



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

VALORACION DE ATRIBUTOS DE VIVIENDA Y BARRIO EN EL CENTRO DE SANTIAGO

IGNACIA TORRES RIVAS

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

Profesor Supervisor:
JUAN DE DIOS ORTUZAR SALAS

Santiago de Chile, (Marzo, 2009)

© 2009, Ignacia Torres Rivas



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

VALORACION DE ATRIBUTOS DE VIVIENDA Y BARRIO EN EL CENTRO DE SANTIAGO

IGNACIA TORRES RIVAS

Tesis presentada a la Comisión integrada por los profesores:

JUAN DE DIOS ORTUZAR SALAS

LUIS IGNACIO RIZZI CAMPANELLA

MARGARITA GREENE ZUÑIGA

JOSE MIGUEL AGUILERA RADIC

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

Santiago de Chile, (Marzo, 2009)

A Rolando por acompañarme,
ayudarme y animarme durante todo
el Magíster y a mi familia por su
apoyo incondicional.

AGRADECIMIENTOS

En primer lugar, quiero agradecer a mi profesor guía, Juan de Dios Ortúzar, por su apoyo, paciencia y completa disposición a todos los requerimientos que tuve en la elaboración de la tesis. Su dedicación y cariño me acompañaron en todo momento y dejaron su marca en cada etapa de la investigación.

También agradezco a Margarita Greene y Luis Rizzi, miembros muy activos de mi comisión de defensa. Ellos participaron desde el comienzo de la investigación con valiosas respuestas a consultas y opiniones que enriquecieron en todo momento la tesis.

Además, me gustaría destacar la participación de Elisabetta Cherchi y John Rose, quienes, desde Cagliari y Sydney respectivamente, entregaron orientación y valiosísimo *feedback* en las partes más complicadas y áridas de la investigación. Tanto la sección de diseño de la encuesta como la de formulación de los modelos finales no habrían resultado exitosas si no fuera por los innumerables mails que me enviaron con detalladas respuestas y propuestas a mis consultas.

De igual forma, los aportes de Paula Iglesias, María Francisca Yáñez y Joffre Swait siempre fueron entregados con mucha dedicación, paciencia y buena disposición. La manera en que trataron cada una de mis dudas, como si fuera propia, y el tiempo que dedicaron a pensar cómo ayudarme hicieron que varios aspectos de la tesis fueran pulidos con éxito.

Por otra parte, deseo destacar el apoyo dado por el Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT), el cual entregó fondos para la realización de esta tesis a través del Proyecto 1050928. De la misma forma, agradezco a Francisco Tagle por su aporte y buena disposición para compartir conmigo todo el material que necesité, y a Rolando Castillo por asesorarme en el diseño de la interfaz de la encuesta.

Finalmente, quisiera agradecer a todo el Departamento de Ingeniería de Transporte y Logística, tanto profesores, administrativos y alumnos, por acogerme, acompañarme y hacer del postgrado una etapa muy especial para mi.

INDICE GENERAL

	Pág.
RESUMEN.....	xiii
ABSTRACT.....	xiv
1. INTRODUCCION	1
2. CONTEXTO Y OBJETIVOS DEL ESTUDIO.....	3
3. MARCO TEORICO	6
3.1. Modelos de Elección Discreta.....	6
3.1.1. <i>Modelo Logit Simple (MNL)</i>	7
3.1.2. <i>Modelo Logit Mixto (ML)</i>	10
3.2. Disposición al Pago por Atributos	11
3.3. Encuestas de Preferencias Declaradas	13
3.4. Diseño de Experimentos de Elección.....	15
3.4.1. <i>Fundamentación Teórica</i>	15
3.4.2. <i>Metodología Utilizada</i>	17
3.5. Tamaño de Muestra.....	19
4. DISEÑO DEL EXPERIMENTO.....	21
4.1. Características Generales de la Encuesta	21
4.2. Selección de las Variables del Experimento	22

4.3.	Entrevistas en Profundidad	25
4.4.	Diseño Preliminar Encuesta	29
4.4.1.	<i>Generación del Experimento de Preferencias Declaradas</i>	30
4.4.2.	<i>Diseño e Implementación de Interfaz en Microsoft Excel</i>	37
4.5.	Encuestas Piloto	40
4.5.1.	<i>Encuesta Piloto 1 (EP1)</i>	41
4.5.2.	<i>Encuesta Piloto 2 (EP2)</i>	43
4.5.3.	<i>Encuesta Piloto 3 (EP3)</i>	45
4.5.4.	<i>Simulación para Recuperar Parámetros del Experimento Final</i>	51
4.6.	Diseño Final de la Encuesta	56
4.7.	Costos de Diseño y Realización de la Encuesta.....	58
5.	ANALISIS DE LA BASE DE DATOS	60
5.1.	Características Socio-Económicas	61
5.2.	Características de la Nueva Vivienda y sus Habitantes	68
6.	ESTIMACION DE MODELOS DE ELECCION DISCRETA	76
6.1.	Modelos Logit Simple.....	76
6.1.1.	<i>Modelo MNL1</i>	77
6.1.2.	<i>Modelo MNL2</i>	80

6.2. Modelo Logit Mixto.....	94
7. DISPOSICIÓN AL PAGO POR ATRIBUTOS.....	99
8. CONCLUSIONES.....	107
9. REFERENCIAS.....	111
ANEXOS.....	118
ANEXO A : Encuesta Facultad de Arquitectura	119
ANEXO B : Diseño Factorial de la Encuesta Inicial.....	120
ANEXO C : Interfaz en Microsoft Excel de la Encuesta Inicial	122
ANEXO D : Fotografías Iniciales Atributo <i>Limpieza_Calle</i>	127
ANEXO E : Niveles Atributo <i>Limpieza_Calle</i> , Encuesta Inicial.....	128
ANEXO F : Explicación y Elección Tipo Juego 1, luego de EP1	129
ANEXO G : Niveles Atributo <i>Limpieza_Calle</i> , luego de EP1	130
ANEXO H : Diseño Factorial para EP3.....	131
ANEXO I : Correo Electrónico de Contacto.....	133
ANEXO J : Diseño Factorial para Encuesta Final.....	135
ANEXO K : Interfaz en Internet de la Encuesta Final.....	137
ANEXO L : Modelo MNL2	142
ANEXO M : Parámetros <i>Cultura</i> , Modelo MNL2.....	144

INDICE DE TABLAS

	Pág.
Tabla 4-1: Lista Preliminar de Atributos, Encuesta Facultad de Arquitectura	23
Tabla 4-2: Cantidad de Personas que Respondieron Entrevistas en Profundidad	25
Tabla 4-3: Explicación de Atributos del Experimento.....	31
Tabla 4-4: Primer Valor de Parámetros Previos por Atributo.....	33
Tabla 4-5: Niveles Atributo <i>Superficie</i>	35
Tabla 4-6: Estructura de la Interfaz de la Encuesta.....	37
Tabla 4-7: Modelos Estimados con EP2	44
Tabla 4-8: Modelos Estimados con EP3	50
Tabla 4-9: Probabilidad Generada de Cantidad Requerida de Dormitorios	51
Tabla 4-10: Resultado Simulación para Recuperación de Parámetros	54
Tabla 4-11: Preguntas de Características Socioeconómicas de la Encuesta Final.....	57
Tabla 4-12: Límite Inferior Teórico de Cantidad de Encuestados, por Parámetro	58
Tabla 4-13: Costos de la Encuesta, Sin Incluir Trabajo de Tesista.....	59
Tabla 5-1: Distribución Variable <i>Ocupación</i>	63
Tabla 5-2: Análisis de <i>Outliers</i> de la Variable <i>Ingreso_Personal</i>	68
Tabla 5-3: Análisis de <i>Outliers</i> de la Variable <i>IFAM</i>	69

Tabla 6-1: Modelo MNL1	77
Tabla 6-2: Características del Modelo MNL2	81
Tabla 6-3: Parámetros de <i>Precio</i> , modelo MNL2	82
Tabla 6-4: Parámetros de <i>Metro</i> , modelo MNL2	84
Tabla 6-5: Parámetros de <i>Areas_Verdes</i> , modelo MNL2	85
Tabla 6-6: Parámetros de <i>Servicios</i> , modelo MNL2	89
Tabla 6-7: Parámetros de <i>Espacio_Común</i> , modelo MNL2	90
Tabla 6-8: Parámetros de <i>Limpieza_Calle</i> , modelo MNL2	92
Tabla 6-9: Parámetros de <i>Superficie</i> , modelo MNL2	93
Tabla 6-10: Modelo ML1	96
Tabla 6-11: Análisis de las Distribuciones de los Parámetros del Modelo ML1	98
Tabla 7-1: Disposición al Pago por Atributos, Modelo MNL1	99
Tabla 7-2: Disposición al Pago por Atributos, Modelo MNL2	100
Tabla 7-3: Disposición al Pago por Atributos, Modelo ML1, Método 1	100
Tabla 7-4: Disposición al Pago por Atributos, Modelo ML1, Métodos 2 y 3	101

INDICE DE FIGURAS

	Pág.
Figura 3-1: Estructura de la Matriz de Fisher para Experimento con Dos Juegos.....	19
Figura 4-1: Efecto de Pérdida de Encuestados por Página de la Encuesta	49
Figura 4-2: Error Cuadrático Medio entre Parámetros Estimados y Simulados, según la Desviación Estándar.....	53
Figura 4-3: Porcentaje de Cambio de Elección al Agregar Componente Aleatorio a la Utilidad, según Desviación Estándar	55
Figura 5-1: Distribución de Variable <i>Sexo</i>	61
Figura 5-2: Histograma de Variable <i>Edad</i>	62
Figura 5-3: Distribución Variable <i>Estudios</i>	63
Figura 5-4: Distribución Variable <i>Ocupación</i> en Detalle	64
Figura 5-5: Porcentaje Variable <i>Ocupación</i> por Sexo	64
Figura 5-6: Porcentaje de Respuestas por Comuna de Origen de Encuestados.....	65
Figura 5-7: Histograma de la Variable <i>Ingreso_Personal</i>	67
Figura 5-8: Histograma de la Variable <i>IFAM</i>	69
Figura 5-9: Gráfico de Dispersión entre <i>IFAM</i> e <i>Ingreso_Personal</i>	70
Figura 5-10: Distribución de Variable <i>N°_Autos</i>	71

Figura 5-11: Distribución Porcentual de Variable <i>N°_Habitantes</i>	71
Figura 5-12: Porcentaje de Respuestas según Número de Dormitorios.....	72
Figura 5-13: Porcentaje de Número de Dormitorios según <i>N°_Habitantes</i>	73
Figura 5-14: Porcentaje de Respuestas según <i>Años</i>	74
Figura 5-15: Porcentaje de <i>Años</i> según <i>N°_Habitantes</i>	74
Figura 5-16: Histograma Porcentual Variable <i>Edad</i> según <i>N°_Habitantes</i>	75
Figura 6-1: Comparación entre Parámetros <i>a priori</i> y Parámetros de MNL1	78
Figura 6-2: Comparación entre Parámetros <i>a priori</i> y Parámetros de MNL1	78
Figura 6-3: Comparación entre Parámetros <i>a priori</i> y Parámetros de MNL1	79
Figura 6-4: Parámetros de <i>Precio</i> , comunas de Maipú, Puente Alto y La Florida	83
Figura 6-5: Parámetros de <i>Precio</i> , otras comunas	83
Figura 6-6: Parámetros de <i>Metro</i>	84
Figura 6-7: Parámetros de <i>Areas_Verdes</i> , en Maipú, Puente Alto, La Florida, Providencia y Ñuñoa	85
Figura 6-8: Parámetros de <i>Areas_Verdes</i> , otras comunas.....	86
Figura 6-9: Parámetros de <i>Cultura</i> , <i>IFAM</i> < \$600.000	86
Figura 6-10: Parámetros de <i>Cultura</i> , \$600.000 < <i>IFAM</i> < \$1.000.000.....	87
Figura 6-11: Parámetros de <i>Cultura</i> , \$1.000.000 < <i>IFAM</i>	87

Figura 6-12: Parámetros de <i>Servicios</i>	89
Figura 6-13: Parámetros de <i>Espacio_Común</i> , comunas fuera de la ciudad de Santiago .	90
Figura 6-14: Parámetros de <i>Espacio_Común</i> , comuna de Santiago	91
Figura 6-15: Parámetros de <i>Espacio_Común</i> , otras comunas.....	91
Figura 6-16: Parámetros de <i>Limpieza_Calle</i>	93
Figura 6-17: Parámetros de <i>Superficie</i>	94
Figura 7-1: Disposición al Pago por Atributos, Comparación entre Modelos.....	102
Figura 7-2: Distribuciones WTP según Método 2 y 3, Atributo <i>Metro</i>	103
Figura 7-3: Distribuciones WTP según Métodos 2 y 3, Atributo <i>Cultura</i>	104

RESUMEN

El centro de Santiago tiene una serie de atributos, en cuanto a su calidad como lugar de residencia, que son difíciles de encontrar en otros lugares de la ciudad. El objetivo de esta tesis es identificarlos y determinar de qué manera son valorados por los nuevos pobladores del sector; éstos se han definido como compradores de departamentos construidos a partir del año 2006.

Se diseñó un ejercicio de preferencias declaradas con ocho atributos de interés, en dos juegos con sólo seis situaciones de elección cada uno, para disminuir la carga cognitiva de los encuestados. El primer juego se asoció al *barrio* (accesibilidad a transporte público, áreas verdes, servicios comerciales básicos y servicios culturales) y el segundo a la *vivienda* (tamaño, existencia de espacios comunes en el edificio, y mantención y estado de conservación de la calle); el precio se mantuvo como atributo común en ambos juegos. En la etapa de diseño experimental se optó por la técnica de diseños eficientes; y específicamente se buscó minimizar el indicador denominado Error-D. Como para calcularlo es necesario contar con una estimación inicial de los parámetros del modelo, este problema se resolvió en forma heurística en base a los resultados de sucesivas encuestas piloto.

Con los datos obtenidos de la encuesta final, realizada a través de una página *web*, se estimaron modelos logit simple y mixto. Éstos arrojaron como resultado disposiciones al pago mayores por los atributos relacionados con necesidades prácticas, y menores para aquellos ligados al esparcimiento. Dos de los atributos de este último tipo incluso cuentan con disposición al pago negativa para ciertos individuos (espacios comunes en el edificio y acceso a servicios culturales). Los dos atributos más valorados resultaron la accesibilidad a metro y la mantención y estado de conservación de la calle.

Palabras Claves: experimentos de preferencias declaradas, diseños eficientes, elección residencial, repoblamiento de centros urbanos.

ABSTRACT

The center of Santiago City has some attributes as residence area that are difficult to find combined in other places of the city. The objective of this thesis is to identify them and determine how the new residents value them. This group is defined as people who have bought apartments built since 2006.

A stated preference experiment was designed with eight attributes of interests, separated into two games of six choice situations each, to lower the cognitive stress of the respondents. The first game was associated to *neighborhood* (accessibility to public transport, green areas, basic commercial services and cultural services) and the second one to the *dwelling* (size, existence of gym and pool in the building, and cleanness and state of the street); the price was the only attribute present in both games. In the stage of experimental design the efficient designs technique was used; this was done through the minimization of an indicator called *D-error*. To obtain its value an initial estimation of the model's parameters is needed; an heuristic was used to overcome this problem, through the results of successive pilot surveys.

With the data base obtained in the final survey, which was done through a web page, simple and mixed logit models were estimated. The results showed higher willingness to pay for attributes that cover practical needs, and lower ones for attributes related to recreation. Two of the attributes of this type even have negative willingness to pay for some individuals (existence of gym and pool in the building, and accessibility to cultural services). The attributes that were more highly valued are the accessibility to public transport, and cleanness and state of the street.

Keywords: stated choice experiments, efficient designs, residential choice, repopulation of urban centers.

1. INTRODUCCION

Los modelos de elección discreta, y en particular su estimación utilizando experimentos de preferencias declaradas (PD), han adquirido gran importancia en áreas relacionadas con la provisión de servicios a la comunidad, tales como transporte, salud y vivienda. Estos modelos plantean un protocolo de decisión para elegir la mejor alternativa disponible para un individuo que se comporta como el “consumidor racional” de la microeconomía clásica (ver Ortúzar, 2000a). Al modelar, se estima la probabilidad de que se escoja una alternativa como función de los atributos del conjunto de opciones disponibles y de sus características socioeconómicas (ingreso, género, edad). Luego, las razones entre los coeficientes de cada atributo (utilidades marginales) y el de la variable costos, se interpretan como valores subjetivos o disposiciones a pagar (Espino *et al*, 2006; Jara-Díaz, 2007; Sillano y Ortúzar, 2005), y tienen gran importancia en evaluación social de proyectos.

Las PD reflejan una intención de comportamiento de los individuos, que se determina a través de la aplicación de una encuesta conformada por un conjunto de situaciones hipotéticas que los entrevistados deben evaluar; en ésta responden cómo actuarían frente a cada situación, si se les presentara en la realidad (Ortúzar, 2000b). Una manera eficiente de desarrollar una encuesta de PD es utilizando un computador y un programa interactivo que permita diseñar un conjunto de situaciones hipotéticas de acuerdo a las características de cada entrevistado. Este sistema fue utilizado para estudiar la valoración subjetiva que hacen distintos sujetos de los atributos de una vivienda mínima (Greene y Ortúzar, 2002).

En esta ocasión, se utilizó la técnica para poder conocer en profundidad las características de los individuos que están comprando departamentos en el centro de la ciudad de Santiago. El sector ha experimentado un acelerado proceso de renovación urbana que incluye la construcción de numerosos edificios de departamentos con

subsidios gubernamentales. Resulta interesante, tanto desde el punto de vista de las inmobiliarias como del Estado propiciador del proceso, investigar cuáles son las características del centro de Santiago que son valoradas por los nuevos propietarios y aquellas que son prescindibles. De esta forma, los recursos pueden ser mejor dirigidos hacia aquellos que interesan a la población, logrando una mejor renovación urbana.

La tesis se encuentra estructurada comenzando en el capítulo dos con una descripción más detallada de los objetivos. En el capítulo tres se explica la fundamentación teórica sobre la que se basa la metodología utilizada en la investigación, y el capítulo cuatro detalla el proceso de diseño del experimento, incluyendo las tres encuestas piloto realizadas. Por último, los capítulos finales se refieren a los resultados obtenidos de la encuesta definitiva. En el capítulo cinco se analizan la base de dato, caracterizando a los individuos que respondieron la encuesta. Los modelos estimados son detallados en el capítulo seis y las disposiciones al pago calculadas para cada atributo se reportan en el capítulo siete. Finalmente, en el capítulo ocho se comentan las principales conclusiones de la investigación.

2. CONTEXTO Y OBJETIVOS DEL ESTUDIO

Los Juegos Olímpicos de 1992 en Barcelona fueron utilizados por el gobierno español para otorgar a la ciudad y el país el carácter de modernos y competitivos en el mundo. Para ello se trabajó en un proyecto de ciudad definido, con estrategias y políticas urbanísticas adecuadas, acompañadas de inversiones importantes en infraestructura vial, deportiva, de espacios públicos y reconstrucción de barrios (Alcalá, 2006; Benach, 1993). Esta utilización del urbanismo como imagen de marca dio paso a una redefinición del concepto de ciudad imperante hasta dicho momento. La ciudad como estructura física pasó de ser el guardián de la memoria colectiva y cuna de la cultura urbana a adquirir valor como motor del desarrollo económico de los países.

En paralelo, ya existía conciencia de los problemas que conlleva la falta de planificación urbana, tales como que las ciudades crezcan de manera periférica, que se concentren espacialmente los distintos grupos socioeconómicos (estratificación) y que se subutilicen las zonas céntricas. La conjugación de estos hechos ha provocado un proceso mundial de renovación urbana que ha tenido como resultado una transformación radical de los espacios urbanos más postergados, en particular los centros históricos.

Como parte de este proceso, y luego de más de cincuenta años de abandono, las zonas centrales de las principales ciudades latinoamericanas se han visto envueltas en diversos proyectos de recuperación acompañados de fuertes inversiones tanto privadas como estatales (Carrión, 2001; Rojas, 2004). Sin embargo, a diferencia de las estrategias y planes exitosos de Barcelona o Bilbao, por ejemplo, en la mayoría de los casos latinoamericanos no se ha contado con un proyecto urbano claro donde se coordine a los diferentes actores del proceso de renovación, ni se ha realizado un seguimiento posterior de los resultados, ni análisis adecuado de los nuevos pasos a seguir. Luego, parece vital definir una metodología que permita promover la recuperación de los centros históricos logrando, de manera simultánea, la renovación y la protección del cuerpo edificado y de

la población del área. Este corresponde al objetivo del proyecto en que se enmarca esta tesis.

Existen dos condiciones necesarias, pero no suficientes, para que se lleven a cabo procesos exitosos de renovación urbana. En primer lugar, el Estado debe ser el iniciador del proceso, asumiendo el costo de intervenciones como mejoramiento de los espacios públicos, infraestructura y sistema de transporte. El *timing* de estos procesos debe ser adecuado en cuanto a la coordinación de sus diferentes etapas, puesto que de otro modo el esfuerzo puede no entregar frutos (Rojas, 1999); por ejemplo, es necesario que la facilitación de subsidios se realice una vez que se comience la construcción de nuevas edificaciones o la renovación de las antiguas, y no antes. La segunda condición corresponde al repoblamiento de la zona, el cual permite la consolidación del proceso de renovación. Los nuevos pobladores deben conformar un grupo mixto, que permita el ingreso de nuevos recursos a la zona y que no desplacen en su totalidad a los antiguos habitantes (Rojas, 1999).

Actualmente el centro de la ciudad de Santiago se enfrenta a un proceso de renovación acelerado, con grandes inversiones para la construcción de nuevos edificios de departamentos y entrega de subsidios habitacionales especiales para el sector. Esto ha provocado una alta migración hacia esta zona histórica de la ciudad. Luego, resulta interesante caracterizar a estos nuevos pobladores y determinar cómo afectarán el futuro de la zona a través de sus preferencias. El sector céntrico de Santiago conjuga características que en general no se encuentran en otras zonas de la ciudad, de modo que parece importante determinar si ellas son valoradas por sus nuevos habitantes.

Así, los objetivos primarios de esta tesis fueron los siguientes:

- Definir, a través de una encuesta de PD, la manera en que son valorados los atributos del centro de Santiago por sus nuevos pobladores.

- Utilizando el mismo instrumento, caracterizar a esta nueva población.

Por otro lado, como objetivos adicionales de carácter metodológico, se encuentran:

- Analizar la metodología de diseños eficientes para encuestas de PD en un contexto donde la mayoría de los atributos corresponden a variables mudas (*dummy*).
- Adaptar la búsqueda de diseños eficientes de encuestas de PD a un caso con dos juegos de preguntas.

Para comprender estos dos objetivos secundarios, resulta necesaria una explicación de la fundamentación teórica sobre la que se basa el experimento, la que se encuentra en el capítulo 3.

3. MARCO TEORICO

3.1. Modelos de Elección Discreta

Esta investigación utilizó como base metodológica los modelos de elección discreta, usados ampliamente en la literatura para estimar la disposición al pago por atributos no transados en el mercado (Boxall *et al.*, 1996; Goett *et al.*, 2000; Greene y Ortúzar, 2002; Hanley *et al.*, 1998; Munizaga *et al.*, 2006; Ortúzar *et al.*, 2000). Básicamente, estos modelos estiman la probabilidad de que un individuo con ciertas características escoja un bien discreto sobre otro; los bienes se describen en términos de los atributos que los constituyen y se supone que la elección es también afectada por las características del consumidor.

La teoría de la utilidad aleatoria (McFadden, 1974), sobre la que se sustentan los modelos de elección discreta, considera, en su versión más sencilla, que los individuos poseen información perfecta y procuran maximizar su utilidad personal; por otro lado, como el modelador es sólo un observador del sistema, desconoce una fracción de la realidad, y esto le impide describir las funciones de utilidad personales en su totalidad. Por ello, las funciones de utilidad se desglosan en dos partes: una determinística o representativa, que es conocida por el modelador, y una componente aleatoria, que representa todos los elementos desconocidos por el modelador, como son gustos particulares, errores de medición, etc. Así, si un individuo q debe escoger entre un conjunto de alternativas disponibles A_q , el modelador describe a la utilidad U_{iq} de la alternativa A_i como:

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

donde V_{iq} es la porción determinística o conocida y ε_{iq} el error aleatorio. La manera más común y sencilla de representar V_{iq} es a través de la siguiente función lineal:

$$(3.2) \quad V_{iq} = \sum_k \theta_{ik} \cdot X_{ikq}$$

Aquí X_{ikq} representa el atributo k-ésimo de la opción A_i para el individuo q , y θ_{ik} es un parámetro a estimar que, en los modelos más sencillos, se asume constante entre los individuos, pero puede variar entre alternativas.

Bajo el supuesto de que los individuos son racionales y, por ende, escogen siempre la alternativa que les entrega mayor utilidad, su elección será la alternativa A_i cuando (Ortúzar, 2000a):

$$(3.3) \quad U_{iq} > U_{jq} \quad \forall A_j \in A_q$$

$$(3.4) \quad \Rightarrow V_{iq} - V_{jq} > \varepsilon_{jq} - \varepsilon_{iq} \quad \forall A_j \in A_q$$

Puesto que el término $\varepsilon_{jq} - \varepsilon_{iq}$ es desconocido para el modelador, no puede determinar con seguridad si es que el individuo escoge una alternativa o no, sino que sólo puede estimar una probabilidad P_{iq} de que ello ocurra:

$$(3.5) \quad P_{iq} = \Pr ob(\varepsilon_{jq} \leq \varepsilon_{iq} + V_{iq} - V_{jq}, \forall A_j \in A_q)$$

Dado que los residuos ε_{iq} son variables aleatorias con algún tipo de distribución $f(\varepsilon) = f(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N)$, se pueden obtener diversos modelos dependiendo de cual sea la distribución escogida. A continuación se presentan brevemente los modelos de elección discreta utilizados en esta investigación.

3.1.1. Modelo Logit Simple (MNL)

Este es el modelo más sencillo y popular en esta materia, y se genera a partir del supuesto de que el componente aleatorio de la función de utilidad, ε_{iq} , distribuye

Gumbel independiente e idéntico (IID) con media cero y varianza σ^2 . La probabilidad de que el individuo q escoja la alternativa A_i está dada por (Ortúzar 2000a):

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

donde el parámetro de escala $\beta = \frac{\pi}{\sigma\sqrt{6}}$ que caracteriza a la distribución Gumbel es no identificable, por lo que debe ser normalizado. Si bien este modelo es sencillo de estimar, tiene varias limitaciones, como no permitir variaciones aleatorias en los gustos, correlación entre las alternativas, ni heterocedasticidad.

El método más común para estimar los parámetros θ_{ik} es máxima verosimilitud, que se basa en la idea de que aun cuando una muestra podría ser generada por varias poblaciones, algunas son más probables que otras; los estimadores máximo verosímiles $\hat{\theta}_{jk}$ corresponden al conjunto de parámetros que generarán más a menudo la muestra observada. En el caso de un modelo con probabilidades de elección P_{iq} , la función de verosimilitud $L(\theta)$ está dada por la expresión (Ortúzar 2000a):

$$(3.7) \quad L(\theta) = \prod_{q=1}^Q \prod_{A_j \in A_q} P_{jq}^{g_{jq}}$$

donde $g_{jq} = \begin{cases} 1, & \text{si } q \text{ elige } A_j \\ 0, & \text{en otro caso} \end{cases}$ y Q es el total de individuos encuestados.

Al aplicar logaritmo se obtiene la función a maximizar (log-verosimilitud), la que se deriva e iguala a cero para obtener los valores de los parámetros:

$$(3.8) \quad \ln(L(\theta)) = l(\theta) = \sum_{q=1}^Q \sum_{A_j \in A_q} g_{jq} \cdot \ln(P_{jq})$$

Los estimadores máximo verosímiles $\hat{\theta}_{jk}$ distribuyen asintóticamente Normal (θ_{jk}, S^2) , esto es, su media son los parámetros poblacionales; su matriz de varianzas - covarianzas S^2 está dada por:

$$(3.9) \quad S^2 = - \left[E \left(\frac{\partial^2 l(\theta)}{\partial \theta^2} \right) \right]^{-1}$$

Esta matriz corresponde a la inversa de la matriz de información, también conocida como matriz de Fisher. A partir de S^2 se pueden obtener los datos para aplicar el test de significancia de un parámetro, utilizado durante la búsqueda de la mejor especificación de los modelos. Este test otorga información sobre la evidencia existente para rechazar o no la hipótesis nula de que un parámetro determinado sea igual a cero. Como s_k (raíz de la diagonal de la matriz S^2) es el error estándar del parámetro $\hat{\theta}_{jk}$ estimado, el test se define a partir de que:

$$(3.10) \quad t_{\alpha} = \frac{\hat{\theta}_{jk} - \theta_{jk}}{s_k} \sim N(0,1)$$

Dado que en este caso la hipótesis nula está dada por $H_0 : \theta_{jk} = 0$, se obtiene $t_{\alpha} = \frac{\hat{\theta}_{jk}}{s_k}$,

y se rechaza la hipótesis nula si $t_{\alpha} \geq t_{crit,\alpha}$, con α típicamente igual a 5%.

Otro test utilizado ampliamente es el de Razón de Verosimilitud (LR), el cual facilita la comparación entre dos modelos, uno de los cuales debe ser una versión restringida del otro. Si $l(\hat{\theta})$ y $l(\hat{\theta}_r)$ son las log-verosimilitudes del modelo general y el restringido

respectivamente, entonces el estadígrafo LR distribuye asintóticamente χ^2 con tantos grados de libertad como número de restricciones lineales tenga el modelo restringido respecto al general:

$$(3.11) \quad LR = -2 \cdot \{l(\hat{\theta}_r) - l(\hat{\theta})\}$$

Se rechaza la hipótesis nula si $LR > \chi^2_{r,\alpha}$, lo que quiere decir que el modelo restringido no sería adecuado.

Para información más detallada sobre la derivación matemática del modelo, sus propiedades y restricciones, y otros tests asociados, revisar Ortúzar (2000a).

3.1.2. Modelo Logit Mixto (ML)

El modelo Logit Mixto (ML) es más general que el MNL, puesto que permite correlación entre alternativas, variación aleatoria en los gustos, heterocedasticidad, etc. Si bien el modelo permite una gran libertad en cuanto a su especificación, dos formulaciones son las más utilizadas (Train, 2003):

- a) Modelo de componentes de error (*error components* EC): en esta especificación la función de utilidad U_{iq} se compone de un elemento determinístico V_{iq} , un componente aleatorio IID $\varepsilon_{iq} \sim \text{Gumbel}(0, \sigma^2)$ y otro término aleatorio η_{iq} con una distribución cualquiera. La función de utilidad toma la siguiente forma:

$$(3.12) \quad U_{iq} = V_{iq} + \eta_{iq} + \varepsilon_{iq} = \sum_k \theta_{ik} \cdot X_{ikq} + \eta_{iq} + \varepsilon_{iq}$$

El componente aleatorio η_{iq} puede representar, entre otras cosas, correlación entre alternativas o entre respuestas de los individuos (efecto panel). Para más detalles de esta formulación ver Munizaga y Alvarez-Daziano (2002).

b) Modelo de parámetros aleatorios (*random parameters* RP): en este caso la especificación es igual a la del modelo MNL, pero los parámetros θ_{jkq} no son fijos, sino que aleatorios con alguna distribución. De esta forma, varían en la población y representan las diferencias en los gustos de los individuos no capturada por la parte determinística de la función de utilidad. Asumiendo entonces parámetros θ_{jkq} con media poblacional $\bar{\theta}_{jk}$ y desviaciones individuales γ_{qj} , la estructura de la función de utilidad resulta (Train, 2003):

$$(3.13) \quad U_{iq} = \sum_k \bar{\theta}_{ik} \cdot X_{ikq} + \sum_k \gamma_{qk} \cdot X_{ikq} + \varepsilon_{iq}$$

Para mayor detalle sobre dicha especificación del modelo, ver Sillano y Ortúzar (2005) o Train (2003).

La estimación de los modelos ML es más compleja que la de los modelos MNL, debido a que dependen de la solución de una serie de integrales multivariadas sin forma cerrada. Luego, el modelo generalmente es estimado mediante máxima verosimilitud simulada o incluso métodos Bayesianos (Godoy, 2004; Train, 2003).

3.2. Disposición al Pago por Atributos

Para derivar la disposición al pago (*willingness to pay* WTP) de la población por los atributos que describen a las alternativas de un experimento, es necesario que uno de los atributos corresponda al precio o costo directo. La disposición al pago corresponde a la tasa marginal de sustitución entre el atributo de interés y el costo (X_{ic}):

$$(3.14) \quad WTP(X_{ik}) = - \frac{dV_i/dX_{ik}}{dV_i/dX_{ic}}$$

Para modelos simples, la WTP se puede derivar de manera sencilla. Por ejemplo, para un modelo MNL lineal en los parámetros resulta:

$$(3.15) \quad WTP(X_{ik}) = -\frac{\theta_{ik}}{\theta_{ic}}$$

donde θ_{ik} y θ_{ic} con los parámetros (utilidades marginales) de los atributos de interés y del costo, respectivamente (Jara-Díaz, 2007). Debido a que la WTP se construye a partir de estimadores θ_{ik} que tienen una distribución asintótica Normal, no es razonable trabajar con la razón entre los estimadores puntuales y es recomendable construir un intervalo de confianza. Entre los dos métodos propuestos en Armstrong *et al.* (2001), con resultados similares, se utilizó el del test-t por su fácil aplicación:

$$(3.16) \quad WTP(X_{ik}) = \left(\frac{\theta_{ik}}{\theta_{ic}} \frac{t_c}{t_k} \right) \frac{(t_c t_k - \rho t_\alpha^2)}{(t_c^2 - t_\alpha^2)} \pm \left(\frac{\theta_{ik}}{\theta_{ic}} \frac{t_c}{t_k} \right) \frac{\sqrt{(\rho t_\alpha^2 - t_c t_k)^2 - (t_k^2 - t_\alpha^2)(t_c^2 - t_\alpha^2)}}{(t_c^2 - t_\alpha^2)}$$

donde t_k y t_c corresponden a los t-estadísticos de θ_{ik} y θ_{ic} , respectivamente; t_α es el valor crítico de t para el nivel de confianza α , y ρ es el coeficiente de correlación entre ambos parámetros estimados.

Finalmente, la estimación de la WTP por atributos en modelos ML es bastante más compleja, debido a que en este caso no sólo los estimadores son variables aleatorias, sino que también lo son los parámetros mismos. En esta tesis se utilizaron tres métodos diferentes explicados en Sillano y Ortúzar (2005):

1. Método 1, radio entre medias: al igual que con los modelos MNL, se dividen las medias de los parámetros y se genera el intervalo de confianza según la Ecuación 3.16. El problema de este método es que no se considera las desviaciones estándar estimadas para los parámetros.
2. Método 2, simulación: se genera una instancia aleatoria del parámetro del costo y del otro atributo a partir de sus distribuciones, y se dividen esos dos valores. Este

procedimiento se repite muchas veces, derivándose luego la media y la desviación estándar del total de valores calculados. No es posible estimar un intervalo de confianza en este caso, dado que éstos se construyen a partir de los errores estándar de las distribuciones normales asintóticas de los estimadores, diferentes de las desviaciones estándar estimadas en este caso, que son indicadores de la varianza de la división entre los parámetros sobre la población simulada. En otras palabras, un intervalo de confianza basado en las desviaciones estándar estimadas indicaría en qué intervalo se encuentra el 95% de las divisiones de parámetros simuladas.

3. Método 3, parámetro del costo fijo: para este caso, al estimar el modelo logit mixto con parámetros aleatorios, el parámetro del costo se considera fijo. De esta forma, la WTP se obtiene al dividir el parámetro del atributo en análisis por el del costo, y tiene la misma distribución del parámetro del primero. Por ejemplo, si el parámetro del atributo tiene distribución normal, la media y la desviación estándar de la WTP serán las del atributo divididas en el parámetro fijo del costo.

3.3. Encuestas de Preferencias Declaradas

Para estimar modelos de elección discreta, es necesario contar con una base de datos correspondiente a elecciones individuales. Una manera común de obtener estos datos es a través de encuestas, ya sea sobre elecciones reales hechas por los encuestados (Preferencias Reveladas, PR) o sobre situaciones ficticias (Preferencias Declaradas, PD). Este último método, ampliamente utilizado en la literatura, también se ha usado en diversas tesis anteriores en el Departamento de Ingeniería de Transporte y Logística (Galilea, 2002; Iglesias, 2007; Perez, 1998; Rizzi, 2001), y es el que se ocupó para esta investigación. El enfoque consiste en presentar al encuestado una serie de situaciones de elección ficticias, donde cada una de las alternativas a elegir es descrita en base a atributos y de la manera más realista posible (Hensher *et al*, 2005).

El enfoque de PD permite solucionar dos importantes inconvenientes del enfoque PR, que son la dificultad para evaluar nuevas alternativas y la incidencia de factores cualitativos difíciles de medir en la práctica, como por ejemplo, la comodidad en un viaje. Otro aspecto positivo de utilizar encuestas de PD es que, a través de un diseño adecuado, es posible aislar atributos secundarios o que no son de interés primario para aquellos que se desea estudiar, evitándose sesgos por falta de información. Además, este tipo de encuestas es considerablemente menos costosa de realizar y se puede obtener mayor información a partir de la misma cantidad de encuestados, puesto que es posible plantear muchas situaciones de elección ficticias a cada individuo. Este punto resulta especialmente relevante al modelar, puesto que se tendrán más observaciones que individuos, y por lo tanto si se utiliza el modelo MNL, se supondrá una independencia no real entre las observaciones de cada individuo. Afortunadamente, el modelo ML permite atacar este problema incluyendo un componente de error que correlacione las respuestas de cada encuestado.

Por otro lado, entre las limitaciones de las encuestas de PD se encuentran el no saber si el encuestado se comportará en la realidad de la misma manera cómo responde y la poca confiabilidad de los datos para hacer predicciones. Su mayor utilización radica en obtener la disposición al pago por los atributos de las alternativas.

Finalmente, cabe destacar que existen tres tipos de encuestas de PD¹:

- a) Jerarquización o *ranking*: el individuo debe ordenar una serie de situaciones hipotéticas, entregando su elección de manera indirecta.

¹ Para mayor detalle sobre las características y modelación de cada tipo ver Ortúzar (2000a).

- b) Escalamiento o *rating*: el encuestado debe clasificar las alternativas dentro de una escala semántica.
- c) Elección: la persona escoge entre las alternativas de manera directa. Este fue el tipo utilizado en la tesis.

3.4. Diseño de Experimentos de Elección

3.4.1. Fundamentación Teórica

El experimento de elección corresponde a la serie de elecciones hipotéticas que se presenta al encuestado, incluyendo los niveles de los atributos de cada alternativa en cada situación de elección. Un diseño adecuado resulta relevante puesto que de él depende la calidad de los datos con que se modelará posteriormente. Si bien la confiabilidad de los resultados depende de la cantidad de encuestados y de las respuestas entregadas por cada uno, también se ve fuertemente afectada por las situaciones planteadas, el tipo de modelo de elección utilizado y los supuestos sobre los que éste se basa (Cherchi y Ortúzar, 2008). En definitiva, el diseño del experimento debe entregar datos lo más informativos posibles.

Hasta fines de los '90, el diseño de experimentos de preferencias declaradas se realizó usando técnicas basadas en principios que aplican a modelos lineales, siendo los diseños factoriales ortogonales los más comunes (Louviere *et al.*, 2000). Sin embargo, la utilización de este tipo de diseño puede llevar a resultados ineficientes, puesto que el interés se enfoca en obtener diseños con características que no son relevantes para modelos de elección (por ejemplo, ortogonalidad). Una explicación detallada de los problemas generados por utilizar estas técnicas se encuentra en Iglesias y Ortúzar (2008).

Otra opción es utilizar diseños eficientes, definidos como aquellos que dan como resultado la estimación de modelos con parámetros que tienen bajos errores estándar. Recordemos que estos errores se pueden derivar de la matriz asintótica de varianza - covarianza de los parámetros, S^2 (ecuación 3.9). Se han postulado diferentes medidas de eficiencia obtenidas a partir de S^2 , como el Error-D y el Error-A, entre otros (Sandor y Wedel, 2001). El más común, y que fue el utilizado en esta investigación, corresponde al Error-D, el cual se calcula como la raíz k-ésima del determinante de la matriz S^2 , donde k corresponde al número de parámetros.

Desgraciadamente, para calcular la matriz S^2 , y con ello el Error-D, es necesario conocer previamente los valores de los parámetros del modelo a estimar. Luego, para obtener la medida de eficiencia se debe, en primer lugar, asumir la estructura de las funciones de utilidad y, por otro lado, suponer valores iniciales de los parámetros (parámetros *a priori*). En cuanto a estos últimos se pueden presentar tres casos, y existen recomendaciones sobre cómo atacar el problema en cada uno (Bliemer y Rose, 2006):

- a) No hay información sobre los valores de los parámetros. En este caso es posible fijarlos iguales a cero.
- b) Hay información bastante precisa sobre los valores de los parámetros. Se asumen esos valores como *priors* fijos.
- c) Hay información sobre los valores de los parámetros, pero no es precisa o existe incertidumbre sobre su validez. En lugar de asumir valores fijos de los parámetros, se utiliza una aproximación Bayesiana considerando una determinada distribución de probabilidad para cada parámetro.

En la práctica, dependiendo de los algoritmos y restricciones que se consideren, resulta muy complejo encontrar el diseño con menor Error-D, llamado diseño D-Optimo, por lo

que se utiliza un diseño con Error-D lo suficientemente bajo, llamado diseño D-Eficiente (Bliemer y Rose, 2006).

3.4.2. Metodología Utilizada

En todos los casos se utilizó una aproximación Bayesiana para los parámetros en lugar de asumirlos fijos.

La heurística ocupada en esta tesis para buscar un diseño eficiente fue la siguiente:

1. Diseño Inicial - Encuesta Piloto 1: se buscaron valores para los parámetros *a priori* en la literatura y se realizaron supuestos para aquellos valores no encontrados. A partir de ellos, se buscó un diseño D-Eficiente, es decir, que tuviera un Error-D bajo, con el cual se generó el primer experimento.
2. Encuesta Piloto 2: con las respuestas obtenidas de la encuesta anterior se estimaron modelos, los cuales permitieron obtener nuevos parámetros *a priori*. Se generó una nueva búsqueda de diseño D-Eficiente, y con ello se obtuvo otro experimento de elección.
3. Encuesta Piloto 3: se estimaron nuevamente modelos, esta vez a partir de las respuestas obtenidas en la segunda encuesta piloto. Con ello, y a partir de nuevos valores iniciales de los parámetros, se buscó un diseño D-Eficiente final con el que se generó el experimento de la encuesta definitiva.

En cada uno de los tres pasos fue necesario encontrar un diseño D-Eficiente a través de la minimización de la medida de eficiencia Error-D. Para esto se utilizó una mecánica iterativa, basada en una Macro en Visual Basic de un documento Microsoft Excel. Además, se optó por utilizar una aproximación Bayesiana de los parámetros en lugar de asumirlos fijos, puesto que la información sobre sus valores no era confiable. La forma de calcular el Error-D para cada diseño fue:

- a. Se define la cantidad de diseños que se evaluarán, N , y el número de instancias que se generarán de cada parámetro, I .
- b. Se genera aleatoriamente un diseño.
 - b1. A través de números cuasi-aleatorios (secuencia Halton), se genera una instancia de cada parámetro aleatorio del modelo.
 - b2. Se evalúa y almacena el Error-D del diseño con dichos valores de los parámetros.
 - b3. Los pasos b1. y b2. se repiten I veces, es decir, se almacenan I valores para el Error-D del diseño.
 - b4. Se obtiene el promedio entre los I Error-D. Este corresponde al Error-D del diseño generado en b.
- c. El paso b. (y todos sus pasos intermedios) se repite N veces, almacenándose cada diseño y su Error-D respectivo.
- d. Se escoge entre los N diseños cuál es el que tiene menor Error-D.

El último aspecto relevante de mencionar corresponde a la aplicación de esta metodología a un experimento con dos juegos y una variable común, como se utilizó en esta tesis. La construcción de la matriz de Fisher, a partir de la que se obtiene la matriz S^2 y el Error-D, se basa en cálculos sobre los niveles de cada atributo, la función de utilidad y la probabilidad de elección de cada alternativa en cada situación de elección. Luego, todos los atributos del Juego 2 toman nivel igual a cero en las situaciones de elección del Juego 1 y viceversa. Sólo el atributo común en los dos juegos tendrá valores diferentes de cero en ambos. Debido a ello, la matriz de Fisher cuenta con dos bloques de valores nulos en cuadrantes opuestos, como se muestra en la Figura 3-1.

	Atributos Juego 1				At. Común	Atributos Juego 2		
Atributos Juego 1	x	x	x	x	x	0	0	0
	x	x	x	x	x	0	0	0
	x	x	x	x	x	0	0	0
	x	x	x	x	x	0	0	0
At. Común	x	x	x	x	x	x	x	x
Atributos Juego 2	0	0	0	0	x	x	x	x
	0	0	0	0	x	x	x	x
	0	0	0	0	x	x	x	x

Figura 3-1: Estructura de la Matriz de Fisher para Experimento con Dos Juegos

Eventualmente, dependiendo de los valores que tomen las celdas distintas de cero, podría haber problemas de invertibilidad con la matriz de Fisher, en cuyo caso no se podría calcular la matriz de varianzas-covarianzas S^2 .

3.5. Tamaño de Muestra

Comúnmente, se utilizan técnicas de muestreo probabilísticas para analizar el tamaño de la base de datos requerido para realizar una encuesta (ver distintos métodos en Henry 1990).

Sin embargo, estos métodos no son adecuados para el diseño de encuestas de preferencias declaradas, porque asumen conocimiento previo sobre las proporciones de elección, información que no está disponible en este tipo de experimentos (en especial si se trabaja con alternativas genéricas). Una discusión más detallada sobre este tema se puede encontrar en Bliemer y Rose (2005). En ese artículo, los autores proponen otra manera de encontrar un límite inferior teórico para el tamaño de la muestra, a partir de los valores iniciales de los parámetros con los que se diseña el experimento:

$$(3.17) \quad N = \left(\frac{t_{\alpha} \cdot se_k(X, \hat{\theta})}{\hat{\theta}_{jk}} \right)^2$$

Este método permite obtener un límite inferior para cada parámetro, por lo que la recomendación es calcular la muestra requerida por cada parámetro y usar el mayor de esos valores como límite inferior del tamaño que debe tener la base de datos de la encuesta.

4. DISEÑO DEL EXPERIMENTO

El diseño del experimento fue la fase de mayor longitud en la investigación. Se requirió de diversas etapas para lograr un diseño adecuado, considerando la correcta especificación de las variables, la eficiencia del diseño y su implementación tanto en Microsoft Excel, al comienzo, como vía página *web* al final.

4.1. Características Generales de la Encuesta

En primer lugar, se definió que la encuesta debía realizarse a personas con intención de comprar un departamento en la comuna de Santiago, las cuales estarían familiarizadas en alguna medida con el proceso de escoger una vivienda y, seguramente, ya habrían reflexionado sobre qué aspectos les parecían importantes para su compra. Se optó por no encuestar a personas que ya hubieran comprado departamento, porque éstas podrían, involuntariamente, responder la encuesta reafirmando la elección ya hecha o recordando el proceso de elección de manera que éste se ajustara a sus preferencias actuales (Buckingham y Saunders, 2004). Se trataba que los individuos escogidos no presentaran apego a su antigua vivienda (por estar buscando una nueva) ni a la nueva (por no haber realizado aun la compra) evitándose sesgos.

En una primera etapa, se deseaba las preguntas iniciales de la encuesta correspondieran a información socioeconómica y probablemente a características tanto de la vivienda actual como del departamento deseado. De hecho, se consideró la opción de desarrollar el experimento de preferencias declaradas haciendo que las personas escogieran entre su vivienda actual y un departamento presentado en la encuesta. Sin embargo, esto presentaba el inconveniente de consultar a las personas mucha información de sus viviendas actuales e implicaba dificultades en el diseño del experimento (por ejemplo, complicaciones para encontrar parámetros previos o para definir funciones de utilidad para el diseño del experimento), por lo que se descartó esta opción.

Finalmente, se optó por realizar un experimento donde los encuestados escogieran entre dos departamentos ficticios, definidos de acuerdo a variables controladas por el modelador. De esta manera, se podría definir mejor las funciones de utilidad, por saber qué variables las componen y sus niveles. Además, sería menos complejo buscar los valores iniciales de los parámetros del experimento.

4.2. Selección de las Variables del Experimento

Como parte del Proyecto FONDECYT en que se inserta esta tesis, los alumnos de la Facultad de Arquitectura, Diseño y Estudios Urbanos de esta Universidad desarrollaron una encuesta que fue aplicada a 166 habitantes de la comuna de Santiago, en los barrios República, Brasil y Lira. Del total de personas encuestadas, 123 habían cambiado de residencia en los últimos 10 años, y se les preguntó sobre la incidencia de determinados factores en dicho cambio (ver ANEXO A).

A partir de las respuestas obtenidas, se generó un ranking de los factores, para definir cuáles serían incorporados en el modelo. Tomando en consideración tanto las respuestas de las personas como la factibilidad de incluir ciertos factores en la encuesta y los objetivos de la investigación, se definió la lista de atributos potenciales de la encuesta, que se listan en la Tabla 4-1. Estos son atributos asociados a los factores más relevantes de la encuesta comentada, agrupando aquellos que son similares o están relacionados.

A continuación se procedió a definir la manera en que estas variables serían incluidas en el experimento. Para cada una de ellas existían diferentes opciones, que se evaluaron a través de entrevistas en profundidad y de encuestas piloto.

Para el tamaño de la vivienda (*Superficie*) se consideró necesario consultar al encuestado la cantidad de dormitorios que prefería para su nuevo departamento y definir los niveles de la variable de acuerdo a esa información. No se sabía cuál era la mejor manera de definir la variable y las opciones que se evaluaron fueron:

Tabla 4-1: Lista Preliminar de Atributos, Encuesta Facultad de Arquitectura

Atributo	Simbología
Tamaño de la vivienda.	<i>Superficie</i>
Precio de arriendo o venta.	<i>Precio</i>
Existencia de espacios comunes en el edificio, como piscina y gimnasio.	<i>Espacio_Común</i>
Limpieza y estado de calles y veredas.	<i>Limpieza_Calle</i>
Accesibilidad y locomoción.	<i>Metro</i>
Existencia o cercanía de comercio-almacenes-supermercados/ de servicios de tramite en general: bancos, cuentas, etc.	<i>Servicios</i>
Barrio interesante y diverso/ Actividad cultural/ Existencia o cercanía de bares, restaurantes, cines, cafés, etc. Esparcimiento en general.	<i>Cultura</i>
Existencia o cercanía de áreas verdes, parques y plazas/ de espacios deportivos y de recreación.	<i>Areas_Verdes</i>

- Usar la cantidad de dormitorios deseada como unidad de medida, y luego utilizar pivotes en torno a este valor. Por ejemplo, si la persona dice que desea dos dormitorios, los niveles del atributos sería uno, dos o tres dormitorios. El problema de usar esta especificación era que se sospechaba que los encuestados se encontrarían cautivos a la cantidad de dormitorios que requerían, por lo que no optarían por un departamento con una habitación menos o una más.

- Utilizar m² como unidad de medida. En este caso se fijaría un promedio de mercado para cada cantidad de dormitorios y se utilizarían dos pivotes en torno a dicho valor (uno superior y otro inferior). Por ejemplo, si el promedio de mercado para departamentos de dos habitaciones es de 40 m², los niveles presentados a personas que desean esa cantidad de habitaciones serían 35, 40 y 45 m². El problema aquí era que no se tenía certeza si los encuestados comprendían el concepto de m² a cabalidad y si serían capaces de captar las diferencias entre los distintos niveles del atributo.

Respecto al precio (*Precio*), existía la posibilidad de incluir el precio de venta (o de arriendo), ya sea en UF o en pesos chilenos. En ambos casos se optaría por estimar un promedio de mercado para la cantidad de dormitorios requeridos por la persona, y se utilizaría un pivote superior y uno inferior. Además, no se tenía claridad sobre si se debía incluir una estimación de los gastos comunes mensuales o no.

Para la existencia de espacios comunes (*Espacio_Común*) en el edificio se tenía la opción de incluir una imagen o sólo una descripción escrita. Esto dependía de la manera en que se diseñaran las otras variables, puesto que varias de ellas podían ser incluidas a través de imágenes en el experimento, pero esto podía conducir a un experimento muy complejo de entender. Por otro lado, no se tenía claridad sobre cuáles espacios comunes era más pertinente incluir, pero de todas formas se definió que la variable debía tener sólo dos niveles: ausencia y presencia de espacios comunes.

Se consideró que la mejor opción para incluir el atributo limpieza y estado de calles y veredas (*Limpieza_calle*) en el experimento, sería a través de imágenes. Es importante mencionar que si bien esta variable representa la labor del municipio, es un atributo del departamento; por lo tanto, resultaba vital que la imagen mostrara la situación real de la comuna de Santiago en este sentido, evitando ideas preconcebidas por parte de los nuevos habitantes.

Por otro lado, se decidió representar a la variable *Metro* a través de la distancia a una estación de metro. Esto se debió principalmente a que al momento de diseñar la encuesta el sistema de transporte Transantiago había sido lanzado recientemente, por lo que la percepción de las micros (buses urbanos) era muy negativa dentro de la población. Así, si bien toda la comuna de Santiago se puede considerar conectada mediante micros, no está completamente cubierta por la malla de metro, por lo que podría ser una variable de decisión. Esta podría encontrarse cerca o lejos del departamento, y una tarea adicional consistió en definir la medida de cercanía (metros, cuadras, minutos, etc.).

Las tres últimas variables (*Servicios*, *Cultura* y *Areas_Verdes*) corresponden a atributos del barrio donde se encuentra el departamento. Se postuló que todas deberían tener dos niveles (presencia y ausencia) y ser representadas a través de imágenes, con una definición previa. Estas tres variables resultan de especial interés, porque son justamente aquellas que hacen de Santiago una comuna distintiva dentro de la ciudad. Por ejemplo, es la segunda comuna del país (y primera de la capital) con mayor superficie de áreas verdes con mantenimiento por habitante².

4.3. Entrevistas en Profundidad

Se decidió realizar entrevistas en profundidad previas a las encuestas pilotos con el objeto de tener mayor claridad sobre cómo definir cada atributo en la encuesta. Dado que el objetivo de las encuestas era examinar el manejo de conceptos por parte de los compradores de viviendas, se decidió segmentar de acuerdo a dos dimensiones: edad y nivel educacional alcanzado (Tabla 4-2).

Tabla 4-2: Cantidad de Personas que Respondieron Entrevistas en Profundidad

	Jóvenes 18 – 35 años	Adultos 36 – 55 años	Adultos Mayores 56 años o más
Enseñanza Media o niveles inferiores	1	2	1
Enseñanza Universitaria completa e incompleta	2	2	2
Total	3	4	3

² Fuente: página web www.observatoriourbano.cl, el día 14 de Julio de 2008.

La edad puede constituir una distinción importante por varios motivos. Entre las distintas generaciones no sólo hay diferencias en términos de gustos y uso del lenguaje, sino que además existe una experiencia previa muy distinta en cuanto a buscar una vivienda; esto puede determinar manejos conceptuales diferentes en esta área, además de las obvias diferencias de inquietudes. El nivel educacional aparece como otro elemento interesante de observar, ya que es un indicador indirecto de la posesión de distintos niveles de capital cultural, social y económico.

Las entrevistas arrojaron luz sobre varias interrogantes tanto del comportamiento y gustos, como de la comprensión de conceptos por parte de los individuos. En primer lugar, existe consenso en que buscar departamento es un trámite que debe realizarse con tiempo y calma, e incluso fue definido como un “paseo”. Ello permitió que se decidiera realizar las encuestas piloto en las mismas salas de ventas, puesto que la actitud relajada de las personas permitiría una alta tasa de respuestas.

Los criterios de elección de departamentos se centran preferentemente en la ubicación del edificio, el espacio del departamento y su valor. Esto permitió confirmar que la elección de atributos parecía adecuada. Por otro lado, se determinó que para la mayoría de los entrevistados el proceso de selección de la vivienda se desarrollaba bajo el siguiente esquema: primero se elige el barrio o algún sector más amplio; luego, teniendo esta localización espacial, se procede a evaluar las características de las viviendas disponibles y se elige la más conveniente. Esta información fue vital en el desarrollo de la encuesta, puesto que se tenía una disyuntiva relativa al número de atributos del experimento: por un lado, se pensaba que la cantidad de atributos podía ser muy alta, de acuerdo a las recomendaciones de Caussade *et al* (2005); pero por otro lado, no se quería quitar realismo a un proceso de elección de por sí complejo y en donde siempre se encuentran involucradas muchas variables. De hecho, se analizaron 90 avisos publicitarios en revistas relativos a ventas de departamentos y en promedio contaban con

seis atributos o características, por lo que los ocho considerados en el experimento parecían ser realistas, pero simultáneamente complejos de analizar.

En vista de la evidencia que el proceso se realizaba de manera secuencial, se optó por diseñar un experimento con dos juegos. De esta forma, los encuestados se verían enfrentados, en primer lugar, a un conjunto de situaciones de elección con atributos relativos al barrio (*Metro, Servicios, Cultura, Areas_Verdes*), y luego se les solicitaría escoger entre alternativas con los atributos restantes (*Superficie, Espacio_Común, Limpieza_Calle*). El *Precio* sería el único atributo presente en todas las situaciones de elección.

Para la mayoría de las personas, el principal criterio para evaluar el tamaño de las viviendas resultó ser el número de habitaciones del departamento y la amplitud de ellas. El concepto de “amplitud” fue fundamental y transversal para todos, pero sin embargo cada uno lo interpretó a su manera, de forma subjetiva y personal. Por otro lado, se constató que los m² no corresponden a una medida entendida ni utilizada por las personas entrevistadas; no parece existir una noción clara de cuánto equivale en espacio físico una cantidad determinada de m². Por lo mismo, los entrevistados manifestaron que los m² no les entregan información que facilite y aporte en su proceso de búsqueda de nueva vivienda. Este hecho determinó que se optase por utilizar una variable de tipo cualitativo, en lugar de cuantitativo, para la variable *Superficie*, decidiéndose que hiciera referencia a dormitorios “amplios” versus dormitorios de tamaño normal.

Respecto al precio de la vivienda, los entrevistados indicaron que les era indiferente si se les entregaba el valor del dividendo en pesos chilenos o en UF. Sin embargo, declararon que transforman el valor a pesos chilenos y a un monto mensual, para poder comprender de manera más fácil cómo se vería afectado su presupuesto familiar. Debido a esto, se optó justamente por incluir la variable *Precio* como dividendo mensual en pesos chilenos, con tres niveles que difieren según la cantidad de habitaciones que desea el

encuestado. Además, se decidió no incluir los gastos comunes, debido a que las personas reportaron que generalmente los datos entregados por las inmobiliarias les resultaban poco creíbles y muy variables, lo que llevaba a que no los tomaran en cuenta al elegir qué comprar.

Las respuestas que se obtuvieron respecto a los espacios comunes en el edificio fueron muy variadas, y no presentaron una tendencia clara. Si bien, resultaban atractivos para algunos entrevistados, otros indicaron que sólo implicaban mayores gastos comunes y que les darían poco uso. De todas formas, mencionaron que los que mayormente se ofrecían eran piscina y gimnasio.

Se consultó luego a los encuestados sobre los servicios que sería conveniente tener en el barrio de sus departamentos:

- a) Servicios Comerciales: el más importante resultó tener cerca locales que satisfagan las necesidades diarias (almacén, panadería verdulería). Algunos hicieron mención a supermercados, pero todos dieron a entender que era más importante el acceso a artículos de necesidades básicas que el tamaño del establecimiento. Otros servicios mencionados, por todos los segmentos, fueron la presencia de un cajero automático o banco y de algún lugar donde pagar cuentas (Servipag). Se puso además especial interés en la disponibilidad de locomoción para poder acceder a otros servicios que estén fuera del barrio.
- b) Servicios Culturales: el más mencionado fue el cine, seguido del teatro. Sin embargo, la mayoría de quienes identificaron estos servicios los mencionaron como algo “ideal”, pero que no determinaría la elección de su vivienda. Al respecto, se comentó que era más importante la existencia de locomoción para poder acceder a estos servicios e otros lugares.

- c) Áreas Verdes: varios entrevistados consideraron a parques y plazas como servicios culturales. En todos los segmentos se mencionó que en la medida que existan estos espacios de reunión, de todas formas se van a generar actividades culturales. Además, todos mencionaron interés en la presencia de áreas verdes cerca de sus viviendas.
- d) Espacios Deportivos: las personas no fueron capaces de identificar espacios deportivos en la comuna de Santiago, por lo que no se consideró realista incluirlos junto a las áreas verdes, como se había pensado en un principio.

Se consultó también a los entrevistados sobre cuáles medidas utilizaban para medir distancia, constatándose que el número de cuadras es la más común. Nadie indicó utilizar metros y un par de personas dijeron que cuadras o minutos. Respecto a cuál sería el límite dentro del cuál encontrarían que un servicio se encuentra cerca de su vivienda, se entregaron diversas respuestas, con un límite inferior de seis cuadras. Con esto, se decidió que cinco cuadras sería el límite para definir servicios en el barrio del departamento en la encuesta.

Algo importante de mencionar es que las entrevistas en profundidad se realizaron de manera segmentada con el fin de detectar diferencias significativas entre los distintos grupos; como no se encontró diferencias importantes, se optó por realizar un solo tipo de encuesta PD, en lugar de variaciones por segmento. Tampoco se presentaron distintas interfaces según segmento debido a los resultados reportados por Stern *et al* (2007), quienes indican que el diseño y presentación de las encuestas afecta de manera similar a distintos grupos demográficos.

4.4. Diseño Preliminar Encuesta

Para elaborar la encuesta se debió trabajar varios aspectos: el diseño de la interfaz, la definición de las preguntas relativas a recopilar información socioeconómica y la

generación del experimento de preferencias declaradas propiamente tal. A continuación se describe cómo se llevaron a cabo estos procesos.

4.4.1. Generación del Experimento de Preferencias Declaradas

En primer lugar, se estableció que se utilizaría un método de elección binario entre dos pares de departamentos genéricos descritos a partir de los atributos ya señalados. Se optó por este método, en desmedro de jerarquización o escalamiento, principalmente debido a la cantidad de atributos del experimento y a la complejidad que de por sí tiene asociada la elección de una vivienda. Ello implicaría diseños de encuesta de mucha longitud para experimentos de jerarquización o escalamiento, los cuales además resultan menos intuitivos de responder que una elección (Ortúzar y Garrido, 1991).

A continuación fue necesario definir la función de utilidad representativa básica del modelo MNL a partir del cual se buscaría un diseño eficiente para la encuesta. Si bien al modelar se estimaría también un modelo ML, como el esfuerzo computacional requerido para diseñar un experimento de acuerdo a ese tipo de modelos es altísimo y la pérdida de eficiencia no es considerable (Bliemer y Rose, 2008), se optó por buscar la eficiencia a través de un modelo MNL. Dado que en este experimento las elecciones se realizan entre pares de departamentos genéricos, es decir, no se cuenta con alternativas rotuladas, la función de utilidad tiene la misma forma para ambas alternativas, no hay constante específica y los parámetros son genéricos. De esta forma, la función de utilidad sistemática se especificó con la estructura descrita en (4.1), y la explicación de todos los atributos se muestra en la Tabla 4-3:

$$\begin{aligned}
 (4.1) \quad V = & \theta_{\text{precio}} \cdot \text{Precio} + \theta_{\text{superficie}} \cdot \text{Superficie} + \theta_{\text{limpieza_calle}} \cdot \text{Limpieza_Calle} \\
 & + \theta_{\text{espacio_comun}} \cdot \text{Espacio_Comun} + \theta_{\text{metro}} \cdot \text{Metro} \\
 & + \theta_{\text{servicios}} \cdot \text{Servicios} + \theta_{\text{cultura}} \cdot \text{Cultura} + \theta_{\text{areas_verdes}} \cdot \text{Areas_Verdes}
 \end{aligned}$$

Tabla 4-3: Explicación de Atributos del Experimento

Atributo	Significado	Signo Esperado Parámetro
<i>Precio</i>	Precio. Igual al dividendo mensual por el departamento, en miles de pesos chilenos. Variable continua.	Negativo
<i>Superficie</i>	Superficie. Variable muda que vale uno si los dormitorios son amplios, y cero si tienen tamaño normal.	Positivo
<i>Limpieza_Calle</i>	Limpieza y estado de calles. Variable muda que vale uno si las calles se encuentran en mal estado y con suciedad, y cero si se encuentran en estado normal.	Negativo
<i>Espacio_Común</i>	Espacios comunes. Variable muda que vale uno si existen espacios comunes en el edificio y cero si no.	No se sabe
<i>Metro</i>	Estación de metro. Variable muda que vale uno si existe una estación de metro en el barrio y cero si no.	Positivo
<i>Servicios</i>	Servicios comerciales. Variable muda que vale uno si existen servicios comerciales en el barrio y cero si no.	Positivo
<i>Cultura</i>	Servicios culturales. Variable muda que vale uno si existen servicios culturales en el barrio y cero si no.	No se sabe
<i>Areas_Verdes</i>	Áreas verdes. Variable muda que vale uno si existen áreas verdes en el barrio y cero si no.	Positivo

El experimento cuenta entonces con ocho atributos, siete de los cuales corresponden a variables mudas (con dos niveles) y una variable continua (*Precio*) que puede tomar tres valores. Se decidió presentar 12 situaciones de elección a los encuestados; esta cantidad fue definida según dos criterios (Bliemer y Rose, 2006):

- a) Balance entre los niveles de los atributos, por lo que se deben utilizar sólo múltiplos de seis (3×2).

- b) El número de situaciones de elección no debe ser inferior a los grados de libertad, que en este caso corresponden a 9 ($= 8 + 1$), ya que se está utilizando una función de utilidad lineal sólo con efectos principales (Rose y Bliemer, 2004, Tabla 1).

Otorgar valores iniciales a los parámetros del modelo no resultó trivial. De todas formas, como punto de inicio se contaba con los signos que ellos debían tener (ver Tabla 4-3), lo que ya implica mejores diseños que fijarlos todos en cero (Bliemer y Rose, 2006). Se utilizó aproximación Bayesiana de los parámetros por no tener éstos valores confiables (ver sección 3.4.2); se asumió que los parámetros derivados de referencias en la literatura tendrían una distribución de probabilidad Normal, mientras que los parámetros sin base en la literatura tendrían distribución de probabilidad Uniforme, debido a que la información sobre ellos debiera ser menos confiable y por lo tanto no habría fundamentos para considerar que su distribución se encuentra concentrada en una determinada media. En la Tabla 4-4 se muestran los parámetros iniciales utilizados en el experimento con sus respectivas fuentes y distribuciones de probabilidad supuestas. Específicamente, la metodología utilizada para obtener los valores que se muestran en la Tabla 4-4 fue la siguiente:

- Medias (distribución Normal): se utilizaron valores de distintas fuentes, y la única manera de parametrizarlos fue usar disposiciones al pago asumiendo parámetros fijos (ver Sección 3.2). Luego, se buscó, primero en la literatura, un valor para el parámetro de *Precio* y valores para las disposiciones al pago por el resto de los atributos. Sin embargo, se encontró sólo valores para *Limpieza_Calle* y *Metro* (Iglesias, 2007; Varela, 1997). A partir de las disposiciones al pago y el parámetro del *Precio*, se derivaron las medias de los atributos. Por ejemplo, si se encontró que la disposición al pago por un atributo es \$20.000 y el parámetro del *Precio* es -0,0299; luego la media del parámetro de ese atributo será $\left(\frac{20.000}{1.000} \right) \cdot 0,0299 = 0,598$.

Tabla 4-4: Primer Valor de Parámetros Previos por Atributo

Atributo	Media	Desv. Std.	Distrib.	Fuente
<i>Precio</i>	-0,0299	0,0090	Normal	Rodríguez (2000) ³
<i>Limpieza_Calle</i>	-0,1153	0,0355		Iglesias (2007) ⁴
<i>Metro</i>	0,2640	0,0680		Varela (1997) ⁵
Atributo	Lim. Inf	Lim. Sup.	Distrib.	Fuente
<i>Cultura</i>	0	0,2988	Uniforme	-
<i>Espacio_Común</i>	0	0,8964		-
<i>Superficie</i>	0,0896	0,8964		-
<i>Servicios</i>	0,0598	0,5976		-
<i>Areas_Verdes</i>	0,0598	0,4482		-

³ La variable se encontraba definida como arriendo mensual en miles de pesos, aplicada a arrendatarios de la Región Metropolitana, y tenía tres niveles.

⁴ Este atributo, representado con una imagen, se refiere al estado de la calle, veredas, limpieza y cuidado de áreas verdes. Tiene dos variables mudas, una relativa a cambiar de un nivel malo a uno bueno y otra donde se pasa de nivel normal a bueno. Se utilizó la correspondiente a pasar de estado normal a bueno por tener mayor relación con la variable de este experimento.

⁵ La variable se encontraba definida como cuerdas al metro o paradero, y tenía signo negativo (más cuerdas = menos utilidad). Puesto que en este caso la distancia es fija e igual a cinco cuerdas, y representa estar cerca del metro, se tuvo que adaptar la disposición al pago a ello. Se calculó el valor subjetivo de estar “cerca” (cinco cuerdas) y de estar “lejos” (diez cuerdas). La diferencia entre ambos se consideró como la disposición al pago por pasar de lejos a cerca. Este fue el valor utilizado para calcular el parámetro del atributo.

- Desviaciones estándar (distribución Normal): la manera de escoger las desviaciones estándar fue asegurando que para 3000 instancias generadas de cada parámetro se cumpliera su respectiva restricción de signo.
- Límite superior e inferior (distribución Uniforme): estos valores se obtuvieron en base a suponer la disposición al pago de cada atributo, considerando un límite inferior y uno superior. Luego, también ocupando el parámetro del *Precio*, se derivaron los parámetros mínimo y máximo de cada distribución. Las disposiciones al pago mínima y máxima que se asumieron por atributo fueron: *Cultura* (\$0 y \$10.000); *Espacio_Común* (\$0 y \$30.000); *Superficie* (\$3.000 y \$30.000); *Servicios* (\$2.000 y \$20.000); *Areas_Verdes* (\$2.000 y \$15.000). La realización de las sucesivas encuestas piloto fue necesaria especialmente debido a que estos valores no tienen fundamentación teórica.

Otro aspecto considerado en la búsqueda de un diseño eficiente, fue el hecho que el atributo *Precio* correspondía a una variable de tipo pivote, es decir, no tomaba el mismo valor para todos los encuestados, sino que se generaba tomando como referencia una respuesta anterior de los mismos. En este caso, se optó por proponer distintos dividendos, de acuerdo a la cantidad de habitaciones que deseara el individuo, para otorgarle realismo al experimento y lograr que los encuestados relacionaran los valores con aquellos que ya habían visto en su proceso de búsqueda de vivienda. Los tres niveles correspondieron a la media de mercado del dividendo según cantidad de habitaciones, un valor 15% menor y otro 15% mayor (denominados niveles Bajo, Medio y Alto).

Los valores de mercado se obtuvieron del sitio web www.portalinmobiliario.com en Diciembre de 2007. Se buscaron los precios de todos los departamentos en venta en ese momento en la comuna de Santiago y se cotizó el dividendo a 20 años con un 20% de pie, tal como sería presentado posteriormente en la encuesta, por ser la opción de compra más común. Luego se promediaron los valores, clasificándolos según cantidad

de dormitorios. En la Tabla 4-5 se observan los niveles del atributo utilizados finalmente en la encuesta.

Tabla 4-5: Niveles Atributo *Superficie*

Dormitorios	Nivel Inferior	Nivel Medio	Nivel Superior
Ambiente Único	\$93.500	\$110.000	\$126.500
1	\$110.500	\$130.000	\$149.500
2	\$157.250	\$185.000	\$212.750
3	\$199.750	\$253.000	\$270.250
4	\$269.450	\$317.000	\$364.550

En Rose *et al* (2008) se analizan distintos métodos para generar diseños eficientes con atributos en pivotes; esto es relevante, puesto que las respuestas entregadas por los encuestados afectan la eficiencia del diseño y no se conocen *a priori*. Los autores trabajaron a partir de cuatro estrategias diferentes, obteniendo modelos fuertemente más eficientes que con diseños ortogonales en todos los casos; en particular, los dos métodos que entregaron mayor eficiencia fueron:

- Generar un diseño eficiente en el momento para cada encuestado. Para ello, la persona responde preguntas iniciales que determinan el experimento, se busca *in situ* un diseño eficiente, e inmediatamente se le presentan las situaciones de elección generadas con el mejor diseño encontrado. Esto requiere de muchos recursos computacionales para que ocurra de manera rápida, por lo que no pudo ser llevado a la práctica.
- Desarrollar un proceso en dos etapas, consistente en hacer un cuestionario inicial a cada individuo, a partir del cual se generan los diseños eficientes, y luego realizar el resto de la encuesta. Desgraciadamente este método implica volver a contactar a los

encuestados, lo que tiene un alto riesgo de rechazo, por lo que tampoco se optó por este método.

Por otro lado, los otros dos métodos entregaron resultados lo suficientemente satisfactorios como para que se recomendara su utilización en caso de no optar por los dos previos. Ellos son:

- Segmentar la población en base a las posibles respuestas y generar un diseño eficiente para cada una de ellas. Este método es sólo ligeramente superior que el siguiente, pero requería de mayores esfuerzos en la implementación, por lo que no fue escogido.
- Generar sólo un diseño a partir de la media poblacional para aquella pregunta que otorga información al atributo pivote. Por ejemplo, si la pregunta pivote es la talla de zapatos del encuestado, y se sabe que la media poblacional es 40, se genera el diseño considerando el nivel 40 y sus respectivos pivotes. Este fue el método utilizado.

No fue posible calcular un promedio poblacional de la variable *Precio*, debido a que no se conoce el tamaño de la población. No obstante, como la mayor oferta de departamentos encontrada en el sitio *web* de Portal Inmobiliario fue de dos dormitorios, se asumió que esta sería la opción más escogida por los encuestados y se utilizaron sus niveles de *Precio* para generar los diseños eficientes.

El diseño del experimento con que se realizó las dos primeras encuestas piloto se puede observar en el ANEXO B. En él, al igual que en el resto de los diseños utilizados en la investigación, se buscaron de manera simultánea los diseños de los dos juegos del experimento. Se utilizó el procedimiento de minimización de Error-D descrito en la sección 3.4.2, con $N = 250.000$ e $I = 500$, el cual arrojó un valor de la medida de eficiencia de 0,558 para el diseño seleccionado. Resulta interesante mencionar que el hecho que el experimento incluya tantas variables mudas implicó dificultades en la

búsqueda de diseños eficientes. En efecto, debido a la estructura del experimento, conformado por dos juegos (ver Figura 3-1) y muchas variables mudas, los valores distintos de cero en la matriz de Fisher son bastante escasos, lo que puede provocar que ella sea singular (no invertible), y por lo tanto, que no sea posible construir la matriz asintótica de varianzas-covarianzas S^2 . Esto hizo que la búsqueda del diseño fuera realizada con especial cuidado. Al calcular el Error-D de cada diseño, era necesario que éste fuera calculado con $I = 500$ extracciones diferentes de los parámetros (ver sección 3.4.2) y luego se obtenía el promedio entre ellos. Sin embargo, en muchos diseños ocurría que no se podía invertir la matriz de Fisher, lo que implicaba que tampoco se pudiera calcular el Error-D. Estos diseños eran descartados, como resguardo para evitar posteriores problemas de estimación con la base de datos obtenida de la encuesta. Algunos diseños presentaron el problema con sólo una de las extracciones, pero fueron descartados de todas formas.

4.4.2. Diseño e Implementación de Interfaz en Microsoft Excel

La encuesta propiamente tal se diagramó en un archivo Microsoft Excel de la manera descrita en la Tabla 4-6.

Tabla 4-6: Estructura de la Interfaz de la Encuesta

Sección	Hojas
Datos personales	1
Datos familiares, vivienda actual y nueva vivienda	2
Explicación Juego 1	3
Juego 1 – 6 situaciones de elección	4-9
Explicación Juego 2	10
Juego 2 – 6 situaciones de elección	11-16
Preguntas de ingreso, agradecimientos	17

Ella constaba de 16 preguntas sobre datos de la persona, su familia y vivienda, 14 de las cuales se encontraban al inicio de la encuesta y dos al final. Estas dos últimas se referían al ingreso personal y familiar, tema que puede ser sensible para algunas personas, por lo que se prefirió incluirlas en la hoja final para evitar un rechazo prematuro de la encuesta. Todas las preguntas se incluyeron con la intención de caracterizar la muestra y luego poder segmentar a los entrevistados de acuerdo a sus respuestas.

En el ANEXO C se pueden observar imágenes de la interfaz diseñada. Se solicitó la asesoría de un diseñador en la elección de colores, letra y formato en la encuesta. Se optó por usar un color azul claro de fondo por tres razones:

- La población chilena relaciona dicho color con la Pontificia Universidad Católica de Chile, debido a que es el utilizado por su club deportivo y equipo de fútbol, que son conocidos en todo el país. Dado el prestigio de la universidad, se asegura mayor seriedad en las respuestas.
- El color azul expresa serenidad, profundidad, comodidad y racionalidad, por lo que crea un ambiente propicio para responder la encuesta.
- El hecho de tener un fondo de color hace que las hojas que sólo contienen preguntas se vean más atractivas y por ende capten mejor la atención del encuestado.

Se incluyó el logo de la Pontificia Universidad Católica de Chile en las diapositivas iniciales para otorgar identidad a la encuesta y reforzar su seriedad e importancia. En cuanto a las preguntas, se redactaron de manera que sus respuestas fueran simples y cortas. Se utilizó menús desplegables para dar agilidad a la encuesta y facilitar el análisis posterior de las respuestas. Esto era muy importante debido a la longitud del cuestionario inicial y a la cantidad de situaciones de elección. Las únicas preguntas que se dejaron abiertas fueron aquellas que implicaban una ventana desplegable demasiado larga (edad y tiempo al trabajo).

La explicación del Juego 1 describe, en primer lugar, el proceso al cual se enfrentarán los encuestados, luego la forma de pago del dividendo y finalmente los atributos que tienen los departamentos alternativos. Las descripciones son breves y van acompañadas de un ícono que representa el respectivo atributo en las situaciones de elección. Se puso especial cuidado en definir los atributos del barrio según los resultados de las entrevistas en profundidad.

Debido a que el Juego 1 sólo incluye características del barrio donde se localiza el departamento, se optó por usar planos para presentar las alternativas. Estos muestran la información a partir de los íconos identificadores de cada atributo y del edificio donde se compraría el departamento. La ausencia de un determinado ícono en el plano implica que el atributo no está presente en el barrio. Los planos son simples, con calles representadas por segmentos y sin ninguna otra información que pudiera resultar distractora (ver ejemplo de elección del Juego 1 en ANEXO C). Como base, se utilizaron calles reales de la comuna de Santiago, con el fin de dotar de mayor realismo al experimento. Además, se localizaron los elementos de barrio (atributos) a menos de cinco cuadras del edificio donde se encuentra el departamento en todos los casos, para que la persona percibiera una real cercanía a ellos, según las percepciones obtenidas en las entrevistas en profundidad.

La explicación del Juego 2 se hizo a través de frases simples y concisas para acotar el significado de las variables involucradas. Se indicó que el proceso de elección sería igual al de la sección anterior, pero cambiando los atributos de los departamentos, los cuales nuevamente se definieron tomando en consideración los resultados de las entrevistas en profundidad.

En el Juego 2 el único atributo descrito a través de una imagen fue *Limpieza_Calle*. Resultaba importante que los encuestados tuvieran una concepción homogénea de esta variable y coincidente con la realidad de la comuna de Santiago. De no utilizarse una

imagen que la representara, las personas podrían imaginar niveles de suciedad y mala mantención de las calles inexistentes en la comuna. Para esto se tomaron fotos de distintas calles del centro de Santiago con el objeto de definir cuáles eran los niveles posibles de mantención y limpieza. Luego, mediante el software Adobe PhotoShop, se compuso una calle “tipo” que tuviera los siguientes elementos: un edificio en construcción o a la venta, y edificaciones tanto antiguas como más actuales. Esta dicotomía entre ciudad nueva y vieja se encuentra fuertemente presente en la comuna de Santiago. En el ANEXO D se encuentran las fotografías a partir de las cuales se trabajó. Para representar los elementos de suciedad y mal estado de la calle se hicieron composiciones entre los elementos de distintas fotos, tratando de lograr realismo y tomando en consideración la luz y ángulo de la foto. En el ANEXO E se observan las dos fotografías que representaron los niveles del atributo *Limpieza_Calle* en la encuesta inicial.

Para finalizar, en la última hoja se agradecía a los encuestados y se les solicitaba la información de ingreso, tanto personal como de todas las personas que habitarían el departamento.

4.5. Encuestas Piloto

Para lograr un diseño final efectivo de la encuesta se realizó un proceso secuencial que obligó a realizar tres encuestas piloto. A partir de ellas se realizaron mejoras en la encuesta definitiva, tanto desde el punto de vista práctico como econométrico. La implementación sufrió diversas modificaciones durante este proceso pero además se pudo mejorar los valores iniciales de los parámetros a estimar, a fin de tener un diseño más eficiente de la encuesta.

4.5.1. Encuesta Piloto 1 (EP1)

La EP1 no fue utilizada para mejorar los valores iniciales de los parámetros del modelo, sino que para evaluar el funcionamiento de la interfaz en Microsoft Excel, la comprensión por parte de los encuestados y su recepción en general. Esta se realizó a diez personas, de manera presencial y monitoreando su interacción con la encuesta. Se les indicó que debían señalar todas aquellas secciones o frases que les parecieran confusas y de difícil comprensión.

El primer resultado que se obtuvo es que los individuos no tuvieron problemas para comprender el proceso de elección al cual debían enfrentarse. A todos les pareció claro e incluso algunos sostuvieron que lo encontraron interesante y que les hizo reflexionar en profundidad.

Se tenía especial interés en analizar dos puntos: la presentación del Juego 1 a partir de planos y la inclusión de imágenes para el atributo *Limpieza_Calle* en el Juego 2.

Se llegó a resultados interesantes al conversar con los individuos respecto al Juego 1. Si bien los planos les parecían fáciles de comprender y rápidos para evaluar la situación, la mayoría admitió que de alguna manera se veían influenciados por la distancia que se mostraba entre los íconos y el edificio. Pero el interés de la encuesta no era cuantificar el valor de las cuadras de distancia entre cada atributo de barrio y el departamento, sino que sólo su presencia o ausencia. Así, a pesar de que todos los íconos se localizaron a una distancia menor a cinco cuadras del edificio, resultaba imposible que la disposición espacial fuera exactamente igual para todos. Esto llevó a que se modificara la manera de presentar el Juego 1, cambiando los planos por un cuadro que contenía sólo los íconos de los atributos presentes en el barrio. Esta nueva configuración superaba la dificultad provocada por la distribución espacial y además permitía incluir una pequeña descripción de cada atributo que ayudaría a la efectiva recordación de su significado. En

el ANEXO F se observa la nueva manera en que se explicó el Juego 1 en la encuesta y también un ejemplo del nuevo formato de elección.

A cinco de las personas se les realizó la encuesta sin presentar fotografías para describir la variable *Limpieza_Calle*. Al consultarles, al final, su percepción general de la encuesta, cuatro declararon que les pareció muy larga y un poco tediosa. Además se les mostró al final las fotografías representativas de los dos niveles de *Limpieza_Calle* y tres declararon que se habían imaginado un lugar más sucio que el de las fotografías.

Por otro lado, los otros cinco individuos que contestaron la encuesta con *Limpieza_Calle* representada por fotografías declararon que, si bien la encuesta era larga, no la consideraron tediosa ni les pareció haber dedicado mucho tiempo a contestarla. Se les consultó si la imagen de *Limpieza_Calle* había influenciado su respuesta y tres declararon que posiblemente consideraron en mayor medida este atributo que si no hubiera fotografía. Los otros dos dijeron que este atributo en general no les parecía importante y que no les influyó por el hecho de que apareciera en una fotografía.

De estos resultados se concluyó que era necesario incluir la fotografía en el Juego 2 para mantener la atención de las personas y facilitar su procesamiento de información. Sin embargo, también quedó claro que existía el riesgo de potenciar el valor otorgado a dicha variable para aquellas personas que la consideraban importante, por lo que se optó por modificar levemente las imágenes representativas de los dos niveles, para que fueran más similares (ver ANEXO G). Así, en el nivel bajo se eliminó un poco de basura y se cubrió el graffiti de la vivienda de la derecha, ya que hacía percibir el lugar como “menos seguro” para algunos encuestados. Para el nivel bueno se había cubierto el pavimento de la calle, pero se eliminó esta capa para aumentar realismo y dejar el nivel como “normal” en lugar de “alto”.

4.5.2. Encuesta Piloto 2 (EP2)

Dado que el público objetivo de la encuesta eran personas que se encontraban en el proceso de comprar un departamento, se solicitó a una inmobiliaria, con compromiso de absoluta confidencialidad, permiso para realizar las encuestas en salas de venta de la comuna de Santiago. Al contactar a la inmobiliaria, se recibieron varios consejos:

- Las personas que visitaban sus diferentes salas de venta eran en su mayoría los mismos, puesto que es bastante común que los buscadores de departamento realicen recorridos por las diferentes opciones en un mismo día.
- El flujo de personas entre los días lunes y viernes es bastante bajo. Incluso hay días en que no se presenta ningún visitante.
- La mayor cantidad de visitas se concentran en el fin de semana, en especial cerca de la hora en que se abre la oficina de venta (11 AM) y la hora de cierre (7 PM).

A partir de lo anterior, se optó por concentrar la toma de encuestas en una sala de ventas y sólo durante los días sábados y domingos. Dado que el flujo de gente era bajo, no fue necesario contar con más de un encuestador.

En esta etapa se encuestó a 30 personas, lo que constituye una base de tamaño suficiente para mejorar la estimación inicial de los parámetros. Para esto se utilizó el software BIOGEME (Bierlaire, 2003) y se estimó un modelo Logit Multinomial con la estructura descrita en la ecuación (4.1). Se realizaron dos estimaciones: una tradicional (M1-EP2) y otra con parámetros de escala diferentes para los dos juegos del experimento (M2-EP2), utilizando para ello la función *Scale* de BIOGEME (ver Bierlaire, 2008). Dado que es posible estimar sólo uno de los dos parámetros de escala, el del Juego 1 se fijó como igual a uno y se estimó el del Juego 2.

En la Tabla 4-7 se observan los modelos obtenidos. Se puede ver que el modelo M2 – EP2 es bastante distinto y, en general, superior al M1 – EP2; además, si bien el parámetro de escala λ_{J2} no resultó significativamente distinto de uno para $\alpha = 0,05$, sí lo es para $\alpha = 0,1$, por lo que se esperó que con una muestra de mayor tamaño resultara más significativo. El modelo M1-EP2 es una versión restringida de M2-EP2, con $\lambda_{J2} = 1$, por lo que se realizó un test LR entre ellos. Se rechaza la hipótesis nula para $\alpha = 0,25$, es decir, el modelo M2-EP2 es más informativo sobre M1-EP2 al 75%. Por lo tanto, se decidió utilizar el modelo M2-EP2 como base para diseñar la Encuesta Piloto 3.

Tabla 4-7: Modelos Estimados con EP2

	M1-EP2		M2-EP2	
	Parámetro	test-t	Parámetro	test-t
θ_{metro}	1,530	4,87	1,600	5,01
θ_{areas_verdes}	0,582	1,71	0,856	1,98
$\theta_{cultura}$	0,612	2,48	0,592	2,39
θ_{precio}	-0,032	-1,66	-0,049	-1,92
$\theta_{espacio_comun}$	0,362	1,09	0,523	0,56
$\theta_{limpieza_calle}$	-2,330	-7,06	-8,410	-0,64
$\theta_{servicios}$	0,453	1,88	0,630	2,12
$\theta_{superficie}$	0,419	1,75	1,340	0,65
λ_{J2}	-	-	0,262	1,74 (respecto a 1)
$l(\theta)$	-177,618		-176,95	
N° de Observaciones	360		360	

Un factor importante que se descubrió en la EP2, fue que la manera de contactar a las personas resultaba extremadamente ineficiente, ya que sólo se lograban entre 5 y 10 encuestas por fin de semana. Si bien la disposición general a contestarla era positiva, el flujo de gente era muy bajo, lo que llevaba a una lenta obtención de respuestas.

4.5.3. Encuesta Piloto 3 (EP3)

Para buscar el diseño eficiente de la EP3 nuevamente se utilizó el método Bayesiano de estimación, utilizando 500 extracciones de cada parámetro para evaluar el Error-D de cada diseño. A diferencia del diseño original donde las extracciones de los parámetros provinieron de distribuciones tanto Normales como Uniformes dependiendo de su fuente, en esta ocasión se asumieron distribuciones Normales para todos los parámetros, con las medias señaladas en la Tabla 4-7 para el modelo M2-EP2, y varianzas que aseguraran el signo correcto del parámetro para 3000 extracciones aleatorias. El diseño eficiente utilizado para la EP3 se puede observar en el ANEXO H, y tuvo un Error-D de 0,85.

Debido a los resultados obtenidos en la EP2, se decidió cambiar la estrategia de contacto a los encuestados, puesto que llevaría demasiado tiempo recolectar la cantidad de individuos suficientes para tener una muestra suficiente. La inmobiliaria que había permitido la toma de encuestas en sus salas de venta facilitó una base de datos absolutamente confidencial de personas que hubieran buscado información sobre departamentos en la comuna de Santiago en una determinada página *web* entre Abril de 2007 y 2008. Fue necesario filtrar y ordenar la información recibida, puesto que había muchísimas personas repetidas varias veces y también un número importante de datos evidentemente falsos (por ejemplo: 111@111.cl). La base final, de 5.664 personas, entregaba nombre completo, correo electrónico, RUT, dirección y teléfono. Dado que son las mismas personas quienes completan estos datos al acceder a la página web, la mayoría de ellas no entregaba su información privada, como dirección y teléfono, pero sí inscribían su nombre y correo electrónico. Éste último resulta ser un modo de contacto bastante efectivo, por lo que se optó por desarrollar una encuesta que fuera posible enviar por correo electrónico.

Se temía que se podía incurrir en un error de cobertura (Couper, 2000) por realizar la encuesta sólo a personas que tuvieran acceso a Internet, pero la inmobiliaria pregunta a los interesados en comprar departamentos de dónde obtienen información sobre ellos, y un 90% responde que desde Internet. Además, Chile presenta índices de conectividad a Internet muy altos y similares a los de Italia y España, con un 86,7% de la población catalogada como usuario directo o indirecto⁶ al año 2006 (Godoy, 2006). Finalmente, se ha demostrado en investigaciones previas (Fleming y Bowden, 2007) que la tasa de respuesta para encuestas realizadas por e-mail y por correo tradicional vía interceptación, no son significativamente diferentes, así como tampoco las valoraciones subjetivas estimadas con ellas. Debido a todos estos puntos, se decidió realizar la encuesta a través de correo electrónico.

Por otro lado, existen muchas ventajas asociadas a realizar una encuesta no presencial, siendo las principales su bajo costo y el hecho que la persona no se ve afectada por la presencia del encuestador (Buckingham y Saunders, 2004). Esto último puede ocurrir de dos maneras: la persona puede no querer entregar información por vergüenza o recelo, o puede necesitar pensar una respuesta mayor tiempo de lo normal. El hecho que no se encuentre el encuestador presente implica una mayor privacidad y una mayor disposición de tiempo para contestar preguntas. Sin embargo, también se deben tomar en cuenta algunas desventajas, como una baja tasa de respuesta o la posible no comprensión de alguna pregunta (Fink, 1995). Algunas recomendaciones para asegurar una alta tasa de respuestas son: enviar una carta de contacto que logre motivar a las personas y que

⁶ Usuarios directos (48,6%): utilizaron Internet en los últimos 90 días, al realizarse el estudio. Usuarios indirectos (38,1%): dependen de alguien cercano para revisar su correo, bajar información o realizar trámites en Internet.

sea simple de comprender, realizar la encuesta de la manera más corta y llamativa posible, y asegurar la confidencialidad de las respuestas. Todas estas recomendaciones fueron estrictamente tomadas en cuenta al realizar la encuesta.

El contacto a las personas seleccionadas se realizó a través de un correo electrónico enviado desde la cuenta del profesor Juan de Dios Ortúzar, en lugar de la de la tesista, debido a que él podía ser buscado en Internet e inmediatamente aparecería como académico de la Pontificia Universidad Católica de Chile. En el ANEXO I se muestra el correo electrónico de contacto utilizado. Resultaba importante enfatizar la importancia de que la persona contestara y los incentivos que habían para ello (Moskowitz y Martin, 2008), por lo que se recalcó que el estudio provocaría un bien a la comunidad; también se comentó que se recibiría un premio y se otorgó una clave, indicativa de que no cualquier persona podía realizar la encuesta, sino que ésta formaba parte de una investigación seria.

Si bien ya se ha utilizado Microsoft Excel para encuestas realizadas a través de correo electrónico (Black *et al*, 2005), se tenía el temor de que los archivos no fueran devueltos con las respuestas, que las personas no supieran ocuparlo (puesto que había que habilitar algunas Macros para iniciar el proceso), que no la contestaran por temor a que fuera un virus, etc. Luego, se tenían dos opciones para realizar la encuesta por Internet: contratar a una persona que la programara, como se había hecho en experiencias previas (Iragüen, 2002), o utilizar un servicio *on-line* donde se pudieran diseñar y realizar encuestas. Se optó por esto último, por ser menos costoso y porque permitía un mayor control sobre el proceso (ver costos finales en sección 4.7).

Es importante notar que, como el diseño del experimento de la EP3 es diferente al de la encuesta final, hubiera sido necesario programar parte de la encuesta dos veces si es que no se hubiera utilizado un servicio *on-line* para realizarla. Para esto se trabajó con la compañía Webtools S.L. a través de su página web www.e-encuesta.com, debido a que

en ella se pudieron elaborar las encuestas de manera fácil, incluyendo imágenes en las preguntas, y recibiendo los resultados posteriormente en formato *.cvs, el cual es muy cómodo de procesar. La interfaz utilizada es análoga a la de Microsoft Excel, pero se tomó en consideración las excelentes recomendaciones de diseño descritas en Dillman (2000), como por ejemplo, utilizar botones de navegación que permitan comprender el orden lógico de la encuesta; utilizar siempre el mismo fondo; enlistar de manera vertical en lugar de horizontal; usar un color oscuro o negrita para las preguntas y uno más claro para las respuestas; diseñar cada hoja de la encuesta de manera que sea vista en una pantalla, sin que sea necesario utilizar *scroll-down* o la tecla *left* para ver toda la información; indicar el porcentaje de la encuesta ya contestado; utilizar íconos o letras de igual tamaño para elementos de igual importancia; usar simplicidad, regularidad y simetría, etc. La única pregunta que fue modificada respecto de la EP2, fue aquella donde se solicitaba el ingreso personal y familiar, la que se transformó a una pregunta abierta en lugar de menú desplegable. Este cambio se realizó porque para la modelación posterior resultaba más útil contar con el ingreso exacto de los individuos.

Como se mencionó anteriormente, la encuesta presenta diferentes valores para los departamentos en cada situación de elección dependiendo de la cantidad de dormitorios que desea cada persona. En la interfaz desarrollada en Microsoft Excel, era posible asignar el precio del departamento según lo que contestara el individuo, pero ello no fue posible de implementar en la interfaz *online*. Por esto se elaboraron cinco encuestas iguales, una para cada número de habitaciones (ambiente único, 1, 2, 3 y 4), que solamente se diferenciaban en el valor del atributo *Precio*. Las personas contactadas debían ingresar sólo a un enlace según la cantidad de dormitorios que desearan (ver ANEXO I).

El primer aspecto que se deseaba estudiar con la EP3 era qué porcentaje aproximado de las personas contactadas contestaba la encuesta. El día 9 de Mayo de 2008 se envió un correo electrónico a 200 destinatarios con un encabezado genérico; es decir, comenzaba

con “Estimado/a”. Los correos electrónicos de los destinatarios se remitieron en copia oculta, para respetar la confidencialidad de las personas. Por otro lado, el mismo día se enviaron otros 200 correos individuales donde el nombre del destinatario aparecía en el encabezado, por ejemplo, “Estimada María”. El objetivo de este ejercicio era determinar si se reportaba una mayor cantidad de respuestas al dirigirse de manera personal a cada individuo. La tasa de respuesta de las personas contactadas individualmente fue de 15%, mientras que la de aquellas contactadas de manera genérica alcanzó sólo un 8,5%.

Por otro lado, se constató que aún cuando 74 personas comenzaron a responder la encuesta, ésta fue terminada por sólo 48 de ellas. Al analizar las respuestas, se observó que la mayor pérdida de encuestados ocurría al cambiar de página (ver Figura 4-1); por esto, se encontró necesario disminuir el número de preguntas e información solicitada a lo estrictamente necesario para evitar este efecto.

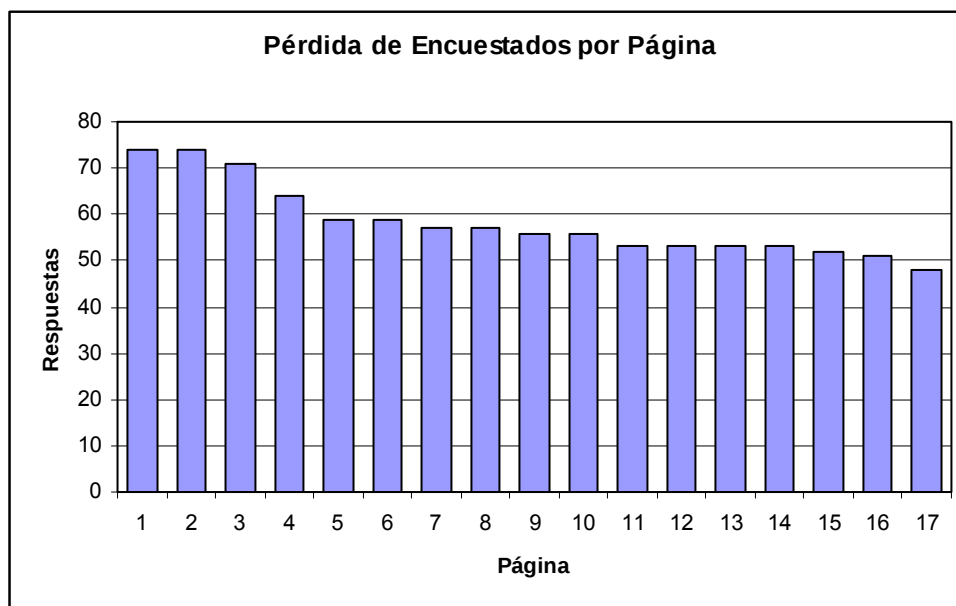


Figura 4-1: Efecto de Pérdida de Encuestados por Página de la Encuesta

Con los datos obtenidos, nuevamente se estimaron dos modelos, análogos a los de la sección 4.5.2, cuyos resultados se muestran en la Tabla 4-8.

Tabla 4-8: Modelos Estimados con EP3

	M1-EP3		M2-EP3	
	Parámetro	test-t	Parámetro	test-t
θ_{metro}	1,130	4,97	1,370	4,18
$\theta_{\text{areas_verdes}}$	0,631	3,23	0,651	3,27
θ_{cultura}	0,130	0,55	0,592	1,51
θ_{precio}	-0,016	-3,51	-0,045	-2,55
$\theta_{\text{espacio_comun}}$	0,384	2,03	0,600	1,62
$\theta_{\text{limpieza_calle}}$	-1,730	-5,69	-3,000	-2,04
θ_{servicio}	0,198	1,05	0,305	1,14
$\theta_{\text{superficie}}$	0,547	3,38	0,899	2,06
λ_{J2}	-	-	0,545	1,53 (respecto a 1)
$l(\theta)$	-391,52		-390,99	
Nº de Observaciones	612		612	

Si bien nuevamente el parámetro de escala no es significativamente distinto de uno, se juzgó otra vez al segundo modelo (M2-EP3) como superior; esto se debió a que los parámetros con test – t más bajos del modelo M1-EP3 son más elevados en M2-EP3, y a que se supuso que con una muestra mayor definitivamente se tendrían parámetros de escala diferentes entre ambos juegos. En definitiva, se utilizaron los valores de los parámetros del modelo M2-EP3 como *priors* para poder buscar un diseño eficiente para la encuesta final. Debido a la importancia de estos últimos parámetros en el diseño de la encuesta final, se decidió tomar otra medida de resguardo adicional y realizar una simulación de la muestra, para verificar que se pudieran recuperar los parámetros de diseño del experimento.

4.5.4. Simulación para Recuperar Parámetros del Experimento Final

Esta simulación tenía como objetivo verificar que los valores de los parámetros del modelo que se utilizarían para generar el experimento de PD final pudieran obtenerse como resultado de modelar las respuestas de una muestra simulada. Se simuló 5.400 individuos, cada uno de los cuales debió “contestar” 12 preguntas; con esto, el tamaño final de la muestra de preudoindividuos fue de 64.800. En este proceso se siguieron los siguientes pasos:

1. Se calculó la probabilidad de que un individuo quisiera un departamento de 1, 2, 3, 4 dormitorios o ambiente único, a partir de las respuestas obtenidas en las tres encuestas piloto (ver Tabla 4-9). Se asignó a cada uno de los 5.400 individuos simulados el número de dormitorios que deseaba, según estas probabilidades.

Tabla 4-9: Probabilidad Generada de Cantidad Requerida de Dormitorios

Cantidad de Dormitorios	Probabilidad
Ambiente Único	0,0196
1	0,1961
2	0,3529
3	0,3922
4	0,0392

2. Para cada una de las 12 situaciones de elección enfrentadas por cada individuo, se calculó la utilidad representativa de cada alternativa utilizando los valores de los parámetros del modelo M2-EP3 (Tabla 4-8) y el valor de *Precio* según los dormitorios deseados.

3. Para construir la utilidad percibida de cada alternativa, era necesario agregar un error ε con distribución Gumbel independiente e idéntica, media cero y varianza desconocida. La función de distribución Gumbel es:

$$(4.2) \quad F(\varepsilon) = e^{e^{-\mu(\varepsilon-\eta)}}$$

con μ el parámetro de escala y η la moda de la función. Para generar sus valores se utilizó el Método de la Transformada Inversa (Johnson 1987) que permite generar variables aleatorias con esta distribución a partir de números r_i con distribución Uniforme entre 0 y 1 evaluando (4.3):

$$(4.3) \quad F^{-1}(\varepsilon) = \eta - \frac{1}{\mu} \ln(-\ln(r_i))$$

Además se sabe que la media cumple con la condición (Williams, 1977):

$$0 = \eta + \frac{\gamma}{\mu} \Rightarrow \eta = -\frac{\gamma}{\mu} \text{ y } \mu = -\frac{\pi}{\sigma\sqrt{6}} \text{ donde } \gamma \text{ es la constante de Euler } (\gamma \approx 0,577).$$

Luego:

$$(4.4) \quad \Rightarrow F^{-1}(\varepsilon) = \frac{-\gamma\sigma\sqrt{6}}{\pi} - \frac{\sigma\sqrt{6}}{\pi} \ln(-\ln(r_i))$$

De esta forma, el único valor desconocido de la función (4.4) es la desviación estándar σ . Garrido (1991) recomienda que el valor a utilizar no sea muy pequeño, porque podría forzar la elección a ser completamente determinística, ni demasiado grande, para que el componente aleatorio no determine completamente la elección.

4. Se sumó el componente determinístico y el aleatorio de la función de utilidad de cada alternativa y en base a ese valor se determinó cuál alternativa era escogida en cada situación de elección.

5. Se estimó un modelo MNL a partir de los valores de los atributos y las elecciones obtenidas en el paso 4. Posteriormente se comparó los parámetros obtenidos con los parámetros originales.

Este proceso se repitió muchas veces modificando el valor de σ . Para poder comparar los modelos generados se utilizó el siguiente indicador: se calcula el porcentaje de diferencia entre cada parámetro estimado con la simulación y su valor inicial, se eleva este porcentaje al cuadrado, se suma sobre todos los parámetros del modelo y se divide por el total de parámetros. Este indicador se llama “Error cuadrático medio” y en la Figura 4-2 se observa el valor que toma en la medida que se utilizan distintos valores de la desviación estándar σ . Es claro que los valores más bajos de este indicador, que se busca minimizar, se encuentran para rangos de σ entre -1 y -2 o entre 1 y 2.

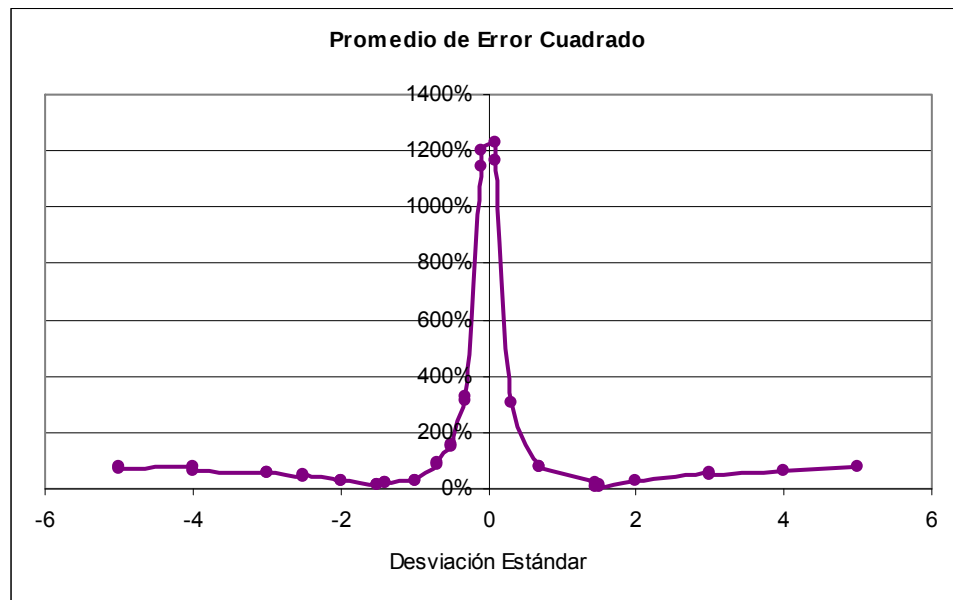


Figura 4-2: Error Cuadrático Medio entre Parámetros Estimados y Simulados, según la Desviación Estándar

De hecho, para la simulación realizada con $\sigma = 1,5$ se obtuvieron valores estimados de los parámetros muy similares a los iniciales, con diferencias menores a 10% para todos ellos. Este resultado, que se detalla en la Tabla 4-10, es completamente satisfactorio como experimento de recuperación de parámetros a través de simulación. Casi todos los parámetros iniciales se encuentran en dentro del intervalo de confianza de los estimados, a excepción de dos de ellos. Si se toma en consideración la diferencia en cuanto a cantidad de observaciones, y lo ajustados que son los intervalos de confianza de los parámetros simulados, se deduce que los parámetros previos a utilizar son apropiados.

Tabla 4-10: Resultado Simulación para Recuperación de Parámetros

Atributo	Parámetro Inicial	Parámetro Recuperado	¿Está el Inicial en Interv. de Confianza del Simulado?	Diferencia
θ_{metro}	1,370	1,270	Sí	7,3%
θ_{areas_verdes}	0,651	0,642	No	1,4%
$\theta_{cultura}$	0,592	0,624	No	-5,4%
θ_{precio}	-0,045	-0,041	Sí	9,5%
$\theta_{espacio_comun}$	0,600	0,549	Sí	8,5%
$\theta_{limpieza_calle}$	-3,000	-2,740	Sí	8,7%
$\theta_{servicios}$	0,305	0,297	Sí	2,6%
$\theta_{superficie}$	0,899	0,820	Sí	8,8%
λ_{J2}	0,545	0,502	Sí	7,9%
N° de Observaciones	612	64.800		

Es posible analizar otro aspecto de la simulación que da luz sobre futuras aplicaciones de la misma. Cuando se realiza el paso 2 descrito anteriormente, es posible obtener cuales serían las elecciones de los individuos si sus utilidades fueran completamente

determinísticas. Luego, al agregar el componente estocástico a las utilidades, se obtienen las elecciones finales. Con estos datos se calculó el porcentaje de observaciones donde cambiaba la elección por efecto del componente estocástico de la utilidad, obteniéndose los resultados de la Figura 4-3. Los rangos de σ donde se minimiza la diferencia entre los parámetros simulados y los originales, corresponde al área de la curva de la Figura 4-3 donde comienza el descenso más acelerado, es decir, donde entre un 35% y 40% de las observaciones cambian la elección por efecto del error estocástico.

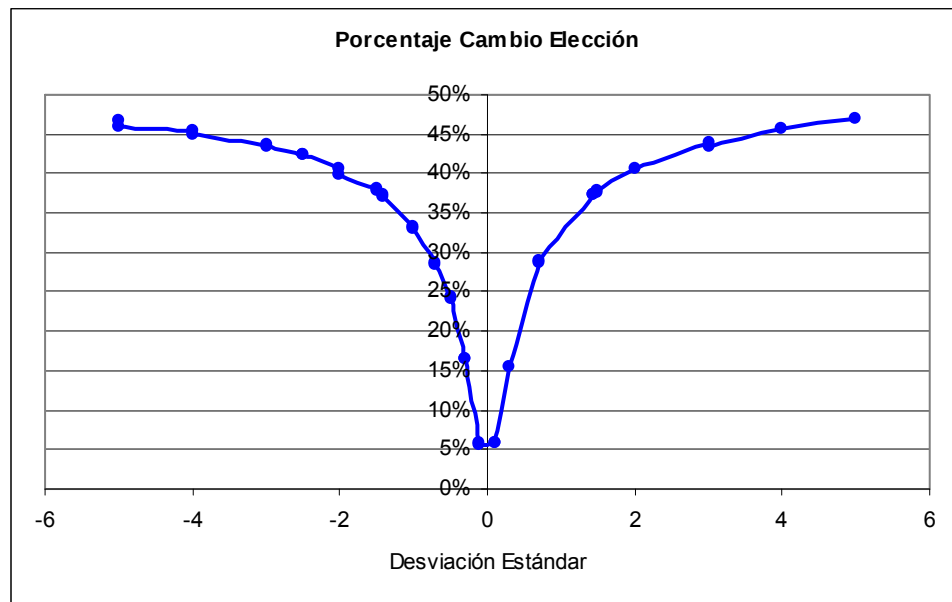


Figura 4-3: Porcentaje de Cambio de Elección al Agregar Componente Aleatorio a la Utilidad, según Desviación Estándar

Este resultado permite dar una referencia cuantitativa a la recomendación entregada en Garrido (1991) sobre el valor que se debe utilizar de la desviación estándar, la que debe permitir que la elección no sea gobernada completamente por el componente determinístico ni por el aleatorio.

4.6. Diseño Final de la Encuesta

Finalmente, se realizó el proceso de selección del diseño eficiente para la encuesta final, a partir de los valores de los parámetros obtenidos en el modelo M2-EP3 de la Tabla 4-8. Al igual que en las encuestas pilotos previas, se utilizó un método Bayesiano de estimación para calcular el Error-D, usando 500 extracciones de cada parámetro, todos con distribución Normal. El diseño final del experimento se encuentra en el ANEXO J y tiene un Error-D de 0,711.

Debido al efecto de pérdida de encuestados por páginas de la encuesta presentado en la EP3, se optó por eliminar las preguntas que no se consideraran estrictamente necesarias, esto es, estado civil, licencia de conducir, cuál es el *status* de la persona en la vivienda actual (arrendatario, allegado, etc.) y cantidad de habitantes menores de 15 años en la nueva vivienda. Con ello, las preguntas iniciales definitivas de la encuesta, relativas a características socioeconómicas, de la vivienda actual y de la nueva vivienda, se pudieron incluir en sólo dos páginas. En la Tabla 4-11 se adjuntan las preguntas del comienzo de la encuesta, así como la simbología que se utilizó para representarlas. La encuesta final en Internet, incluyendo todos estos ajustes, se encuentra en el ANEXO K. Ésta tardaba entre 10 y 15 minutos en ser contestada.

La encuesta definitiva fue enviada por correo electrónico a 3.500 personas entre los días 27 y 28 de Junio de 2008, nuevamente desde la cuenta del profesor supervisor. En base a los resultados obtenidos en la EP3, se optó por enviar correos personales con el nombre de cada destinatario en el encabezado. Para hacer este proceso más rápido, se agrupó a las personas con el mismo nombre, remitiendo su correo electrónico en copia oculta. Así, por ejemplo, para el nombre “Alejandra” se envió un solo correo con el encabezado “Estimada Alejandra”, en lugar de los 64 correspondiente a la cantidad de personas con dicho nombre. Si bien este método permitió acelerar el proceso del envío de *e-mails*, de todas formas se necesitaron aproximadamente 12 horas para completar la tarea.

Tabla 4-11: Preguntas de Características Socioeconómicas de la Encuesta Final

Pregunta	Simbología
Edad	<i>Edad</i>
Sexo	<i>Sexo</i>
¿Qué estudios tiene?	<i>Estudios</i>
¿A qué se dedica?	<i>Ocupación</i>
¿Dónde vive usted actualmente?	<i>Comuna</i>
¿Cuántos automóviles habrá en su nuevo hogar?	<i>N°_Autos</i>
Número de personas que habitarán en su nuevo departamento	<i>N°_Habitantes</i>
¿Cuántos años espera vivir en su nuevo departamento?	<i>Años</i>
Por favor indique su ingreso líquido personal en pesos chilenos (\$)	<i>Ingreso_Personal</i>
Por favor indique la suma de los ingresos líquidos (\$) de todas las personas que vivirán en su nuevo hogar (incluido usted)	<i>IFAM</i>

De las 3.500 personas contactadas, un 61% correspondía a mujeres y 39%, a hombres. Luego de ser remitidos los mensajes, se recibieron respuestas de error correspondientes a 120 correos electrónicos (3,4% del total), indicando que éstos no habían sido recibidos. Esto se puede haber debido a que algunas personas entregaron correos erróneos, a que sus cuentas se encontraban con la capacidad completa o a que los *e-mails* estaban obsoletos.

Se obtuvo un total de 615 respuestas de la encuesta, correspondiente a un 18,2% del total de correos enviados, si se descuentan aquellos que fueron devueltos con mensaje de error. Sin embargo, muchos de ellos no contestaron la encuesta completa, por lo que fue necesario realizar una limpieza de la base de datos. Finalmente, la base de datos utilizada para modelar dispuso de 532 individuos (15,7% del total contactado). De ellos, 37 no entregaron información sobre su *Ingreso_Personal* e *IFAM*; se volvió a contactar a cinco de estos individuos, los cuales habían entregado su correo electrónico al

responder la encuesta. Dos personas respondieron entregando sus ingresos, una comunicó que no sabía el monto familiar y que él no tenía ingresos por ser estudiante, y los otros dos no respondieron el correo electrónico. Con ello, finalmente se utilizaron sólo 497 datos para aquellos modelos que incluyeran las variables de ingreso y 532 observaciones para los modelos sin este tipo de variable.

En la Tabla 4-12 se muestra el tamaño mínimo de la base de datos requerido para cada uno de los parámetros (ver sección 3.5). Luego, como límite inferior teórico se utiliza 130, correspondiente a *Servicios*, por ser el mayor de todos. Dado que la muestra cuenta con 497 individuos, el límite se encuentra sobrepasado ampliamente, por lo que el tamaño se considera satisfactorio.

Tabla 4-12: Límite Inferior Teórico de Cantidad de Encuestados, por Parámetro

Parámetro	Cantidad de Individuos (N)
θ_{metro}	7
$\theta_{\text{areas_verdes}}$	17
θ_{cultura}	57
θ_{precio}	10
$\theta_{\text{espacio_comun}}$	86
$\theta_{\text{limpieza_calle}}$	6
$\theta_{\text{servicios}}$	130
$\theta_{\text{superficie}}$	23

4.7. Costos de Diseño y Realización de la Encuesta

Finalmente, en la Tabla 4-13 se detallan los costos en que se incurrió para realizar las encuestas. Éstos no incluyen las horas de trabajo invertidas por la tesista ni su profesor guía.

Tabla 4-13: Costos de la Encuesta, Sin Incluir Trabajo de Tesista

Ítem	Monto
Entrevistas en Profundidad	\$500.000
Habilitación de Cuenta Profesional con saldo de 1000 respuestas en www.e-encuesta.com	65 € (Abril 2008) ≈ \$46.000
Premios (2) de \$50.000 rifados entre los individuos encuestados	\$100.000
Total	\$646.000

Como se puede ver, el menor costo fue de desarrollo de la encuesta en Internet, por lo que se recomienda utilizar este tipo de soluciones, en lugar de contratar a una persona que la programe. El elevado costo de las entrevistas en profundidad se debió a que se contaba con un plazo muy acotado para realizarlas, por lo que se debió contratar a dos personas en vez de una. De todas formas, el costos es sustancialmente menor al de una encuesta de preferencias reveladas, que puede superar los \$2.000.000 (Yáñez, 2008).

5. ANALISIS DE LA BASE DE DATOS

Los datos fueron revisados con el fin de identificar posibles inconsistencias en ellos o problemas en las respuestas entregadas. Se encontró que las únicas observaciones que presentaron inconsistencias fueron:

- Tres datos en donde se reportó un *Ingreso_Personal* mayor al ingreso familiar (*IFAM*). Se contactó a dos de las personas con el problema (puesto que la tercera no reportó su correo electrónico, lo que era optativo) y ambas respondieron que se confundieron al momento de entregar su *IFAM* y no consideraron su propio ingreso en este valor. Luego, se sumaron los dos ingresos reportados como variable *IFAM* para estas personas, y se asumió que el tercer encuestado había tenido el mismo problema.
- Diez de los usuarios presentaron un comportamiento lexicográfico en la variable *Precio* (Foster y Mourato, 2002), puesto que escogieron la alternativa de menor valor en todas las situaciones de elección. Estas observaciones no fueron eliminadas, pero sí se tabularon como lexicográficos en caso de se presentaran problemas al momento de estimar modelos.
- 286 usuarios evidenciaron comportamiento “semi-lexicográfico” en la variable *Limpieza_Calle*, y 171 en la variable *Metro*. Este comportamiento no pudo ser catalogado como lexicográfico puesto que no todas las situaciones de elección contaban con estas variables. De hecho, dado que el experimento consistía en dos juegos con distintos atributos y que en algunas situaciones de elección las variables tenían el mismo nivel para ambas alternativas, se pudo analizar el comportamiento lexicográfico en sólo cinco situaciones para la *Metro* y en cuatro para *Limpieza_Calle*. Debido a ello y al alto número de observaciones con este rótulo, se optó por no excluirlas al estimar los modelos; esto habría podido disminuir de manera importante la significancia estadística de los parámetros. Sin embargo,

pareció relevante contar con esta información para analizar posteriormente los resultados de los modelos.

Se revisaron otros posibles errores, por ejemplo verificar que N° _Autos fuera menor o igual que N° _Habitantes, pero los datos presentaron coherencia por lo que no se eliminó ninguna observación. A continuación se presenta un análisis estadístico de las respuestas de la muestra.

5.1. Características Socio-Económicas

La primera variable a estudiar fue la distribución de la población encuestada según *Sexo*, constatándose 348 mujeres (65,4%) y 184 hombres (34,6%). Esta distribución es bastante similar a la de la base completa de 3.500 personas contactadas, tal como se muestra en la Figura 5-1.

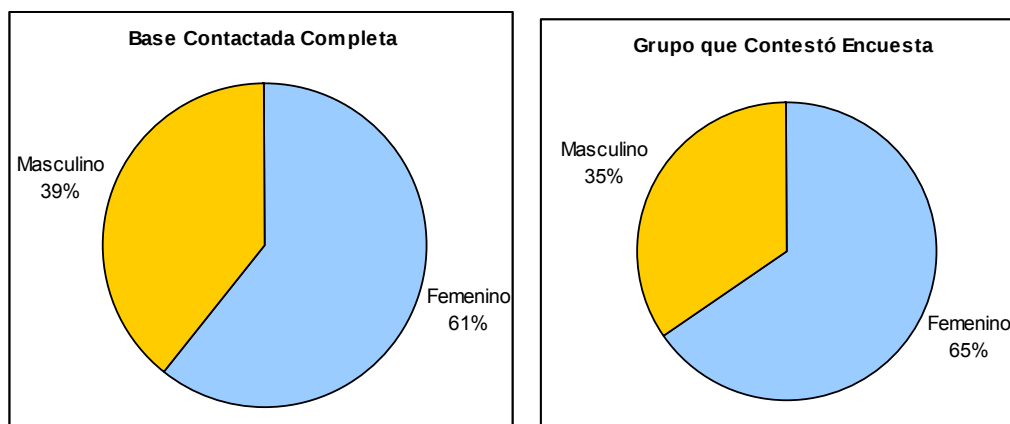


Figura 5-1: Distribución de Variable Sexo

Al parecer, las mujeres tienen menos objeciones en entregar sus datos personales en una página *web*, puesto que la mayoría de los datos falsos eliminados al recibir la base completa correspondían a hombres. De hecho, al realizar las encuestas piloto cara a cara, en la sala de ventas, muchas veces la encuesta era respondida por parejas, pero la mujer

era quien entregaba sus datos. La variable *Sexo* debiera ingresar interactuando con otras en los modelos, y no es posible definir *a priori* cuál será el signo que debiera llevar.

A continuación se estudió la variable *Edad*, cuyo histograma según sexo del encuestado se muestra en la Figura 5-2. Claramente, la población se encuentra concentrada en el rango entre 26 y 35 años, dado que el 54% de los hombres y el 51% de las mujeres se encuentran en dicha categoría. Dicha distribución sugiere que la mayoría de los encuestados se disponen a adquirir su primera vivienda.

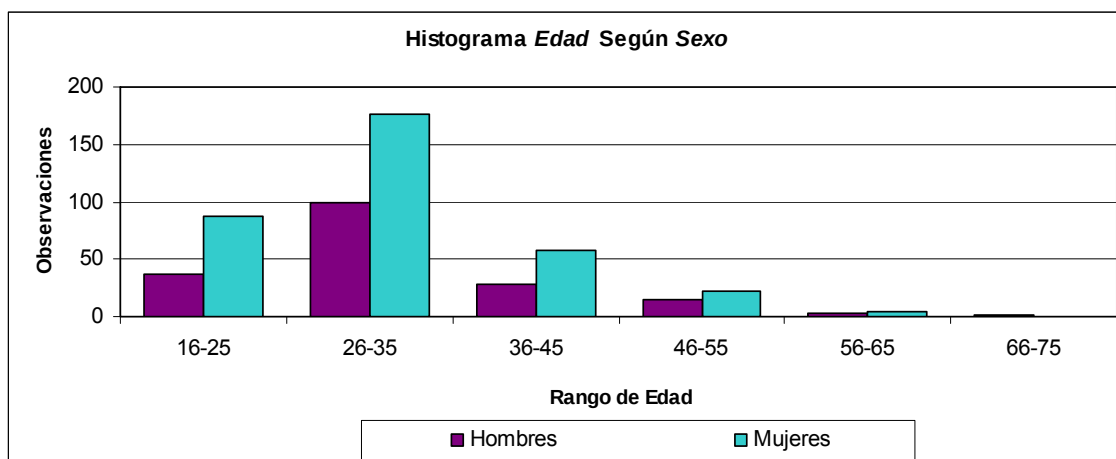


Figura 5-2: Histograma de Variable *Edad*

En la Figura 5-3 se observa la distribución de la variable *Estudios* en toda la muestra. Ninguno de los encuestados reportó no tener estudios o haber alcanzado sólo un nivel primario. Si además se considera que sólo un 4% tiene nivel secundario, prácticamente toda la población cuenta con estudios superiores, completos o incompletos. Es por ello que no se espera una gran incidencia de esta variable en los modelos.

La variable *Ocupación* se distribuye en el total de la muestra de la manera indicada en la Tabla 5-1. La mayoría de las personas trabajan, pero cerca de un 20% se distribuye entre personas que estudian y tienen otras actividades.

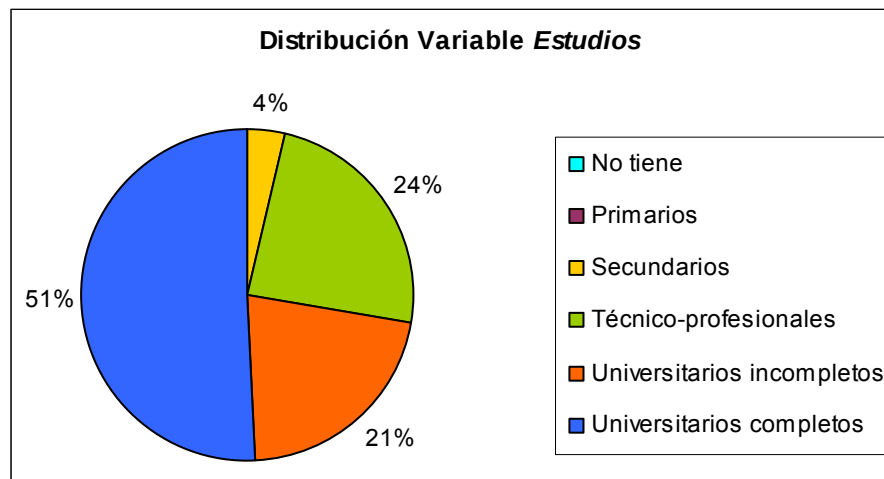


Figura 5-3: Distribución Variable *Estudios*

Tabla 5-1: Distribución Variable *Ocupación*

Ocupación	Cantidad	Porcentaje
Trabaja	434	81,6%
Estudia	59	11,1%
Otros	39	7,3%

Cabe recordar que un departamento no es necesariamente comprado por una sola persona, por lo que la aparente falta de recursos que podría tener una persona que no trabaja se puede suplir con los aportes del resto de la familia o del grupo que esté comprando. Además, pueden tener aportes extra, como podría ser el caso de estudiantes, herencias, etc.

Si se analiza el detalle de las ocupaciones (Figura 5-4), se ve que un 76% de las personas trabaja a tiempo completo, un 10% estudia en la universidad, alrededor de 5% trabaja en jornada parcial y el resto de las ocupaciones alcanzan menos de un 4% de la muestra.

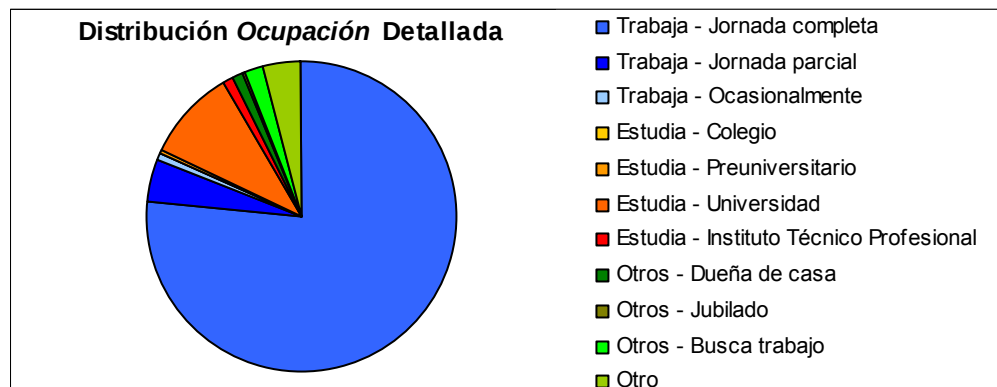


Figura 5-4: Distribución Variable *Ocupación* en Detalle

Finalmente, la distribución de la ocupación según sexo es bastante homogénea entre hombres y mujeres. En la Figura 5-5 se observa que en ambos géneros la *Ocupación* “trabajo” supera el 80%, siendo en el caso de las mujeres un par de puntos porcentuales menor que para los hombres. Esta variable no será utilizada de manera desagregada en los modelos, sino que como variable muda para trabajo, estudio y otros.

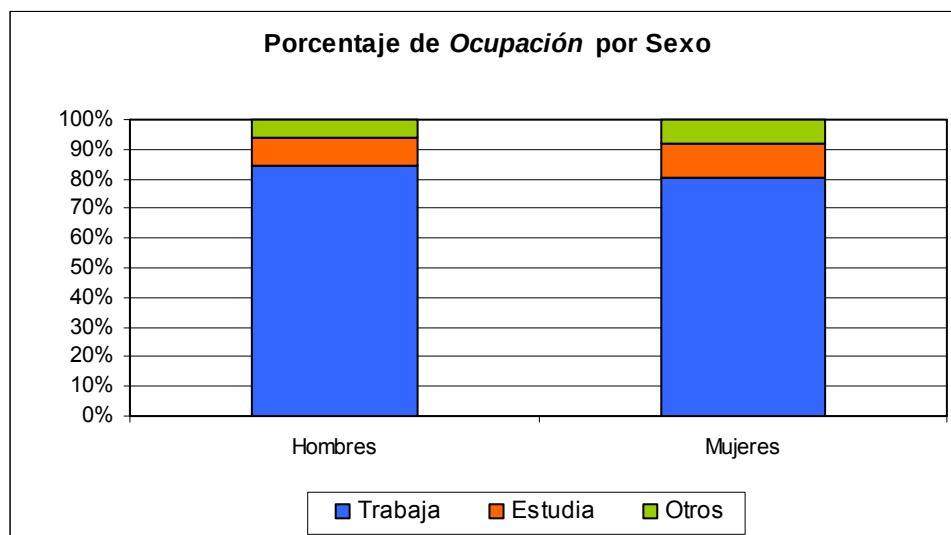


Figura 5-5: Porcentaje Variable *Ocupación* por Sexo

Antes de realizar la encuesta se había obtenido la información, tanto de la inmobiliaria que facilitó las salas de venta como de los investigadores de la Facultad de Arquitectura, Diseño y Estudios Urbanos, que muchas de las personas que tenían intención de comprar un departamento en el centro de Santiago provenían de la misma comuna; esto resultó confirmado con las respuestas obtenidas. En la Figura 5-6 se observa el porcentaje de observaciones proveniente de cada comuna.

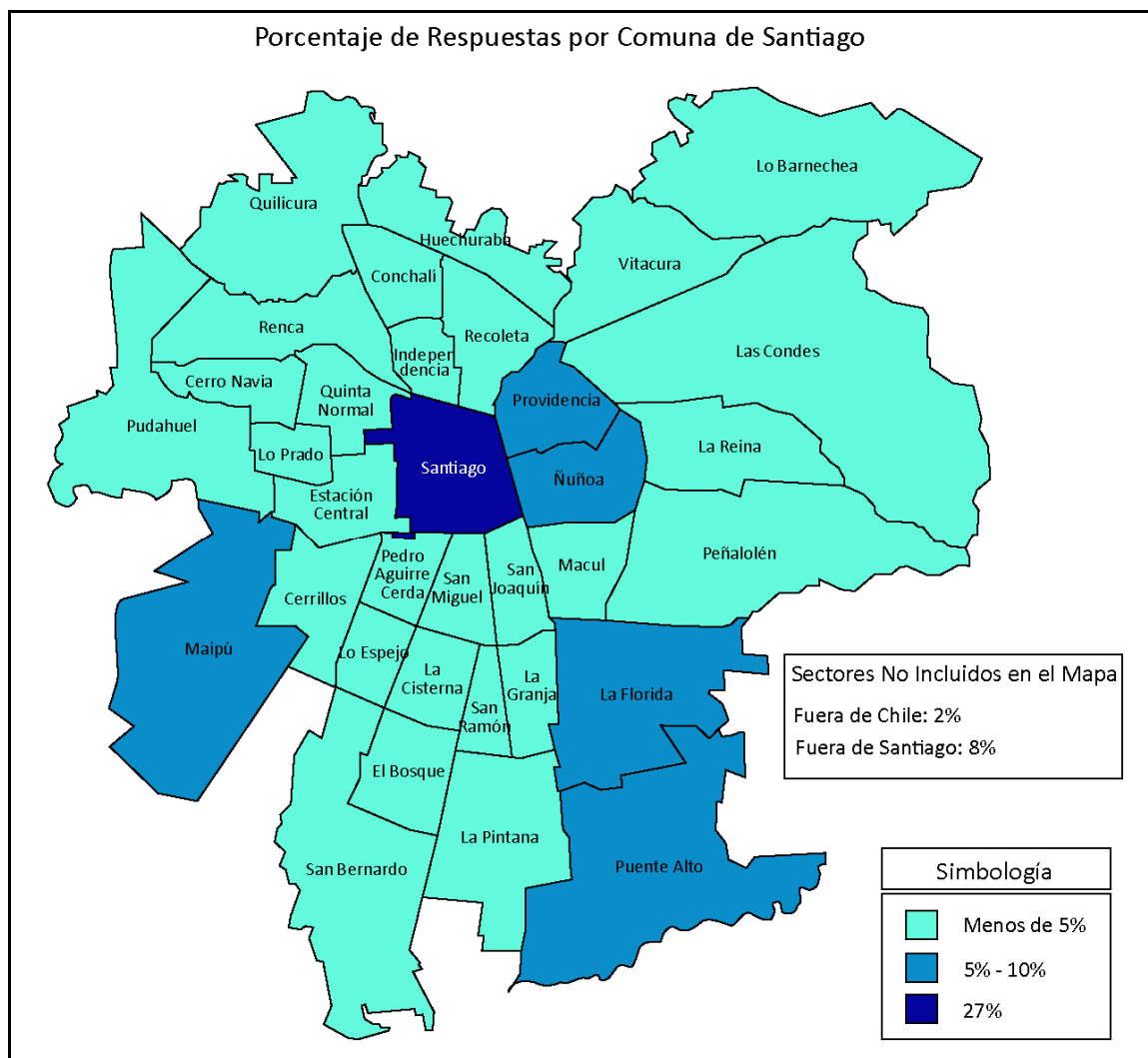


Figura 5-6: Porcentaje de Respuestas por Comuna de Origen de Encuestados

La comuna de Santiago aporta un 27% del total, seguido por las respuestas de Fuera de Santiago (8%), y de las comunas de Maipú (7%), Providencia (6%), Ñuñoa (5%), Puente Alto (5%) y La Florida (5%). Resulta interesante constatar que cada comuna tiene al menos una persona entrevistada, es decir, el fenómeno de migración hacia el centro de Santiago incluye a toda la ciudad. Es posible clasificar a las comunas que cuentan con más entrevistados en dos grupos:

- Maipú, Puente Alto y La Florida: estas son las tres comunas con mayor población del país. Además, Puente Alto y Maipú son las dos comunas del país donde se forman mayor cantidad de hogares nuevos al año, y La Florida es la onceava. Por otro lado, están conformadas mayormente por población de estrato medio y medio-alto, es decir, que cuenta con el poder adquisitivo para comprar vivienda propia. En definitiva, el traslado de habitantes resulta acorde a la gran cantidad de residentes en ellas y su estrato socioeconómico, en especial si se considera que la comuna de Santiago cuenta el mayor número de viviendas nuevas con permiso de edificación otorgado.
- Santiago, Providencia y Ñuñoa: estas tres comunas se encuentran dentro de las diez con mayores ingresos por hogar, con mayor porcentaje de población con título universitario y con mayor Índice de Desarrollo Humano⁷ del país. Finalmente, considerando todas las características en común y el hecho de que son vecinas directas, no debiera resultar extraña la transferencia de población entre ellas⁸.

⁷ Índice compuesto que contiene tres variables que definen la situación de salud, educación e ingreso de las personas.

⁸ Todos los indicadores comunales se obtuvieron de la página web www.observatoriourbano.cl, el día 14 de Julio de 2008.

El *Ingreso_Personal* reportado por los encuestados tiene una media de \$636.222 y una desviación estándar de \$550.246. Los valores mínimo y máximo entregados fueron \$0 y \$7.000.000 respectivamente.

En la Figura 5-7 se muestra el histograma del *Ingreso_Personal*, en miles de pesos chilenos. La mayor parte de la muestra se concentra en valores menores a \$1.200.000, siendo el rango entre \$200.000 y \$800.000 el que tiene mayor cantidad de observaciones (74,2% del total). La distribución es asimétrica, con una cola larga hacia los valores altos, por lo que se realizó un análisis de *outliers*.

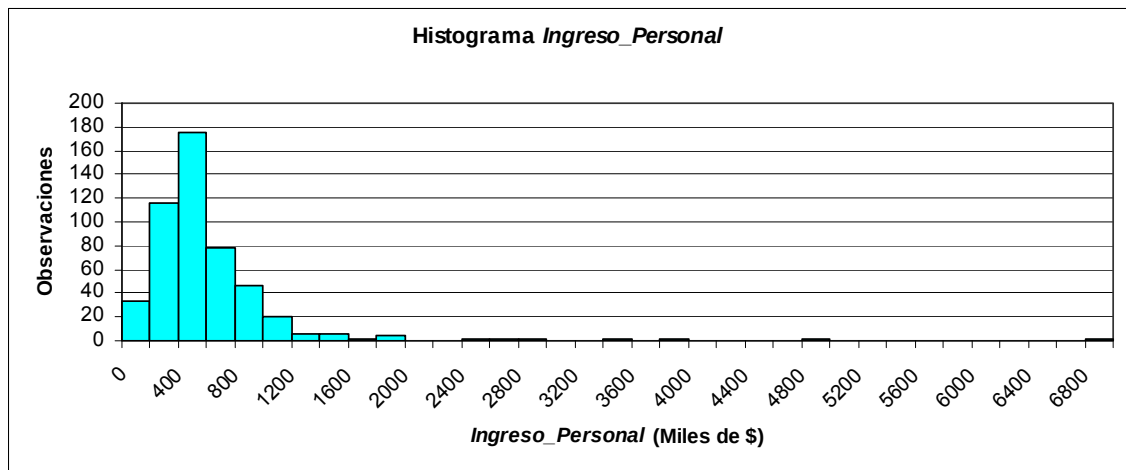


Figura 5-7: Histograma de la Variable *Ingreso_Personal*

El método utilizado es el de gráfico de caja o *Box Plot* (Walfish, 2006), útil para datos con cualquier distribución. Los *outliers* leves (O_L) y *outliers* extremos (O_E) se definen a partir del primer cuartil (q_1), el tercer cuartil (q_3) y el rango intercuartil ($q_3 - q_1$). Se fija un límite externo a partir de dos puntos: $q_1 - 3 \cdot (q_3 - q_1)$ y $q_3 + 3 \cdot (q_3 - q_1)$; y un límite interno que se define con: $q_1 - 1,5 \cdot (q_3 - q_1)$ y $q_3 + 1,5 \cdot (q_3 - q_1)$. Luego, se consideran como O_E a aquellas observaciones que queden fuera del límite externo, y como O_L a las

que se encuentran entre el límite interno y el externo. En la Tabla 5-2 se muestran los límites calculados.

Tabla 5-2: Análisis de *Outliers* de la Variable *Ingreso_Personal*

Estadígrafos	Valor	
Primer Cuartil (q_1)	360.000	
Tercer Cuartil (q_3)	720.000	
Rango Intercuartil ($q_3 - q_1$)	360.000	
Límite <i>Outlier</i> Leve (O_L)	Inferior: -180.000	Superior: 1.260.000
Límite <i>Outlier</i> Extremo (O_E)	Inferior: - 720.000	Superior: 1.800.000

El análisis arrojó 13 *outliers* leves y 14 extremos, todos en la cola de la derecha. Resulta interesante constatar que siete de los *outliers* extremos corresponden a personas provenientes de fuera de Santiago. El análisis de *outliers* es relevante para la caracterización de la nueva población que habitará la comuna de Santiago, que es uno de los objetivos de la investigación. Además, fue importante identificarlos para tomarlos en consideración en la estimación de modelos.

5.2. Características de la Nueva Vivienda y sus Habitantes

Los encuestados entregaron información sobre características que debieran tener o esperan que tengan sus nuevos departamentos y los habitantes de ellos. La primera que se analizó es el ingreso que tendrán en conjunto todos los residentes de la vivienda (IFAM). Esta variable tiene una media de \$1.000.056 y una desviación estándar de \$636.426. Su mínimo es \$120.000 y su máximo \$7.000.000.

El histograma (ver Figura 5-8) tiene una forma similar al de la variable *Ingreso_Personal*, en el sentido de que es asimétrico con una cola extendida hacia la

derecha. El 59,4% de las observaciones se encuentra en el rango entre \$400.000 y \$1.000.000, es decir, corresponden a hogares de clase media y media-alta.

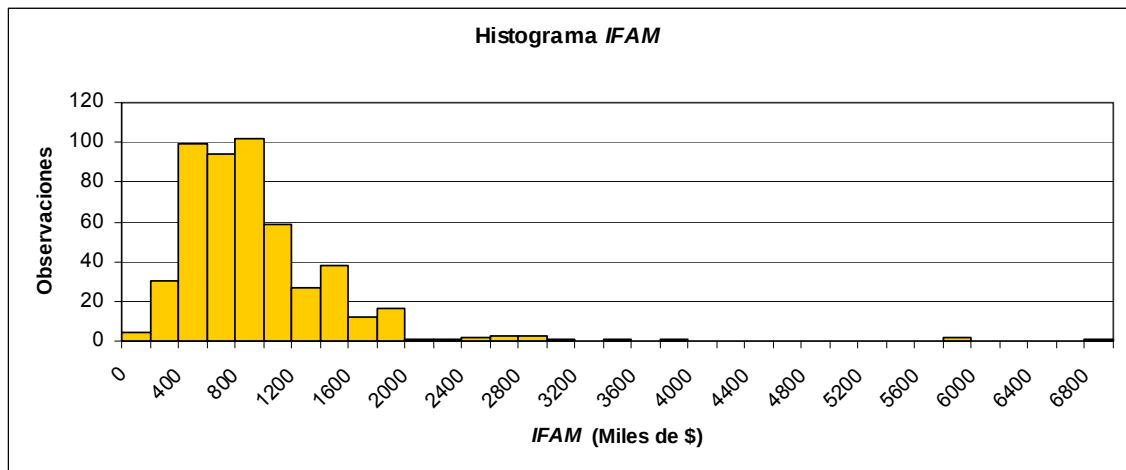


Figura 5-8: Histograma de la Variable IFAM

También se consideró necesario realizar un análisis de *outliers* en este caso debido a la presencia de valores muy altos en la cola derecha del histograma. En la Tabla 5-3 se observan los estadígrafos utilizados para este análisis, a través del cual se obtuvo siete *outliers* leves y nueve extremos.

Tabla 5-3: Análisis de Outliers de la Variable IFAM

Estadígrafos	Valor	
Primer Cuartil	600.000	
Tercer Cuartil	1.200.000	
Rango Intercuartil	600.000	
Límite <i>Outlier</i> Leve	Inferior: -300.000	Superior: 2.100.000
Límite <i>Outlier</i> Extremo	Inferior: - 1.200.00	Superior: 3.000.000

Los nueve *outliers* extremos de esta variable son coincidentes con siete extremos y uno leve de la variable *Ingreso_Personal*, mientras que los siete *outliers* leves coinciden con dos extremos y uno leve de dicha variable. El interés por analizar los *outliers* de esta variable es análogo al del *Ingreso_Personal*, es decir, se relaciona con la caracterización de la población y la estimación de modelos.

En la Figura 5-9 se presenta un gráfico de dispersión entre las dos variables de ingreso, las cuales se encuentran bastante correlacionadas (*Ingreso_Personal* es el límite inferior de *IFAM*). Es esperable que se obtengan mejores resultados en los modelos al utilizar el *IFAM* en lugar que el *Ingreso_Personal*.

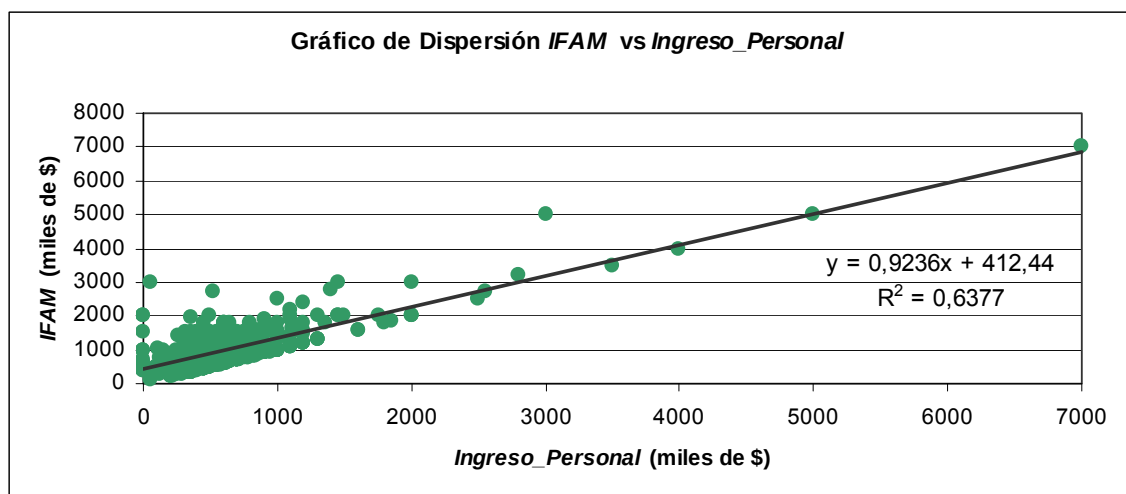


Figura 5-9: Gráfico de Dispersión entre IFAM e Ingreso_Personal

La variable *N°_Autos* presenta la distribución que se muestra en la Figura 5-10. Casi un 60% de los hogares tiene un automóvil, y ninguno de los encuestados reportó tres o más vehículos. Es de esperar que esta variable influya en el interés de la persona por encontrarse cerca (o lejos) de una estación de metro. Los hogares sin automóvil debieran presentar una disposición al pago mayor por accesibilidad a estaciones de metro que aquellos con auto.

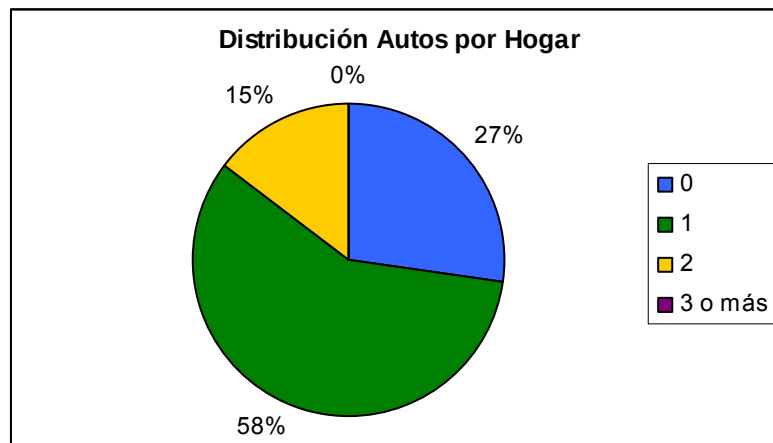


Figura 5-10: Distribución de Variable N° _Autos

La distribución de la variable N° _Habitantes, correspondiente a la cantidad de habitantes por hogar, se presenta en la Figura 5-11. La primera mayoría se da para dos personas, seguida de tres y luego un habitante. Sólo un 15% de las nuevas viviendas alojará cuatro o más habitantes, siendo siete personas el máximo reportado (dos observaciones).

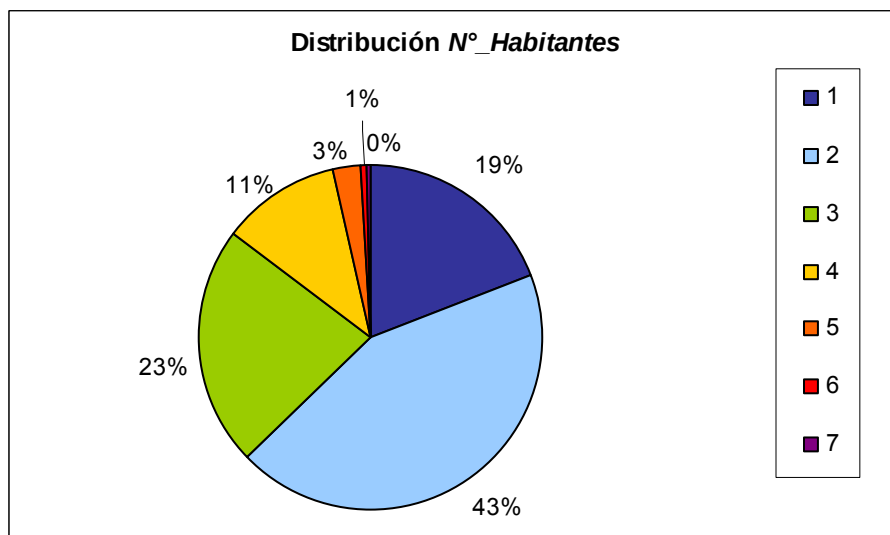


Figura 5-11: Distribución Porcentual de Variable N° _Habitantes

En cuanto a la cantidad de dormitorios requeridos por departamento, la distribución se observa en la Figura 5-12. La cantidad más frecuentemente señalada es dos, seguida por cerca de tres y luego un dormitorio. Las opciones de cuatro dormitorios y ambiente único fueron escogidas por un porcentaje bajo de la muestra. Cabe recordar que el diseño eficiente fue buscado a partir del dividendo para dos dormitorios, asumiendo que este sería el más requerido, lo que resultó cierto.

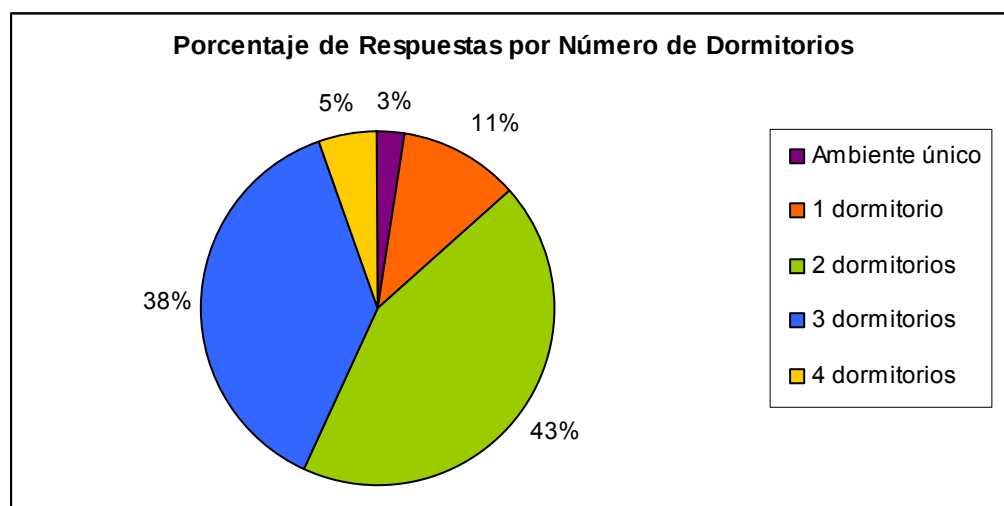


Figura 5-12: Porcentaje de Respuestas según Número de Dormitorios

Si ahora se analiza esta variable, pero separando la muestra según la cantidad de habitantes que vivirían en los departamentos, se obtienen las distribuciones de la Figura 5-13. Se puede ver que la opción de un dormitorio es casi inexistente para observaciones con N° *Habitantes* mayor o igual a tres. Sin embargo, llama la atención que el porcentaje de respuestas que opta por ambiente único es casi igual en ambos grupos de observaciones. Para el caso de uno o dos habitantes, la respuesta más común fue dos dormitorios (57%), mientras que para tres o más personas se optó mayormente por tres dormitorios (65%). Esta variable no se encuentra correlacionada de manera importante con ninguna otra.

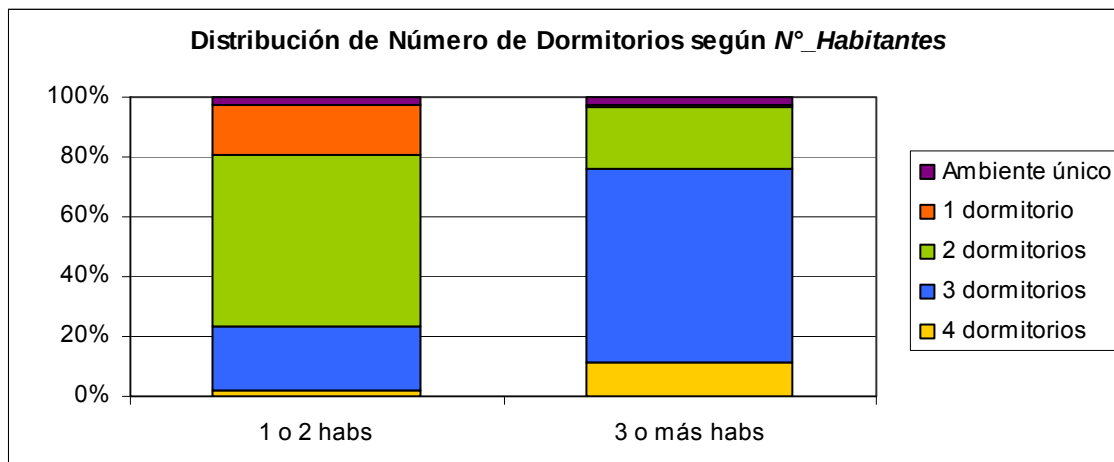


Figura 5-13: Porcentaje de Número de Dormitorios según N°_Habitantes

Si bien puede parecer contra intuitivo, no se consideró como observaciones erróneas aquellas familias de tres o más personas que escogieron uno o dos dormitorios. Esto se podría dar, por ejemplo, para parejas con hijos pequeños que desean vivir sólo un par de años en el departamento y luego arrendarlo.

Finalmente, la última pregunta hecha a los encuestados se refiere a la cantidad de años que creen que vivirían en el nuevo departamento (*Años*). Se obtuvo que casi un quinto de los encuestados desearía vivir menos de cinco años en su departamento, dos quintos desearían vivir entre cinco y 10 años, y los dos quintos restantes, más de 10 años (ver Figura 5-14). La distribución varía si se analizan las observaciones clasificando según cantidad de habitantes en el departamento, tal como se muestra en la Figura 5-15. En el caso de uno o dos habitantes, un 25% desearía vivir menos de cinco años en su departamento, mientras que para tres o más habitantes este porcentaje disminuye a sólo un 9%. De la misma forma, el porcentaje de encuestados que desearía vivir más de 10 años es de un 60% para tres o más habitantes, considerablemente mayor al 33% para uno o dos habitantes.

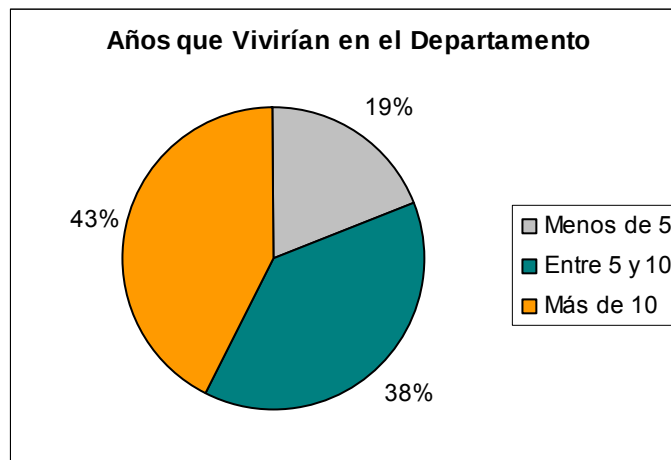


Figura 5-14: Porcentaje de Respuestas según Años

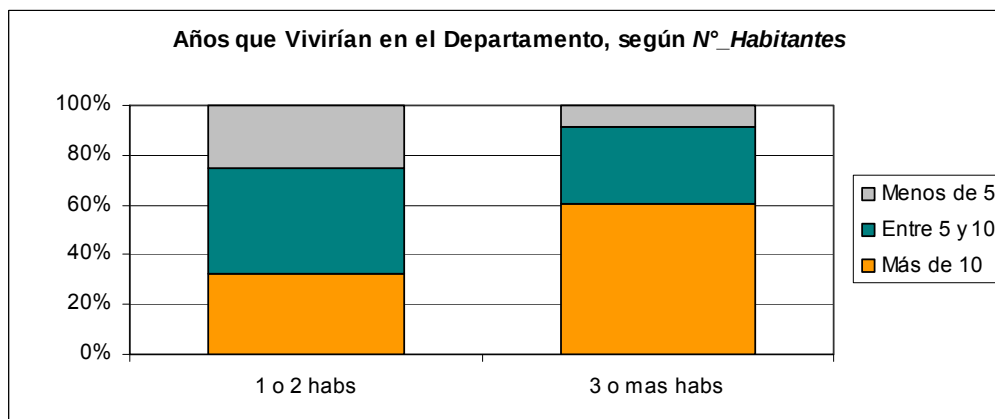


Figura 5-15: Porcentaje de Años según N°_Habitantes

Por ejemplo, es posible asumir que la mayoría de las observaciones con *N°_Habitantes* mayor o igual a tres corresponden a estructuras familiares más afianzadas, mientras que si *N°_Habitantes* es menor que tres puede corresponder a personas solteras o parejas jóvenes, que asumen cambios futuros en su estructura familiar, como por ejemplo la llegada de hijos. Esto se deduce de la estructura de la variable *Edad* de acuerdo a la cantidad de habitantes de los departamentos (Figura 5-16).

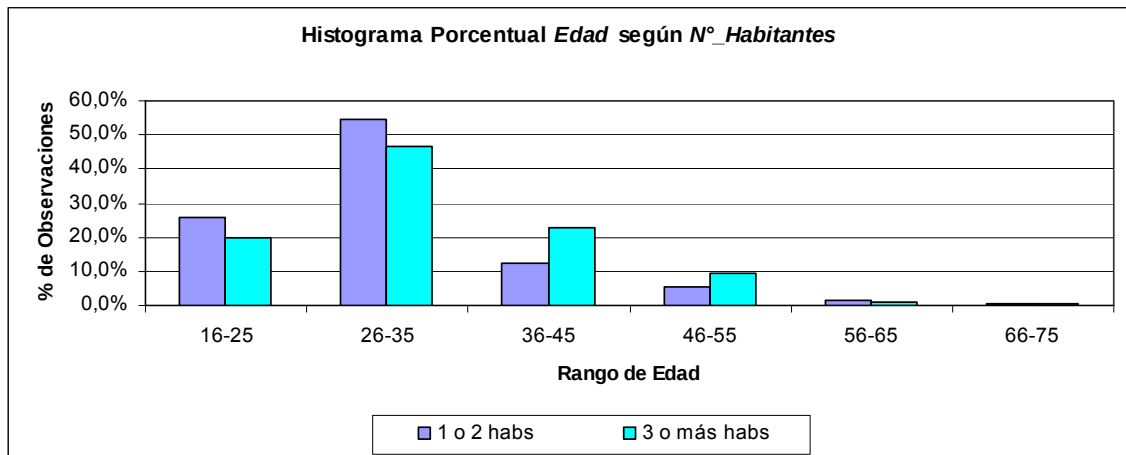


Figura 5-16: Histograma Porcentual Variable *Edad* según *N°_Habitantes*

Para uno o dos habitantes, un 81% de los encuestados tiene menos de 36 años, mientras que para tres o más habitantes dicho porcentaje disminuye a 66% en el mismo rango; también se aprecia un histograma levemente más plano, con porcentajes mayores en los rangos de edad sobre 35 años.

6. ESTIMACION DE MODELOS DE ELECCION DISCRETA

6.1. Modelos Logit Simple

La estimación de los modelos MNL se realizó a través del software BIOGEME, al igual que en las encuestas piloto. Se intentó estimar modelos con un factor de escala diferente para cada juego del experimento, pero éstos no resultaron significativamente distintos para ninguno de los modelos. Este resultado fue sorpresivo, puesto que indicó que el supuesto tomado en consideración en las encuestas piloto, de factor de escala diferente para ambos juegos, fue erróneo. Es posible que, si se hubiera optado por los modelos sin parámetro de escala distintos para los dos juegos en ambas encuestas piloto, el diseño del experimento final habría sido aun más eficiente que el obtenido. El parámetro de escala estimado se muestra sólo en el primer modelo, de manera de ilustrar los resultados obtenidos en todo el proceso.

Por otra parte, se estimó modelos con y sin aquellos datos rotulados como lexicográficos en la variable *Precio*, para identificar si éstos producían algún ruido en los valores de los parámetros. En todos los casos, la diferencia entre los valores de los parámetros estimados nunca superó el 5%, lo que probablemente se debe a que sólo 10 observaciones tenían este rótulo; debido a este resultado se resolvió no extraer las observaciones lexicográficas.

En primer lugar se estimó un modelo MNL básico (MNL1), con la función de utilidad presentada en la ecuación (4.1), al igual que en todas las encuestas piloto. Luego se modeló el efecto de variables socioeconómicas en interacción con los atributos del experimento (variaciones sistemáticas de gustos), resultando el modelo MNL2.

6.1.1. Modelo MNL1

En la Tabla 6-1 se observan los estadígrafos obtenidos para el modelo más básico. Es claro que el modelo con factor de escala diferente para cada juego no aporta información, tanto en términos de estimación de parámetros como en log-verosimilitud.

Tabla 6-1: Modelo MNL1

	MNL1		MNL1 con factor de escala	
	Parámetro	test-t	Parámetro	test-t
θ_{metro}	1,691	21,69	1,690	22,96
$\theta_{\text{areas_verdes}}$	0,597	10,72	0,598	9,36
θ_{cultura}	0,051	0,70	0,051	0,65
θ_{precio}	-0,022	-10,03	-0,022	-12,08
$\theta_{\text{espacio_comun}}$	0,373	4,84	0,373	5,10
$\theta_{\text{limpieza_calle}}$	-2,222	-21,66	-2,220	-24,60
$\theta_{\text{servicios}}$	0,796	13,37	0,796	11,96
$\theta_{\text{superficie}}$	1,021	13,60	1,020	13,89
λ_{j2}	-	-	1,040	0,16 (respecto a 1)
$l(\theta)$	-3337,85		-3337,85	
Tamaño muestra	5964		5964	
Cantidad de parámetros	8		9	

Todos los signos obtenidos para los parámetros son consistentes con los supuestos previos y con los modelos estimados con las encuestas piloto. El único atributo con parámetro no significativo es *Cultura*, pero no se extrajo del modelo por ser considerada una variable de política (necesaria), al igual que en los modelos de las encuestas piloto. Los resultados obtenidos posteriormente, con modelos más completos, ayudaron a comprender el problema con este atributo. Por otro lado, llaman la atención los altos valores absolutos de los parámetros de *Limpieza_Calle* y *Metro* en comparación con los

otros atributos. Un mayor análisis de este aspecto se realiza al calcular las disposiciones al pago por cada atributo (capítulo 7).

Dado que esta especificación de modelo fue la que se utilizó en las encuestas piloto, resulta relevante estudiar si los parámetros obtenidos son similares o diferentes a los valores utilizados como parámetros *a priori*. En las Figuras 6-1, 6-2 y 6-3 se muestran los parámetros *a priori* y los parámetros del modelo MNL1 de los atributos y del factor de escala del segundo juego. A todos se les muestra también su intervalo de confianza, puesto que resulta más adecuado realizar la comparación tomándolos en consideración.

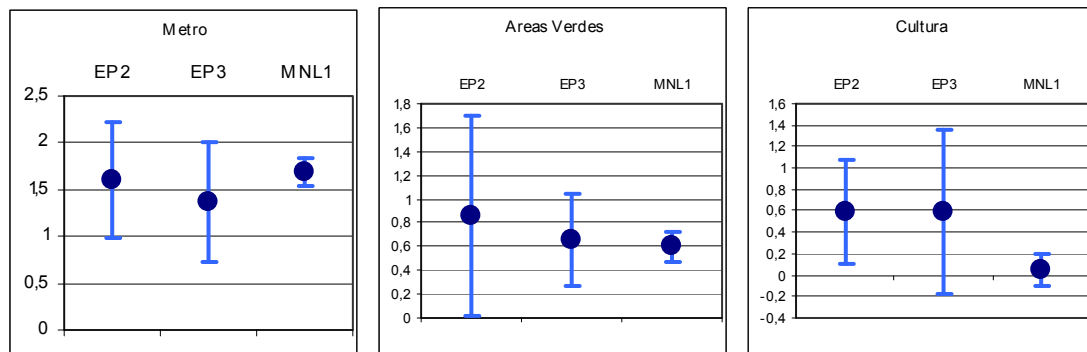


Figura 6-1: Comparación entre Parámetros *a priori* y Parámetros de MNL1

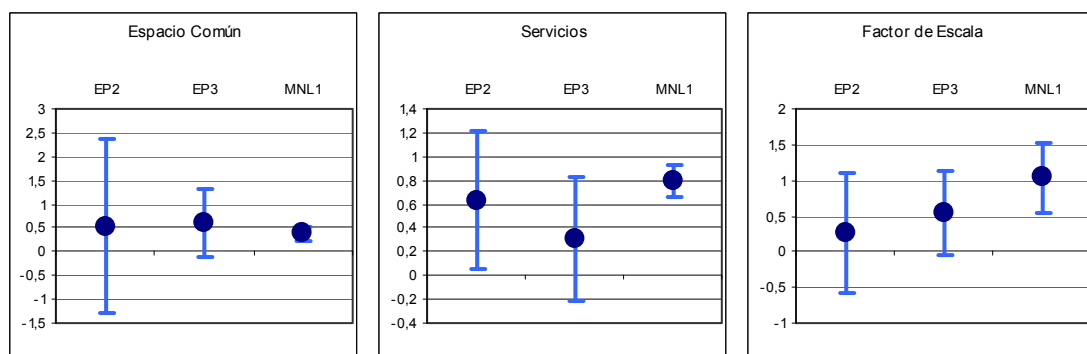


Figura 6-2: Comparación entre Parámetros *a priori* y Parámetros de MNL1

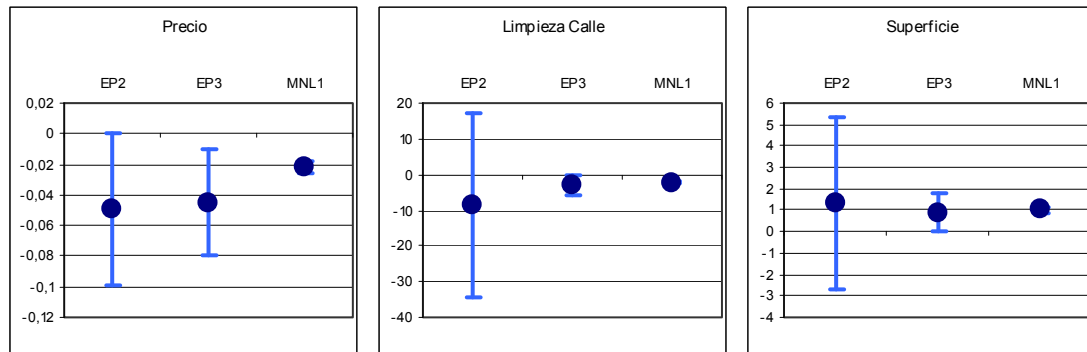


Figura 6-3: Comparación entre Parámetros *a priori* y Parámetros de MNL1

Para los atributos *Metro*, *Areas_Verdes* y *Espacio_Comun* se ve claramente cómo el intervalo de confianza se va ajustando de una encuesta piloto a la siguiente en torno al valor final que tendrían en MNL1. Para el atributo *Cultura* ello no es tan claro, puesto que el intervalo de confianza aumentó en la tercera encuesta piloto. Sin embargo, ello llevó a que incluyera valores negativos, tal como ocurrió finalmente en el modelo MNL1. Además tanto la media como el intervalo de confianza estimados con MNL1 se encuentran dentro del intervalo de confianza obtenido en la EP3, cosa que no ocurre con el de la EP2. Por otro lado, el atributo *Servicios* lamentablemente cuenta con un parámetro *a priori* más ajustado al parámetro obtenido en MNL1 con la EP2 que con la EP3. Este es el único atributo que presentó dicho problema. Finalmente, con el factor de escala del segundo juego, el intervalo de confianza del parámetro estimado en MNL1 no se encuentra contenido completamente dentro de los estimados por las encuestas piloto, pero al comparar las medias sí es clara la mejoría entre la segunda y la tercera encuesta piloto. Finalmente, en los tres atributos restantes, *Precio*, *Limpieza_Calle* y *Superficie* resulta claro que los intervalos obtenidos en la EP3 se ajustan mejor a los del modelo MNL1 que los de la EP2.

Luego, en términos generales la utilización de las encuestas piloto resultó satisfactoria para mejorar los valores de los parámetros *a priori* ocupados para diseñar el

experimento. Si bien puede resultar engañosa la utilización de los intervalos de confianza, porque las bases de dato con las que se estimaron los modelos tienen distintos tamaños, el hecho de que los parámetros de MNL1 con sus intervalos se encuentren contenidos en los de los parámetros de las encuestas piloto ya indica que se lograron mejorías en los valores previos de los parámetros.

6.1.2. Modelo MNL2

A continuación se estimaron modelos que incluyeran la interacción de las variables socioeconómicas con los distintos atributos. La estructura de la función de utilidad cambió entonces de una lineal en los atributos (ecuaciones 3.2 o 4.1), a una de la siguiente forma (Ortúzar y Willumsen, 2001):

$$(6.1) \quad V_{iq} = \sum_k \left(\theta_k + \theta_{kSE1} \cdot SE_{1q} + \theta_{kSE2} \cdot SE_{2q} + \dots + \theta_{kSEn} \cdot SE_{nq} \right) \cdot X_{ikq}$$

donde SE_{wq} corresponde al valor de la variable socioeconómica w para el individuo q . Debido al desconocimiento que se tenía sobre cómo podrían influir las variables socioeconómicas en los atributos de servicio, se optó por incluir todos los SE para cada atributo, y luego se realizó un proceso de eliminación secuencial de variables en base a distintos tests estadísticos.

Si bien la mayoría de los atributos se modelaron mediante variables mudas binarias (0 y 1), con algunas se probaron diferentes especificaciones:

- a) N°_Autos : en este caso se probaron dos especificaciones. La primera fue como variable muda binaria, donde el nivel 0 representaba no poseer automóvil; por otro lado, se modeló a través de una variable muda con tres niveles: 0, 1 auto y 2 o más autos.

- b) *IFAM*: se modeló como una variable continua, como una variable muda de dos y tres niveles y también como variable que dividía al costo (atributo *Precio*). En los casos en que se modeló como una variable continua se evaluaron modelos con y sin los *outliers* detectados en el análisis preliminar de la base de datos. Si bien se obtuvieron modelos más consistentes cuando no se incluyeron los *outliers*, de todas formas los modelos con variables mudas obtuvieron mejores estadígrafos y se prefirieron por sobre ellos. Además, se estimaron modelos con la variable *Ingreso_Personal* en lugar de *IFAM*, utilizando especificaciones similares, pero tampoco se obtuvieron resultados satisfactorios.
- c) *Edad*: esta variable se modeló como continua, como variable muda de tres y cuatro niveles, y de manera cuadrática.
- d) *Nº_Habitantes*: se modeló como variable muda de tres niveles (1, 2 y 3 o más habitantes).

Se estimaron más de 50 modelos con SE, y se escogió al MNL2 como el más informativo y con mejores estadígrafos. En el ANEXO L se detalla su función de utilidad y los parámetros estimados. Además, en la Tabla 6-2 se muestra el resto de la información sobre el modelo.

Tabla 6-2: Características del Modelo MNL2

$l(\theta)$	-3235,7
Nº de Observaciones	5964
Cantidad de parámetros	39

Al realizar el test de LR entre MNL1 y MNL2 se obtiene que los modelos no son equivalentes, es decir, se prefiere MNL2 por sobre MNL1. La formulación de MNL2 es

compleja y por ello se analizaron los resultados de cada atributo a fin de facilitar su comprensión.

a) *Precio*: en la Tabla 6-3 se encuentran los valores de los parámetros asociados a esta variable para los distintos estratos socioeconómicos en que se encontró una interacción significativa y en las figuras 6-4 y 6-5 se muestran los gráficos de estos parámetros.

El signo negativo del parámetro es consistente con la teoría microeconómica y los supuestos previos. Se observa que la magnitud del parámetro disminuye con el aumento del ingreso y con la posesión de automóvil, lo que resulta consistente con una menor aversión al pago al presentarse esas características.

Por otro lado, las comunas de Maipú, Puente Alto y La Florida presentan menores magnitudes que otras comunas, es decir, tienen una mejor disposición a pagar dividendos mayores que personas provenientes de otras comunas.

Finalmente, los varones siempre presentan parámetros de mayor magnitud, es decir, las mujeres están dispuestas a gastar una mayor cifra en dividendo en todos los casos, independiente de su comuna de origen, posesión de automóvil o *IFAM*.

Tabla 6-3: Parámetros de *Precio*, modelo MNL2

IFAM		< \$600.000		≥ \$600.000	
Comuna	Sexo	N°_Autos			
		Sin auto	Con auto	Sin auto	Con auto
Maipú, Puente Alto, La Florida	Masculino	-0,0388	-0,0336	-0,0246	-0,0194
	Femenino	-0,0315	-0,0263	-0,0173	-0,0121
Otras comunas	Masculino	-0,0448	-0,0396	-0,0305	-0,0253
	Femenino	-0,0375	-0,0322	-0,0232	-0,0180

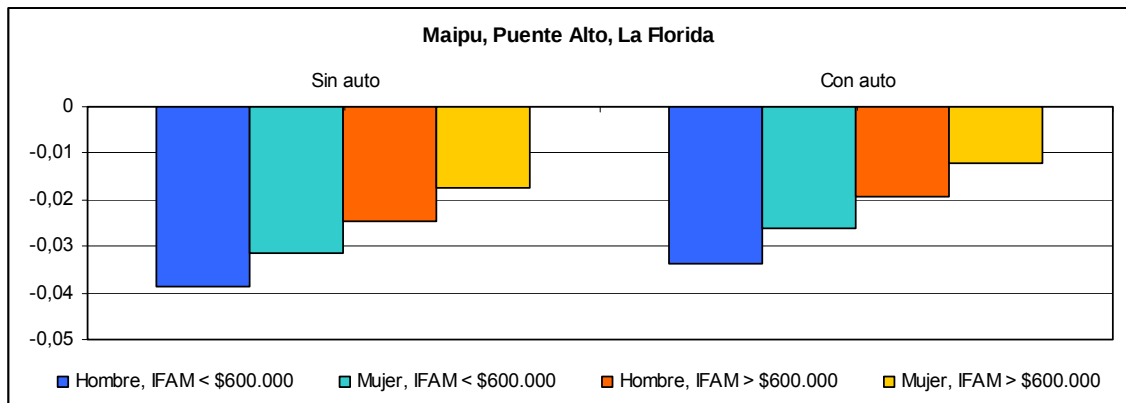


Figura 6-4: Parámetros de Precio, comunas de Maipú, Puente Alto y La Florida

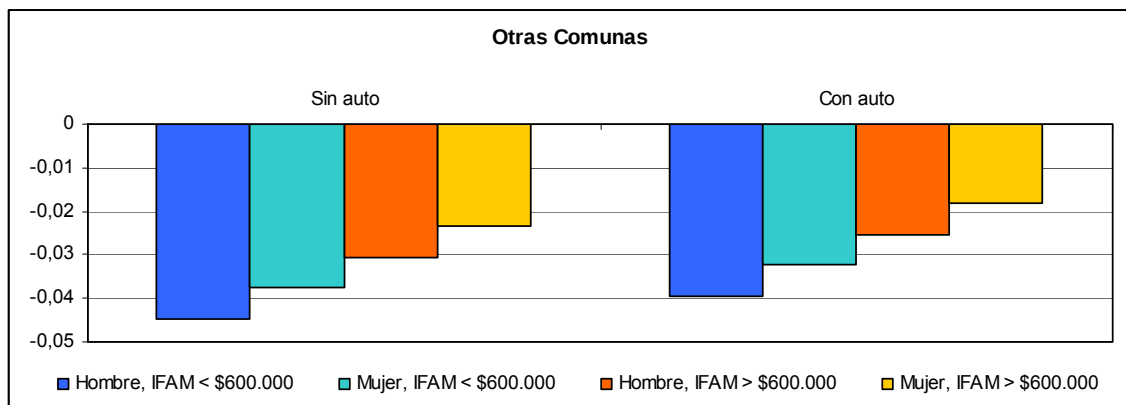


Figura 6-5: Parámetros de Precio, otras comunas

b) *Metro*: el atributo tiene parámetros positivos en todas sus interacciones con variables SE, como se puede observar en los valores entregados en la Tabla 6-4 y en la Figura 6-6. Los individuos con menor ingreso presentan parámetros más altos que aquellos encuestados con mayor *IFAM*, lo que resulta adecuado si se considera que las personas de menor ingreso utilizan más el transporte público. Por otro lado, las personas que vivirán menos de 10 años en el departamento tienen parámetros de mayor valor, lo que puede estar justificado en un interés posterior por arrendarlo, ya que la cercanía al metro implicaría mayor monto de arriendo. Se observa, además, que se estimaron

parámetros menores para los hombres que para las mujeres; es decir, ellas valoran más el acceso al metro.

Tabla 6-4: Parámetros de *Metro*, modelo MNL2

<i>IFAM</i>	< \$600.000		≥ \$600.000	
<i>Años</i>	<i>Sexo</i>			
	Masculino	Femenino	Masculino	Femenino
10 o menos años	1,8899	2,1666	1,6670	1,9436
Más de 10 años	1,6296	1,9062	1,4066	1,6832

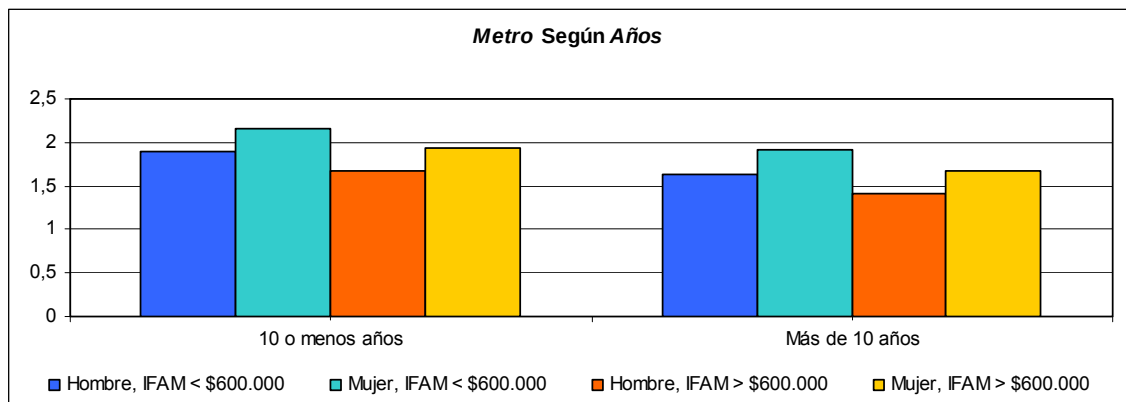


Figura 6-6: Parámetros de *Metro*

c) *Areas_Verdes*: los parámetros estimados para el atributo se muestran en la Tabla 6-5 y en los gráficos de las figuras 6-7 y 6-8. Las comunas de Maipú, Puente Alto, La Florida, Providencia y Ñuñoa tienen parámetros más altos que el resto de las comunas, de lo que se deduce que en estas comunas se otorga mayor valoración a las *Areas_Verdes*. Los encuestados que vivirán solos en sus departamentos también tienen parámetros mayores al resto de los individuos, seguramente porque las *Areas_Verdes* son un espacio para compartir con más personas.

Tabla 6-5: Parámetros de *Areas_Verdes*, modelo MNL2

IFAM		< \$1.000.000		≥ \$1.000.000	
Comuna	Nº_Habitantes	Años			
		Años < 10	Años > 10	Años < 10	Años > 10
Maipú, Puente Alto, La Florida, Providencia y Ñuñoa	Uno	0,8309	1,0020	1,1331	1,3041
	Más de uno	0,5925	0,7636	0,8947	1,0658
Otras comunas	Uno	0,5530	0,7240	0,8551	1,0262
	Más de uno	0,3146	0,4857	0,6168	0,7878

Por otro lado, las personas con *IFAM* superior a \$1.000.000 y aquellos que vivirán en el departamento más de 10 años, presentan tendencia a parámetros más altos. Es posible que en ambos casos los encuestados consideren no sólo los aspectos más básicos de los departamentos, sino también otros ya sea por que pueden pagar por ellos o por la cantidad de tiempo que habitarán ahí.

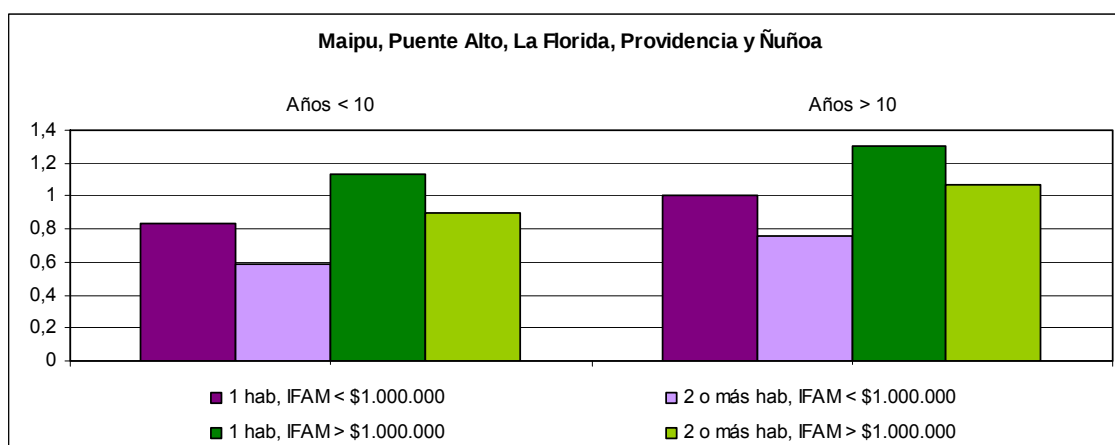


Figura 6-7: Parámetros de *Areas_Verdes*, en Maipú, Puente Alto, La Florida, Providencia y

Ñuñoa

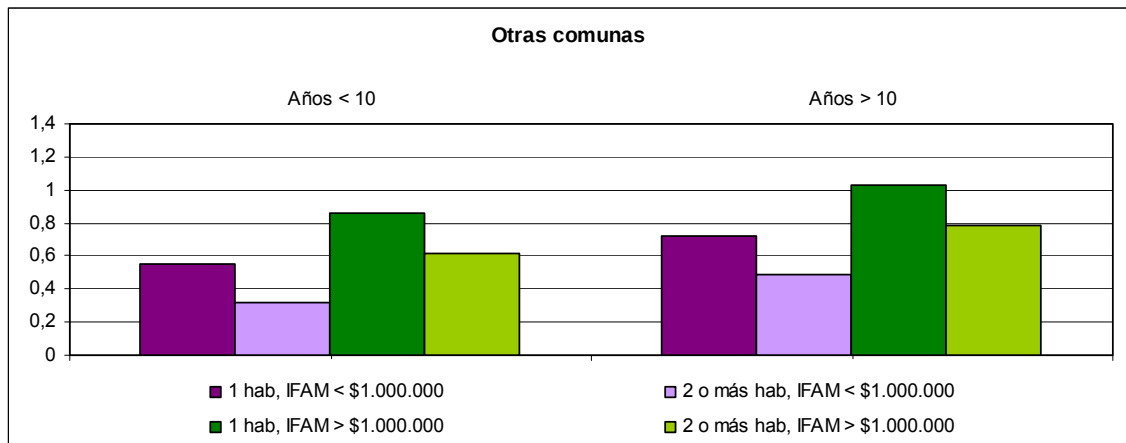


Figura 6-8: Parámetros de *Areas_Verdes*, otras comunas

d) *Cultura*: este atributo interactúa con muchas variables socioeconómicas en el modelo, y debido a la longitud de los resultados las tablas con los valores de los parámetros se adjuntan en el ANEXO M. Los resultados gráficos se pueden observar en las figuras 6-9, 6-10 y 6-11.

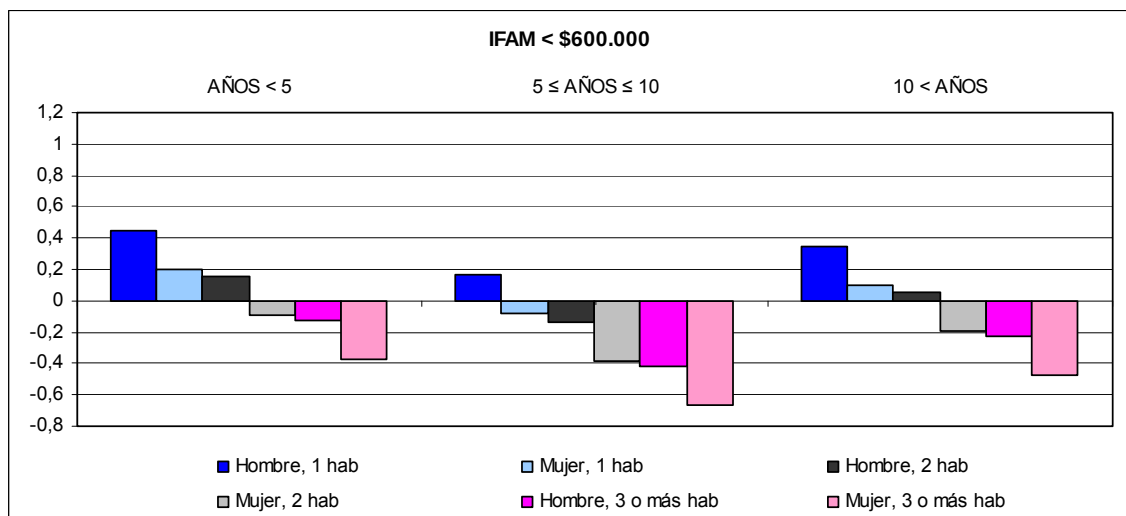


Figura 6-9: Parámetros de *Cultura*, IFAM < \$600.000

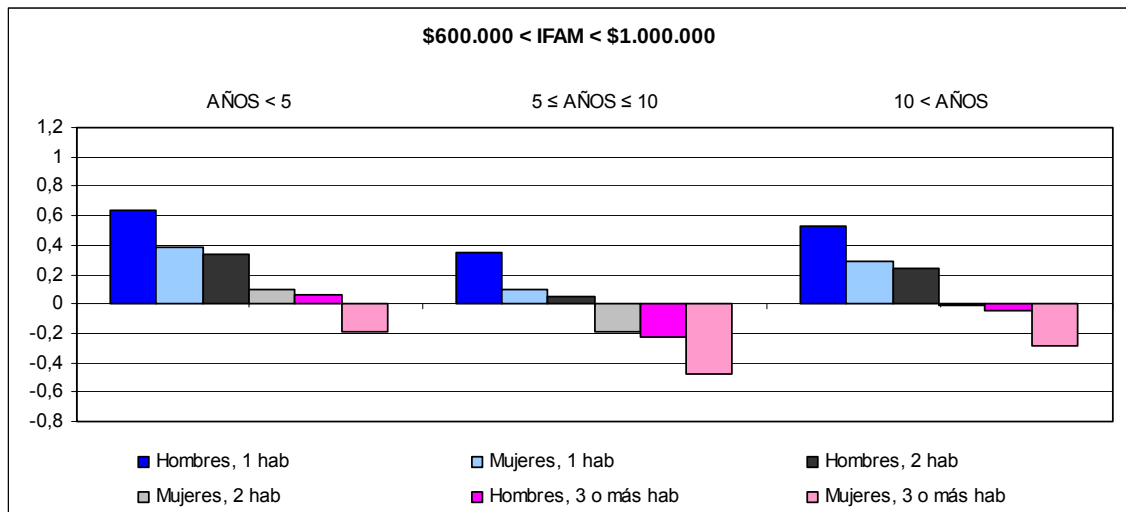


Figura 6-10: Parámetros de Cultura, \$600.000 < IFAM < \$1.000.000

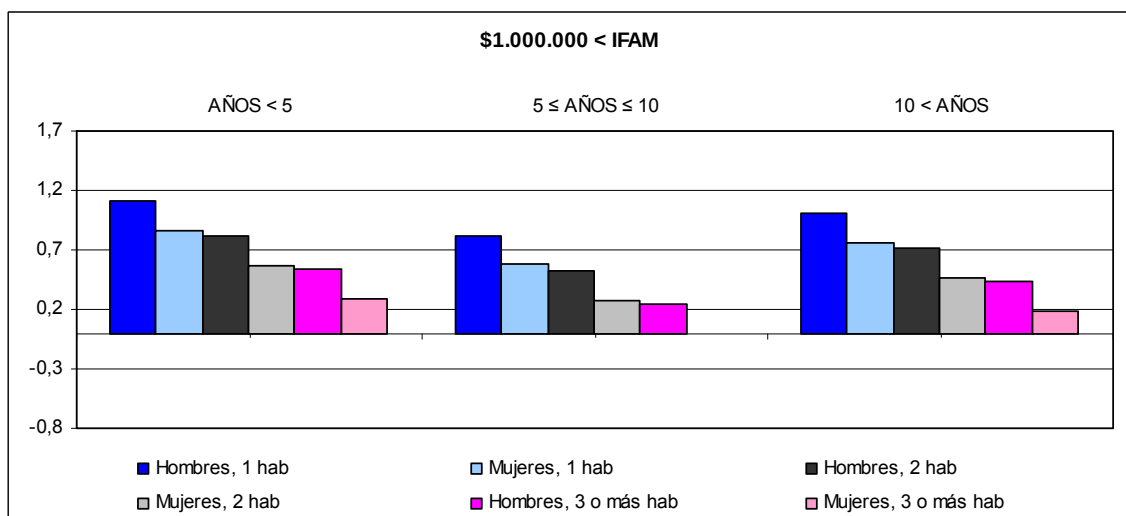


Figura 6-11: Parámetros de Cultura, \$1.000.000 < IFAM

Este atributo tuvo un comportamiento particular durante toda la investigación. Ya durante las entrevistas en profundidad se observó que algunos individuos lo consideraban relevante, y otros preferían no contar con este tipo de servicio cerca de sus hogares.

Luego, en el modelo estimado en la segunda encuesta piloto y en el MNL1 resultó no significativamente distinto de cero. Los resultados aquí obtenidos pueden arrojar luz sobre este hecho: en efecto, se observa que el parámetro toma signo positivo para algunos individuos y negativo para otros, muchos de los cuales además son cercanos a cero. Esta particularidad provoca que el parámetro estimado para toda la muestra tenga una media muy cercana a cero.

Al analizar en detalle los resultados se observa, en primer lugar, la tendencia que tienen los valores del parámetro a aumentar su magnitud en la medida que crece el *IFAM*. De esta forma, en el tramo de $IFAM < \$600.000$ hay más valores negativos que positivos, en el segmento entre $\$600.000$ y $\$1.000.000$ hay más positivos que negativos y en el tramo superior sólo se estimaron valores positivos. Se puede asumir que existe una correlación entre estas dos características, pero no se puede determinar cuál es la causa y cuál el efecto, o quizás la existencia de una tercera variable que las relaciona. Por otro lado, la mujer obtuvo parámetros de menor magnitud que el hombre en todos los casos; es decir, pareciera valorar menos el acceso a servicios culturales y de entretenimiento que el hombre. Así mismo, de manera similar a lo ocurrido con el atributo *Areas_Verdes*, mientras menor es el tamaño del hogar, más se prefiere la cercanía a servicios de entretenimiento y cultura. Finalmente, en la interacción con la variable *Años* no existe continuidad entre los tres tramos: los mayores parámetros son para $Años < 5$, luego para $10 < Años$ y finalmente para $5 \leq Años \leq 10$. El fundamento puede ser que las personas que vivirán muy pocos años estiman que el acceso a *Cultura* les traerá beneficios al arrendar el departamento y los que vivirán más de 10 años consideran atractivo el acceso a dichos servicios en el largo plazo.

e) *Servicios*: en la Tabla 6-6 y en la Figura 6-12, se observan los parámetros estimados para este atributo, los cuales cuentan adecuadamente con signo positivo. Las personas procedentes de fuera de Santiago valoran en mayor medida la accesibilidad a este servicio, seguidas de aquellas que ya vivían en la comuna de Santiago o provienen

de Providencia y Ñuñoa, las cuales seguramente están acostumbradas a contar con estos servicios cerca. Por otro lado, se observa que aquellos encuestados que no contarán con vehículo en el hogar tienen parámetros de mayor magnitud; esto tiene sentido si se considera, por ejemplo, que cargar mercadería de almacén o verdulerías resulta mucho más complejo en transporte público o caminando, que en auto.

Tabla 6-6: Parámetros de Servicios, modelo MNL2

<i>Comuna</i>	<i>N°_Autos</i>	
	Sin auto	Con auto
Providencia, Ñuñoa, Santiago	1,1551	0,9589
Fuera de la ciudad de Santiago	1,3084	1,1123
Otras comunas	0,8467	0,6505

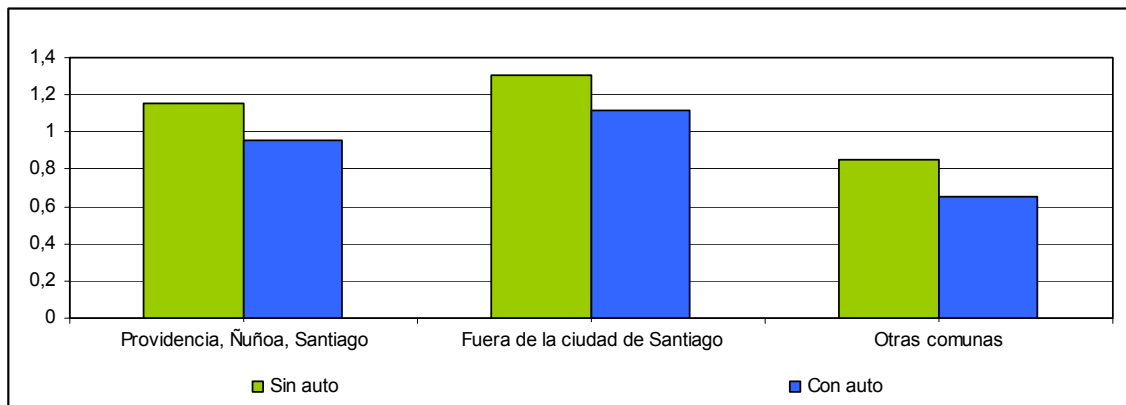


Figura 6-12: Parámetros de Servicios

f) *Espacio_Común*: si bien este atributo resultó significativamente distinto de cero en el modelo MNL1, es el que tuvo test-t más bajo a continuación de *Cultura*. Además, en las entrevistas en profundidad algunas personas lo consideraron favorable, mientras que otras no, ya que temían darle poco uso y que significara mayores gastos comunes.

En base a esto, los resultados que se observan en la Tabla 6-7 y en los gráficos de las figuras 6-13, 6-14 y 6-15 parecen consistentes. El parámetro toma valores cercanos a cero en algunos casos y en uno en particular (decimal) resulta negativo.

Tabla 6-7: Parámetros de *Espacio_Común*, modelo MNL2

<i>IFAM</i>		< \$600.000		\$600.000 ≤ x < \$1.000.000		\$1.000.000 ≤	
<i>Comuna</i>	<i>Sexo</i>	<i>N°_Habitantes</i>					
		Dos	otro	Dos	otro	Dos	otro
Fuera de la ciudad de Santiago	Masculino	0,3555	0,6080	0,6355	0,8881	0,9054	1,1579
	Femenino	0,1371	0,3896	0,4171	0,6697	0,6869	0,9395
Comuna de Santiago	Masculino	0,3088	0,5613	0,5888	0,8414	0,8587	1,1112
	Femenino	0,0904	0,3429	0,3704	0,6230	0,6402	0,8928
otras comunas	Masculino	0,0073	0,2598	0,2873	0,5399	0,5571	0,8097
	Femenino	-0,2112	0,0414	0,0689	0,3214	0,3387	0,5913

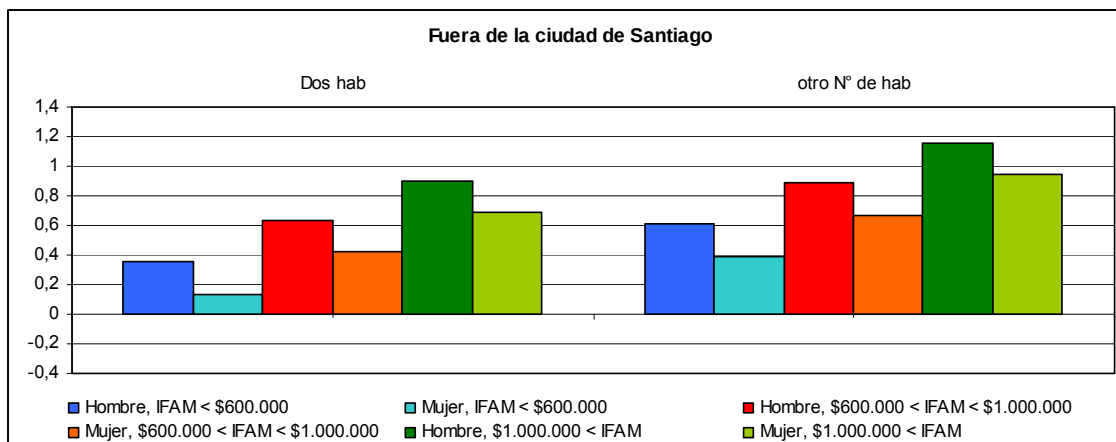


Figura 6-13: Parámetros de *Espacio_Común*, comunas fuera de la ciudad de Santiago

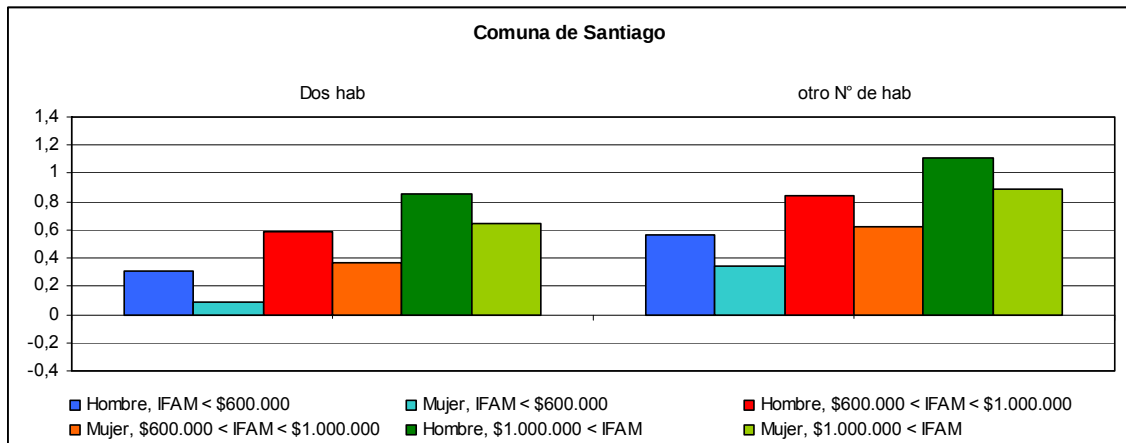


Figura 6-14: Parámetros de *Espacio Común*, comuna de Santiago

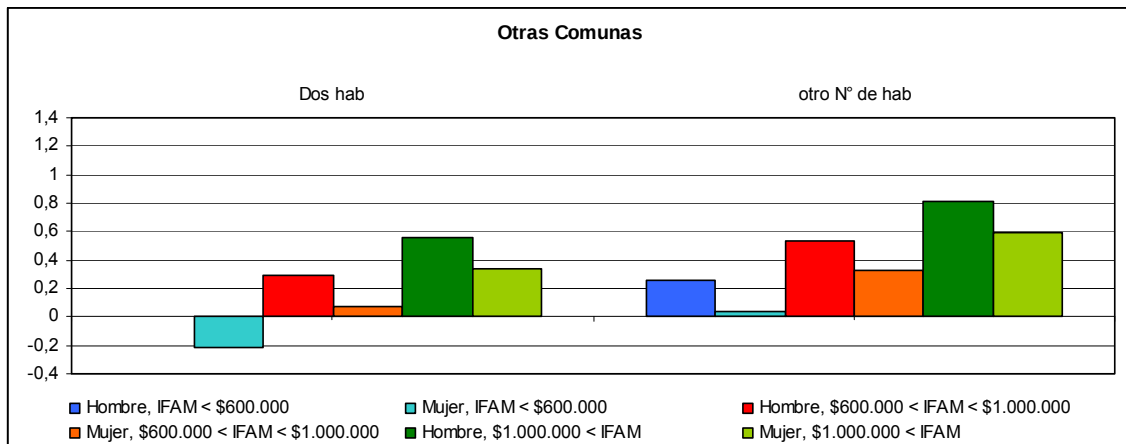


Figura 6-15: Parámetros de *Espacio Común*, otras comunas

Los valores más altos corresponden a personas procedentes de fuera de la ciudad de Santiago, luego los de la comuna de Santiago y a continuación el resto de las comunas. Esto resulta interesante para las empresas constructoras e inmobiliarias, puesto que, al menos dentro del universo de personas encuestadas, la mayoría era originaria de la comuna de Santiago y luego de comunas fuera de la ciudad. De esta forma, en esta

variable es donde resulta más clara una diferencia entre antiguos y nuevos habitantes de la comuna de Santiago.

Por otro lado, las mujeres tienen parámetros menores que los hombres en esta variable, indicando que les parece un servicio menos relevante en sus departamentos.

En la medida que aumenta el *IFAM* también crece la magnitud de los parámetros, lo que nuevamente se puede relacionar como una tendencia de las personas con mayores ingresos a interesarse por más aspectos que sólo los básicos con que debe contar una vivienda.

En cuanto a la cantidad de habitantes por departamento, se presentó una situación curiosa, puesto que aquellos con dos habitantes presentan parámetros con valor menor que los que tienen uno, y tres o más habitantes. Se puede postular que en aquellos departamentos con un habitante el *Espacio_Común* es valorado como una instancia para compartir y relacionarse con otras personas, mientras que en aquellos con tres o más habitantes, el *Espacio_Común* puede ser utilizado por los niños, puesto que una familia de padres e hijos constituye la estructura familiar más común para esta cantidad de habitantes.

g) *Limpieza_Calle*: esta variable debiera tener parámetros de signo negativo, puesto que la variable muda que vale uno representa suciedad y mal estado de la calle. Esta condición se cumple para todos los parámetros estimados (ver Tabla 6-8 y Figura 6-16).

Tabla 6-8: Parámetros de *Limpieza_Calle*, modelo MNL2

<i>Años</i>	<i>Edad</i>	
	< 36	36 ≤
10 o menos años	-2,3910	-2,7720
Más de 10 años	-2,0628	-2,4438

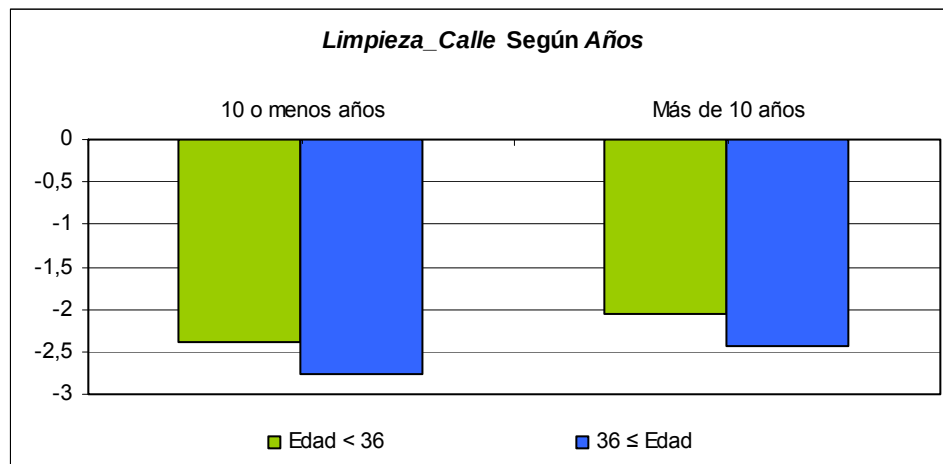


Figura 6-16: Parámetros de Limpieza_Calle

Las personas menores de 36 años tienen parámetros de menor magnitud, lo que resulta consecuente con que los jóvenes se vean menos afectados por este atributo que las personas de mayor edad. Esta fue la única variable donde se apreció un comportamiento diferente según el rango de edad de los encuestados. En cuanto al tiempo que vivirán las personas en el departamento, aquellas que lo harán más de 10 años tienen magnitudes menores que el resto. Esto se puede interpretar como una confianza en que en el largo plazo se mejorará este aspecto en el barrio, tarea correspondiente a la municipalidad.

h) *Superficie*: los parámetros obtenidos para este último atributo se muestran en la Tabla 6-9 y en la Figura 6-17, y cuentan adecuadamente con signo positivo.

Tabla 6-9: Parámetros de Superficie, modelo MNL2

Comuna	Maipú, Puente Alto, La Florida		Otras comunas	
Años	N°_Habitantes			
	Uno o dos	tres o más	Uno o dos	tres o más
5 o menos años	1,1052	0,8851	1,5529	1,3328
Más de 5 años	0,7436	0,5236	1,1913	0,9712

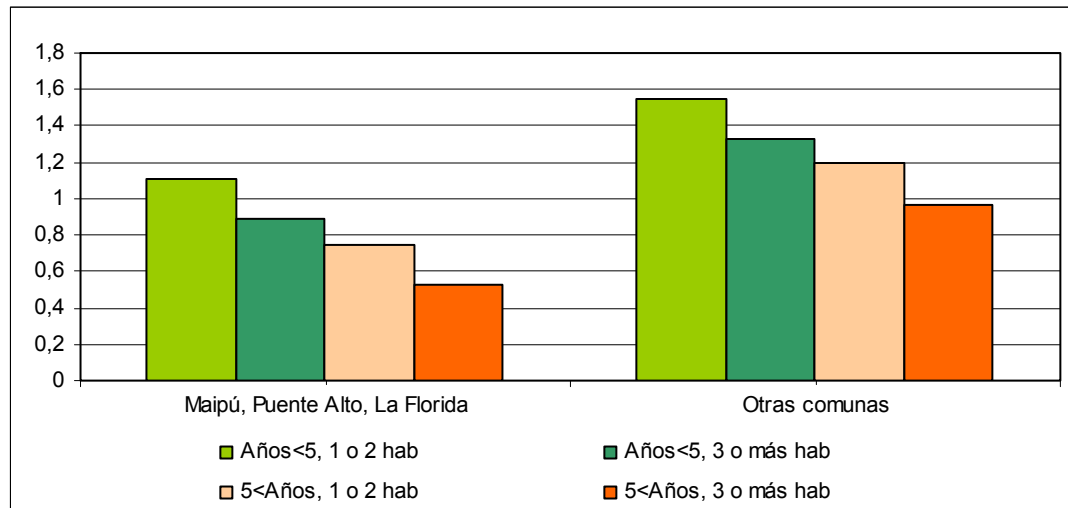


Figura 6-17: Parámetros de Superficie

El comportamiento de los parámetros respecto a las comunas de origen es consistente con lo observado en *Areas_Verdes*; esto es, en Maipú, Puente Alto y La Florida se valorarían menos los espacios interiores y más los exteriores. Por ello los valores de los parámetros estimados resultan menores en esas comunas. Por otro lado, los encuestados que residirán menos de cinco años tienen mayor tendencia a preferir espacios amplios, lo que nuevamente se asume que traería beneficios al momento de arrendar el departamento. Finalmente, para los departamentos con uno o dos habitantes se obtuvieron parámetros con mayor valor que en aquellos donde residirían tres o más personas. La justificación de ello sería que los individuos que no tendrán hijos valorarían más contar con un espacio propio de mayor tamaño, mientras que los que sí tienen hijos deben considerar muchos otros aspectos, por lo que estarían dispuestos a ceder en espacio.

6.2. Modelo Logit Mixto

Se decidió que era necesario estimar modelos ML principalmente por dos razones: en primer lugar, porque en los modelos MNL se asumió que las observaciones son

independientes, pero en la realidad existe una correlación entre todas las respuestas de cada individuo; en segundo lugar, porque el análisis realizado con el modelo MNL2 sugiere que podría haber variación en los gustos más allá de los detectados de manera sistemática.

La estimación de los modelos ML se realizó a través del código de programación desarrollado en Gauss por Train *et al.* (1999). Los primeros modelos que se intentaron estimar fueron dos con las mismas funciones de utilidad que los modelos MNL1 y MNL2, pero agregando un componente de error (EC) en cada uno que incluyera el efecto panel, es decir, que relacionara todas las respuestas de un mismo individuo a través de un error aleatorio. Se utilizaron para ello tanto distribuciones uniformes como triangulares para el EC. Es muy interesante señalar que en todos los modelos planteados se obtuvieron parámetros no significativamente distintos de cero para la desviación estándar del error. Además, no aportaban información de acuerdo al test LR. Si bien este resultado puede parecer contra intuitivo - ¿cómo es posible que en un panel de datos no se pueda modelar el efecto panel? – es importante destacar que el efecto primario que se rescata a través de EC es la heterogeneidad en la elección de alternativas respecto a una de referencia. Dado que en este experimento no se cuenta con alternativas específicas, aparentemente no es posible detectar dicha heterogeneidad. Por otro lado, se intentó modelar con EC pero sin efecto panel y se llegó a resultados análogos en cuanto a log-verosimilitud y significancia del error aleatorio. De esta forma, no se incluyen modelos con esta estructura, por no haberse obtenido resultados relevantes.

El modelo de tipo Logit Mixto estimado fue llamado ML1, y cuenta con la misma función de utilidad que el MNL1, pero esta vez con parámetros aleatorios (*random parameters* RP) con distribución Normal. Se intentó estimar dos versiones de este modelo: una tomando en cuenta el efecto panel y otra no. En términos prácticos, el modelo sin efecto panel estima un parámetro para cada una de las 5964 observaciones, mientras que el modelo con efecto panel estima un parámetro por individuo,

relacionando las 12 respuestas de cada encuestado. El modelo sin efecto panel no fue posible de estimar, ya sea con el código de Gauss como con el software BIOGEME, debido a limitaciones de la estructura de datos (Hessiano singular). Los resultados del modelo con RP considerando el efecto panel se muestran en la Tabla 6-10.

Tabla 6-10: Modelo ML1

Parámetro	$\bar{\theta}_k$	test-t	$\hat{\sigma}_k$	test-t
θ_{metro}	3,656	15,65	2,091	10,22
$\theta_{\text{areas_verdes}}$	1,201	10,19	1,047	5,37
θ_{cultura}	0,444	2,85	0,958	3,17
θ_{precio}	-0,053	-9,46	0,047	14,29
$\theta_{\text{espacio_comun}}$	0,722	4,6	1,813	7,19
$\theta_{\text{limpieza_calle}}$	-4,958	-14,45	2,577	10,05
$\theta_{\text{servicios}}$	1,878	12,78	1,12	5,02
$\theta_{\text{superficie}}$	2,4	11,8	1,455	5,64
$l(\theta)$	-2905,5			
N° de Observaciones	5964			
Individuos	497			
Cantidad de parámetros	16			

Para la estimación de este modelo se probó, al igual que para los MNL, una estimación con las observaciones rotuladas como lexicográficas y otra sin ellas. Se obtuvieron resultados análogos al caso anterior, esta vez con valores de los parámetros diferentes a lo más en un 7%; por lo tanto, nuevamente se optó por no eliminar las observaciones de carácter lexicográfico.

Resulta relevante que para este modelo todos los test-t de significancia son aprobados, a diferencia del modelo MNL1, donde el atributo *Cultura* no lo hizo. Este modelo resulta

complementario al MNL2, donde ya se había demostrado que existía variación sistemática en los gustos.

El test LR entre este modelo y MNL1 da un resultado ampliamente favorable hacia ML1, obteniéndose una disminución de 432 en la magnitud de la log-verosimilitud a expensas de sólo ocho parámetros extra. Este resultado es común, debido al elevado poder explicativo de la especificación logit mixta (Hensher, 2001). Por otro lado, se puede observar que las medias de los parámetros tienen valores bastante más altos que los obtenidos para el modelo MNL1; esto se debe a que el parámetro de escala de un modelo ML con parámetros aleatorios es mayor por construcción que el de un modelo MNL con función de utilidad análoga y parámetros fijos (Sillano y Ortúzar, 2005).

En cuanto a los signos de los parámetros, las medias cumplen con las condiciones mencionadas a lo largo de la investigación. Sin embargo, como en este caso los parámetros tienen distribución normal, cuyo dominio son todos los valores reales, resulta relevante estudiar cómo se comportan sus distribuciones. En la Tabla 6-11 se observa qué porcentaje de las distribuciones se encuentra bajo y sobre cero, calculadas como el área bajo la curva de densidad de probabilidad de los parámetros entre menos infinito y cero, y entre cero e infinito, respectivamente. Los resultados son plenamente satisfactorios, puesto que seis de los parámetros cuentan con sobre un 87% de su distribución con el signo adecuado. Los restantes dos parámetros, correspondientes a *Cultura* y *Espacio_Común*, tienen sobre 66% de sus distribuciones en la porción que corresponde. Si bien este valor no es tan alto como en el resto de los casos, es importante recordar que estas dos variables no parecen ser percibidas de manera homogénea por toda la muestra, tanto de acuerdo al modelo MNL2 como a la información que se recolectó en las entrevistas en profundidad.

Se analizó la posibilidad de estimar parámetros con funciones log-normal, de manera de asegurar que tuvieran el signo adecuado. Sin embargo, ello hubiera implicado otros

problemas, como un valor sesgado de la media o la varianza (Hess y Polak, 2004) debido a su larga cola, o problemas de convergencia, puesto que las distribuciones log-normal tienden a producir funciones de log-verosimilitud muy planas (Algers *et al*, 1999).

Tabla 6-11: Análisis de las Distribuciones de los Parámetros del Modelo ML1

	Signo esperado	% distrib. bajo cero	% distrib. sobre cero
θ_{metro}	positivo	4%	96%
$\theta_{\text{areas_verdes}}$	positivo	13%	87%
θ_{cultura}	¿positivo?	32%	68%
θ_{precio}	negativo	87%	13%
$\theta_{\text{espacio_comun}}$	¿positivo?	34%	66%
$\theta_{\text{limpieza_calle}}$	negativo	97%	3%
$\theta_{\text{servicios}}$	positivo	5%	95%
$\theta_{\text{superficie}}$	positivo	5%	95%

Los atributos *Metro* y *Limpieza_Calle* son los que tienen parámetros con medias de mayor magnitud. Resulta importante recordar que se encontró un comportamiento pseudo-lexicográfico para esas dos variables en un número importante de observaciones, lo que posiblemente afecta los resultados obtenidos.

7. DISPOSICIÓN AL PAGO POR ATRIBUTOS

Se calculó la disposición al pago (WTP) por los atributos a partir de los modelos MNL1, MNL2 y ML1. Los valores que se obtuvieron representan cuánto están dispuestos a pagar mensualmente los habitantes de los departamentos por los atributos. La unidad reportada es pesos chilenos.

El cálculo de la WTP por los atributos de acuerdo al modelo MNL1 se realizó ocupando la fórmula de test-t propuesta por Armstrong *et al.* (2001) (ver sección 3.2). En la Tabla 7-1 se muestran los valores obtenidos.

Tabla 7-1: Disposición al Pago por Atributos, Modelo MNL1

Atributo	WTP (\$)	Límite Inferior	Límite Superior
<i>Metro</i>	76.900	65.900	92.400
<i>Areas_Verdes</i>	27.100	20.800	35.600
<i>Cultura</i>	2.300	-4.800	8.000
<i>Espacio_Común</i>	17.000	11.100	22.500
<i>Limpieza_Calle</i>	-101.000	-119.500	-88.000
<i>Servicios</i>	36.200	30.900	43.000
<i>Superficie</i>	46.400	39.600	55.300

En el caso del modelo MNL2, para calcular la WTP por atributo fue necesario, en primer lugar, obtener las WTP por atributo y por tipo de usuario, según las características socioeconómicas incluidas en el modelo. Luego, se calculó un promedio ponderado por la cantidad de individuos clasificados de acuerdo a dichas características. Los resultados obtenidos se encuentran en la Tabla 7-2. Los valores son muy similares en magnitud a los del modelo MNL2, con una leve tendencia a ser mayores.

Tabla 7-2: Disposición al Pago por Atributos, Modelo MNL2

Atributo	WTP (\$)
<i>Metro</i>	79.700
<i>Areas_Verdes</i>	26.800
<i>Cultura</i>	4.500
<i>Espacio_Común</i>	19.200
<i>Limpieza_Calle</i>	-104.300
<i>Servicios</i>	37.500
<i>Superficie</i>	48.300

En cuanto a la WTP derivada del modelo ML1, se utilizaron los tres métodos descritos en la sección 3.2. Los resultados del primer método, consistente en dividir las medias poblacionales y calcular el intervalo de confianza con la Ecuación 3.16, se muestran en la Tabla 7-3.

Tabla 7-3: Disposición al Pago por Atributos, Modelo ML1, Método 1

Atributo	WTP (\$)	Límite Inferior	Límite Superior
<i>Metro</i>	69.000	58.300	84.000
<i>Areas_Verdes</i>	22.700	17.400	29.700
<i>Cultura</i>	8.400	3.000	13.100
<i>Espacio_Común</i>	13.600	8.400	19.000
<i>Limpieza_Calle</i>	-93.500	-112.900	-79.500
<i>Servicios</i>	35.400	30.200	42.400
<i>Superficie</i>	45.300	38.100	54.700

Por otro lado, los valores obtenidos del segundo método, simulación, y el tercer método, parámetros del costo fijo, se pueden ver en la Tabla 7-4. Los resultados de estos dos métodos no cuentan con intervalo de confianza como los dos primeros.

Tabla 7-4: Disposición al Pago por Atributos, Modelo ML1, Métodos 2 y 3

Atributo	Simulación		Parámetro del costo fijo	
	WTP (\$)	Desv Std	WTP (\$)	Desv Std
<i>Metro</i>	59.200	87.500	98.600	71.800
<i>Areas_Verdes</i>	19.100	32.300	30.100	22.900
<i>Cultura</i>	6.900	22.300	-7.900	64.800
<i>Espacio_Común</i>	11.100	41.700	19.100	42.300
<i>Limpieza_Calle</i>	-80.600	117.100	-118.000	67.300
<i>Servicios</i>	30.400	45.200	43.400	38.900
<i>Superficie</i>	38.800	57.900	48.700	42.500

Para facilitar el análisis, se graficaron los valores obtenidos por los cuatro métodos en la Figura 7-1. En primer lugar, resulta interesante que, si bien las medias poblacionales de los parámetros obtenidas con el modelo ML1 eran bastante mayores que las del modelo MNL1, al calcular las WTP se observa bastante estabilidad entre los valores arrojados por los dos modelos, y también por MNL2. Existen diferencias leves entre los resultados de MNL1 y ML1 utilizando el primer método. De hecho, las medias de las WTP de todos los atributos se encuentran contenidas en el intervalo de confianza del otro modelo, a excepción del atributo *Cultura*. Los valores calculados con el modelo MNL2 también se encuentran contenidos en dichos intervalos de confianza.

Para el método de simulación, se generaron 20.000 instancias de cada parámetro, cantidad que se consideró suficiente en base a los resultados obtenidos en Espino *et al.* (2006), donde a partir de 10.000 instancias la WTP se encontraba estable.

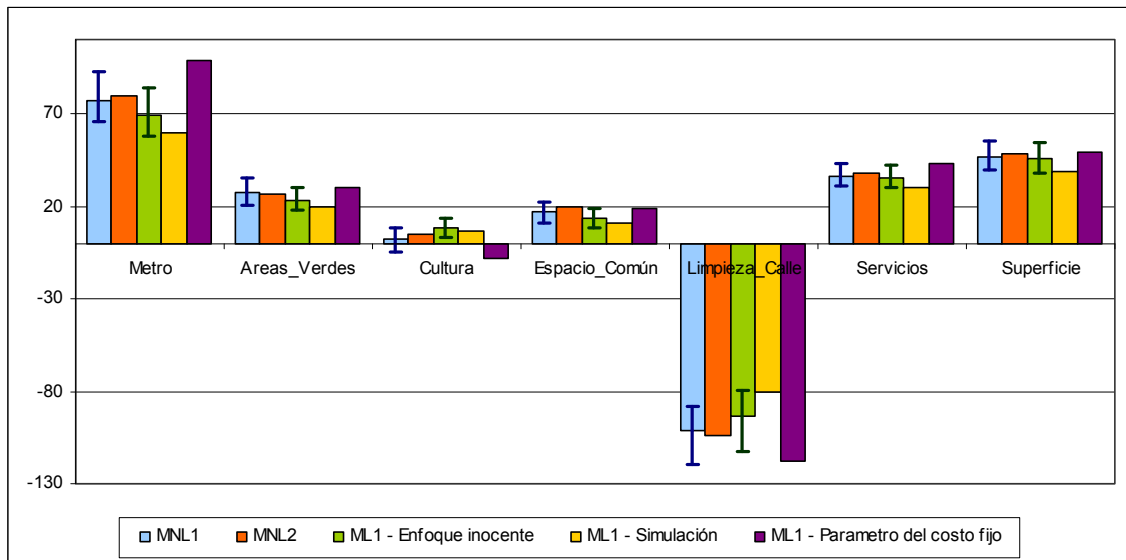


Figura 7-1: Disposición al Pago por Atributos, Comparación entre Modelos

Dado que la división entre dos normales genera distribuciones inestables (Meijer y Rouwendal, 2006), no se hicieron supuestos sobre la distribución que tendrían las WTP. Al realizar la simulación, se obtuvieron valores extremadamente grandes en magnitud, tanto positivos como negativos, por lo que las desviaciones estándar resultaron muy elevadas. Ello se debe a que las distribuciones normales tienen su dominio en todos los reales; luego, cuando se generan instancias muy cercanas a cero para el parámetro del costo o muy elevadas para los parámetros de los otros atributos, la división entre ambos tiende a infinito. Como solución a este problema, y dado que en la realidad no existen individuos con WTP tan elevadas, se optó por eliminar datos de las colas de las distribuciones de las WTP, tal como se hiciera en Sillano y Ortúzar, 2005. Un 10% de cada cola de las distribuciones fue eliminado y a partir de los 18.000 datos restantes se obtuvieron los estimadores indicados en la Tabla 7-4. Algo interesante de notar, es que todas las medias poblacionales obtenidas con este segundo método son menores en magnitud que los estimadores resultantes del primer método, pero todas se encuentran dentro del intervalo de confianza de estos últimos.

Finalmente, los resultados obtenidos con el tercer método fueron los más diferentes en comparación al resto, siendo sus magnitudes considerablemente mayores. Estos resultados no son nuevos (Algers *et al.*, 1999; Revelt y Train, 1998), y en Sillano y Ortúzar (2005) se analizó la razón de estos resultados. Previamente se mencionó que las medias estimadas de los modelos ML son mayores estructuralmente que las de los modelos MNL. Al fijar en un modelo logit mixto un parámetro que sí varía en la población, éste tendrá una media “menos mayor” que el resto de los atributos en comparación al modelo MNL equivalente. Luego, si se fija el parámetro del costo y este sí varía en la población, tal como se demostró con el modelo MNL2, su estimador será menor en magnitud que como debería ser; por lo tanto, las distribuciones de las WTP tendrán medias mayores a las reales.

Además de lo anterior, existe otro inconveniente de este método: supone que la WTP tiene la misma distribución del parámetro del atributo analizado. Como ejemplo, en la Figura 7-2 se comparan las distribuciones obtenidas para la WTP de la variable *Metro* por el segundo y el tercer método.

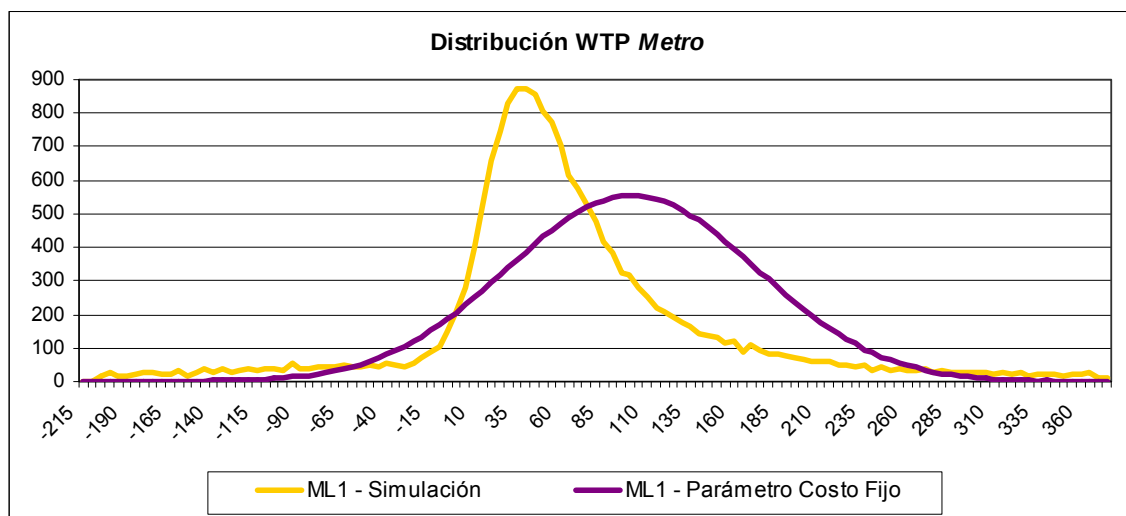


Figura 7-2: Distribuciones WTP según Método 2 y 3, Atributo *Metro*

Para el método de simulación, se observa que los valores se encuentran más concentrados en torno a su media, en comparación con el tercer método, cuya curva es más plana. Además, en el caso del segundo método, la curva desciende de manera brusca al llegar a los números negativos, pero su pendiente es bastante más suave hacia los valores positivos, lo que es coherente con la fundamentación microeconómica. Esto es diferente para el tercer método, puesto que la curva es simétrica en torno a la media.

Se graficó también las distribuciones para la WTP del atributo *Cultura* (Figura 7-3), para realizar análisis similares. En este caso, se sabía que no toda la población tenía WTP positivas, lo que es correctamente incorporado por ambas curvas. Sin embargo, al ser la distribución del parámetro de la *Cultura* muy plana, también lo es la de su WTP, obteniéndose muchas observaciones irreales en cuanto a su alta magnitud. En este aspecto, los valores de la distribución obtenida por simulación se encuentran más concentrados, de manera más coherente con los resultados obtenidos en el resto del experimento.

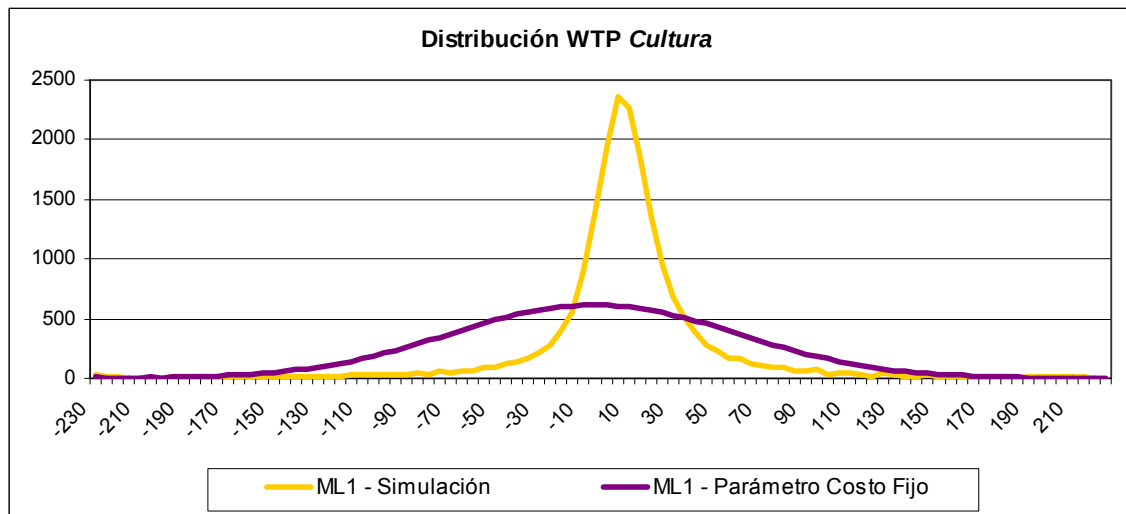


Figura 7-3: Distribuciones WTP según Métodos 2 y 3, Atributo *Cultura*

Debido a todos los aspectos antes mencionados, se consideran los valores obtenidos por el segundo método, de simulación, los más adecuados para utilizar, por sobre los derivados de los modelos MNL1 y MNL2, y los del modelo ML1 ocupando otros métodos. Considerando los valores obtenidos por simulación, es posible elaborar un ranking entre las magnitudes de las WTP de cada atributo, el cual tendría el siguiente orden: *Limpieza_Calle*, *Metro*, *Superficie*, *Servicios*, *Areas_Verdes*, *Espacio_Comun*, *Cultura*. En este orden, resulta interesante que en general tienen mayor valoración los atributos relativos a necesidades básicas y luego los ligados al esparcimiento (*Areas_Verdes*, *Espacio_Comun*, *Cultura*).

Los dos atributos que obtuvieron WTP con mayor magnitud fueron *Limpieza_Calle* y *Metro*. Justamente estas dos variables presentaron un elevado número de observaciones pseudo-lexicográficas (ver capítulo 5). Sin embargo, se cree que las razones por las cuales se presentó este comportamiento fueron diferentes. Para la *Limpieza_Calle* se presume que afectó de manera importante el hecho de que fuera la única variable presentada con una imagen en el segundo juego del experimento. Si bien se cree que era necesario incluir la variable con imagen para que fuera adecuadamente evaluada, se estima que la presentación general del segundo juego podría haber sido mejorada aun más para disminuir el efecto de atención provocado por la imagen. Por otra parte, es posible que los encuestados relacionen esta variable con otros componentes como la seguridad, y esos aspectos sean los que impliquen que tenga una WTP tan elevada. De hecho, resulta bastante irreal que una persona esté dispuesta a pagar alrededor de \$100.000 mensuales sólo para tener limpieza en la calle del edificio, pero quizás sí podría ser una inversión en seguridad.

Por otro lado, la variable *Metro* también presentó comportamiento pseudo-lexicográfico, pero en menor cantidad de observaciones que la *Limpieza_Calle*. Esta variable fue identificada como muy valiosa en las entrevistas en profundidad. De hecho, 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. Luego, se cree que la elevada WTP obtenida para la *Metro* es coherente con el comportamiento de los encuestados. En vista de lo anterior, no se cree que el valor esté sesgado.

Las WTP de los atributos *Superficie*, *Servicios*, *Areas_Verdes* y *Espacio_Comun* se consideran consistentes con la información recogida previa al experimento y dentro de rangos realistas. El atributo *Espacio_Comun*, que había presentado un parámetro en el modelo ML1 con distribución bastante plana y con más valores negativos que el resto de los atributos, tiene una WTP con media positiva según todos los métodos.

Por último, el atributo *Cultura* es el que presenta un comportamiento menos estable en su WTP. Tanto el intervalo de confianza obtenido por el modelo MNL1, como la media poblacional del método de parámetro de costo fijo del modelo ML1, tienen valores negativos. En los otros dos casos, se presenta un valor positivo. De todas formas, todos los valores obtenidos son cercanos a cero, es decir, la disposición al pago por este atributo es el más bajo dentro de la población, infiriéndose que no constituye un elemento de interés importante para los nuevos habitantes del sector.

8. CONCLUSIONES

El primer objetivo de la investigación, consistente en rescatar las preferencias de los nuevos propietarios del centro de Santiago, se logró de manera satisfactoria. La definición de los atributos y el diseño de la encuesta se llevaron a cabo a través de un proceso iterativo y de larga duración, para poder obtener un experimento que fuera tanto comprensible y rápido de contestar por los encuestados, como eficiente desde un punto de vista econométrico. Fue posible derivar disposiciones al pago por todos los atributos, a través de distintos métodos, para asegurar congruencia y robustez en los resultados.

El atributo con mayor WTP resultó ser *Limpieza_Calle*, con un valor cercano a los \$80.000 mensuales. Se cree que lamentablemente este valor puede estar sesgado por la forma en que se incluyó el atributo en la encuesta, a través de una imagen. Sin embargo, este es el único atributo con dicho problema. Por otra parte, es posible que los individuos relacionaran la limpieza con otros atributos, en especial con la seguridad. Esto también podría explicar su valoración tan elevada, al corresponder a dos atributos en lugar de uno. Además, este fue el único atributo que en el modelo MNL2 tuvo diferentes parámetros por rango etáreo, tomando menores magnitudes para personas más jóvenes. También esto podría reafirmar el hecho de que la variable se asocie a la seguridad, la cual constituye una preocupación mayor para personas de mayor edad.

La siguiente variable es *Metro* con WTP cercana a los \$60.000 mensual. Este valor puede parecer levemente elevado, pero se sustenta en la información cualitativa recopilada previamente a la encuesta. Además, se sabe que efectivamente las viviendas experimentan un aumento en su valor de compra o arriendo al construirse cerca de ellas una línea de metro, por lo que el resultado no debiera extrañar.

Los atributos que siguen en el ranking son *Superficie* y *Servicios*, con WTP de aproximadamente \$40.000 y \$30.000 mensuales respectivamente. Estos dos atributos también se detectaron como relevantes para los encuestados en la etapa de entrevistas en

profundidad. Dado que el 81% de la muestra indicó que deseaba departamentos con dos o tres dormitorios, presentados en un rango de valores entre \$160.000 y \$270.000, las WTP de estos atributos representa entre un 15% y un 20% del valor de arriendo; este valor no es pequeño, pero se cree que el rango es adecuado en relación a la información entregada verbalmente por los participantes de las entrevistas en profundidad.

Los tres atributos restantes (*Areas_Verdes*, *Espacio_Comun* y *Cultura*) tienen relación con el esparcimiento, en contraste con los antes mencionados cuyo carácter está ligado a necesidades prácticas. El que obtuvo WTP más alta fue *Areas_Verdes*, con aproximadamente \$20.000 al mes. En la comuna de Santiago existe una alta densidad de áreas verdes, por lo que resulta coherente que ellas constituyan un elemento de interés para los habitantes del sector. Finalmente, los atributos *Espacio_Comun* y *Cultura* cuentan con WTP muy bajas en contraste con el resto, con valores aproximados de \$10.000 y \$7.000 mensuales respectivamente. Este hallazgo es relevante desde dos puntos de vista: en primer lugar, una gran parte de los edificios nuevos o en construcción en el centro de Santiago cuentan con espacios comunes como gimnasio y piscina, los cuales son ampliamente publicitados por las inmobiliarias. Sin embargo, ellos no parecen constituir una necesidad de primer orden para los nuevos propietarios, e incluso se obtuvieron WTP muy cercanas a cero para este atributo en algunos tipos de individuos. De todas formas, resulta relevante que en el modelo MNL2, se encontraron parámetros mucho más elevados de esta variable para habitantes antiguos de la comuna y para personas provenientes de fuera de la ciudad de Santiago, lo que implica que sí corresponde un atributo de interés para ellos. Por otro lado el centro de Santiago cuenta con diversos barrios ligados a servicios culturales y de entretenimiento, como barrios Bellavista, Brasil, Bellas Artes, República, etc. A partir de las disposiciones al pago obtenidas por contar con este tipo de servicios cerca de los hogares, se infiere que los nuevos propietarios prefieren que sus barrios residenciales se encuentren separados de

los ligados al entretenimiento. Esto constituye información muy relevante al momento de decidir dónde construir nuevos edificios o dónde otorgar subsidios a la vivienda.

Estos resultados evidencian que estos nuevos habitantes desean habitar en barrios residenciales, donde puedan cubrir sus necesidades básicas, pero al mismo tiempo se encuentren bien conectados con servicios de entretenimiento o de cualquier otro tipo que requieran. La alta disposición al pago por encontrarse cerca del metro, acompañada de la preferencia por calles limpias y en buen estado, con servicios y áreas verdes, indican que los nuevos residentes aun quieren habitar en sectores de características residenciales, sólo que no tan alejados como serían los nuevos barrios de este tipo que se encuentran cada vez más cerca de la periferia.

En cuanto al segundo objetivo de la encuesta, correspondiente a la caracterización de los nuevos habitantes del centro de Santiago, fue lograda con éxito a partir de las preguntas de la encuesta ligadas a las características socioeconómicas. En general, la población se constituye por individuos de entre 20 y 45 años, de clase media y media-alta, con estudios secundarios y estructuras familiares pequeñas. Se detectó que el proceso de migración hacia el centro de Santiago proviene de todas las comunas de la ciudad, a pesar de que el grupo de mayor tamaño provenga de la misma comuna de Santiago. Además, alrededor de un 20% de la muestra indicó que vivirá en el departamento por menos de cinco años, lo que implica que una porción importante está adquiriendo este tipo de departamentos a modo de inversión.

En cuando al diseño mismo del experimento de PD, se recomienda la utilización de un proceso secuencial de encuestas piloto para lograr una mayor eficiencia en el diseño. En particular, si la información sobre los parámetros *a priori* es muy incierta, como lo fue en este caso, la utilización de un método Bayesiano para estimar el Error-D se ve complementada de manera muy exitosa con los modelos derivados de las encuestas piloto. Luego, éstas no sólo sirven para mejorar la interfaz, definición de variables y

maneras de contactar a los encuestados, sino que además tienen un rol muy importante en el diseño mismo del experimento.

Por último, se propone para investigaciones futuras el análisis de dos situaciones puntuales encontradas al estimar modelos en esta investigación. En primer lugar, queda pendiente un análisis más profundo de la utilización de modelos logit mixto con componentes de error para los casos en que las alternativas son genéricas. Resulta relevante determinar en qué medida los componentes de error capturan la correlación entre individuos en contraste con la heterogeneidad en la elección de alternativas respecto a una de referencia. Por otro lado, se debe investigar hasta qué punto es posible capturar la correlación entre las respuestas de un individuo a partir de la utilización de efectos socioeconómicos en modelos logit simple. Resultaría interesante contrastar dichos resultados con modelos mixto de parámetros aleatorios con efecto panel y estudiar si las diferencias o similitudes dependen mucho de los datos.

9. REFERENCIAS

- Alcalá, L. I. (2006). Reflexiones sobre algunas estrategias urbanísticas en la ciudad de Barcelona durante el último cuarto del siglo XX. *Cuaderno Urbano*, **5**, 75 – 104.
- Algers, S., Bergstrom, M., Dahlberg, M. y Dillen, J. (1999). Mixed logit estimation of the value of travel time. Working paper, Uppsala University, Department of Economics.
- Armstrong, P. M., Garrido, R.A. y Ortúzar, J. de D. (2001). Confidence intervals to bound the value of time. *Transportation Research*, **37E**(2 - 3), 143 – 161.
- Benach, N. (1996). Producción de imagen en la Barcelona del 92. *Estudios Geográficos*, **LIV**(212), 453 – 504.
- Bierlaire, M. (2003). BIOGEME: A free package for the estimation of discrete choice models. *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Marzo.
- Bierlaire, M. (2008). An Introduction to BIOGEME Version 1.7. biogeme.epfl.ch.
- Black, I.R., Efron, A., Ioannou, C. y Rose, J.M. (2005). Designing and implementing internet questionnaires using Microsoft Excel. *Australasian Marketing Journal*, **13**(2), 62 – 73.
- Bliemer, M.C.J. y Rose, J.M. (2005). Efficiency and sample size requirements for stated choice studies. Working Paper, University of Sydney, Institute of Transport and Logistics Studies, Abril.
- Bliemer, M.C.J. y Rose, J.M. (2006). Designing stated choice experiments: state-of-the-art. *11th International Conference on Travel Behavior Research*, Kyoto, Agosto.

Bliemer, M.C.J. y Rose, J.M. (2008). Construction of experimental for mixed logit models allowing for correlation across choice observations. *Proceedings of the 87th Annual Meeting of the Transportation Research Board*, Washington D.C., Enero.

Boxall, P.C., Adamowicz, W., Swait, J., Williams, M. y Louviere, J. (1996). A comparison of stated preference methods for environmental valuation. *Ecological Economics*, **18**(3), 243 – 253.

Buckingham, A. y Saunders, P. (2004). *The Survey Methods Workbook*. Polity Press Ltd., Cambridge.

Carrión, F. (2001) (Ed.). *Centros Históricos de América Latina y El Caribe*. Banco Interamericano de Desarrollo, Quito.

Caussade, S., Ortúzar, J. de D., Rizzi, L.I. y Hensher, D.A. (2005). Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research*, **39B**(7), 621 – 640.

Cherchi, E. y Ortúzar, J. de D. (2008). Empirical identification in the mixed logit: analyzing the effect of data richness. *Networks and Spatial Economics*, **8**(2-3), 109 – 124.

Couper, M.P. (2000). Web surveys: a review of issues and approaches. *The Public Opinion Quarterly*, **64**(4), 464 – 494.

Dillman, D.A. (2000). *Mail and Internet Surveys: The Tailored Design Method*. John Willey & Sons, Nueva York.

Espino, R., Ortúzar, J. de D. y Román, C. (2006). Confidence intervals for willingness to pay measures in mode choice models. *Networks and Spatial Economics*, **6**, 81 – 96.

Fink, A. (1995). *How to Ask Survey Question*. Sage Publications, California.

Fleming, C.M. y Bowden, M. (2007). Web-based surveys as an alternative to traditional mail methods. *Journal of Environmental Management*, en imprenta, disponible online desde 21 de Diciembre de 2007.

Foster, V. y Mourato, S. (2002). Testing for consistency in contingent ranking experiments. *Journal of Environmental Economics and Management*, **44**(2), 309 – 328.

Galilea, P. (2002). *Preferencias Declaradas en la Valoración del Nivel de Ruido en un Contexto de Elección Residencial*. Tesis de Magíster, Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile.

Garrido, R.A. (1991). *Preferencias Declaradas en la Modelación de Demanda por Nuevas Alternativas de Transporte*. Tesis de Magíster, Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile.

Godoy, G. (2004). *Estimación Clásica y Bayesiana del Modelo Logit Mixto*. Tesis de Magíster, Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile.

Godoy, S. (editor) (2006). *Monitoreando el Futuro Digital: Resultados Encuesta WIP-Chile*. Instituto de Estudios Mediales UC, Instituto de Sociología UC, Escuela de Ingeniería UC/CETIUC, Centro de Estudios de la Economía Digital CCS, Santiago.

Goett, A., Hudson, K. y Train, K. (2000). Costumer choice among retail energy suppliers: the willingness-to-pay for service attributes. *Energy Journal*, **21**(4),

Greene, M. y Ortúzar, J. de D. (2002). Willingness to pay for social housing attributes: a case study from Chile. *International Planning Studies*, **7**(1), 55 – 87.

Hanley, N., Wright, R.E., Adamowicz, V. (1998). Using choice experiments to value the environment. *Environmental and Resource Economics*, **11**(3-4), 413 – 428.

Henry, G.T. (1990). *Practical Sampling*. Sage Publications, Newbury Park.

Hensher, D. (2001). The sensitivity of the valuation of travel time savings to the specification of unobserved effects. *Transportation Research*, **37E**(2 – 3), 129 – 142.

Hensher, D., Shore, N. y Train, K. (2005). Households' willingness to pay for water service attributes. *Environmental and Resource Economics*, **32**, 509 – 531.

Hess, S. y Polak, J.W. (2004). Mixed logit estimation of parking type choice. 83rd *Annual Meeting of the Transportation Research Board*, Washington D.C., January.

Iglesias, P.A. (2007). *Valoración de Atributos de Diseño Urbano en la Percepción de Seguridad Ciudadana en Barrios de Bajo Ingreso*. Tesis de Magíster, Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile.

Iglesias, P.A. y Ortúzar, J. de D. (2008). Diseño experimental en estudios de preferencias declaradas: criterios de eficiencia para modelos logit. *Actas del XV Congreso Panamericano de Ingeniería de Tránsito y Transporte*, Cartagena de Indias, Septiembre.

Iragüen, P.A. (2002). *Valoración de Reducciones en el Riesgo de Muerte por Accidente en Areas Urbanas: Una Aplicación de Internet como Medio para Realizar Encuestas*. Tesis de Magíster, Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile.

Jara – Díaz, S. (2007). *Transport Economic Theory*. Elsevier Science, Amsterdam.

Johnson, M.E. (1987). *Multivariate Statistical Simulation*. John Wiley & Sons, Nueva York.

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

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

Meijer, E. y Rouwendal, J. (2006). Measuring welfare effects in models with random coefficients. *Journal of Applied Econometrics*, **21**, 227 – 244.

Moskowitz, H.R. y Martin B. (2008). Optimising the language of email survey invitations. *International Journal of Market Research*, **50**(4), 491 – 510.

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

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*, **XXXIII**(2), 69 – 86.

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

Ortúzar, J. de D. (2000b). Fundamentals of stated preference. En J. de D. Ortúzar (ed.), *Stated Preference Modelling Techniques*, PTRC, Londres.

Ortúzar, J. de D. y Garrido, R.A. (1991). Rank, rate or choice? An evaluation of SP methods in Santiago. *Proceedings 19th PTRC Summer Annual Meeting*, University of Sussex, Inglaterra, Septiembre.

Ortúzar, J. de D., Martínez, F. J. y Varela, F. J. (2000). Stated preferences in modelling accessibility. *International Planning Studies*, **5**(1), 65 – 85.

Ortúzar, J. de D. y Willumsen, L.G. (2001). *Modelling Transport*. 3º Edición, John Wiley & Sons, Chichester.

- Perez, P. (1998). *Preferencias Declaradas en la Elección por Localización Residencial*. Tesis de Magíster, Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile.
- Revelt, D. y Train, K. (1998). Mixed logit with repeated choices: households' choice of appliance efficiency level. *Review of Economics and Statistics*, **80**, 647 – 657.
- Rizzi, L.I. (2001). *Economía de los Accidentes Fatales: Una Aplicación al Caso de Seguridad Vial en Carreteras*. Tesis de Doctorado, Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile.
- Rodriguez, G. (2000). *Preferencias Declaradas en la Valoración de la Contaminación Atmosférica en un Contexto de Elección Residencial*. Tesis de Magíster, Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile.
- Rojas, E. (1999). *Old Cities, New Assets: Preserving Latin America's Urban Heritage*. Banco Interamericano de Desarrollo, Washington DC.
- Rojas, E. (2004). *Volver al Centro: La Recuperación de Áreas Urbanas Centrales*. Banco Interamericano de Desarrollo, Washington DC.
- Rose, J.M. y Bliemer, M.C.J. (2004). The design of stated choice experiments: the state of practice and future challenges. Working Paper, University of Sydney, Institute of Transport and Logistics Studies.
- Rose, J.M., Bliemer, M.C.J., Hensher, D.A. y Collins, A.T. (2008). Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research*, **42B**(4), 395 – 406.
- Sandor, Z. y Wedel, M. (2001). Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research*, **38**(4), 430 – 444.

Sillano, M. y Ortúzar, J. de D. (2005). Willingness-to-pay estimation with mixed logit models: some new evidence. *Environment and Planning*, **37**(A), 525 – 550.

Stern, M.J., Dillman, D.A., Smyth, J.D. (2007). Visual design, order effects, and respondent characteristics in a self-administred survey. *Survey Research Methods*, **1**(3), 121 – 138.

Train, K. (2003). *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge.

Train, K., Revelt, D., Ruud, P. (1999). Mixed Logit Estimation Routine for Panel Data. GAUSS Code.

Varela, F.J. (1997). *Preferencias Declaradas en la Modelación de la Variable Metro*. Tesis de Magíster, Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile.

Walfish, S. (2006). A review of statistical outlier methods. *Pharmaceutical Technology*, **30**(11), 82 – 88.

Williams, H.C.W.L. (1977). On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning*, **9**(A), 285 – 344.

Yañez, M.F. (2008). *Un Panel de Datos para Capturar los Cambios Introducidos por Transantiago*. Memoria de Título, Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile.

ANEXOS

ANEXO A : Encuesta Facultad de Arquitectura

A continuación se encuentra la pregunta sobre cambio de vivienda que se realizó en la encuesta desarrollada por alumnos de la Facultad de Arquitectura, Diseño y Estudios Urbanos PUC.

FACTORES ELECCIÓN ACTUAL RESIDENCIA			
6.3. ¿Cual de los siguientes factores influyó en la elección de su actual vivienda?	mucho	poco	No influyó
Distancia al trabajo			
Accesibilidad y Locomoción colectiva			
Seguridad			
Condiciones ambientales: olores, aire, ruido, vegetación			
Existencia de áreas verdes, parques y plazas			
Limpieza y estado de calles y veredas			
Existencia o cercanía de espacios deportivos y recreación			
Existencia o cercanía de servicios de salud			
Existencia o cercanía de Comercio-almacenes-supermercados			
Existencia o cercanía de Universidades-Colegios-liceos-jardines infantiles			
Existencia o cercanía de servicios de trámite en general: bancos, cuentas, etc.			
Posibilidad de comprar o arrendar una vivienda a un precio accesible			
Cercanía a familiares o amigos			
Prestigio y calidad del barrio			
Barrio interesante y diverso			
Actividad cultural			
Existencia o cercanía de bares, restaurantes, cafés, cines, etc. Esparcimiento en general			
Existencia de una comunidad amistosa y amable			
Tamaño de la vivienda			
Calidad de la vivienda			

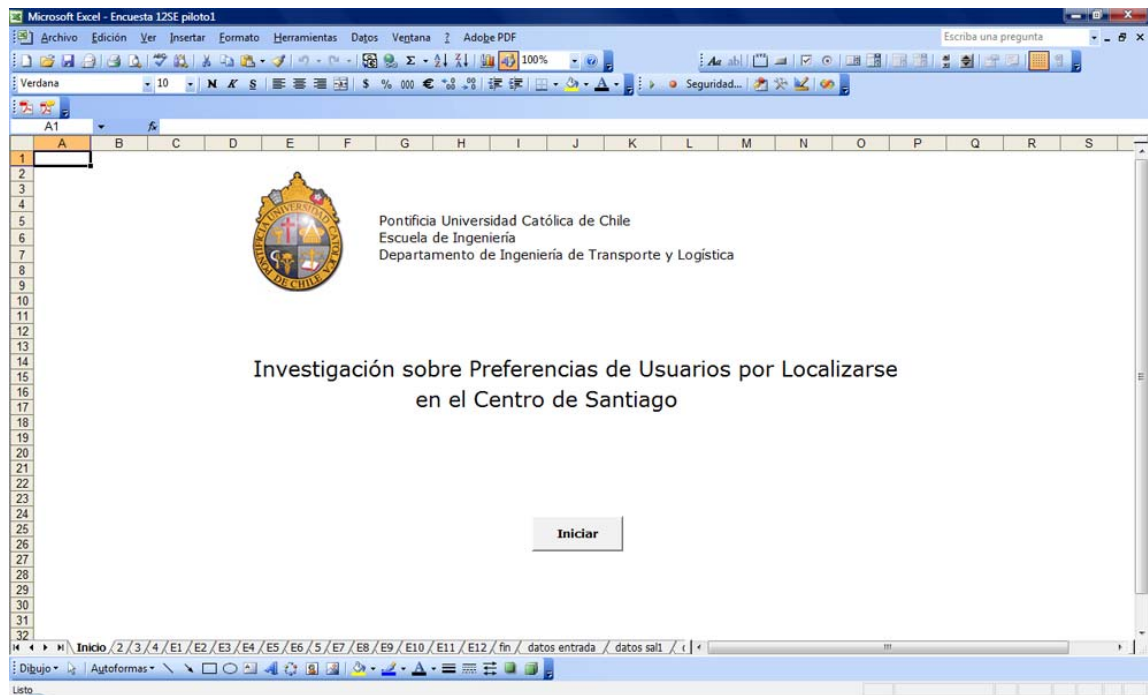
ANEXO B : Diseño Factorial de la Encuesta Inicial

JUEGO 1						
Situación de Elección	Alternativa	Niveles de las Variables				
		<i>Precio</i>	<i>Metro</i>	<i>Servicios</i>	<i>Cultura</i>	<i>Areas_Verdes</i>
1	1	Bajo	1	0	1	0
	2	Alto	1	1	1	1
2	1	Medio	1	1	0	1
	2	Bajo	0	0	0	1
3	1	Bajo	0	0	1	1
	2	Bajo	0	1	0	0
4	1	Bajo	0	1	1	0
	2	Medio	1	0	0	1
5	1	Medio	0	0	0	0
	2	Alto	1	1	1	0
6	1	Medio	1	1	0	1
	2	Bajo	0	0	1	0

JUEGO 2					
Situación de Elección	Alternativa	Niveles de las Variables			
		<i>Precio</i>	<i>Superficie</i>	<i>Limpieza_Calle</i>	<i>Espacio_Común</i>
7	1	Medio	0	0	0
	2	Alto	1	0	1
8	1	Alto	1	1	1
	2	Medio	0	0	1
9	1	Alto	1	0	1
	2	Bajo	0	1	0
10	1	Alto	0	1	1
	2	Medio	1	1	0
11	1	Bajo	1	1	0
	2	Medio	1	0	0
12	1	Alto	0	0	0
	2	Alto	0	1	1

ANEXO C : Interfaz en Microsoft Excel de la Encuesta Inicial

Se incluyen en este anexo la primera versión de la encuesta, previa a las pruebas piloto. La información se presentaba en pantalla completa de un computador portátil a los encuestados. No se incluyen aquí todas las situaciones de elección, sino que una representativa de cada juego.



Archivo Edición Ver Insertar Formato Herramientas Datos Ventana Adobe PDF

A B C D E F G H I J K L M N O P Q R

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44



Muchas gracias por participar en este estudio de la Pontificia Universidad Católica de Chile y por el tiempo que está invirtiendo en él.
En esta encuesta no existen respuestas correctas o incorrectas, sólo nos interesa conocer su opinión.
Es muy importante para nosotros que conteste de la manera más seria posible.

Para comenzar la encuesta necesitamos saber algunos de sus datos personales:

Edad:

Sexo:

Estado civil:

¿Qué estudios tiene?

¿A qué se dedica actualmente?

¿Posee licencia de conducir?

Siguiente

Archivo Edición Ver Insertar Formato Herramientas Datos Ventana Adobe PDF

A B C D E F G H I J K L M N O P Q

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44



Ahora necesitamos algunos datos familiares y de su vivienda actual:

Número de personas que habitan regularmente su vivienda actual:

Número de personas que habitarán en su nuevo departamento:

Número de personas menores de 18 años que habitarán en su nuevo departamento:

¿Dónde vive usted actualmente?

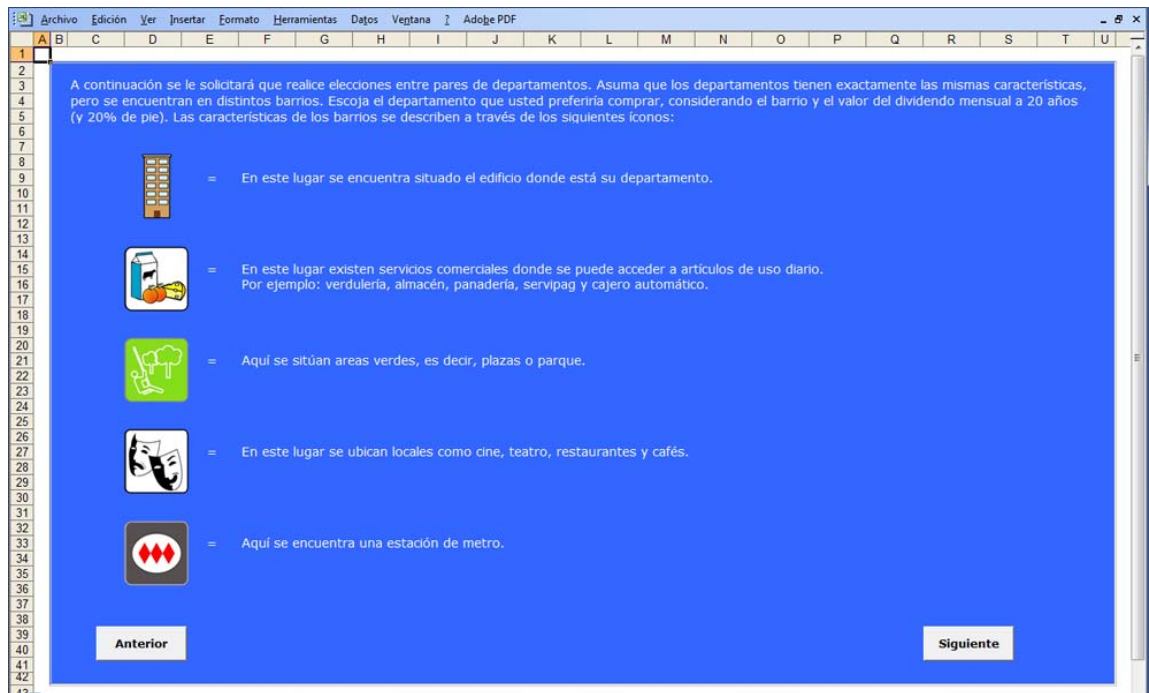
En su vivienda actual, usted o su pareja es:

¿Cuántos minutos demora actualmente en viajar a su trabajo?

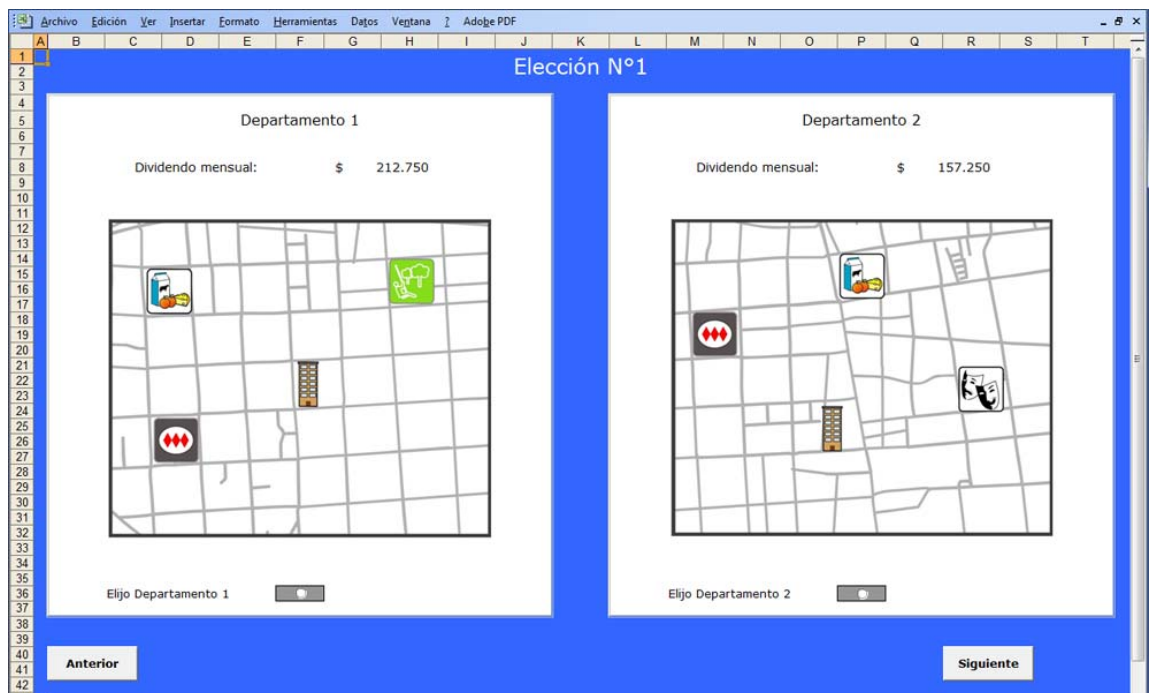
¿Cuántos dormitorios desea que tenga su nuevo departamento?

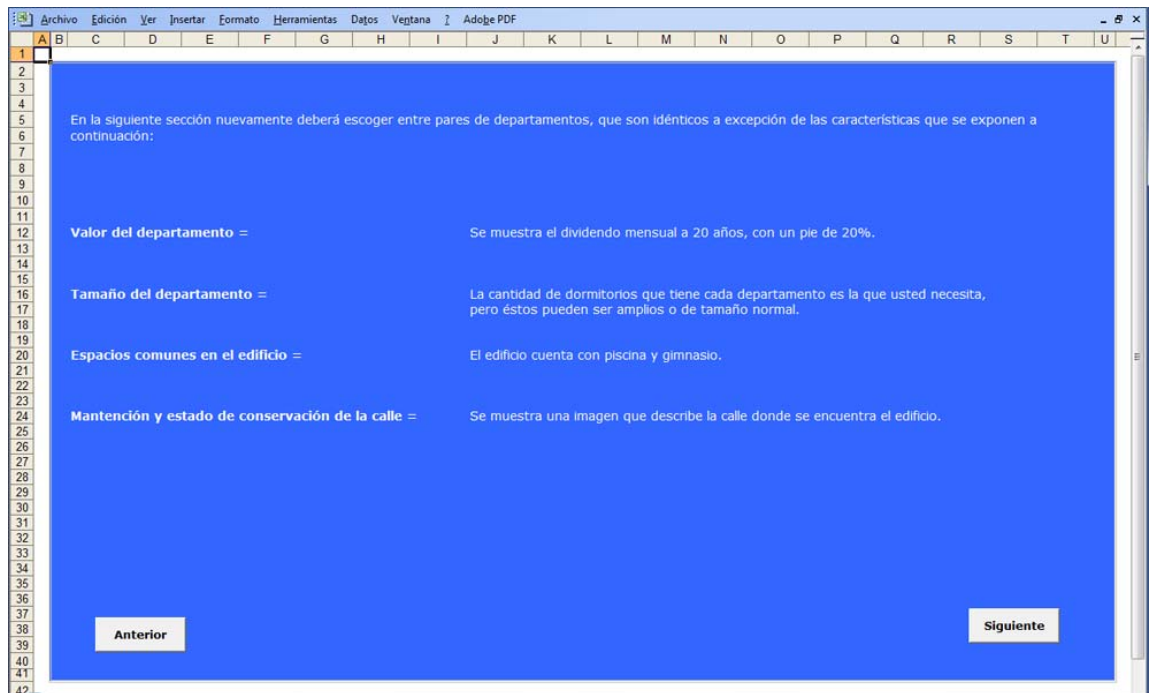
¿Cuántos años espera vivir en su nuevo departamento?

Anterior Siguiente

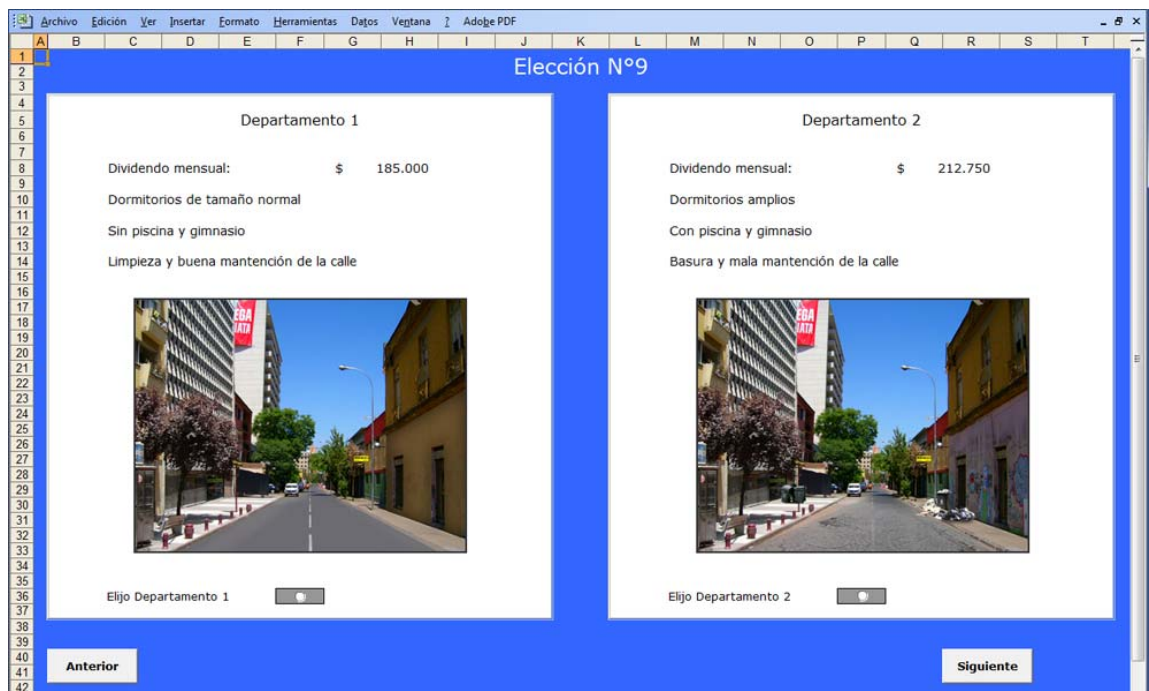


Situación de elección tipo, Juego 1:





Situación de elección tipo, Juego 2:



Archivo Edición Ver Insertar Formato Herramientas Datos Ventana ? Adobe PDF

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1																			
2																			
3																			
4																			
5																			
6																			
7																			
8																			
9																			
10																			
11																			
12																			
13																			
14																			
15																			
16																			
17																			
18																			
19																			
20																			
21																			
22																			
23																			
24																			
25																			
26																			
27																			
28																			
29																			
30																			
31																			
32																			
33																			
34																			
35																			
36																			
37																			
38																			
39																			
40																			
41																			
42																			



Muchas gracias por su ayuda en esta investigación.

Para finalizar, necesitamos pedirle que responda dos preguntas relativas a su ingreso que nos permitirán analizar mejor los resultados de la investigación. Recuerde que esta encuesta es totalmente confidencial y que no conocemos su nombre y su dirección, pero necesitamos estos datos para clasificar la información recopilada.

Por favor indique el ingreso líquido aproximado de su nuevo hogar:

Por favor indique su ingreso líquido personal:

Muchas gracias por su participación. Sus respuestas serán de mucha utilidad en la investigación. Le agradecemos el tiempo y la dedicación que ha invertido en ella.

[Anterior](#)[Fin](#)

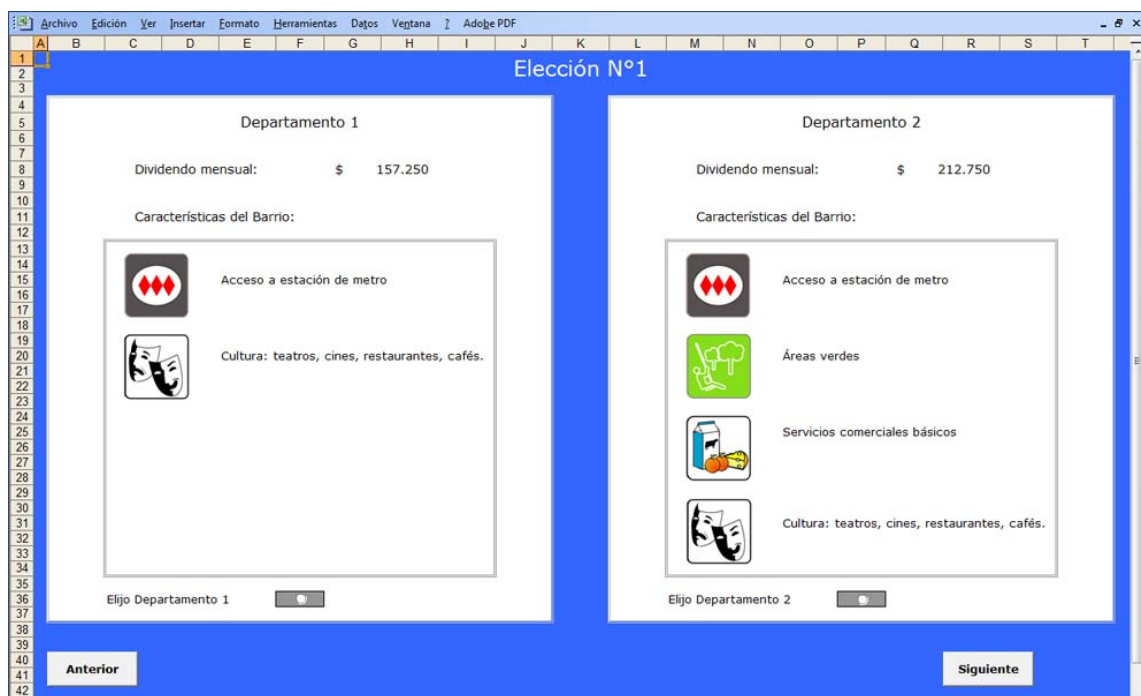
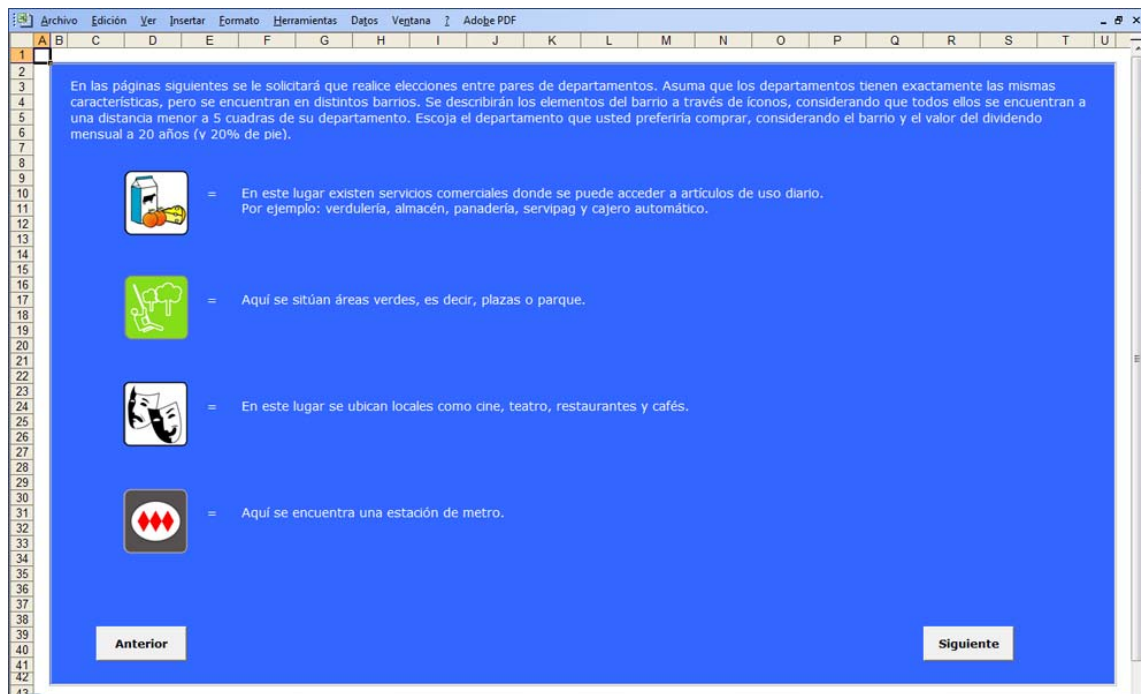
ANEXO D: Fotografías Iniciales Atributo *Limpieza_Calle*



ANEXO E : Niveles Atributo *Limpieza_Calle*, Encuesta Inicial



ANEXO F : Explicación y Elección Tipo Juego 1, luego de EP1



ANEXO G: Niveles Atributo *Limpieza_Calle*, luego de EP1



ANEXO H: Diseño Factorial para EP3

JUEGO 1						
Situación de Elección	Alternativa	Niveles de las Variables				
		<i>Precio</i>	<i>Metro</i>	<i>Servicios</i>	<i>Cultura</i>	<i>Areas_Verdes</i>
1	1	Medio	1	0	0	0
	2	Bajo	0	1	0	0
2	1	Alto	1	1	0	1
	2	Medio	0	1	0	1
3	1	Medio	0	0	1	1
	2	Alto	1	1	1	0
4	1	Bajo	1	0	1	0
	2	Bajo	1	0	0	1
5	1	Bajo	0	1	1	1
	2	Bajo	1	0	1	0
6	1	Alto	0	1	0	0
	2	Medio	0	0	1	1

JUEGO 2					
Situación de Elección	Alternativa	Niveles de las Variables			
		<i>Precio</i>	<i>Superficie</i>	<i>Limpieza_Calle</i>	<i>Espacio_Común</i>
7	1	Medio	0	0	1
	2	Alto	1	0	1
8	1	Medio	1	1	0
	2	Alto	0	0	0
9	1	Bajo	1	0	0
	2	Medio	0	0	1
10	1	Alto	1	0	1
	2	Bajo	0	1	0
11	1	Bajo	0	1	0
	2	Alto	1	1	1
12	1	Medio	0	1	1
	2	Alto	1	1	0

ANEXO I : Correo Electrónico de Contacto



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE ESCUELA DE ARQUITECTURA – ESCUELA DE INGENIERIA

Estimado/a:

Las Escuelas de Arquitectura e Ingeniería de esta universidad estamos colaborando en la realización de un proyecto de investigación, cuyo fin es contribuir a mejorar las políticas de construcción de edificios y equipamiento comunitario en barrios residenciales del centro de Santiago.

Por esto creemos que los resultados de nuestro estudio pueden ser de gran beneficio no sólo para Ud. y su familia, sino que en general para toda la ciudadanía.

Para seguir adelante y tener éxito con este proyecto, necesitamos su participación respondiendo una encuesta que no le tomará más de ocho minutos. Hemos podido contactarlo porque Ud. se encuentra, o se encontraba, en un proceso de búsqueda de departamento en la comuna de Santiago. No tenemos otros datos que su nombre y su correo electrónico.

Es muy importante señalar que la encuesta es totalmente anónima (al punto que no guardamos información de quienes la contestan), que la información que Ud. nos entregue está respaldada por la cláusula de *total confidencialidad* de esta universidad, y que sólo se utilizará para los fines exclusivos de esta investigación. Finalmente, y con el fin de agradecer su participación, les haremos partícipes de un incentivo especial.

Por favor seleccione la encuesta que le corresponde, según la cantidad de dormitorios que desea que tenga su departamento. La contraseña para entrar a las encuestas es “encuesta-uc”.

Ambiente único	Encuesta ambiente único
1 dormitorio	Encuesta 1 dormitorio
2 dormitorios	Encuesta 2 dormitorios
3 dormitorios	Encuesta 3 dormitorios
4 dormitorios	Encuesta 4 dormitorios

En caso de cualquier consulta o problema con respecto a la encuesta, por favor comuníquese con Ignacia Torres, a los teléfonos 354-4270 o 08-5007408, o por correo electrónico (iqtorres@puc.cl).

Agradeciendo de antemano su tiempo y consideración, le saludan atentamente,

Margarita Greene Z.

Directora Investigación y Postgrado
Escuela de Arquitectura

Tel: 354-5603; e-mail: mgreenez@uc.cl

Juan de Dios Ortúzar S.

Jefe Departamento Ingeniería de Transporte
Escuela de Ingeniería

Tel: 354-4822; e-mail: jos@ing.puc.cl

ANEXO J : Diseño Factorial para Encuesta Final

JUEGO 1						
Situación de Elección	Alternativa	Niveles de las Variables				
		<i>Precio</i>	<i>Metro</i>	<i>Servicios</i>	<i>Cultura</i>	<i>Areas_Verdes</i>
1	1	Bajo	0	0	0	1
	2	Alto	1	1	1	0
2	1	Medio	1	1	0	1
	2	Bajo	0	1	0	0
3	1	Alto	1	0	1	0
	2	Medio	0	1	0	1
4	1	Medio	1	0	0	1
	2	Alto	1	0	1	1
5	1	Alto	0	1	1	0
	2	Medio	0	0	1	1
6	1	Alto	0	1	1	0
	2	Medio	1	0	0	0

JUEGO 2					
Situación de Elección	Alternativa	Niveles de las Variables			
		<i>Precio</i>	<i>Superficie</i>	<i>Limpieza_Calle</i>	<i>Espacio_Común</i>
7	1	Bajo	0	1	1
	2	Medio	1	0	0
8	1	Bajo	0	0	1
	2	Bajo	1	1	1
9	1	Alto	1	1	0
	2	Alto	0	0	0
10	1	Bajo	1	0	0
	2	Alto	1	0	1
11	1	Medio	1	1	1
	2	Bajo	0	1	0
12	1	Medio	0	0	0
	2	Bajo	0	1	1

ANEXO K: Interfaz en Internet de la Encuesta Final



Pontificia Universidad Católica de Chile
Escuela de Arquitectura - Escuela de Ingeniería

Introducción

Muchas gracias por participar en este estudio de la Pontificia Universidad Católica de Chile y por el tiempo que está invirtiendo en él.

En esta encuesta no existen respuestas correctas o incorrectas, sólo nos interesa conocer su opinión.

Es muy importante para nosotros que conteste de la manera más seria posible.

5%

[Siguiente](#)

Listo





Pontificia Universidad Católica de Chile
Escuela de Arquitectura - Escuela de Ingeniería

Datos Personales

Para comenzar la encuesta necesitamos saber algunos de sus datos personales:

Edad (*)

Sexo (*)

¿Qué estudios tiene? (*)

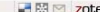
¿A qué se dedica? (*)

¿Dónde vive usted actualmente? (*)

11%

[Anterior](#) [Siguiente](#)

Listo





Pontificia Universidad Católica de Chile
Escuela de Arquitectura - Escuela de Ingeniería

Datos nuevo departamento

¿Cuántos automóviles habrá en su nuevo hogar? (*)

Número de personas que habitarán en su nuevo departamento (*)

¿Cuántos años espera vivir en su nuevo departamento? (*)

[Anterior](#)
[Siguiente](#)

Listo




Pontificia Universidad Católica de Chile
Escuela de Arquitectura - Escuela de Ingeniería

Explicación - 2 Dormitorios

En las páginas siguientes se le solicitará que realice elecciones entre pares de departamentos. Asuma que los departamentos tienen exactamente las mismas características, pero se encuentran en distintos barrios. Se describirán los elementos del barrio a través de íconos, considerando que todos ellos se encuentran a una distancia menor a 5 cuadras de su departamento. escoja el departamento que usted preferiría comprar, considerando el barrio y el valor del dividendo mensual a 20 años (y 20% de pie).



= En este lugar existen servicios comerciales donde se puede acceder a artículos de uso diario. Por ejemplo: verdulería, almacén, panadería, servipag y cajero automático.



= Aquí se sitúan áreas verdes, es decir, plazas o parque.



= En este lugar se ubican locales como cine, teatro, restaurantes y cafés.



= Aquí se encuentra una estación de metro.

[Anterior](#)
[Siguiente](#)

Listo


Ejemplo de situación de elección, Juego 1.

Elección N° 1

¿Cuál departamento escogería? (*)

Departamento N° 1

Dividendo mensual: \$157.250

Características del barrio:

 Areas verdes

Departamento N° 2

Dividendo mensual: \$212.750

Características del barrio:

 Acceso a estación de metro

 Servicios comerciales básicos

 Cultura: teatros, cines, restaurantes, cafés

Departamento N° 1 ☐ Departamento N° 2 ☐

26%

[Anterior](#) [Siguiente](#)

Listo 

 Pontificia Universidad Católica de Chile
Escuela de Arquitectura - Escuela de Ingeniería

Explicación

En la siguiente sección nuevamente deberá escoger entre pares de departamentos, que son idénticos a excepción de las características que se exponen a continuación:

Valor del departamento = Se muestra el dividendo mensual a 20 años, con un pie de 20%.

Tamaño del departamento = La cantidad de dormitorios que tiene cada departamento es la que usted necesita, pero éstos pueden ser amplios o de tamaño normal.

Espacios comunes en el edificio = El edificio cuenta con piscina y gimnasio.

Mantenimiento y estado de conservación de la calle = Se muestra una imagen que describe la calle donde se encuentra el edificio.

58%

[Anterior](#) [Siguiente](#)

Listo 

Ejemplo de situación de elección, Juego 2.

Elección N° 7

¿Cuál departamento escogería? (*)

Departamento N° 1

Dividendo mensual: \$157.250

Dormitorios de tamaño normal

Con piscina y gimnasio

Basura y mala mantención de la calle



Departamento N° 2

Dividendo mensual: \$185.000

Dormitorios amplios

Sin piscina y gimnasio

Limpieza y buena mantención de la calle



Departamento N° 1 ☐ Departamento N° 2 ☐

63%

Listo

 **Pontificia Universidad Católica de Chile**
Escuela de Arquitectura - Escuela de Ingeniería

Preguntas Finales

Para finalizar, necesitamos pedirle que responda dos preguntas relativas a su ingreso que nos permitirán analizar mejor los resultados de la investigación.

Por favor indique su ingreso líquido personal en pesos chilenos (\$) (*)

Por favor indique la suma de los ingresos líquidos (\$) de todas las personas que vivirán en su nuevo hogar (incluido usted) (*)

Al finalizar la encuesta sortharemos dos premios de \$50.000 cada uno entre los participantes. Necesitamos su correo electrónico para poder incluirlo en el sorteo. Esta información se encuentra respaldada por la cláusula de total confidencialidad de la Pontificia Universidad Católica de Chile.

Si usted desea concursar, ingrese su correo electrónico aquí:

95%

[Anterior](#) [Siguiente](#)

Listo



Pontificia Universidad Católica de Chile
Escuela de Arquitectura - Escuela de Ingeniería

Muchas Gracias

Muchas gracias por su participación.

Sus respuestas serán de mucha utilidad en la investigación.

Le agradecemos el tiempo y la dedicación que ha invertido en ella.

100%

[Anterior](#) [Fin](#)

Listo

ANEXO L : Modelo MNL2

Función de Utilidad:

$$\begin{aligned}
 V_j = & \left(\begin{array}{c} -0,0232 + 0,0059 \cdot D_{\text{Maipu}} + 0,0052 \cdot D_{\text{Car}} - 0,0142 \cdot \text{IFAM1} - 0,0073 \cdot D_{\text{Sexo}} \\ (-6,93) \quad (1,63) \quad (1,69) \quad (-3,83) \quad (-2,33) \end{array} \right) \cdot \text{Precio} + \\
 & \left(\begin{array}{c} 1,943 - 0,2604 \cdot D_{10\text{años}} + 0,223 \cdot \text{IFAM1} - 0,2766 \cdot D_{\text{Sexo}} \\ (15,37) \quad (-1,86) \quad (1,35) \quad (-1,92) \end{array} \right) \cdot \text{Metro} + \\
 & \left(\begin{array}{c} 0,6168 + 0,2384 \cdot D_{\text{Hab1}} + 0,278 \cdot (D_{\text{Maipu}} + D_{\text{Provi}}) + 0,1711 \cdot D_{10\text{años}} \\ (5,45) \quad (1,66) \quad (2,05) \quad (1,46) \\ -0,3022 \cdot (\text{IFAM1} + \text{IFAM2}) \\ (-2,49) \end{array} \right) \cdot \text{Metro} + \\
 & \left(\begin{array}{c} 0,576 \cdot D_{\text{Hab1}} + 0,2798 \cdot D_{\text{Hab2}} + 0,2879 \cdot D_{5\text{años}} + 0,1854 \cdot D_{10\text{años}} \\ (2,72) \quad (2,15) \quad (1,67) \quad (1,47) \\ -0,6617 \cdot \text{IFAM1} - 0,4742 \cdot \text{IFAM2} + 0,2478 \cdot D_{\text{Sexo}} \\ (-3,58) \quad (-3,35) \quad (2,06) \end{array} \right) \cdot \text{Cultura} + \\
 & \left(\begin{array}{c} 0,8467 + 0,4618 \cdot D_{\text{Fuera}} + 0,3084 \cdot (D_{\text{Prov}} + D_{\text{Stgo}}) - 0,1961 \cdot D_{\text{Car}} \\ (6,93) \quad (2,07) \quad (2,71) \quad (-1,58) \end{array} \right) \cdot \text{Servicios} + \\
 & \left(\begin{array}{c} 0,5913 - 0,2526 \cdot D_{\text{Hab2}} + 0,3482 \cdot D_{\text{Fuera}} + 0,3015 \cdot D_{\text{Stgo}} \\ (4,07) \quad (-1,92) \quad (1,39) \quad (2,1) \\ -0,5499 \cdot \text{IFAM1} - 0,2698 \cdot \text{IFAM2} + 0,2184 \cdot D_{\text{Sexo}} \\ (-3,13) \quad (-1,83) \quad (1,56) \end{array} \right) \cdot \text{Espacio _ Comun} + \\
 & \left(\begin{array}{c} -2,772 + 0,3282 \cdot D_{10\text{años}} + 0,381 \cdot \text{EDAD1} \\ (-13,06) \quad (1,96) \quad (1,97) \end{array} \right) \cdot \text{Limpieza _ calle} + \\
 & \left(\begin{array}{c} 0,9712 + 0,2201 \cdot (D_{\text{Hab1}} + D_{\text{Hab2}}) - 0,4477 \cdot D_{\text{Maipu}} + 0,3616 \cdot D_{5\text{años}} \\ (8,83) \quad (1,64) \quad (-2,6) \quad (2,03) \end{array} \right) \cdot \text{Superficie}
 \end{aligned}$$

Significado de variables:

Variable SE	Definición
D _{5años}	Vale 1 si vivirán en el departamento por menos de 5 años
D _{10años}	Vale 1 si vivirán en el departamento por más de 10 años
D _{Car}	Vale 1 si el nuevo departamento tendrá 1 o 2 autos
D _{Fuera}	Vale 1 si la persona vive fuera de la ciudad de Santiago
D _{Maipu}	Vale 1 si la persona vive en la comuna de Maipú, Puente Alto o La Florida
D _{Hab1}	Vale 1 si vivirá 1 persona en el departamento
D _{Hab2}	Vale 1 si vivirán 2 personas en el departamento
D _{Provi}	Vale 1 si la persona vive en la comuna de Providencia o Ñuñoa
D _{Sexo}	Vale 1 si la persona es hombre
D _{Stgo}	Vale 1 si la persona vive en la comuna de Santiago
EDAD1	Vale 1 si el encuestado tiene menos de 36 años
IFAM1	Vale 1 si el <i>IFAM</i> será menor o igual que \$600.000
IFAM2	Vale 1 si el <i>IFAM</i> será entre \$600.001 y \$1.000.000

ANEXO M : Parámetros *Cultura*, Modelo MNL2

Parámetros Asociados a *Cultura*, Modelo MNL2, $IFAM < \$600.000$

Habitantes	Género	Tiempo que vivirá en depto.		
		<i>Años</i> < 5	$5 \leq \text{Años} \leq 10$	$10 < \text{Años}$
Uno	Masculino	0,4500	0,1620	0,3475
	Femenino	0,2022	-0,0857	0,0997
Dos	Masculino	0,1537	-0,1342	0,0512
	Femenino	-0,0941	-0,3820	-0,1966
Tres o más	Masculino	-0,1260	-0,4140	-0,2285
	Femenino	-0,3738	-0,6617	-0,4763

Parámetros Asociados a *Cultura*, Modelo MNL2, $\$600.000 \leq IFAM \leq \$1.000.000$

Habitantes	Género	Tiempo que vivirá en depto.		
		<i>Años</i> < 5	$5 \leq \text{Años} \leq 10$	$10 < \text{Años}$
Uno	Masculino	0,6375	0,3495	0,5349
	Femenino	0,3897	0,1017	0,2872
Dos	Masculino	0,3412	0,0533	0,2387
	Femenino	0,0934	-0,1945	-0,0091
Tres o más	Masculino	0,0615	-0,2265	-0,0411
	Femenino	-0,1863	-0,4742	-0,2888

Parámetros Asociados a *Cultura*, Modelo MNL2, \$1.000.000 < IFAM

Habitantes	Género	Tiempo que vivirá en depto.		
		<i>Años</i> < 5	5 ≤ <i>Años</i> ≤ 10	10 < <i>Años</i>
Uno	Masculino	1,1117	0,8238	1,0092
	Femenino	0,8639	0,5760	0,7614
Dos	Masculino	0,8155	0,5275	0,7129
	Femenino	0,5677	0,2798	0,4652
Tres o más	Masculino	0,5357	0,2478	0,4332
	Femenino	0,2879	0	0,1854