

IMPORTANCIA DE LOS ATRIBUTOS DE VIVIENDA Y BARRIO EN LOCALIZACIÓN RESIDENCIAL: UNA APLICACIÓN DEL MÉTODO *BEST-WORST* AL CENTRO DE SANTIAGO

Camila Balbontin, Pontificia Universidad Católica de Chile, cbalbon1@uc.cl
Juan de Dios Ortúzar, Pontificia Universidad Católica de Chile, jos@ing.puc.cl
Joffre Swait, UTS Business School, joffre.swait@uts.edu.au

RESUMEN

Las encuestas de preferencias tradicionales utilizadas para modelar elecciones entre alternativas discretas no permiten diferenciar entre la importancia de un atributo y la importancia de los niveles que éste presenta en la encuesta. Recientemente se ha planteado que las encuestas del tipo *best-worst* permiten hacer esta diferenciación. En este artículo se estudia la posibilidad de combinar respuestas de esta naturaleza, respecto a los atributos de distintas alternativas de vivienda y barrio, con elecciones binarias (arrendaría o no la vivienda ofrecida) en el centro de Santiago. El mejor modelo resultó de la combinación de las respuestas de elección discreta con las respuestas *best* (mejor atributo), considerando dos atributos específicos y seis comunes a cada tipo de encuesta.

Palabras clave: experimentos de elección discreta, valoración de atributos, elección best-worst

ABSTRACT

Traditional discrete choice experiments, designed to study individual choices, do not differentiate between the importance of an attribute and that associated with its levels of variation. It has been suggested recently that the best-worst task allows this differentiation. In this paper we pool best worst answers with binary stated choices to study the importance of several dwelling and neighbourhood attributes of apartments located in the centre of Santiago, Chile. The best resulting model was obtained by pooling only the best answers with the binary stated choices, under the assumption of six common and two specific attributes to each dataset.

Keywords: discrete choice experiments, attribute importance, best-worst task

1. INTRODUCCIÓN

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

Uno de los problemas asociados a los modelos de elección estimados utilizando encuestas de preferencias tradicionales, ha sido como detectar correctamente el impacto que tiene la diferencia entre los niveles de un atributo en su significancia estadística. En este marco, Louviere y Swait (1997) desarrollaron un enfoque de encuestas llamada *best-worst*, cuyo objetivo es precisamente entregar información respecto a la significancia de cada atributo y, de forma separada, la significancia de los niveles presentados al encuestado. En su investigación, Torres (2009), también realizó una encuesta de tipo *best-worst* en conjunto con una del tipo elección binaria (arrendaría o no la alternativa presentada, descrita en términos del conjunto de atributos). Esta encuesta fue aplicada a individuos que buscaban arrendar un departamento en el centro de Santiago, y no fue analizada en su estudio. En este artículo se usan los datos de dicha encuesta para mostrar la potencial complementariedad del enfoque *best-worst* con el enfoque de elección discreta.

El resto del artículo está organizado de la siguiente forma. En la sección 2, se explica la teoría *best-worst*, y también se analizan las diferencias entre la función de utilidad estimada a partir de las respuestas a este tipo de encuestas con la función utilidad estimada a partir de las respuestas de elección discreta binaria. En la sección 3 se explica con mayor detalle las encuestas utilizadas en esta investigación, se definen los atributos utilizados y la codificación de cada uno y en la cuarta se exponen los principales resultados obtenidos. En la sección 5 se validan los resultados utilizando una muestra de elección discreta tradicional aplicada a individuos con similares características y en la sección 6 se discuten y analizan los resultados obtenidos. Finalmente, en la sección 7 se presentan las conclusiones y alcances del trabajo.

2. MARCO TEÓRICO

2.1. Teoría Best-Worst

En encuestas del tipo *best-worst* se presenta una situación de elección, con alternativas definidas en base a tres o más atributos, y para cada una los individuos deben elegir el mejor (*best*) y peor (*worst*) de los atributos estipulados. Además, se permite que cada individuo procese la elección de forma distinta; por ejemplo, el proceso de elección puede ser simultáneo o secuencial. El primero se conoce como *Max-Diff* o “diferencia máxima”, y en este caso el individuo busca simultáneamente un par de atributos que tengan la mayor diferencia de utilidad reportada, esto es, la mejor y peor opción. El proceso secuencial, por otro lado, puede realizarse de distintas formas, por ejemplo: (1) primero buscar el mejor atributo, luego el segundo mejor y así hasta llegar al peor; (2) primero buscar el peor, luego el segunda peor y así hasta llegar al mejor, etc.

Finn y Louviere (1992) definen tres tipos (casos) de encuesta *best-worst*:

Caso 1 (Objeto): Aquí se presentan al individuo distintos objetos y éste sólo debe elegir cuál es el mejor y cuál es el peor.

Caso 2 (Nivel de atributo): En este caso, que se utilizará en este artículo, se presenta al individuo un solo conjunto de atributos (perfil de elección o una descripción de un producto), y éste debe elegir cuáles le parecen mejor y peor. Cada atributo posee un cierto número de niveles n , pero en cada conjunto de atributos (alternativa) sólo se muestra un nivel de cada atributo. En la literatura se pueden encontrar aplicaciones de este caso especialmente en áreas de la salud (Flynn *et al.* 2013; Couzner *et al.* 2013; Flynn *et al.* 2010).

Este caso permite diferenciar entre la importancia de cada atributo y la importancia de la escala de sus niveles. En un experimento tradicional de elección discreta, puede que un atributo parezca no significativo debido a que los niveles presentados en el diseño no varían lo suficiente. Lo mismo puede ocurrir al revés, donde un atributo que en realidad no sea relevante parezca serlo porque la variación en sus niveles es tan grande que lo hace aparecer como significativo.

Caso 3 (Perfil): En este caso se presenta al individuo más de un perfil de atributos, y éste debe elegir los perfiles mejor y peor. Este caso es el más similar al de modelos de elección discreta tradicional.

2.2. Funciones Utilidad Elección *Best-Worst* vs. Elección Discreta Binaria

En nuestro análisis utilizaremos un modelo logit simple para obtener las funciones de utilidad tanto para las respuestas obtenidas en la elección *best-worst* como en las de elección discreta binaria. No obstante, es importante entender que las funciones de utilidad obtenidas serán distintas para cada tipo de encuesta. En efecto, a través de las respuestas correspondientes a la elección *best-worst*, se obtendrá una función de utilidad para cada atributo (ya que estas son las opciones de elección). De esta forma, cada atributo tendrá un *intercepto* que representa su significancia y es un valor promedio para los niveles definidos. A su vez, cada atributo tendrá una pendiente, que representa su *sensibilidad* a los niveles utilizados. Es importante destacar que esto se debe a la codificación utilizada (*effects code* para los atributos discretos con dos niveles y normalización entre -1 y +1 para el resto, ver Tabla 1 para más detalles); vale decir, de no codificarse de la forma planteada, no se podrían deducir las mismas conclusiones.

Por otro lado, la función de utilidad que se puede obtener de las respuestas de elección binaria no tiene la misma forma, ya que se basa en un conjunto de atributos (y no en cada uno por separado). Por esto, se puede obtener un intercepto común para las respuestas “si elegiría la alternativa presentada” (o para la respuesta “no elegiría la alternativa presentada”, dependiendo cuál se quiera tomar como base¹). A su vez, se obtendrá una utilidad marginal para cada atributo, que representa la *sensibilidad* a los niveles utilizados. La identificación de la función de utilidad en modelos de elección se basa en las diferencias entre las utilidades de las distintas alternativas. En nuestro caso, a partir de las respuestas de elección binaria, ambas alternativas (si/no en este caso) se posicionan en forma relativa en la misma escala de utilidad. Por lo tanto, no es necesario incluir un factor de escala.

¹ Esto se puede ver con más claridad en el capítulo siguiente, donde se explicita el diseño de la encuesta.

Entonces, es posible comparar las sensibilidades obtenidas para cada atributo a partir de las respuestas de elección *best-worst* y las respuestas de elección binaria. No obstante, es fundamental entender que la función de utilidad también se explicita de forma distinta para el caso *best* que para el *worst*, así como también es distinta para el caso “si elegiría la alternativa presentada” que cuando “no elegiría la alternativa presentada”, ya que cambia la probabilidad en los distintos tipos de elección.

Como explican Louviere y Swait (1997), la probabilidad de elegir un atributo como el mejor (*best*) corresponde a la probabilidad de que la función de utilidad de ese atributo sea mayor que la del resto. A su vez, la probabilidad de elegir un atributo como el peor (*worst*), es igual a la probabilidad que su función utilidad sea menor que la del resto. Entonces, teóricamente la probabilidad de elegir un atributo como el mejor debería ser la función inversa de elegirlo como el peor. Por lo mismo, para las respuestas *worst* se codificará los niveles de los atributos y sus interceptos como el inverso del caso *best*. No obstante, el proceso mental que lleva a cabo una persona para elegir un atributo como mejor no es necesariamente el mismo que para escoger el peor; esto se puede reflejar en diferencias entre los coeficientes de las diferentes funciones de utilidad y será examinado cuando se presenten los resultados de estos modelos en la sección 4.

Explícitamente, las funciones de utilidad para las distintas respuestas tomaron la siguiente forma:

Encuesta Best:

$$U_{metro} = C_{metro} + \theta_{metro} * X_{metro}$$

$$U_{áreas\ verdes} = C_{áreas\ verdes} + \theta_{áreas\ verdes} * X_{áreas\ verdes}$$

...

y así sucesivamente para cada atributo.

Encuesta Worst:

$$U_{metro} = -C_{metro} + \theta_{metro} * -X_{metro}$$

$$U_{áreas\ verdes} = -C_{áreas\ verdes} + \theta_{áreas\ verdes} * -X_{áreas\ verdes}$$

...

etc.

Encuesta ED:

$$U_{Si\ lo\ arrendaría} = C_{si} + \theta_{metro} * X_{metro} + \theta_{áreas\ verdes} * X_{áreas\ verdes} + \dots$$

$$U_{No\ lo\ arrendaría} = 0$$

Donde θ representa la *sensibilidad* descrita anteriormente, que se considera común para todas las respuestas y C representa el *intercepto* de cada atributo. En el caso de la Encuesta ED los parámetros de sensibilidad son equivalentes a las utilidades marginales de cada atributo.

3. DATOS DISPONIBLES

3.1. Atributos Presentados en la Encuesta

La encuesta utilizada en este artículo fue aplicada a individuos que planeaban arrendar un departamento en la comuna de Santiago (Torres, 2009; Torres *et al.*, 2012). Con el fin de estudiar

y comparar la metodología *best-worst* con la de elección binaria, se consideró siete atributos con dos niveles cada uno, y un atributo (*dividendo del departamento*) con ocho niveles. Estos atributos, que se describen en la Tabla 1, están agrupados en torno a: (i) características del barrio (*acceso al metro, áreas verdes, servicios comerciales básicos y servicios culturales*) y (ii) características de la vivienda (*tamaño de dormitorios, disponibilidad de piscina en el edificio, limpieza y estado de conservación de la calle* en que está ubicado el edificio y, por supuesto, una variable precio, especificada como el costo de arriendo o *dividendo del departamento*).

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

3.1. Diseño de la Encuesta

En la Figura 1 se muestra la encuesta presentada a los individuos. En primer lugar debían elegir cuál atributo les parecía mejor (*best*), luego el que les parecía peor o menos atractivo (*worst*), y finalmente debían señalar si arrendarían o no el departamento mostrado. Esta última respuesta es la que hemos denominado “elección binaria”.

Figura 1: Diseño encuesta B/W y Elección Binaria

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

Atributo más atractivo	Atributos del Departamento	Valor	Atributo menos atractivo
<input type="checkbox"/>	Dividendo departamento	\$\$\$	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Tamaño de los dormitorios (m2)	400	<input type="checkbox"/>
<input type="checkbox"/>	Gimnasio y piscina en el edificio	No tiene	<input type="checkbox"/>
<input type="checkbox"/>	Mantenimiento y estado de conservación de la calle	Mala	<input type="checkbox"/>
<input checked="" type="checkbox"/>	Acceso a metro	Cerca	<input type="checkbox"/>
<input type="checkbox"/>	Servicios comerciales básicos	Lejos	<input type="checkbox"/>
<input type="checkbox"/>	Servicios culturales	Lejos	<input type="checkbox"/>
<input type="checkbox"/>	Áreas verdes	Cerca	<input type="checkbox"/>

Arrendaría este departamento? Si No

Elección B/W

Elección Discreta Binaria

3.1. Codificación

Para comparar los coeficientes obtenidos a partir de la elección *best-worst* y la elección binaria (como se explica en la sección siguiente), se debió codificar las respuestas utilizando el sistema *effects code* para las variables discretas (Louviere y Swait, 1997), lo que se muestra en la última columna de la Tabla 1. El atributo *Dividendo del departamento* no fue codificado de la misma forma, ya que primero se normalizó sus valores con el objetivo que su menor nivel (\$ 120.000) tomara el valor -1 y su mayor nivel (\$ 341.000) tomara el valor +1. Los niveles entre estos dos valores fueron normalizados siguiendo esta escala.

Tabla 1: Atributos utilizados en la encuesta

Atributo	Niveles	Código
Acceso a metro	Lejos, a más de 5 cuadras de su departamento	+1
	Cerca, a menos de 5 cuadras de su departamento	-1
Áreas verdes	Lejos, a más de 5 cuadras de su departamento	+1
	Cerca, a menos de 5 cuadras de su departamento	-1
Dividendo departamento	\$120.000	-1
	\$140.000	-0,819
	\$171.000	-0,538
	\$199.000	-0,285
	\$217.000	-0,122
	\$253.000	+0,204
	\$293.000	+0,567
	\$341.000	+1
Gimnasio y piscina en el edificio	Sí tiene	+1
	No tiene	-1
Mantención y estado de conservación de la calle	Buena	+1
	Mala	-1
Servicios comerciales básicos	Lejos, a más de 5 cuadras de su departamento	+1
	Cerca, a menos de 5 cuadras de su departamento	-1
Servicios culturales	Lejos, a más de 5 cuadras de su departamento	+1
	Cerca, a menos de 5 cuadras de su departamento	-1
Tamaño de los dormitorios	Amplios	+1
	Normales	-1

3.1. Base de Datos y Muestra de Validación

En total la encuesta fue contestada por 202 individuos y cada uno enfrentó ocho escenarios. Como todos los individuos contestaron adecuadamente la pregunta de elección discreta binaria, se cuenta con 1.616 respuestas de este tipo. No obstante, algunos no contestaron la pregunta de cuál atributo era el más o menos atractivo en algunos de los escenarios, por lo que en total se obtuvo 1.596 respuestas del tipo *best* y 1.506 para el tipo *worst*.

Torres (2009) realizó un estudio sobre valoración de atributos en este contexto utilizando una encuesta de preferencias declaradas clásica que fue aplicada a una muestra distinta de individuos. En ese estudio se encuestó a 497 individuos, quienes también debieron enfrentar ocho escenarios consistentes en dos perfiles de elección. En total se logró 5.965 observaciones. Esta muestra será utilizada como muestra de validación para analizar la capacidad predictiva de nuestros modelos.

4. RESULTADOS

Luego de entender el aporte que tiene la función de utilidad obtenida a través de la encuesta *best-worst* y cómo se puede comparar con la función de utilidad obtenida a partir de las respuestas de elección binaria, se buscó el mejor modelo combinando las respuestas a ambas encuestas. Para lograr esto, primero se investigó si en realidad las respuestas *best* (mejor atributo) y *worst* (peor atributo) eran efectivamente inversas. Para esto, se estimó las sensibilidades (pendientes) de cada atributo con: (a) las respuestas sobre cuál era el mejor atributo (*best*); (b) las respuestas sobre cuál era el peor (*worst*) y (c) las respuestas sobre el mejor (*best*) y peor (*worst*) atributos simultáneamente, considerando un factor de escala para las respuestas *worst*. En la Tabla 2 se presentan estos resultados y en la Figura 2 se muestran en forma gráfica.

Tabla 2: Sensibilidades (pendientes) respuestas *Best*, *Worst* y *Best-Worst*

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

Se puede observar que el factor de escala considerado para las respuestas *worst* es significativamente distinto de uno, por lo que es necesario incluirlo.

Para verificar si las respuestas del tipo *best* y *worst* permiten considerar los ocho atributos como comunes, se puede utilizar el estadístico de razón de verosimilitud siguiente (Ortúzar y Willumsen, 2011, pag. 325):

$$\lambda_a = -2(L_1 + L_2) - L_\mu \quad (1)$$

donde L_μ representa la log-verosimilitud del modelo conjunto, L_i representa la log-verosimilitud de cada modelo por separado; si los a coeficientes comunes son efectivamente iguales, λ_a distribuye χ^2 con a grados de libertad. Si λ_a es mayor que el valor crítico de χ^2 con a grados de libertad para el nivel de confianza requerido, se rechaza la hipótesis nula que el modelo conjunto es equivalente a los modelos por separado, indicando que no todos los parámetros pueden ser

considerados comunes. En nuestro caso resulta que $\lambda_a = 41$ y $\chi^2_8 = 15,51$, por lo que claramente se rechaza la hipótesis nula que todos los parámetros son comunes. Esto implica que las funciones de utilidad de las respuestas *best* y *worst* no son inversas, ya que algunos de los atributos se comportan de manera distinta.

En este artículo nos interesa encontrar el mejor modelo resultante de combinar las respuestas de elección discreta binaria con las de la elección *best-worst*, y no las de *best-worst* por sí solas. Como se constató que la hipótesis de considerar los ocho atributos comunes en las respuestas *best* y *worst* se rechazaba, se estudiará cuál de ellas por separado se combina mejor con las de elección discreta binaria.

En la Tabla 3 se presentan las sensibilidades obtenidas utilizando: (a) sólo las respuestas de elección binaria; (b) las respuestas de elección binaria junto con las respuestas *best* (mejor atributo); (c) las respuestas de elección binaria junto con las respuestas *worst* (peor atributo). En todos estos modelos se tomó las respuestas de elección discreta como la base (factor de escala igual a uno) y para las otras respuestas se estimó un factor de escala.

Tabla 3: Coeficientes (pendientes) Modelo BW y Elección Discreta Combinados

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

Con los resultados de la Tabla 3 se puede aplicar nuevamente el test estadístico anterior (1), obteniéndose los valores que se muestran en Tabla 4 (es importante mencionar que los factores de escala, en todos los escenarios, resultaron significativamente distintos de uno, por lo que es necesario incluirlos). Como se puede ver, la muestra de elección discreta parece complementarse de mejor forma con las respuestas *best* (el valor λ_a es más parecido al de χ^2_8), pero aun así se rechaza la hipótesis que los ocho atributos son comunes. Es interesante mencionar que aun cuando el test formal es rechazado, la diferencia entre χ^2_8 y λ_a no es tan grande como se ha encontrado en otras aplicaciones (Flynn *et al.*, 2013); esto es indicativo de un grado de similitud sorprendente entre las pendientes comparables de los dos tipos de encuesta², como también se ha visto en otros casos (Louviere y Swait; 1993).

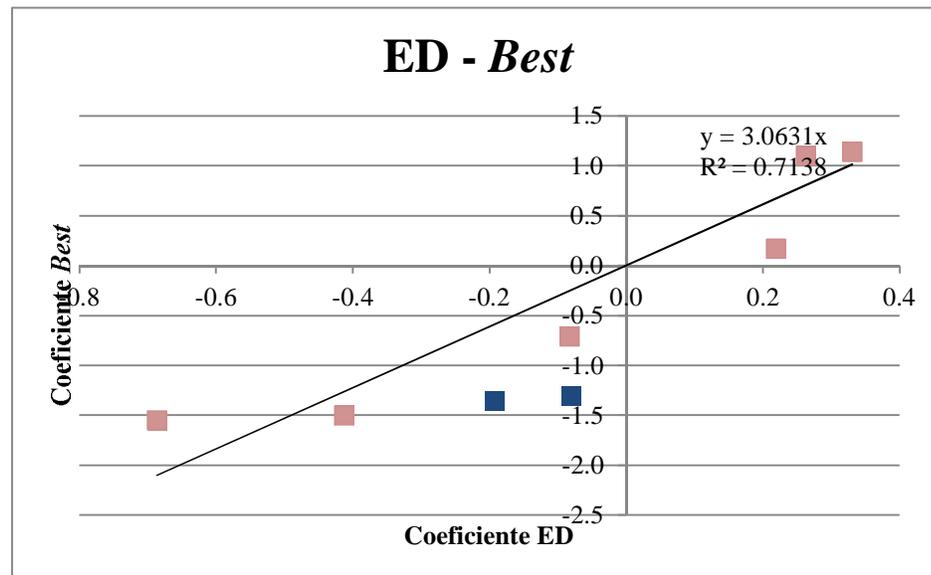
² Louviere *et al.* (2000, Cap. 13), demuestran empíricamente que este tipo de regularidad entre métodos de elicitación y contexto no es inusual.

Tabla 4: Test de Razón de Verosimilitud para Modelos de Elección Discreta y *Best-Worst*

	ED, Best, ED+B	ED, Worst, ED+W
L_1	-1.003	-1.003
L_2	-2.151	-2.156
L_μ	-3.169	-3.179
λ_a	30	40
χ^2_8	15,51	15,51
Test	Rechaza	Rechaza

Observando la Figura 2 es posible intuir cuáles serían los atributos que parecen no ser comunes entre las muestras de elección discreta binaria y *best*. De hecho, el coeficiente del atributo *Distancia a áreas verdes* parece ser el más diferente en ambos tipos de encuesta, seguido por el de *Distancia a servicios básicos* (ambos marcados en azul en la Figura 2).

Figura 2: Gráfico de Coeficientes Elección Discreta y Best



Se comenzó por considerar el primer atributo como específico a cada muestra, obteniéndose una log verosimilitud (L_μ) de -3.164 y un $\lambda_a = 19$, mayor que $\chi^2_7 = 14,07$, por lo que se continuaba rechazando la hipótesis nula. Se procedió, entonces, a considerar a ambos atributos como específicos a cada muestra, obteniéndose una log verosimilitud (L_μ) de -3.161 y $\lambda_a = 12$, para $\chi^2_6 = 12,59$, por lo que se aceptó la hipótesis nula que este modelo efectivamente era apropiado.

La Tabla 5 presenta los coeficientes del modelo combinado de elección discreta binaria y *best*. El *Intercepto* “*Si lo arrendaría*” corresponde a la constante específica asociada a las respuestas de elección binaria. Para cada atributo, se presenta el coeficiente de su intercepto (e.g. *Intercepto Metro*) y de su sensibilidad (e.g. *Metro*). Para el atributo *Servicios comerciales básicos* no hay un intercepto debido a que éste es considerado como base para estimar el modelo. La última fila de la Tabla 5 muestra el factor de escala agregado a las respuestas *best*.

Tabla 5: Resultados Modelo Elección Discreta + *Best*

	Coefficiente	Error Estándar	Test t
Intercepto "Si lo arrendaría"	-0,543	0,058	-9,4
Intercepto Metro	0,269	0,060	4,5
Metro	-0,413	0,041	-10,0
Intercepto Áreas Verdes	-0,085	0,067	-1,3
Áreas Verdes (Elección Discreta)	-0,087	0,054	-1,6
Áreas Verdes (<i>Best</i>)	-0,359	0,063	-5,7
Intercepto Dividendo	0,044	0,055	0,8
Dividendo	-0,472	0,075	-6,3
Intercepto Piscina	-0,177	0,063	-2,8
Piscina	0,115	0,036	3,2
Intercepto Limpieza	-0,273	0,069	-4,0
Limpieza	0,324	0,045	7,3
Servicios Básicos (Elección Discreta)	-0,186	0,054	-3,5
Servicios Básicos (<i>Best</i>)	-0,375	0,058	-6,5
Intercepto Cultura	-0,086	0,057	-1,5
Cultura	-0,149	0,034	-4,5
Intercepto Tamaño	0,504	0,069	7,3
Tamaño	0,299	0,032	9,4
Factor de Escala <i>Best</i>	3,628	0,366	7,2

5. VALIDACIÓN

Utilizando la muestra de validación se puede analizar la capacidad predictiva del modelo ED+*Best*. Además, se puede ver cuál de los coeficientes estimados para los atributos *Distancia a áreas verdes* y *Distancia a servicios básicos* sería más adecuado para ella. Es necesario mencionar que debido como la muestra de validación considera distintos individuos que la base de datos utilizada para estimar los modelos, es pertinente estimar un factor de escala para las utilidades obtenidas mediante los coeficientes del modelo original ED+*Best* para lograr una mejor adaptación a la muestra.

En primer lugar se calculó la log verosimilitud asociada a estimar los parámetros del modelo ED usando la muestra de validación; esta tomó el valor de -3.337. Este valor será comparado más adelante con la log verosimilitud obtenida al aplicar el modelo ED+*Best* a esta muestra.

Luego, se examinó que coeficientes (del par no común) se adaptaban mejor a la muestra de validación, obteniéndose que para el atributo *Distancia a áreas verdes* es mejor utilizar el coeficiente estimado mediante las respuestas *Best* y para el atributo *Distancia a servicios básicos* es mejor utilizar el coeficiente estimado mediante las respuestas ED. En la Tabla 6 se muestran los principales indicadores obtenidos de predecir las elecciones observadas en la muestra de validación mediante el modelo ED+*Best* escalado.

Tabla 6: Indicadores de la Predicción Modelo ED+Best en Muestra de Validación

	Factor de Escala (test t)	Log-Verosimilitud	FPR	ER	CR	χ^2 FPR (gl)
ED+Best	1,56 (9,3)	-3.690	0,695	0,641	0,5	10,65 (4)
ED	1,66 (9,1)	-3762	0,659	0,631	0,5	144,17 (4)

La log verosimilitud del modelo estimado con la muestra original tiene una diferencia de sólo 10,6% respecto a la óptima (-3.337) para la muestra de validación. Además, el modelo ED+Best recupera un 69,5% de las primeras preferencias de la muestra (FPR), siendo este valor similar a su valor esperado (ER) y mayor que el de una predicción al azar (CR), lo que permitiría concluir que el modelo es razonable e informativo (Ortúzar y Willumsen, 2011). No obstante, el test χ^2_{FPR} ³ entrega un valor de 10,65 (para cuatro grados de libertad) que es levemente superior al valor crítico ($\chi^2_4 = 9,49$) con un índice de valor crítico $\alpha = 0,05$; de hecho, el test se aprobaría para un valor crítico de $\alpha = 0,025$.

En la segunda fila de la Tabla 6 se puede ver la predicción mediante los coeficientes estimados a través de las respuestas ED en la muestra de validación. Se puede ver que efectivamente la unión de las respuestas Best y ED mejora la capacidad predictiva en todos los indicadores.

6. ANÁLISIS DE RESULTADOS

Con la información anterior se puede construir la escala mostrada en la Figura 3, que muestra tanto la significancia de cada atributo (intercepto) como la del coeficiente de cada nivel de los atributos mostrados al encuestado. Para simplificar la figura, los coeficientes de los niveles del atributo *Dividendo del departamento* se muestran sólo como dos niveles: bajo (intercepto - sensibilidad) y alto (intercepto + sensibilidad), en vez de presentar los ocho niveles distintos.

En la figura, los atributos más cercanos al origen (ya sea como intercepto o coeficiente) son aquellos menos relevantes (tienen un peso menor en la función de utilidad), al contrario de los que se encuentran más lejos del origen. Así, se pueden ver claramente las diferencias entre los pesos de los atributos y los pesos de los niveles utilizados. Por ejemplo, uno de los casos más interesantes corresponde al *Dividendo del departamento*, que como atributo no resulta muy importante para el individuo (el valor de su intercepto es cercano a cero), pero debido a los niveles presentados en la encuesta resulta finalmente ser un atributo de gran impacto (los coeficientes del nivel alto y bajo se encuentran muy lejos del origen). De hecho, el impacto que tienen los niveles presentados al individuo es el más significativo de todos los atributos.

Hay otros atributos, como el *Tamaño de dormitorios*, que resultan ser significativos tanto como atributos como por su escala, ya que su intercepto está lejano al origen y los niveles (amplios y normales) se encuentran lejos de donde se encuentra el intercepto. Esto mismo ocurre con *Mantención y estado de conservación de la calle*, y *Acceso al metro*. Por otro lado, *Gimnasio y piscina en el edificio* resultó ser uno de los atributos más cercanos al origen en ambos casos, esto es, tanto en su propio peso como por la variación en sus niveles.

³ Para detalles sobre el test χ^2_{FPR} y la relación entre FPR, ER y CR, ver Ortúzar y Willumsen (2011, Cap. 8).

Finalmente, también resulta interesante analizar las diferencias en la valoración de atributos ocasionadas por aplicar dos tipos de encuestas, y estudiar sus posibles causas. En la Figura 3 se muestran los coeficientes que mejor se adaptaron a la muestra de validación, pero esta diferencia se puede ver a través de los coeficientes estimados en los dos tipos de encuesta para los atributos específicos. En este caso dos atributos – que se muestran en la Tabla 7 - resultaron ser específicos a cada tipo de encuesta y por lo tanto resultaron tener una distinta valoración.

Figura 3: Escala de Importancia de Atributos y Niveles de los Atributos

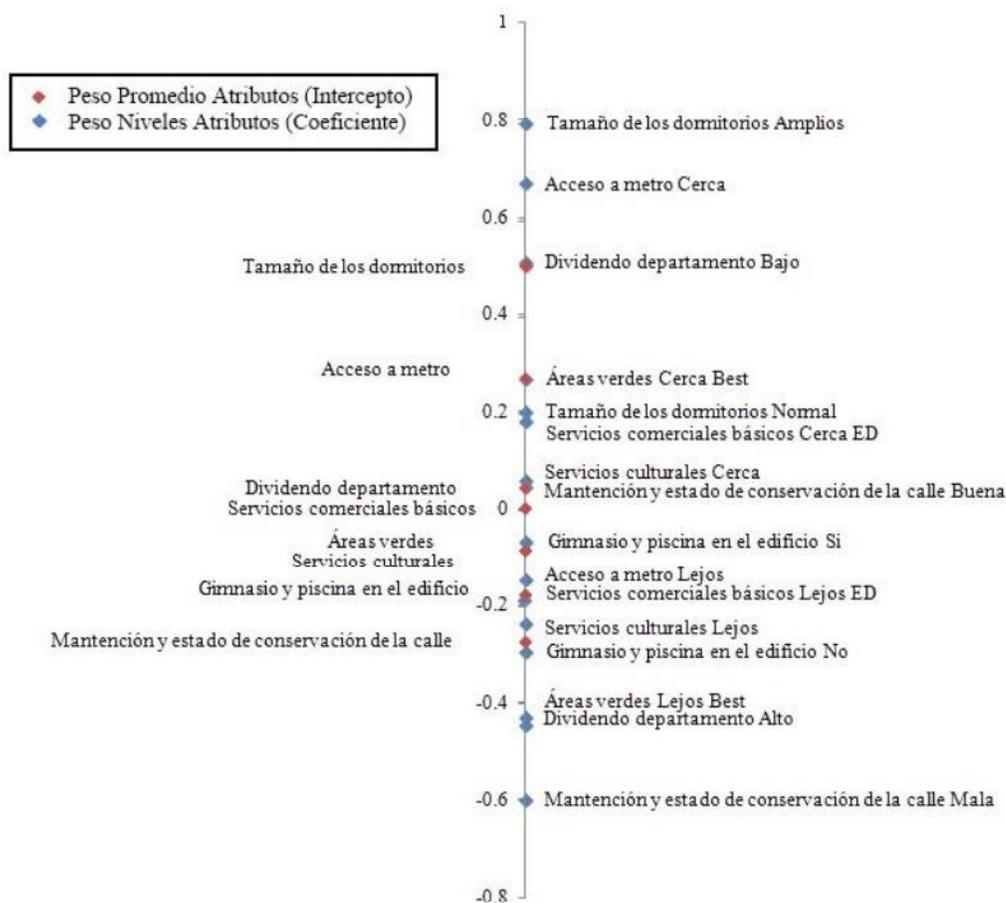


Tabla 7: Sensibilidades Atributos Específicos en Encuestas *Best* y ED

Sensibilidad	<i>Best</i>	ED
<i>Distancia a áreas verdes</i>	-0,359	-0,087
<i>Distancia a servicios comerciales básicos</i>	-0,375	-0,186

En el caso de las *Áreas verdes* el atributo recibió mayor importancia en las respuestas *best* que en las de elección binaria. Esto podría mostrar que cuando a un individuo le preguntan cuánto le importa tener áreas verdes, muestra mayor interés (o le asigna mayor importancia, eligiéndolo como el mejor atributo) que cuando le piden escoger entre dos alternativas de departamento (en

que uno tiene áreas verdes más cerca que el otro). Es interesante notar que además la capacidad predictiva del modelo en la muestra de validación mejoró al considerar este atributo como el estimado mediante las respuestas *best*.

Algo similar, pero con menor diferencia, ocurre en el caso de *Servicios comerciales básicos*, aunque en este caso resultó ser mejor considerar el coeficiente estimado por las respuestas ED al predecir el comportamiento de los individuos en la muestra de validación.

7. CONCLUSIONES

El gran aporte que hace la encuesta *best-worst* es permitir diferenciar entre la importancia de los atributos y la importancia de los niveles utilizados en su presentación. En la Figura 3 se muestran estas importancias por separado en una escala, habiendo logrado así complementar la encuesta *best-worst* con la de elección binaria para agregar más información a la valoración de atributos de las viviendas del centro de Santiago.

Nuestro objetivo en este artículo era estudiar la potencial complementariedad de la encuesta *best-worst* con la de elección binaria. Por este motivo, en los modelos planteados no se incluyó variables socio-económicas que podrían explicar parte del comportamiento de los individuos. En este caso interesaba únicamente mostrar cómo se pueden unir estas dos bases de datos para maximizar la información deducible del modelo estimado. Así, queda propuesta como futura investigación la inclusión de variables adicionales que podrían explicar parte de los errores, preferencias y gustos de los encuestados.

Encontramos que las respuestas *best* se complementaban de mejor manera con las de elección binaria. La complementariedad de respuestas se logró usando seis atributos comunes y dos, *Distancia a áreas verdes* y *Distancia a servicios básicos* específicos a cada conjunto de datos. Es decir, estos dos atributos presentaron diferentes valoraciones dependiendo de la manera en que se hacía la elección. En ambos casos, el atributo se valoraba más cuando se le pedía a la persona que escogiera el mejor y peor atributo, que cuando su valor se deducía a partir de las respuestas de elección binaria.

Por otro lado, se encontró que el atributo *Dividendo del departamento* tenía mayor importancia debido a la diferencia en sus niveles que como atributo en sí. Los atributos *Tamaño de dormitorios*, *Mantención y estado de conservación de la calle*, y *Acceso al metro* resultaron ser significativos tanto en sí como por su escala, siendo el primero el más significativo de todos. Finalmente, el atributo *Gimnasio y piscina en el edificio* resultó ser poco significativo tanto como atributo como por sus niveles.

Agradecimientos

Agradecemos al Fondo Nacional de Desarrollo Científico y Tecnológico (Proyecto FONDECYT 1100068), al Instituto Milenio en Sistemas Complejos de Ingeniería (ICM: P05-004F; FONDECYT: FB016) y al Centro de Desarrollo Urbano Sustentable–CEDEUS (Conicyt/Fondap/15110020), por el apoyo otorgado a esta investigación. También deseamos agradecer el apoyo de Ignacia Torres quien, además, recolectó tanto la base de datos utilizada en esta investigación como la muestra de validación.

Referencias

- Couzner, L., Ratcliffe, J., Lester, L., Flynn, T. y Crotty, M. (2013) Measuring and valuing quality of life for public health research: application of the ICECAP-O capability index in the Australian general population. *International Journal of Public Health* **58**, 367-376.
- Finn, A. y Louviere, J.J. (1992) Determining the appropriate response to evidence of public concern: the case of food safety. *Journal of Public Policy and Marketing* **11**, 12-25.
- Flynn, T.N., Louviere, J.J., Peters, T.J. y Coast, J. (2010) Using discrete choice experiments to investigate heterogeneity in preferences for quality of life: variance scale heterogeneity matters. *Social Science and Medicine* **70**, 1957-1965.
- Flynn, T.N., Peters, T.J. y Coast, J. (2013) Quantifying response shift or adaptation effects in quality of life by synthesising best-worst scaling and discrete choice data. *Journal of Choice Modelling* **6**, 34-43.
- Louviere, J.J., Hensher, D.A. y Swait, J.D. (2000) *Stated Choice Methods: Analysis and Applications*. Cambridge University Press, Cambridge.
- Louviere, J.J., Street, D., Burgess, L., Wasi, N., Islam, T. y Marley, A.A. (2008) Modelling the choices of individual decision - makers by combining efficient choice experiment designs with extra preference information. *Journal of Choice Modelling* **1**, 128-163.
- Louviere, J.J. y Swait, J.D. (1993) The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research* **30**, 305-314.
- Louviere, J.J. y Swait, J.D. (1997) Separating weights and scale values in conjoint tasks using choices of best and worst attribute levels by separating weights and scale values in conjoint tasks using choices of best and worst attribute levels abstract. Working Paper, Centre for the Study of Choice, University of Technology Sydney.
- Marley, A.A. y Louviere, J.J. (2005) Some probabilistic models of best, worst, and best-worst choices. *Journal of Mathematical Psychology* **49**, 464-480.
- Ortúzar, J. de D. y Willumsen, L.G. (2011) *Modelling Transport*. Cuarta Edición, John Wiley and Sons, Chichester.
- Torres, I. (2009) Valoración de Atributos de Vivienda y Barrio en el Centro de Santiago. *Tesis de Magister en Ciencias de la Ingeniería*, Departamento de Ingeniería de Transporte y Logística, Pontificia Universidad Católica de Chile.
- Torres, I., Greene, M. y Ortúzar, J. de D. (2013) Valuation of housing and neighbourhood attributes for city centre location: a case study in Santiago. *Habitat International* **39**, 62-74.