedasticidad de forma concluyente.

Finalmente el test de Independencia de los residuos, entrega un valor para el estadístico de Durbin Watson de 1,013. Como se mencionó anteriormente, este valor debiese estar entre 1,5 y 2,5 para aceptar independencia entre los residuos.

4.4.5 Modelo de Validaciones de Metro

A continuación se muestra en la Tabla 6 el modelo cuya forma funcional se muestra en la Ecuación 13 y que corresponde a validaciones de Metro. En este modelo se cambiaron aquellas variables de oferta de Buses por oferta de metro, y a diferencia del modelo agregado y sólo de buses, el Desempleo se observa como variable predictora a un nivel de significancia del 99%.

El R cuadrado corregido es de 0,942, por lo que se puede decir que un 94,2% de la variación de la demanda de transacciones en Metro es explicada por estas variables.

$$Y = a_0 + a_e \delta_e + a_f \delta_f + a_j \delta_j + a_{mm} \delta_{mm} + a_{ckmM} CkmM + a_{dl} DL + a_{des} Des + a_{tar} Tar + \varepsilon \quad (12)$$

En donde la variable dependiente Y corresponde a las validaciones o transacciones de buses, y las variables independientes y parámetros están definidos a continuación.

CkmM: Coche Km Metro (millones de km)

DL: Días Laborales

Des: Desempleo (Porcentaje)

Tar: Tarifa (Pesos)

Las variables mudas son las mismas utilizadas en el modelo agregado (buses + metro).

$$\delta_e = \begin{cases} 1 & \text{si mes es enero} \\ 0 & \sim \end{cases}$$

$$\delta_f = \begin{cases} 1 & \text{si mes es febrero} \\ 0 & \sim \end{cases}$$

$$\delta_j = \begin{cases} 1 & \text{si mes es junio} \\ 0 & \sim \end{cases}$$

$$\delta_{mm} = \begin{cases} 1 & \text{si mes es mes movilitado} \\ 0 & \sim \end{cases}$$

Escriba aquí la ecuación.

Tabla 6 – Modelo Validaciones Metro

	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.	Estadísticos de colinealidad	
	B	Error típ.	Beta			Tolerancia	FIV
(Constante)	42,76	3,84		11,12	,00		
Enero	-9,24	,61	-,45	-15,24	,00	,96	1,04
Febrero	-15,58	,66	-,76	-23,60	,00	,81	1,23
Julio	-1,81	,56	-,10	-3,24	,00	,96	1,04
MesMovilizado	-,86	,70	-,04	-1,23	,22	,88	1,13
Desempleo	-106,92	18,40	-,24	-5,81	,00	,50	1,98
DiasLaborales	,77	,13	,19	5,98	,00	,87	1,15
CocheKmMetro	1,67	,32	,38	5,24	,00	,17	6,05
TarifaCorregida	-,02	,01	-,31	-3,74	,00	,13	7,83

Observaciones	68
R cuadrado	0,949
R cuadrado corregida	0,942
Error típ. de la estimación	1,280
F(9 , 59)	122,482
P-Valor	0

El signo de los coeficientes se dan de acuerdo a lo esperado por la teoría económica, y no se presentan grandes sorpresas correspondientes a los signos. Aquellas variables dummies que representan los meses donde hay vacaciones, dan las tres de signo negativo, es decir, existe una relación de esos meses con una baja en la demanda de Metro.

Se observa además una relación positiva en la oferta Coche-Km de Metro con respecto a la demanda, situación que había sido advertida con anterioridad en estudios realizados por Metro S.A. y comentado en reuniones con ellos.

El desempleo entrega un signo negativo como es de esperar, debido a que a un mayor desempleo existe una menor cantidad de viajes obligatorios que hacer. De este modo un aumento en la cifra de desempleo de un 1% se relaciona con una baja en la demanda mensual de Metro de 1,06 millones de validaciones.

Finalmente se observa una correlación negativa entre el precio y el número de validaciones, como era de esperar.

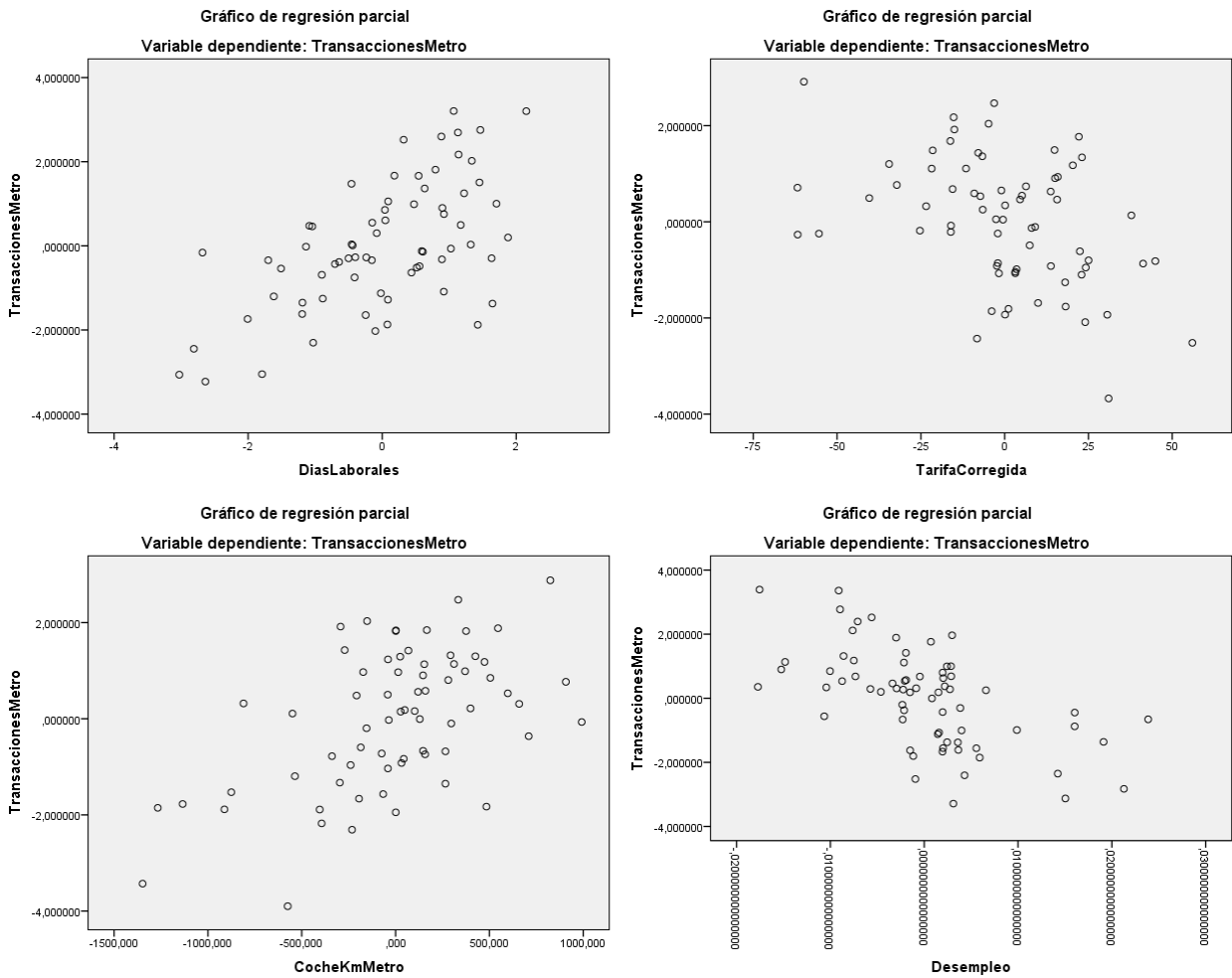


Figura 27 – Gráficos de regresiones parciales modelo de validaciones Metro

Los gráficos de regresión parcial muestran que existe una relación lineal y positiva entre las transacciones Metro y los días laborales y con la oferta Coche kilómetro de Metro. Como se mencionó se logra ver además el signo negativo de la relación entre transacciones Metro y Desempleo, y entre transacciones y Tarifa.

Al observar los valores de FIV, se observan que todos los valores son menores a 10, y los únicos valores que presentan un factor mayor a 5 son Coche-Km Metro y Tarifa. Si bien se observa que existe cierto nivel de correlación entre ambas variables, no es lo suficiente como para que el modelo presente problemas de multicolinealidad.

El estadístico de Durbin Watson para este modelo da un valor de 1,853, por lo que se acepta el supuesto de que existe independencia entre los residuos.

De la figura 28 se desprende que no existen problemas evidentes con el supuesto de normalidad de los residuos en este modelo. El histograma pudiese dar para dudar de este supuesto, sin embargo el gráfico P-P de regresión Residuo tipificado no es concluyente.

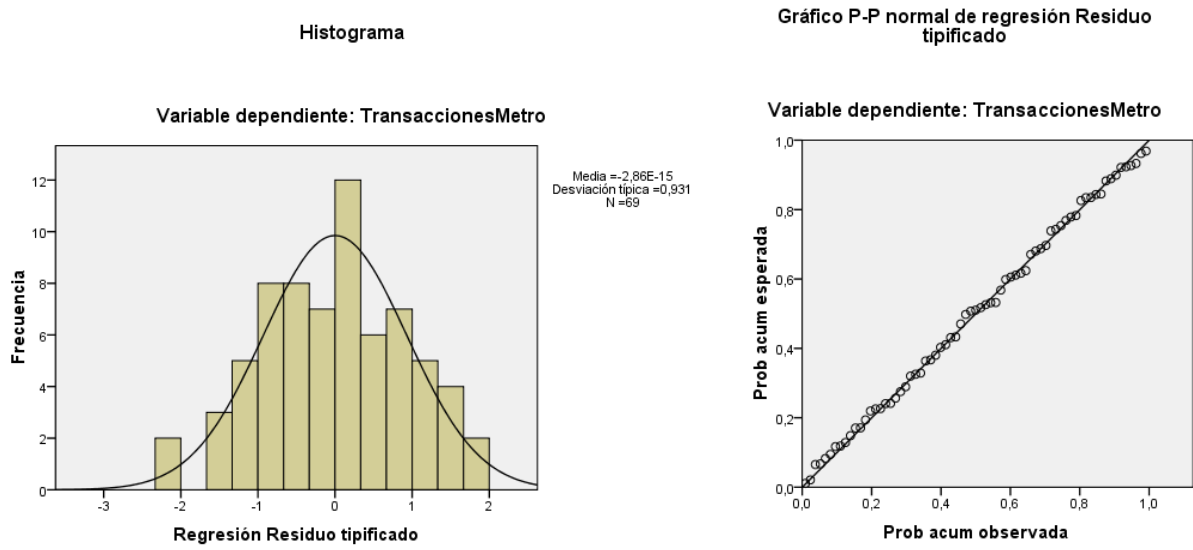


Figura 28 – Histograma y Gráfico P-P normal modelo validaciones Metro

De la Figura 29 se puede desprender que existe un cierto nivel de heterogeneidad en las varianzas de los residuos.

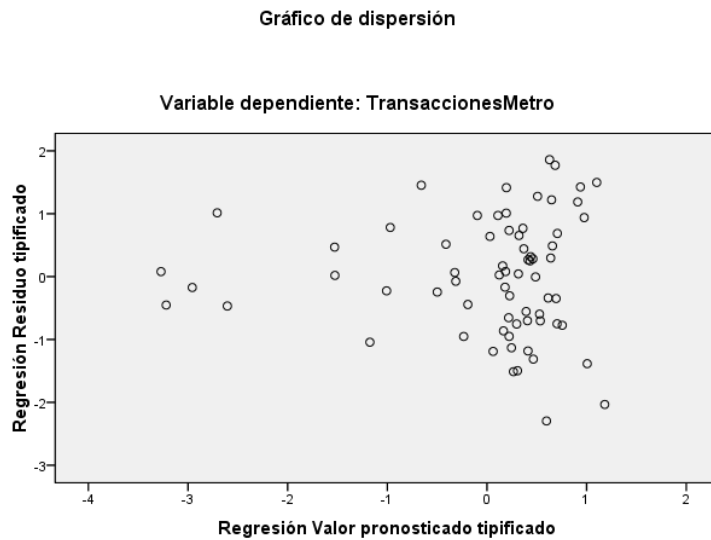


Figura 29 – Gráfico de Dispersión modelo de Validaciones Metro

4.4.6 Comparación modelos Lineales

Tabla 7 – Comparación modelos lineales de validaciones

	Agregado (buses+metro)			Buses			Metro		
	Coeficientes no estandarizados		Sig.	Coeficientes no estandarizados		Sig.	Coeficientes no estandarizados		Sig.
	B	Error típ.		B	Error típ.		B	Error típ.	
(Constante)	44,153	14,291	0,003	17,513	12,218	0,157	42,76	3,84	0,000
Enero	-18,899	2,068	0,000	-10,719	1,768	0,000	-9,24	0,62	0,000
Febrero	-40,603	2,047	0,000	-23,944	1,750	0,000	-15,58	0,64	0,000
Julio	-6,735	1,861	0,001	-3,990	1,591	0,015	-1,81	0,56	0,002
LongredBus	0,018	0,003	0,000	0,021	0,002	0,000			
DiasLaborales	2,418	0,406	0,000	1,538	0,347	0,000	0,77	0,13	0,000
Frecuencia buses	1,532	0,361	0,000	1,319	0,309	0,000			
Tarifa	-0,047	0,011	0,000	-0,070	0,010	0,000	-0,023	0,01	0,001
MesMovilizado	-4,907	2,221	0,031	-4,223	1,899	0,030	-0,86	0,70	0,257
MesLargo	3,293	1,171	0,007	2,463	1,001	0,017			
Desempleo							-106,92	18,621	0,000
Coche-km Metro							1,67	0,000	0,000
Observaciones	65			65			65		
R cuadrado	0,933			0,905			0,949		
R cuadrado corregida	0,922			0,889			0,942		
Error típ. de la estimación	4,049			3,462			1,280		
F(9 , 56)	86,829			58,958			122,482		
P-Valor	0			0			0		

De la Tabla 7 es posible desprender varios resultados con respecto a los coeficientes y ajustes obtenidos en los distintos modelos. El primero de estos resultados sería notar los ajustes de las regresiones, obteniendo desde mayor a menor ajuste el modelo de Metro, modelo agregado y finalmente modelo de Buses.

Otro resultado importante está relacionado con los coeficientes de la tarifa. Se observa que el modelo de buses es mucho más susceptible a cambios en la tarifa (-0.07) que el modelo agregado (-0.047) y únicamente de metro (-0.023). Una posible explicación para esto son los niveles de evasión que existen en los buses que no se replican del mismo modo en el Metro.

Se aprecian además que las variables mudas tienen el mismo efecto en los tres modelos, y el único cambio es en cuanto a la magnitud de este efecto. Lo mismo sucede con los días laborales.

4.4.7 Comparación modelos normalizados y sin normalizar

En esta sección se comparan los modelos utilizando datos normalizados y sin normalizar. Los datos normalizados intentan homogenizar los meses, normalizando la demanda a meses equivalentes de 30 días. De este modo no se utiliza variables que expliquen el tamaño del mes (30 o 31 días), y las variables normalizadas son la variable dependiente junto con la cantidad de días laborales mensuales, y la oferta coche kilometro mensual de Metro. Las Tablas 8, 9 y 10 muestran los resultados de esta comparación en los modelos.

Tabla 8 – Comparación modelos agregado de validaciones Datos normalizado y sin normalizar

Modelo Agregado	No Normalizado				Normalizado			
	B	Error típ.	Sig.	FIV	B	Error típ.	Sig.	FIV
(Constante)	44,15	14,29	,003		46,02	14,20	,002	
Enero	-18,90	2,06	,000	1,20	-21,54	1,90	,000	1,06
Febrero	-40,60	2,04	,000	1,18	-37,33	2,01	,000	1,18
Julio	-6,74	1,86	,001	1,15	-6,50	1,74	,000	1,05
Días Laborales	2,42	,40	,000	1,10	2,44	,40	,000	1,15
Frecuencia	1,53	,36	,000	2,10	1,46	,35	,000	2,09
Tarifa	-,047	,01	,000	2,48	-,048	,01	,000	2,47
LongredBus	,018	,00	,000	1,62	,018	,002	,000	1,62
MesMovilizado	-4,90	2,22	,031	1,13	-4,78	2,17	,032	1,12
MesLargo	3,29	1,17	,007	1,38				
R	0,966				0,957			
R cuadrado	0,933				0,915			
R cuadrado corregida	0,922				0,903			
Error típ. de la estimación	4,050				3,982			
Durbin-Watson	1,436				1,380			

Se puede observar que los coeficientes no varían en gran magnitud ni tampoco en su signo, lo que lleva a pensar que ambas metodologías son correctas en cuanto a la teoría económica que hay detrás. Sin embargo es importante notar que el modelo no normalizado muestra un mejor ajuste (0,922 contra 0,903), un mejor coeficiente de Durbin-Watson (1,436 contra 1,380). La diferencia entre los factores de inflación de varianza es baja en favor al modelo normalizado, pero ninguno de los modelos presenta problemas aparentes de colinealidad. Por estos motivos es que se optó por utilizar el modelo agregado no normalizado por sobre el modelo agregado normalizado.

Tabla 9 – Comparación modelo validaciones de buses datos no normalizado y sin normalizar

Modelo	Normalizado				No Normalizado			
	B	Error típ.	Sig.	FIV	B	Error típ.	Sig.	FIV
(Constante)	19,075	12,125	,121		17,513	12,218	,157	
Enero	-12,646	1,629	,000	1,061	-10,719	1,768	,000	1,205
Febrero	-21,914	1,720	,000	1,183	-23,944	1,750	,000	1,180
Julio	-3,667	1,491	,017	1,049	-3,990	1,591	,015	1,152
DiasLaborales	1,551	,348	,000	1,157	1,538	,347	,000	1,107
LongredBus	,021	,002	,000	1,622	,021	,002	,000	1,624
Frecuencia	1,263	,302	,000	2,092	1,319	,309	,000	2,106
Tarifa	-,070	,009	,000	2,477	-,070	,010	,000	2,484
MesMovilizado	-4,189	1,858	,028	1,122	-4,223	1,899	,030	1,130
MesLargo					2,463	1,001	,017	1,380
R	,941 ^a				,951 ^a			
R cuadrado	,885				,905			
R cuadrado corregida	,869				,889			
Error típ. de la estimación	3,400				3,463			
Durbin-Watson	,969				1,013			

Para el modelo de validaciones de buses (Tabla 9), se puede observar, al igual que en el caso anterior, una mejora en el estadístico de Durbin-Watson y en el ajuste del modelo no normalizado por sobre el modelo normalizado (1,013 contra 0.969 y 0.889 contra 0.869 respectivamente). La magnitud y el signo de los coeficientes son muy parecidos, al igual que los valores de FIV. Por estos motivos se decidió utilizar el modelo no normalizado en desmedro del modelo normalizado para buses.

Del mismo modo que las dos comparaciones mostradas, para el caso del modelo de validaciones de Metro (Tabla 9), se puede observar que el modelo de Metro no normalizado presenta un mayor ajuste (0,946 contra 0,931), un mejor estadístico de Durbin-Watson (1,993 contra 1,937), un menor FIV en aquellas variables que podrían ser candidatas a tener problemas de colinealidad (Coche kilómetro Metro y Tarifa) y del mismo modo que los modelos anteriores, los coeficientes presentan magnitudes y

signos parecidos para ambos modelos, por lo que para este caso también se considera que el modelo no normalizado es superior al modelo con variables normalizadas.

Tabla 10 – Comparación modelo validaciones de Metro datos normalizados y sin normalizar

Modelo Metro	Normalizado				No Normalizado			
	B	Error típ.	Sig.	FIV	B	Error típ.	Sig.	FIV
(Constante)	42,279	3,650	,000		43,309	3,670	,000	
Enero	-9,053	,586	,000	1,045	-9,055	,591	,000	1,049
Febrero	-14,608	,617	,000	1,159	-15,199	,659	,000	1,306
Julio	-2,339	,541	,000	1,052	-2,109	,546	,000	1,059
DiasLabNorm	,674	,131	,000	1,283	,670	,129	,000	1,240
LongredBus	-,003	,001	,027	4,616	-,003	,001	,024	4,484
Tarifa	-,020	,006	,002	8,426	-,021	,006	,001	8,093
ChKmMetNorm	,002	,000	,000	9,398	,002	,000	,000	9,047
Desempleo	-68,641	23,586	,005	3,507	-72,230	23,753	,003	3,511
R	,969 ^a				,976 ^a			
R cuadrado	,939				,952			
R cuadrado corregida	,931				,946			
Error típ. de la estimación	1,234				1,242			
Durbin-Watson	1,937				1,993			

4.4.8 Comparación de formas funcionales para modelo de validaciones

Modelo Agregado de Validaciones

Una vez analizados los modelos lineales como base, a continuación se agregan modelos Lineal-Log, Log-Log y Log-Lineal. Las distintas formas funcionales de los modelos fueron evaluadas en base a su bondad de ajuste y el test de Ramsay RESET (Regression specification error test). El ajuste del modelo, en término de R cuadrado, es únicamente comparable cuando la variable dependiente es la misma. Es por este motivo que se podrán comparar de mediante esta forma el modelo lineal con el lineal-log, y los modelos log-log con el modelo log-lineal.

Tabla 11 – Comparación Formas Funcionales modelo agregado de Validaciones

	Lineal	Lineal-Log	Log-Log	Log-Lineal
(Constante)	44,152 ** <i>14,29</i>	-305,164 *** <i>64,076</i>	1,891 *** <i>0,433</i>	4,283 *** <i>0,097</i>
Enero	-18,899 *** <i>2,067</i>	-18,718 *** <i>2,049</i>	-0,130 *** <i>0,014</i>	-0,132 *** <i>0,014</i>
Febrero	-40,603 *** <i>2,046</i>	-40,425 *** <i>2,031</i>	-0,315 *** <i>0,014</i>	-0,317 *** <i>0,014</i>
Julio	-6,735 *** <i>1,861</i>	-6,795 *** <i>1,834</i>	-0,046 *** <i>0,012</i>	-0,046 *** <i>0,013</i>
LongredBus	0,018 *** <i>0,002</i>	41,759 *** <i>5,870</i>	0,279 *** <i>0,040</i>	0,000 *** <i>0,000</i>
DiasLaborales	2,418 *** <i>0,405</i>	49,390 *** <i>8,065</i>	0,333 *** <i>0,054</i>	0,016 *** <i>0,003</i>
FrecnoCorreg	1,531 *** <i>0,361</i>	35,295 *** <i>7,998</i>	0,251 *** <i>4,638</i>	0,011 *** <i>0,002</i>
TarifaCorregida	-0,0466 *** <i>0,011</i>	-20,916 *** <i>5,643</i>	-0,131 *** <i>-3,562</i>	0,000 *** <i>0,000</i>
MesMovilizado	-4,906 ** <i>2,221</i>	-5,035 ** <i>2,190</i>	-0,032 ** <i>-2,172</i>	-0,031 ** <i>0,015</i>
MesLargo	3,292 *** <i>1,171</i>	3,287 *** <i>1,154</i>	0,022 *** <i>2,838</i>	0,022 *** <i>0,008</i>
Observaciones	65	65	65	65
R cuadrado	,933	0,935	0,946	0,944
R cuadrado corregida	,922	0,925	0,937	0,935
Error típ. de la estimación	4,049	3,99	0,0269	0,02742
F(9 , 56)	86,829	89,637	108,943	105,084
P-Valor	,000	0	0	0

** P<0.05, *** P<0.01; Valores en cursiva corresponden a Desviación estándar

Tomando el primer par de modelos, Lineal y Lineal-Log, se puede observar una pequeña mejora en el ajuste del segundo modelo con respecto al primero (0.925 en comparación a 0.922).

Realizando esta misma comparación entre los modelos Log-Log y Log-Lineal, es posible observar una pequeña diferencia en el ajuste de R cuadrado en el modelo Log-Log, siendo este 0.937 en comparación a R cuadrado del modelo Log-Lineal de 0.935.

Es importante notar que la significancia de las variables para los 4 modelos son muy similares, observándose una única diferencia en la significancia de la constante del modelo lineal (significativa al 95% de confianza en dicho modelo, en comparación a el 99% de los otros modelos).

Siendo superiores a sus pares el modelo lineal-log y Log-Log, se procedió a realizar un test de Ramsay RESET por sus siglas en inglés (Regression Equation Specification Error Test), el cual permite saber si alguna combinación no lineal de los valores pronosticados ayudan a explicar la variable dependiente.

Lo que se busca con este test, es saber si los modelos se encuentran bien especificados, y además sirve como herramienta adicional para decidir si alguno de los modelos es superior al otro en este aspecto. Para ambos modelos resultó ser no significativa la adición de nuevas variables no lineales, obteniendo para el modelo lineal-log una $Prob > F$ de 0,739 y para el modelo Log-Log una $Prob > F$ de 0,966.

Se desprende de este resultado que el modelo logarítmico tiene una menor probabilidad de estar mal especificado y de que se hayan omitido variables no lineales significativas para el modelo, por lo que se decide utilizar este mismo modelo para los escenarios mostrados en las secciones siguientes de este trabajo.

4.4.9 Modelo Agregado Log-Log de Validaciones

A continuación se muestra que este modelo cumple con todos los supuestos de las regresiones, motivo por el cual fue elegido para realizar las predicciones y modelar distintos escenarios en el capítulo 5.

La ecuación de este modelo se muestra a continuación. Las variables mudas mantienen la misma notación que la utilizada en los modelos lineales.

$$\ln(Y) = a_0 + a_e \delta_e + a_f \delta_f + a_j \delta_j + a_{mm} \delta_{mm} + a_{ml} \delta_{ml} + a_{lrb} \ln(LRB) + a_{dl} \ln(DL) + a_{frec} \ln(Frec) + a_{tar} \ln(Tar) + \varepsilon \quad (13)$$

Tabla 12 – Modelo agregado (Bus + metro) Log-Log de Validaciones

	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.	Estadísticos de colinealidad	
	B	Error típ.	Beta			Tolerancia	FIV
(Constante) a_0	1,891	0,433		4,369	0,000		
Enero a_e	-0,130	0,014	-0,323	-9,415	0,000	0,821	1,218
Febrero a_f	-0,315	0,014	-0,781	-22,987	0,000	0,835	1,197
Julio a_j	-0,046	0,012	-0,124	-3,727	0,000	0,868	1,152
MesMovilizado a_{mm}	-0,032	0,015	-0,072	-2,172	0,034	0,883	1,132
MesLargo a_{ml}	0,022	0,008	0,104	2,838	0,006	0,725	1,380
LongRedBusLog a_{trb}	0,279	0,040	0,273	7,028	0,000	0,638	1,567
DiasLabLog a_{dl}	0,333	0,054	0,200	6,107	0,000	0,903	1,107
TarifaCorregLog a_{tar}	-0,136	0,038	-0,177	-3,562	0,001	0,392	2,548
FrecNoCorrLog a_{frec}	0,251	0,054	0,213	4,638	0,000	0,458	2,185
Observaciones	65						
R cuadrado	0,946						
R cuadrado corregida	0,937						
Error típ. de la estimación	0,0269						
F(9 , 56)	108,943						
P-Valor	0						

Se puede ver el P-valor de 0,000 que indica que las variables explicativas están relacionadas con la variable dependiente. Además se observa que la variación de la demanda está explicada en un 93.7% por las variables seleccionadas.

Además de esto, todas las variables son significativas a un 99% de confianza, menos el MesMovilizado que es significativa a un 95% de confianza.

Al igual que en el modelo lineal, la interpretación económica de los signos de las correlaciones son intuitivas y están de acuerdo a la literatura.

Es posible observar que no hay problemas de colinealidad, debido a que todos los valores FIV son menores a 10.

En la Figura 31 se observa que no es posible rechazar el test de normalidad de los residuos. En el primer gráfico existe una pequeña discrepancia, sin embargo no es concluyente, al igual que en el segundo gráfico. Sin embargo no es suficiente para rechazar la hipótesis de normalidad de los residuos.

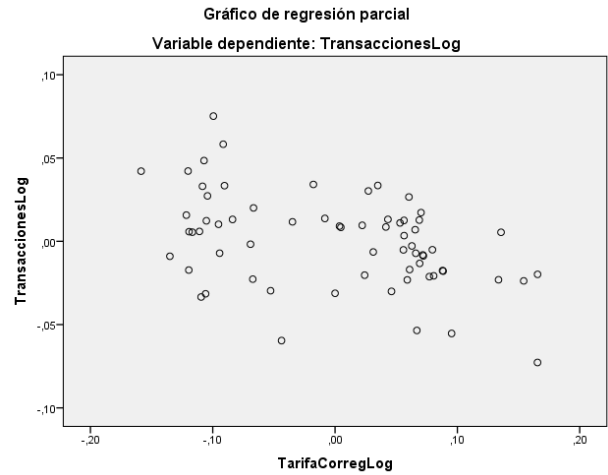
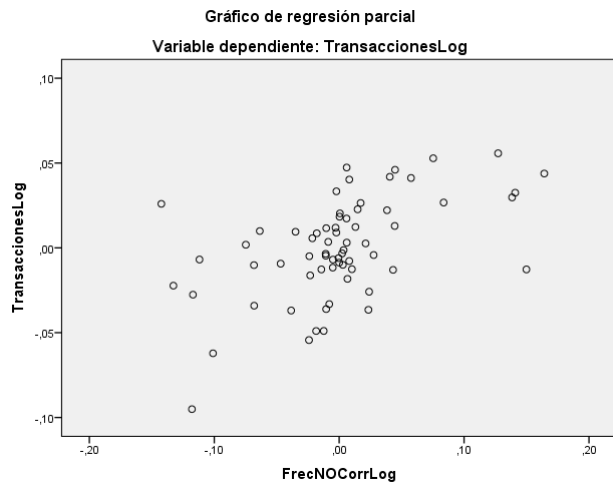
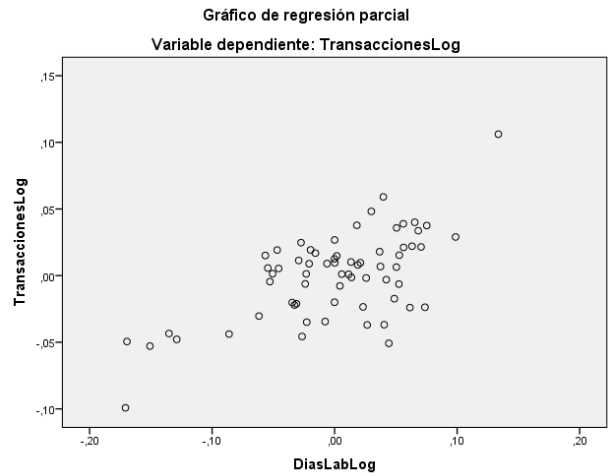
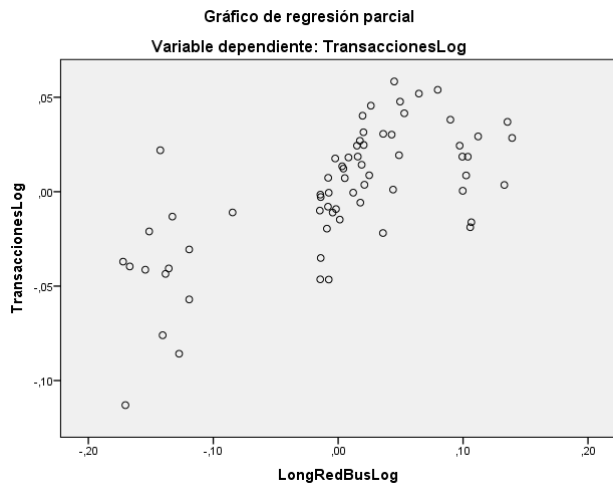


Figura 30 – Gráficos de regresiones parciales modelo logarítmico de Validaciones

Los gráficos de regresión parcial muestran que no se viola el supuesto de linealidad de la regresión múltiple, observándose relaciones líneas positivas para transacciones con días laborales, transacciones con frecuencia y transacciones con longitud red bus, y una relación negativa entre transacciones y tarifa.

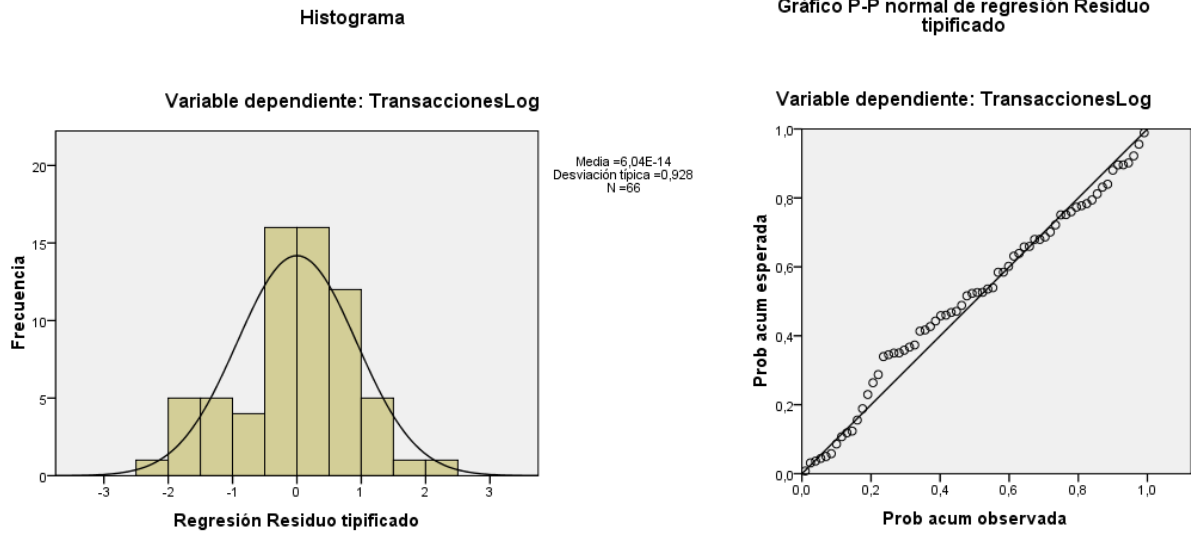


Figura 31 – Histograma y gráfico P-P normal modelo logarítmico de Validaciones

El estadístico Durbin Watson es de 1.488, valor muy cercano a 1,5. De este modo no se rechaza la hipótesis de independencia de los residuos.

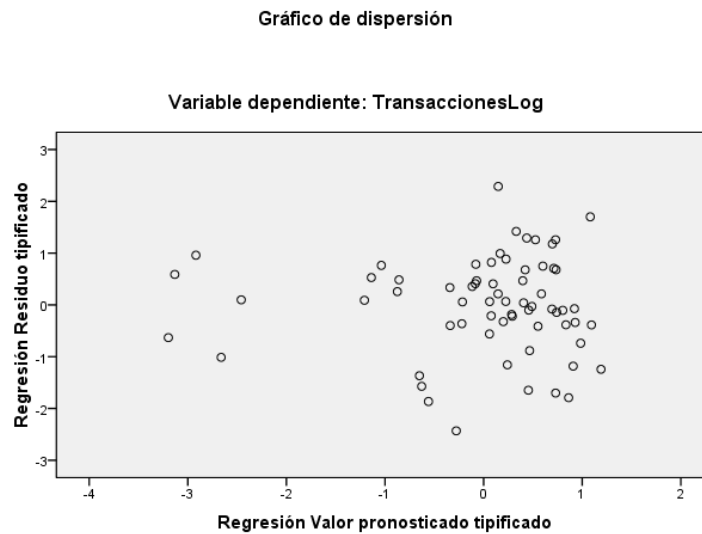


Figura 32 – Gráfico de Dispersión modelo logarítmico de Validaciones

Se puede observar en la Figura 32 que el gráfico de dispersión no muestra una tendencia clara, sin embargo es posible observar 5 claros outliers que no debiesen afectar de sobremanera la distribución. De este modo no se rechaza tampoco la hipótesis de homocedasticidad.

Junto con este modelo, y en los que vienen a continuación, se mostrarán los resultados del modelo realizando un cambio con respecto a las variables explicativas. Los Kilómetros Comerciales ofrecidos por Transantiago dependen de la frecuencia de los servicios así como de la longitud de la red de buses. En conversaciones con Transantiago se pidieron los resultados no utilizando la variable frecuencia, debido a que el cálculo de la misma no es directo, por lo que la utilización de kilómetros comerciales es de más fácil y directo acceso. Los resultados de este modelo se muestran en la Tabla 13. Al realizar este cambio se debe además eliminar la variable muda de mes largo, debido a que la oferta de kilómetros comerciales incluye esto.

Tabla 13 – Modelo log-log validaciones (Km Comerciales)

	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.	Intervalo de confianza de 95,0% para B		Estadísticos de colinealidad	
	B	Error típ.	Beta			Límite inferior	Límite superior	Tolerancia	FIV
(Constante)	4,84	,53		9,16	,00	3,79	5,90		
Enero	-,14	,01	-,35	-11,29	,00	-,17	-,12	,90	1,12
Febrero	-,26	,01	-,63	-19,24	,00	-,29	-,23	,79	1,27
Julio	-,04	,01	-,11	-3,67	,00	-,06	-,02	,95	1,05
MesMovilizado	-,04	,01	-,08	-2,64	,01	-,07	-,01	,88	1,14
DiasLabLog	,21	,05	,12	4,02	,00	,10	,31	,89	1,13
logKmCom	,35	,08	,21	4,29	,00	,19	,52	,34	2,94
logCocheKm	,43	,08	,49	5,30	,00	,27	,59	,10	10,25
TarifaCorregLog	-,44	,07	-,56	-6,63	,00	-,57	-,31	,12	8,42

R 0,974
R cuadrado 0,949
R cuadrado corregida 0,942
Error típ. de la estimación 0,026

Comparando ambos modelos es posible observar que el modelo utilizando kilómetros comerciales presenta un mejor ajuste, pero esta comparación no es directa, debido a que el modelo que utiliza los kilómetros comerciales también se le agregó la variable de oferta de Metro (Coche Kilómetros). Esto se hizo debido a que los valores de FIV están en el límite de lo aceptable, mientras que en el modelo base estos valores sobrepasaban este límite por un amplio margen.

En caso de agregar la variable de oferta de metro al modelo presentado en la Tabla 12, el ajuste de dicho modelo es 0.966, superior al 0.942 obtenido por el modelo observado en la Tabla 13.

4.4.10 Modelo log-log de Validaciones de Buses

En la Tabla 14, se muestran los resultados del modelo log-log de validaciones de buses. Al igual que en el caso de los modelos agregados (buses + metro), se encuentra con esta forma funcional un mejor ajuste en comparación al modelo lineal.

Tabla 14 – Modelo log-log de validaciones de buses

	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.	Intervalo de confianza de 95,0% para B	
	B	Error típ.	Beta			Límite inferior	Límite superior
(Constante)	,637	,596		1,069	,290	-,557	1,832
Enero	-,112	,019	-,256	-5,884	,000	-,150	-,074
Febrero	-,286	,019	-,653	-15,124	,000	-,324	-,248
Julio	-,044	,017	-,108	-2,558	,013	-,078	-,009
MesMovilizado	-,043	,020	-,088	-2,092	,041	-,083	-,002
LongRedBusLog	,501	,055	,454	9,181	,000	,392	,611
DiasLaborales Log	,329	,075	,182	4,382	,000	,179	,479
Tarifa Log	-,332	,053	-,399	-6,331	,000	-,438	-,227
Frecuencia log	,345	,074	,270	4,629	,000	,195	,494
MesLargo	,026	,011	,113	2,438	,018	,005	,048

Observaciones 65
R cuadrado 0,912
R cuadrado corregida 0,898
Error típ. de la estimación 0,037
F(9 , 56)
P-Valor

Del mismo modo que en el caso de validaciones de metro y buses, se realizó un modelo que utilizara la variable de kilómetros comerciales en vez de frecuencia y longitud red de bus. Este modelo se muestra a continuación en la Tabla 15.

En este caso, el modelo que utiliza frecuencia y longitud de red buses muestran un mejor resultado, obteniendo un ajuste de 0.898 en comparación al ajuste del modelo de la Tabla 15 que es de 0.872.

Tabla 15 – Modelo log-log de validaciones de buses (Km Comerciales)

	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.	Estadísticos de colinealidad	
	B	Error típ.	Beta			Tolerancia	FIV
(Constante)	1,63	,45		3,60	,00		
Enero	-,10	,02	-,22	-4,91	,00	,95	1,05
Febrero	-,19	,02	-,43	-8,76	,00	,79	1,27
Julio	-,05	,02	-,12	-2,75	,01	,96	1,04
MesMovilizado	-,03	,02	-,06	-1,42	,16	,90	1,12
DiasLabLog	,19	,08	,10	2,25	,03	,91	1,09
TarifaCorregLog	-,25	,04	-,28	-6,29	,00	,92	1,09
logKmCom	1,05	,09	,58	11,59	,00	,75	1,33

R 0,941
R cuadrado 0,886
R cuadrado corregida 0,872
Error típ. de la estimación 0,042

4.4.11 Modelo log-log de validaciones de Metro

El resultado del modelo log-log de validaciones de Metro se muestra en la Tabla 16. Con un R cuadrado de 0.956 tiene un mejor ajuste que la forma funcional base de este modelo.

Tabla 16 – Modelo log log de validaciones de Metro

	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.	Intervalo de confianza de 95,0% para B	
	B	Error típ.	Beta			Límite inferior	Límite superior
(Constante)	1,425	,268		5,323	,000	,890	1,960
Enero	-,184	,011	-,426	-16,421	,000	-,206	-,161
Febrero	-,350	,012	-,811	-29,107	,000	-,375	-,326
Julio	-,034	,010	-,086	-3,337	,001	-,055	-,014
MesMovilizado	-,012	,013	-,026	-,946	,348	-,038	,014
DiasLabLog	,309	,047	,177	6,544	,000	,215	,404
TarifaCorregLog	-,209	,055	-,254	-3,794	,000	-,319	-,099
CocheKmMetroLog	,270	,052	,297	5,182	,000	,166	,374
DesempleoLog	-,178	,028	-,228	-6,255	,000	-,235	-,121

Observaciones	65
R cuadrado	0,961
R cuadrado corregida	0,956
Error típ. de la estimación	0,023

4.4.12 Elasticidades

Una de las ventajas de utilizar el modelo log-log, es la directa obtención de las elasticidades, correspondientes al parámetro β de la regresión.

La elasticidad de la frecuencia obtenida en el modelo agregado es de 0,251, mientras que la correspondiente a la tarifa es de -0,136. Finalmente la elasticidad atribuida al alcance espacial de la red de bus es de 0,279.

En cuanto a las elasticidades del modelo sólo buses, encontramos una elasticidad precio más elevada, de -0.332. Lo mismo ocurre con la frecuencia de buses, cuya elasticidad es de 0.345. Que ésta elasticidades sea más elevada que en el modelo agregado no es algo raro, ya que la frecuencia de buses si bien afecta a todo el sistema, está ligada a los buses en particular, por lo que tiene un impacto más grande en el modelo propio de estos.

Para el modelo de Metro, se observa una elasticidad precio de -0.209, y la elasticidad correspondiente al desempleo es de -0.178.

Es importante notar que el uso y la comparación de estas elasticidades está restringido por una serie de propiedades, por lo que tiene sus limitaciones. Entender cuáles son estas limitaciones es importante para una correcta aplicación (Balcombe et al, 2004 y Littman, 2012) Dentro de éstas podemos nombrar que las elasticidades pueden variar según el propósito de viaje, por lo que en un modelo agregado su correcta comparación sería con un modelo similar. Del mismo modo es importante el periodo de tiempo en que se midió la demanda, las características individuales de la población estudiada. Las elasticidades además no son siempre transferibles en el tiempo y espacio, debido a que las preferencias pueden cambiar en el tiempo. Es por esto que una comparación con modelos separados por intervalos demasiado grandes puede no ser justa. Finalmente la forma funcional del modelo influye en la elasticidad, por lo que es deseable que la comparación de elasticidades sean entre modelos que compartan la misma forma funcional.

Además de esto, es importante notar que las elasticidades nombradas en el capítulo 2, corresponden a elasticidades con respecto a los viajes, y las acá nombradas son con respecto a las validaciones, otro motivo adicional para tener cuidado con respecto a su utilización y comparación. El concepto de validaciones no es algo que se encuentre en la literatura, pero resulta de vital importancia para las proyecciones financieras de Transantiago el poder estimarlo.

4.4.13 Resultados

Es importante notar la consistencia de los signos de las variables explicativas en los modelos, puesto que en gran parte de los modelos dan de forma acorde a lo esperado según la teoría económica. Es así como es posible observar para todos los modelos una relación negativa entre la Tarifa del transporte público y la cantidad de transacciones realizadas. Si bien esto se puede deber a un cambio en el modo de transporte del usuario, no es posible descartar que este sea el único motivo, sino que la disponibilidad a pagar puede disminuir, aumentando la evasión.

Con respecto a la Frecuencia de los buses y su signo positivo en relación a la demanda de validaciones, es de esperar que una mayor frecuencia disminuya los tiempos de espera y por ende aumente la calidad de servicio del transporte público, lo que a su vez puede conllevar a un aumento en la demanda o una disminución en la evasión (al cambiar la percepción de la calidad de servicio de la gente).

En reunión con la gente de Transantiago se pidió que se estudiara la baja permanente en cuanto a validaciones que se registraron desde la segunda semestre del año 2011. Motivo aparente de dicha disminución eran las marchas estudiantiles, las cuales agregaban incertidumbre tanto a que las rutas comerciales pudiesen desarrollarse de forma normal, como al tiempo de viaje de las mismas. Sin embargo luego de utilizar variables dummies para explicar este efecto, no se consiguió un modelo que diese una correlación negativa y significativa al respecto (ni a sus rezagos). Sin embargo un análisis al comportamiento de la frecuencia entrega nuevas luces con respecto a este fenómeno, puesto que se observa una baja continua de la misma. Esta baja se produce principalmente al crear nuevos servicios a partir de servicios antiguos. Esto si bien aumenta el alcance de la red buses, causa una baja en la calidad de servicio percibido por el usuario.

Una particularidad observada es que si bien el precio de la gasolina no dio significativo en los modelos, su signo para todos los casos probados en cuanto a modelos de validaciones de Buses fue negativo, lo que no es algo de esperar ni que vaya de la mano con teoría económica. Sin embargo este valor para los modelos de Metro si da positivo, es decir al haber un aumento en el precio de la gasolina, existe un aumento en los viajes de Metro. Una explicación a este fenómeno es que posiblemente entre los usuarios de buses haya más usuarios cautivos del transporte público, relativo a los usuarios que utilizan Metro, y no tienen como alternativa el uso de transporte particular para movilizarse. Este fenómeno no se presenta de forma tan marcado en los usuarios de Metro, y entre ellos puede existir un cambio de modo al haber una variación en el precio de la gasolina.

La inclusión de días laborales como variable generadora de viajes logró resultados de acuerdo a lo esperado, siendo significativo en un nivel de confianza de 99% para todos los modelos. Sin embargo el utilizar la cantidad de días sábados e interfiados como variable, además de los domingos y festivos no resulto de utilidad, debido a que solían tener valores muy cercanos en el mismo mes ambos valores, y el modelo es incapaz de distinguirlos entre sí, dando valores no significativos y una alta correlación.

Con respecto a la variable de mes movilizado, fue de gran utilidad para poder modelar los 4 meses en donde una inusual cantidad de colegios y universidades e encontraban

en paro, y si bien la utilidad que tiene la variable para poder estimar la demanda a futuro es baja, si ayuda a poder calibrar el modelo de manera correcta. Esta variable por otro lado si se podría utilizar en caso de que exista a futuro una movilización estudiantil de magnitud similar a la de 2011, ya que serviría como indicación de la magnitud de disminución en validaciones.

El cambiar las variables frecuencia y longitud de red bus por la variable kilómetros comerciales es algo válido debido a la facilidad para Transantiago de utilizar esta variable. Los R^2 son más bajos, sin embargo las diferencias no son tan sustanciales como para invalidar los resultados de estos modelos, y permiten una comparación más directa con respecto a resultados que manejan en Transantiago con respecto a estas variables.

4.5 Modelo de Viajes

En esta sección se mostrará el modelo cuya variable explicada son los Viajes. A diferencia de los modelos de validaciones, la data para estos modelos es menor, debido a que Transantiago sólo cuenta información de viajes, usando la definición actual de los mismos, desde el año 2009.

4.5.1 Matriz de Covarianzas

En la tabla 17 se muestra la matriz de covarianzas de las variables independientes utilizando los datos desde inicios del 2009 hasta diciembre 2012.

Tabla 17 – Matriz de Covarianza Modelo de Viajes

	Desempleo	KMComerciales	LongredBus	LongRedMetro	IA	Gasolina	DiasLaborales	Precipitaciones	PA	CocheKmMetro	Frecuencia	Tarifa
Desempleo	1											
KMComerciales	,47	1										
LongredBus	-,43	-,30	1									
LongRedMetro	-,88	-,40	,61	1								
IA	-,82	-,58	,56	,86	1							
Gasolina	-,84	-,39	,64	,93	,86	1						
DiasLaborales	-,02	,15	,03	-,04	-,12	-,07	1					
Precipitaciones	,14	,14	-,05	-,11	-,06	,00	-,02	1				
PA	-,83	-,57	,57	,87	1	,86	-,12	-,06	1			
CocheKmMetro	-,88	-,40	,52	,94	,84	,85	,10	-,13	,84	1		
Frecuencia	,86	,57	-,48	-,83	-,90	-,76	,00	,16	-,90	-,87	1	
Tarifa	-,92	-,57	,61	,93	,90	,91	-,07	-,06	,90	,91	-,90	1

Es posible observar un fenómeno que no se dio al analizar la matriz de correlaciones utilizando como variable de control las transacciones, y es que al disminuir el tamaño muestral se observan correlaciones mucho más elevadas. Ejemplo de eso es que la tarifa, muestra una correlación superior en valor absoluto a 0,9 con el desempleo, la longitud de la red metro, el índice automotor, el precio de la gasolina, el parque automotriz, la oferta coche-km de metro y la frecuencia de buses.

De manera similar ocurre con una serie de variables que fueron de gran importancia a la hora de explicar las validaciones, como es el caso de la frecuencia, que da una elevada correlación con desempleo, índice automotor, parque automotriz, oferta coche-km metro y la tarifa.

4.5.2 Resultado del Modelo de Viajes

En una primera aproximación por realizar un modelo capaz de explicar el impacto de distintas variables en la demanda de Transantiago, es posible observar problemas serios de colinealidad entre algunas de estas variables. Es por este motivo que se abordó este modelo realizando una comparación a la metodología actual que utiliza Transantiago para predecir la demanda, la cual consideran que es propia de la estacionalidad, y por tal motivo se asume constante en comparación al mismo mes del año anterior, con respecto a una nueva metodología que propone mejorar el resultado de este pronóstico.

De este modo, se presentan dos modelos en la Tabla 18, el primer modelo corresponde al modelo que utiliza únicamente variables que no dependen de decisiones de Transantiago, sino que son propias de la estacionalidad. Estas serían los días laborales, y las variables dummies representando los meses de enero, febrero y julio, y si correspondió a un mes movilizadado.

Es posible observar como en el modelo 1, todas las variables explicativas resultan significativas al 95%, no se observan problemas de multicolinealidad, y en cierto modo, es un modelo similar al utilizado por Transantiago hoy en día, dado que se asume una demanda mensual propia del mes en cuestión. Este modelo en si presentaría una pequeña mejora dado que agrega además variables que dan una mayor especificidad como la cantidad de días laborales.

Este modelo es capaz de explicar un 93.6% de la varianza en la demanda de los viajes (R^2 de 0,936).

Sin embargo, si a este mismo modelo se agregan las variables de oferta de buses y Metro, además de la tarifa, exista una mejora con respecto al modelo (R^2 de 0,964).

Es necesario hacer la pregunta, si esta mejora ¿es estadísticamente significativa o si únicamente aumenta el ajuste del modelo al aumentar la cantidad de variables explicativas? Para responder esto, se realizó un Test F-anidado, en donde se testea si la variable agregada al modelo 2, (que es igual al modelo 1 pero con una variable extra agregada), produce un cambio significativo en el resultado del modelo. Los resultados se muestran a continuación en la Tabla 19.

Es posible observar que en el modelo con variables de oferta de metro y buses, casi todas las variables son significativas al 95%, mientras que la tarifa es significativa al 90%.

Tabla 18 – Modelos lineales anidados de Viajes Transantiago

	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.	Estadísticos de colinealidad	
	B	Error típ.	Beta			Tolerancia	FIV
1 (Constante)	65,34	6,10		10,72	,00		
Enero	-14,12	1,25	-,43	-11,30	,00	,97	1,03
Febrero	-29,19	1,26	-,88	-23,18	,00	,96	1,04
Julio	-3,10	1,26	-,09	-2,45	,02	,95	1,05
MesMovilizado	-3,31	1,27	-,10	-2,60	,01	,94	1,06
DiasLaborales	1,47	,29	,19	5,00	,00	,97	1,03
2 (Constante)	39,49	9,46		4,18	,00		
Enero	-13,12	1,08	-,40	-12,17	,00	,88	1,14
Febrero	-23,59	1,56	-,71	-15,12	,00	,42	2,38
Julio	-3,46	1,04	-,10	-3,33	,00	,95	1,05
MesMovilizado	-3,53	1,08	-,11	-3,27	,00	,88	1,14
DiasLaborales	,88	,27	,11	3,21	,00	,75	1,33
KMComerciales	,61	,21	,14	2,85	,01	,39	2,56
CocheKmMetro	2,81	1,00	,31	2,79	,01	,07	13,46
TarifaCorregida	-,03	,02	-,21	-1,93	,06	,08	12,76

Tabla 19 – Cuadro comparativo resultados de modelos anidados

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio				
					Cambio en R cuadrado	Cambio en F	gl1	gl2	Sig. Cambio en F
1	,971 ^a	,943	,936	2,360120413	,943	134,997	5	41	,000
2	,982 ^b	,964	,957	1,933266449	,022	7,701	3	38	,000

La significancia en el cambio del estadístico F del modelo 1, indica que existe una mejora con respecto al modelo base (sin variables). Sin embargo la del modelo 2, es en comparación al modelo 1, es decir, se rechaza la hipótesis nula de que la nueva variable no mejora significativamente el modelo.

Notemos además que este modelo no presenta problemas de independencia de los residuos (estadístico DW de 1,818).

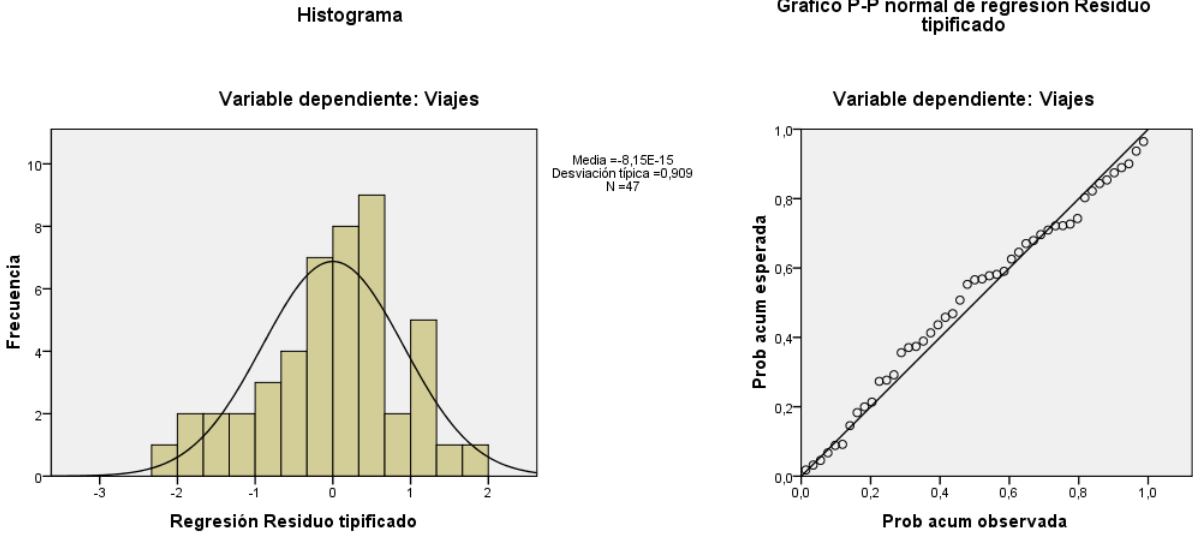


Figura 33 – Histograma y Gráfico P-P normal modelo de Viajes

En la Figura 33 se puede observar que existe normalidad en los residuos, y al revisar el gráfico de dispersión (Figura 34), se observa una nube sin una clara tendencia, dejando de lado los puntos de más a la izquierda correspondientes a los meses de Febrero y Enero. De este modo se acepta el supuesto de homocedasticidad.

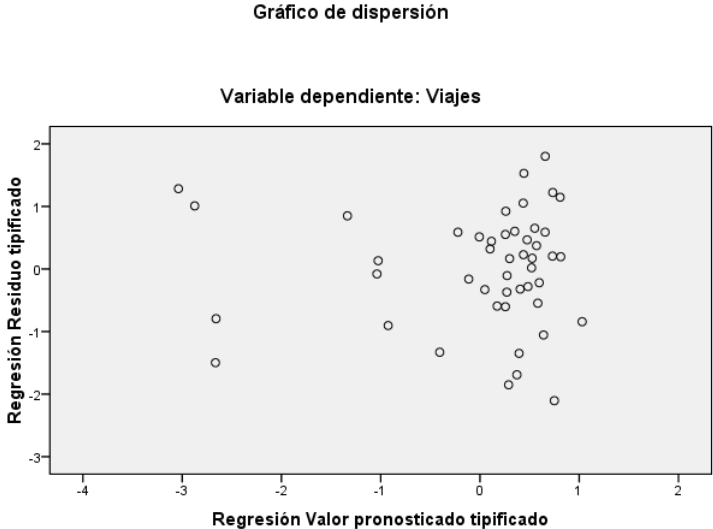


Figura 34 – Gráfico de dispersión modelo de Viajes

Del mismo modo que los modelos de validaciones o transacciones, la forma funcional log-log de los modelos entregó mejores resultados en término de bondad de ajuste. El modelo que utiliza únicamente variables propias de la estacionalidad es capaz de explicar un 95.2% de la varianza en la demanda de viajes, mientras que en el modelo que incorpora variables de oferta de buses y Metro, además de la tarifa, se evidencia una mejora de este valor a un 96.6%.

Tabla 20 – Modelos log-log anidados de Viajes Transantiago

	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.	Estadísticos de colinealidad	
	B	Error típ.	Beta			Tolerancia	FIV
1 (Constante)	3,59	,19		19,04	,00		
Enero	-,16	,01	-,40	-12,06	,00	,97	1,03
Febrero	-,37	,01	-,91	-27,59	,00	,96	1,04
Julio	-,03	,01	-,08	-2,50	,02	,95	1,05
MesMovilizado	-,04	,01	-,09	-2,60	,01	,94	1,06
DiasLabLog	,32	,06	,17	5,14	,00	,97	1,03
2 (Constante)	3,38	,58		5,87	,00		
Enero	-,15	,01	-,37	-12,78	,00	,87	1,15
Febrero	-,31	,02	-,76	-17,41	,00	,39	2,58
Julio	-,04	,01	-,09	-3,32	,00	,95	1,06
MesMovilizado	-,04	,01	-,09	-3,22	,00	,88	1,13
DiasLabLog	,20	,06	,11	3,40	,00	,75	1,33
TarifaCorregLog	-,16	,09	-,19	-1,84	,07	,07	14,08
logKmCom	,24	,09	,11	2,59	,01	,38	2,61
logCocheKm	,30	,11	,28	2,62	,01	,07	15,10

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio				
					Cambio en R cuadrado	Cambio en F	gl1	gl2	Sig. Cambio en F
1	,978	,957	,952	,02497	,957	181,961	5	41	,000
2	,986	,972	,966	,02098	,015	6,698	3	38	,001

5. Pronóstico y modelación distintos escenarios

En este capítulo se validarán los modelos encontrados en el capítulo anterior, utilizando los datos de demanda del año 2013. En primer lugar se enseñará la validación del modelo de transacciones, seguido con la presentación y análisis de distintos escenarios en un horizonte de 3 años. Finalmente se revisará y validará el modelo de viajes, junto con su estimación a 3 años y los distintos escenarios.

5.1 Predicción modelo de Transacciones

5.1.1 Predicción con datos utilizados para estimación: 2007-2012

Las Figura 35 y 36 muestran las validaciones reales y las validaciones estimadas por el modelo (años 2007-2009 y 2010-2012)

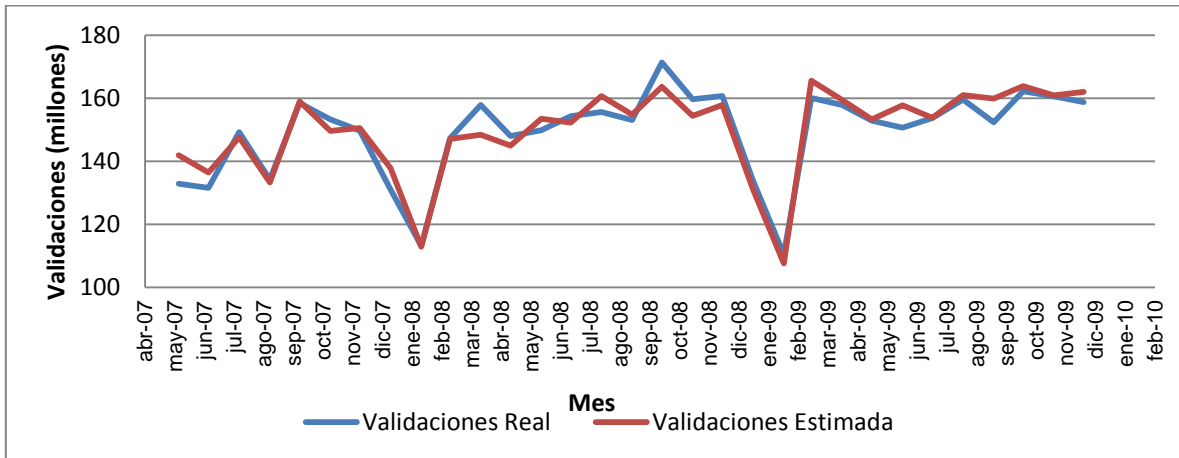


Figura 35 – Validaciones Real vs Validaciones Estimada (2007-2009)

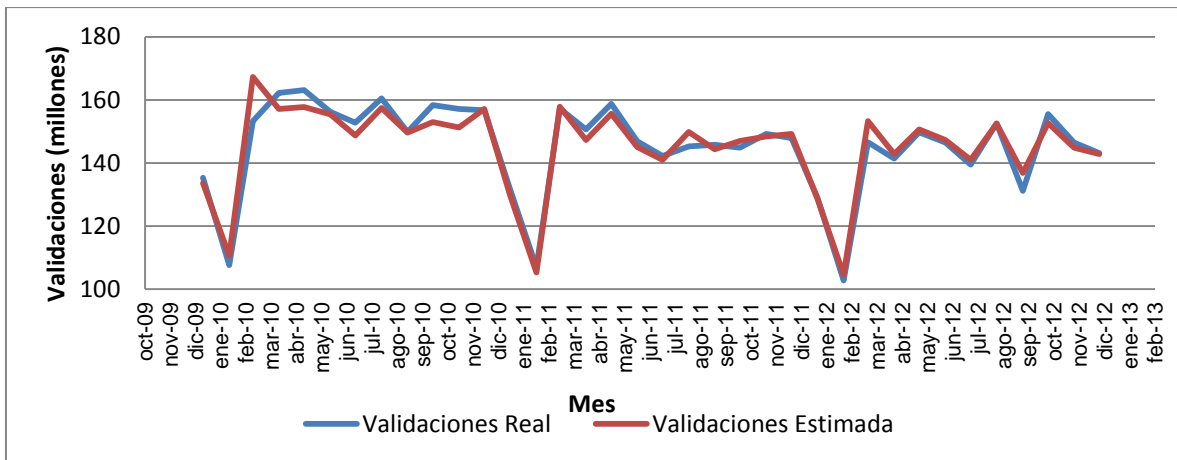


Figura 36 - Validaciones Real vs Validaciones Estimada (2010-2012)

Es importante notar que el modelo logra estimar de manera correcta las tendencias de la demanda, teniendo las mayores diferencias entre valores reales y estimados en aquellos meses en donde existe un factor externo importante que cambia la estructura

de viajes pero no logro ser determinada por el modelo. Ejemplo de ellos es la sobrestimación de demanda para el mes de Marzo de 2010, producto del terremoto sufrido a fines del mes anterior.

Es posible observar de forma adicional una subestimación por parte del modelo en meses de Septiembre y algunos Mayo. Esto se debe a que si bien el modelo no encontró significativo la adición como variable de domingos y festivos, ni de sábados e interferiados, estos meses tienen más días feriados e interferiados, en donde una gran cantidad de personas viaja menos o sale de Santiago.

Los valores de MAE y MAPE son 2.99 millones y 2.03% respectivamente.

5.1.2 Predicción con datos no utilizados para estimación: 2013

Centrándose en los datos de 2013, es posible observar en la Figura 37 una buena estimación, tanto en los meses de vacaciones como en el resto del año. Las peores estimaciones corresponden a los meses de marzo, en donde se sobreestimó el número de validaciones por algo más de 7 millones de validaciones, correspondientes a un error de 5,12% y Septiembre, en donde la demanda fue sobreestimada por 8,25 millones de validaciones, equivalentes a un 6,33% de las validaciones del mes. Al mes de Septiembre es posible atribuirle gran parte del error de estimación a las fiestas patrias, dado que en esa oportunidad los días feriados fueron a mitad de semana (miércoles y jueves), añadiéndose el día viernes como feriado por ley, lo que produce que los días lunes y martes no se comporten como días laborales normales, ni tampoco dicho fin de semana tenga un comportamiento normal en cuanto a viajes realizados por los usuarios.

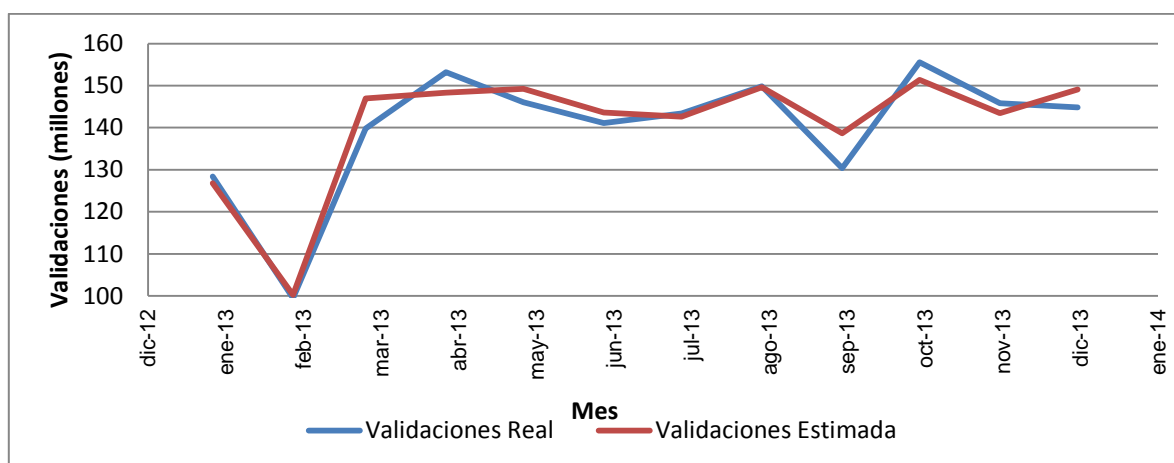


Figura 37 – Validación modelo validaciones con datos 2013

Tabla 21 – Resultados Validación modelo datos 2013

Fecha	Dda Real	Dda Estimada	Error Absoluto	Error Porcentual
ene-13	128,407	126,790	1,617	1,26%
feb-13	99,524	100,279	0,754	0,76%
mar-13	139,811	146,969	7,158	5,12%
abr-13	153,179	148,292	4,887	3,19%
may-13	146,058	149,256	3,198	2,19%
jun-13	141,080	143,639	2,559	1,81%
jul-13	143,406	142,592	0,814	0,57%
ago-13	149,877	149,590	0,287	0,19%
sep-13	130,406	138,663	8,256	6,33%
oct-13	155,552	151,384	4,168	2,68%
nov-13	145,811	143,449	2,362	1,62%
dic-13	144,836	149,113	4,278	2,95%

El error absoluto promedio o MAE por sus siglas en inglés (mean absolute error), fue de 3,36 millones de validaciones, y el error absoluto porcentual promedio o MAPE (mean absolute percentual error), fue de un 2,38%, siendo valores bajo el 10% considerado muy buenos en las predicciones.

5.1.3 Escenarios y pronóstico a 3 años

Existen distintos posibles escenarios en función de las distintas variables explicativas que utiliza el modelo de validaciones. Para esta sección se utilizarán ciertos supuestos en base a cómo se han comportado estas variables en los años anteriores, y cómo se podrían comportar en los años 2014, 2015 y 2016.

Cómo se observó en el capítulo anterior, el modelo de validaciones posee 3 variables que son posibles de manipular. Estas serían la longitud de red de buses, o el alcance espacial de Transantiago, la frecuencia de los buses y la tarifa.

Se sensibilizarán dos de estas variables, la tarifa y la longitud de la red de buses. No se realizarán pronósticos a partir de cambios en la frecuencia debido a que esta variable si bien es posible atribuirle diversos impactos, tales como tiempos de espera, oferta de buses y nivel de servicio, no es posible separar del todo cuánto de la variación de la frecuencia corresponde a decisiones operacionales y cuánto es en respuesta de la demanda misma.

Para la variable tarifa, se propondrá un aumento gradual desde enero 2014 a diciembre 2016 de 100 pesos a realizarse en los meses de vacaciones principalmente, como han sido la mayoría de las alzas de tarifas, llegando a un valor de 700 pesos. Esto no es algo descabellado considerando las alzas constantes del precio del petróleo, y recordando además que desde enero del 2009 a enero del 2012, la tarifa tuvo un

aumento desde los 400 pesos a 580 pesos. En la Figura 38 se muestra la comparación manteniendo el precio actual de la tarifa en 600 pesos.

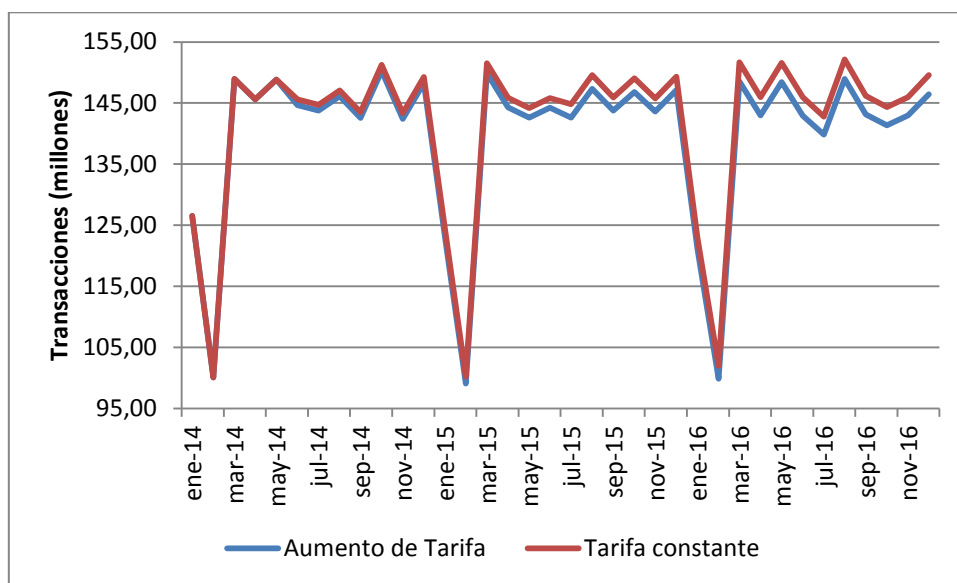


Figura 38 - Pronóstico a 3 años bajo escenarios de cambio en la Tarifa

Es posible observar en el gráfico el impacto que tiene un aumento en la tarifa a 3 años. Al ser un alza gradual ambas estimaciones se separan cada vez más hacia el 2016. Como se observa en la Tabla 22, una política que apunte a no subir las tarifas podría llevar a un alza en el total de validaciones para el año 2016 de un 2%. A lo largo de 3 años esto se traduce en más de 60 millones de validaciones atribuibles al alza de la tarifa.

Tabla 22 – Resumen Resultados distintos escenarios Tarifa

	2014	2015	2016	Total
Aumento de Tarifa	1688,14	1674,80	1666,54	5029,48
Tarifa Constante	1694,91	1696,74	1701,06	5092,71

A continuación se mostrará cual es el impacto de un aumento en la longitud de red de bus. Esto se logra o alargando recorridos o creando nuevos servicios, y como el modelo muestra, tiene un impacto positivo en la cantidad de validaciones mensuales.

A comienzos del 2007, la red de buses tenía una longitud de 2100 km, y a fines del 2013 la red abarcaba 2770 km. Se mostrará en la Figura 39 aumentos graduales de longitud (creación de nuevos recorridos), de manera de sumar para fines del año 2016 100 nuevos kilómetros en un escenario, y 200 km en el segundo escenario.

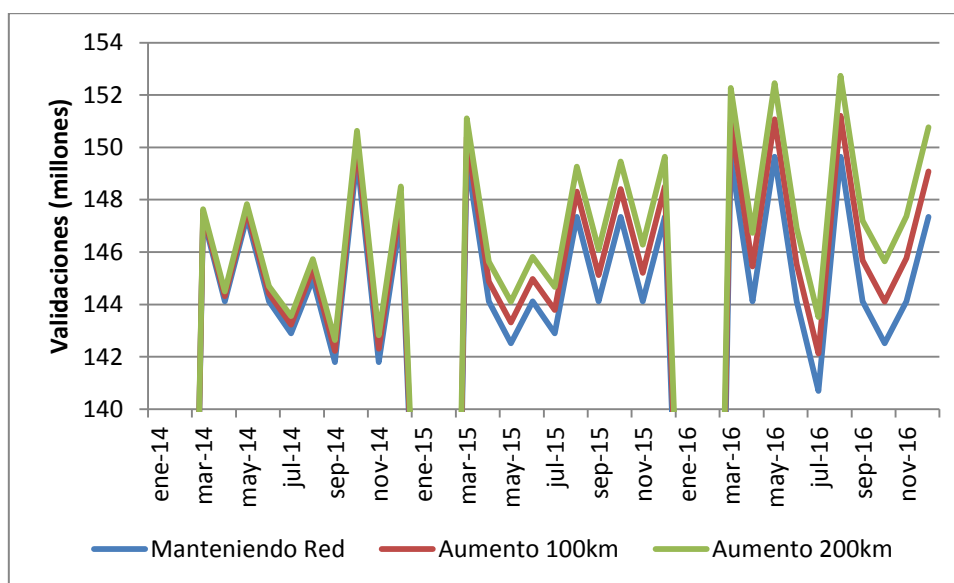


Figura 39 – Pronóstico a 3 años bajo escenarios de cambio en la longitud de red de buses

Es posible observar que los impactos de un aumento en la red de buses no es tan grande como en el caso de la tarifa en cuanto al total de validaciones, pero aun así existe un aumento en el total de las mismas al aumentar la red. Al aumentar la longitud de la red, se crean nuevos recorridos lo que permite viajes más directos (menos transbordos), lo que a su vez se traduce en una menor cantidad de validaciones.

A pesar de esto, como se observa en la Tabla 23, de manera agregada existe un aumento en la cantidad de validaciones a pesar de este fenómeno, ante estos escenarios en el año 2016 se espera un aumento en la cantidad de validaciones de 1% con un aumento de 100km en la red, y a su vez de un 2% con un aumento de 200 km.

Tabla 23 – Resumen resultados distintos escenarios Longitud Red Buses

	2014	2015	2016	Total
Manteniendo Red	1683,3	1683,52	1685,58	5052,41
Aumento 100 km	1687,02	1693,87	1702,52	5083,42
Aumento 200 km	1690,70	1704,06	1719,03	5113,80

Se puede desprender de esto que es necesario realizar esfuerzos con el fin de poder aumentar la demanda anual en el mediano plazo, debido a que actualmente se puede observar una tendencia a la baja de la misma. Este fenómeno no es algo de extrañar en una sociedad como la chilena, en donde junto con un elevado crecimiento económico viene de la mano una mayor posesión de vehículos, aumentando los índices de motorización del país y lleva a un menor uso de transporte público.

5.1.4 Aumento de la red de Metro

Si bien no se realizó un modelo que logre predecir el impacto de la apertura de las líneas 3 y 6 del metro, si fue posible realizar modelos que muestren el impacto o la correlación que ha existido con la apertura de las nuevas estaciones de las líneas 1 y 5 con la cantidad de validaciones tanto en buses, metro y de manera agregada (buses y metro).

Varios de los modelos de validaciones mostrados en el Capítulo 4 utilizan la variable de oferta de Buses (Km Comerciales) y de Metro (Coche Kilómetro). Ahora bien, en esta sección se realiza un estudio del impacto cruzado que tienen estas variables de oferta en el otro modo de transporte. Es así como se estudiará cuál es el efecto de un aumento en la oferta de Metro en el total de validaciones (buses y metro), sólo en metro y sólo en buses. Esto sirve como un buen indicador para el impacto de la longitud de la red de metro debido a que ambas variables están altamente correlacionadas (0.96).

Para el caso de Metro, es directo ver la correlación de que un aumento en los Coche Km ha significado un aumento en las transacciones o validaciones correspondientes a dicho modo de transporte (Tabla 6). En este caso la variable de oferta de buses en el modelo de Metro no dio significativa, por lo que no es posible concluir una relación positiva o negativa de esta variable en el total de validaciones de Metro.

En el caso del modelo de buses, el impacto de la variable coche kilómetro de Metro se puede observar en la Tabla 24.

Tabla 24 – Impacto Coche Km Metro en modelo de transacciones de buses

	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.	Estadísticos de colinealidad	
	B	Error típ.	Beta			Tolerancia	FIV
(Constante)	27,69	12,61		2,20	,03		
Enero	-10,59	1,77	-,26	-5,99	,00	,91	1,10
Febrero	-14,88	1,90	-,37	-7,84	,00	,79	1,27
Julio	-4,22	1,59	-,11	-2,66	,01	,95	1,05
MesMovilizado	-4,42	2,00	-,10	-2,22	,03	,88	1,14
DiasLaborales	,57	,36	,07	1,57	,12	,88	1,14
KMComercialesmillones	1,74	,29	,42	6,00	,00	,36	2,80
CocheKmMetro	4,60	1,25	,52	3,68	,00	,09	11,49
TarifaCorregida	-,11	,02	-,74	-5,74	,00	,10	9,75

De este modelo se puede apreciar que los kilómetros comerciales de Metro tienen un impacto positivo en la cantidad de validaciones de buses, en donde cada millón de coche-km mensual esta correlacionado de forma positiva a un aumento de 4.6 millones de validaciones en buses. Esto es lo contrario a lo descrito en Matas (2004), en donde

un aumento de la red de Metro se correlaciona de manera negativa con la demanda de buses.

Sin embargo, si al modelo presentado en la Tabla 24 se elimina la variable de tarifa, entonces el signo de este coeficiente cambia, como se muestra en Tabla 25.

Tabla 25 - Impacto Coche Km Metro en modelo de transacciones de buses (sin tarifa)

	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.	Estadísticos de colinealidad	
	B	Error típ.	Beta			Tolerancia	FIV
(Constante)	-19,43	11,81		-1,64	,11		
Enero	-8,46	2,13	-,21	-3,97	,00	,95	1,05
Febrero	-15,63	2,34	-,38	-6,69	,00	,79	1,26
Julio	-4,99	1,95	-,13	-2,55	,01	,96	1,04
MesMovilizado	-3,80	2,46	-,08	-1,54	,13	,88	1,14
DiasLaborales	1,03	,43	,13	2,37	,02	,92	1,08
KMComercialesmillones	2,88	,26	,69	11,00	,00	,67	1,49
CocheKmMetro	-2,15	,51	-,24	-4,19	,00	,78	1,28

Este cambio de signo y la alta correlación entre las variables de Coche Km Metro y Tarifa, no permite sacar una conclusión clara con respecto al real impacto de dicha variable con respecto al total de validaciones en buses. Si bien en el modelo mostrado en Tabla 25, todos los valores son significativos, la magnitud del impacto del coeficiente de Coche Km levanta sospechas al compararlo con la magnitud de la misma variable mostrada en el modelo de Metro (Tabla 6), 4.6 millones de validaciones de bus versus 1.67 millones de metro. Es decir que el aumento de la oferta de metro tendría un mayor impacto en los buses que en el metro (en número de validaciones), lo que a priori no parece un resultado lógico ni esperado.

Mismo problema es evidenciado en el caso del modelo agregado de validaciones. En el modelo que incluye tarifa además de las variables de oferta de los distintos modos, es posible observar un impacto positivo en el total de validaciones del sistema (Tabla 26), mientras que en el modelo en donde se elimina la variable tarifa, existe un cambio de signo en el impacto (Tabla 27).

De este modo no es posible concluir de manera categórica el impacto de la red de metro en el total de validaciones de buses ni en el modelo agregado, pero si es posible hacerlo para el modelo de Metro en donde tiene un impacto positivo. Es importante recordar que los modelos mostrados en esta sección corresponde a modelos de validaciones y no de viajes, por lo que es posible que un aumento en la oferta de metro disminuya el total de validaciones en el modelo agregado (efecto de realizar viajes mas directos), sin embargo como ya se mencionó, estos resultados no son concluyentes.

Tabla 26 – Impacto Coche Km Metro en modelo agregado de transacciones

	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.	Estadísticos de colinealidad	
	B	Error típ.	Beta			Tolerancia	FIV
(Constante)	67,13	13,40		5,01	,00		
Enero	-20,02	1,88	-,36	-10,65	,00	,91	1,10
Febrero	-31,74	2,02	-,57	-15,74	,00	,79	1,27
Julio	-6,08	1,69	-,12	-3,60	,00	,95	1,05
MesMovilizado	-5,84	2,12	-,09	-2,75	,01	,88	1,14
DiasLaborales	1,41	,38	,13	3,68	,00	,88	1,14
KMComercialesmillones	1,39	,31	,24	4,52	,00	,36	2,80
CocheKmMetro	6,77	1,33	,56	5,10	,00	,09	11,49
TarifaCorregida	-,13	,02	-,63	-6,28	,00	,10	9,75

Tabla 26 – Impacto Coche Km Metro en modelo agregado de transacciones (sin tarifa)

	Coeficientes no estandarizados	Error típ.	Coeficientes tipificados	t	Sig.	Estadísticos de colinealidad	FIV
	B		Beta			Tolerancia	
(Constante)	12,33	12,98		,95	,35		
Enero	-17,54	2,34	-,31	-7,48	,00	,95	1,05
Febrero	-32,62	2,57	-,59	-12,70	,00	,79	1,26
Julio	-6,97	2,15	-,14	-3,25	,00	,96	1,04
MesMovilizado	-5,12	2,70	-,08	-1,89	,06	,88	1,14
DiasLaborales	1,95	,48	,17	4,09	,00	,92	1,08
KMComercialesmillones	2,71	,29	,47	9,45	,00	,67	1,49
CocheKmMetro	-1,08	,56	-,09	-1,91	,06	,78	1,28

5.2 Predicción Modelo Viajes

5.2.1 Predicción con datos utilizados para estimación: 2009-2012

En la Figura 38 se observa el ajuste del modelo de ambos modelos de viajes mostrados en el Capítulo 4. El modelo que cuenta únicamente con variables estacionales y el modelo que además tiene tarifa y variables de oferta de los distintos modos de transporte. Es posible observar que ambos modelos predicen de manera similar, prueba de esto se observa en la Tabla 27, en dónde el MAE del modelo estacional es

de 1.68 millones y el MAE del modelo con variable de ofertas es de 1.62 millones. En el caso del error porcentual, esto equivale a un 1.84% para el modelo con variables estacionales, y un MAPE de 1.82% para el otro modelo.

Existen únicamente 2 puntos en donde el modelo no fue capaz de seguir la tendencia con respecto al mes anterior, es decir, pronosticó una baja en vez de un alza en la demanda, o viceversa. Uno de ellos fue producto del terremoto en el año 2010, por lo que era de esperar un cambio importante en la cantidad de viajes realizados.

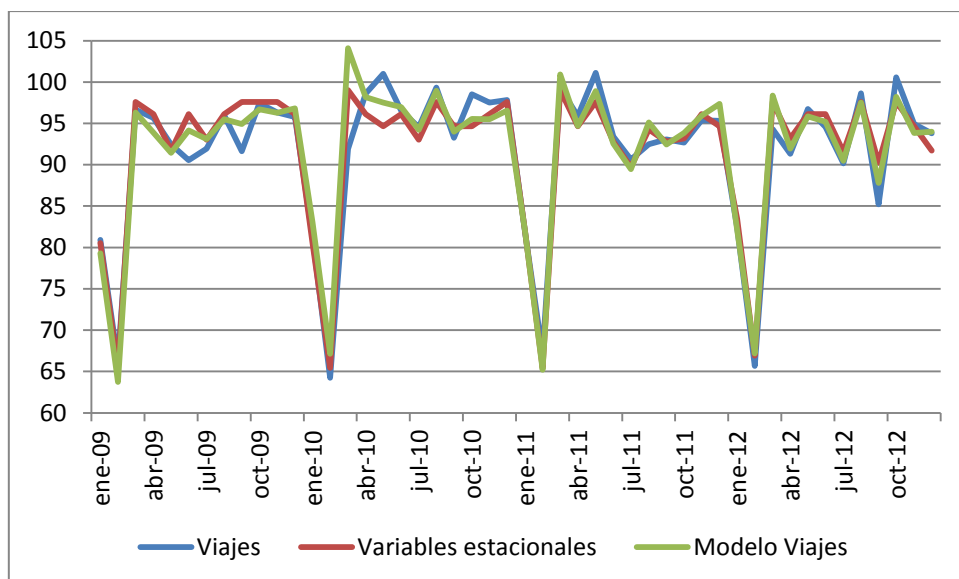


Figura 38 – Demanda Real vs Demanda estimada modelo de Viajes

Tabla 27 - MAE y MAPE modelos de Viajes

	MAPE	MAE
Modelo Variables estacionales	1,84%	1,688
Modelo de Viajes	1,82%	1,625

5.2.2 Predicción con datos no utilizados para estimación: 2013

En la Figura 39 y Tabla 28 se muestran los datos y el gráfico de validación del modelo, con los datos reales del año 2013 (datos no utilizados para la creación del mismo). Acá es posible observar que para el año 2013 el modelo que utiliza datos únicamente estacionales pronosticó de mejor manera. Esto es principalmente por el fuerte aumento de la oferta de Metro en ese año que no estuvo relacionada con un aumento en la longitud de la red, sino que con decisiones operacionales como la extensión de los servicios expreso en las líneas 2 y 5.

Para la Tabla 28, el modelo 1 corresponde al modelo que utiliza únicamente variables estacionales, y el modelo 2 al modelo anidado del mismo, en dónde se agregan variables de oferta y tarifa.

Tabla 28 – Validación modelo de Viajes datos 2013

Fecha	Demand a Real	Demanda estimada Modelo 1	Demanda estimada modelo 2	MAE 1	MAE 2	MAPE 1	MAPE 2
ene-13	82,973	83,450	82,581	0,478	0,391	0,58%	0,47%
feb-13	64,090	65,453	65,262	1,362	1,172	2,13%	1,83%
mar-13	90,928	94,644	96,788	3,715	5,860	4,09%	6,44%
abr-13	99,904	97,574	99,575	2,330	0,329	2,33%	0,33%
may-13	95,568	96,109	99,051	0,541	3,484	0,57%	3,65%
jun-13	92,061	94,644	96,437	2,583	4,376	2,81%	4,75%
jul-13	93,312	93,008	96,577	0,304	3,265	0,33%	3,50%
ago-13	97,815	96,109	99,157	1,706	1,342	1,74%	1,37%
sep-13	85,262	91,713	92,057	6,451	6,795	7,57%	7,97%
oct-13	101,554	97,574	100,737	3,980	0,817	3,92%	0,80%
nov-13	95,361	94,644	96,015	0,718	0,653	0,75%	0,68%
dic-13	95,119	96,109	98,519	0,990	3,400	1,04%	3,57%

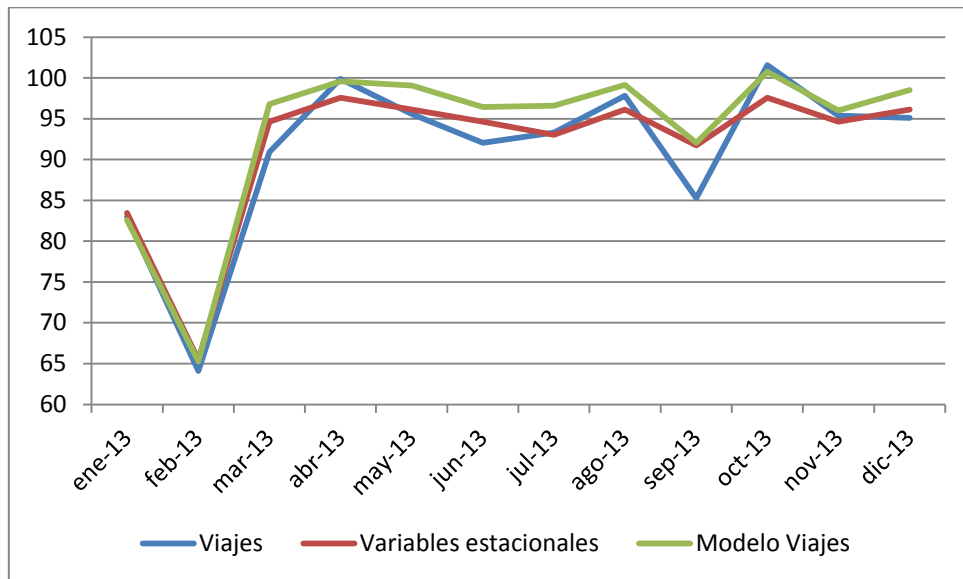


Figura 39 – Validación modelo de viajes con datos 2013

Es claro observar de la Figura 39, que el modelo de viajes (línea verde) sobreestima gran parte de la demanda, mientras que el modelo con variables estacionales presenta los mayores problemas en los meses de marzo y septiembre.

Los errores MAE y MAPE de estos pronósticos, son mostrados en la Tabla 29.

Tabla 29- MAE y MAPE modelos de viajes (2013)

	MAPE	MAE
Modelo Variables estacionales	2,32%	2,097
Modelo de Viajes	2,95%	2,657

Para los datos del año 2013, los errores (absolutos y porcentuales) de la predicción son menores en el modelo que no utiliza las variables de oferta ni la tarifa.

6. Conclusiones

La implementación de Transantiago en el año 2007 ha influido de alguna u otra forma en todas las personas que viven en Santiago. Para que sea una verdadera solución es necesario lograr introducir un sistema de transporte integrado y de alta calidad, sin embargo en las condiciones actuales, la ciudad de Santiago no está cerca de lograrlo aún.

Existen una gran cantidad de desafíos y pasos intermedios, y parte fundamental conseguirlo es la aplicación de políticas de transporte público acordes con la realidad de la ciudad, pero que además estén alineadas con un objetivo mayor. Indiscutiblemente para lograr realizar esto es necesario contar con un conocimiento acabado del impacto que estas políticas de transporte puedan causar en la demanda y en el comportamiento de los usuarios, que recordemos, no es comparable directamente con los usuarios de otros sistemas de transporte público.

No sólo es necesario saber los impactos que estas políticas tendrán, sino que además es importante poder hacer proyecciones certeras que permitan tomar decisiones y medidas a tiempo.

En base a estas necesidades se han presentado en esta memoria varias herramientas y modelos capaces de ayudar a conseguir un conocimiento más acabado de las implicancias e impactos que distintas variables tienen sobre las validaciones y viajes en Transantiago. Modelos que son una mejora al actual modo de estimación utilizado, permitiendo tomar decisiones operacionales y financieras con un mayor grado de conocimiento, o si se quiere ver así, con un menor nivel de incertidumbre.

Una serie de variables resultó ser significativa para todos los modelos mostrados en esta memoria. Además de las variables mudas correspondientes a los meses de enero, febrero, julio y mes movilizad, la variable días laborales siempre resulto ser significativa (tanto en modelos de viajes como de transacciones), así como la tarifa.

En el modelo de validaciones de buses y el agregado, además resultaron ser significativas las variables de frecuencia y longitud red de buses. Estas variables pueden ser reemplazadas por los kilómetros comerciales ofrecidos por Transantiago, obteniendo resultados un poco peores, pero de más fácil uso para Transantiago.

Para el modelo de validaciones de Metro, resultaron además significativas las variables de oferta de Metro (Coche-Km) y el desempleo. En el modelo de viajes, ambas variables de oferta resultaron ser significativas, así como la tarifa.

Dentro de las variables que resultaron no ser significativas se pueden encontrar el precio del combustible, la cantidad de precipitaciones, el parque automotriz, el índice automotor y la evasión.

Dentro de los resultados obtenidos, importante es destacar el impacto de la oferta de Metro en las validaciones de buses y del sistema agregado. A diferencia de lo expresado en Matas (2004), en donde el Metro de Madrid era competencia de los buses, el Metro de Santiago funciona como complemento al sistema de buses. Es por

esto que un aumento en la oferta podría tener un impacto positivo en la cantidad de validaciones de buses, además de aumentar la cantidad de viajes en el sistema de transporte público. En este aspecto, el impacto final de la oferta en el número de validaciones en bus debe ser estudiado más detalladamente en estudios futuros.

Este impacto en la cantidad de validaciones de buses impacta directamente a los distintos operadores, debido a que se les paga por validaciones. Si bien las realidades de la extensión de las líneas 1 y 5 del Metro son distintas a la apertura de las nuevas líneas (3 y 6) del Metro, es posible esperar, ya sea en menor o mayor magnitud, un efecto similar al encontrado.

Las elasticidades encontradas, no son del todo comparables en a lo que aparece en la literatura, ya que existe muy poco en esta con respecto a validaciones. Sin embargo los signos y los valores tienen un sentido económico lógico y están dentro de lo que se podría esperar. Que la elasticidad precio de las validaciones sea de -0.14 indica en cierta manera las pocas alternativas de modo que presenta mucha gente que utiliza el transporte público como alternativa de viaje. Al realizar un análisis más desagregado y separar buses con respecto de Metro, es posible observar que en los usuarios de buses su elasticidad es de -0.33 , en comparación con la del Metro que es -0.21 . Esta mayor elasticidad precio en las validaciones de buses con respecto al metro en parte es posible explicarla por la posibilidad de no pago en el bus (evasión), es mayor a la posibilidad de evadir en el Metro.

En cuanto al uso de los modelos encontrados para predecir la demanda, las validaciones obtenidas corroboran los resultados de García Ferrer et al (2006), en donde se obtiene que los modelos de demanda que utilizaban datos mensuales eran capaces de predecir bien a corto y mediano plazo utilizando únicamente variables propias de estacionalidad y propias del sistema de transporte público, y no así variables demográficas y económicas. Para los modelos estimados se encontró un error absoluto porcentual de predicción de tan solo un 2.38% para los datos de 2013 (validaciones), y de 2.95% para viajes, pudiéndose mejorar estos valores con la inserción de nuevas variables dummies capaces de explicar el comportamiento de meses como septiembre, en donde la cantidad de días laborales de calendario no representa a cabalidad la realidad. Este enfoque fue recomendado y utilizado por Gaudry (1975).

El sistema de la tarjeta Bip! junto con los GPS instalados en los buses entregan un sinnúmero de datos que permiten desarrollar modelos u obtener información que algunos años atrás eran impensado. Bajo este panorama es que a continuación se presentará las líneas de investigación a futuro que permite este trabajo con el fin de complementar y agregar nueva información relevante para un entendimiento más acabado del sistema.

Muchas de las variables que se presentaron como candidatas no pudieron ser utilizadas por distintos motivos, tanto por sus altas correlaciones como por no explicar los fenómenos que se creía que explicarían con este nivel de agregación.

Es por esto que se justifica un cambio en la periodicidad de los datos, que si bien cambiaría el enfoque del estudio, puede entregar nuevas luces sobre el impacto de muchas variables omitidas o no significativas en los modelos presentados a lo largo del estudio.

Ciertas variables como la cantidad de días feriados o interferidos, si existen o no precipitaciones durante el día, o eventos externos como marchas, paros o incidentes puntuales, no muestran un impacto en cuanto a la cantidad total de viajes o validaciones realizadas en un mes, y si lo muestran es marginal en comparaciones al total del mes. Esto puede ser tanto porque la magnitud del cambio es muy pequeña, o porque los usuarios deciden postergar su viaje. De este modo en un día que presente precipitaciones, los usuarios pueden decidir postergar ciertos compromisos o viajes que tenían pensado hacer a otro momento de la semana, y dicho cambio no es posible de medir en base a los datos utilizados en el estudio realizado anteriormente.

Este es el motivo por el que cambiar los datos de forma mensual a un análisis diario puede entregar información importante en el comportamiento de los usuarios, así como entender como algunas de estas variables que no tuvieron importancia sobre la variable dependiente.

Es importante mencionar que este nuevo enfoque entregaría nueva información que podría complementarse con la ya entregada con el fin de entender de mejor manera el impacto de un mayor número de variables ya no sólo en un carácter mensual, sino que en las decisiones que día a día toma la gente con respecto al uso de transporte público

Si bien Transantiago ha pasado por diversos indicadores y metodologías que buscan el medir el desempeño de los operadores y mejorar la calidad de los servicios, esta búsqueda a mejorar no permite que exista una toma de datos homogénea o constante en el tiempo, que lleva a que datos como la velocidad, y los índices de calidad de servicio no existan con fechas anteriores a 2009, y algunos incluso de 2012. Es importante poder llegar a indicadores y metodologías definitivas con el fin de poder utilizar estos datos que se están obteniendo en nuevos modelos y análisis que permitan mejorar el trabajo realizado y poder estudiar ciertas variables que se cree tienen un impacto importante en la demanda y en la calidad de servicio, pero por motivos de falta de datos no han podido ser utilizados en los modelos.

Siendo estos modelos los primeros modelos econométricos de Transantiago desde su puesta en marcha en el año 2007, sin lugar a dudas queda aún espacio para mejorarlos, recordando que los valores obtenidos son dinámicos y evolucionarán junto con el sistema de transporte y la sociedad en los próximos años.

A modo de cierre es importante recalcar que para lograr un sistema de transporte público integrado y de alta calidad, es necesario el realizar políticas de acorde a los objetivos buscados, y en ese aspecto la investigación realizada, las herramientas entregadas y el aprendizaje generado pueden resultar ser un factor decisivo en una buena planificación y proyecciones a futuro, de manera de lograr una mejora continua en Transantiago, que sin lugar a dudas tendrá efectos positivos en toda la comunidad.

Bibliografía

Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman, M. & White, P. (2004). The demand for public transport: a practical guide. TRL Report 593

Beltrán, P., Gschwender, A., & Palma, C. (2013). The impact of compliance measures on the operation of a bus system: The case of Transantiago. *Research in Transportation Economics*, 39(1), 79-89.

Bresson, G., Dargay, J., Madre, J. L., & Pirotte, A. (2003). The main determinants of the demand for public transport: a comparative analysis of England and France using shrinkage estimators. *Transportation Research part A: policy and practice*, 37(7), 605-627.

Dargay, J. M., & Hanly, M. (2002). The demand for local bus services in England. *Journal of Transport Economics and Policy*, 73-91.

De Rus, G. (1990). Public transport demand elasticities in Spain. *Journal of Transport Economics and Policy*, 189-201.

Durbin, J., & Watson, G. S. (1951). Testing for serial correlation in least squares regression. II. *Biometrika*, 159-177.

Garcia-Ferrer, A., Bujosa, M., de Juan, A., & Poncela, P. (2006). Demand forecast and elasticities estimation of public transport. *Journal of Transport Economics and Policy*, 45-67.

Gaudry, M. (1975). An aggregate time-series analysis of urban transit demand: the Montreal case. *Transportation Research*, 9(4), 249-258.

Goodwin, P. B. (1992). A review of new demand elasticities with special reference to short and long run effects of price changes. *Journal of transport economics and policy*, 155-169.

Gschwender, A. (2005). Improving the urban public transport in developing countries: the design of a new integrated system in Santiago de Chile. *9th Conference on Competition and Ownership in Land Transport*. Lisbon, Portugal

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1984). Multivariate data analysis with readings, 1995. *New York, NY Macmillan*.

Hidalgo, D., & Graftieux, P. (2008). Planning and implementation issues of a large scale transit modernization plan: The case of Transantiago, Chile. In *Transportation Research Board, 87th Annual Meeting, Washington*.

Holmgren, J. (2007) "Meta-analysis of public transport demand". *Transportation Research Part A*, Vol. 41, No. 10, pp. 1021-1035

- Jara-Díaz S. (1995) Evolución Zonal de la afluencia al Metro de Santiago: Un enfoque agregado multivariado
- Jara-Díaz S. y Paredes A.(1993) Modelo Estructural para la Demanda por Viajes en el Metro de Santiago. Actas del VI Congreso Chileno de Ingeniería de Transporte
- Litman, T. (2013). Understanding transport demands and elasticities. *How prices and other factors affect travel behavior.*(Victoria Transport Policy Institute: Litman) Available at <http://www.vtpi.org/elasticities.pdf> [Verified 22 November 2013].
- Mackett, R., Paulley, N., Preston, J. M., Shires, J., Titheridge, H., Wardman, M., & White, P. (2004). The demand for public transport: a practical guide.
- Marquardt, D. W. (1970). Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics*, 12(3), 591-612.
- Matas, A. (2004). Demand and revenue implications of an integrated public transport policy: The case of Madrid. *Transport Reviews*, 24(2), 195-217.
- McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics*, 3(4), 303-328.
- Muñoz, J. C., & Gschwender, A. (2008). Transantiago: A tale of two cities. *Research in Transportation Economics*, 22(1), 45-53.
- Neter, J., Wasserman, W., & Kutner, M. G. (1989). Applied linear regression analysis. *Homewood, IL: Irwin*.
- Nijkamp, P., & Pepping, G. (1998). Meta-analysis for explaining the variance in public transport demand elasticities in Europe. *Journal of Transportation and Statistics*, 1(1), 1-14.
- Oum, T. H., Waters, W. G., & Yong, J. S. (1992). Concepts of price elasticities of transport demand and recent empirical estimates: an interpretative survey. *Journal of Transport Economics and policy*, 139-154.
- Pucher, J., & Kurth, S. (1995). Making transit irresistible: Lessons from Europe. *Transportation Quarterly*, 49(1).
- Siimonetti C. (2012). Cambio del Modelo de Negocio y Ajuste Financiero Transantiago. *Sochitran*, 2013.
- Siimonetti C. (2013). Transantiago: Historia y evolución de los nuevos contratos. *Congreso SIBRT, León 2012*.
- Tabachnick, B. G., & Fidell, L. S. (2001). Using multivariate statistics.
- Tsai, Chi-Hong (2013). A longitudinal study on the linkage between public transport demand and land use characteristics: A pseudo panel approach. Tesis de Doctorado

Velásquez, J. M. (2009). Análisis de factores que inciden en la demanda del sistema TransMilenio en Bogotá, Colombia. *Revista de Ingeniería*, (30), 56-64.

Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2010). *Statistical and econometric methods for transportation data analysis*. CRC press.

Yáñez, M. F., Mansilla, P., & de Dios Ortúzar, J. (2010). The Santiago Panel: measuring the effects of implementing Transantiago. *Transportation*, 37(1), 125-149.