

MODELOS CON ESCALAMIENTO BEST-WORST Y CONSIDERACIONES DE ENDOGENEIDAD: UNA APLICACIÓN AL DESARROLLO INMOBILIARIO EN DENSIDAD CON INTEGRACIÓN SOCIAL¹

Natan Waintrub, Pontificia Universidad Católica de Chile nwaintrub@puc.cl

Juan de Dios Ortúzar, Pontificia Universidad Católica de Chile jos@ing.puc.cl

Luis Ignacio Rizzi, Pontificia Universidad Católica de Chile lir@ing.puc.cl

Joffre Swait, University of South Australia joffre.swait.jr@gmail.com

RESUMEN

Los métodos escalamiento *best-worst* han sido usados para estimar modelos y atributos con mayor significancia a los modelos tradicionales. Por otro lado, la endogeneidad entre respuestas interrelacionadas puede producir un sesgo en los modelos. Utilizando una encuesta de preferencias declaradas (con preguntas de escalamiento *best-worst*) sobre la disposición a densificar, con y sin integración social, por parte de desarrolladores inmobiliarios, se estiman modelos que logran juntar las respuestas de distinta procedencia. También se estima un modelo que soluciona el problema de endogeneidad intrínseco en la encuesta, a fin de obtener la probabilidad para que desarrolladores inmobiliarios densifiquen con integración social.

Palabras clave: *escalamiento best-worst, endogeneidad, modelos de elección discreta*

Best-worst scaling had been used for estimating models with more significant attributes than traditional discrete choice models. On the other side, endogeneity between inter-related responses may bias the models. With a stated preference survey (with best-worst scaling questions) of the disposition of real-estate developers for density developments, with and without social integration, we estimate models that mix answers from different background. We also estimate a model that solves the endogeneity problem of the survey, so we could obtain the probability that real-estate developers create density projects with social integration.

Keywords: *best-worst scaling, endogeneity, discrete choice models*

¹ Parte de este trabajo está basado en la tesis para optar al grado de Magíster en Ciencias de la Ingeniería de Natan Waintrub Santibáñez “Disposición a recibir incentivos para desarrollo inmobiliario en densidad en torno a estaciones de metro y corredores de buses”, 2015, Pontificia Universidad Católica de Chile.

1. INTRODUCCIÓN

Los métodos de elección discreta han sido ampliamente utilizados para comprender las elecciones individuales, por medio de la estimación de modelos que representan el proceso de decisión. Basados en la teoría de la utilidad aleatoria (McFadden, 1974; Ortuzar y Willumsen, 2011), donde cada individuo escoge la alternativa que le genera mayor utilidad dados los parámetros de su función de utilidad. Así, estos métodos estiman el valor de los parámetros teóricamente utilizados en las elecciones individuales, a fin de cuantificar la importancia de cada parámetro en la decisión. En esta materia, han surgido distintos métodos que persiguen identificar con mayor precisión estos parámetros, como el método de escalamiento *best-worst* propuesto por Louvier y Swait (1997).

El método de escalamiento *best-worst* permite determinar los atributos y niveles de los modelos con mayor significancia que los modelos tradicionales de elección discreta, cuyos atributos pueden resultar poco significativos debido a que sus niveles no varían lo suficiente sobre el perfil de elección. Posibles ventajas sobre los modelos de elección discreta tradicionales son (Marley y Louviere, 2005): (i) la gran cantidad de información obtenida sobre el ranking de los individuos, (ii) la propensión del individuo a identificar y responder sobre los valores extremos, y (iii) la elección resulta ser una tarea fácil para los individuos.

Los métodos best-worst han sido usados en estudios muy diversos, por ejemplo, relacionados con la salud (Flynn et al., 2007; Whitty et al. 2013; Lancsar et al. 2013), marketing (Louviere et al. 2013), elección de productos (Louviere et al., 2008; Lagerkvist, 2013) y elección de departamentos en el centro de Santiago (Balbontín et al., 2014).

En general, los modelos de elección discreta tradicional, así como modelos con escalamiento *best-worst*, hacen el supuesto que las variables explicativas son independientes de los factores no observados (Train, 2009). Sin embargo, es posible que las variables explicativas estén correlacionadas entre sí o con factores no observados. Existen varios tipos de endogeneidad. Por ejemplo, la asociada a individuos que toman decisiones de forma estratégica (Louviere et al., 2005). Este es el caso de encuestas cuyo resultado podría tener consecuencias de política reales.

En particular, la endogeneidad puede considerarse como la influencia de decisiones pasadas sobre la elección actual (Train, 2009). En este sentido, Train y Wilson (2008) sugieren la base teórica para estimar modelos pivoteados y contrarrestar ese sesgo. En profundidad, los modelos pivoteados corresponden inicialmente en una elección de preferencias reveladas (PR), para luego realizar una serie de preguntas de tipo preferencias declaradas (PD), siendo las alternativas de esta última pivoteadas sobre los valores de la alternativa PR.

La presente investigación tiene por objetivo: (i) estimar modelos conjuntos entre respuestas *best-worst* y elección discreta, e (ii) incorporar los efectos de endogeneidad entre decisiones tomadas *a priori*. No obstante, la estimación de los modelos antes señalados no consideran el sesgo que se asume frente a modelos más complejos que consideren, por ejemplo, la dependencia entre respuestas relacionadas de un individuo (i.e. pseudo efecto panel, Ortúzar y Willumsen, 2011).

2. MARCO TEORICO

Los modelos best-worst desarrollados en esta sección corresponden a una aplicación diferente de los modelos de elección discreta tradicional, por lo que se dará por entendido su versión tradicional (ver Ortúzar y Willumsen, 2011 para su versión tradicional). Por otro lado, en esta

sección se explicita una versión particular de la endogeneidad (el lector interesado puede revisar Train, 2009 y Louviere *et al.*, 2005).

2.1. Modelos *best-worst*

El método *best-worst* explicita tres casos en que puede ser utilizado (Marley y Louviere, 2005). De forma general, cada individuo enfrentado a una elección deberá escoger el mejor (*best*) y peor (*worst*) bien hedónico definido por un perfil o conjunto de atributos. De esta forma, se logra reportar gran cantidad de información sobre las preferencias individuales. Además, tiene la ventaja de que las elecciones *best-worst* son sencillas de responder y ello aliviana la carga cognitiva a la que se enfrenta el encuestado.

La presente investigación se enfoca en la aplicación del segundo caso de escalamiento *best-worst*. Cada individuo se enfrentará a una situación con atributos y niveles definidos, debiendo escoger los atributos *best* y *worst*, obviando el perfil de atributos mostrado. Es decir, el individuo escogerá por un nivel determinado de cada atributo. En general, cada atributo posee una cantidad definida de niveles, pero en cada situación se muestra únicamente uno de ellos.

Este enfoque permite diferenciar la importancia relativa de los atributos y la variación (escala) debida a sus niveles. Tradicionalmente, los modelos de elección discreta pueden estimar un atributo que resulte poco significativo. Si bien es posible que el atributo no sea relevante, puede suceder que los niveles del atributo no varíen lo suficiente en todo el perfil de elección. Así, aun cuando el atributo fuera relevante en realidad, no se podría rescatar su importancia en el modelo. Alternativamente en marketing, por ejemplo, un producto que tenga un atributo con niveles muy distintos (e.g. marca mal posicionada y bien posicionada), podría recibir una importancia relativa muy grande. Sin embargo, si se invierte en mejorar su escala (i.e. mejorar su posicionamiento), disminuiría la diferencia entre los niveles del atributo con los del mercado, pero probablemente su participación de mercado no aumentaría.

Los modelos de tipo *best-worst*, a diferencia de los modelos de elección discreta tradicional (ED), estima una función de utilidad única por atributo. Cada función de utilidad tiene un intercepto, que representa su significancia relativa al conjunto de atributos, y puede ser el promedio de los niveles definidos según la codificación utilizada. Por otro lado, cada función tiene una pendiente (escala) que representa la sensibilidad de la utilidad a la variación de los niveles del atributo.

Así, las funciones de utilidad por atributo en modelos *best-worst* poseen diferencias sobre la modelación tradicional ED. En la primera, el intercepto señala la preferencia promedio por ese atributo; mientras que en la segunda, el intercepto es común al conjunto de atributos, debido a que se elige una alternativa correspondiente a un paquete de atributos. Por otro lado, la utilidad marginal de cada atributo en las funciones de utilidad ED es igual a la sensibilidad (pendiente) de los niveles en las funciones de utilidad *best-worst*. Finalmente, la estimación de las funciones de utilidad de los atributos en *best* o *worst* se posicionan en la misma escala de utilidad relativa, siendo innecesaria la inclusión de un factor de escala para relacionarlos.

A escala individual, Louviere y Swait (1997) señalan que un individuo q escogerá un atributo como *best* si la función de utilidad de ese atributo es mayor que la de los restantes del perfil. De forma inversa, el individuo q escogerá un atributo como *worst* si ese atributo le reporta la menor utilidad al perfil. Por lo tanto, en teoría, la probabilidad de elegir un atributo *best* es la función inversa de elegir ese atributo como *worst*. Así, los niveles de los atributos en el caso *worst* deberán codificarse con signo contrario al caso *best*. Finalmente, es importante considerar que el proceso mental que el individuo realiza al momento de escoger los atributos *best* puede ser distinto de cuando escoge *worst*. Esto se ve reflejado en el valor de los parámetros de la función

de utilidad en ambos casos. Para comprobar la viabilidad de utilizar los mismos parámetros, se puede realizar el test de razón verosimilitud (Ortúzar, 2000).

Si bien las respuestas *best*, *worst* y ED son estimadas por separado, es posible comparar sus sensibilidades bajo el supuesto que miden la misma variable. Más aún, estudios recientes (Flynn *et al.*, 2013) señalan que no existe evidencia estadística para estimar que estos parámetros deban ser distintos para cada tipo de datos.

En relación a lo anterior, las funciones de utilidad para las respuestas son :

Best (B)

$$\begin{aligned} U_{B_1} &= \theta_{B_1} + \beta_{B_1} \cdot X_1 + \varepsilon_B \\ &\vdots \\ U_{B_m} &= \theta_{B_m} + \beta_{B_m} \cdot X_m + \varepsilon_B \end{aligned} \quad (1)$$

Worst (W)

$$\begin{aligned} U_{W_1} &= -\theta_{W_1} + \beta_{W_1} \cdot (-X_1) + \varepsilon_W \\ &\vdots \\ U_{W_m} &= -\theta_{W_m} + \beta_{W_m} \cdot (-X_m) + \varepsilon_W \end{aligned} \quad (2)$$

Elección discreta (ED)

$$\begin{aligned} U_{ED_A} &= \beta_{ED_0} + \beta_{ED_1} \cdot X_1 + \cdots + \beta_{ED_m} \cdot X_m + \varepsilon_{ED} \\ U_{ED_B} &= \varepsilon_{ED} \end{aligned} \quad (3)$$

donde m es la cantidad de atributos presentados en cada perfil, θ es el *intercepto* de cada alternativa y β es la *sensibilidad* de los atributos. Ahora, si el test de razón verosimilitud no rechaza la hipótesis nula, se concluye que no existe evidencia estadística suficiente para rechazar la similitud entre los *interceptos* y/o *sensibilidades*. Entonces, es posible relacionar los *interceptos* de las funciones *best* y *worst*, y las sensibilidades de las respuestas *best*, *worst* y ED. Es decir:

Interceptos

$$\begin{aligned} \theta_{B_1} &= \theta_{W_1} \\ &\vdots \\ \theta_{B_m} &= \theta_{W_m} \end{aligned} \quad (4)$$

Sensibilidades

$$\begin{aligned} \beta_{B_1} &= \beta_{W_1} = \beta_{ED_1} \\ &\vdots \\ \beta_{ED_m} &= \beta_{ED_m} = \beta_{ED_m} \end{aligned} \quad (5)$$

Además, se asume que los errores $\varepsilon_B, \varepsilon_W, \varepsilon_{ED} \sim IID Gumbel(0, \lambda)$.

2.2. Endogeneidad en respuestas interrelacionadas

En general, los modelos de elección discreta tradicional, y escalamiento *best-worst*, hacen el supuesto que las variables explicativas son independientes de los factores no observados (Train, 2009). Sin embargo, es posible que las variables explicativas estén correlacionadas entre sí o con factores no observados y, finalmente, con el término de error ε . Una forma particular de la endogeneidad considera el efecto de decisiones interrelacionadas, las que ocurren cuando la decisión de elección actual de una persona está relacionada con una elección realizada anteriormente por ella. Por ejemplo, Train (2009) señala que esto sucede en la elección de modo de transporte y de vivienda; un individuo puede tener mayor propensión a escoger transporte público si, en una decisión previa, escogió una vivienda cerca de una estación de Metro.

Train y Wilson (2008) comentan la relevancia del problema de endogeneidad anterior para los casos de preguntas PR que conducen a preguntas de PD. Estas preguntas pueden estar

interrelacionadas de dos formas; respuestas pivoteadas y sp-off-rp . El primer caso se utiliza, por ejemplo, para evaluar opciones de ruta por medio de una pregunta PR. El individuo elige la alternativa i , correspondiente a la ruta revelada, y luego se enfrenta a una elección de tipo PD. Dentro de las alternativas a la pregunta PD, se utiliza la respuesta a la pregunta PR (alternativa i) como “su viaje reciente”, junto con algunas alternativas hipotéticas. El supuesto principal es que el individuo, al verse enfrentado nuevamente a la misma alternativa del caso PR en la elección PD, considerará los mismo coeficientes y sus errores. En particular, el individuo evaluará la pregunta PR con una función de utilidad del tipo (Train y Wilson, 2008):

$$U_i = \beta \cdot x_i + \varepsilon_i \quad (6)$$

donde $\varepsilon_i \sim IID Gumbel (0, \lambda)$ e i corresponde a la alternativa escogida. Bajo el supuesto que el individuo evalúa la alternativa k de la pregunta PD de forma idéntica a la alternativa i de PR, se obtiene la siguiente función de utilidad para la alternativa i en PD:

$$W_k = \alpha \cdot \beta \cdot x'_k + \alpha \cdot \delta_{ki} \cdot \varepsilon_i + \eta_k \quad (7)$$

donde α captura la variación de los parámetros PD en relación a los parámetros PR, η_i es valor extremo iid con escala $1/\alpha$, ε_i corresponde a los atributos no observados de la alternativa i escogida en la respuesta RP y δ_{ji} toma el valor uno si la alternativa k es idéntica a la alternativa i , y cero en caso contrario. Debido a que ε_i no puede ser observado, las probabilidades deben integrarse sobre su distribución condicional como:

$$P_{k|i} = \int \frac{e^{\alpha\beta x'_k + \alpha\delta_{ki}\varepsilon_i + \eta_k}}{\sum_l e^{\alpha\beta x'_l + \alpha\delta_{li}\varepsilon_i + \eta_l}} f(\varepsilon_i | \beta x_i + \varepsilon_i > \beta x_j + \varepsilon_j \forall j \neq i) d\varepsilon_i \quad (8)$$

Esta probabilidad corresponde a un modelo Logit Mixto (ML), condicionado a la densidad de ε , y puede ser simulado en programas computacionales calculando la probabilidad logit de cada extracción², y promediando el resultado.

Finalmente, para obtener la probabilidad P_{ki} se debe resolver la siguiente ecuación:

$$P_{ik} = P_{k|i} \cdot P_i \quad (9)$$

Este trabajo considera una pregunta PD inicial que luego será pivoteadas para obtener otra pregunta PD posterior, manteniendo la alternativa inicial escogida por el individuo.

2.3. Estimación de modelos con datos mixtos

Como se mencionó anteriormente, los parámetros de las respuestas best, worst y ED pueden ser considerados iguales. Sin embargo, debido a que las respuestas provienen de tipos de elección distintas, debe considerarse la inclusión de un factor de escala λ , que permite suponer independencia entre los errores ε de las funciones best, worst y ED, según:

² Corresponde a los valores de las variables aleatorias cuando se realiza una simulación en modelos Logit.

$$\begin{aligned}\varepsilon_B &\sim \text{IID Gumbel}(0, \lambda_B) \\ \varepsilon_W &\sim \text{IID Gumbel}(0, \lambda_W) \\ \varepsilon_{ED} &\sim \text{IID Gumbel}(0, \lambda_{ED})\end{aligned}$$

obteniendo las utilidades:

Best (B)

$$\begin{aligned}U_{B_1} &= \lambda_B \cdot (\theta_{B_1} + \beta_{B_1} \cdot X_1 + \varepsilon_B) \\ &\vdots \\ U_{B_m} &= \lambda_B \cdot (\theta_{B_m} + \beta_{B_m} \cdot X_m + \varepsilon_B)\end{aligned}\tag{10}$$

Worst (W)

$$\begin{aligned}U_{W_1} &= \lambda_W \cdot (-\theta_{W_1} + \beta_{W_1} \cdot (-X_1) + \varepsilon_W) \\ &\vdots \\ U_{W_m} &= \lambda_W \cdot (-\theta_{W_m} + \beta_{W_m} \cdot (-X_m) + \varepsilon_W)\end{aligned}\tag{11}$$

Elección discreta (ED)

$$\begin{aligned}U_{ED_A} &= \lambda_{ED} \cdot (\beta_{ED_0} + \beta_{ED_1} \cdot X_1 + \dots + \beta_{ED_m} \cdot X_m + \varepsilon_{ED}) \\ U_{ED_B} &= \lambda_{ED} \cdot (\varepsilon_{ED})\end{aligned}\tag{12}$$

donde λ_i es un factor de dispersión relacionada a la desviación estándar σ_i (Ortúzar, 2000):

$$\lambda_i = \frac{\pi}{\sigma_i \sqrt{6}}\tag{13}$$

con i independiente para *best*, *worst* y ED.

De la ecuación (13) se desprende que, a mayor escala, menor es la desviación estándar. Sin embargo, la estimación conjunta de todos los factores de escala λ_i presenta un problema de identificabilidad³. Este puede solucionarse considerando alguno de los λ_i con valor fijo e igual a uno y estimando los factores de escala restantes. Separar los valores de escala λ_i permite explicar los errores ε_i provenientes de fuentes de datos con desviación estándar σ_i distinta.

Juntar las respuestas de origen diferente permite estudiar si los parámetros señalados anteriormente pueden considerarse comunes. De esta forma, para determinar si se pueden juntar, se utiliza el test de razón verosimilitud (Ortúzar, 2000)

3. BANCO DE DATOS

Se utilizó una base de datos recolectada por DICTUC (2014) y Waintrub (2015), donde fueron entrevistados 75 desarrolladores inmobiliarios sobre su disposición a construir proyectos de densificación en altura, que consideren integración social, en torno a estaciones de Metro y corredores de buses en la ciudad de Santiago. En el primer estudio fueron encuestados 52 desarrolladores inmobiliarios respecto a tres estaciones de Metro, de un pool de 10 estaciones. Estas estaciones fueron seleccionadas en conjunto con un grupo de expertos urbanistas. En el segundo estudio fueron encuestados 25 desarrolladores inmobiliarios sobre tres “estaciones” de corredores de buses, de un pool de cuatro estaciones; del total de encuestados, 22 correspondían a los mismos desarrolladores del primer estudio.

Cada encuesta tiene como variables siete incentivos relacionados al desarrollo de vivienda en densidad, donde: (i) dos incentivos tienen efecto directo en la demanda por viviendas (i.e. las familias que desean comprar una vivienda), (ii) cuatro incentivos tienen efecto directo a la oferta

³ Este problema aparece cuando los programas de estimación son incapaces de estimar algún parámetro ya que se confunde el efecto de otros parámetros.

de vivienda (i.e. desarrolladores inmobiliarios), y (iii) un incentivo es indirecto a la oferta⁴. Los primeros corresponden a bonos que el comprador de vivienda puede adquirir; los segundos son modificaciones a la norma de construcción en la zona encuestada y un incentivo correspondiente a un bono para la oferta sujeto a ciertas condiciones. El último corresponde a inversión estatal en forma de áreas verdes y mobiliario público en el entorno de la estación.

Además, en la encuesta se realiza una pregunta sobre la disposición a construir con integración social, entendida como un 30% mínimo que el desarrollador inmobiliario debe comprometer de su proyecto inmobiliario si desea un aumento relativo en los incentivos otorgados.

Los incentives fueron seleccionados con la ayuda de expertos (del gobierno y urbanistas) y fueron categorizados como: (i) incentivos directos a la demanda, (ii) incentivos directos a la oferta, y (iii) incentivos indirectos a la oferta. Todos estos incentivos son de tipo “zanahoria” (Walle *et al.*, 2004), por lo que *a priori*, se espera que sean atractivos para los desarrolladores inmobiliarios.

Incentivos directos a la demanda:

- a. Bono a cualquier comprador (BCC). Se entrega un monto en UF a la familia que desee comprar un departamento en torno a un radio de 500 metros de la estación, siempre que su valor sea menor o igual a 2000 UF.
- b. Bono a cualquier comprador con subsidio (BCS). Este bono es un monto en UF que se entrega (por sobre el subsidio habitacional) a familias que ya cuenten con subsidio, o puedan obtenerlo. Las viviendas asociadas a este bono no deben exceder las 1400 UF.

Incentivos directos a la oferta:

- c. Bono por tiempo limitado (BTL). En torno a un radio de 500 metros de la estación, se entrega un monto en UF por cada departamento construido del proyecto al desarrollador inmobiliario, siempre que obtenga el permiso de edificación en un plazo menor a dos años de lanzada la norma.
- d. Aumento en el factor de constructibilidad (CON). Considera un aumento en los metros cuadrados que un proyecto puede construir, por medio de un aumento porcentual en este factor. El aumento debe ser traducido en forma de mayor altura edificable u otra medida, como variación en el porcentaje de utilización del suelo del proyecto.
- e. Aumento en densidad (DEN). Este incentivo implica una mayor cantidad de unidades en el departamento.
- f. Disminución del mínimo de estacionamientos (EST). En Santiago, la normativa de estacionamientos exige que cada departamento cuente con, al menos, un número de estacionamientos proporcional a la cantidad de departamentos construidos. Este incentivo permite reducir el número de estacionamientos en forma proporcional a la exigencia total.

Incentivo indirectos a la oferta:

- g. Inversión pública en espacio público (IPU). Consiste en un monto en UF, por cada departamento construido en el proyecto inmobiliario, en forma de inversión en áreas verdes y/o infraestructura (i.e. iluminación, arborización, mobiliario, entre otros) por parte del Estado en torno a la estación.

La Tabla 1 muestra los niveles de los incentivos utilizados en el experimento. Los incentivos directos a la demanda se basaron en el subsidio de renovación urbana (utilizado ampliamente en la ciudad de Santiago) con un límite de 200 UF. Especialistas en urbanismo, en base a los criterios del MINVU, determinaron los niveles de los incentivos directos e indirectos a la oferta. Todos los niveles fueron validados por medio de grupos focales y experiencias piloto. Como la

⁴ Un detalle de los incentivos se muestra en el apéndice A.

integración social era considerada un desincentivo al emprendimiento inmobiliario, las alternativas que la incluían tenían niveles iguales o superiores a aquellas que no incluían integración social.

Tabla 1: Niveles de los incentivos

	Incentivo	Niveles sin integración social	Niveles con integración social
<i>Directos a la Demanda</i>	Bono a cualquier comprador	50 - 100 - 150 UF/dpto	50 - 100 - 150 - 200 UF/dpto
	Bono a cualquier comprador con subsidio	50 - 100 - 150 UF/dpto	50 - 100 - 150 - 200 UF/dpto
	Bono por tiempo limitado	40 - 80 - 120 UF/dpto	40 - 80 - 120 - 160 UF/dpto
<i>Directo a la Oferta</i>	Aumento en constructibilidad	0 - 10 - 20 - 30%	10 - 20 - 30 - 40%
	Aumento en densidad	0 - 50%	0 - 100%
	Disminución del mínimo de estacionamientos	0 - 50%	0 - 100%
<i>Indirectos a la Oferta</i>	Inversión pública en espacio público	0 - 25 - 50 - 75 UF/dpto	25 - 50 - 75 - 100 UF/dpto

Como fue mencionado anteriormente, el experimento de elección utiliza dos técnicas de elección discreta; escalamiento *best-worst* y preferencias declaradas tradicional (ver figura 1).

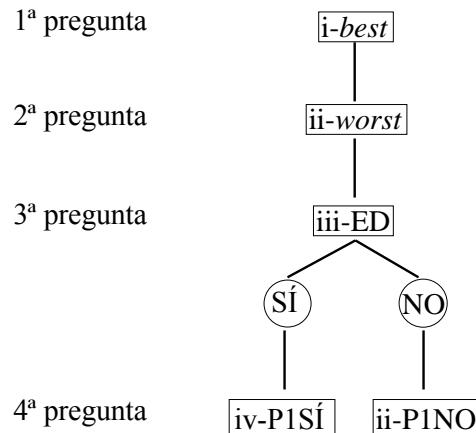


Figura 1: Diagrama del proceso de elección

A cada persona encuestada se le presentaba una alternativa de construcción con un valor particular para cada uno de los atributos descritos anteriormente. Así, para una estación en específico el encuestado debía realizar cuatro elecciones:

- (i) el atributo considerado más atractivo (*best*) en el paquete de incentivos presentado,
- (ii) el atributo menos atractivo (*worst*),
- (iii) la disposición del desarrollador a construir utilizando los incentivos y niveles mostrados en la alternativa, y
- (iv) su disposición a construir con integración social.

Las preguntas (i) y (ii) corresponden a la elección de un incentivo –y su nivel- más y menos atractivo, respectivamente. La pregunta (iii) corresponde a la elección de construir o no construir con un perfil de incentivos y niveles señalados. Dependiendo de la respuesta a la pregunta (iii), el desarrollador inmobiliario debe contestar una –de dos- preguntas; de responder “SÍ”, debe elegir entre construir sin integración social con los incentivos y niveles señalados en la pregunta (iii), o construir con integración social e incentivos, en cierta medida, mayores a la opción sin integración social (iv-P1SI). En caso de responder “NO”, el inmobiliario debe elegir si construiría

con integración social, con un aumento del nivel de los incentivos, o no construir en ningún caso (iv-P1NO)⁵. Este ejercicio se repitió 12 y 15 veces con cada desarrollador inmobiliario, para las estaciones de Metro y corredores de buses, respectivamente.

Así, las respuestas (iv-P1SI) y (iv-P1NO) poseen un sesgo de endogeneidad debido a que ambas son preguntas que dependen de una decisión anterior; la respuesta en la pregunta (iii). El tratamiento de este problema permitirá entender de forma adecuada la disposición de los desarrolladores inmobiliarios a integrar socialmente en torno a las estaciones de Metro y corredores de buses.

Finalmente, el banco de datos cuenta con 2.428 y 1.396 observaciones para las estaciones de Metro y corredores de buses, respectivamente; esto es, un total de 3.824 observaciones. También se cuenta con información sobre las características de cada estación seleccionada, como características demográficas y socioeconómicas, normativas, historial de construcciones en altura, entre otras.

4. MODELACIÓN **BEST-WORST-ELECCIÓN DISCRETA Y ENDOGENEIDAD**

Los datos fueron modelados en el software PythonBiogeme, Bierliere (2003). En primera instancia, se modelaron los datos *best* y *worst* de manera independiente para luego, se estimar un modelo conjunto que considerara comunes las sensibilidades e interceptos para ambas bases, pero con signo contrario.

Luego, se estimó un modelo que lograra reunir las respuestas de tipo *best*, *worst* y elección discreta tradicional (BWED). Los parámetros estimados de esta forma, así como los modelos independientes se muestran en la Tabla 2.

Tabla 2: Estimación de los modelos BEST, WORST, ED y BWED

Sensibilidad	Parámetros	BEST		WORST		ED		BWED	
		Valor	test-t	Valor	test-t	Valor	test-t	Valor	test-t
	Constante P1SI					0,605	3,9	0,676	5,44
Sensibilidad	Bono a cualquier comprador	0,453	7,33	0,424	3,86	0,140	2,64	0,098	4,79
	Bono a cualquier comprador con subsidio	0,436	7,11	-0,056	-0,6	0,058	1,05	0,081	4,58
	Bono por tiempo limitado	0,459	8,89	0,140	2,45	0,134	3,06	0,094	5,1
	Aumento en constructibilidad	1,26	9,54	0,137	1,63	0,166	2,4	0,218	5,82
	Aumento en densidad			0,732	3,52	0,379	2,65	0,287	3,62
	Disminución del mínimo de estacionamientos			-0,173	-1,14	-0,096	-0,68	0,080	1,14
	Inversión pública en espacio público	0,938	4,4	-0,074	-0,98	0,016	0,25	0,086	2,86
Intercepto	Bono a cualquier comprador	0,892	6,3	0,811	2,73			0,185	4,31
	Bono a cualquier comprador con subsidio	0,669	4,9	0,200	0,93			0,130	3,71
	Bono por tiempo limitado (fijo)								

⁵ En el Apéndice A se puede ver un ejemplo de una situación de elección.

Aumento en constructibilidad	-0,807	-5,47	-0,269	-1,65		-0,148	-3,99	
Aumento en densidad	-0,681	-5,8	0,197	0,95		-0,128	-4,04	
Disminución del mínimo de estacionamientos	-	-	-1,640	11,19	-0,772	-4,93	-0,354	-5,03
Inversión pública en espacio público	-	-	-2,190	10,29	-0,715	-4,84	-0,434	-5,1
Factor de escala λ_{BEST} (distinto de 1)						5,050	5,48	
λ_{WORST} (distinto de 1)						4,39		
# Parámetros	11		13		8		16	

Para poder considerar comunes los parámetros de sensibilidad e interceptos, se estimaron los datos considerando un factor de escala entre los distintos tipos de modelo. Como no es posible estimar un factor de escala para cada modelo, se consideró el factor de escala del modelo de elección discreta tradicional (λ_{ED}) con valor igual a uno. El factor de escala de las respuestas i-best y ii-worst resultó significativo al 95% de confianza, mostrando la importancia de incluirlo. Por otro lado, el test de razón verosimilitud resulta $\lambda_a = 17,84 < \chi^2_{0,95;16} = 26,3$, por lo que no existe evidencia estadística para rechazar la hipótesis nula. Es decir, los interceptos y sensibilidades se pueden considerar comunes para todas las respuestas.

Una vez obtenido este modelo, se modelaron los datos anteriores junto con las respuestas iv-P1SÍ y iv-P1NO, considerando el sesgo debido a la endogeneidad. Como se mencionó anteriormente, el problema de endogeneidad de la base de datos aparece como un error asociado a la primera alternativa (iii-ED) que, de alguna forma, está presente en la alternativa posterior (iv-P1SÍ o iv-P1NO).

Los incentivos relacionados a las funciones de utilidad de las alternativas iv-P1SÍ y iv-P1NO deben ser los mismos (aunque pueden tener distintos niveles) que en la función de utilidad presentada en iii-ED. Sin embargo, podría rechazarse la similitud estadísticas de sus parámetros estimados. Así, las funciones de utilidad tienen la siguiente forma:

Elección discreta (ED)

$$\begin{aligned} U_{ED_{P1SI}} &= \lambda_{ED} \cdot (\beta_{ED_{SI}} + \beta_{ED_1} \cdot X_1 + \dots + \beta_{ED_m} \cdot X_M + \varepsilon_{ED}) \\ U_{ED_{P1NO}} &= \lambda_{ED} \cdot (\varepsilon_{ED}) \end{aligned} \quad (14)$$

P1SI

$$\begin{aligned} U_{P1SI_{SINIS}} &= \lambda_{P1SI} \cdot (\beta_{P1SI_1} \cdot X_1 + \dots + \beta_{P1SI_m} \cdot X_M + \delta_{P1SI} \cdot \varepsilon_{ED| \varepsilon_{ED_{P1SI}} > \varepsilon_{ED_{P1NO}} + v_{P1NO} - v_{P1SI}}) \\ &\quad + \eta_{P1SI} \\ U_{P1SI_{CONIS}} &= \lambda_{P1SI} \cdot (\beta_{P1SI_{CONIS}} + \beta_{P1SI_{IS_1}} \cdot \widetilde{X}_1 + \dots + \beta_{P1SI_{IS_M}} \cdot \widetilde{X}_M) + \eta_{P1SI} \end{aligned} \quad (15)$$

P1NO

$$\begin{aligned} U_{P1NO_{CONIS}} &= \lambda_{P1NO} \cdot (\beta_{P1NO_{SICONIS}} + \beta_{P1NO_{IS_1}} \cdot \widetilde{X}_1 + \dots + \beta_{P1NO_{IS_M}} \cdot \widetilde{X}_M) + \eta_{P1NO} \\ U_{P1NO_{NO}} &= \lambda_{P1NO} \cdot (\delta_{P1NO} \cdot \varepsilon_{ED| \varepsilon_{ED_{P1NO}} > \varepsilon_{ED_{P1SI}} + v_{P1SI} - v_{P1NO}}) + \eta_{P1NO} \end{aligned} \quad (16)$$

donde λ_i es el factor de escala para las respuestas i (ED, P1SÍ y P1NO), M son los incentivos estudiados, ε_{ED} es el error asociado a las alternativas de ED, δ_{P1SI} y δ_{P1NO} son factores que determinan la inclusión del error asociado a la alternativa anterior si son significativamente distintos de cero (Train y Wilson, 2008) y η_j es el error asociado a las alternativas j (P1SÍ, P1NO). Y, al igual que para el modelo BWED, se busca que las sensibilidades de los parámetros sean iguales; es decir, se espera que los β_i de todas las alternativas sean iguales, es decir:

$$\beta_{ED_m} = \beta_{P1SI_m} = \beta_{P1SI_{IS_m}} = \beta_{P1NO_{IS_m}} \quad (17)$$

para todo incentivo m .

Considerando lo anterior, el modelo estimado que considera endogeneidad (ENDO) se puede ver en la Tabla 3.

Tabla 3: Estimación del modelo ENDO

ENDO			
	Parámetros	Valor	test-t
	Constante P1SI	0,703	5,54
	Constante P1SI_CONIS	0,030	0,04
	Constante P1NO_SICONIS	0,313	0,62
Sensibilidad	Bono a cualquier comprador	0,094	4,82
	Bono a cualquier comprador con subsidio	0,087	4,73
	Bono por tiempo limitado	0,088	5,07
	Aumento en constructibilidad	0,204	5,78
	Aumento en densidad	0,342	4,46
Intercesto	Disminución del mínimo de estacionamientos	0,100	1,77
	Inversión pública en espacio público	0,090	3,25
	Bono a cualquier comprador	0,181	4,31
	Bono a cualquier comprador con subsidio	0,138	3,86
	Bono por tiempo limitado		
Factor de escala	Aumento en constructibilidad	-0,143	-3,95
	Aumento en densidad	-0,120	-3,99
	Disminución del mínimo de estacionamientos	-0,341	-5,04
	Inversión pública en espacio público	-0,421	-5,06
	λ_{BEST}	5,200	5,45
Errores	(distinto de 1)		4,40
	λ_{WORST}	1,420	4,98
	(distinto de 1)		1,47
	λ_{P1SI}	0,976	1,55
	(distinto de 1)		-0,04
Endogeneidad	λ_{P1NO}	1,850	1,11
	(distinto de 1)		0,51
	δ_{P1SI}	0,735	1,04
	δ_{P1NO}	0,847	1,7

Log-verosimilitud (0)	-6948,392
Log-verosimilitud	-4106,544
ρ^2	0,409
$\bar{\rho}^2$	0,406
# Parámetros	22

En primer lugar, los parámetros asociados a la inclusión de endogeneidad, δ_{P1SI} y δ_{P1NO} , son significativos al 70% y un 90% de confianza, respectivamente.

Por medio del test razón verosimilitud, se puede comparar este modelo con una versión restringida, consistente en no considerar el error asociado a la endogeneidad. En este caso, $\lambda_a = 4,28 > \chi^2_{0,88;2} = 4,24$, por lo que se rechaza la hipótesis nula, y se concluye que el modelo que considera endogeneidad es superior.

4.1. Probabilidad de desarrollo inmobiliario con integración social

Utilizando los resultados anteriores junto con la ecuación (9), se puede estimar la probabilidad de que un desarrollador inmobiliario utilice alguno de los incentivos para la construcción de un proyecto que incluya integración social. Estas probabilidades se muestran en la Tabla 4.

Tabla 4: Probabilidades de densificar con integración social

	Sin incentivos	2%
con un incentivo		
Bono a cualquier comprador	50 UF/dpto 100 UF/dpto 150 UF/dpto 200 UF/dpto	2% 3% 3% 4%
Bono a cualquier comprador con subsidio	50 UF/dpto 100 UF/dpto 150 UF/dpto 200 UF/dpto	2% 3% 3% 4%
Bono por tiempo limitado	40 UF/dpto 80 UF/dpto 120 UF/dpto 160 UF/dpto	2% 3% 3% 4%
Aumento en constructibilidad	10% 20% 30% 40%	2% 3% 4% 5%
Aumento en densidad	50% 100%	2% 4%
Disminución del mínimo de estacionamientos	50% 100%	2% 2%
Inversión pública en espacio público	50 UF/dpto 75 UF/dpto 100 UF/dpto	2% 2% 3%

Para estimar las probabilidades anteriores se debió estimar el error asociado a la endogeneidad para las probabilidades iv-P1SI y iv-P1NO. Este error tiene una media distinta de cero y está asociado a la elección iii-ED. Por lo tanto, se simuló el error con 1000 iteraciones, promediando las probabilidades obtenidas de la misma forma en que son estimadas (Train y Wilson, 2008).

5. CONCLUSIONES

Esta investigación es pionera en la utilización de encuestas de preferencias y modelos de elección discreta para comprender la disposición de desarrolladores inmobiliarios a construir viviendas con inclusión social. En ella se utilizan métodos innovadores de elección *best-worst*, para intentar comprender como los atributos incentivan o inhiben tales emprendimientos, y también trata el problema de endogeneidad inherente a la encuesta.

Los resultados obtenidos muestran que el modelo conjunto entre las respuestas *best*, *worst* y elección discreta tradicional posee interceptos y las sensibilidades inversas entre los modelos *best* y *worst*, como sugiere la teoría. Por otro lado, incluir un factor de escala para relacionar ambos modelos resultó significativo, demostrando la existencia de diferencias entre los errores asociados al tipo de respuesta.

El modelo conjunto para las respuestas *best*, *worst*, ED, D-P1SI y D-P1NO, considerando el error asociado a la endogeneidad de las respuestas relacionadas, resultó ser estadísticamente relevante; la inclusión del parámetro de la endogeneidad fue significativo al 70% y 90% para las respuestas iv-P1SI y iv-P1NO, respectivamente. Además, los modelos permitieron estimar la probabilidad con la que un desarrollador inmobiliario estaría dispuesto construir con integración social.

Finalmente, sería interesante estimar modelos que permitieran entender la disposición de los desarrolladores inmobiliarios a densificar con integración social en torno a cada estación particular. Es decir, estimar modelos que considerando variaciones sistemáticas de gusto, solucionaran el problema de endogeneidad asociado a la encuesta y consideraran simultáneamente el efecto panel. A partir de ellos se podrían obtener modelos superiores, que hicieran posible diseñar paquetes de incentivos para su utilización en materias de política pública.

AGRADECIMIENTOS

Los autores agradecemos a Angelo Guevara, David Hensher, John Rose por su valiosa ayuda en la concepción e interpretación de los modelos y a DICTUC S.A. por compartir la base de datos para el desarrollo de este proyecto. También dar las gracias al Fondo para el Desarrollo Científico y Tecnológico (FONDECYT proyecto 1141096), al Instituto de Sistemas Complejos de Ingeniería y Sistemas (ICM: P-05-004-F; CONICYT: FB016) y CEDEUS, por haber financiado parcialmente este trabajo.

REFERENCIAS

- Balbontin, C., Ortúzar, J.D. y Swait, J.D. (2014) Importance of Dwelling, Neighbourhood Attributes in Residential Location Modelling: Best Worst Scaling vs. Discrete Choice. **XI Congreso de Ingenieria del Transporte.** 19 de Diciembre de 2014, Santander, 92 – 101.

Bierlaire, M. (2003). BIOGEME: a free package for the estimation of discrete choice models. **3rd Swiss Transportation Research Conference**. Marzo 2003, Ascona, Switzerland.

DICTUC (2014) **Estudio de Densificación En Torno a la Red de Metro de Santiago**. Informe Final. Ministerio de Vivienda y Urbanismo. Informe N°:1191398.

Flynn, T.N., Louviere, J.J., Peters, T.J. y Coast, J. (2007). Best-worst scaling: What it can do for health care research and how to do it. **Journal of Health Economics**, 26, 171 – 189.

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.

Lagerkvist, C.J. (2013). Consumer preferences for food labelling attributes: Comparing direct ranking and best-worst scaling for measurement of attribute importance, preference intensity and attribute dominance. **Food Quality and Preference**, 29, 77–88.

Lancsar, E., Louviere, J.J., Donaldson, C., Currie, G. y Burgess, L. (2013). Best worst discrete choice experiments in health: Methods and an application. **Social Science & Medicine**, 76(1), 74–82.

Louviere, J.J., Lings, I., Islam, T., Gudergan, S. y Flynn, T. (2013). An introduction to the application of (case 1) best-worst scaling in marketing research. **International Journal of Research in Marketing**, 30, 292–303.

Louviere, J.J., Train, K., Ben-Akiva, M., Bhat, C., Brownstone, D., Cameron, T.A., Carson, R.T., Deshazo, J.R., Fiebig, D., Greene, W., Hensher, D. y Waldman, D. (2005). Recent Progress on Endogeneity in Choice Modeling. **Marketing Letters**, 16(3-4), 255–265.

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.

Louviere, J.J., Street D., Burgess, L., Wasi1, N., Islam, T. y Marley, A.A. J. (2008). Modeling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information. **Journal of Choice Modelling**, 1(1), 128–164.

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

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

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

Ortúzar, J. de. D. y Willumsen, L.G. (2011). **Modelling Transport**. Cuarta Edición. Chichester: John Wiley and Sons.

Train, K. y Wilson, W.W. (2008) Estimation on stated-preference experiments constructed from revealed-preference choices. **Transportation Research B**, 42, 191 – 203.

Train, K. (2009). **Discrete choice methods with simulation**. 2nd ed. Cambridge: Cambridge University Press.

Waintrub, N. (2015) **Disposición a Recibir Incentivos para Desarrollo Inmobiliario en Densidad en Torno a Estaciones de Metro y Corredores de Buses**. Tesis de Magíster en Ciencias de la Ingeniería, Departamento de Ingeniería en Transporte y Logística, Pontificia Universidad Católica de Chile.

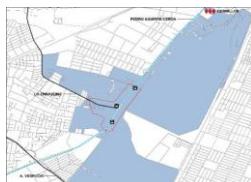
Walle, V.S., Steenberghen, T., Paulley, N., Pedler, A. y Martens, M. (2004). The role of indicators in the assessment of integrated land-use and transport policies in European cities. **International Planning Studies** 9, 173–196.

Whitty, J., Ratcliffe, J., Chen, G. y Scuffham, P. (2013). A comparison of Discrete Choice and Best Worst Scaling methods to assess Australian public preferences for the funding of new health technologies. **International Choice Modelling Conference**. Sydney Australia.

APÉNDICE A

Cada encuesta cuenta con 12 o 15 situaciones de elección. Un ejemplo de cada situación de elección sobre una estación se muestra continuación. La Figura 2 muestra la elección del incentivo mejor (*i-best*) y peor (*ii-worst*); además, el inmobiliario debe decidir si la situación señalada, con todos los incentivos, es atractiva para construir (*iii-ED*).

PEDRO AGUIRRE CERDA



(1) Indique cuál de los incentivos ofrecidos le parece el más atractivo y cuál el menos atractivo.
 (2) Indique si considera que este conjunto de incentivos genera un escenario atractivo para construir en este entorno.

INCENTIVO MAS ATRACTIVO		INCENTIVO MENOS ATRACTIVO
<input type="checkbox"/>	Bono por cercanía al BRT a cualquiera que cargue con subsidio (adicional al subsidio ya obtenido)	150 UF/beneficiario
<input type="checkbox"/>	Bono cercanía al BRT por tiempo limitado	80 UF/depto
<input type="checkbox"/>	Aumento en constructibilidad	0%
<input type="checkbox"/>	Aumento en densidad	0%
<input type="checkbox"/>	Reducción del número mínimo de estacionamientos	0%
<input type="checkbox"/>	Inversión pública	0 UF/depto

Atractivo para construir

Si No

Figura 2: Ejemplo de la situación de elección *i-best*, *ii-worst* y *iii-ED*

Si el desarrollado inmobiliario responde que el paquete de incentivos “SÍ” es atractivo para construir, se pasa a la pregunta iv-P1SÍ de la Figura 3.

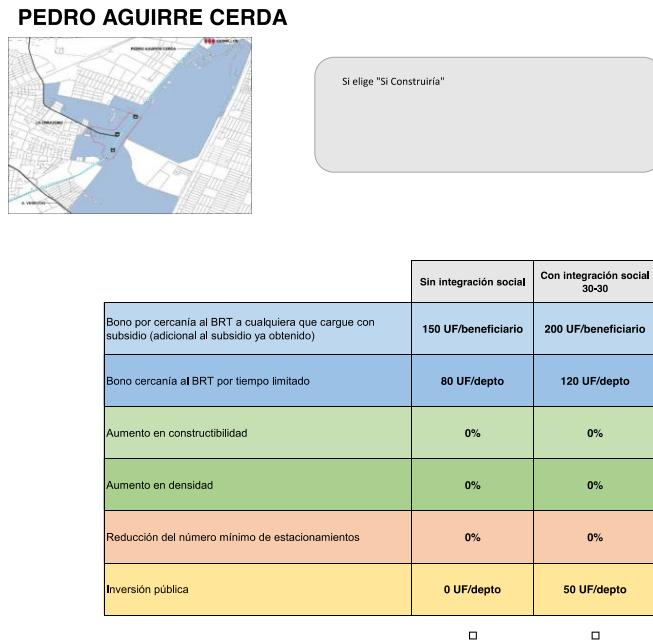


Figura 3: Ejemplo de la situación de elección P1SÍ

Si el desarrollador inmobiliario contesta que “NO” es atractivo construir con el paquete de incentivos mostrado, entonces se pasa a la pregunta iv-P1NO.

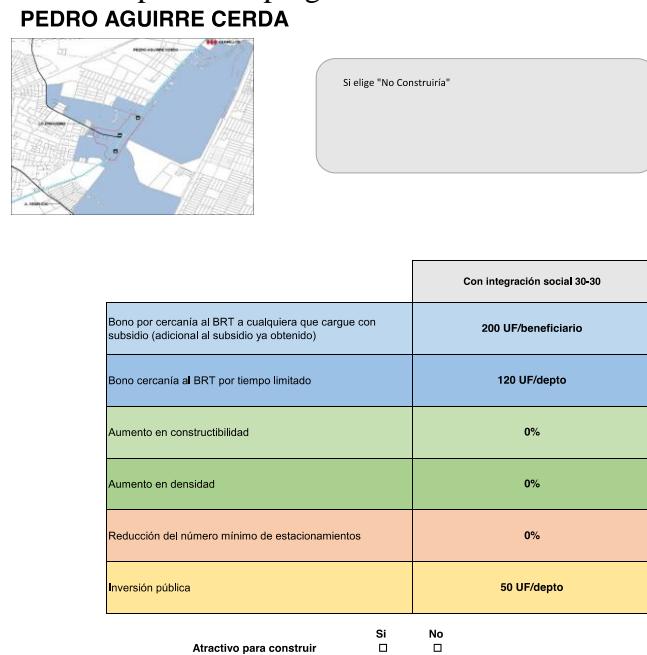


Figura 4: Ejemplo se de la situación de elección P1NO