



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE  
ESCUELA DE INGENIERIA

**IMPORTANCIA DE LOS ATRIBUTOS DE  
VIVIENDA Y BARRIO EN LOCALIZACION  
RESIDENCIAL: UNA APLICACION DEL  
METODO *BEST-WORST* AL CENTRO DE  
SANTIAGO**

**CAMILA BALBONTIN TANHNUZ**

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

Profesor Supervisor:  
**JUAN DE DIOS ORTUZAR SALAS**

Santiago de Chile, Noviembre, 2013

© 2013, Camila Balbontin Tanhnuz



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE  
ESCUELA DE INGENIERIA

# **IMPORTANCIA DE LOS ATRIBUTOS DE VIVIENDA Y BARRIO EN LOCALIZACION RESIDENCIAL: UNA APLICACION DEL METODO *BEST-WORST* AL CENTRO DE SANTIAGO**

**CAMILA BALBONTIN TANHNUZ**

Tesis presentada a la Comisión integrada por los profesores:

**JUAN DE DIOS ORTUZAR SALAS**

**LUIS IGNACIO RIZZI CAMPANELLA**

**MARGARITA GREENE ZUÑIGA**

**ALFREDO SERPELL BLEY**

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

Santiago de Chile, Noviembre, 2013

A mis hermanas, padres y amigos  
por alentarme y apoyarme durante  
todo el Magíster.

## AGRADECIMIENTOS

Me gustaría comenzar agradeciendo el apoyo incondicional de mi profesor supervisor, Juan de Dios Ortúzar. Su total disposición, dedicación, cariño y paciencia durante todo este período fueron esenciales en la realización de esta investigación.

En segundo lugar, me gustaría agradecer a Joffre Swait que se comportó como un co-supervisor en esta investigación. Pese a encontrarse en otro lugar, tuvo una increíble disposición a enseñarme todo lo referido al método *Best-Worst Scaling* que fue fundamental para la tesis. Asimismo, en todo momento se mostró disponible para resolver mis dudas referidas tanto a la teoría como a la parte de modelación de esta investigación.

Además, agradezco a Ignacia Torres quien recolectó tanto la base de datos utilizada en esta investigación como la muestra de validación. A su vez, me gustaría agradecer a Luis Rizzi y Margarita Greene, que como parte de la comisión, me entregaron comentarios que enriquecieron esta investigación.

Por otro lado, quiero agradecer al Fondo Nacional de Desarrollo Científico y Tecnológico (Proyecto FONDECYT 1100068), al Instituto Milenio en Sistemas Complejos de Ingeniería (ICM: P05-004F; FONDECYT: FB016) y al Centro de Desarrollo Urbano Sustentable – CEDEUS (Conicyt/Fondap/15110020), por el apoyo otorgado a esta investigación.

Finalmente, quisiera agradecer a todo el Departamento de Ingeniería de Transporte y Logística, tanto profesores, administrativos y alumnos, por acompañarme y apoyarme durante todo este período.

## INDICE GENERAL

	Pág.
1 INTRODUCCIÓN .....	1
2 OBJETIVOS DEL ESTUDIO .....	3
3 MARCO TEÓRICO .....	4
3.1 Modelos de Elección Discreta .....	4
3.1.1 Modelo Logit Simple o Multinomial (MNL) .....	6
3.1.2 Modelo Logit Mixto (ML) .....	7
3.2 Teoría Best-Worst .....	8
3.3 Investigaciones Anteriores sobre Encuestas Best-Worst Caso 2 .....	11
3.4 Funciones de Utilidad <i>B-W</i> Caso 2 vs. Elección Discreta Sin Heterogeneidad .....	12
3.5 Combinación de Respuestas de Dos Bancos de Datos .....	17
3.6 Inclusión de Heterogeneidad .....	19
3.6.1 Variaciones Sistemáticas en los Gustos (VSG) .....	19
3.6.2 Efecto Panel .....	20
3.7 Tests Utilizados .....	20
3.7.1 Tests para Modelos <i>Best</i> , <i>Worst</i> , ED .....	20
3.7.2 Indicadores para Analizar Capacidad Predictiva .....	21
3.8 Disposición al Pago por Atributos .....	22
4 PRESENTACIÓN Y ANÁLISIS DE LOS DATOS DISPONIBLES .....	24
4.1 Atributos Presentados en la Encuesta .....	24
4.2 Diseño de la Encuesta .....	24
4.3 Codificación .....	25
4.4 Muestra de Validación .....	26
4.5 Análisis de Variables Socio-Económicas en el Banco de Datos .....	26
4.6 Comparación de Características Individuales en el Banco de Datos y Muestra de Validación .....	33

5	ESTIMACIÓN DE MODELOS SIN HETEROGENEIDAD .....	39
5.1	Modelos $B_{SINHET}$ , $W_{SINHET}$ y $B+W_{SINHET}$ .....	39
5.2	Modelos Mixtos $ED_{SINHET}$ , $B_{SINHET}$ y $W_{SINHET}$ .....	40
5.3	Modelo $ED+Best_{SINHET}$ .....	41
5.4	Modelo $ED+Worst_{SINHET}$ .....	42
5.5	Modelos $ED+Best+Worst_{SINHET}$ .....	43
5.6	Comportamiento Predictivo de Modelos en la Muestra de Validación.....	45
5.6.1	$ED_{SINHET}$ .....	45
5.6.2	$ED+Best_{SINHET}$ .....	46
5.6.3	$ED+Worst_{SINHET}$ .....	46
5.6.4	$ED+Best+Worst_{SINHET}$ .....	47
5.6.5	Comparación del Comportamiento Predictivo de los Modelos .....	48
6	ESTIMACIÓN DE MODELOS CON HETEROGENEIDAD .....	50
6.1	Heterogeneidad Mediante Componentes de Error .....	50
6.1.1	Resultados Modelos $Best_{EP}$ , $Worst_{EP}$ y $ED_{EP}$ .....	50
6.1.2	Resultados Modelo $Best+Worst_{EP}$ .....	50
6.1.3	Resultados Modelo $ED+Best_{EP}$ .....	52
6.1.4	Resultados Modelo $ED+Worst_{EP}$ .....	53
6.1.5	Comportamiento Predictivo en la Muestra de Validación.....	54
6.2	Variaciones Sistemáticas en los Gustos (VSG).....	55
6.2.1	Análisis de Variables .....	55
6.2.2	Modelo $ED_{VSG}$ .....	56
6.2.3	Modelo $ED+Best_{VSG}$ .....	57
6.2.4	Modelo $ED+Best+Worst_{VSG}$ .....	59
6.2.5	Comportamiento Predictivo en la Muestra de Validación.....	65
6.3	Variaciones Sistemáticas en los Gustos y Efecto Panel .....	66
6.3.1	Modelo $ED_{EP+VSG}$ .....	67
6.3.2	Modelo $ED+Best_{EP+VSG}$ .....	68
6.3.3	Comportamiento Predictivo en Muestra de Validación.....	70
6.4	Análisis de Inclusión de Heterogeneidad .....	71
6.4.1	Log Verosimilitud.....	71

6.4.2	Comportamiento Predictivo en la Muestra de Validación.....	72
6.4.3	Coeficientes Estimados.....	73
6.4.4	Análisis de Resultados.....	81
6.5	Análisis por Atributo y Cálculo de Disposiciones al Pago .....	84
6.5.1	Acceso al Metro.....	84
6.5.2	Áreas Verdes.....	85
6.5.3	Dividendo Departamento.....	85
6.5.4	Gimnasio y Piscina en el Edificio.....	86
6.5.5	Mantenimiento y Estado de Conservación de la Calle .....	87
6.5.6	Servicios Comerciales Básicos .....	88
6.5.7	Servicios Culturales .....	88
6.5.8	Tamaño de los Dormitorios .....	89
6.6	Cálculo de Disposiciones al Pago (WTP) .....	90
7	CONCLUSIONES.....	95
	BIBLIOGRAFÍA .....	99
	ANEXOS .....	102
	Anexo A: Tests Utilizados para Analizar Capacidad Predictiva en Muestra de Validación .....	103
	Anexo B: Diseño de la Encuesta .....	107
	Anexo C: Variables Socioeconómicas de los Individuos en el Banco de Datos ..	108
	Anexo D: Coeficientes Estimados Modelos Sin Heterogeneidad .....	113
	Anexo E: Coeficientes Estimados Modelos con Efecto Panel (EP) .....	116

## INDICE DE TABLAS

	Pág.
Tabla 4-1: Atributos Presentados en la Encuesta y Codificación de sus Niveles .....	25
Tabla 4-2: Niveles del Atributo <i>Dividendo</i> según Número de Habitaciones Requeridas .....	33
Tabla 5-1: Test de Razón de Verosimilitud para Modelos de ED junto con <i>Best</i> y <i>Worst</i> .....	41
Tabla 5-2: Predicción Modelo ED <sub>SINHET</sub> en Muestra de Validación.....	45
Tabla 5-3: Predicción Modelo ED+ <i>Best</i> <sub>SINHET</sub> en Muestra de Validación .....	46
Tabla 5-4: Predicción Modelo ED+ <i>Worst</i> <sub>SINHET</sub> en Muestra de Validación.....	47
Tabla 5-5: Predicción Modelo ED+ <i>Best</i> + <i>Worst</i> <sub>SINHET</sub> en Muestra de Validación.....	48
Tabla 5-6: Indicadores de Predicción Mejores Modelos ED con <i>Best</i> y/o <i>Worst</i> .....	48
Tabla 6-1: Capacidad Predictiva Modelo ED+ <i>Best</i> <sub>EP</sub> y ED <sub>EP</sub> en Muestra de Validación .....	55
Tabla 6-2: Función de Utilidad Desagregada por Atributos Modelo ED <sub>VSG</sub> .....	57
Tabla 6-3: Función de Utilidad Desagregada Modelo ED+ <i>Best</i> <sub>VSG</sub> .....	60
Tabla 6-4: Función Utilidad Desagregada por Atributo Modelo ED+ <i>Best</i> + <i>Worst</i> <sub>VSG</sub> ...	64
Tabla 6-5: Indicadores Predicción Modelos con VSG en Muestra de Validación .....	65
Tabla 6-6: Función de Utilidad Desagregada Modelo ED <sub>EP+VSG</sub> .....	67

Tabla 6-7: Función de Utilidad Modelo ED+ <i>Best</i> <sub>VSG+EP</sub> .....	70
Tabla 6-8: Indicadores de Predicción Modelo ED <sub>EP+VSG</sub> y Modelo ED+ <i>Best</i> <sub>EP+VSG</sub> .....	71
Tabla 6-9: Log Verosimilitudes de los Modelos Estimados .....	71
Tabla 6-10: Log-Verosimilitud de Predecir en la Muestra de Validación .....	73
Tabla 6-11: Intercepto de Atributos en Modelos ED+ <i>Best</i> sin y con Heterogeneidad ...	75
Tabla 6-12: Sensibilidades de Atributos Modelo ED+ <i>Best</i> , sin y con Efecto Panel .....	77
Tabla 6-13: Desviación Estándar Sensibilidades para Modelos sin y con Heterogeneidad .....	78
Tabla 6-14: Parámetros del atributo <i>Acceso al Metro</i> .....	84
Tabla 6-15: Parámetros del atributo <i>Dividendo departamento</i> respuestas <i>best</i> .....	86
Tabla 6-16: Parámetros Promedio del <i>Dividendo departamento</i> en Respuestas ED .....	86
Tabla 6-17: Parámetros del Atributo <i>Servicios comerciales básicos</i> .....	88
Tabla 6-18: Parámetros del Atributo <i>Servicios culturales</i> .....	89
Tabla 6-19: Parámetros del Atributo <i>Tamaño de los dormitorios</i> .....	90
Tabla 6-20: Disposición al Pago (WTP) por Atributos Modelo ED+ <i>Best</i> <sub>EP+VSG</sub> .....	91
Tabla 6-21: WTP Estimadas por Torres (2009).....	92
Tabla A-1: Sensibilidades (pendientes) respuestas <i>Best</i> , <i>Worst</i> y <i>Best+Worst</i> .....	113
Tabla A-2: Sensibilidades (pendientes) Modelo Final <i>B+W</i> sin Heterogeneidad .....	114

Tabla A-3: Coeficientes (pendientes) Modelos Sin Heterogeneidad Estimados con  
Respuestas *Best* o *Worst* combinadas con las ED ..... 115

Tabla A-4: Coeficientes (sensibilidades) Modelo *Best*, *Worst* y ED Con Efecto Panel 116

## INDICE DE FIGURAS

	Pág.
Figura 4-1: Habitantes vs. Número de Autos en el Hogar .....	30
Figura 4-2: Ingreso Personal Mensual de los Individuos en el Banco de Datos.....	31
Figura 4-3: Ingreso Familiar Mensual de los Hogares en el Banco de Datos.....	31
Figura 4-4: Dispersión IFAM vs Ingreso Personal .....	32
Figura 4-5: Distribución Número de Dormitorios según <i>Número Habitantes</i> .....	34
Figura 4-6: Promedio IFAM por Comuna de Origen en MV y BD.....	35
Figura 4-7: Distribución Variable Sexo Femenino por Comuna de Origen en MV y BD .....	36
Figura 4-8: Promedio de Autos en el Hogar en Relación al Sexo en BD y MV.....	37
Figura 4-9: Cociente N°Autos/N°Habitantes en Relación al Sexo en BD y MV .....	38
Figura 5-1: Sensibilidades (pendientes) de Modelos Sin Heterogeneidad.....	40
Figura 5-2: Coeficientes (sensibilidades) de Modelos Sin Heterogeneidad para Respuestas ED versus Respuestas <i>Best</i> .....	42
Figura 5-3: Coeficientes (sensibilidades) de Modelos Sin Heterogeneidad para Respuestas ED versus Respuestas <i>Worst</i> .....	43
Figura 5-4: Coeficientes (sensibilidades) de Modelos Sin Heterogeneidad para Respuestas ED versus Parámetros Modelo <i>Best+Worst</i> <sub>SINHET</sub> .....	44

Figura 6-1: Coeficientes (sensibilidades) de Modelos con EP para Respuestas <i>Best</i> versus Respuestas <i>Worst</i> .....	51
Figura 6-2: Coeficientes (sensibilidades) de Modelos con EP para Respuestas <i>Best</i> versus Respuestas ED.....	52
Figura 6-3: Coeficientes (sensibilidades) de Modelos con EP para Respuestas <i>Worst</i> versus Respuestas ED.....	53
Figura 6-4: Coeficientes (sensibilidades) de Modelos con VSG para Respuestas <i>Best</i> versus Respuestas ED.....	58
Figura 6-5: Coeficientes (sensibilidades) del modelo ED+ <i>Best</i> + <i>Worst</i> <sub>VSGinicial</sub> para Respuestas <i>Best</i> versus las Respuestas <i>Worst</i> .....	61
Figura 6-6: Coeficientes (sensibilidades) de Respuestas ED y <i>Best</i> + <i>Worst</i> <sub>VSG</sub> considerando las VSG estimadas en el modelo ED+ <i>Best</i> + <i>Worst</i> <sub>VSGinicial</sub> .....	63
Figura 6-7: Coeficientes (sensibilidades) de Modelos con VSG y EP para Respuestas <i>Best</i> versus Respuestas ED .....	68
Figura 6-8: Escala Relativa Modelo ED+ <i>Best</i> <sub>EP+VSG</sub> .....	82
Figura 6-9: Parámetros del atributo <i>Áreas Verdes</i> .....	85
Figura 6-10: Parámetros del atributo <i>Gimnasio y piscina en el edificio</i> .....	87
Figura 6-11: Parámetros del Atributo <i>Mantenimiento y estado de conservación de la calle</i> .....	87
Figura 6-12: Escala WTP Estimadas por Nuestro Modelo ED+ <i>Best</i> versus las Estimadas en Torres (2009) .....	93

Figura A-1: Diseño de la Encuesta.....	107
Figura A-2: Distribución Variable <i>Sexo</i> en el Banco de Datos.....	108
Figura A-3: Distribución Variable <i>Edad</i> en el Banco de Datos.....	108
Figura A-4: Distribución por Comuna de Origen de los Individuos el Banco de Datos .....	109
Figura A-5: Distribución Variable <i>Número de Autos</i> en el Banco de Datos .....	109
Figura A-6: Distribución Variable <i>Número de Habitantes</i> en el Banco de Datos .....	110
Figura A-7: Distribución Variable <i>Tiempo de Estadía</i> en el Banco de Datos .....	110
Figura A-8: Distribución Variable <i>Estudios</i> el Banco de Datos .....	111
Figura A-9: Distribución Agregada Variable <i>Ocupación</i> en el Banco de Datos .....	111
Figura A-10: Distribución Variable <i>Ocupación</i> el Banco de Datos .....	112
Figura A-11: Distribución Variable <i>Número de Dormitorios</i> el Banco de Datos.....	112

## RESUMEN

En el marco de estudiar la valoración de atributos de vivienda y barrio en el centro de Santiago, por parte de los usuarios, se estimó modelos de elección discreta utilizando encuestas del tipo *Best-Worst Scaling* (Caso 2) en conjunto con una más clásica de elección discreta (ED). Se estudió además la significancia de incluir heterogeneidad mediante variaciones sistemáticas de los gustos y, además, tratando correctamente el efecto panel mediante componentes de error. De acuerdo a la literatura estudiada, esto no se había aplicado anteriormente. Nuestra base de datos corresponde a individuos que planeaban arrendar un departamento en la comuna de Santiago centro. Adicionalmente, se dispone de una muestra de validación que proviene de un experimento de preferencias declaradas realizado en este mismo contexto (Torres 2009). El modelo elegido incluyó ambos tipos de heterogeneidad y se estimó uniendo las respuestas del tipo *Best* con ED. De los ocho atributos presentados en la encuesta, se pudo considerar siete comunes entre estas respuestas, pero el atributo *Dividendo departamento* debió considerarse específico. Con este modelo, se pudo ver la diferencia entre el impacto que tiene un atributo en sí en la función de utilidad, y el impacto que tiene la variación en sus niveles mostrados a los encuestados. El atributo más interesante fue el *Dividendo departamento*, con un impacto intrínseco bajo en la función de utilidad, pero debido a la variación de sus niveles alcanzaba el impacto mayor. Además, se estimó las disposiciones al pago por cada atributo, concluyendo que los más relevantes eran el *Acceso al Metro*, *Mantenimiento y estado de conservación de la calle*, y *Tamaño de los Dormitorios*. Los atributos menos relevantes fueron *Servicios culturales* y *Gimnasio y piscina en el edificio*.

*Palabras Clave:* encuestas de elección discreta; localización residencial; impacto de los atributos; *Best-Worst Scaling*; heterogeneidad.

## ABSTRACT

In the context of finding the importance of dwelling and neighbourhood attributes in city centre location, we estimated discrete choice models mixing Best-Worst (B-W) Scaling (Case 2) and Stated Choice (SC) data. We allowed for observed heterogeneity through systematic taste variations and also for unobserved heterogeneity by treating the pseudo panel nature of the data through error components (separately and jointly). To the best of our knowledge, this has not been done before. Our data set corresponds to a sample of individuals interested in acquiring a flat in the centre of Santiago. In addition, we had available a validation sample corresponding to the responses of an even larger set of individuals to a more traditional SC exercise in the same context. We found that the best statistical model was obtained by pooling the ‘best’ and SC responses including observed and unobserved heterogeneity. This model considers seven of the eight attributes presented in the survey as common between the ‘best’ and ED answers, and the attribute Rent as specific to each environment. The use of B-W information allows distinguishing between the intrinsic importance of an attribute and that which is due to its variation levels in the experiment. For example, in the data used it was found that the *Rent* appeared to have a relatively small intrinsic impact, but the variation across its range had the largest impact among all attributes (the levels presented in the survey were based on actual observations). With this model we also estimated the users’ willingness to pay (WTP) for each attribute. The results showed a higher WTP for *Access to Metro, Cleanliness and state of maintenance of streets & sidewalks* and *Size*. The lower WTP were found for the attributes *Distance to cultural/gastronomic areas* and *Swimming Pool/Gym*.

*Keywords:* discrete choice experiments; residential location; attribute importance; best-worst scaling; heterogeneity

## 1 INTRODUCCIÓN

En Chile, al igual que en otros países, ha surgido la necesidad de revitalizar los centros urbanos que han ido perdiendo residentes en el transcurso del tiempo. Un foco de investigación asociado a esta necesidad es la valoración subjetiva, por parte de los usuarios, de los distintos atributos que permiten definir o caracterizar a viviendas y barrios en el centro de Santiago. En este marco, Torres (2009) realizó un estudio en la comuna de Santiago, entrevistando a individuos que habían considerado la posibilidad de comprar un departamento allí, utilizando encuestas de preferencias declaradas (PD) y posteriormente estimando modelos de elección discreta.

Uno de los problemas asociados a los modelos de elección estimados utilizando encuestas de PD tradicionales, ha sido cómo detectar correctamente el impacto que tiene la diferencia entre los niveles de un atributo en su significancia estadística. En este marco, Louviere y Swait (1997) desarrollaron un enfoque de encuestas llamada *best-worst scaling* (B-W), cuyo objetivo es precisamente entregar información respecto a la significancia de cada atributo y, de forma separada, la significancia de los niveles presentados al encuestado. Como parte de su tesis, Torres (2009), también realizó una encuesta de tipo B-W en conjunto con una del tipo elección discreta binaria (ED) a una muestra adicional de individuos; la encuesta ED consultaba si la persona arrendaría o no la alternativa presentada, que se describía en términos de un conjunto de atributos. Estas dos encuestas no fueron analizadas en su estudio. En esta investigación se utilizan los datos de estas dos encuestas para analizar si existe complementariedad entre el enfoque B-W con el enfoque de elección discreta. A su vez, se utiliza la encuesta de elección discreta tradicional de Torres (2009) para validar los modelos estimados.

Adicionalmente, se estudia el efecto que tiene la inclusión de heterogeneidad al estimar modelos para las respuestas conjuntas de una encuesta B-W y de una de ED. En esta investigación se incluye heterogeneidad mediante variaciones sistemáticas de los gustos

y también se incluye efecto panel (observaciones repetidas por parte de cada individuo) mediante componentes de error.

El resto de la tesis está organizada de la siguiente forma. En el capítulo dos se explican los principales objetivos del estudio. El capítulo tres corresponde al marco teórico de esta tesis. Primero, se explican los modelos de elección discreta. Luego, se expone la teoría *best-worst* y se describen investigaciones previas que han aplicado este enfoque. Luego, se analizan las diferencias entre la función de utilidad estimada a partir de las respuestas a este tipo de encuestas con la función utilidad estimada a partir de las respuestas de elección discreta binaria. A su vez, se analiza cómo se pueden combinar estos tipos de respuesta. Al final de este capítulo se explica la manera en que se incluirá heterogeneidad en los modelos y se muestran los *tests* e indicadores relevantes para esta investigación.

En el cuarto capítulo se describen los datos disponibles y la manera en que se presentó la encuesta a los individuos, se definen los atributos utilizados y la codificación de cada uno. Asimismo, se realiza un análisis de la base de datos que será utilizada para la posterior estimación de los modelos. Esta se compara con la muestra de validación que será utilizada para analizar la capacidad predictiva de los modelos estimados. En el capítulo quinto se estiman los modelos sin incluir heterogeneidad. En el capítulo seis se estiman los modelos con heterogeneidad, ya sea sólo con VSG, sólo con EP, o con una combinación de ambos. Luego, se realiza un análisis sobre el efecto que tuvo la inclusión de esta heterogeneidad, ya sea en su log verosimilitud, en su capacidad predictiva, o en los parámetros de la función utilidad estimada, y se elige el modelo más adecuado. Con este modelo se estudia el impacto que tiene cada atributo por separado y se finaliza este capítulo estimando las disposiciones al pago por cada uno de ellos. Finalmente en el capítulo 7 se exponen las principales conclusiones de esta investigación.

## 2 OBJETIVOS DEL ESTUDIO

El objetivo principal de esta investigación es evaluar la complementariedad de las encuestas *best worst* con las de ED. Para lograr este objetivo, se estimará distintos modelos del tipo *logit simple* y *logit mixto* intentando verificar si efectivamente estas respuestas se alinean correctamente. En primer lugar, se utilizará un modelo *logit simple*, uniendo las respuestas *best worst* con las de ED, para evaluar qué atributos se pueden considerar comunes entre ellas. Luego, se agregará heterogeneidad para ver su efecto en la unión de las respuestas. Se evaluará la posibilidad de incluir heterogeneidad mediante variaciones sistemáticas de los gustos (VSG) y también se tratará el efecto panel (EP) inherente a estos datos (múltiples observaciones por individuo) mediante componentes de error, estimando modelos *logit mixto*. Se analizará el impacto que tiene la inclusión de cada tipo de heterogeneidad por separado, y luego, se verá su efecto en conjunto. Es importante mencionar que no se han encontrado en la literatura estudios anteriores sobre encuestas *best worst* junto con una pregunta de elección discreta binaria que incluyan heterogeneidad mediante VSG y EP en conjunto.

Habiendo estimado los distintos modelos, el objetivo será verificar si efectivamente las respuestas *best worst* y las de ED pueden unirse; también se examinará si los modelos conjuntos predicen mejor (en una muestra de validación diferente a la utilizada para estimar los modelos) que si sólo se hubieran tenido las respuestas de ED.

Posteriormente, se analizará el efecto que tiene la inclusión de heterogeneidad en los coeficientes estimados. A partir de esto, se elegirá el modelo más adecuado para todas las respuestas y con éste se analizará las preferencias de los individuos por cada atributo y se determinarán las disposiciones al pago por mejorarlos.

### 3 MARCO TEÓRICO

#### 3.1 Modelos de Elección Discreta

En esta investigación los modelos de elección discreta son usados como base para unir las respuestas B-W con las de ED. Estos modelos estiman la probabilidad de elección para distintas alternativas en función de una serie de características o atributos que las definen. Además, se supone que las características de los individuos también pueden afectar la elección. De esta forma, estos modelos permiten estimar la valoración de cada atributo por separado y así derivar la disposición al pago correspondiente a atributos de bienes que no son necesariamente transados en el mercado (Greene y Ortúzar, 2002; Goett *et al.*, 2000; Munizaga *et al.*, 2006; Ortúzar *et al.*, 2000; Torres *et al.*, 2013).

La teoría de la utilidad aleatoria (McFadden, 1974) asume que los individuos se comportan de manera racional, es decir, eligen la alternativa que maximiza su utilidad. El nivel de utilidad de cada alternativa es una combinación definida a partir de sus atributos, en que cada uno es ponderado por su importancia relativa. Así, se supone que cada individuo  $q$  tiene una función utilidad  $U_{iq}$  definida para cada alternativa  $A_i$ . Esta función de utilidad tiene dos componentes: una sistemática (representativa o medible)  $V_{iq}$ , que es función de los atributos del bien mostrado (que a su vez pueden variar entre individuos) y una componente aleatoria  $\varepsilon_{iq}$ . Se supone la existencia de esta componente aleatoria por posibles variaciones en los gustos particulares de cada individuo, y también por posibles errores de medición y/u observación por parte del modelador. Notar que el modelador sólo es capaz de observar una fracción de los elementos de juicio de los individuos, y por tanto, en principio no es capaz de incluir todos los atributos que podrían afectar la función de utilidad individual.

$$(3.1) \quad U_{iq} = V_{iq} + \varepsilon_{iq}$$

La forma más simple y popular de representar la parte sistemática de la función de utilidad, es como una función lineal de los parámetros (Ortúzar 2000):

$$(3.2) \quad V_{iq} = \sum_{k=1}^K \theta_{ik} \cdot x_{ikq}$$

donde  $\theta_{ik}$  es un parámetro a estimar, que se asume constante para todos los individuos, pero puede variar entre alternativas  $A_i$ , y  $x_{ikq}$  representa el atributo  $k$ -ésimo de la alternativa  $A_i$  para el individuo  $q$ .

Como se explicó anteriormente, se asume que el individuo elige la opción que le reporta un mayor nivel de utilidad, por lo que elegirá la alternativa  $A_i$  si y sólo si su utilidad es mayor que el de las otras alternativas ( $A_j$ ):

$$(3.3) \quad U_{iq} \geq U_{jq} \quad \forall A_j \in A(q)$$

Si se reemplaza la ecuación (3.1) en la anterior, se obtiene que el individuo  $q$  elegiría la alternativa  $A_i$  si:

$$(3.4) \quad V_{iq} - V_{jq} \geq \varepsilon_{jq} - \varepsilon_{iq} \quad \forall A_j \in A(q)$$

pero como el modelador desconoce los componentes de error (y por ende su diferencia), no es capaz de determinar con certeza si el individuo efectivamente elegiría esa opción, y sólo se puede estimar la probabilidad de elegir  $A_i$  como:

$$(3.5) \quad P_{iq} = \text{Prob}\{\varepsilon_{jq} \leq \varepsilon_{iq} + V_{iq} - V_{jq}, \quad \forall A_j \in A(q)\}$$

Si no se conoce la distribución de los errores de las alternativas, no es posible obtener una expresión analítica para el modelo.

A continuación se presentarán dos modelos paradigmáticos usados en la investigación, y sus hipótesis sobre la distribución de la componente aleatoria de la función de utilidad.

### 3.1.1 Modelo Logit Simple o Multinomial (MNL)

Este modelo supone que la componente aleatoria de la función de utilidad distribuye independiente e idénticamente (IID) Gumbel con media cero y varianza  $\sigma^2$  (Ortúzar 2000). Uno de sus problemas es que en principio no incluye heterogeneidad alguna en las preferencias individuales. En este caso, extremadamente sencillo, se puede demostrar que la probabilidad que el individuo  $q$  elija la alternativa  $A_i$  está dada por:

$$(3.6) \quad P_{iq} = \frac{e^{\lambda \cdot V_{iq}}}{\sum_{A_j \in A(q)} e^{\lambda \cdot V_{jq}}}$$

donde la varianza de la función densidad de los errores está dada por:

$$(3.7) \quad \sigma^2 = \frac{\pi^2}{6\lambda^2}$$

Es fácil ver que en (3.6) existe un problema de identificabilidad que afecta al parámetro de escala  $\lambda$ , por lo que debe ser normalizado (Ortúzar y Willumsen, 2011).

Para estimar los coeficientes  $\theta_{ik}$  que aparecen en la ecuación (3.2) se utiliza el método de máxima verosimilitud. Ortúzar (2000) explica que la idea detrás de este método es que si bien una muestra observada puede provenir de distintas poblaciones, existe una para la cual es más probable que esto ocurra. En el fondo, el estimador máximo-verosímil es aquel que generaría más a menudo la muestra observada (dada la función densidad de probabilidad supuesta). De esta forma, la función de verosimilitud puede escribirse de la siguiente forma:

$$(3.8) \quad L(\theta) = \prod_{q=1}^Q \prod_{A_j \in A(q)} P_{jq}^{g_{jq}}$$

donde  $g_{jq} = \begin{cases} 1, & \text{si } q \text{ escoge } A_j \\ 0, & \text{en otros casos} \end{cases}$

Como es más conveniente operar con sumas que con productos, se suele trabajar con el logaritmo de este valor (que conduce al mismo máximo):

$$(3.9) \quad \ln L(\theta) = \ell(\theta) = \sum_{q=1}^Q \sum_{A_j \in A(q)} g_{jq} \cdot \ln(P_{jq})$$

Así, en este método se maximiza  $\ell(\theta)$  para encontrar el conjunto de estimadores  $\theta_{ik}$ , máximo verosímiles; estos distribuyen asintóticamente Normal  $(\underline{\theta}, \underline{V})$ , donde  $\underline{\theta}$  representa el vector de parámetros  $\theta_{ik}$ , y  $\underline{V}$  es su matriz de covarianza que, en este caso, está dada por (Ortúzar y Willumsen, 2011):

$$(3.10) \quad \underline{V} = - \left\{ E \left( \frac{\partial^2 \ell(\theta)}{\partial \underline{\theta}^2} \right) \right\}^{-1}$$

Esto permite desarrollar una serie de test estadísticos para analizar la bondad de ajuste del modelo.

### 3.1.2 Modelo Logit Mixto (ML)

El modelo ML tiene la ventaja de permitir agregar heterogeneidad en las preferencias individuales, así como correlación entre alternativas. Las formulaciones más comunes de este modelo son la de *componentes de error* (EC), y la de *coeficientes aleatorios* (RC). En esta investigación sólo será utilizada la primera, por lo que a continuación se profundiza algo más su formulación.

La función de utilidad  $U_{iq}$  en el modelo EC está compuesta de una parte determinística,  $V_{iq}$ , un componente aleatorio  $\varepsilon_{iq}$ , que distribuye IID Gumbel(0,  $\sigma^2$ ) y un segundo término aleatorio  $\eta_{iq}$  con una distribución a especificar (Train, 2009):

$$(3.11) \quad U_{iq} = V_{iq} + \eta_{iq} + \varepsilon_{iq} = \sum_{k=1}^K \theta_{ik} \cdot x_{ikq} + \eta_{iq} + \varepsilon_{iq}$$

En el fondo, este modelo agrega una segunda componente aleatoria  $\eta_{iq}$  a la función de utilidad del MNL, que permite representar diversos efectos de interés, como correlación entre alternativas (Munizaga y Alvarez-Daziano 2002) o entre las respuestas de los mismos individuos (*efecto panel*), entre otras (Train 2009).

La estimación de modelos ML es mucho más compleja que la de modelos MNL. En general, se utiliza máxima verosimilitud simulada (Ortúzar y Willumsen 2011, cap. 8), como se hará en esta investigación.

### 3.2 Teoría Best-Worst

En encuestas del tipo *best-worst* (*B-W*) se presenta una situación de elección, con tres o más alternativas, y en cada una los individuos deben elegir la mejor (*best*) y peor (*worst*) alternativa (Finn y Louviere 1992). Además, se permite que cada individuo procese la elección de forma distinta; por ejemplo, el proceso de elección puede ser simultáneo o secuencial. En el primer caso, conocido como *Max-Diff* o “diferencia máxima”, el individuo busca simultáneamente el par de alternativas (mejor y peor) que tengan la mayor diferencia de utilidad reportada (Louviere 2013). El proceso secuencial, por otro lado, puede realizarse de distintas formas, por ejemplo: (i) primero buscar la mejor alternativa, luego la segunda mejor y así hasta llegar la peor; (ii) primero buscar la peor, luego la segunda peor y así hasta llegar a la mejor, etc. (Louviere 2013).

Finn y Louviere (1992) desarrollan este enfoque que puede ser aplicado de tres formas (Louviere 2013):

1. Caso 1 (Objeto)

En este caso, se presentan distintos objetos al individuo, quien debe elegir cuál es el mejor y cuál es el peor dentro de las posibilidades. Por objetos se refiere a que no

necesariamente poseen atributos ni niveles de atributos. Este enfoque ha sido utilizado en el mercado de vinos, por ejemplo, donde no se definen distintos niveles para los atributos (como textura o color), sino que sólo se evalúa la importancia relativa de cada atributo para los individuos (Louviere 2013). Otra aplicación común se relaciona con evaluar mejoras al sistema de transporte, donde se plantea a los individuos posibles iniciativas como agregar ciclovías, o aumentar los estacionamientos para bicicletas en las cercanías de estaciones de transferencia de transporte público, y el encuestado debe indicar cuál de esas opciones constituiría la mejor inversión (Louviere 2013).

## 2. Caso 2 (Nivel de atributo)

En este caso se presenta al individuo una secuencia de perfiles de elección; cada perfil contiene distintos atributos y cada atributo aparece con un cierto nivel. El individuo debe elegir cuál atributo (nivel) le parece mejor y cuál peor. Cada atributo posee, en total, un cierto número de niveles  $n$ , pero en cada situación (o perfil) de elección sólo se muestra uno de estos niveles. Así, cada elección individual se basa en un nivel determinado de cada atributo (Louviere 2013).

Este caso permite diferenciar la importancia intrínseca de cada atributo de la importancia que tiene la escala de los niveles elegidos. Esta diferencia se puede detectar ya que la elección individual es, en este caso, en base a los niveles del atributo y, por tanto, es posible estimar una función de utilidad para cada atributo por separado (como se verá en mayor detalle en la sección 3.4). En un experimento tradicional de elección discreta, es posible que un atributo parezca ser no significativo porque los niveles presentados en el diseño no varían lo suficiente (aun cuando el atributo efectivamente tenga importancia en la realidad). Lo mismo puede ocurrir al revés, cuando un atributo no relevante parece serlo porque la variación en sus niveles es muy grande. Como mencionan Louviere y Swait (1997), esta distinción resulta fundamental tanto en el área de investigación como en la práctica (por ejemplo, en marketing). Si en un mercado en

que un producto tiene un atributo con niveles muy distintos (marcas muy bien posicionadas versus otras mal posicionadas) se utilizan modelos de elección discreta tradicional, posiblemente se encontraría que ese atributo tiene una importancia relativa muy grande. En ese caso, una marca que se encuentra mal posicionada podría decidir invertir para posicionarse mejor en la escala de ese atributo, a fin de aumentar su participación de mercado. Al hacer esto, la diferencia entre los niveles del atributo existentes en el mercado va a disminuir, y en ese caso la importancia del atributo podría verse reducida y probablemente la marca no aumentaría su participación de mercado de la forma predicha por el modelo.

### 3. Caso 3 (Perfil)

En este caso se presentan al individuo varias alternativas (perfiles) simultáneamente, que son definidas por distintos valores de sus atributos, y este debe elegir cuáles perfiles de elección le resultan más y menos atractivos (en cada situación de elección). Este es el caso más similar al de elección discreta tradicional, ya que la respuesta está asociada a un perfil completo que contiene más de un atributo (en estudios de preferencias declaradas en base a *rankings*, el encuestado debe señalar un orden de preferencia entre distintos perfiles de elección). Además, la respuesta *best*, es decir cuál perfil le parece más atractivo al individuo es comparable (e incluso idéntica) a la elección realizada en una encuesta de elección discreta tradicional. Whitty *et al.* (2013) aplican este caso de B-W para estudiar las preferencias de la población para financiar nuevas tecnologías relacionadas con la salud. Louviere (2013), por otro lado, discute una investigación donde se preguntó a un grupo de estudiantes sus preferencias en vuelos internacionales, para analizar si se obtendría mayor información aplicando este método.

### 3.3 Investigaciones Anteriores sobre Encuestas Best-Worst Caso 2

En la literatura, el Caso 2 *Best-Worst* – que se utiliza en esta tesis – ha sido aplicado principalmente en investigaciones en el área de salud para realizar análisis de costo-utilidad. Flynn (2010), por ejemplo, estudia los principales experimentos que se han diseñado para estimar los valores de QALY (*quality adjusted life year*) incluyendo elección discreta tradicional, rankings y también el caso 2 de *Best-Worst*. El estudio enfatiza las contribuciones de éste último enfoque en su capacidad de permitir separar la importancia de un atributo de los niveles presentados en la encuesta.

Flynn *et al.* (2013) estudian la valoración de atributos sobre la calidad de vida por parte de personas mayores. Se busca comparar las valoraciones estimadas mediante dos experimentos de elección discreta, donde se altera la forma en que la persona se enfrenta a la toma de decisión. Un atributo cualquiera podría ser valorado de manera distinta si la persona se enfrenta, por un lado, a una encuesta que se asemeja a su experiencia personal (por lo que se ve afectado por sus expectativas), y por otro lado, una encuesta que no se asemeja a su propio contexto. El contexto en que se encuentran los encuestados es uno de los focos de investigación de este artículo, que discute un tema importante en el área de psicología matemática, esto es, que la importancia total de un atributo está determinada por el producto entre la utilidad del nivel del atributo y la importancia del atributo en sí (que se ve afectada por el contexto). El Caso 2 de *Best-Worst* permite estimar la importancia del atributo, definida como la utilidad promedio generada por sus niveles.

La encuesta utilizada en Flynn *et al.* (2013), al igual que en otros estudios sobre salud en personas mayores (Coast *et al.* 2008; Couznet *et al.* 2013; Flynn *et al.* 2010;

Szeinbach *et al.* 1999;), se basa en el instrumento ICECAP-O<sup>1</sup>. Este incluye cinco atributos: medida de unión (amor y amistad), seguridad (pensar en el futuro sin preocupaciones), rol (realizar cosas que lo hacen sentir valorado), disfrute (gozo y placer) y control (independencia). Cada atributo posee cuatro niveles. El artículo además plantea una pregunta de elección discreta binaria, en que se solicita comparar la situación descrita en la encuesta con la del individuo en la actualidad. Esto se hace preguntándole si la situación descrita es igual, mejor o peor que la suya.

Louviere (2013) expone cómo se aplicó este caso de B-W para analizar las preferencias sobre planes para tarificar emisiones de contaminantes en Australia (el Gobierno Federal de Australia propuso un proyecto legislativo para cobrar impuestos por la emisión de carbono).

### **3.4 Funciones de Utilidad B-W Caso 2 vs. Elección Discreta Sin Heterogeneidad**

En nuestro análisis, sin heterogeneidad, utilizaremos un modelo *logit* simple (MNL) para estimar las funciones de utilidad tanto de respuestas obtenidas en elecciones B-W como de elección discreta binaria (ED). No obstante, es importante entender que las funciones de utilidad obtenidas serán distintas para cada tipo de encuesta. En efecto, a través de las respuestas correspondientes a la elección B-W, se obtiene una función de utilidad para cada atributo (ya que estas son las opciones de elección).

De esta forma, mediante las respuestas B-W se puede estimar un *intercepto* para cada atributo, que representa el impacto en la función de utilidad como atributo en sí (ya que

---

<sup>1</sup> Ver <http://www.icecap.bham.ac.uk/documents/icecapquest.doc>

es un valor promedio para los niveles definidos), y una pendiente, que representa su *sensibilidad* a los niveles utilizados. Es importante destacar que en el caso de atributos discretos, esto se debe a la codificación utilizada (*effects code*, que para los atributos de dos niveles toma el valor de -1 para el nivel más pequeño y +1 para el nivel mayor, ver Tabla 4-1 para más detalles); vale decir, de no codificarse los datos de la forma planteada, no se podrían deducir las mismas conclusiones. En el caso de atributos continuos usualmente se realiza una normalización entre -1 y +1 con el fin de lograr que las pendientes (sensibilidades) entre los atributos sean comparables en términos de su magnitud. También se debe considerar que si todas las variables tienen el mismo orden de magnitud, se facilita la optimización mediante algoritmos Quasi-Newton (en los cuales se basan muchos programas de estimación).

La función de utilidad que se puede obtener a partir de respuestas de ED no tiene la misma forma, ya que se basa en un conjunto de atributos (y no en cada uno por separado). Por esto, lo que se obtiene es un intercepto común para las respuestas, por ejemplo, “si elegiría la alternativa presentada” (o “no elegiría la alternativa presentada”, dependiendo cuál se tome como base<sup>2</sup>) y una utilidad marginal para cada atributo, que representa nuevamente la *sensibilidad* a los niveles utilizados. La identificación de la función de utilidad en modelos de ED se basa en las diferencias entre las utilidades de las distintas alternativas. En nuestro caso, a partir de las respuestas de elección binaria, ambas alternativas (si/no en este caso) se posicionan en forma relativa en la misma escala de utilidad.

---

<sup>2</sup> Esto se puede ver con más claridad en el capítulo siguiente, donde se explicita el diseño de la encuesta.

Entonces, es posible comparar las sensibilidades obtenidas para cada atributo a partir de las respuestas de elección *B-W* y las respuestas de ED, suponiendo que los dos tipos de respuesta (*B-W* y ED) miden la misma variable latente. Esto se puede realizar ya que en ambos casos las sensibilidades están relacionadas con los niveles de los atributos presentados en la encuesta. Por lo mismo, ambas representan el impacto de los niveles del atributo en la función de utilidad. Aun así, no existe una equivalencia en las respuestas de ED para el intercepto estimado para cada atributo mediante las respuestas *B-W*. Esto ocurre ya que, como fue mencionado anteriormente, la elección en la encuesta de ED se basa en el perfil completo y, por lo mismo, no se puede estimar un promedio entre los niveles presentados para cada atributo, por separado.

Es fundamental entender que la función de utilidad también se explicita de forma distinta para el caso *best* que para el *worst*, así como también es distinta para el caso “si elegiría la alternativa presentada” que cuando “no elegiría la alternativa presentada”, ya que cambia la probabilidad en los distintos tipos de elección.

Como explican Louviere y Swait (1997), la probabilidad de elegir un atributo como el mejor (*best*) corresponde a la probabilidad de que la utilidad de ese atributo sea mayor que la del resto. A su vez, la probabilidad de elegir un atributo como el peor (*worst*), es igual a la probabilidad de que su utilidad sea menor que la del resto. Entonces, teóricamente la probabilidad de elegir un atributo como el mejor debería ser la función inversa de elegirlo como el peor. Por lo mismo, para las respuestas *worst* se codificará los niveles de los atributos y sus interceptos como el inverso del caso *best*. Aun así, el proceso mental que lleva a cabo una persona para elegir un atributo como mejor no es necesariamente el mismo que para escoger el peor; esto se puede reflejar en diferencias entre los coeficientes de las diferentes funciones de utilidad y será examinado cuando se presenten los resultados de estos modelos en los capítulos 5 y 6, incluyendo un factor de escala entre ambos tipos de respuesta.

Explícitamente, para los datos disponibles en esta tesis, las funciones de utilidad de los distintos tipos de respuesta toman la siguiente forma:

**Encuesta *Best*:**

$$U_{\text{metro}} = C_{\text{metro\_Best}} + \theta_{\text{metro\_Best}} \cdot X_{\text{metro}} + \varepsilon_{\text{Best}}$$

$$U_{\text{áreas verdes}} = C_{\text{áreas verdes\_Best}} + \theta_{\text{áreas verdes\_Best}} \cdot X_{\text{áreas verdes}} + \varepsilon_{\text{Best}}$$

y así sucesivamente para cada atributo.

**Encuesta *Worst*:**

$$U_{\text{metro}} = -C_{\text{metro\_Worst}} + \theta_{\text{metro\_Worst}} \cdot (-X_{\text{metro}}) + \varepsilon_{\text{Worst}}$$

$$U_{\text{áreas verdes}} = -C_{\text{áreas verdes\_Worst}} + \theta_{\text{áreas verdes\_Worst}} \cdot (-X_{\text{áreas verdes}}) + \varepsilon_{\text{Worst}}$$

y así sucesivamente para cada atributo.

**Encuesta *ED*:**

$$U_{\text{Si lo arrendaría}} = C_{\text{si}} + \theta_{\text{metro\_ED}} \cdot X_{\text{metro}} + \theta_{\text{áreas verdes\_ED}} \cdot X_{\text{áreas verdes}} + \dots + \varepsilon_{\text{ED}}$$

$$U_{\text{No lo arrendaría}} = 0 + \varepsilon_{\text{ED}}$$

donde  $C$  representa el *intercepto* de cada atributo,  $\theta$  representa la *sensibilidad* descrita anteriormente y los  $\varepsilon_i$  distribuyen IID Gumbel  $(0, \lambda_i)$ . Como se verá en la aplicación empírica, si se supone que las distribuciones para las respuestas *best* y *worst* son idénticas, exceptuando su signo, y que las sensibilidades de los atributos son iguales entre ellas e iguales a las obtenidas mediante las respuestas de ED (en cuyo caso los parámetros de sensibilidad son equivalentes a las utilidades marginales de cada atributo), se debe agregar las siguientes restricciones:

$$\theta_{\text{metro\_Best}} = \theta_{\text{metro\_Worst}} = \theta_{\text{metro\_ED}}$$

$$\theta_{\text{áreas verdes\_Best}} = \theta_{\text{áreas verdes\_Worst}} = \theta_{\text{áreas verdes\_ED}}$$

y así sucesivamente para cada atributo. El cumplimiento de estas restricciones, que se puede probar estadísticamente, significa que efectivamente los atributos se pueden

considerar comunes entre las respuestas. Si no se cumplen en algunos casos, los atributos en cuestión deben considerarse específicos para cada tipo de respuesta.

Si además, se asume que los interceptos son idénticos entre ellos, excepto por su signo, se agregarían las siguientes restricciones:

$$C_{\text{metro\_Best}} = C_{\text{metro\_Worst}}$$

$$C_{\text{áreas verdes\_Best}} = C_{\text{áreas verdes\_Worst}}$$

y así sucesivamente para cada atributo.

Notar que al hacer estos supuestos, también se asume que todos los errores Gumbel distribuyen con la misma varianza (un único valor de  $\lambda$ ).

Luego de haber estimado el intercepto (mediante las respuestas *best* y/o *worst*) y la sensibilidad de cada atributo (mediante las respuestas ED, *best* y/o *worst*), estos se pueden posicionar en una escala relativa. Esta se utiliza para analizar el impacto del atributo por sí solo y el impacto de la diferencia entre los niveles presentados en la encuesta. Es necesario mencionar que, como la estimación se hace a partir de las diferencias entre las funciones de utilidad, el intercepto de uno de los atributos se debe considerar como base (esto es, igual a cero). De esta forma, el impacto que tiene el resto de los atributos por sí solos (intercepto), estará estimado en relación al atributo considerado como base. Además, cada nivel de atributo se puede posicionar en la misma escala relativa, y su impacto se calcula como la suma entre el intercepto y la sensibilidad multiplicada por el nivel del atributo. Así, la diferencia entre las posiciones de los niveles representa el impacto de la variación entre los niveles del atributo. Finalmente, es interesante también analizar el grado de dispersión de los atributos (por sí solos o debido a sus niveles) en la escala relativa. Esto se puede analizar mediante la desviación estándar de la escala relativa, que representa su grado de dispersión.

### 3.5 Combinación de Respuestas de Dos Bancos de Datos

Las respuestas *best*, *worst* y ED tienen distintas funciones de utilidad y cada una constituye un banco de datos distinto. Así, es importante analizar si los coeficientes estimados en las tres fuentes de datos pueden considerarse iguales, o si sería necesario incluir uno o más factores de escala ( $\lambda$ ). Incluir un factor de escala es esencialmente un supuesto sobre los términos de error considerados en las ecuaciones anteriores, que en principio son independientes entre sí:

$$\varepsilon_{\text{Best}} \sim \text{IID Gumbel}(0, \lambda_{\text{Best}})$$

$$\varepsilon_{\text{Worst}} \sim \text{IID Gumbel}(0, \lambda_{\text{Worst}})$$

$$\varepsilon_{\text{ED}} \sim \text{IID Gumbel}(0, \lambda_{\text{ED}})$$

De esta forma, al considerar la posibilidad de combinar distintas fuentes de datos, las funciones de utilidad quedan representadas como (esto no ocurría en el caso anterior, pues el factor de escala no es identificable al considerar un solo tipo de datos):

#### Encuesta Best:

$$U_{\text{metro}} = \lambda_{\text{Best}} \cdot (C_{\text{metro\_Best}} + \theta_{\text{metro\_Best}} \cdot X_{\text{metro}}) + \varepsilon_{\text{Best}}$$

$$U_{\text{áreas verdes}} = \lambda_{\text{Best}} \cdot (C_{\text{áreas verdes\_Best}} + \theta_{\text{áreas verdes\_Best}} \cdot X_{\text{áreas verdes}}) + \varepsilon_{\text{Best}}$$

y así sucesivamente para cada atributo.

#### Encuesta Worst:

$$U_{\text{metro}} = \lambda_{\text{Worst}} \cdot (-C_{\text{metro\_Worst}} + \theta_{\text{metro\_Worst}} \cdot -X_{\text{metro}}) + \varepsilon_{\text{Worst}}$$

$$U_{\text{áreas verdes}} = \lambda_{\text{Worst}} \cdot (-C_{\text{áreas verdes\_Worst}} + \theta_{\text{áreas verdes\_Worst}} \cdot -X_{\text{áreas verdes}}) + \varepsilon_{\text{Worst}}$$

y así sucesivamente para cada atributo.

#### Encuesta ED:

$$U_{\text{Si lo arrendaría}} = \lambda_{\text{ED}} \cdot (C_{\text{si}} + \theta_{\text{metro\_ED}} \cdot X_{\text{metro}} + \theta_{\text{áreas verdes\_ED}} \cdot X_{\text{áreas verdes}} + \dots) + \varepsilon_{\text{ED}}$$

$$U_{\text{No lo arrendaría}} = 0 + \varepsilon_{\text{ED}}$$

donde la varianza ( $\sigma_i^2$ ) de la componente de error para el tipo de respuesta  $i$  (ED, *best*, *worst*) se puede expresar de la siguiente forma (Ortúzar y Willumsen, 2011, cap. 7):

$$(3.12) \quad \sigma_i^2 = \frac{\pi^2}{6\lambda_i^2}$$

Por esto, un mayor factor de escala implica una menor varianza del error. Debido al problema de identificabilidad, no es posible estimar todos los factores de escala pertenecientes a cada tipo de respuesta en un modelo conjunto. Por eso, se considerará que el correspondiente a las respuestas ED es igual a uno ( $\lambda_{ED}=1$ ), lo que permite estimar el factor de escala para las respuestas restantes,  $\lambda_{Best}$  y  $\lambda_{Worst}$ . Se estudiará si es necesario agregar factores de escala entre las respuestas ED, *best* y *worst* y, a través de ellos, se analizará si efectivamente sus componentes de error tienen distinta varianza.

Al combinar fuentes de datos distintas, además se verificará si los coeficientes estimados pueden, en realidad, considerarse comunes entre las respuestas. Para esto, se utilizará el estadístico de razón de verosimilitud (Ortúzar y Willumsen, 2011, pag. 325):

$$(3.13) \quad \lambda_a = -2(L_1 + L_2 - L_\mu)$$

donde  $L_\mu$  representa la log-verosimilitud del modelo conjunto y  $L_i$  la log-verosimilitud de cada modelo por separado;  $\lambda_a$  distribuye  $\chi^2$  con  $a$  grados de libertad, donde  $a$  es el número de restricciones lineales necesarias para pasar del modelo general (estimación de ambos modelos por separado) al modelo restringido (esto es, que ciertas sensibilidades de los atributos sean iguales para las distintas respuestas); en este caso, sería el número de parámetros considerados comunes entre las respuestas.

Si  $\lambda_a$  supera el valor de tablas de  $\chi^2$  con  $a$  grados de libertad para el nivel de confianza requerido, se rechaza la hipótesis nula que el modelo conjunto es equivalente a los modelos por separado, indicando que no todos los parámetros considerados como comunes a ambas fuentes de datos lo son en realidad.

Es importante mencionar que como se desea saber si los modelos son efectivamente equivalentes (i.e. que considerar los  $a$  atributos comunes entre respuestas sea válido), esperamos que no exista suficiente evidencia para rechazar la hipótesis nula. En general, se buscará que la hipótesis nula no sea rechazada al 95% de nivel de confianza, pero en caso de ser necesario se buscará con qué nivel de confianza no sería rechazada. Para esto, se necesitará aumentar el nivel de confianza, ya que eso implica que la hipótesis nula tenga menor probabilidad de ser rechazada. Esto resulta contra intuitivo en una primera mirada, ya que al aumentar el nivel de confianza puede ser posible lograr no rechazar la hipótesis nula (que es lo buscado); por esto es importante entender el significado del *test* y del nivel de confianza postulado al analizar los resultados.

### **3.6 Inclusión de Heterogeneidad**

En esta tesis se procederá, en primer lugar, a incluir heterogeneidad mediante componentes observables (variables socio-económicas). Además, se incluirá una componente no observable que permita heterogeneidad en las preferencias tratando el efecto panel en los datos mediante componentes de error (ML visto en sección 3.1.2). Hasta donde llega nuestro conocimiento, no hay investigaciones anteriores que hayan incluido ambos tipos de heterogeneidad de manera conjunta en datos B-W.

#### **3.6.1 Variaciones Sistemáticas en los Gustos (VSG)**

Las variaciones sistemáticas en los gustos permiten incluir una componente de aleatoriedad a través de hacer interactuar variables socio-económicas con los atributos de los modelos, con el fin de que los coeficientes estimados no sean fijos para toda la población, sino que varíen por estrato (individuos con distintas características socio-económicas, ver discusión en Ortúzar y Willumsen 2011, p. 279).

Se considerará posibles VSG en la sensibilidad de los atributos, es decir, en las variables que dependen del nivel de cada uno. Así, se podrá estudiar la heterogeneidad en las preferencias medida a través de encuestas ED en conjunto con encuestas *best* y/o *worst*.

Es importante mencionar que al unir las respuestas ED, *best* y *worst*, se analizará qué coeficientes se pueden considerar comunes entre ellas. En el caso de atributos con VSG se estimará el conjunto de coeficientes (del atributo básico y sus interacciones) como específicos (o comunes) para los distintos tipos de respuesta (Louviere *et al.* 2000).

### 3.6.2 Efecto Panel

El correcto tratamiento del *efecto panel* inherente a los datos utilizados en esta tesis, permite incluir heterogeneidad que no puede ser observada mediante variables, al posibilitar detectar la correlación entre las respuestas de un mismo individuo. Para esto se debe especificar un modelo logit mixto (ML) muy sencillo, que incluye componentes de error, como se vio con más detalle en la sección 3.1.2.

## 3.7 Tests Utilizados

### 3.7.1 Tests para Modelos *Best*, *Worst*, ED

#### *Test de Razón de Verosimilitud (LR)*

El test de razón de verosimilitud (LR) se puede aplicar en el caso que un modelo sea una versión restringida de otro (Ortúzar 2000). En este caso, considerando que  $l(\hat{\theta})$  es la log verosimilitud del modelo general y  $l(\hat{\theta}_r)$  es la log verosimilitud del modelo restringido, se define al estadígrafo LR como:

$$(3.14) \quad LR = -2 \cdot (l(\hat{\theta}_r) - l(\hat{\theta}))$$

Este estadígrafo distribuye asintóticamente  $\chi^2$  con  $r$  grados de libertad, donde  $r$  representa el número de restricciones lineales que tiene el modelo restringido en relación al general. La hipótesis nula de este test dice que el modelo restringido es equivalente al modelo general, en cuyo caso sería adecuado para la muestra. La hipótesis nula se rechaza si  $LR > \chi^2_{n,\alpha}$ , donde  $\alpha$  es el nivel de confianza usado. Es importante entender que este *test* es el mismo descrito en la sección 3.5 para analizar si algunos parámetros estimados se pueden considerar comunes entre dos fuentes de datos distintas. Sin embargo, cuando se aplique en este último escenario será descrito como test  $\lambda_a$ , y cuando se aplique para verificar si el modelo mejora al estimar más coeficientes (lo descrito en esta sección) será descrito como test LR.

### 3.7.2 Indicadores para Analizar Capacidad Predictiva

Primero, se utilizarán directamente los coeficientes encontrados y luego agregaremos un factor de escala (FE) para estas utilidades, que permite que los modelos estimados se adapten mejor a la muestra de validación. Se considera necesario incluir este factor de escala pues las muestras pertenecen a individuos distintos (aunque de características similares, ya que ambas corresponden a personas que buscaban arrendar departamento en el centro de Santiago). En el Anexo A se describen otros *tests* e indicadores utilizados para analizar la capacidad predictiva de los modelos en la MV, entre ellos el FPR, ER, CR,  $\chi^2_{\text{FPR}}$ , e índice de éxito normalizado ( $\sigma_{\text{norm}}$ ). Acá, sólo se presentarán los indicadores que se considera más relevantes para discriminar entre los modelos en base a su capacidad predictiva.

#### *Log verosimilitud y errores residuales*

La log verosimilitud se explicó en más detalle en la sección 3.1.1. Este es el indicador más relevante al analizar el comportamiento de modelos en la muestra de validación, pues representa una medida exacta que no está sujeta a restricciones.

Los errores residuales son una medida más desagregada que la log verosimilitud, ya que consideran un promedio del error de predicción para cada individuo. Estos errores pueden ser calculados de distintas formas (McCullagh y Nelder 1989); en esta tesis se utilizará su forma más general:

$$(3.15) \quad \text{Errores Residuales} = \sqrt{\frac{\sum_{q=1}^Q \sum_{A_i \in A(q)} (Po_{iq} - Pe_{iq})^2}{Q}}$$

donde  $Po_{iq}$  representa la probabilidad con que el individuo  $q$  eligió la alternativa  $A_i$ , es decir, toma valor uno para la alternativa elegida y cero para el resto;  $Pe_{iq}$  representa la probabilidad que el modelo asignó a que el individuo  $q$  eligiera la alternativa  $A_i$ . Así, se saca un promedio para el error residual de cada alternativa para cada individuo. A priori no se esperan grandes variaciones en este indicador.

### 3.8 Disposición al Pago por Atributos

Para estimar la disposición al pago por mejorar distintos atributos (*willingness to pay*, WTP) de la población o muestra observada, se requiere que uno de los atributos considerados sea el precio o costo directo. La WTP se define como la tasa marginal de sustitución entre el atributo de interés ( $X_{ik}$ ) y el costo ( $X_{ic}$ ):

$$(3.16) \quad \text{WTP}(X_{ik}) = \frac{\partial V_i / \partial X_{ik}}{\partial V_i / \partial X_{ic}}$$

La WTP en un modelo MNL se puede calcular simplemente de manera sencilla, ya que se estima la función utilidad directamente y por lo tanto resulta ser igual a la razón entre el coeficiente estimado para el atributo de interés y el coeficiente estimado para el atributo que representa el costo (Gaudry et al. 1989).

Debido a que en esta investigación no se utilizarán parámetros aleatorios, no es necesario indagar más sobre cómo estimar las WTP en estos casos, ya que para los ML con errores compuestos se puede aplicar las mismas ecuaciones que para el MNL.

## 4 PRESENTACIÓN Y ANÁLISIS DE LOS DATOS DISPONIBLES

### 4.1 Atributos Presentados en la Encuesta

La encuesta utilizada en esta tesis fue aplicada por Ignacia Torres a individuos que planeaban arrendar un departamento en la comuna de Santiago (Torres, 2009). Con el fin de estudiar y comparar la metodología *best-worst* (B-W) con la de elección discreta binaria (ED), se consideró siete atributos con dos niveles cada uno, y un atributo (*Dividendo del departamento*) con ocho niveles. Estos atributos, que se describen en la Tabla 4-1, están agrupados en torno a: (i) características del barrio (*Acceso al metro, Áreas verdes, Servicios comerciales básicos y Servicios culturales, Mantención y estado de conservación de la calle*) y (ii) características de la vivienda (*Tamaño de dormitorios, Gimnasio y piscina en el edificio*, y, por supuesto, una variable precio, especificada como el costo de arriendo o *Dividendo del departamento*).

A cada individuo se le presentaron ocho perfiles de elección y en cada uno de ellos cambiaban los niveles mostrados para cada atributo. Así, todos observaron en alguna instancia todos los niveles de los distintos atributos, exceptuando el valor del dividendo; en este último caso, los encuestados sólo veían dos niveles que dependían del número de habitaciones que habían señalado requerir en la parte inicial de la encuesta. Los dos valores mostrados dependían del precio de mercado de un departamento con ese número de habitaciones (un valor estaba por debajo y el otro sobre el precio de mercado).

### 4.2 Diseño de la Encuesta

En el Anexo B se muestra la encuesta presentada a los individuos. Ellos debían elegir cuál nivel de atributo les parecía mejor o más atractivo (*best*), cuál les parecía peor (*worst*), y finalmente señalar si arrendarían o no el departamento mostrado. Esta última respuesta es la que hemos denominado elección discreta binaria (ED).

Tabla 4-1: Atributos Presentados en la Encuesta y Codificación de sus Niveles

<b>Atributo</b>	<b>Niveles</b>	<b>Código</b>
Acceso a metro	Lejos, a más de cinco cuadras del departamento	+1
	Cerca, a menos de cinco cuadras del departamento	-1
Áreas verdes	Lejos, a más de cinco cuadras de su departamento	+1
	Cerca, a menos de cinco cuadras del departamento	-1
Dividendo departamento	\$120.000	-1
	\$140.000	-0,819
	\$171.000	-0,538
	\$199.000	-0,285
	\$217.000	-0,122
	\$253.000	+0,204
	\$293.000	+0,567
	\$341.000	+1
Gimnasio y piscina en el edificio	Sí tiene	+1
	No tiene	-1
Mantenimiento/estado de conservación de la calle	Buena	+1
	Mala	-1
Servicios comerciales básicos	Lejos, a más de cinco cuadras del departamento	+1
	Cerca, a menos de cinco cuadras del departamento	-1
Servicios culturales	Lejos, a más de cinco cuadras del departamento	+1
	Cerca, a menos de cinco cuadras del departamento	-1
Tamaño de dormitorios	Amplios	+1
	Normal	-1

### 4.3 Codificación

Para comparar los coeficientes obtenidos a partir de la elección *B-W* y ED (como se explica en la sección siguiente), se debió codificar las respuestas utilizando el sistema

*effects code* para las variables discretas (Louviere y Swait, 1997); las variables continuas se codificaron normalizando sus valores entre -1 y +1. El atributo *Dividendo del departamento* es el único continuo en este caso, por lo que fue codificado con una normalización de los parámetros, donde su menor nivel (\$ 120.000) tomó el valor -1 y su mayor nivel (\$ 341.000) el valor +1. Los niveles entre estos dos valores fueron normalizados siguiendo esta escala (Tabla 4-1).

#### **4.4 Muestra de Validación**

Torres *et al.* (2013) realizaron un estudio en la comuna de Santiago Centro sobre las disposiciones al pago por atributos de barrio y vivienda. Se preguntó a cada individuo, sus características socioeconómicas y luego se aplicó una encuesta de preferencias declaradas (PD) en que debían elegir una de dos viviendas en doce escenarios. Cada vivienda mostrada estaba descrita por ocho atributos con un nivel cada uno. Estos niveles son los mismos que se definió para la encuesta *B-W*. Esta encuesta se utilizará como muestra de validación para los modelos planteados en esta tesis.

Torres (2009) entrevistó a 497 individuos quienes debieron considerar 12 perfiles de elección, por lo que este banco de datos cuenta con 5.965 observaciones. Es importante mencionar que estos 497 individuos son distintos a los que respondieron las encuestas *B-W* y ED, por lo que será necesario incluir un factor de escala a la hora de validar nuestros modelos, como se comentó en la sección 3.7.2.

#### **4.5 Análisis de Variables Socio-Económicas en el Banco de Datos**

Antes de estimar modelos, es necesario analizar las observaciones que serán utilizadas para este fin. En las encuestas que no fueron utilizadas en la tesis de Ignacia Torres, no sólo se pregunta por los atributos más y menos atractivos y si las personas arrendarían o

no el departamento considerado, sino también por características socioeconómicas del individuo encuestado. Entre ellas se pregunta:

- Edad: todos contestan (las respuestas varían entre 19 y 69 años)
- Sexo: todos contestan
- Educación: todos contestan (las respuestas varían entre universitaria completa, incompleta, técnico-profesional y secundaria)
- Ocupación: todos contestan (las opciones son estudia, trabaja u otros).
- Comuna de origen: todos contestan (las opciones son fuera de Chile, fuera de Santiago, Santiago Centro, y prácticamente todas las comunas de la capital)
- Número de autos: todos contestan (las respuestas varían entre 0, 1, 2 y 3 o más)
- Número de individuos que habitarían el nuevo departamento: todos contestan (las respuestas varían entre 1 y 5)
- Cuánto tiempo planean vivir en el nuevo departamento: todos contestan (las opciones son menos de 5 años, 5 a 10 años, y más de 10 años)
- Número de habitaciones: todos contestan (las respuestas varían entre 1 y 4)
- Ingreso personal: 15 de las 217 personas no contestan (afectando 120 observaciones) y cinco personas ponen cero (afectando 40 observaciones)
- Ingreso total: 13 personas no contestan (afectando 104 observaciones) y una persona puso cero (afectando ocho observaciones).

Se buscaron inconsistencias en la muestra y se encontró que un individuo declaró tener 3 o más autos en el hogar en circunstancias que vivía sólo una persona. Este dato se eliminó, pese a no tener un valor erróneo necesariamente, ya que constituía un *outlier*. Se encontraron, además, seis individuos que declararon un ingreso familiar menor que su ingreso personal. Torres (2009) contactó a dos de tres individuos que presentaron

esta misma característica en su muestra y ambos reconocieron que no habían sumado su ingreso personal al de la familia. Considerando que este banco de datos fue obtenido de la misma forma y que la encuesta fue diseñada por las mismas personas (pese a que no fue aplicada a los mismos individuos), se decidió que cinco de estos seis individuos posiblemente habían cometido el mismo error que los mencionados en Torres (2009) y por ende se sumó su ingreso personal al declarado como familiar para obtener su verdadero ingreso familiar mensual total (IFAM). El individuo remanente reportó un ingreso personal muy alto (\$7.500.000) para una persona promedio, y un ingreso familiar (\$3.500.000) que de sumarse al anterior daría un IFAM muy superior al del resto de la muestra. Por eso, se consideró que esta persona se había equivocado al contestar, y se mantuvo su “ingreso personal” como el IFAM.

Como el ingreso familiar mensual se considera un atributo importante al evaluar las variaciones sistemáticas de gusto (VSG), se procedió a eliminar las observaciones que no declararon ingreso familiar mensual. Bajo este criterio se eliminó las respuestas de otros 14 de los 217 individuos (aproximadamente 6,5% de las respuestas). Por ende, en total se eliminó las respuestas de 15 individuos (aproximadamente 6,9% del total).

Los 202 individuos finalmente considerados para estimar modelos, respondieron adecuadamente la pregunta ED en las ocho situaciones de elección, por lo que se obtuvo 1.616 observaciones de este tipo. Sin embargo, no todos contestaron cuál atributo les parecía más y/o menos atractivo en los ocho escenarios presentados. De hecho, sólo se obtuvo 1.569 observaciones sobre el atributo más atractivo (*best*), y 1.506 sobre el menos atractivo (*worst*). En total se obtuvo 4.691 observaciones incluyendo los tres tipos de respuestas.

Como la encuesta se aplicó aleatoriamente a individuos que buscaban arrendar un departamento en la comuna de Santiago, parece relevante analizar sus características socioeconómicas con el fin de analizar qué variables podrían ser incluidas en el análisis

de VSG. En primer lugar, se pudo constatar que un mayor porcentaje de mujeres (64%) contestó la encuesta (ver Figura A-2 del Anexo C). Por otro lado, la distribución de la variable *Edad* de los individuos encuestados se muestra en la Figura A-3 del Anexo C. Es destacable que casi la mitad de los encuestados (49%) tiene entre 26 y 35 años de edad; el segundo grupo etario está formado por las personas entre 16 y 25 años que alcanzan un 30% de la muestra. Las personas entre 36 y 45 años de edad son un 13% de la muestra, y las mayores a 46 años alcanzan sólo al 8% de la muestra.

En la Figura A-4 del Anexo C se puede ver la distribución por comuna de origen de los encuestados. La comuna de Santiago aporta un 34% del total, seguido por las comunas de Ñuñoa y Maipú con un 8% cada una, luego La Florida (7%) y Providencia (6%). Las comunas de Lo Barnechea y Quilicura son las únicas sin personas entrevistadas; todas las restantes tienen por lo menos 10 personas encuestadas, lo que sugiere que existe migración desde toda la ciudad a la comuna de Santiago. Se espera que las comunas con mayores porcentajes de encuestados tengan mayor impacto en términos de posibles variaciones sistemáticas de gustos. Por otro lado, la Figura A-5 del Anexo C muestra que la mayoría de los individuos tiene un auto en su hogar (55%), seguido por la gente que no tiene autos (34%) y las personas con dos autos (11%); ningún encuestado tenía más de dos autos en el hogar. El número de autos en el hogar debiera afectar la utilidad asociada a la cercanía a distintos servicios del barrio; por ejemplo, si hay autos en el hogar la utilidad de estar cerca del metro debiera disminuir. Ahora bien, como la disponibilidad de autos en un hogar no sólo depende del número de autos sino también del número de habitantes en el hogar, parece interesante analizar la interacción entre estas variables.

En la Figura 4-1 se puede ver que no existen hogares con un habitante que tengan dos autos. Además, se puede ver que la mayoría de los hogares con un auto tienen dos habitantes. Así, la disponibilidad de autos en un hogar puede ser analizada dependiendo de su número de habitantes. Por lo mismo, cuando se consideren posibles variaciones

sistemáticas en los gustos, se incluirá el cociente entre las variables número de autos y número de habitantes, para estudiar sus posibles efectos en la valoración de los distintos atributos de vivienda y barrio.

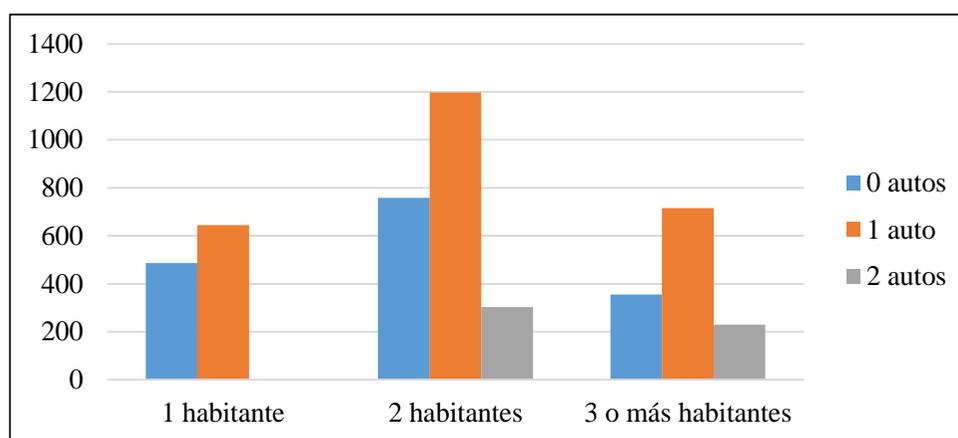


Figura 4-1: Habitantes vs. Número de Autos en el Hogar

En la Figura A-6 del Anexo C se puede ver la distribución de la variable *Número de Habitantes* en la muestra. Casi la mitad (48%) de los hogares encuestados tiene dos habitantes, el 24% tiene un habitante, el 19% tres habitantes y el 9% restante considera hogares con cuatro y cinco habitantes. Como estos últimos dos porcentajes no alcanzan el 10% de la muestra, al analizar VSG se podrán agrupar con el intervalo anterior para lograr un valor más representativo. Es decir, se podrán agrupar los hogares con tres o más habitantes.

En la Figura A-7 del Anexo C se ve que el mayor porcentaje de los encuestados planea quedarse entre 5 a 10 años en el departamento (41%); un porcentaje muy cercano planea quedarse más de 10 años (38%) y el menor porcentaje corresponde a quienes planean quedarse en el departamento menos de 5 años (21%).

En la Figura 4-2 se muestra la distribución de la variable *Ingreso Personal Mensual* en el banco de datos. Se puede ver que la mayoría de la población concentra sus ingresos personales entre \$ 200.000 y \$ 800.000.

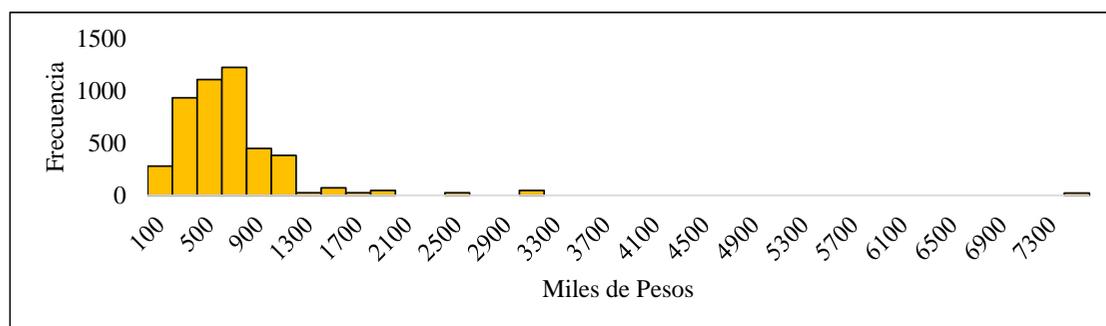


Figura 4-2: Ingreso Personal Mensual de los Individuos en el Banco de Datos

En la Figura 4-3 se muestra la distribución de la variable IFAM; en este caso la mayoría de la población se concentra entre los \$ 200.000 y \$ 1.300.000. Para comparar, se procedió a graficar ambas distribuciones en conjunto en la Figura 4-4, mostrando que – obviamente - la variable *Ingreso Personal* es siempre una cota inferior del *IFAM*. Además, se puede afirmar que la variable *Ingreso Personal* no es siempre un buen indicador del ingreso del hogar porque depende de quién contestó la encuesta, y esa persona en algunos casos puede haber sido el proveedor primario del hogar pero en otros, algún integrante de la familia que no aporte económicamente. Por lo mismo, al considerar variaciones sistemáticas de los gustos se optará por usar la variable *IFAM*.

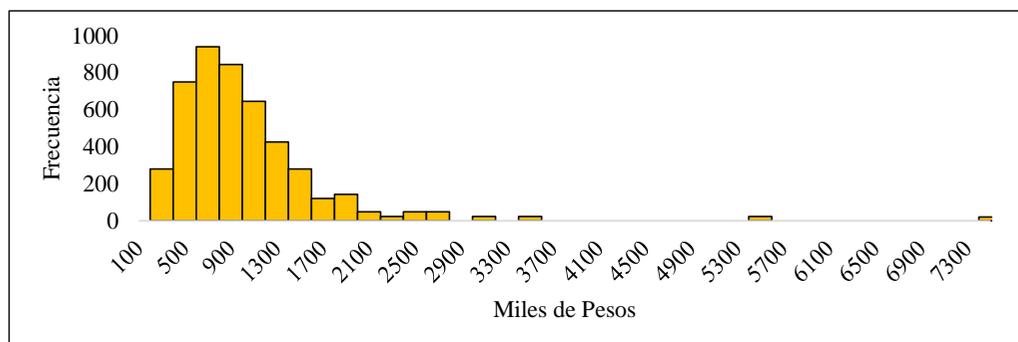


Figura 4-3: Ingreso Familiar Mensual de los Hogares en el Banco de Datos

En la Figura A-8 del Anexo C se muestra la distribución de la variable *Estudios*. Ninguno de los encuestados reportó no tener estudios o haber alcanzado sólo un nivel primario. Además, sólo un 3% de la población tiene estudios secundarios, esto es, prácticamente toda la población cuenta con estudios superiores (ya sea completos o incompletos), por lo que es esperable que esta variable no tenga gran impacto al probar VSG.

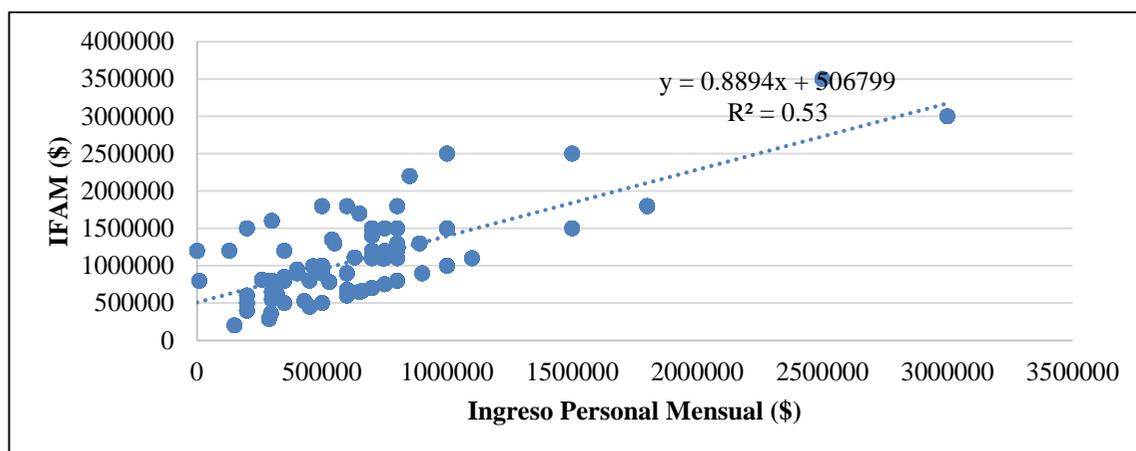


Figura 4-4: Dispersión IFAM vs Ingreso Personal

En las Figura A-9 y Figura A-10 del Anexo C se muestra la distribución de la variable *Ocupación*. Se puede ver que la mayoría de los encuestados trabaja (84%), mientras que sólo un 9% estudia y en “otros” hay sólo un 7%. Por lo mismo tampoco se espera gran incidencia de esta variable en los modelos.

En la Figura A-11 del Anexo C se muestra la distribución del *Número de dormitorios* que las personas requerían. Se puede ver que un 45% de los encuestados necesitaba un departamento con dos dormitorios, seguido por un 42% de los encuestados que necesitaba tres dormitorios. El 11% de la muestra requería sólo un dormitorio, y el 2% de la muestra cuatro dormitorios. Es importante mencionar que los valores de los niveles del atributo *Dividendo* considerados en la encuesta dependían de la cantidad de

dormitorios que requería la familia; los niveles mostrados en las encuestas, dependiendo del número de habitaciones, se presentan en la Tabla 4-2 .

Tabla 4-2: Niveles del Atributo *Dividendo* según Número de Habitaciones Requeridas

Número de Habitaciones	Nivel Inferior Atributo <i>Dividendo</i>	Nivel Superior Atributo <i>Dividendo</i>
1	\$ 120.000	\$ 140.000
2	\$ 171.000	\$ 199.000
3	\$ 217.000	\$ 253.000
4	\$ 293.000	\$ 341.000

En la Figura 4-5 se muestra la distribución del número de dormitorios según la variable *Número de habitantes*. Se puede ver que la mayoría de los hogares con uno y dos habitantes aspiraban a un departamento con dos habitaciones, y que los hogares con tres o más habitantes requerían, en su mayoría, un departamento con tres dormitorios.

#### **4.6 Comparación de Características Individuales en el Banco de Datos y Muestra de Validación**

Debido a que se utilizará una muestra de validación (MV) para analizar la capacidad predictiva de los modelos estimados con el banco de datos (BD), será relevante analizar las características socioeconómicas de los individuos en ambos casos. De todas formas, ambas muestras corresponden a personas que buscaban arrendar departamento en la comuna de Santiago en el 2009, por lo que se podría esperar un comportamiento similar.

Aun así, ambas muestras no son grandes; de hecho, en la MV se cuenta con las respuestas de 497 individuos, mientras que en el BD - tras remover inconsistencias y a los individuos que no entregaron su ingreso - se cuenta con las respuestas de sólo 202

individuos. Por lo mismo, se podría esperar que existan diferencias en las características socio-económicas y resultaría interesante detectarlas.

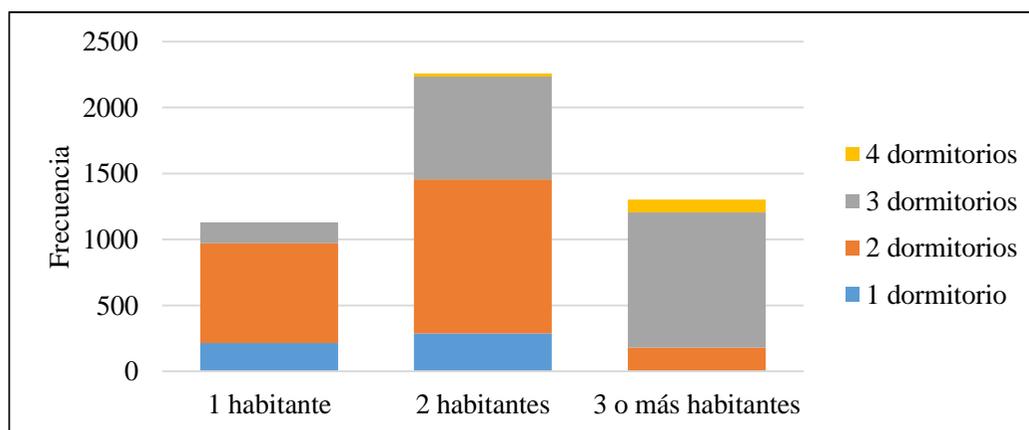


Figura 4-5: Distribución Número de Dormitorios según *Número Habitantes*

Aun así, ambas muestras no son grandes; de hecho, en la MV se cuenta con las respuestas de 497 individuos, mientras que en el BD - tras remover inconsistencias y a los individuos que no entregaron su ingreso - se cuenta con las respuestas de sólo 202 individuos. Por lo mismo, se podría esperar que existan diferencias en las características socio-económicas y resultaría interesante detectarlas.

En primer lugar se revisó la distribución de las variables *Sexo*, *Edad*, *Comuna de origen*, *Autos*, *Tiempo de estadía*, *Ingreso*, *Dormitorios* y *Habitantes*, no encontrando diferencias significativas en sus distribuciones. Aun así, no se considera que esto sea suficiente para sostener que las muestras podrían tener un comportamiento similar, por lo que es necesario realizar un análisis más profundo. Se procedió, entonces, a analizar las características de personas provenientes de distintas comunas y, por otro lado, las características de mujeres y hombres de ambas muestras.

Al analizar las características individuales por comuna de origen, se detectaron diferencias principalmente en el ingreso familiar mensual (IFAM), aunque el IFAM promedio en ambas muestras es similar (\$ 1.008.362 para el BD y \$ 1.000.056 para la

MV). Aun así existen diferencias dependiendo de la comuna de origen. En la Figura 4-6 se presentan los promedios del IFAM para BD y MV por comuna de origen. Se puede ver que las mayores diferencias se encuentran en los individuos que provienen de fuera de Santiago y en los que vienen de Providencia.

En el caso de personas provenientes de fuera de Santiago, se puede ver que en la MV el ingreso es muy superior (\$ 1.478.000) que en el BD (\$ 991.220). De hecho, si se comparan con los promedios de ambas muestras, se ve que en el caso de la MV los individuos provenientes de fuera de Santiago tienen mayor ingreso que el promedio, en cambio en el BD tienen menor ingreso que el promedio. Además, en el BD, los de fuera de Santiago representan un 5,4% de la muestra, y en la MV un 8,0%.

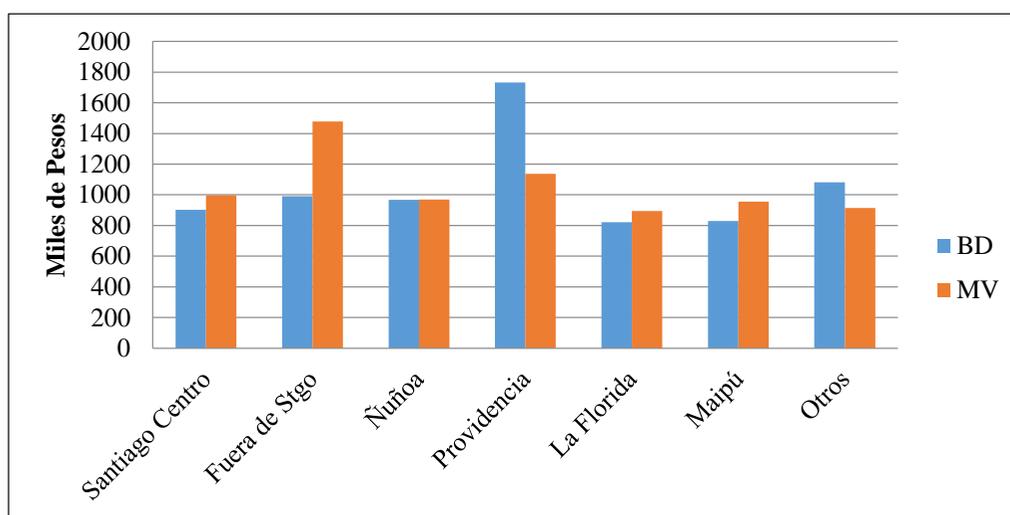


Figura 4-6: Promedio IFAM por Comuna de Origen en MV y BD

En el caso de personas provenientes de la comuna de Providencia ocurre lo contrario, es decir, en el BD el ingreso es muy superior (\$ 1.733.000) al de la MV (\$ 1.136.667), pero en ambos casos estos ingresos superan a sus promedios. Los individuos que vienen de Providencia representan un 6,1% de la muestra del BD y un 6,0% de la MV. Finalmente, una diferencia en los IFAM menos notoria, pero más relevante en cuanto al porcentaje de la muestra, ocurre en el caso de personas que vienen de “otras comunas”,

ya que el BD tiene un IFAM promedio de \$ 1.080.916, que es superior al de la MV (\$913.493). En el caso del BD la gente proveniente de otras comunas (32%) tiene mayor ingreso que el promedio de la muestra, y en la MV (42,5%) ocurre lo contrario.

En segundo lugar, otra diferencia importante que se encontró dependiente de la comuna de origen fue la distribución de la variable Sexo. En la Figura 4-7 se presenta el porcentaje de mujeres en ambas muestras por comuna de origen. Las principales diferencias se ven en las comunas de La Florida, Providencia y Ñuñoa.

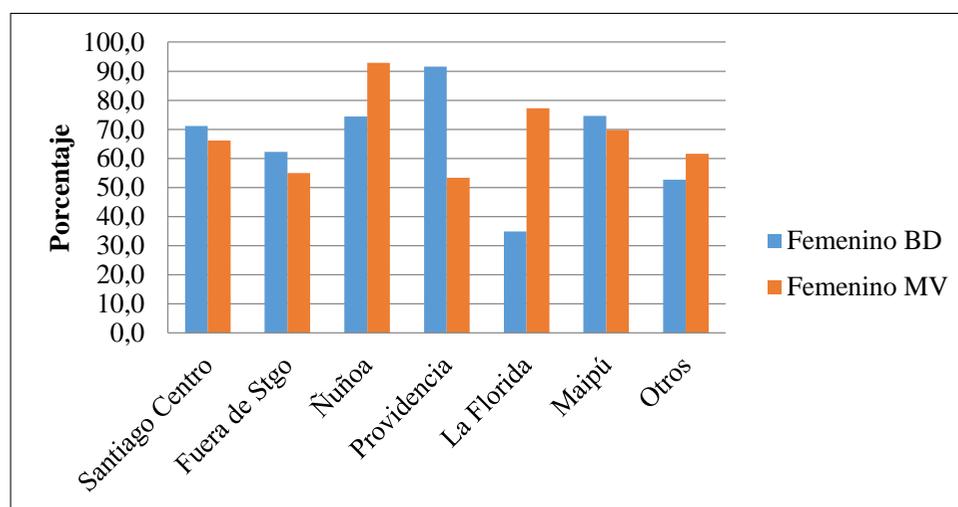


Figura 4-7: Distribución Variable Sexo Femenino por Comuna de Origen en MV y BD

En el primer caso, en la MV una gran mayoría (77,3%) son mujeres pero en el BD representan la minoría (34,9%). En el caso de encuestados provenientes de Providencia, se ve que en el BD la casi la totalidad son mujeres (91,5%) mientras que en la MV representan un poco más de la mitad (53,3%). Finalmente, casi la totalidad de los encuestados que provienen de la comuna de Ñuñoa son mujeres (92,9%) en la MV, y este porcentaje disminuye, aun cuando continua siendo mayoritario (74,5%) en el BD.

Así, se puede afirmar que las características socioeconómicas varían significativamente de acuerdo a la comuna de origen. Por lo mismo, se puede anticipar que al agregar

VSG, las variables que dependan de la comuna de origen podrían oscilar de manera significativa entre ambas muestras, lo que se debiera ver reflejado en una peor adaptación de los modelos estimados a la muestra de validación.

La variable Sexo también muestra diferencias importantes en relación al número de autos (Figura 4-8) y en relación al cociente entre el número de autos y el número de habitantes (Figura 4-9). Es importante mencionar que en la MV y BD las mujeres representan aproximadamente un 64% del total.

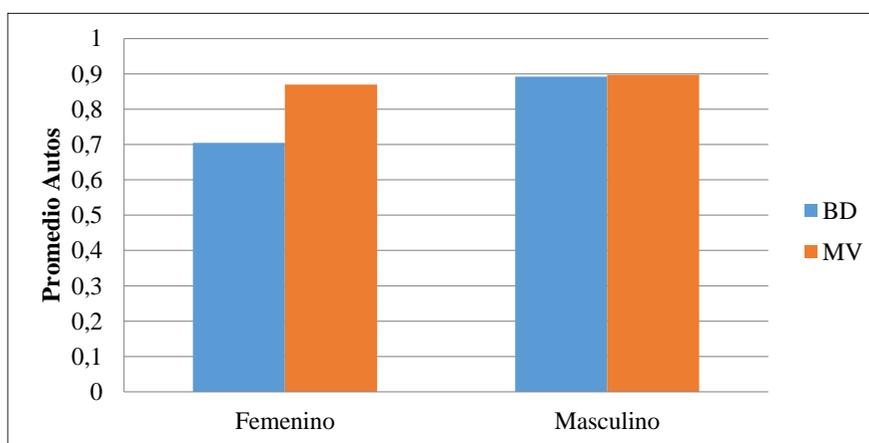


Figura 4-8: Promedio de Autos en el Hogar en Relación al Sexo en BD y MV

La Figura 4-8 permite ver que los hombres no tienen grandes diferencias en el número promedio de autos en el BD y en la MV, pero en el BD las mujeres tienen significativamente menos autos en promedio que en la MV. Resulta interesante analizar estos datos en relación al cociente entre número de autos en el hogar y número de habitantes en el hogar, lo que se muestra en la Figura 4-9.

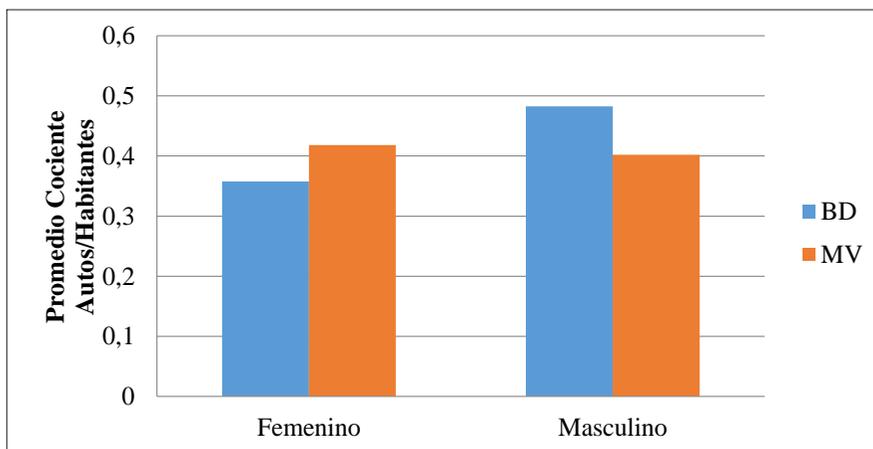


Figura 4-9: Cociente N°Autos/N°Habitantes en Relación al Sexo en BD y MV

Se puede ver que en el BD las mujeres tienen menos autos por habitante que los hombres, mientras que en la MV esto es al revés. Este factor resulta interesante, ya que representa la disponibilidad de autos por habitante en el hogar y es contrario para hombres y para mujeres en ambas muestras. Por lo mismo, a partir de esto se puede intuir que la variable Sexo también puede tener un comportamiento distinto en el BD y en la MV al interactuar con los distintos atributos, sobre todo con los que tienen que ver con cercanía o lejanía, ya que la disponibilidad de autos podría implicar una valoración distinta de las distancias.

## 5 ESTIMACIÓN DE MODELOS SIN HETEROGENEIDAD

### 5.1 Modelos $B_{SINHET}$ , $W_{SINHET}$ y $B+W_{SINHET}$

El primer paso será investigar si en realidad las respuestas *best* (mejor atributo) y *worst* (peor atributo) son efectivamente inversas, y si es válido considerar los ocho atributos de la encuesta como comunes. Para esto, se estimó las sensibilidades (pendientes) de cada atributo con: (a) las respuestas sobre cuál era el mejor atributo (*best*); (b) las respuestas sobre cuál era el peor (*worst*) y (c) las respuestas sobre el mejor (*best*) y peor (*worst*) atributo simultáneamente, considerando un factor de escala para las respuestas *worst*. En la Tabla A-1 del Anexo D se presentan los resultados obtenidos.

El factor de escala estimado al unir las respuestas *best* y *worst*, es significativo con un 99% de confiabilidad (por lo que es necesaria su inclusión); es decir, las respuestas *worst* no son exactamente inversas a las respuestas *best*. Además, se puede ver que la varianza correspondiente al error de las respuestas *worst* es menor que el de las *best*.

Para comprobar si se pueden considerar los ocho atributos comunes a los dos tipos de respuesta, se debe aplicar el test de razón de verosimilitud explicado en la sección 3.5. En este caso resulta que  $\lambda_a=41$  y como  $\chi^2_{8;0,05}=15,51$  se rechaza la hipótesis que los ocho atributos puedan considerarse comunes.

Para analizar qué atributos deben considerarse específicos a cada tipo de respuesta, se recomienda graficar las sensibilidades de las respuestas *best* y *worst* y verificar cuáles atributos parecen alejarse de la tendencia lineal (Swait y Bernardino 2000). Observando la Figura 5-1 se concluye que los atributos que parecen alejarse más de la recta son: *Mantenimiento y estado de conservación de la calle* (marcado en negro), *Servicios culturales* (marcado en celeste) y *Dividendo departamento* (marcado en morado). Además se estudiará si el atributo *Gimnasio y piscina en el edificio* (marcado en verde)

debiera también ser considerado como específico a cada respuesta, dado que su parámetro estimado tiene signo opuesto en ambos casos.

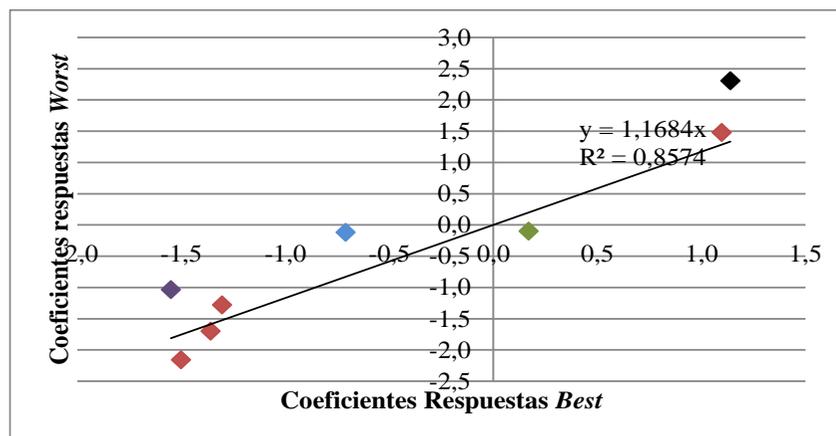


Figura 5-1: Sensibilidades (pendientes) de Modelos Sin Heterogeneidad

Probando las distintas combinaciones, se llega a que los atributos *Mantenimiento y estado de conservación de la calle* y *Servicios culturales* debieran ser considerados específicos. En ese caso, el modelo *Best+Worst* tiene una log verosimilitud de -4.313. Con esto, se obtiene un  $\lambda_a=12$  y como  $\chi^2_{6;0,05}=12,59$ , no se rechaza la hipótesis que los restantes seis atributos puedan ser considerados comunes a ambas encuestas. En la Tabla A-2 del Anexo D se muestran las sensibilidades obtenidas para el modelo final  $B+W_{\text{SINHET}}$ .

## 5.2 Modelos Mixtos $ED_{\text{SINHET}}$ , $B_{\text{SINHET}}$ y $W_{\text{SINHET}}$

A continuación consideraremos la unión de las respuestas *best* o *worst* con las de ED agregando un factor de escala. Se estimó modelos utilizando (a) sólo las respuestas de elección binaria; (b) las respuestas de elección binaria junto con las respuestas *best* y (c) las respuestas de elección binaria junto con las respuestas *worst*. En los dos últimos casos se tomó las respuestas ED como la base (factor de escala igual a uno) y se incluyó un factor de escala que resultó ser significativo al 95% de confianza en ambos modelos. En la Tabla A-3 del Anexo D se presentan las sensibilidades obtenidas.

Con las log verosimilitudes obtenidas se aplicó nuevamente el test LR para verificar si los ocho atributos se podían considerar comunes y los valores obtenidos se muestran en la Tabla 5-1. Es interesante mencionar que aun cuando el test formal es rechazado, la diferencia entre  $\chi^2_8$  y  $\lambda_a$  es pequeña; esto es indicativo de un grado de similitud sorprendente entre las pendientes comparables de los dos tipos de encuesta, como se ha visto en otros artículos (Swait y Louviere; 1993). De todas formas, se evaluará qué atributos podrían considerarse comunes entre las encuestas en ambos escenarios.

Tabla 5-1: Test de Razón de Verosimilitud para Modelos de ED junto con *Best* y *Worst*

	ED, <i>Best</i> , ED+B	ED, <i>Worst</i> , ED+W
$L_1$	-1.003	-1.003
$L_2$	-2.151	-2.156
$L_\mu$	-3.169	-3.179
$\lambda_a$	30	40
$\chi^2_8$	15,51	15,51
Test	Rechaza	Rechaza

Se analizará además cómo se podrían combinar las tres respuestas, examinando qué atributos se podrían considerar comunes en la encuesta ED con el modelo  $B+W_{\text{SINHET}}$  estimado anteriormente (que consideraba seis atributos comunes a las dos respuestas).

### 5.3 Modelo ED+ $Best_{\text{SINHET}}$

Para verificar cuáles parámetros se deben considerar específicos a cada respuesta, se pueden graficar los estimadores para cada tipo de respuesta por separado (Figura 5-2). Se puede ver que el atributo *Dividendo departamento* (marcado en verde) es el que parece alejarse más de la línea de tendencia. Aun así, el considerarlo específico no aumentó significativamente la log verosimilitud. Probando distintas combinaciones, se concluyó que los atributos *Áreas verdes* y *Servicios comerciales básicos* (ambos puntos marcados azul en la Figura 5-2) se debían considerar específicos a cada respuesta. En

este caso, el modelo ED+*Best*<sub>SINHET</sub> obtiene una log verosimilitud ( $L_\mu$ ) de -3.161 y  $\lambda_a=12$ , para  $\chi^2_{6;0,05}=12,59$ , y no se rechaza la hipótesis nula que acepta considerar estos seis parámetros comunes entre las respuestas.

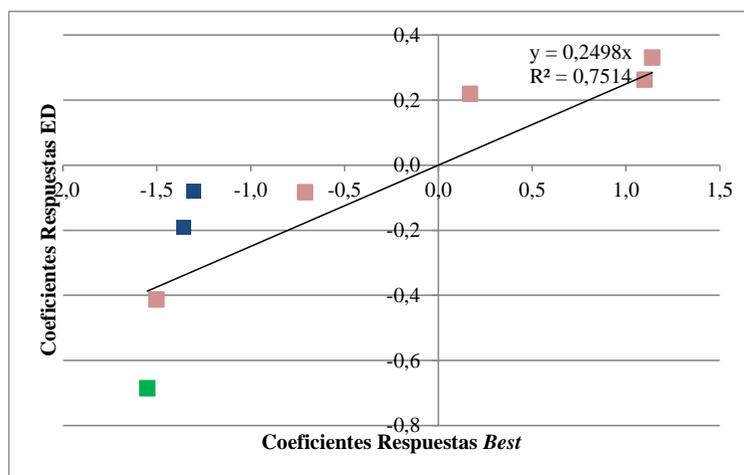


Figura 5-2: Coeficientes (sensibilidades) de Modelos Sin Heterogeneidad para Respuestas ED versus Respuestas *Best*

#### 5.4 Modelo ED+*Worst*<sub>SINHET</sub>

En la Figura 5-3 se grafican los coeficientes de los modelos estimados con respuestas ED versus respuestas *worst*. La figura permite intuir qué atributos parecerían no ser comunes entre las respuestas de ED y las respuestas *worst* y así probar distintas alternativas. Luego de varias pruebas, se concluyó que los atributos *Gimnasio* y *piscina en el edificio* y *Dividendo departamento* debían considerarse específicos (ambos marcados en burdeo en la Figura 5-3). Con esto, el modelo final ED+*Worst*<sub>SINHET</sub> tiene una log verosimilitud ( $L_\mu$ ) de -3.164 y aplicando el test LR se obtiene que  $\lambda_a=9$ ; como  $\chi^2_{6;0,05}=12,59$ , no se rechaza la hipótesis que los restantes seis atributos sean comunes entre las respuestas de ED y *worst* con un 95% de confiabilidad.

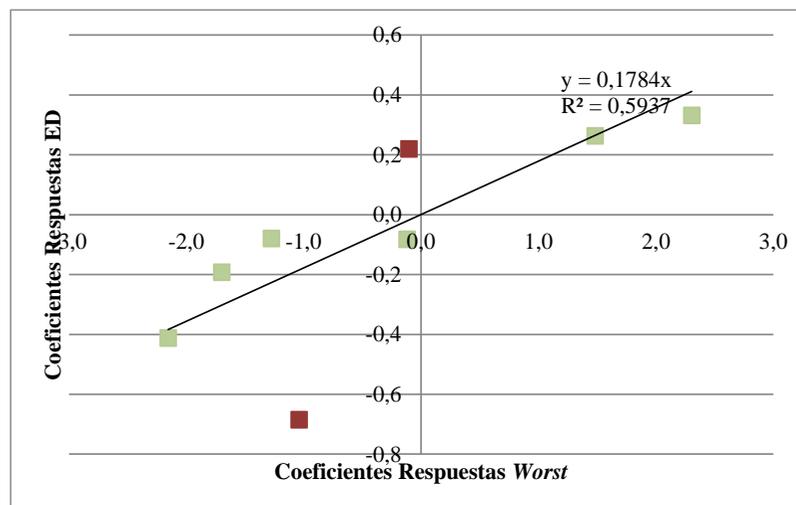


Figura 5-3: Coeficientes (sensibilidades) de Modelos Sin Heterogeneidad para Respuestas ED versus Respuestas Worst

### 5.5 Modelos ED+Best+Worst<sub>SINHET</sub>

Cabe recordar que el modelo *Best+Worst<sub>SINHET</sub>* consideró el atributo *Mantenición* y *estado de conservación de la calle* y el atributo *Servicios culturales* específicos para cada respuesta. Por lo mismo, para buscar el modelo ED+*Best+Worst<sub>SINHET</sub>* se debe analizar si los parámetros estimados para los atributos *Mantenición* y *estado de conservación de la calle* y *Servicios culturales* para las respuestas ED pueden considerarse comunes con los estimados para las respuestas *best* o para las *worst*.

Probando las distintas combinaciones, el valor de la log verosimilitud más cercana a cero se obtiene considerando el atributo *Mantenición* y *estado de conservación de la calle* igual al estimado para las respuestas *worst*, y el atributo *Servicios culturales* igual al estimado para las respuestas *best*. Con esto se obtiene una log verosimilitud de -5.336, un  $\lambda_a=40$ , y como  $\chi^2_{8;0,05}=15,51$  se rechaza la hipótesis nula que estos ocho atributos deban considerarse comunes entre las respuestas.

De esta forma, se debe analizar cuál(es) atributo(s) en realidad debe(n) ser estimado(s) específico(s) para las respuestas ED. En la Figura 5-4 se grafican los coeficientes del modelo ED versus los coeficientes del modelo *Best+Worst*<sub>SINHET</sub>. El atributo que parece ser más distinto es *Dividendo departamento* (marcado en amarillo). Además, el atributo *Gimnasio y piscina en el edificio* (marcado en verde) presenta signos opuestos, por lo que también se verificará si efectivamente debería ser considerado específico para las respuestas ED.

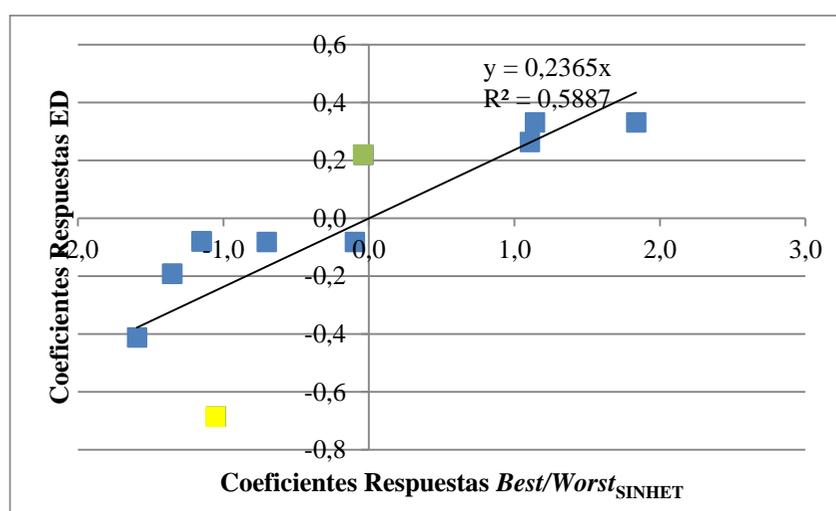


Figura 5-4: Coeficientes (sensibilidades) de Modelos Sin Heterogeneidad para Respuestas ED versus Parámetros Modelo *Best+Worst*<sub>SINHET</sub>

Analizando posibles combinaciones de cuáles atributos considerar comunes en las respuestas, se llegó nuevamente a la conclusión que se debían considerar atributos específicos a cada tipo de respuesta: (i) para las respuestas ED éstos son *Dividendo departamento* y *Gimnasio y piscina en el edificio*; (ii) para las respuestas *best* se debe considerar específico el atributo *Mantenimiento y estado de conservación de la calle*, y (iii) para las respuestas *worst* se debe considerar específico el atributo *Servicios culturales*. Con esto, el modelo *ED+Best+Worst*<sub>SINHET</sub> tiene una log verosimilitud de -5.323,

un  $\lambda_a=12,54$  y como  $\chi^2_{8;0,05}=12,56$  se acepta la hipótesis nula que seis atributos se pueden considerar comunes al 95% de confianza.

## 5.6 Comportamiento Predictivo de Modelos en la Muestra de Validación

Para ver cuál modelo se ajusta mejor a la realidad, se analizarán sus capacidades predictivas en la muestra de validación (MV), y también se compararán con la predicción del modelo ED por sí solo con el objeto de verificar si en realidad las respuestas *best* y/o *worst* aportan información a la ED. Se debe recordar, que como se vio en la sección 3.7.2, se deberá incluir un factor de escala en las utilidades de los modelos estimados ya que las muestras pertenecen a distintos individuos.

### 5.6.1 ED<sub>SINHET</sub>

En la Tabla 5-2 se muestran los indicadores más importantes asociado a predecir el comportamiento de los usuarios en la muestra de validación usando los coeficientes del modelo ED<sub>SINHET</sub>. Los tests FPR, ER, CR,  $\chi^2_{FPR}$  y  $\sigma_{norm}$  se definen en el Anexo A.

Tabla 5-2: Predicción Modelo ED<sub>SINHET</sub> en Muestra de Validación

	Factor de Escala (test t)	$\ell(\theta)$	Error Residual	FPR	ER	CR	$\chi^2_{FPR}$ (GL)	$\sigma_{norm}$
ED <sub>SINHET</sub>	-	-3.815	0,473	0,659	0,584	0,5	264,64 (1)	0,304
ED <sub>SINHET</sub> FE	1,66 (9,1)	-3.762	0,469	0,659	0,631	0,5	144,17 (4)	0,304

A partir de estos resultados, lo primero que se puede afirmar es que la inclusión de un factor de escala es efectivamente necesaria, ya que es significativamente distinto de uno con una confiabilidad del 99%; además, el modelo con factor de escala parece ser más adecuado en todos los indicadores (lo que es lógico). Predice correctamente un 66% de la MV y de acuerdo a los valores de FPR, ER y CR el modelo sería razonable e informativo (Ortúzar y Willumsen, 2011). No obstante, de acuerdo al test  $\chi^2_{FPR}$  el

modelo no parece ser consistente con los datos ( $\chi^2_{4;0,05} = 9,49$ ) a un nivel de confianza del 95%.

### 5.6.2 ED+Best<sub>SINHET</sub>

Como vimos en la sección 5.3, el modelo ED+Best<sub>SINHET</sub> consideraba dos atributos específicos a cada muestra: *Cercanía a servicios comerciales básicos* y *Cercanía a áreas verdes*. Por lo mismo, primero se analizará cuál parámetro (si perteneciente a las respuestas ED o a las Best) obtiene una mejor predicción en la muestra de validación. Probando las distintas combinaciones, se encontró que la mejor log-verosimilitud se obtenía considerando el coeficiente de *Cercanía a servicios comerciales básicos* estimado por las respuestas ED y el de *Cercanía a áreas verdes* de las respuestas best.

Tabla 5-3: Predicción Modelo ED+Best<sub>SINHET</sub> en Muestra de Validación

	Factor de Escala (test t)	$\ell(\theta)$	Error Residual	FPR	ER	CR	$\chi^2_{FPR}$ (GL)	$\sigma_{norm}$
ED+Best <sub>SINHET</sub>	-	-3.738	0,467	0,695	0,597	0,5	96,21 (1)	0,368
ED+Best <sub>SINHET</sub> FE	1,56 (9,3)	-3.690	0,463	0,695	0,641	0,5	10,65 (4)	0,368

Los principales indicadores de predicción para el modelo ED+Best<sub>SINHET</sub> se muestran en la Tabla 5-3. Nuevamente, el modelo con factor de escala es superior, predice correctamente un 70% de la MV y parece ser razonable e informativo. Además, de acuerdo al test  $\chi^2_{FPR}$  el modelo con factor de escala parece ser consistente con los datos de la MV con un nivel de confianza de 97.5% ( $\chi^2_{4;0,05} = 11,14$ ).

### 5.6.3 ED+Worst<sub>SINHET</sub>

Como se vio en la sección 5.4, el modelo ED+Worst<sub>SINHET</sub> consideraba dos atributos específicos a cada muestra: *Gimnasio y piscina en el edificio* y *Dividendo departamento*. Además, en el primer caso se obtuvo coeficientes de signo contrario para

los dos tipos de respuesta. Normalmente se considera válido utilizar aquel coeficiente que tenga un signo teóricamente correcto, pero en este caso no se puede saber *a priori* qué signo debería tener este atributo. Analizando las cuatro permutaciones posibles, se encontró que la mejor log-verosimilitud se obtenía considerando el coeficiente de *Gimnasio y piscina en el edificio* estimado mediante las respuestas *worst* y el de *Dividendo departamento* estimado mediante las respuestas ED.

Los principales indicadores de las predicciones realizadas este modelo se muestran en la Tabla 5-4. El modelo con factor de escala (nuevamente muy significativo), predice correctamente un 73% de la MV y parece ser razonable e informativo. No obstante, de acuerdo al test  $\chi^2_{FPR}$  no parece ser consistente con los datos de la MV ( $\chi^2_{4;0,05} = 9,49$ ) para un nivel de confianza del 95%.

Tabla 5-4: Predicción Modelo ED+Worst<sub>SINHET</sub> en Muestra de Validación

	Factor de Escala (test t)	$l(\theta)$	Error Residual	FPR	ER	CR	$\chi^2_{FPR}$ (GL)	$\sigma_{norm}$
ED+Worst <sub>SINHET</sub>	-	-3.657	0,460	0,73	0,595	0,5	9,12 (1)	0,466
ED+Worst <sub>SINHET</sub> FE	2,34 (18,0)	-3.461	0,444	0,73	0,697	0,5	27,87 (4)	0,466

#### 5.6.4 ED+Best+Worst<sub>SINHET</sub>

En la sección 5.5 se concluyó que los atributos que debían ser considerados específicos para las respuestas ED en este caso eran *Dividendo departamento* y *Gimnasio y piscina en el edificio*; además, para las respuestas *best* se debía considerar específico el atributo *Mantenimiento y estado de conservación de la calle* y para las respuestas *worst* el atributo *Servicios culturales*. Al analizar qué coeficientes (para qué tipo de respuesta) debían considerarse para realizar la mejor predicción, se concluyó que la mejor log-verosimilitud se obtenía con el coeficiente de *Mantenimiento y Estado de conservación de*

la calle obtenido de las respuestas *worst*/ED, el de *Servicios Culturales* obtenido de las respuestas *worst*, el de *Dividendo departamento* de las respuestas ED y de *Gimnasio y piscina en el edificio* proveniente de las respuestas *best/worst*.

En la Tabla 5-5 se muestran los indicadores de predicción. El modelo con factor de escala predice correctamente un 73% de la MV y parece ser razonable e informativo. Sin embargo, de acuerdo al test  $\chi^2_{FPR}$  no parece ser consistente con los datos de la MV ( $\chi^2_{4;0,05} = 9,49$ ) para un nivel de confianza del 95%.

Tabla 5-5: Predicción Modelo ED+*Best*+*Worst*<sub>SINHET</sub> en Muestra de Validación

	Factor de Escala (test t)	$l(\theta)$	Error Residual	FPR	ER	CR	$\chi^2_{FPR}$ (GL)	$\sigma_{norm}$
ED+B+W <sub>SINHET</sub>	-	-3.665	0,460	0,73	0,594	0,5	9,12 (1)	0,466
ED+B+W <sub>SINHET</sub> FE	2,44 (18,9)	-3.451	0,443	0,73	0,704	0,5	22,71 (4)	0,466

### 5.6.5 Comparación del Comportamiento Predictivo de los Modelos

En la Tabla 5-6 se muestran los indicadores de predicción en la MV de los mejores modelos obtenidos al unificar las respuestas ED y *Best* y/o *Worst*.

Tabla 5-6: Indicadores de Predicción Mejores Modelos ED con *Best* y/o *Worst*

	Factor de Escala (test t)	$l(\theta)$	Error Residual	FPR	ER	CR	$\chi^2_{FPR}$ (GL)	$\sigma_{norm}$
ED <sub>SINHET</sub> FE	1,66 (9,1)	-3.762	0,469	0,659	0,631	0,5	144,17 (4)	0,304
ED+ <i>Best</i> <sub>SINHET</sub> FE	1,56 (9,3)	-3.690	0,463	0,695	0,641	0,5	10,65 (4)	0,368
ED+ <i>Worst</i> <sub>SINHET</sub> FE	2,34 (18,0)	-3.461	0,444	0,733	0,697	0,5	27,87 (4)	0,466
ED+B+W <sub>SINHET</sub> FE	2,44 (18,9)	-3.451	0,443	0,733	0,704	0,5	22,71 (4)	0,466

Observando la Tabla 5-6 es evidente que todos los indicadores para los tres primeros modelos son superiores al del modelo ED, por lo que efectivamente las respuestas *best* y/o *worst* mejoran la capacidad predictiva del modelo conjunto. En relación a la log verosimilitud, errores residuales, FPR, ER, CR e índice de éxito normalizado, el modelo que presenta mejores indicadores es el  $ED+B+W_{SINHET}$  seguido por el modelo  $ED+Worst_{SINHET}$ . Además, el modelo  $ED+Best_{SINHET}$  es el único que aprueba el test  $\chi^2_{FPR}$  al 97.5% nivel de confianza (i.e. es consistente con los datos). No obstante, este test no se considera tan relevante a la hora de tomar una decisión, ya que presenta complicaciones al ser aplicado, y por ende no es tan preciso a la hora de discriminar entre modelos (Ortúzar, 2000). Por todo esto, se concluye que el mejor modelo es el  $ED+Best+Worst_{SINHET}$ .

## 6 ESTIMACIÓN DE MODELOS CON HETEROGENEIDAD

### 6.1 Heterogeneidad Mediante Componentes de Error

#### 6.1.1 Resultados Modelos *Best*<sub>EP</sub>, *Worst*<sub>EP</sub> y *ED*<sub>EP</sub>

Se procedió a considerar el efecto panel, inherente a nuestros datos con observaciones repetidas por individuo, mediante la adición de componentes de error (EC) en las funciones de utilidad. En primer lugar, se verificó si la desviación estándar de la EC (tiene media igual a cero) para cada tipo de respuesta era significativa y, por ende, si la EC aportaba información al modelo. Los resultados de esta modelación se presentan en la Tabla A-4 del Anexo D. En los tres casos la adición de EC mejora significativamente la log verosimilitud de los modelos, y los parámetros *Sigma* (desviación estándar) estimados son significativamente distintos de cero al 95% de confianza.

Además, si se aplica el test LR (ecuación (3.14)) para verificar si efectivamente es necesario estimar estos coeficientes en el modelo ED, queda:

$$LR = -2 \cdot (-1.003 + 878) = 250 > \chi^2_{1;0,001} = 10,83$$

constatando que se rechaza la hipótesis nula (los modelos son equivalentes) al 99% de confianza; la gran diferencia implica que estimando tan sólo un coeficiente más el modelo mejora significativamente. Por lo mismo se deduce que es necesario incluir heterogeneidad EP (efecto panel mediante componentes de error) sino se podrían obtener conclusiones erróneas.

#### 6.1.2 Resultados Modelo *Best+Worst*<sub>EP</sub>

Como antes, en primer lugar se buscó analizar si las respuestas *best* y *worst* podían ser estimadas en conjunto considerando algún atributo como común entre ellas. Para comenzar, se estimó un modelo con las respuestas *best* y *worst* considerando todos los

atributos comunes. Este modelo tiene una log verosimilitud de -4.036, con lo que nos queda un  $\lambda_a=71$ , y como  $\chi^2_{8;0,05}=15,51$  claramente se rechaza la hipótesis de que los ocho atributos deben ser considerados comunes. La Figura 6-1 grafica los coeficientes de ambas respuestas, con el fin de ver cuáles atributos se comportan de manera distinta y por ende deberían ser estimados como atributos específicos. Se puede ver que *Mantenición y estado de conservación de la calle* (marcado en verde) es el que parece estar más lejos de la recta, por lo que sería el primero a estimar como específico a cada respuesta. Por otro lado, *Gimnasio y piscina en el edificio* (marcado en rojo) tiene signo contrario en ambos tipos de encuesta, por lo que también se verificará si debería considerarse específico a cada respuesta.

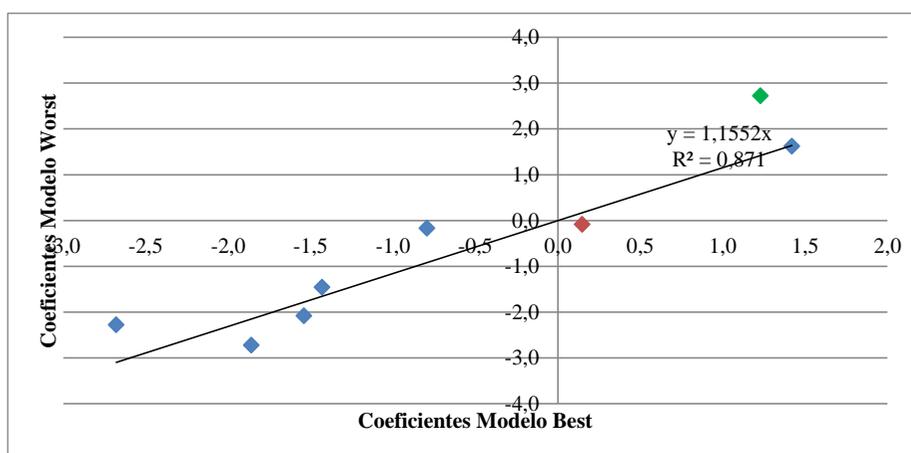


Figura 6-1: Coeficientes (sensibilidades) de Modelos con EP para Respuestas *Best* versus Respuestas *Worst*

Se probó distintas combinaciones, rechazando la hipótesis de atributos comunes en todos los casos. Por esto se concluyó que las respuestas *best* y *worst* se comportaban de manera muy distinta y no debían ser agregadas.

### 6.1.3 Resultados Modelo ED+Best<sub>EP</sub>

El primer paso para decidir qué atributos se pueden considerar comunes entre respuestas ED y *best*, es estimar un modelo considerando los ocho atributos comunes. Este modelo obtiene una log verosimilitud de -2.933, con lo que  $\lambda_a=41$ , y como  $\chi^2_{8;0,05}=15,51$  se rechaza la hipótesis de que los ocho atributos deben ser considerados comunes. En la Figura 6-2 se grafica los coeficientes del modelo *Best*<sub>EP</sub> versus los del modelo ED<sub>EP</sub>.

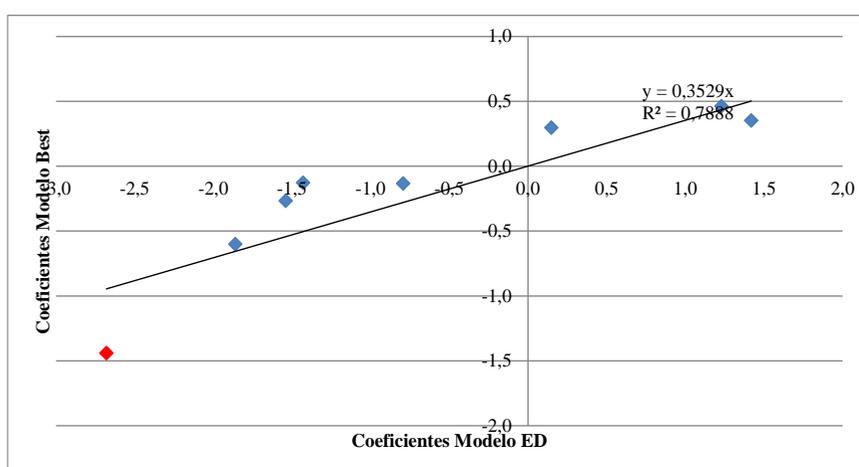


Figura 6-2: Coeficientes (sensibilidades) de Modelos con EP para Respuestas *Best* versus Respuestas ED

El atributo cuyos coeficientes se alejan más de la recta es *Dividendo del departamento* (marcado en rojo), por lo que se procedió a estimarlo como específico; sin embargo, se siguió rechazando el test de que los restantes siete atributos se podían considerar comunes. Debido a esto se probó distintas combinaciones hasta llegar al modelo *Best*+ED<sub>EP</sub> definitivo, que considera cuatro atributos comunes y cuatro específicos. Los que deben considerarse específicos son *Dividendo del departamento*, *Áreas verdes*, *Gimnasio y piscina en el edificio* y *Servicios comerciales básicos*. Este modelo tiene una log verosimilitud de -2.917, por lo que se obtiene un  $\lambda_a=11$ , y como  $\chi^2_{4;0,025}=11,14$ , se acepta que los restantes cuatro atributos se pueden considerar

comunes con un 97.5% de confiabilidad. Además, se puede aplicar el test LR para ver si efectivamente mejora el modelo al estimar los coeficientes que representan la desviación estándar en los errores pertenecientes a las respuestas *best* y ED:

$$LR = -2 \cdot (-3.161 + 2.917) = 487 > \chi^2_{2;0,05} = 5,99$$

Se puede ver que la diferencia es muy grande, por lo que mejora significativamente el modelo estimando dos coeficientes más.

#### 6.1.4 Resultados Modelo ED+Worst<sub>EP</sub>

Al considerar todos los atributos comunes en este caso, se obtuvo una log verosimilitud de -2.878 ( $\lambda_a=66$ ), y como  $\chi^2_{8;0,05}=15,51$  se rechaza la hipótesis de que los ocho atributos pueden ser considerados comunes. Como  $\lambda_a$  y  $\chi^2_{8;0,05}$  son significativamente distintos, el grado de similitud es bajo entre este tipo de respuestas. No obstante, se estudiará si esto se debe a algún atributo en particular y si sería válido considerar ciertos atributos comunes entre las respuestas.

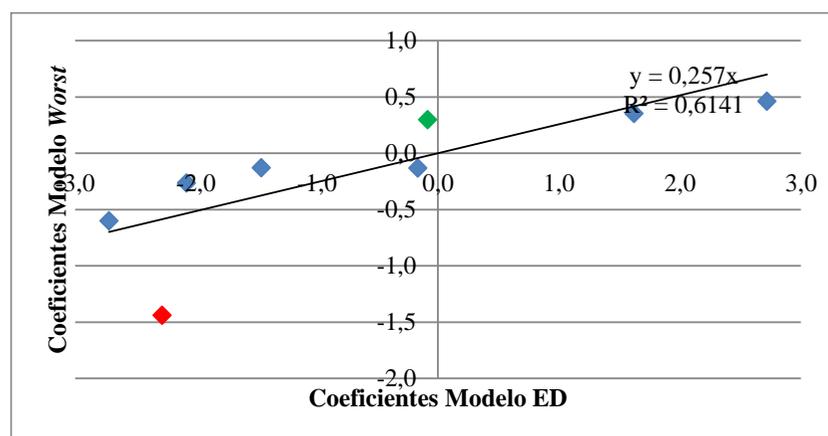


Figura 6-3: Coeficientes (sensibilidades) de Modelos con EP para Respuestas *Worst* versus Respuestas ED

En la Figura 6-3 se grafica los coeficientes del modelo  $Worst_{EP}$  versus los del modelo  $ED_{EP}$ . El atributo con coeficientes más disímiles es el de *Dividendo del departamento* (marcado en rojo), pero también se puede ver que los coeficientes de *Gimnasio y piscina en el edificio* (marcado en verde) tienen signo contrario en ambos casos. Así, se partió evaluando si ambos deberían considerarse específicos pese a que el último no es significativo en las respuestas *worst*. Sin embargo, tras probar distintas combinaciones siempre resultó  $\lambda_a$  significativamente mayor que  $\chi^2_{a;0,05}$ , concluyéndose que las respuestas *worst* se comportaban de manera muy distinta al agregar efecto panel mediante componentes de error y, por lo tanto, no podían ser agregadas.

### 6.1.5 Comportamiento Predictivo en la Muestra de Validación

El modelo  $Best+ED_{EP}$  considera cuatro atributos específicos a cada respuesta y cuatro comunes, por lo que se debe probar - para cada atributo considerado como específico - cuál de los coeficientes estimados tiene mejor capacidad predictiva en la muestra de validación. Se llegó a la conclusión que los coeficientes que predecían mejor el comportamiento en la muestra de validación eran los de *Áreas verdes y Gimnasio y piscina en el edificio* estimados mediante las respuestas *best*, y los de *Dividendo del departamento y Servicios comerciales básicos* estimados mediante las respuestas *ED*. En la Tabla 6-1 se muestra la predicción realizada por el modelo  $ED+Best_{EP}$  y por el modelo  $ED_{EP}$  (ambos con y sin factor de escala, FE). Se puede ver que el modelo  $ED+Best_{EP}$  predice mejor en la muestra de validación (que al sólo incluir las respuestas  $ED_{EP}$ ) en relación a todos los indicadores mostrados. En particular, predice adecuadamente un 72% de la muestra y parece ser razonable e informativo (aunque nuevamente, de acuerdo al test  $\chi^2_{FPR}$ , ninguno de los dos modelos sería consistente con los datos). Además, el modelo  $ED+Best_{EP}$  tiene errores residuales más pequeños y un índice de éxito normalizado mayor que el  $ED_{EP}$ .

Tabla 6-1: Capacidad Predictiva Modelo ED+BestEP y ED<sub>EP</sub> en Muestra de Validación

	Factor de Escala (test t)	$\alpha(0)$	Error Residual	FPR	ER	CR	$\chi^2_{FPR}$ (GL)	$\sigma_{norm}$
ED+BestEP	-	-3.621	0,458	0,716	0,623	0,5	36,49 (1)	0,428
ED+BestEP FE	1,44 (8,8)	-3.579	0,454	0,716	0,665	0,5	17,04 (4)	0,428
ED <sub>EP</sub>		-3.718	0,465	0,681	0,612	0,5	149,59 (1)	0,363
ED <sub>EP</sub> FE	1,45 (8,4)	-3.680	0,462	0,681	0,654	0,5	157,38 (4)	0,363

## 6.2 Variaciones Sistemáticas en los Gustos (VSG)

### 6.2.1 Análisis de Variables

En esta sección se analizará la posible inclusión de VSG al unir las respuestas ED con las best y/o worst, y en las ED por sí solas, para luego analizar si efectivamente la unión de respuestas aporta mayor información al modelo. Para esto, se procederá a incluir interacciones entre las sensibilidades de los atributos (que intentamos especificar como comunes para los distintos tipos de respuesta) y las siguientes variables socio-económicas de los individuos en el banco de datos:

- Edad: Esta variable se modelará como una variable categórica (*dummy*) con cinco niveles (entre 16-25 años, entre 26 y 35 años, entre 36 y 45 años y mayor a 45 años) y también se probará su comportamiento como variable continua.
- Sexo: Esta variable se modelará como una *dummy* con dos niveles (vale uno si es femenino y cero si es masculino).
- Distrito: Se modelará como una variable *dummy* con siete niveles (Santiago Centro, Fuera de Santiago, Ñuñoa, Providencia, La Florida, Maipú y Otras Comunas) que representan las comunas con mayor porcentaje en la muestra.

- Número de autos: Se modelará como una variable *dummy* con tres niveles (0, 1, y 2 autos en el hogar).
- Número de habitantes en el hogar: Se modelará como una variable *dummy* con tres niveles (1, 2 o 3 o más habitantes en el hogar).
- Tiempo de estadía en departamento: Se modelará como una variable *dummy* con tres niveles (menos de 5 años, de 5 a 10 años y más de 10 años).
- Ingreso familiar mensual (IFAM): Se modelará como una variable *dummy* con tres niveles (menos de \$ 600.000, de \$ 600.000 a \$ 1.000.000, y más de \$ 1.000.000). También se probará agregar esta variable como continua en caso que se adapte mejor a los atributos.

### 6.2.2 Modelo ED<sub>vsg</sub>

El objetivo principal de esta investigación es examinar el aporte que pueden hacer las respuestas *best* y *worst* a las respuestas ED. Por lo mismo, en cada caso resulta fundamental examinar qué hubiera pasado si sólo se hubiera contado con las respuestas ED, y cómo cambiarían los coeficientes estimados al combinarlas con las respuestas *best* y/o *worst*. La función utilidad del modelo ED<sub>vsg</sub> se presenta - desagregado por atributos - en la Tabla 6-2.

Este modelo, con 19 coeficientes, obtiene una log verosimilitud de -980. Con esto, se puede aplicar el test de razón de verosimilitud para ver si efectivamente es superior al modelo sin heterogeneidad (recordar que ese modelo tenía nueve coeficientes y obtuvo una log verosimilitud de -1.003):

$$LR = -2 \cdot (-1.003 + 980) = 45 > \chi^2_{9;0,001} = 29,59$$

Por lo tanto, se rechaza la hipótesis nula que los modelos son equivalentes al 99% de confianza, y se concluye que estimar estos 10 parámetros adicionales agrega información útil al modelo estimado.

Tabla 6-2: Función de Utilidad Desagregada por Atributos Modelo ED<sub>vsg</sub>

Atributo	Componente de la Función de Utilidad
Acceso al metro	-0,422 (7,7)
Áreas Verdes	-0,083 (1,5)
Dividendo departamento	$-1,832 + 0,799 \cdot Edad_{16-35} + 0,53 \cdot 10^{-3} \cdot IFAM$ (5,0) (2,7) (2,7)
Gimnasio y piscina en el edificio	$0,297 - 0,220 \cdot StgoCentro$ (4,4) (1,9)
Limpieza y estado de mantención de la calle	$0,498 + 0,338 \cdot FueraStgo + 0,382 \cdot Maipú$ (4,9) (1,4) (1,8) $- 0,448 \cdot Ing_{Bajo} - 0,223 \cdot Ing_{Medio}$ (3,2) (1,7)
Servicios comerciales básicos	$-0,147 - 0,386 \cdot Dorm_{1o2}$ (2,6) (2,2)
Servicios culturales	-0,084 (1,6)
Tamaño de los dormitorios	$0,642 - 0,450 \cdot Auto_{0o1} + 0,208 \cdot Estadia_{< 5 años}$ (3,7) (2,6) (1,8)

### 6.2.3 Modelo ED+Best<sub>vsg</sub>

Pese a que las respuestas ED, *best* y *worst* pudieron ser agregadas al no incluir efecto panel, se probará también la combinación de sólo las respuestas ED y *best*. Esto ya que esa fue la única combinación de respuestas estimable en el caso anterior, por lo que será el modelo base para analizar el impacto de incluir heterogeneidad en los gustos. Esto se verá en más detalle en sección 6.3, donde se encontró que las respuestas *best* y *worst* se comportaban de manera muy distinta y por lo tanto no pudieron ser estimadas en conjunto (no obstante, sí fue posible estimar las respuestas ED en conjunto con las *best*).

Considerando las respuestas ED y *best* en conjunto, se buscó cuáles variables parecían interactuar correctamente con los niveles de los atributos para explicar variaciones sistemáticas de gustos. Se obtuvo un modelo con 36 parámetros comunes entre las

respuestas ED y *best* y cuya log verosimilitud fue -3.081. Asimismo, para estas mismas VSG se estimó los coeficientes para las respuestas *best* (log verosimilitud de -2.067) por separado que para las ED (log verosimilitud de -976). Con estos valores se aplicó el test LR para verificar si efectivamente era válido considerar estos 36 parámetros comunes a ambas respuestas. Se obtuvo que  $\lambda_a=76$ , y como  $\chi^2_{36;0,05}=51,00$  se rechazó la hipótesis nula que los modelos son equivalentes (i.e. no es válido considerar los 36 atributos comunes a ambos tipos de respuesta).

Por esto, se graficó los coeficientes estimados con el fin de poder intuir cuáles deberían considerarse específicos a cada tipo de respuesta (Figura 6-4), y se realizó un análisis equivalente a los ya explicados en las secciones anteriores. Probando distintas combinaciones, se llegó a que los atributos *Áreas verdes*, *Dividendo departamento* y *Gimnasio y piscina en el edificio*, debían considerarse específicos para cada tipo de respuesta.

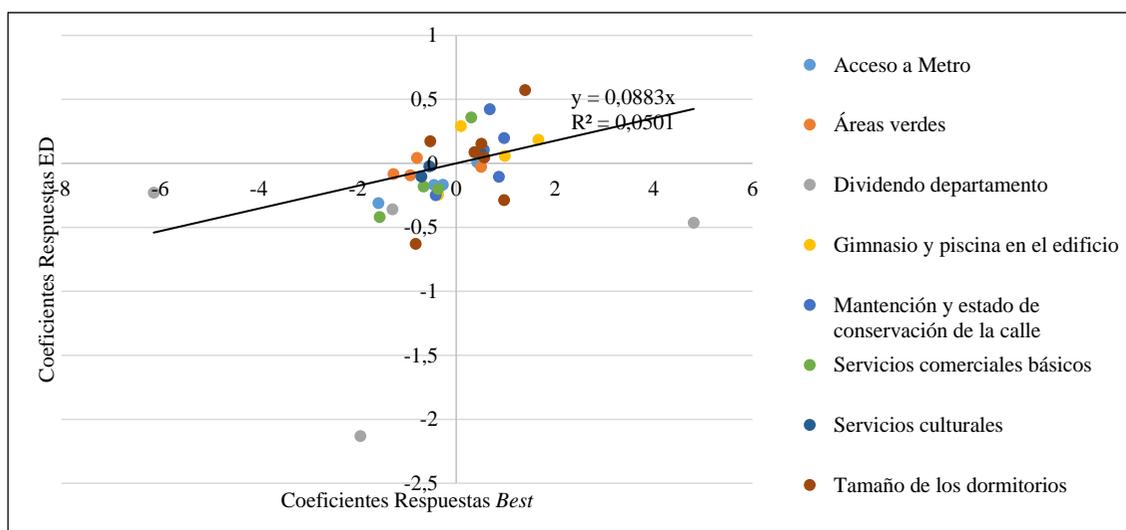


Figura 6-4: Coeficientes (sensibilidades) de Modelos con VSG para Respuestas *Best* versus Respuestas ED

Es importante mencionar que luego de especificar estos tres atributos como específicos se debe volver a probar las interacciones de las variables socioeconómicas con ellos

para cada tipo de respuesta. Tras este largo procedimiento, se obtuvo el mejor modelo  $ED+Best_{VSG}$ , que tiene una log verosimilitud de -3.057 y cuya función utilidad se muestra en la Tabla 6-3.

Posteriormente, para las mismas VSG se estimó un modelo con sólo las respuestas ED (log verosimilitud de -972) y con sólo las respuestas *best* (log verosimilitud de -2.068). Aplicando el test LR se obtuvo  $\lambda_a=35$ , y como  $\chi^2_{24;0,05}=36.42$  no se rechaza la hipótesis nula que los modelos son equivalentes, y se puede concluir que es válido considerar los 24 parámetros restantes como comunes a ambos tipos de respuesta en este caso. En este modelo se estimó 51 coeficientes en total y se puede aplicar el test LR para ver si efectivamente agregar variaciones sistemáticas de los gustos mejora significativamente el modelo (recordar que en el modelo  $ED+Best$  sin VSG se estimó 19 coeficientes y obtuvo una log verosimilitud de -3.161).

$$LR = -2 \cdot (-3.161 + 3.057) = 207 > \chi^2_{32;0,001} = 62,49$$

Claramente se rechaza la hipótesis nula que ambos modelos son equivalentes con un 99% de confiabilidad, y se concluye que agrega información significativa estimar estos 32 parámetros adicionales.

#### **6.2.4 Modelo $ED+Best+Worst_{VSG}$**

Al igual que en los casos anteriores, se comenzó por probar todas las interacciones posibles entre atributos y variables socio-económicas para la totalidad de la muestra ( $ED+Best+Worst$ ), sacando aquellos coeficientes no significativos de acuerdo al test t; se verificó, además, que los signos de las variables fueran los esperados. Con esto, se llegó al modelo  $ED+Best+Worst_{VSGinicial}$  - con 41 parámetros comunes entre los tres tipos de respuesta - y una log verosimilitud de -4.196.

Tabla 6-3: Función de Utilidad Desagregada Modelo ED+Bestvsg

Atributo	Componente de la Función de Utilidad
Acceso al metro	$-0,354 + 0,089 \cdot Sexo - 0,108 \cdot Hab_1 - 0,071 \cdot Hab_2$ (7,2) (2,7) (2,5) (1,9)
Áreas Verdes	<b>Best:</b> $-0,209 - 0,177 \cdot Edad_{16-35} + 0,111 \cdot Sexo - 0,278 \cdot \tilde{Nu\~{n}}oa$ (2,9) (2,6) (2,2) (3,4)
	<b>ED:</b> $-0,082$ (1,5)
Dividendo departamento	<b>Best:</b> $-1,362 - 0,274 \cdot \tilde{Nu\~{n}}oa - 0,425 \cdot LaFlorida + 1,064 \cdot Dorm_{1o2}$ (4,4) (2,3) (2,7) (3,8)
	<b>ED:</b> $-1,732 + 0,833 \cdot Edad_{16-35} - 2,079 \cdot LaFlorida + 4,8 \cdot 10^{-4} \cdot IFAM$ (4,8) (3,5) (2,8) (2,5)
Gimnasio y piscina en el edificio	<b>Best:</b> $0,201 + 0,403 \cdot FueraStgo - 0,212 \cdot Dorm_{2om\acute{a}s}$ (1,9) (3,5) (1,9)
	<b>ED:</b> $0,306 - 0,235 \cdot StgoCentro$ (4,5) (2,1)
Limpieza y estado de mantención de la calle	$0,246 + 0,197 \cdot \tilde{Nu\~{n}}oa - 0,146 \cdot Auto_0 + 0,136 \cdot Estad\acute{a}_{<5a\~{n}}os + 0,105 \cdot Dor_{1o2}$ (4,2) (2,1) (2,3) (2,1) (1,8)
Servicios comerciales básicos	$-0,332 + 0,098 \cdot StgoCentro - 0,151 \cdot Estad\acute{a}_{<5a\~{n}}os - 0,096 \cdot Estad\acute{a}_{5a10a\~{n}}os$ (4,3) (2,3) (2,8) (2,1)
	$+ 0,101 \cdot Dorm_{2om\acute{a}s}$ (1,6)
Servicios culturales	$-0,138 + 0,101 \cdot StgoCentro - 0,099 \cdot Auto_0$ (3,5) (1,9) (1,9)
Tamaño de los dormitorios	$0,334 + 0,117 \cdot Edad_{16-35} + 0,177 \cdot \tilde{Nu\~{n}}oa - 0,220 \cdot Auto_{0o1} + 0,124 \cdot Hab_1$ (5,1) (3,4) (2,9) (3,8) (2,7)
	$+ 0,086 \cdot Hab_2 - 0,101 \cdot Dorm_{1o2}$ (2,3) (2,9)

El siguiente paso fue verificar si todos los atributos se deberían considerar comunes entre las respuestas *best* y *worst*. Para esto, se estimó dos modelos con las mismas VSG que en el modelo ED+Best+Worstvsginicial, el primero con las respuestas *best* (log-verosimilitud de -2.085) y el segundo con las respuestas *worst* (log-verosimilitud

de -2.080). Aplicando el test LR, se obtuvo que  $\lambda_a=62$ , y como  $\chi^2_{41;0,05}=56,94$  se rechazó la hipótesis nula que los 41 coeficientes se podían considerar comunes a ambos tipos de respuesta. Se procedió, entonces, a graficar los coeficientes obtenidos en ambos casos y analizar cuáles parecían ser los que se comportaban de manera distinta (ver Figura 6-5). Es interesante destacar que a priori, no se detectó ningún atributo que se comportara de manera distinta para los tipos de respuesta. Por esto, se procedió a considerar cada atributo específico (en su totalidad, es decir todas las interacciones que lo componen).

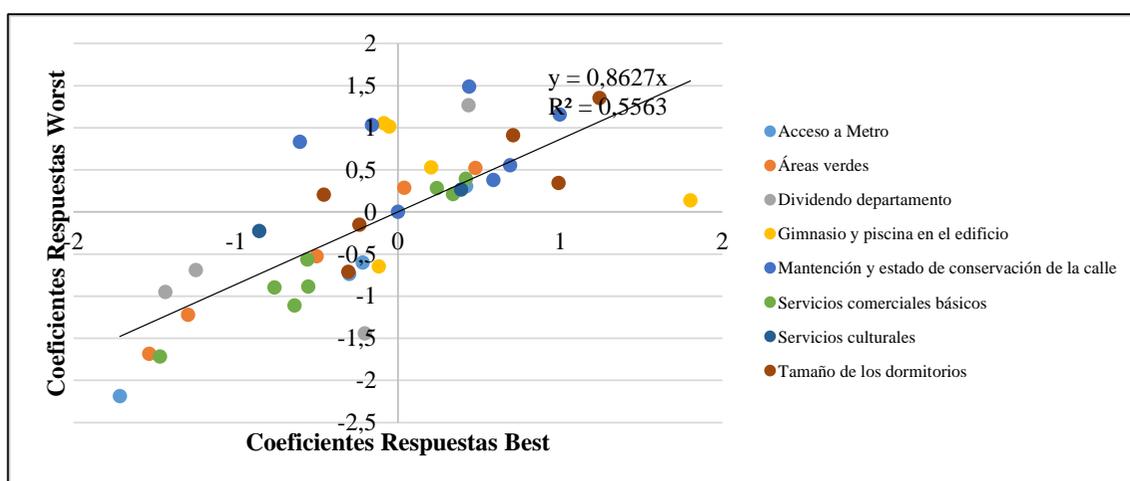


Figura 6-5: Coeficientes (sensibilidades) del modelo ED+Best+Worst vs Giniicial para Respuestas Best versus las Respuestas Worst

De esta forma se detectó que el atributo *Servicios culturales* se comportaba de manera muy distinta para los diferentes tipos de encuesta, ya que la log verosimilitud aumentaba significativamente al estimarlo como específico para cada tipo de respuesta (log verosimilitud de -4.190). Aun así, las interacciones consideradas para este atributo - al unir las respuestas - no parecían ser significativas para cada una por separado. De hecho, para las respuestas *worst* ninguna interacción resultó significativa (log verosimilitud de -2.080), y para las respuestas *best* se mantuvo la interacción entre las personas que provenían de la comuna Santiago Centro y el atributo *Servicios culturales*

(se mantiene la log verosimilitud anterior -2.085, ya que se estiman los mismos parámetros). Este modelo *Best+Worst*<sub>VSG</sub> obtiene un  $\lambda_a=51$ , y como  $\chi^2_{39;0,05}=54,57$ , no se rechaza la hipótesis nula que los modelos son equivalentes y por lo tanto se considera válido estimar los restantes 39 coeficientes como comunes en las dos respuestas.

El siguiente paso fue analizar si al unir las respuestas ED al modelo *Best+Worst*<sub>VSG</sub>, el coeficiente del atributo *Servicios culturales* estimado para las respuestas ED se adaptaba mejor que el estimado para las respuestas *best* o *worst*. Para esto se probaron las dos combinaciones posibles y se examinó cuál log verosimilitud era mayor (más cercana a cero). Esto permitió concluir que considerar el atributo *Servicios culturales* común con las respuestas *best* (log verosimilitud de -5.208) era lo más apropiado.

Luego, se procedió a aplicar el test LR para ver si en realidad era pertinente considerar los 41 atributos comunes entre las respuestas ED y el modelo *Best+Worst*<sub>VSG</sub>. El modelo ED tiene una log verosimilitud de -981 considerando las mismas VSG que en el modelo *ED+Best+Worst*<sub>VSGinicial</sub>, con lo que se obtiene un  $\lambda_a=73$ , y como  $\chi^2_{41;0,05}=56,94$  se rechazó la hipótesis que los 41 coeficientes debían considerarse comunes. Al igual que antes, se graficó los coeficientes para ver cuáles parecían comportarse de manera distinta, lo que se puede ver en la Figura 6-6.

Observando la figura se puede ver, en primera instancia, que dos puntos correspondientes al atributo *Dividendo departamento* se alejan bastante de la línea de tendencia, por lo que se procedió a considerarlo como específico para las respuestas ED. Con esto, se obtuvo una log verosimilitud de -5.199 y un  $\lambda_a=56$ , y como  $\chi^2_{37;0,05}=52,19$  se rechazó la hipótesis que los 37 coeficientes podían considerarse comunes (cabe mencionar que a un nivel de confianza de 97.5% también se rechazaría la hipótesis nula).

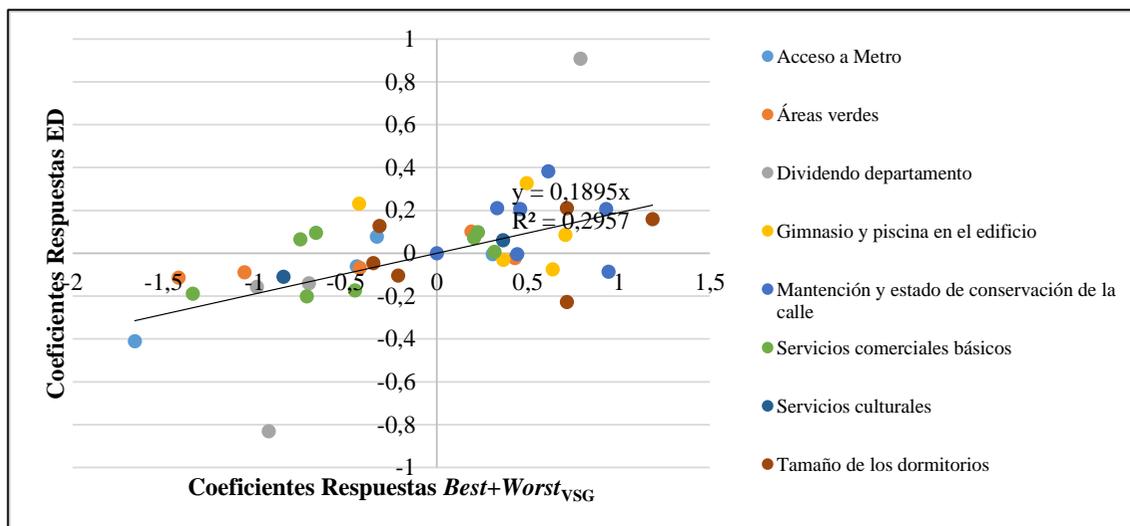


Figura 6-6: Coeficientes (sensibilidades) de Respuestas ED y  $Best+Worst_{VSG}$  considerando las VSG estimadas en el modelo  $ED+Best+Worst_{VSGinicial}$

Se procedió a analizar cuál de los restantes siete atributos se comportaban de manera distinta, verificando si uno por sí solo o en conjunto con el atributo *Dividendo departamento*, se deberían considerar específicos. Analizando posibles combinaciones, se concluyó que los coeficientes de los atributos *Dividendo departamento* y *Gimnasio y piscina en el edificio* se debían considerar específicos para las respuestas ED. Al hacer esto, se debió probar nuevas interacciones para estos atributos en estas respuestas, ya que no parecían ser significativas las mismas anteriores. Con lo anterior, se obtuvo el modelo  $ED+Best+Worst_{VSG}$  con una log verosimilitud de -5.188. Considerando las nuevas interacciones para los atributos específicos que resultaron ser significativas en las respuestas ED, también se estimó este modelo en forma separada, obteniéndose una log verosimilitud de -978 (el modelo  $Best+Worst$  se mantiene igual ya que las mismas interacciones resultaron significativas).

Con esto,  $\lambda_a=32$ , y como  $\chi^2_{32;0,05}=46,19$  se considera válido estimar los restantes 32 parámetros comunes para los distintos tipos de respuesta (es interesante mencionar que por tabla  $\chi^2_{32;0,2}=38,47$  por lo que, de hecho, no se rechazaría la hipótesis nula ni con un

nivel de confianza del 80%). Los parámetros e interacciones del modelo final ED+Best+Worst<sub>VSG</sub> se muestran en la Tabla 6-4.

Tabla 6-4: Función Utilidad Desagregada por Atributo Modelo ED+Best+Worst<sub>VSG</sub>

Atributo	Componente de la Función de Utilidad
Acceso al metro	$-0,324 + 0,057 \cdot \text{Sexo} - 0,060 \cdot \text{Maipú} - 0,084 \cdot \text{Hab}_1$ (8,9) (3,1) (1,8) (3,9)
Áreas Verdes	$-0,250 + 0,076 \cdot \text{Sexo} - 0,192 \cdot \text{Ñuñoa} - 0,079 \cdot \text{Maipú} + 0,039 \cdot \text{Dorm}_{1o2}$ (7,0) (2,9) (4,4) (1,7) (1,6)
Dividendo departamento	<b>BW:</b> $-0,177 + 0,151 \cdot \text{Edad}_{16-25} - 0,182 \cdot \text{Ñuñoa} - 0,134 \cdot \text{Estadía}_{<5 \text{ años}}$ (5,1) (2,8) (2,4) (2,6) <b>ED:</b> $-1,783 + 0,782 \cdot \text{Edad}_{16-35} + 0,50 \cdot 10^{-4} \cdot \text{IFAM}$ (5,0) (2,7) (2,7)
Gimnasio y piscina en el edificio	<b>BW:</b> $-0,082 + 0,094 \cdot \text{FueraStgo} + 0,132 \cdot \text{Ñuñoa} + 0,122 \cdot \text{Prov} + 0,070 \cdot \text{Hab}_{2o \text{ más}}$ (3,1) (1,8) (2,7) (2,6) (2,5) <b>ED:</b> $0,295 - 0,228 \cdot \text{StgoCentro}$ (4,4) (2,0)
Limpieza y estado de mantención de la calle	$0,173 + 0,094 \cdot \text{FueraStgo} + 0,175 \cdot \text{Ñuñoa} + 0,125 \cdot \text{Maipú} + 0,082 \cdot \text{Estadía}_{<10 \text{ años}}$ (5,1) (2,1) (4,2) (3,0) (3,4) $+0,357 \cdot 10^{-4} \cdot \text{IFAM} + 0,070 \cdot \frac{\text{Auto}}{\text{Hab}}$ (2,4) (2,2)
Servicios comerciales básicos	$-0,247 + 0,043 \cdot \text{Sexo} + 0,041 \cdot \text{StgoCentro} - 0,116 \cdot \text{Ñuñoa} - 0,137 \cdot \text{Maipú}$ (6,7) (2,0) (1,8) (3,0) (3,5) $+0,059 \cdot \text{Auto}_{1o \text{ más}} - 0,137 \cdot \text{Estadía}_{<5 \text{ años}} - 0,089 \cdot \text{Estadía}_{5a10 \text{ años}}$ (2,7) (4,7) (3,7)
Servicios culturales	<b>B-ED:</b> $-0,150 + 0,068 \cdot \text{StgoCentro}$ (4,4) (1,5) <b>Worst:</b> $-0,0185$ (0,9)
Tamaño de los dormitorios	$0,225 + 0,119 \cdot \text{Ñuñoa} - 0,064 \cdot \text{Estadía}_{>10 \text{ años}} - 0,054 \cdot \text{Dorm}_{1o2} - 0,044 \cdot \text{Ing}_{\text{Bajo}}$ (6,9) (2,9) (2,8) (2,5) (1,9) $+0,141 \cdot \frac{\text{Auto}}{\text{Hab}}$ (4,4)

En este modelo se estiman 64 coeficientes en total y se puede aplicar el test LR para verificar si efectivamente mejora al agregar VSG (recordar que en el modelo ED+Best+Worst se estimó 29 coeficientes en total con una log verosimilitud de -5323):

$$LR = -2 \cdot (-5.323 + 5.188) = 269 > \chi^2_{35;0,001} = 66,62$$

Por lo que se rechaza la hipótesis nula que los modelos son equivalentes al 99% de confianza, y se concluye que agrega información significativa al modelo haber estimado estos 35 parámetros adicionales.

### 6.2.5 Comportamiento Predictivo en la Muestra de Validación

Tal como se hizo para el modelo sin VSG, se examinará la capacidad predictiva considerando los distintos parámetros estimados para los atributos específicos en cada modelo. Así, se verá cuáles de estos realizan una mejor predicción. Los mejores resultados se muestran en la Tabla 6-5.

Tabla 6-5: Indicadores Predicción Modelos con VSG en Muestra de Validación

	Factor de Escala (test t)	$\lambda(\theta)$	Error Residual	FPR	ER	CR	$\chi^2_{FPR}$ (GL)	$\sigma_{norm}$
ED <sub>VSG</sub>	-	-3994	0,488	0,579	0,626	0,5	121,26 (4)	0,156
ED <sub>VSG</sub> FE	0,71 (7,2)	-3969	0,486	0,579	0,592	0,5	6,12 (4)	0,156
ED+Best <sub>VSG</sub>	-	-3793	0,472	0,635	0,606	0,5	19,99 (4)	0,266
ED+Best <sub>VSG</sub> FE	1,31 (5,7)	-3776	0,471	0,635	0,634	0,5	23,21 (4)	0,266
ED+Best+Worst <sub>VSG</sub>	-	-3689	0,463	0,715	0,594	0,5	70,15 (4)	0,428
ED+Best+Worst <sub>VSG</sub> FE	2,11 (15,6)	-3547	0,451	0,715	0,678	0,5	57,47 (4)	0,428

El modelo ED+Best<sub>VSG</sub> logra una mejor predicción en la muestra de validación al considerar la valoración del atributo *Gimnasio y piscina en el edificio* y de *Áreas verdes* estimadas mediante las respuestas *best*, y la del atributo *Dividendo departamento*

estimada mediante las respuestas ED. Por otro lado, el modelo  $ED+Best+Worst_{VSG}$  predice mejor al considerar el coeficiente de *Dividendo departamento* estimado mediante las respuestas ED, de *Gimnasio y piscina en el edificio* estimado mediante las respuestas  $Best+Worst$ , y de *Servicios culturales* estimado mediante las respuestas *Worst*.

A partir de la Tabla 6-5 se pueden comparar las predicciones de los modelos. Primero, se puede ver que en los tres casos es necesario incluir un factor de escala para que los modelos se adapten de mejor manera a la muestra (en todos los casos el factor es significativamente distinto de uno al 99% de confianza). Se puede ver que la mejor predicción la realiza el modelo  $ED+Best+Worst_{VSG}$  (tiene la log verosimilitud mayor y predice correctamente un 72% de la muestra). La segunda log verosimilitud mayor corresponde al modelo  $ED+Best_{VSG}$  que predice correctamente un 64% de la muestra. La peor log verosimilitud corresponde al modelo  $ED_{VSG}$  que sólo predice correctamente un 58% de la muestra.

En la penúltima columna de la Tabla 6-5 se puede ver además que, de acuerdo al test  $\chi^2_{FPR}$ , el único modelo consistente con los datos sería el  $ED_{VSG}$  ( $\chi^2_{4;0,05}=9,49$ ), pero debido a las limitaciones que tiene la aplicación de este test, no se considerará tan relevante al discriminar entre modelos. A su vez, se puede ver que los errores residuales y el índice de éxito normalizado también tienen un mejor valor en el modelo  $ED+Best+Worst_{VSG}$ , seguido por el  $ED+Best_{VSG}$ . La comparación entre la capacidad predictiva de los modelos sin heterogeneidad y con heterogeneidad se realizará en la sección 6.4.2.

### **6.3 Variaciones Sistemáticas en los Gustos y Efecto Panel**

En las secciones 6.1.2 y 6.1.4, se vio que al considerar heterogeneidad en las preferencias tratando correctamente el EP, tanto las respuestas *best* con respecto a las

*worst* y las respuestas ED con respecto a las *worst* mostraron tener un comportamiento muy distinto y por lo tanto no fue posible estimar un modelo conjunto para ellas. Se probó agregando VSG para investigar si al hacer esta inclusión se podía estimar un modelo conjunto, pero debido al alto grado de significancia de la desviación estándar al considerar efecto panel mediante componentes de error, tampoco se logró la estimación conjunta. Por este motivo, no se incluirán en esta sección.

### 6.3.1 Modelo $ED_{EP+VSG}$

Se estimó un modelo ( $ED_{EP+VSG}$ ) incluyendo heterogeneidad a través del efecto panel (componentes de error) y VSG, con 17 parámetros y una log verosimilitud de -854. La función de utilidad de este modelo se presenta en la Tabla 6-6.

Tabla 6-6: Función de Utilidad Desagregada Modelo  $ED_{EP+VSG}$

Atributo	Componente de la Función de Utilidad
Acceso al metro	-0,625 (9,0)
Áreas Verdes	-0,143 (2,2)
Dividendo departamento	$-3,46 + 1,16 \cdot Edad_{16-35} + 1,05 \cdot 10^{-3} \cdot IFAM$ (4,2) (1,7) (2,5)
Gimnasio y piscina en el edificio	$0,433 - 0,349 \cdot StgoCentro$ (5,2) (2,5)
Limpieza y estado de mantenimiento de la calle	$0,110 + 3,44 \cdot 10^{-4} \cdot IFAM$ (0,9) (3,1)
Servicios comerciales básicos	$-0,441 + 0,394 \cdot Dorm_3$ (5,0) (2,9)
Servicios culturales	-0,145 (2,2)
Tamaño de los dormitorios	$1,05 - 0,325 \cdot Sexo - 0,542 \cdot Auto_{0,1}$ (4,4) (2,3) (2,2)

Aplicando el test LR para verificar si efectivamente mejora el modelo al estimar interacciones entre variables socioeconómicas y atributos en las respuestas ED, queda:

$$LR = -2 \cdot (-878 + 854) = 48 > \chi^2_{7;0,001} = 24,33$$

Por lo tanto se rechaza la hipótesis nula que los modelos  $ED_{EP}$  y  $ED_{EP+VSG}$  son equivalentes y se concluye que incorporar los siete parámetros adicionales (que incluyen VSG) aporta información significativa.

### 6.3.2 Modelo $ED+Best_{EP+VSG}$

El modelo estimado para las respuestas  $ED+Best$  considerando un componente de error e incluyendo VSG, tiene 36 parámetros comunes y log verosimilitud de -2.840. Al considerar la misma especificación para las respuestas *best* solamente, se obtiene una log verosimilitud de -1.959 y para las ED una log verosimilitud de -855. Con estos valores se puede aplicar el test LR para verificar si es aceptable considerar estos 36 parámetros comunes a ambas respuestas, obteniéndose  $\lambda_a=54$ , y como  $\chi^2_{36;0,05}=51,00$  se rechaza ajustadamente la hipótesis nula que los modelos son equivalentes. Como antes, se procedió a graficar los coeficientes estimados para cada tipo de respuesta para detectar cuáles parecían comportarse de manera distinta y considerarse específicos a cada fuente de datos (Figura 6-7).

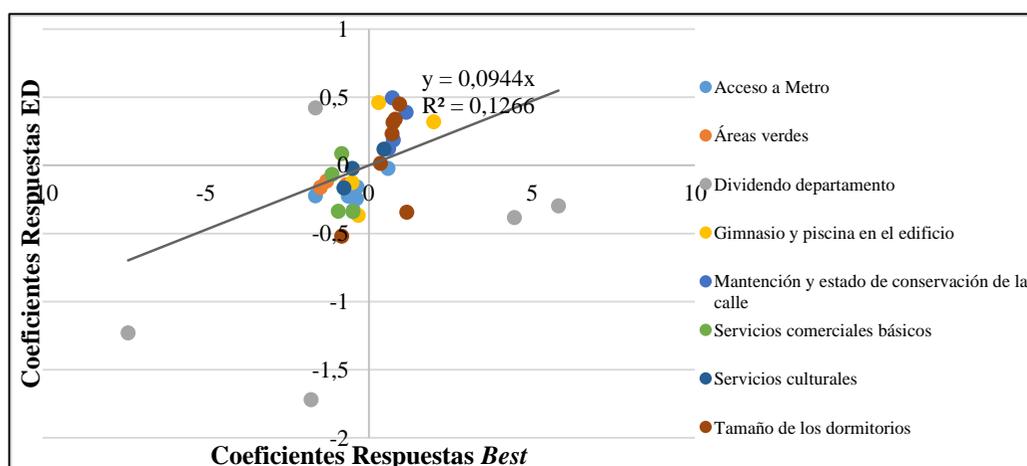


Figura 6-7: Coeficientes (sensibilidades) de Modelos con VSG y EP para Respuestas *Best* versus Respuestas ED

En la Figura 6-7 se puede ver que el atributo *Dividendo departamento* (marcado en gris) parece comportarse de manera más distinta en ambos casos, por lo que se procedió a analizar su consideración como específico en cada respuesta. Al investigar cuáles eran las VSG apropiadas para este atributo en cada tipo de respuesta, se obtuvo un modelo con 31 parámetros comunes y una log verosimilitud de -2.835. Al estimar esta especificación en cada banco de datos por separado, se obtuvo una log verosimilitud de -1.959 para las respuestas *best* y -852 para las ED. Con esto, se aplicó el test LR obteniéndose  $\lambda_a=48$ , y como  $\chi^2_{31;0,025}=48,23$ , no se rechazó la hipótesis nula que los modelos eran equivalentes con un 97,5% de confiabilidad.

Es decir, es adecuado especificar los parámetros del atributo *Dividendo departamento* como específicos para cada tipo de respuesta y los 31 parámetros correspondientes a los siete atributos restantes como comunes. La función de utilidad del modelo resultante, que llamaremos  $ED+Best_{VSG+EP}$ , se muestra en la Tabla 6-7.

Interesa aplicar el test LR para verificar si el modelo  $ED+Best_{EP}$  (que incluye sólo EP, tiene 23 parámetros y una log verosimilitud de -2.917) se podría considerar equivalente a este modelo con 27 parámetros adicionales:

$$LR = -2 \cdot (-2.917 + 2.835) = 164 > \chi^2_{27;0,001} = 55,48$$

Como se ve, se rechaza contundentemente la hipótesis nula de que ambos modelos son equivalentes y, por lo tanto, se concluye que incluir heterogeneidad mediante VSG adicionalmente a considerar correctamente el efecto panel, aporta mayor información al modelo.

Tabla 6-7: Función de Utilidad Modelo ED+BestVSG+EP

Atributo	Componente de la Función de Utilidad
Acceso al metro	$-0,401 - 0,094 \cdot Edad_{16-25} - 0,145 \cdot Edad_{26-35} + 0,133 \cdot Sexo - 0,173 \cdot Hab_1$ (5,7) (1,4) (2,4) (2,6) (2,6) $- 0,114 \cdot Hab_2$ (2,1)
Áreas Verdes	$-0,258 - 0,315 \cdot \tilde{N}u\tilde{ñ}oa - 0,154 \cdot Maipú$ (6,7) (2,8) (1,4)
Dividendo departamento	<b>Best:</b> $-1,89 - 0,362 \cdot \tilde{N}u\tilde{ñ}oa - 0,442 \cdot La\ Florida + 1,480 \cdot Dorm_1$ (4,8) (1,8) (1,6) (3,9) $+ 1,140 \cdot Dorm_2$ (3,1) <b>ED:</b> $-2,01 - 1,00 \cdot LaFlorida + 7,11 \cdot 10^{-4} \cdot IFAM$ (4,0) (1,7) (1,9)
Gimnasio y piscina en el edificio	$0,193 - 0,159 \cdot StgoCentro + 0,432 \cdot FueraStgo - 0,151 \cdot Hab_3$ (3,3) (2,0) (3,0) (1,9)
Limpieza y estado de mantención de la calle	$0,210 + 0,296 \cdot \tilde{N}u\tilde{ñ}oa + 0,315 \cdot \frac{Auto}{Hab} + 0,154 \cdot Estadía_{< de 5 años}$ (3,1) (2,2) (3,0) (1,7)
Servicios comerciales básicos	$-0,246 - 0,177 \cdot Maipú - 0,255 \cdot Estadía_{< de 5 años} - 0,136 \cdot Estadía_{5 a 10 años}$ (5,1) (1,8) (3,0) (2,2)
Servicios culturales	$-0,184 + 0,112 \cdot StgoCentro - 0,103 \cdot Auto_0$ (3,7) (1,6) (1,5)
Tamaño de los dormitorios	$0,270 + 0,213 \cdot Edad_{16-35} + 0,206 \cdot FueraStgo + 0,231 \cdot \tilde{N}u\tilde{ñ}oa$ (2,8) (4,5) (2,1) (2,8) $- 0,244 \cdot Auto_{0,1} + 0,184 \cdot \frac{Auto}{Hab} + 0,0817 \cdot Ing_{Medio/Alto}$ (3,1) (3,2) (2,0)

### 6.3.3 Comportamiento Predictivo en Muestra de Validación

Los modelos ED<sub>EP+VSG</sub> y ED+Best<sub>EP+VSG</sub> se aplicarán a predecir el comportamiento en la muestra de validación. En el segundo caso, se debe analizar cuál de los coeficientes estimados para el atributo *Dividendo departamento* (que es específico a cada respuesta) logra una mejor predicción. Se determinó que la mejor predicción correspondía al caso en que se valoraba el atributo *Dividendo departamento* como en las respuestas ED. Los resultados de las predicciones se pueden ver en la Tabla 6-8.

Tabla 6-8: Indicadores de Predicción Modelo  $ED_{EP+VSG}$  y Modelo  $ED+Best_{EP+VSG}$ 

	Factor de Escala (test t)	$\ell(\theta)$	Error Residual	FPR	ER	CR	$\chi^2_{FPR}$ (GL)	$\sigma_{norm}$
$ED_{EP+VSG}$	-	-3926	0,482	0,620	0,643	0,5	164,63 (4)	0,237
$ED_{EP+VSG}$ FE	0,73 (7,5)	-3899	0,480	0,620	0,609	0,5	38,39 (4)	0,237
$ED+Best_{EP+VSG}$	-	-3781	0,471	0,637	0,620	0,5	28,24 (4)	0,272
$ED+Best_{EP+VSG}$ FE	1,1 (2,3)	-3778	0,471	0,637	0,630	0,5	29,63 (4)	0,272

Se puede ver que todos los indicadores mejoran al considerar las respuestas ED en conjunto con las *best*. El modelo  $ED+Best_{EP+VSG}$  predice correctamente un 64% de las preferencias en la muestra de validación.

## 6.4 Análisis de Inclusión de Heterogeneidad

### 6.4.1 Log Verosimilitud

En la Tabla 6-9 se presentan las log verosimilitudes de todos los modelos discutidos en esta tesis. Se puede ver que al incluir heterogeneidad las log verosimilitudes van aumentando. El test de razón de verosimilitud pertinente fue aplicado en cada capítulo donde se estimaron los modelos, demostrando en todos los casos que incluir tanto la heterogeneidad observada (mediante VSG) como la no observada (efecto panel mediante componentes de error) agregaba información al modelo.

Tabla 6-9: Log Verosimilitudes de los Modelos Estimados

	<b>ED</b>	<b>ED+Best</b>	<b>ED+Worst</b>	<b>ED+Best+Worst</b>
<b>Sin Heterogeneidad</b>	-1.003	-3.161	-3.164	-5.323
<b>Con Heterogeneidad</b>	<b>EP</b>	-878	-2.917	-
	<b>VSG</b>	-980	-3.057	-
	<b>VSG y EP</b>	-854	-2.835	-

Al considerar el efecto panel (EP) mediante componentes de error, no fue posible agregar las respuestas ED y *worst* ni las respuestas *best* y *worst*. Por esto, no se pudo estimar modelos conjuntos y en la tabla aparecen las celdas correspondientes en blanco. Como en los modelos pertenecientes a cada tipo de respuesta la desviación estándar del componente de error resultó altamente significativa, se puede concluir que las respuestas ED y *best* son las únicas que pueden alinearse adecuadamente. Finalmente, el modelo ED+Best<sub>EP</sub>+VSG parece ser claramente la mejor especificación posible con los datos disponibles.

#### **6.4.2 Comportamiento Predictivo en la Muestra de Validación**

En la sección 4.6 se comparó las características de los individuos pertenecientes al banco de datos utilizado para la estimación de modelos con la de los individuos encuestados en la muestra de validación. Se constató la existencia de diferencias en relación a las variables *Comuna de origen* y *Sexo*. Por este motivo, se intuyó que si estas variables tenían importancia en explicar las preferencias en el banco de datos, posiblemente las variaciones sistemáticas en los gustos estimadas no explicarían las de la muestra de validación. En los capítulos anteriores se vio que, efectivamente, las variables *Sexo* y *Comuna de origen* aparecían como significativas al interaccionarse con distintos atributos (al incluir VSG). Además, como se anticipó, la estimación de preferencias individuales en la muestra de validación empeoró al incluir VSG.

En la Tabla 6-10 se muestran las log verosimilitudes asociadas a predecir el comportamiento observado en la muestra de validación utilizando los distintos modelos estimados. Se puede ver que los modelos que incluyen EP predicen mejor que los que no consideran heterogeneidad (en los dos casos comparables). También se puede ver que los modelos que incluyen sólo VSG predicen incluso peor que los modelos que no consideran heterogeneidad. Esto era relativamente esperable debido a las diferencias en las características de los individuos en el banco de datos y muestra de validación.

Tabla 6-10: Log-Verosimilitud de Predecir en la Muestra de Validación

		<b>ED</b>	<b>ED+Best</b>	<b>ED+Worst</b>	<b>ED+Best+Worst</b>
<b>Sin Heterogeneidad</b>		-3.762	-3.690	-3.461	-3.451
<b>Con Heterogeneidad</b>	<b>EP</b>	-3.680	-3.579	-	-
	<b>VSG</b>	-3.969	-3.776	-	-3.547
	<b>VSG y EP</b>	-3.900	-3.778	-	-

Por otro lado, resulta interesante notar que el modelo que sólo utiliza las respuestas ED logra una mejor predicción en la muestra de validación cuando se incluyen VSG y EP en conjunto que cuando se incluyen sólo VSG. Esto no ocurre para el modelo *ED+Best*, que consigue una predicción muy similar para el modelo incluyendo EP en conjunto con VSG que para el modelo con sólo VSG (de hecho, el modelo incluyendo ambos tipos de heterogeneidad predice peor, pese a que la diferencia es insignificante).

Es importante mencionar que debido a las diferencias en las características de los individuos pertenecientes a la MV y el banco de datos, no es ilustrativo comparar las predicciones de modelos con y sin heterogeneidad, ya que las variaciones sistemáticas de gustos son distintas y las estimadas no se adaptan correctamente a la muestra de validación. Aun así, resulta interesante notar que la unión entre las respuestas *best* y/o *worst* con las ED logra una mejor predicción que las respuestas ED por sí solas en todos los casos.

### 6.4.3 Coeficientes Estimados

Como ya se ha comentado, en la exhaustiva revisión bibliográfica realizada no se detectó investigaciones previas que hubieran incluido VSG y EP de forma conjunta en encuestas del tipo *best worst*. Por lo mismo, resulta relevante analizar cómo afectan los parámetros estimados de los modelos estos dos tipos de heterogeneidad.

### *Intercepto*

A continuación se analiza el efecto que tiene incluir heterogeneidad en el intercepto de los atributos, es decir, en la importancia que tienen los atributos por sí solos (distinguiéndolo del impacto en la variación de los niveles expuestos en la encuesta). Se debe recordar que como el atributo *Servicios comerciales básicos* se consideró como base (esto es, su intercepto es siempre cero) para estimar el resto, los correspondientes a los demás atributos se encuentran en relación a éste; por esto, es interesante analizar la escala relativa y sus cambios. Los valores del intercepto de cada atributo para los distintos modelos estimados se presentan en la Tabla 6-11.

Se puede ver que el intercepto del atributo *Dividendo departamento* en los tres primeros modelos (representados en las tres primeras columnas) resultó ser el más cercano al origen (con la valoración más cercana al atributo *Servicios comerciales básicos*) pese a la variación en su signo. Es más, en el modelo que incluye VSG y EP resultó ser el segundo más cercano, siendo el intercepto de *Áreas verdes* el más cercano (y, en este modelo, cuando este intercepto resultó menos significativo). En este sentido, es interesante notar que para el modelo sin heterogeneidad y para el modelo que sólo incluía EP, el intercepto de *Áreas verdes* fue el segundo más cercano al origen y de acuerdo al test t no resultó significativamente distinto de cero al 90% de confiabilidad. Para el modelo que sólo incluía VSG, el intercepto de *Áreas verdes* resultó ser significativamente distinto de cero con 95% de confiabilidad y en ese caso fue el tercero más cercano al origen (superado por el del atributo *Servicios culturales*).

El intercepto de *Servicios culturales* resultó ser significativamente distinto de cero al 95% de confianza en todos los modelos exceptuando el que no incluía heterogeneidad (en éste resultó ser significativamente distinto de cero al 80% de confianza). A su vez, en todos los modelos, menos en el que incluía VSG, resultó ser el tercero más cercano al

origen; en cambio, en el modelo que incluía sólo VSG resultó ser el segundo más cercano al origen.

Tabla 6-11: Intercepto de Atributos en Modelos ED+Best sin y con Heterogeneidad

<b>Intercepto</b>	<b>Sin Heterogeneidad</b>	<b>EP</b>	<b>VSG</b>	<b>EP+VSG</b>
Acceso a metro	0,269 (4,5)	0,312 (4,1)	0,189 (4,8)	0,226 (4,0)
Áreas verdes	-0,085 (1,3)	-0,095 (1,2)	-0,105 (2,0)	-0,055 (0,9)
Dividendo departamento	0,044 (0,8)	-0,018 (0,2)	0,026 (0,6)	-0,059 (0,9)
Gimnasio y piscina en el edificio	-0,177 (2,8)	-0,212 (2,7)	-0,168 (3,2)	-0,265 (3,6)
Mantenimiento y estado de conservación de la calle	-0,273 (4,0)	-0,380 (4,3)	-0,278 (4,6)	-0,386 (4,8)
Servicios comerciales básicos	0	0	0	0
Servicios culturales	-0,086 (1,5)	-0,166 (2,2)	-0,100 (2,2)	-0,177 (2,7)
Tamaño de los dormitorios	0,504 (7,3)	0,626 (7,2)	0,387 (8,4)	0,508 (8,0)

En resumen, los tres atributos antes mencionados (*Servicios culturales*, *Áreas verdes* y *Dividendo departamento*) resultaron ser siempre los más cercanos al origen, cambiando sus posiciones relativas entre ellos.

Los valores más negativos, correspondientes a los atributos *Mantenimiento y estado de conservación de la calle* y *Gimnasio y piscina en el edificio*, se mantuvieron como el menor y segundo menor respectivamente en todos los modelos. Además, mostraron ser significativamente distintos de cero con un 99% de confiabilidad, por lo que parece adecuado que no hayan cambiado de posición en la escala relativa. Por otro lado, los valores más positivos, correspondientes a los atributos *Tamaño de los dormitorios* y *Acceso a metro*, también se mantuvieron en todos los modelos como el valor más lejano y segundo más lejano al origen respectivamente. Además, se puede ver que en todos los casos resultaron ser significativamente distintos de cero con 99% de confiabilidad.

En conclusión, se puede ver que al incluir heterogeneidad la escala relativa del intercepto de los atributos varía, especialmente en el caso de atributos que no resultaron ser significativamente distintos de cero con un 90% de confiabilidad. De hecho, las escalas relativas del modelo sin heterogeneidad y del modelo que incluye tanto VSG como EP, se mantienen en su mayoría exceptuando el intercepto del atributo *Dividendo departamento* que cambia su signo.

### *Sensibilidades*

También resulta interesante analizar cómo se ven afectadas las sensibilidades de los atributos debido a la inclusión de heterogeneidad. Para esto, el caso de modelos con VSG, se calculará un promedio ponderado por las características de los individuos en el banco de datos y su valor se analizará en relación al promedio de la muestra (por esto, no se incluye el test t en los modelos que incluyen VSG).

La Tabla 6-12 muestra cómo cambia la importancia de la sensibilidad de los atributos al incluir heterogeneidad. En primer lugar se puede notar que todas las sensibilidades tienen el signo esperado. Para los atributos que tienen relación con distancia y para el atributo *Dividendo del departamento*, las sensibilidades tienen signo negativo. Y para el atributo *Tamaño de los dormitorios* y *Mantenimiento y estado de conservación de la calle* tienen signo positivo.

Para analizar el cambio en el grado de dispersión en la escala relativa, se puede calcular la desviación estándar de los parámetros. Los resultados se muestran en la Tabla 6-13. Para simplificar el análisis se presenta la desviación estándar de los parámetros que permitían realizar una mejor predicción en la muestra de validación. Esto se hace así, ya que si se considerara todos los parámetros estimados para cada modelo, los que consideraban más parámetros específicos a cada encuesta (por ejemplo, estimaron dos parámetros para la sensibilidad de uno o más atributos), tendrían por este mero hecho una mayor desviación estándar.

Tabla 6-12: Sensibilidades de Atributos Modelo ED+Best, sin y con Efecto Panel

<b>Sensibilidad</b>	<b>Sin Heterogeneidad</b>	<b>EP</b>	<b>VSG</b>	<b>EP+VSG</b>
Acceso a metro	-0,413 (10,0)	-0,556 (10,3)	-0,357	-0,514
Áreas verdes	<i>Best: -0,359 (5,7)</i> ED: -0,087 (1,6)	<i>Best: -0,419 (6,0)</i> ED: -0,129 (2,0)	<i>Best: -0,301</i> ED: -0,082	-0,296
Dividendo departamento	-0,472 (6,3)	<i>Best: -0,788 (5,9)</i> <i>ED: -1,41 (4,5)</i>	<i>Best: -0,821</i> <i>ED: -0,725</i>	<i>Best: -1,279</i> <i>ED: -1,429</i>
Gimnasio y piscina en el edificio	0,115 (3,2)	<i>Best: 0,045 (1,0)</i> ED: 0,298 (4,6)	<i>Best: 0,034</i> ED: 0,227	0,121
Mantenimiento y estado de conservación de la calle	0,324 (7,3)	0,422 (7,5)	0,300	0,394
Servicios comerciales básicos	<i>Best: -0,375 (6,5)</i> <i>ED: -0,186 (3,5)</i>	<i>Best: -0,452 (6,9)</i> <i>ED: -0,265 (4,1)</i>	-0,280	-0,364
Servicios culturales	-0,149 (4,5)	-0,196 (4,9)	-0,138	-0,181
Tamaño de los dormitorios	0,299 (9,4)	0,412 (10,0)	0,262	0,384

\*En itálica se encuentran los valores que se comportaron mejor en la muestra de validación.

Los datos de la tabla anterior muestran que al considerar correctamente el EP, las variaciones entre parámetros aumentan significativamente, haciendo que la escala relativa de sensibilidades tenga una mayor dispersión.

Como en los distintos modelos se debió estimar parámetros específicos (que no siempre eran los mismos) para cada tipo de respuesta, se decidió hacer un análisis más profundo sobre la escala relativa, considerando los parámetros de los atributos que se comportaban mejor en términos de predecir en la MV.

Observando la Tabla 6-12, se puede ver que en todos los casos el atributo con mayor sensibilidad es el *Dividendo departamento*; el valor de su coeficiente siempre se encuentra más lejos del origen en la escala relativa. De hecho, al incluir heterogeneidad la sensibilidad de este atributo aumenta su valor, tanto al ser estimado mediante las

respuestas ED como por las *best* (recordar que el atributo fue considerado específico para ambos tipos de respuesta en los modelos con heterogeneidad). El aumento en su sensibilidad al considerar heterogeneidad mediante EP+VSG es de un 202%; esto permite afirmar que el aumento en la desviación estándar (Tabla 6-13) al incluir los dos tipos de heterogeneidad en las preferencias (VSG+EP) se debe en gran medida al coeficiente del atributo *Dividendo departamento*, ya que su posición en la escala relativa se aleja bastante de la del resto de los atributos al incluir heterogeneidad (tiene una diferencia de 178% respecto a la sensibilidad del atributo que lo sigue, *Acceso a metro*); sin embargo, en el modelo sin heterogeneidad su valor es tan sólo 14% mayor que el del atributo *Acceso a metro* (que le sigue en la posición relativa).

Tabla 6-13: Desviación Estándar Sensibilidades para Modelos sin y con Heterogeneidad

	<b>Desviación Estándar Parámetros MV</b>
<b>Sin Heterogeneidad</b>	0,70
<b>EP</b>	2,45
<b>VSG</b>	0,82
<b>EP+VSG</b>	2,43

También se puede ver que en todos los modelos el coeficiente del atributo *Acceso a metro* tiene la segunda posición más lejana al origen. No obstante, su valor no se aleja tanto de los correspondientes al resto de los atributos al incluir heterogeneidad, como lo hace *Dividendo departamento*. De hecho, se puede ver que para el modelo sin heterogeneidad y para el modelo ED+*Best*<sub>EP+VSG</sub>, los parámetros negativos que le siguen corresponden al atributo *Servicios comerciales básicos* y *Áreas verdes (best)*. Para el modelo sin heterogeneidad la sensibilidad del atributo *Acceso a metro* tiene una diferencia del 10% con el atributo *Servicios comerciales básicos*, mientras que para el modelo ED+*Best*<sub>EP+VSG</sub> tiene una diferencia del 41% con este mismo. El aumento en

esta diferencia también explica el aumento en la desviación estándar al incluir heterogeneidad.

El coeficiente del atributo *Mantenimiento y estado de conservación de la calle* resulta ser el más significativo (de manera positiva) de la escala relativa en todos los modelos. Éste es seguido, en todos los casos, por el atributo *Tamaño de los dormitorios*, pero la diferencia entre los coeficientes estimados es prácticamente despreciable (0,8% en el modelo sin heterogeneidad y 0,3% en el modelo ED+Best<sub>EP+VSG</sub>). Por lo que ambos atributos parecen tener sensibilidades muy similares en todos los modelos.

El atributo *Tamaño de los dormitorios* posee la segunda posición positiva en la escala relativa y en todos los casos es seguido por el atributo *Gimnasio y piscina en el edificio*. Pero esta vez las variaciones entre los coeficientes de estos atributos es muy grande (160% mayor que el del atributo *Gimnasio y piscina en el edificio* para el modelo sin heterogeneidad, y 217% mayor para el modelo ED+Best<sub>EP+VSG</sub>). Este hecho también explica el aumento en la desviación estándar al considerar heterogeneidad, y también ilustra que la importancia de los niveles mostrados para el atributo *Tamaño de los dormitorios* es mucho mayor que la del atributo *Gimnasio y piscina en el edificio*.

De hecho, el atributo *Gimnasio y piscina en el edificio* es el que tiene un valor positivo más cercano al origen en todos los modelos. Su valor no varía tanto, en relación al origen, para el modelo sin heterogeneidad (0,115) y para el modelo con VSG+EP (0,121). Esto muestra que el impacto de incluir ambos tipos de heterogeneidad en las preferencias en este atributo, no afecta en forma importante su sensibilidad. No obstante, al incluir sólo un tipo de heterogeneidad (ya sea EP o VSG por separado), el valor que toma la sensibilidad de este atributo disminuye significativamente (i.e. cuando se considera sólo EP el valor disminuye en un 156% y cuando se incluyen sólo VSG disminuye en un 238%). Esto muestra la importancia de incluir ambos tipos de heterogeneidad en la modelación, con el fin de explicar de mejor forma las variaciones

en las preferencias individuales en el banco de datos. El coeficiente del atributo *Servicios culturales* siempre tiene un valor negativo más cercano al origen. La variación en su valor se ve afectada por la inclusión de heterogeneidad (hay una diferencia de 21% en su sensibilidad entre el modelo sin heterogeneidad y el modelo ED+Best<sub>EP+VSG</sub>).

Para el modelo sin heterogeneidad, el segundo atributo con coeficiente negativo más cercano al origen es *Servicios comerciales básicos* (ED), mientras que para el modelo ED+Best<sub>EP+VSG</sub> es el del atributo *Áreas verdes*. Esto muestra un cambio importante, ya que ambos atributos parecen ser significativos en los modelos, y cambian su orden relativo al incluir heterogeneidad. Es importante mencionar que para los modelos que incluyen sólo un tipo de heterogeneidad (ya sea VSG o EP), la posición relativa de estos atributos se mantiene igual que en la del modelo sin heterogeneidad. Esto muestra, una vez más, el impacto que tiene en la estimación y valoración de atributos la inclusión de heterogeneidad.

Por otro lado, también es importante analizar el cambio en los atributos considerados específicos en los modelos. En el caso del modelo sin heterogeneidad se consideró de esta forma a los atributos *Áreas verdes* y *Servicios comerciales básicos*. Se puede ver, que en el caso del primero, el coeficiente estimado para su sensibilidad mediante las respuestas *best* es significativamente mayor (en valor absoluto) y más significativo que el estimado mediante las respuestas ED; además permite una mejor predicción en la MV. En el caso de *Servicios comerciales básicos*, el coeficiente estimado mediante las respuestas *best* es también significativamente mayor en valor absoluto que el estimado mediante las respuestas ED, aunque son igualmente significativos en ambos casos, y el estimado mediante las respuestas ED permite una mejor predicción en la MV.

En el modelo ED+Best<sub>EP+VSG</sub>, el único atributo específico para las respuestas es el *Dividendo departamento*. En este caso, el coeficiente estimado mediante las respuestas ED es mayor (en valor absoluto) que el de las respuestas *best* y la sensibilidad estimada

mediante estas respuestas permite una mejor predicción en la MV. Es interesante mencionar que en el caso de los modelos que sólo incluían un tipo de heterogeneidad (VSG o EP), el atributo *Dividendo departamento* también fue considerado específico y en todos los casos su valor para las respuestas ED permitió una mejor predicción. Aun así, al incluir sólo VSG el coeficiente estimado mediante las respuestas *best* resultó mayor en valor absoluto que el estimado mediante las respuestas ED, lo que no ocurrió para los otros dos modelos que incluían heterogeneidad.

#### 6.4.4 Análisis de Resultados

El análisis realizado en las secciones anteriores, permite concluir que es necesario incluir ambos tipos de heterogeneidad (VSG+EP) en los modelos. A partir de los coeficientes estimados para el modelo  $ED+Best_{EP+VSG}$ , se puede construir una escala relativa para mostrar la diferencia entre la importancia de los atributos y la significancia de los niveles expuestos en la encuesta (Figura 6-8). Es importante mencionar que los niveles de los atributos en la escala representan la media de las distribuciones, es decir, su intercepto sumado (o restado, dependiendo del nivel del atributo) con su sensibilidad.

Como se puede ver, uno de los atributos más interesantes es el *Dividendo departamento*: por un lado, la importancia que tiene el atributo en sí es la menos significativa (más cercano al origen), pero los niveles mostrados a los encuestados fueron tales, que sus sensibilidades resultaron ser las mayores. Por esto, en los puntos azules de la figura se puede ver que los valores extremos pertenecen a los niveles de este atributo. Esto significa que si los niveles de *Dividendo departamento* en la comuna de Santiago centro fueran similares entre ellos, es decir tuvieran una diferencia pequeña entre ellos, éste no sería un atributo significativo para los usuarios al elegir un departamento.

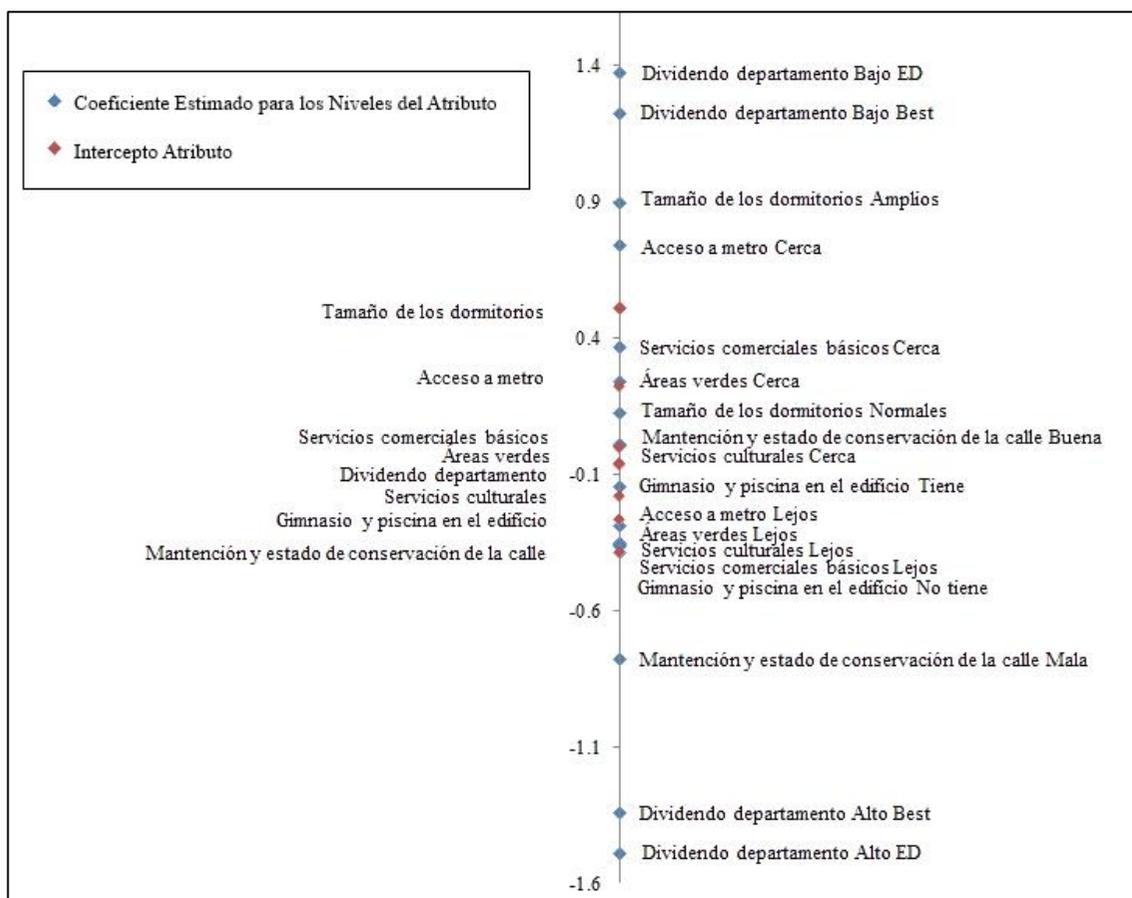


Figura 6-8: Escala Relativa Modelo ED+BestEP+VSG

El atributo *Tamaño de los dormitorios* mostró ser relevante por sí solo y también por la variación en los niveles mostrados al individuo. En efecto, por sí solo es el que tiene un valor positivo más lejano al origen, sugiriendo su gran importancia a la hora de elegir una vivienda. Además, su nivel “amplio” reporta gran utilidad a los usuarios, pero su nivel “normal” se acerca mucho al origen, teniendo un peso menor.

El atributo *Acceso al metro* se comporta de forma similar al *Tamaño de los dormitorios*. Por sí solo, se encuentra en la segunda posición positiva más lejano al origen, y la variación entre cuando está “cerca” (reporta una utilidad importante) y “lejos” (toma un valor negativo, es decir, genera des-utilidad) es alta. Sin embargo, en este último caso su

valor es bastante cercano al origen, por lo que la des-utilidad reportada no es tan relevante en comparación con otros atributos.

Como el atributo *Servicios comerciales básicos* fue tomado como base para estimar los interceptos, se encuentra por defecto en la posición de origen y el resto de los atributos se mide en comparación a su importancia. No obstante, se puede ver que los niveles mostrados sí parecieron ser relevantes, reportando una utilidad positiva para el nivel “cerca” y una des-utilidad para “lejos”; ambos coeficientes se encuentran, además, relativamente lejos del origen. Por otro lado, el intercepto del atributo *Mantenimiento y estado de conservación de la calle* tiene la posición relativa negativa más lejana al origen; esto permite deducir que su importancia en relación a *Servicios comerciales básicos* es significativa pero genera más des-utilidad que éste último. A su vez, los niveles mostrados al individuo parecen ser muy relevantes, es decir, cuando toma el nivel “malo” reporta gran des-utilidad y cuando toma el nivel “bueno” se encuentra más cerca del origen pero reporta utilidad positiva.

El atributo *Gimnasio y piscina en el edificio* resultó ser más importante en sí (intercepto) que *Servicios comerciales básicos*, pero los niveles mostrados al individuo (sensibilidades) no exhiben una variación importante en la escala, ni tampoco una sensibilidad alta en relación al resto de los atributos. El atributo *Áreas verdes*, en sí, tiene una importancia similar al atributo *Servicios comerciales básicos* (muy cercano al origen), pero tiene mayor importancia debido a la variación en los niveles mostrados a los encuestados. Para su nivel “cerca” reporta utilidad y se encuentra relativamente lejos del origen, y para su nivel “lejos” reporta des-utilidad pero se encuentra más cercano al origen. Finalmente, el atributo *Servicios culturales* fue uno de los que mostró ser menos significativo en cuanto a su importancia en sí y en relación a la variación de los niveles mostrados en la encuesta. En efecto, para el nivel “cerca” sólo reporta una utilidad pequeña en comparación con otros atributos, y para el nivel “lejos” también reporta una des-utilidad pequeña.

De esta forma, se puede concluir que los atributos más relevantes por sí solos fueron el *Tamaño de los dormitorios*, *Acceso a metro* y *Mantenimiento y estado de conservación de la calle*. Por otro lado, debido a la variación en los niveles mostrados en la encuesta, los atributos más relevantes fueron el *Dividendo departamento*, *Tamaño de los dormitorios*, *Acceso a metro* y *Mantenimiento y estado de conservación de la calle*.

## 6.5 Análisis por Atributo y Cálculo de Disposiciones al Pago

Para realizar este análisis se utilizarán los coeficientes estimados para el modelo que fue elegido como más apropiado para la muestra, es decir, el  $ED+Best_{EP+VSG}$ .

### 6.5.1 Acceso al Metro

En la Tabla 6-14 se muestran los parámetros estimados para *Acceso al metro* mediante el modelo  $ED+Best_{EP+VSG}$ . Todos los parámetros son negativos (como se vio en la sección 4.3 si el Metro se encuentra “cerca” el atributo toma valor -1); lo que indica que a los individuos les reporta más utilidad la cercanía al Metro.

Tabla 6-14: Parámetros del atributo *Acceso al Metro*

		1 Habitante	2 Habitantes	Más de 2 Habitantes
<b>Masculino</b>	<b>Entre 16 y 25 años</b>	-0,668	-0,609	-0,495
	<b>Entre 26 y 35 años</b>	-0,719	-0,660	-0,546
	<b>Mayores a 35 años</b>	-0,574	-0,515	-0,401
<b>Femenino</b>	<b>Entre 16 y 25 años</b>	-0,535	-0,476	-0,362
	<b>Entre 26 y 35 años</b>	-0,586	-0,527	-0,413
	<b>Mayores a 35 años</b>	-0,441	-0,382	-0,268

Se puede ver que las personas que más valoran la cercanía al Metro son aquellas en el rango de edad entre 26 y 35 años, seguido por las personas menores a 25 años. A su vez,

los hombres parecen valorar la cercanía al Metro más que las mujeres, y esta valoración decrece con el número de habitantes del hogar.

### 6.5.2 Áreas Verdes

El atributo *Áreas verdes* también se codificó de forma que para su nivel “cerca” el atributo tomaba el valor -1; por esto, es esperable que todos los parámetros tengan signo negativo (los parámetros estimados se muestran en la Figura 6-9). Se puede ver que las personas de Ñuñoa valoran de manera más positiva la cercanía a áreas verdes, seguido por las personas provenientes de Maipú.

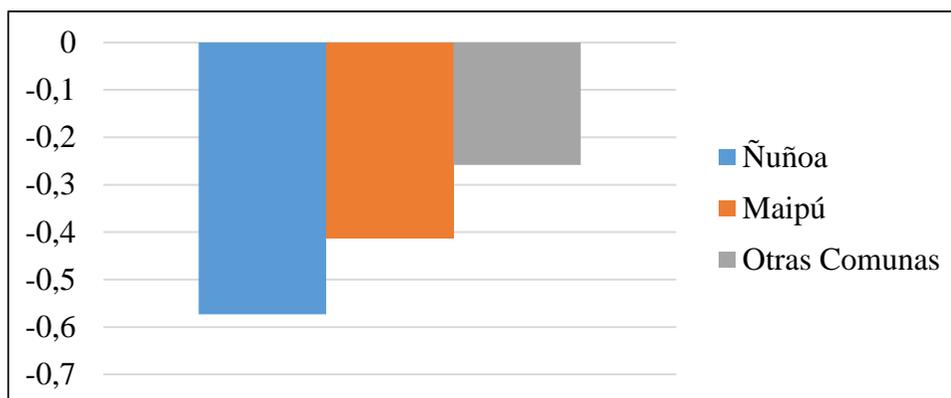


Figura 6-9: Parámetros del atributo *Áreas Verdes*

### 6.5.3 Dividendo Departamento

Este atributo se codificó de tal forma que si era menor tomaba un valor más negativo; por esto se esperaba que todos los parámetros estimados fueran negativos. En este caso, se realizaron dos análisis, ya que este atributo se estimó de manera específica para cada tipo de respuesta. A partir de la Tabla 6-15, que muestra los parámetros estimados mediante las respuestas *best* para el atributo *Dividendo departamento*, se puede ver que las personas provenientes de La Florida son más sensibles a un cambio de precios, seguido por las personas provenientes de Ñuñoa. Además, se puede ver que los

individuos que buscaban un departamento de más de dos dormitorios también resultaron ser más sensibles a una variación en los precios.

Tabla 6-15: Parámetros del atributo *Dividendo departamento* respuestas *best*

	La Florida	Ñuñoa	Otras Comunas
<b>1 dormitorio</b>	-0,852	-0,772	-0,410
<b>2 dormitorios</b>	-1,192	-1,112	-0,750
<b>Más de 2 dormitorios</b>	-2,332	-2,252	-1,890

En la Tabla 6-16, que muestra los parámetros estimados mediante las respuestas ED para el atributo *Dividendo departamento*, se puede ver que a medida que las personas tenían mayor ingreso eran menos sensibles a una variación de precios. Esto tiene sentido teórico, ya que en general las personas con mayor ingreso están dispuestas a pagar más. En este caso, la variable que mostró adaptarse mejor a los datos fue el Ingreso Familiar Mensual - IFAM (como variable continua y no como variable *dummy*). Por eso, en la Tabla 6-16 se muestran los parámetros estimados (promedios) para ciertos rangos de IFAM. Se puede ver que las personas de La Florida parecen tener una sensibilidad mayor hacia un aumento de precios que las personas provenientes de Otras Comunas. Esto último también resultó significativo en el caso de las respuestas *best*.

Tabla 6-16: Parámetros Promedio del *Dividendo departamento* en Respuestas ED

	La Florida	Otras Comunas
0<IFAM<\$600.000	-3,787	-1,797
\$600.000 < IFAM < \$1.000.000	-3,431	-1,441
\$1.000.000 < IFAM < \$2.000.000	-2,934	-0,944

#### 6.5.4 Gimnasio y Piscina en el Edificio

En la Figura 6-10 se muestran los parámetros estimados para este atributo. No fue posible definir el signo esperado, a priori, ya que algunas personas prefieren que el edificio no tenga espacios comunes (Torres, 2009).

Como se puede ver, las personas provenientes de la misma comuna (Santiago centro) y que tienen tres o más habitantes, valoran de manera negativa la existencia de gimnasio y piscina en el edificio (las personas de Santiago centro con menos de tres habitantes en el hogar tienen una valoración más pequeña pero positiva de este atributo). Por otro lado, las personas que provienen de fuera de Santiago tienen una valoración muy positiva de la existencia de estas dependencias. Probablemente esto tiene que ver con que fuera de Santiago suele haber más espacio y podría ser que al estar más acostumbrados a ello, valoren más la existencia de espacios comunes.

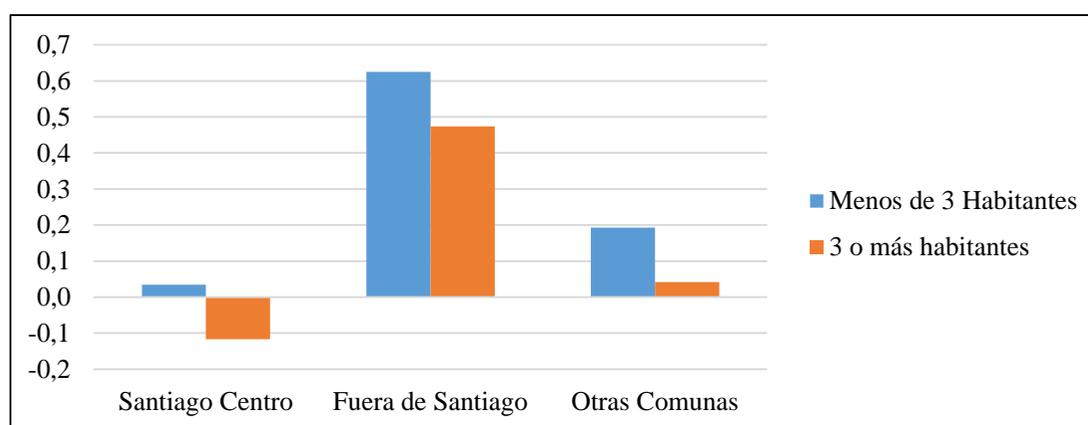


Figura 6-10: Parámetros del atributo *Gimnasio y piscina en el edificio*

### 6.5.5 Mantenimiento y Estado de Conservación de la Calle

Este atributo fue codificado de tal forma, que si el estado era “bueno” tomaba el valor +1, por lo que se espera que todos los parámetros sean positivos (ver Tabla 6-17).

Tabla 6-17: Parámetros del Atributo *Mantenimiento y estado de conservación de la calle*

	Ñuñoa		Otras Comunas	
	$0 < AH < 0,5$	$0,5 < AH < 1$	$0 < AH < 0,5$	$0,5 < AH < 1$
Estadía < 5 años	0,739	0,896	0,443	0,600
Estadía > 5 años	0,585	0,742	0,289	0,446

AH = Cociente entre número de autos y número de habitantes

Se puede ver que los hogares con un mayor cociente entre el número de autos y el número de habitantes (AH) tienen una mayor valoración del atributo. Asimismo, las personas que provienen de la comuna Ñuñoa y las personas que planean estar menos de cinco años en la vivienda, también valoran más este atributo.

### 6.5.6 Servicios Comerciales Básicos

Al igual que otros atributos relacionados con cercanía, fue codificado con valor -1 para el nivel “cerca”. Por esto, se esperaba que todos los parámetros tomaran valor negativo (ver Tabla 6-18). Se puede ver que las personas provenientes de Maipú tienen una mayor valoración de la cercanía a servicios comerciales básicos que las provenientes de Otras Comunas. Asimismo, a menor tiempo que planean quedarse en la vivienda, mayor es la valoración del atributo. Esto podría deberse a que quienes planean quedarse más tiempo están conscientes que las comunas se van desarrollando con el transcurso del tiempo; de esta forma, podría ser que no le den tanta relevancia al atributo como las personas que estarán menos tiempo, porque saben que es probable que más adelante se abran nuevos negocios cerca, o cambien los negocios que se encuentran cerca actualmente.

Tabla 6-18: Parámetros del Atributo *Servicios comerciales básicos*

	Maipú	Otras Comunas
Estadía < 5 años	-0,648	-0,471
0 < Estadía < 10 años	-0,559	-0,382
Estadía > 10 años	-0,423	-0,246

### 6.5.7 Servicios Culturales

Por la misma razón que el atributo anterior, se espera que *Servicios culturales* tenga coeficientes negativos, lo que se puede ver en la Tabla 6-19. Como es esperable, las personas que no tienen autos en el hogar valoran de manera más positiva la cercanía a

servicios culturales. Por otro lado, las personas provenientes de Santiago centro valoran en forma menos positiva la cercanía a servicios culturales que las personas provenientes de Otras Comunas. Esto podría ocurrir ya que estas últimas buscan encontrar un departamento en Santiago centro precisamente por encontrarse en el centro de la ciudad, y participar en más actividades culturales (que es uno de los grandes beneficios que tiene esta comuna en comparación con la mayoría de las restantes).

Tabla 6-19: Parámetros del Atributo *Servicios culturales*

	Sin Autos en el Hogar	Con Autos en el Hogar
Santiago Centro	-0,175	-0,072
Otras Comunas	-0,287	-0,184

#### 6.5.8 Tamaño de los Dormitorios

Se espera que los parámetros de este atributo tomen valores positivos codificó ya que el nivel “amplio” tomó valor +1 y el nivel “normal” -1; en este caso, muchas variaciones sistemáticas resultaron significativas como se muestra en la

Tabla 6-20. Se puede ver que las personas de Ñuñoa reportan mayor utilidad por tener dormitorios más grandes, seguido por la gente fuera de Santiago. Esto podría deberse a que en Ñuñoa existan viviendas de mayor tamaño (al igual que fuera de Santiago), y por ende los individuos valoren el mantener espacios amplios.

A su vez, las personas con mayor ingreso también valoran de manera más positiva la amplitud, lo que parece lógico. Esto también explica por qué los hogares con más de un auto en el hogar y que tienen un mayor cociente entre el número de autos y el número de habitantes (AH) le den mayor importancia a este atributo. Finalmente, también se puede ver que las personas menores de 35 años prefieren tener dormitorios más amplios.

Tabla 6-20: Parámetros del Atributo *Tamaño de los dormitorios*

			IFAM < \$600.000		\$600.000 < IFAM	
			0 < AH < 0,5	0,5 < AH < 1	0 < AH < 0,5	0,5 < AH < 1
Ñuñoa	Edad < 35 años	0 o 1 Auto	0,516	0,608	0,598	0,690
		> de 1 Auto	0,760	0,852	0,842	0,934
	Edad > 35 años	0 o 1 Auto	0,303	0,395	0,385	0,477
		> de 1 Auto	0,547	0,639	0,629	0,721
Fuera de Santiago	Edad < 35 años	0 o 1 Auto	0,491	0,583	0,573	0,665
		> de 1 Auto	0,735	0,827	0,817	0,909
	Edad > 35 años	0 o 1 Auto	0,278	0,370	0,360	0,452
		> de 1 Auto	0,522	0,614	0,604	0,696
Otras Comunas	Edad < 35 años	0 o 1 Auto	0,285	0,377	0,367	0,459
		> de 1 Auto	0,529	0,621	0,611	0,703
	Edad > 35 años	0 o 1 Auto	0,072	0,164	0,154	0,246
		> de 1 Auto	0,316	0,408	0,398	0,490

## 6.6 Cálculo de Disposiciones al Pago (WTP)

Para calcular las WTP de los distintos atributos del modelo  $ED+Best_{EP+VSG}$ , se procedió a computar su valor para cada individuo (considerando sus características socio-económicas), y en base a esto se calculó un promedio para el banco de datos. Es importante mencionar que como el atributo *Dividendo departamento* se codificó variando entre -1 y +1, se debió volver a sus valores numéricos para obtener las WTP. Para esto, se cambió todos los parámetros estimados que interactuaban con el atributo, con el fin de que quedaran en unidades 1/\$.

También es necesario recordar que todos los atributos relacionados con la distancia (es decir, *Acceso metro*, *Áreas verdes*, *Servicios comerciales básicos* y *Servicios culturales*) fueron codificados tal que cuando su nivel era “cerca” tomaban el valor -1 y cuando el nivel era “lejos”, +1. Por este motivo, las WTP por estos atributos debieran ser negativas, ya que los usuarios pagarían por estar cerca de estos atributos. Para el resto

de los atributos era esperable que las WTP quedaran con signo positivo. Así, procediendo de la forma explicada en la sección 3.8, se obtuvo los resultados que se muestran en la Tabla 6-21.

Tabla 6-21: Disposición al Pago (WTP) por Atributos Modelo ED+*Best*<sub>EP+VSG</sub>

Atributo	WTP (\$) considerando <i>Dividendo</i> <i>departamento Best</i>	WTP (\$) considerando <i>Dividendo</i> <i>departamento ED</i>
Acceso a Metro	-46.826	-58.187
Áreas verdes	-26.900	-33.427
Gimnasio y piscina en el edificio	11.001	13.670
Mantenimiento y estado de conservación de la calle	35.852	44.551
Servicios comerciales básicos	-33.140	-41.181
Servicios culturales	-16.510	-20.516
Tamaño de los dormitorios	34.993	43.483

A partir de la Tabla 6-21 se puede ver que la escala de las WTP no varía al utilizar los parámetros estimados tanto a partir de las respuestas *best* como de las ED; sin embargo, las WTP estimadas en base a las respuestas ED son mayores (ya que los parámetros estimados para *Dividendo departamento* mediante las respuestas ED son menores que los estimados para las respuestas *best*). Recordar que para la predicción del comportamiento individual en la MV, se eligió el coeficiente de *Dividendo departamento* estimado mediante las respuestas ED.

El orden de mayor a menor WTP es el siguiente: *Acceso metro*, *Mantenimiento y estado de conservación de la calle*, *Tamaño de los dormitorios*, *Servicios comerciales básicos*, *Áreas verdes*, *Servicios culturales* y, finalmente, *Gimnasio y piscina en el edificio*. Analizando las WTP estimadas en base al coeficiente de *Dividendo departamento* para las respuestas ED, se ve que la WTP por el atributo *Acceso a metro* tiene una diferencia de 31% con la del que le sigue. Asimismo, se puede ver que las disposiciones al pago

por *Tamaño de los dormitorios*, *Mantenimiento y estado de conservación de la calle* y *Servicios comerciales básicos*, son bastante similares (la mayor diferencia entre ellas es 1%). Sin embargo, para los tres atributos restantes las WTP varían bastante. De hecho, *Servicios comerciales básicos* (que es el menor de los tres anteriores) tiene una diferencia de 23% con *Áreas verdes*, y éste último una diferencia de 62% con el que le sigue (*Servicios culturales*). Asimismo, *Servicios culturales* tiene una diferencia de 50% con el atributo que presenta una menor WTP (*Gimnasio y piscina en el edificio*).

Finalmente, resulta interesante comparar las WTP estimadas a partir de nuestro modelo con las obtenidas por Torres (2009). En su trabajo, ella estimó distintos modelos MNL y ML; el más similar en estructura consideraba variaciones sistemáticas de los gustos, por lo que se utilizará para comparar las WTP. Como la manera de codificar los niveles de los atributos en Torres (2009) es distinta a la utilizada en esta investigación, lo correcto es comparar los valores absolutos de las WTP.

En la Tabla 6-22 se muestran las WTP estimadas por Torres (2009). En términos generales, se pueden constatar importantes variaciones respecto a las WTP estimadas con nuestros modelos.

Tabla 6-22: WTP Estimadas por Torres (2009)

	WTP (\$) MNL con VSG
Acceso a metro	79.700
Áreas verdes	26.800
Gimnasio y piscina en el edificio	19.200
Mantenimiento y estado de conservación de la calle	104.300
Servicios comerciales básicos	37.500
Servicios culturales	4.500
Tamaño de los dormitorios	48.300

Fuente: Torres (2009)

Para hacer un análisis más detallado, se puede construir una escala relativa con las WTP de Torres (2009) versus las estimadas por nuestro modelo (por simplicidad, consideraremos las WTP calculadas usando el parámetro de *Dividendo departamento* estimado mediante las respuestas ED, ya que se comportó más apropiadamente en la muestra de validación). En la Figura 6-11 se puede ver la escala construida.

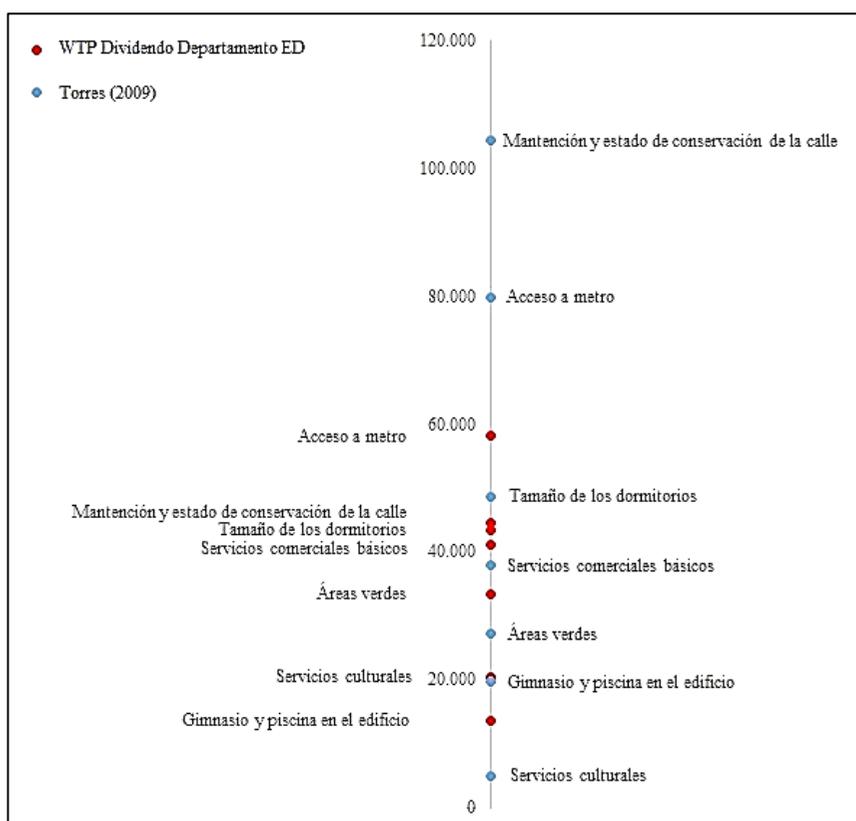


Figura 6-11: Escala WTP Estimadas por Nuestro Modelo ED+Best versus las Estimadas en Torres (2009)

Se ve que la WTP por *Mantenimiento y estado de conservación de la calle* en Torres (2009) es muy grande en comparación al resto de los atributos. En Torres (2009) se comenta que este valor, que parece obviamente sobre-estimado, podría haberse visto afectado por el hecho que el atributo se mostró mediante fotografías a los encuestados,

por lo que éstos pueden haber atribuido otros factores a la fotografía (por ejemplo, la seguridad del barrio). En nuestro modelo, la WTP por *Mantenimiento y estado de conservación de calle* parece más acorde a lo que se esperaría, y esto podría deberse a que en la encuesta B-W no se mostró mediante fotografías. El atributo que tiene segunda mayor valoración en Torres (2009), es el de *Acceso a metro*, que en nuestro modelo resulta ser el que tiene mayor WTP. En Torres (2009) se explica que ... “al consultar sobre otros atributos de barrio a las personas entrevistadas, mencionaron que no les importaría no tener algunos cerca de su hogar, siempre y cuando contaran con medios para acceder a ellos”. Esto es consistente con que este atributo tenga una alta disposición al pago en ambas muestras.

Así, pese a las diferencias de magnitud entre las WTP, los dos atributos con mayor WTP son los mismos en ambas investigaciones. Por otro lado, los dos atributos con menor WTP también son los mismos en ambos casos (*Servicios culturales y Gimnasio y piscina en el edificio*). Los atributos restantes también siguen el mismo orden de mayor a menor WTP: *Tamaño de los dormitorios*, *Servicios comerciales básicos* y *Áreas verdes*. Sin embargo, pese a que las WTP estimadas en ambas investigaciones tienen un orden relativo similar, las magnitudes son muy distintas. La mayor diferencia se encuentra en la WTP por *Servicios culturales*, que en Torres (2009) tiene un valor estimado significativamente menor que en nuestro modelo (356% menor). Otras WTP con diferencias porcentuales menores en el modelo de Torres (2009) son: *Áreas verdes* (25%) y *Servicios comerciales básicos* (10%). Para el resto de los atributos se obtuvo una WTP mayor en Torres (2009) que en nuestro modelo. Por ejemplo, *Mantenimiento y estado de conservación de la calle* (57%), pero como se explicó, esto se puede atribuir a que se presentó de distinta forma en ambas encuestas. La WTP por el atributo *Gimnasio y piscina en el edificio* tuvo una diferencia del 29%, seguido por la WTP de *Acceso a metro* (27%). Finalmente, la WTP por *Tamaño de los dormitorios* sólo tuvo una diferencia de 10% entre ambas investigaciones.

## 7 CONCLUSIONES

El objetivo principal de esta investigación fue estudiar la posible complementariedad entre encuestas *best worst* y de elección discreta binaria (ED). Para esto, se estimó modelos con datos de ambos tipos de encuesta, sin y con considerar heterogeneidad entre observaciones. Para el caso sin heterogeneidad, se encontró que todas las posibles combinaciones entre respuestas (esto es,  $ED+Best_{SINHET}$ ,  $ED+Worst_{SINHET}$  y  $ED+Best+Worst_{SINHET}$ ) eran adecuadas si se consideraba algunos atributos específicos.

Al incluir heterogeneidad del tipo efecto panel (EP) mediante componentes de error, se observó que esto indudablemente agregaba información al modelo. Sin embargo, no fue posible estimar atributos comunes entre las respuestas *worst* y los otros tipos de respuesta; sólo se pudo estimar un modelo combinando las respuestas *best* con las ED ( $ED+Best_{EP}$ ). Al incluir heterogeneidad mediante variaciones sistemáticas de gustos, se constató que esto mejoraba el modelo y, además, se logró estimar atributos comunes entre todas las respuestas (modelos  $ED+Best_{VSG}$  y  $ED+Best+Worst_{VSG}$ ). Finalmente, al incluir ambos tipos de heterogeneidad no fue posible, como se anticipaba, estimar atributos comunes entre las respuestas *worst* y las restantes, pero el modelo con las respuestas *best* y ED ( $ED+Best_{EP+VSG}$ ) resultó satisfactorio.

Un segundo objetivo de esta investigación era analizar el comportamiento predictivo (en la muestra de validación, MV) de los modelos estimados. Se comenzó por analizar las características socioeconómicas de los individuos pertenecientes a ambas muestras (MV y banco de datos), que mostraron tener diferencias relevantes. Por lo mismo, no se consideró apropiado comparar los modelos sin y con heterogeneidad en base a su capacidad predictiva, sino que sólo se examinó si al unir las respuestas *best* y/o *worst* con las ED, mejoraba el comportamiento predictivo en comparación al modelo estimado con las respuestas ED únicamente. En todos los casos, se llegó a la conclusión que al

unir las respuestas *best* y/o *worst* mejoraba el comportamiento predictivo de los modelos (en relación a uno equivalente estimado sólo con las respuestas ED).

Un tercer objetivo era analizar el impacto de incluir heterogeneidad en los modelos y en los coeficientes estimados. Se encontró que el modelo que incluía heterogeneidad EP y VSG ( $ED+Best_{EP+VSG}$ ) tuvo el mejor ajuste (log verosimilitud menor). Se observó que la inclusión de heterogeneidad afectaba tanto al intercepto como a la sensibilidad de los atributos. En conclusión, el aporte de las encuestas *best-worst* fue efectivamente diferenciar entre la importancia de los atributos y la importancia de los niveles utilizados para presentarlos. Con el mejor modelo estimado ( $ED+Best_{EP+VSG}$ ) se confeccionó una escala relativa (Figura 6-8) que permite diferenciar la importancia que tienen los atributos separadamente de la importancia de sus niveles. Así, se logró complementar la encuesta *best* con la de ED para agregar más información a la valoración de atributos de las viviendas del centro de Santiago.

En el caso de los interceptos, se constató que la escala relativa no variaba para los atributos extremos (aquellos con más y menos importancia relativa), pero los que estaban más cerca del origen (el atributo *Servicios comerciales básicos* se tomó como base) cambiaron su posición relativa al incluir heterogeneidad. Al comparar el modelo sin heterogeneidad con el modelo incluyendo ambos tipos de heterogeneidad, el único intercepto que cambió su posición fue el de *Dividendo departamento*; de hecho, este atributo también cambió su signo aunque no resultó significativamente distinto de cero al 70% de confiabilidad.

En resumen, se concluyó que los atributos más relevantes para los individuos, por sí solos (es decir, sin considerar los niveles presentados a los encuestados), son (en orden): *Tamaño de dormitorios* (positivo), *Mantenimiento y conservación de la calle* (negativo), *Gimnasio y piscina en el edificio* (negativo) y *Acceso a metro* (positivo). Asimismo, los atributos menos relevantes para los individuos (en orden) son: *Áreas verdes* (negativo),

*Dividendo departamento* (negativo) y *Servicios culturales* (negativo). Es importante recordar que todos estos interceptos están en una escala relativa en relación al atributo *Servicios comerciales básicos*, por lo que su signo sólo indica su importancia con respecto a ese atributo.

Para el caso de las sensibilidades, que tiene relación con la importancia de la variación en los niveles expuestos a los individuos en la encuesta, se constató que el atributo más afectado al incluir heterogeneidad fue el de *Dividendo departamento*. Este aumentó su valor significativamente en relación a la escala relativa. Además, se pudo ver que la desviación estándar de los parámetros estimados aumentó considerablemente al incluir heterogeneidad mediante VSG y EP.

La escala relativa, permitió ver que la variación en los niveles mostrados a los individuos pareció ser más relevante (en orden) en el caso de los atributos: *Dividendo departamento*, *Acceso a metro*, *Mantenimiento y estado de conservación de la calle* y *Tamaño de los dormitorios*. Asimismo, se vio que resultaron menos significativas las variaciones en los niveles planteados para los atributos *Gimnasio y piscina en el edificio* y *Servicios culturales*.

El último objetivo de esta investigación fue estimar disposiciones al pago por los atributos estimados correspondientes al mejor modelo ( $ED+Best_{EP+VSG}$ ). El orden, de mayor a menor WTP fue el siguiente: *Acceso metro*, *Mantenimiento y estado de conservación de la calle*, *Tamaño de los dormitorios*, *Servicios comerciales básicos*, *Áreas verdes*, *Servicios culturales*, y finalmente *Gimnasio y piscina en el edificio*. El primero de ellos (*Acceso a metro*) obtuvo una valoración muy superior al resto. Los tres atributos siguientes (*Mantenimiento y estado de conservación de la calle*, *Tamaño de los dormitorios*, *Servicios comerciales básicos*) tienen WTP similares. Luego, con una disminución significativa en su WTP, se encuentra el atributo *Áreas verdes*. Finalmente, los dos últimos atributos obtuvieron WTP considerablemente menores que el resto y,

por lo demás, bastante distintas entre sí (50% de diferencia), por lo que se ve claramente una tendencia a pagar menos por tener *Gimnasio y piscina en el edificio*.

En esta investigación se llegó a resultados interesantes y novedosos por dos motivos principales. El primero es metodológico, ya que de acuerdo a lo investigado no existen estudios anteriores que hayan utilizado encuestas B-W (Caso 2) para estimar modelos que incluyan heterogeneidad mediante EP y VSG en conjunto. El segundo tiene relación con el ámbito de políticas públicas, ya que se concluyó – en primer lugar - que los dos atributos más relevantes para elegir una vivienda en el centro de Santiago corresponden a características del barrio: *Acceso a metro y Mantenición y estado de conservación de la calle*. Asimismo, para cada atributo de barrio y vivienda se estimó qué características socioeconómicas de los individuos afectaban la importancia asignada a los distintos atributos.

Esta información complementa la lograda anteriormente por Torres (2009), ya que la posición relativa de la WTP de cada atributo es la misma excepto por el caso de *Mantenición y estado de conservación de la calle*. En la encuesta B-W este atributo se presentó con una descripción verbal, en cambio, en Torres (2009) se presentó mediante fotografías y se sospecha que esto puede haber provocado un alza considerable en su WTP en relación al resto de los atributos. De hecho, Torres (2009) menciona que esta alza se podría haber debido a que los individuos atribuyeron otras significancias a la fotografía, como por ejemplo, seguridad del barrio. Esto sugiere que la WTP obtenida en esta investigación mejora el resultado anterior, con un valor más parecido al del resto de los atributos. Aun así, también se pudo constatar que la magnitud de las WTP de varios atributos varió en relación a los obtenidos por Torres (2009).

## BIBLIOGRAFÍA

Coast, J., Flynn, T.N., Natarajan, L., Sproston, K., Lewis, J., Louviere, J.J., Peters, T.J. (2008) Valuing the ICECAP capability index for older people. *Social Science and Medicine* **67**, 874-82.

Couzner, L., Ratcliffe, J., Lester, L., Flynn, T. y Crotty, M. (2013) Measuring and valuing quality of life for public health research: application of the ICECAP-O capability index in the Australian general population. *International Journal of Public Health* **58**, 367-376.

Finn, A. y Louviere, J.J. (1992) Determining the appropriate response to evidence of public concern: the case of food safety. *Journal of Public Policy and Marketing* **11**, 12–25.

Flynn, T.N. (2010) Using conjoint analysis and choice experiments to estimate quality adjusted life year values: issues to consider. *Pharmacoeconomics* **28**, 711-722.

Flynn, T.N., Louviere, J.J., Peters, T.J. y Coast, J. (2010) Using discrete choice experiments to investigate heterogeneity in preferences for quality of life: variance scale heterogeneity matters. *Social Science and Medicine* **70**, 1957-1965.

Flynn, T.N., Peters, T.J. y Coast, J. (2013) Quantifying response shift or adaptation effects in quality of life by synthesizing best-worst scaling and discrete choice data. *Journal of Choice Modelling* **6**, 34-43.

Goett, A., Hudson, K. y Train, K. (2000). Customer choice among retail energy suppliers: the willingness-to-pay for service attributes. *Energy Journal*, **21**, 1-28.

Greene, M. y Ortúzar, J. de D. (2002). Willingness to pay for social housing attributes: a case study from Chile. *International Planning Studies*, **7**, 55–87.

Louviere, J.J. (2013) Modelling single individuals: the journey from psych lab to the app store. In S. Hess y A.J. Daly (eds.), *Choice Modelling: The State of the Art and the State of Practice*. Edward Elgar Publishing, Inc.

Louviere, J.J. y Swait, J.D. (1997) Separating weights and scale values in conjoint tasks using choices of best and worst attribute levels. Working Paper, Centre for the Study of Choice, University of Technology Sydney.

Louviere, J.J., Hensher, D.A. y Swait, J.D. (2000) *Stated Choice Methods: Analysis and Applications*. Cambridge University Press, Cambridge.

Louviere, J.J., Street, D., Burgess, L., Wasi, N., Islam, T. y Marley, A.A. (2008) Modelling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information. *Journal of Choice Modelling* **1**, 128-163.

Marley, A.A. y Louviere, J.J. (2005) Some probabilistic models of best, worst, and best–worst choices. *Journal of Mathematical Psychology* **49**, 464-480.

McCullagh, P. y Nelder, J.A. (1989) *Generalized Linear Models*. Second Edition. Chapman and Hall, London.

McFadden, D. (1974) The measurement of urban travel demand. *Journal of Public Economics* **3**, 303–328.

Munizaga, M. y Alvarez-Daziano, R. (2002). Evaluation of mixed logit as a practical modelling alternative. *Proceedings European Transport Conference*, Cambridge.

Munizaga, M., Correia, R., Jara-Díaz, S.R. y Ortúzar, J. de D. (2006) Valuing time with a joint mode choice–activity model. *International Journal of Transport Economics* **33**, 69–86.

Ortúzar J. de D. (2000) *Modelos Económicos de Elección Discreta*. Ediciones Universidad Católica de Chile, Santiago.

- Ortúzar, J. de D., Martínez, F.J. y Varela, F.J. (2000) Stated preferences in modelling accessibility. *International Planning Studies* **5**, 65–85.
- Ortúzar, J. de D. y Willumsen, L.G. (2011) *Modelling Transport*. Cuarta Edición, John Wiley and Sons, Chichester.
- Swait, J.D. y Bernardino, A. (2000) Distinguishing taste variation from error structure in discrete choice data. *Transportation Research* **34B**, 1-15.
- Swait, J.D. y Louviere, J.J. (1993) The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research* **30**, 305-314.
- Szeinbach, S.L., Barnes, J.H., McGhan, W.F., Murawski, M.M., Corey, R. (1999) Using conjoint analysis to evaluate health state preferences. *Drug Information Journal* **33**, 849-858.
- Torres, I. (2009) *Valoración de Atributos de Vivienda y Barrio en el Centro de Santiago*. Tesis de Magister en Ciencias de la Ingeniería, Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile.
- Torres, I., Greene, M. y Ortúzar, J. de D. (2013) Valuation of housing and neighbourhood attributes for city centre location: a case study in Santiago. *Habitat International* **39**, 62-74.
- Whitty, J., Ratcliffe, J., Chen, G. y Scuffham, P. (2013). A comparison of the discrete choice and best worst scaling methods to assess Australian public preferences for the funding of new health technologies. *International Choice Modelling Conference*, Sydney.

**ANEXOS**

## Anexo A: Tests Utilizados para Analizar Capacidad Predictiva en Muestra de Validación

### 1. Porcentaje correctamente predicho o FPR

Ortúzar (2000) define la FPR (*first preference recovery*) como el número o proporción de individuos que efectivamente elige la opción de máxima utilidad de acuerdo al modelo estimado. A su vez, la CR (*chance recovery*) se basa en el modelo equiprobable, donde todas las alternativas disponibles para el individuo  $q$  tienen igual probabilidad de ser elegidas. Por lo que se define como:

$$(A.1) \quad CR = \frac{1}{Q} \sum_{q=1}^Q \frac{1}{N_q}$$

donde  $N_q$  es el número de alternativas disponibles para el individuo  $q$ .

Como tercer indicador, se define al valor esperado del FPR como ER:

$$(A.2) \quad ER = E(FPR) = \sum_{q=1}^Q p_q$$

donde  $p_q$  es la probabilidad asociada a la mejor opción para el individuo  $q$  de acuerdo al modelo estimado.

Ortúzar (2000) explica que se puede comparar el valor de FPR para el modelo estimado, en relación al CR y ER concluyendo lo siguiente:

- Si las tres medidas son similares, el modelo es razonable pero poco informativo
- Si FPR y ER son similares y mayores que CR, el modelo es razonable e informativo
- Si FPR y ER no son similares, el modelo no explica la variación en los datos y debiera rechazarse.

Este indicador es bastante general, y por tanto no se considerará tan relevante al elegir los mejores modelos.

## 2. Test $\chi^2_{\text{FPR}}$

Este *test* busca analizar la consistencia del modelo estimado con los datos. El modelo permite predecir una probabilidad de elección para cada alternativa planteada a cada individuo en la muestra. Para realizar este test, se consideran bandas de probabilidad  $i$  (más detalles en Ortúzar, 2000) y se analiza la probabilidad de elección de cada alternativa  $j$ . Entonces, se define a  $N_{ij}$  como el número de observaciones en que el modelo asigna probabilidad en el rango  $i$  a la alternativa  $A_j$ ; por su parte,  $O_{ij}$  es el número observado de elecciones de la alternativa  $A_j$  a las que el modelo asigna una probabilidad de elección en el rango  $i$ . Con estos valores, se puede calcular el número esperado de individuos que elijen la opción  $A_j$  con probabilidad en el rango  $i$ , de acuerdo al modelo, como:

$$(A.3) \quad E_{ij} = N_{ij} \cdot \bar{p}_i$$

donde  $\bar{p}_i$  es la probabilidad promedio del rango  $i$ .

Así, se puede definir:

$$(A.4) \quad \chi^2_{\text{celda}} = \sum_{ij} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

pero se requiere que  $E_{ij} > 5$  para cada celda de la matriz. Como esto habitualmente no es posible, se define un caso menos informativo que permite comparar los totales esperados y observados de cada columna, es decir, para cada rango  $i$  se suma los valores de todas las alternativas, obteniéndose:

$$(A.5) \quad \chi^2_{\text{FPR}} = \sum_{i=1}^m \frac{(O_i - E_i)^2}{E_i}$$

donde  $m$  representa el número de columnas con  $E_i > 5$ . En este caso, se puede comparar  $\chi^2_{\text{FPR}}$  con el valor crítico  $\chi^2_{m-1;\alpha}$ . Si el valor  $\chi^2_{\text{FPR}}$  es mayor al valor de tabla para un nivel de confianza  $\alpha$ , se rechaza la hipótesis nula que el modelo es consistente con los datos. Es importante mencionar que el nivel de confianza definido para la distribución  $\chi^2$  disminuye la probabilidad de rechazar el modelo cuando era verdadero. En este *test* lo que nosotros buscamos es que no se rechace el modelo, y a medida que aumentemos el nivel de confianza, menor va a ser la probabilidad de rechazarlo. Por lo mismo puede parecer contra intuitivo que si se rechaza la hipótesis nula que el modelo es consistente con los datos a un nivel de confianza  $\alpha$  dado, al aumentar este nivel de confianza pueda no rechazarse la hipótesis nula. Por esto, hay que tener claro qué significa el nivel de confianza en la distribución  $\chi^2$  y cuál es la hipótesis nula de este test.

Este test tampoco se considerará un indicador decisivo al analizar la capacidad predictiva de los modelos, ya que tiene bastantes complicaciones. En primer lugar, se deja a criterio del modelador cómo definir los rangos de probabilidad (lo que puede cambiar de manera significativa los resultados). Además, en muchos casos se tiene que utilizar la versión más simple de este test porque no se cumple la condición necesaria que todas las celdas de la matriz sean mayores a cinco.

### 3. Índice de Éxito Normalizado

El índice de éxito es una medida que muestra el éxito predictivo del modelo estimado. Ortúzar (2000) lo define como:

$$(A.6) \quad \sigma = \sum_{i=1}^J \left( \frac{N_{ii}}{N_{i\cdot}} - \left( \frac{N_{\cdot i}}{N_{\cdot\cdot}} \right)^2 \right)$$

donde  $N_{ii}$  es el número esperado de individuos que eligen la alternativa  $A_i$  y que el modelo predecía que elegirían esa alternativa.  $N_{\cdot i}$  es el conteo predicho de la alternativa  $A_i$ , es decir, el número esperado de individuos que elegirían la alternativa  $A_i$  de acuerdo

al modelo estimado.  $N$  representa la suma de todos los conteos predichos (u observados, ya que son iguales) para todas las alternativas  $A_i$ .

Como este índice tiene un máximo teórico, dado por

$$(A.7) \quad \sigma^* = 1 - \left( \frac{N_i}{N} \right)^2$$

se requiere normalizarlo para que su máximo sea igual a uno. Así, entre mayor sea el índice de éxito normalizado  $\sigma_{norm}$  mejor será la predicción del modelo en la muestra de validación.

## Anexo B: Diseño de la Encuesta

Considere el departamento presentado más abajo. Indique en la columna de la izquierda el atributo que considera más atractivo o deseable, y en la columna de la derecha el atributo que considera menos atractivo o deseable.

Atributo más atractivo	Atributo del departamento	Valor	Atributo menos atractivo
<input type="checkbox"/>	Dividendo departamento	\$\$\$	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Tamaño de los dormitorios	Amplios	<input type="checkbox"/>
<input type="checkbox"/>	Gimnasio y piscina en el edificio	No tiene	<input type="checkbox"/>
<input type="checkbox"/>	Mantenimiento y estado de conservación de la calle	Mala	<input type="checkbox"/>
<input checked="" type="checkbox"/>	Acceso a metro	Cerca	<input type="checkbox"/>
<input type="checkbox"/>	Servicios comerciales básicos	Lejos	<input type="checkbox"/>
<input type="checkbox"/>	Servicios culturales	Lejos	<input type="checkbox"/>
<input type="checkbox"/>	Áreas verdes	Cerca	<input type="checkbox"/>

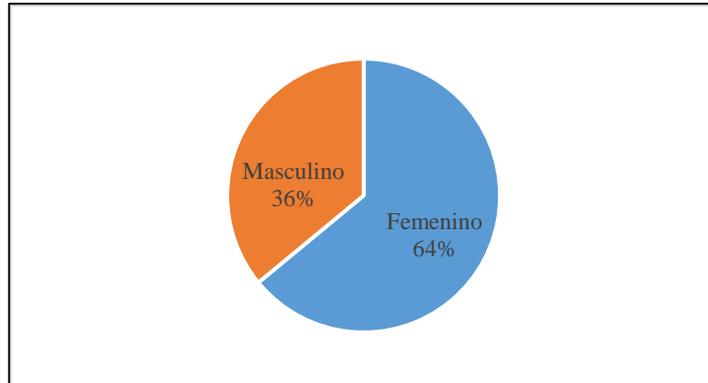
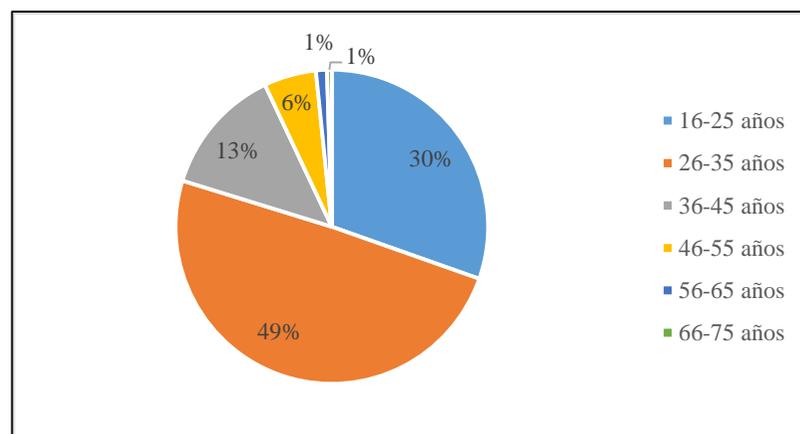
¿Arrendaría este departamento?

Si	No
<input type="checkbox"/>	<input checked="" type="checkbox"/>

Elección  
B-W

Elección  
Discreta  
Binaria

Figura A-1: Diseño de la Encuesta

**Anexo C: Variables Socioeconómicas de los Individuos en el Banco de Datos**Figura A-2: Distribución Variable *Sexo* en el Banco de DatosFigura A-3: Distribución Variable *Edad* en el Banco de Datos

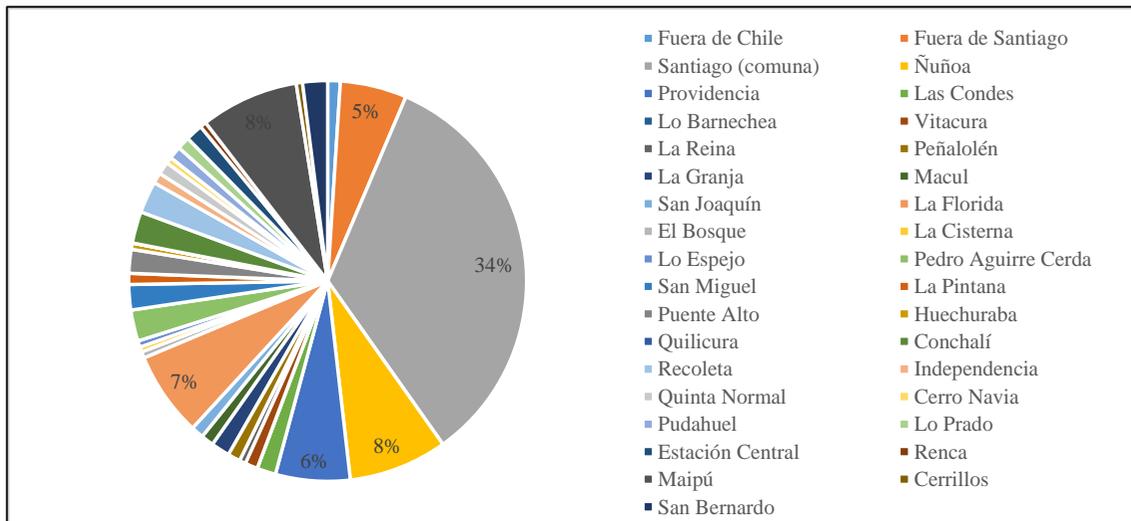


Figura A-4: Distribución por Comuna de Origen de los Individuos el Banco de Datos

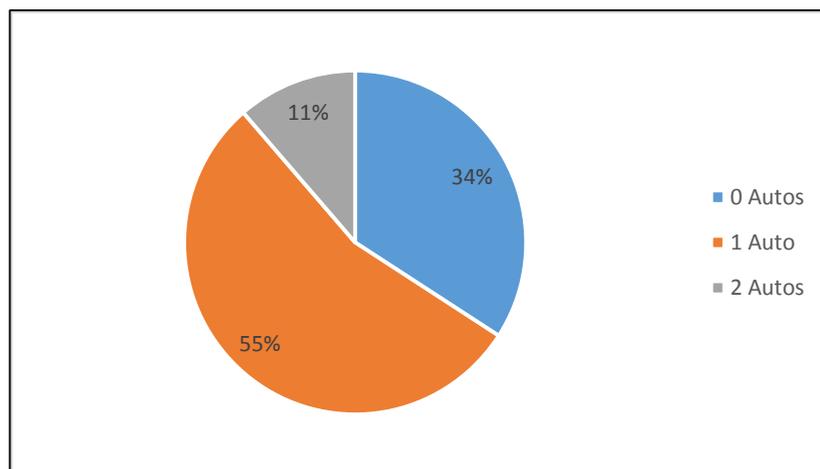


Figura A-5: Distribución Variable *Número de Autos* en el Banco de Datos

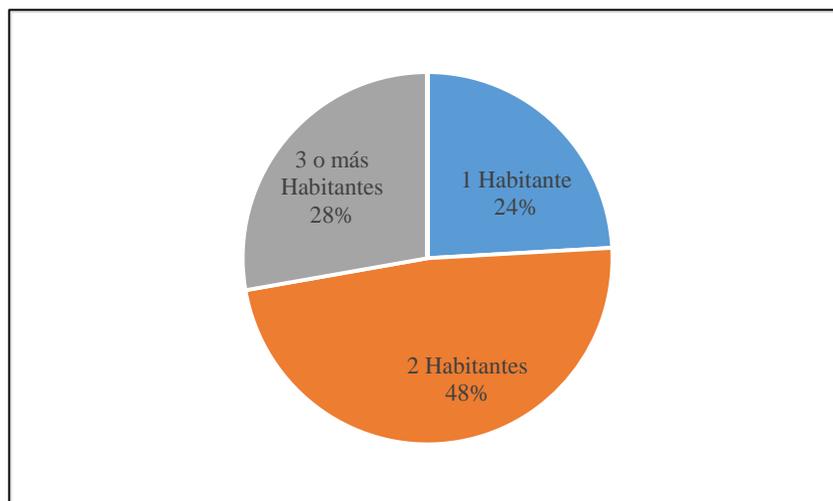


Figura A-6: Distribución Variable *Número de Habitantes* en el Banco de Datos

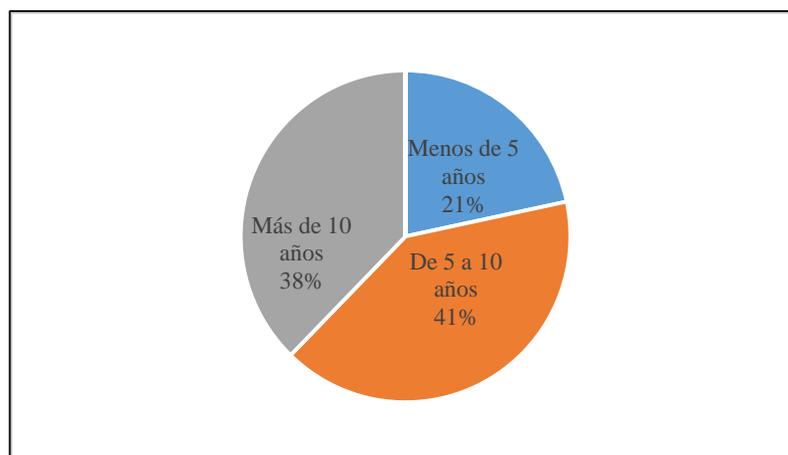


Figura A-7: Distribución Variable *Tiempo de Estadía* en el Banco de Datos

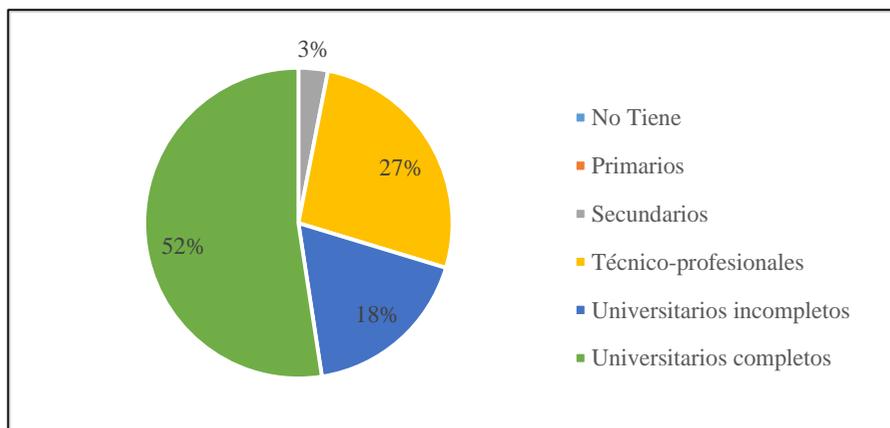


Figura A-8: Distribución Variable *Estudios* el Banco de Datos

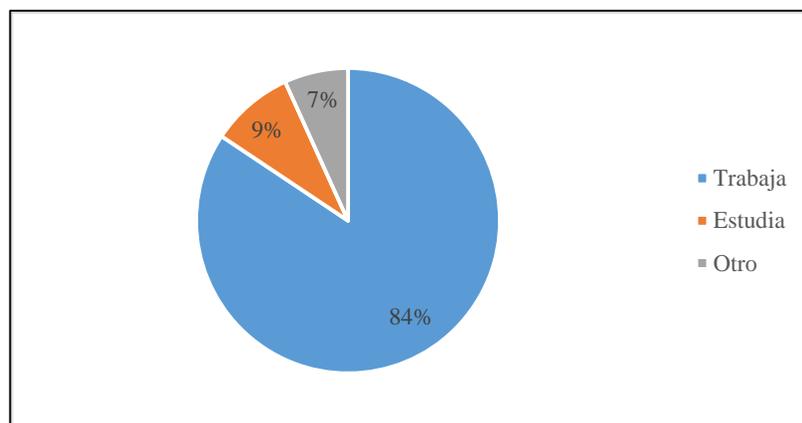


Figura A-9: Distribución Agregada Variable *Ocupación* en el Banco de Datos

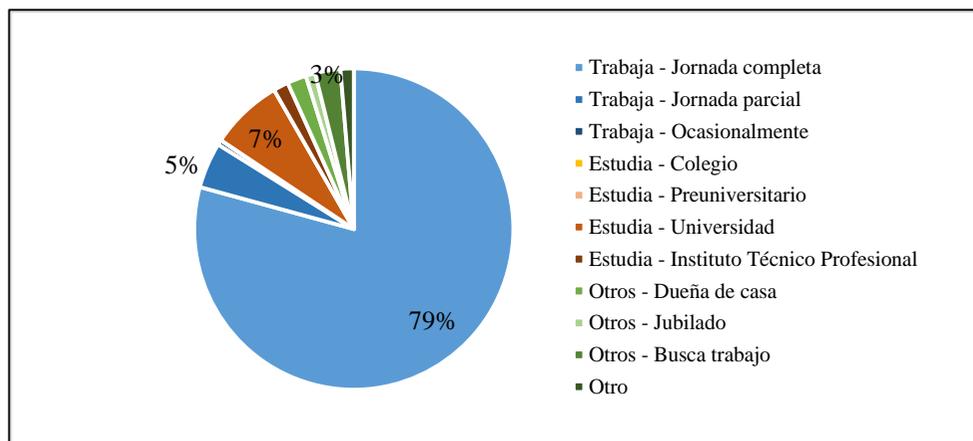


Figura A-10: Distribución Variable *Ocupación* el Banco de Datos

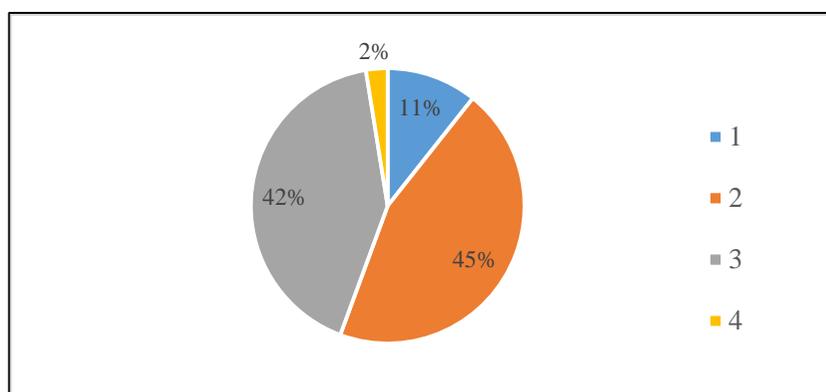


Figura A-11: Distribución Variable *Número de Dormitorios* el Banco de Datos

### Anexo D: Coeficientes Estimados Modelos Sin Heterogeneidad

Tabla A-1: Sensibilidades (pendientes) respuestas *Best*, *Worst* y *Best+Worst*

<b>Atributo/Modelo</b>	<b>Best (test t)</b>	<b>Worst (test t)</b>	<b>B+W (test t)</b>
Acceso a Metro	-1,502 (14,9)	-2,154 (13,8)	-1,591 (17,7)
Áreas verdes	-1,305 (7,0)	-1,276 (6,8)	-1,141 (9,2)
Dividendo departamento	-1,551 (6,5)	-1,038 (5,3)	-1,038 (7,3)
Gimnasio y piscina en el edificio	0,169 (1,1)	-0,101 (1,2)	-0,036 (0,6)
Mantenimiento y estado de conservación de la calle	1,140 (4,8)	2,307 (9,1)	1,623 (9,8)
Servicios comerciales básicos	-1,360 (8,5)	-1,698 (11,2)	-1,339 (12,1)
Servicios culturales	-0,710 (4,3)	-0,120 (0,9)	-0,311 (3,5)
Tamaño de los dormitorios	1,099 (17,1)	1,482 (5,0)	1,116 (18,0)
Factor de Escala	-	-	1,277 (3,3)
Log-verosimilitud	-2.151	-2.156	-4.328

Tabla A-2: Sensibilidades (pendientes) Modelo Final  $B+W$  sin Heterogeneidad

<b>Atributo/Modelo</b>	<b><math>B+W_{\text{SINHET}}</math> (test t)</b>
Acceso a Metro	-1,595 (17,3)
Áreas verdes	-1,150 (9,2)
Dividendo departamento	-1,052 (7,2)
Gimnasio y piscina en el edificio	-0,038 (0,6)
Mantenimiento y estado de conservación de la calle <i>Best</i>	1,140 (4,8)
Mantenimiento y estado de conservación de la calle <i>Worst</i>	1,837 (7,7)
Servicios comerciales básicos	-1,352 (12,0)
Servicios culturales <i>Best</i>	-0,704 (4,2)
Servicios culturales <i>Worst</i>	-0,097 (0,9)
Tamaño de dormitorios	1,106 (17,8)
Factor de Escala	1,257 (14,4)
Log-verosimilitud	-4.313

Tabla A-3: Coeficientes (pendientes) Modelos Sin Heterogeneidad Estimados con Respuestas *Best* o *Worst* combinadas con las ED

<b>Atributo/Modelo</b>	<b>ED (test t)</b>	<b>ED + <i>Best</i> (test t)</b>	<b>ED+<i>Worst</i> (test t)</b>
Acceso a Metro	-0,412 (7,6)	-0,344 (9,1)	-0,342 (8,8)
Áreas verdes	-0,080 (1,5)	-0,223 (7,2)	-0,173 (6,5)
Dividendo departamento	-0,686 (4,7)	-0,384 (6,1)	-0,180 (5,0)
Gimnasio y piscina en el edificio	0,220 (4,1)	0,090 (2,8)	-0,003 (0,2)
Mantenimiento y estado de conservación de la calle	0,331 (6,1)	0,291 (6,7)	0,342 (8,8)
Servicios comerciales básicos	-0,192 (3,6)	-0,267 (8,1)	-0,245 (8,4)
Servicios culturales	-0,082 (1,5)	-0,136 (4,5)	-0,028 (1,4)
Tamaño de dormitorios	0,263 (4,9)	0,244 (9,0)	0,238 (6,1)
Factor de Escala	-	4,54 (7,7)	6,58 (8,0)
Log-verosimilitud	-1.003	-3.169	-3.179

### Anexo E: Coeficientes Estimados Modelos con Efecto Panel (EP)

Tabla A-4: Coeficientes (sensibilidades) Modelo *Best*, *Worst* y ED Con Efecto Panel

<b>Atributo/Modelo</b>	<b><i>Best</i> (test t)</b>	<b><i>Worst</i> (test t)</b>	<b>ED (test t)</b>
Acceso a Metro	-1,86 (15,5)	-2,72 (15,3)	-0,601 (8,9)
Áreas verdes	-1,43 (7,5)	-1,46 (7,4)	-0,128 (2,0)
Dividendo departamento	-2,68 (7,0)	-2,28 (6,2)	-1,440 (4,7)
Gimnasio y piscina en el edificio	0,149 (0,9)	-0,086 (0,9)	0,297 (4,5)
Mantenimiento y estado de conservación de la calle	1,23 (5,0)	2,72 (10,2)	0,462 (6,9)
Servicios comerciales básicos	-1,54 (9,2)	-2,08 (12,5)	-0,266 (4,1)
Servicios culturales	-0,794 (4,6)	-0,167 (1,1)	-0,133 (2,0)
Tamaño de los dormitorios	1,42 (17,8)	1,62 (5,2)	0,353 (5,3)
<i>Sigma</i> (Desviación Estándar)	1,13 (17,8)	-1,34 (5,2)	1,18 (11,6)
Log - Verosimilitud	-2.036	-1.967	-878