



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

MODELO LOGIT BINOMIAL CON COTA SUPERIOR EN LA FUNCIÓN DE
VALOR LATENTE

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL

IGNACIO ANDRÉS EGUILLOR RODRÍGUEZ

PROFESOR GUÍA
PEDRO PABLO DONOSO SIERRA

MIEMBROS DE LA COMISIÓN:
FRANCISCO MARTÍNEZ CONCHA
MARCELA MUNIZAGA MUÑOZ

Este trabajo ha sido parcialmente financiado por FONDEF

SANTIAGO DE CHILE

2014

RESUMEN DE LA MEMORIA PARA
OPTAR AL TÍTULO DE: Ingeniero Civil
POR: Ignacio Andrés Eguillor Rodríguez
FECHA: 14-07-2014
PROFESOR GUIA: Pedro Pablo Donoso
Sierra

MODELO LOGIT BINOMIAL CON COTA SUPERIOR EN LA FUNCIÓN DE VALOR LATENTE

Los modelos de elección discreta basados en utilidades o disposiciones a pagar han sido muy utilizados para representar el comportamiento de los consumidores en sistemas de Uso de Suelo y de Transporte. En este trabajo se desarrolla un modelo de elección que usa disposiciones a pagar para representar la elección en mercados de remate, donde los bienes se asignan a los mejores postores. Los modelos tradicionales no consideran el hecho de que estas funciones están siempre acotadas superiormente porque los consumidores tienen restricciones presupuestarias. Si bien se han generado enfoques para representar esta situación, estas son heurísticas solamente. En este trabajo se deduce de un modo riguroso un nuevo modelo de elección discreta binomial donde las funciones de valor son disposiciones a pagar acotadas superiormente. En particular, se obtiene una expresión analítica en el caso en que estas funciones son variables aleatorias idénticamente distribuidas Gumbel. Este modelo junto con incorporar en su formulación la influencia de la restricción presupuestaria, tiene la interesante propiedad de permitir estimar un número mayor de parámetros que en el caso del modelo Logit Binomial clásico. Junto con caracterizar este modelo, se probó empíricamente usando bases sintéticas que es posible estimar sus parámetros y que en muchos casos, entrega estimaciones más precisas que el modelo Logit Binomial tradicional.

Tabla de contenido

1	INTRODUCCIÓN	1
2	ANTECEDENTES TEÓRICOS Y REVISIÓN BIBLIOGRÁFICA	3
2.1	Paradigma de la maximización de la utilidad	3
2.2	Deducción de la disposición a pagar y de la existencia de la cota superior.....	3
2.3	Usos de la disposición a pagar. Formas en que se ha considerado la restricción presupuestaria.	7
2.4	Modelo probabilístico de elección discreta Logit Multinomial basado en disposición a pagar	8
2.5	Estimación por máxima verosimilitud.....	9
2.6	<i>Constrained Multinomial Logit</i>	11
2.7	Síntesis del <i>Constrained Multinomial Logit</i>	15
3	MODELO PROPUESTO	17
3.1	Formulación matemática de las probabilidades Logit binomiales con cota superior	18
3.2	Condición en que el modelo restringido recupera el irrestricto.	25
3.3	Las cotas pueden inducir un cambio en el orden de las probabilidades.	25
3.4	Caracterización del modelo restringido.....	26
3.4.1	Sensibilidad del modelo de elección respecto al valor de las cotas superiores.	26
3.4.2	Sensibilidad del modelo respecto a las posturas.	29
3.4.3	Sensibilidad del modelo respecto al parámetro μ	31
3.5	Identificabilidad de parámetros.	34
3.5.1	Identificabilidad en el modelo logit irrestricto.....	34
3.5.2	Identificabilidad en el modelo logit restringido.	36
4	SIMULACIÓN.....	40
4.1	El simulador.....	40
4.1.1	Bases de datos sintéticas.....	40
4.1.2	Estimación.	43
4.2	Escenarios estimados.	43
4.2.1	Escenario 1.	43
4.2.2	Escenario 2.....	46

4.2.3 Escenario 3.....	48
5 Conclusiones.....	49
Bibliografía	51
Anexos	53
Anexo A: Proposición 1	53
Anexo B: Proposición 2.....	53

ÍNDICE DE TABLAS

Tabla 1: Parámetros identificables con el modelo restringido.....	39
Tabla 2: Parámetros del escenario 1.	44
Tabla 3: Resultados de estimaciones bajo el escenario 1 variando el tamaño de la muestra.....	45
Tabla 4: Parámetros del escenario 2.....	47
Tabla 5: Resultados de estimaciones bajo el escenario 2 variando el parámetro μ	47
Tabla 6: Parámetros del escenario 3.....	48
Tabla 7: Resultados de la estimación exploratoria bajo el escenario 3.....	48

ÍNDICE DE FIGURAS

Figura 1: Frecuencia de posturas $\mu=1$	18
Figura 2: Probabilidades en función de ΔI , con $I_2 = 1$, $\mu=1$ y $B_1=B_2=1$	27
Figura 3: Probabilidades en función de ΔI , con $I_2 = 1$, $\mu=1$ y $B_1=B_2=0,5$	27
Figura 4: Probabilidades en función de ΔI , con $I_2 = 1$, $\mu=0$ y $B_1=1$ $B_2=0,5$	28
Figura 5: Probabilidades en función de ΔI , con $I_2 = 1$, $\mu=1$ y $B_1=0,5$ $B_2=1$	29
Figura 6: Probabilidades en función de ΔB , con $I_1=2,5$ $I_2= 1,5$ $\mu=1$ y $B_2=1$	30
Figura 7: Probabilidades en función de ΔB , con $I_1=2,5$ $I_2= 1,5$ $\mu=1$ y $B_1=1$	30
Figura 8: Probabilidades en función de μ , con $I_1=2,5$ $I_2= 1,5$ y $B_1=B_2=1$	32
Figura 9: Probabilidades en función de μ , con $I_1=2,5$ $I_2= 1,5$, $B_1=1$ y $B_2=0,5$	33
Figura 10: Probabilidades en función de μ , con $I_1=2,5$ $I_2= 1,5$, $B_1=0,5$ y $B_2=1$	33
Figura 11: Distribución de frecuencias de posturas simuladas con $\mu=1$	42
Figura 12: Numero de iteraciones versus número de observaciones.	46
Figura 13: Tiempo de convergencia versus número de observaciones.	46

1 INTRODUCCIÓN

Los modelos de elección discreta son herramientas muy útiles para caracterizar y predecir el comportamiento del consumidor, en particular, su demanda por bienes y servicios discretos.

Estas herramientas han sido utilizadas en distintas áreas, y en particular en Sistemas Urbanos, como son los sistemas de Uso de Suelo y de Transporte. En los sistemas de Uso de Suelo, se han empleado para estimar la producción espacial de bienes inmuebles y la localización de agentes residenciales (hogares) y no residenciales (firmas) en una ciudad (Martínez, 1996). En los sistemas de transporte se han usado para modelar todas las etapas de decisión de los usuarios, elección de viajar o no (generación de viajes), del destino del viaje (distribución de viajes), del horario del viaje, del modo de transporte (partición modal) y de ruta del viaje (asignación). Estos modelos se han implementado para caracterizar el comportamiento de los usuarios, para proyectar la demanda ante distintos escenarios de políticas y proyectos y para estimar externalidades relevantes en la vida urbana, como son la congestión, contaminación y segregación social, entre otros.

En la especificación funcional de estos modelos se asume que los consumidores tienen un comportamiento “racional”, en el sentido que buscan maximizar su nivel de utilidad dentro de sus posibilidades de consumo, determinadas principalmente por su disponibilidad presupuestaria. Todos estos modelos utilizan una función latente de valor percibido por los consumidores, ya sea la utilidad o la disposición a pagar o una medida dependiente de esta última, como es el excedente del consumidor.

Como se aprecia, la utilidad, como también la disposición a pagar y las restricciones que actúan sobre éstas, son elementos fundamentales para caracterizar el comportamiento de los consumidores. Por este motivo, en este trabajo se modela la influencia de la restricción presupuestaria en la disposición a pagar y su efecto en las elecciones que realizan los consumidores, aspecto menos tratado en la literatura.

Habitualmente, en la especificación de modelos de disposición a pagar y de la elección que realiza un consumidor, se ha prescindido de un elemento central ligado a la propia naturaleza de la disposición a pagar: que el espacio donde esta función define una elección está acotado por la restricción presupuestaria del consumidor. Al obviar esta restricción se puede introducir un error importante en la estimación de las disposiciones a pagar, al explicar o estimar elecciones cuando las posturas usadas para esto son superiores al presupuesto disponible, afectando por tanto, la fiabilidad del modelo de elección. Debe notarse que la estimación de funciones de utilidad no presenta este problema, porque las utilidades no están acotadas superiormente. Por el contrario, la disposición a pagar, al estar expresada en unidades monetarias (y no “útiles” como la función de utilidad) tiene una restricción, cual es la de no puede exceder la restricción presupuestaria del agente, siendo éste un punto relevante que diferencia a ambas funciones.

En la literatura especializada se han formulado diferentes enfoques para representar las restricciones de consumo que enfrentan los agentes (Swait, 2001; Cascetta y Papola, 2001). En particular, es interesante mencionar el modelo *Constrained Multinomial Logit* originalmente propuesto por Martínez *et al* (2009), bajo el cual las restricciones se

representan como funciones de *cutoff* de la función de utilidad indirecta, modificando así el modelo clásico Logit Multinomial. Este enfoque se ha aplicado al modelo de Uso de Suelo de Santiago: MUSSA (Martínez, 1996), penalizando las funciones de disposición a pagar de los agentes en la medida que se aproximen y superen sus restricciones de ingreso. Este enfoque corresponde a una heurística para representar estas restricciones y no una deducción analítica de la imposición de éstas.

En este trabajo, se deduce y analiza un modelo de elección discreta con dos alternativas (binomial) cuando la función de valor de cada una de éstas está acotada superiormente. Al igual que en el caso del modelo Logit Multinomial se obtiene una expresión explícita de la probabilidad de elección en el caso de que los errores aleatorios de las funciones de valor son independientes y distribuidas idénticamente (i.i.d.) Gumbel, obteniéndose el llamado “Logit binomial con cota superior”. Este nuevo modelo tiene interesantes propiedades que se analizan en este trabajo, algunas de las cuales lo asemeja al modelo Logit Binomial clásico, pero otras lo diferencian notablemente.

Además, se ha diseñado una base de datos sintéticos con atributos y elecciones simuladas (asociadas a errores i.i.d. Gumbel) y se han realizado estimaciones de los parámetros del modelo con el método de la máxima verosimilitud, comprobándose la factibilidad de estimar los parámetros de este nuevo modelo.

2 ANTECEDENTES TEÓRICOS Y REVISIÓN BIBLIOGRÁFICA

En este trabajo se aborda el problema de estimar la disposición a pagar de un consumidor considerando explícitamente que está acotada por su disponibilidad presupuestaria. En esta sección se presenta la definición de la disposición a pagar, bajo el paradigma de que el consumidor es racional (McFadden 1980).

2.1 Paradigma de la maximización de la utilidad

El paradigma del consumidor racional se basa en los siguientes supuestos:

1. Cada individuo enfrenta un conjunto de alternativas disponibles de consumo;
2. Cada alternativa le provee al individuo un nivel de utilidad;
3. El individuo elige aquella alternativa que le provee la máxima utilidad dentro de las alternativas disponibles;

Es importante observar que se considera que la utilidad que un bien o servicio le entrega al usuario proviene de sus atributos y no del bien en sí mismo.

2.2 Dedución de la disposición a pagar y de la existencia de la cota superior

En el marco de la teoría de elecciones discretas, el consumidor se enfrenta al problema de elegir cuanto consumir de un conjunto de bienes continuos y cuál elegir de entre un conjunto de bienes discretos, por ejemplo, una casa donde vivir o un modo de transporte para realizar un viaje (Jara-Díaz, Martínez, 1999). Basándose en este enfoque se presenta una deducción para el contexto de este trabajo donde la demanda por los bienes continuos y discretos es considerada como el resultado del problema de maximización, que se presenta a continuación.

Sean:

x_j : Cantidad de unidades consumidas del bien continuo j .

δ_i : 1 si consumidor elige el bien discreto i ; 0 si no.

w_i : Vector de atributos del bien discreto i .

p_j : Precio de una unidad del bien continuo j .

r_i : Precio del bien discreto i .

m_{ij} : Cantidad mínima de unidades consumidas del bien continuo j cuando el consumidor elige el bien discreto i ($m_{ij} \geq 0$)

Luego el problema que resuelve el consumidor se puede escribir como:

$$\begin{aligned}
& \text{Máx}_{x,\delta} \sum_i \delta_i U_i(x, w_i) \\
\text{s. a.} \quad & \sum_j p_j x_j + \sum_i \delta_i r_i = I \\
& x_j \geq \sum_i \delta_i m_{ij} \quad \forall j \\
& \sum_i \delta_i = 1 \\
& x_j \geq 0 \quad \forall j, \quad \delta_i = 0,1 \quad \forall i
\end{aligned} \tag{1}$$

Separando en dos etapas, es equivalente a:

$$\text{Máx}_{\delta_i} \left(\begin{array}{l} \text{Máx}_x U_i(x, w_i) \\ \text{s. a.} \quad \sum_j p_j x_j = I - r_i \\ x_j \geq m_{ij} \quad \forall j \end{array} \right) \tag{2}$$

Sea:

x_{ij}^* : Cantidad óptima de unidades consumidas del bien continuo j , dado que se elige el bien discreto i .

Entonces, reemplazando x_{ij}^* en el problema (2) se convierte en encontrar el bien discreto i , que maximice la utilidad indirecta condicional (V_i).

$$\text{Máx}_i (V_i(x_{ij}^*(I - r_i, p, m_{i.}), w_i)) = V_i(I - r_i, p, m_{i.}, w_i) \tag{3}$$

Se hace notar que los valores óptimos x_{ij}^* cumplen con las siguientes condiciones:

$$\begin{aligned}
\sum_j p_j x_{ij}^* &= I - r_i \quad \forall i \\
x_{ij}^* &\geq m_{ij} \quad \forall j
\end{aligned} \tag{4}$$

De estas dos últimas condiciones (4) se deduce que conforme r_i se acerca a I , el consumo óptimo x_i^* se reduce y V_i disminuye. Notar que, si bien r_i es habitualmente positiva, podría tomar valores nulos o negativos, representando ingresos adicionales que permiten ampliar la capacidad de gastos en otros bienes. También de (4) se deduce que:

$$r_i \leq I - \sum_j p_j m_{ij} \quad (5)$$

Entonces, para que el problema sea factible para el consumidor, el precio del bien discreto i , r_i , no puede superar a su ingreso menos el consumo mínimo de los bienes continuos.

Supongamos que V_i invertible en su primera componente, y G_i es su inversa, luego:

$$I - r_i = G_i(V_i, p, m_i, w_i) \quad (6)$$

Así, $G_i(V_i, p, m_i, w_i)$ es el gasto óptimo en otros bienes.

Luego, se define la disposición a pagar por i (B_i), como el ingreso menos el gasto en otros bienes para alcanzar la utilidad V_i :

$$B_i = I - G_i(V_i, p, m_i, w_i) \quad (7)$$

Es interesante observar que si V_i crece, también lo hace el gasto en otros bienes (o al menos no disminuye) y por lo tanto, disminuye la disposición a pagar por i . Por lo tanto, B_i es no creciente con V_i .

La disposición a pagar por i está acotada superiormente por el ingreso porque el gasto óptimo en otros bienes dado el consumo de i , es no negativo, es decir:

$$\begin{aligned} B_i &\leq I \\ G_i(V_i, p, w_i) &\geq 0 \end{aligned} \quad (8)$$

Obviamente, al existir una cota mínima de consumo de los otros bienes, como se muestra en la segunda restricción del problema (1), entonces hay un nivel de consumo mínimo, denotado por x_i^{min} , entonces, la disposición a pagar por i se puede acotar superiormente de modo más estricto:

$$B_i \leq I - p^T x_i^{min} \quad (9)$$

Para ilustrar el enfoque anterior, se puede considerar que la utilidad indirecta condicional es lineal, $V_i = \alpha_0 + \alpha_1(I - r_i) + \alpha_2 p + \alpha_3 w_i$, donde $\alpha_1 > 0$ y corresponde a la utilidad marginal del ingreso o UMI, que representa la utilidad que reporta marginalmente una unidad más de ingreso. Por lo tanto, se observa que V_i es decreciente con r_i y creciente con $I - r_i$.

Luego, dado por (7) la disposición a pagar en este caso es:

$$B_i = \left(I - \frac{V_i}{\alpha_1} \right) + \left(\frac{\alpha_0}{\alpha_1} \right) + \left(\frac{\alpha_2}{\alpha_1} \right) p + \left(\frac{\alpha_3}{\alpha_1} \right) w_i \quad (10)$$

Notar que w_i se toma como escalar. Supongamos ahora que V_i no es sólo lineal, sino que también aleatoria, por la existencia de un término aditivo de error estocástico: $V_i = \alpha_0 + \alpha_1(I - r_i) + \alpha_2 p + \alpha_3 w_i + \varepsilon_i$. Más aun, si se supone que ε_i son i.i.d. *Gumbel*($0, \mu$) la probabilidad de elección de la opción i , es decir, que $V_i = \text{Máx}_j V_j(I - r_j, p, m_j, w_j)$ es un Logit Multinomial (ver 2.5.).

De la definición de V_i se deduce en este caso que:

$$B_i = \left(I - \frac{V_i}{\alpha_1} \right) + \left(\frac{\alpha_0}{\alpha_1} \right) + \left(\frac{\alpha_2}{\alpha_1} \right) p + \left(\frac{\alpha_3}{\alpha_1} \right) w_i + \left(\frac{1}{\alpha_1} \right) \varepsilon_i \quad (11)$$

Se sabe que la disposición a pagar cumple con la restricción de ingreso, $B_i \leq I$, pero se puede corroborar directamente:

Dado que $I - r_i \geq 0$ y $\alpha_1 \geq 0$ entonces,

$$V_i = \alpha_0 + \alpha_1(I - r_i) + \alpha_2 p + \alpha_3 w_i + \varepsilon_i \geq \alpha_0 + \alpha_2 p + \alpha_3 w_i + \varepsilon_i \quad (12)$$

Al dividir a ambos lados por α_1 se obtiene:

$$\frac{V_i}{\alpha_1} \geq \left(\frac{\alpha_0}{\alpha_1} \right) + \left(\frac{\alpha_2}{\alpha_1} \right) p + \left(\frac{\alpha_3}{\alpha_1} \right) w_i + \left(\frac{1}{\alpha_1} \right) \varepsilon_i \quad (13)$$

Al pasar el término de la izquierda al lado derecho y sumar el ingreso I a ambos lados de la desigualdad se obtiene:

$$I \geq I - \frac{V_i}{\alpha_1} + \left(\frac{\alpha_0}{\alpha_1} \right) + \left(\frac{\alpha_2}{\alpha_1} \right) p + \left(\frac{\alpha_3}{\alpha_1} \right) w_i + \left(\frac{1}{\alpha_1} \right) \varepsilon_i = B_i \quad (14)$$

De esta manera, se aprecia que la disposición a pagar está acotada por el ingreso.

2.3 Usos de la disposición a pagar. Formas en que se ha considerado la restricción presupuestaria.

La disposición a pagar se ha estudiado en diversos sistemas y aspectos de éstos, por ejemplo, en sistemas de transporte para estimar la disposición a pagar por mejoras en cualidades de los diversos modos (McFadden, 1998; Sillano y Ortuzar, 2005), en mercados inmobiliarios, para estimar la demanda por vivienda y sus características o atributos (Gross, 1988; Follain y Jiménez, 1985) y también en otros mercados ajenos al transporte y uso de suelo. En salud se utiliza para evaluar el impacto de programas de asistencia médica en el bienestar de la población (Birch *et al* 1999). En energía se ha utilizado para analizar la demanda por atributos ambientales de servicios eléctricos residenciales mediante una encuesta de preferencias declaradas (Roe *et al* 2001). En la industria alimenticia y también con una encuesta de preferencias declaradas, se mide la disposición a pagar por productos orgánicos (Gil *et al* 2000).

En el mercado inmobiliario la disposición a pagar se presenta de distintas maneras dependiendo del enfoque de modelación del mercado.

El enfoque *choice* (McFadden 1978, Anas 1982) utilizado para modelar la localización de hogares y firmas en una ciudad, bajo el supuesto que los agentes maximizan su utilidad por localización y son tomadores de precios. Estos precios se determinan de manera exógena a través de modelos hedónicos de regresión de precios (Rosen, 1974). Bajo este enfoque el consumidor elige el gasto en bienes continuos y discretos diferentes al de localización que maximiza su utilidad (Jara-Díaz, Martínez, 1999), obteniendo así una función de utilidad indirecta condicional en la elección del bien raíz. La localización elegida por el consumidor es aquella que maximiza esta utilidad indirecta dentro de un conjunto de alternativas. La disposición a pagar asociada al bien inmueble elegido (al nivel de utilidad óptimo) por construcción no viola la restricción presupuestaria. Rosen (1974) arma una función de precios hedónica, producto de la interacción de los distintos hogares tratando de localizarse, y las funciones de oferta de los oferentes de viviendas, es decir una función endógena. La función de precios hedónica $P(H)$, relaciona los gastos en vivienda y las características o atributos del bien raíz $H = f(h_1, h_2, \dots, h_k)$. En particular, se refiere al precio de un atributo k contenido en el vector H , como P_k ($\equiv \partial P(H)/\partial h_k$), la disposición a pagar marginal (indirecta) del atributo.

Alonso (1964) propone que el mercado inmobiliario puede entenderse como un mercado de remates, donde los agentes (hogares y firmas) pujan por un bien raíz en particular. Bajo este proceso, el agente que se localiza en un determinado bien raíz es el que tiene la máxima disposición a pagar y por tanto, este valor define también el precio del bien (*bid-rent*). Ellikson (1981) muestra que la función de postura de los agentes puede escribirse directamente como función de los atributos de la localización, y propone agregarle un término aleatorio que dé cuenta de la heterogeneidad no observada.

La localización se ha modelado también siguiendo un enfoque diferente denominado *bid auction*, donde se reconoce que los bienes en el mercado inmobiliario son únicos por la condición natural de que el suelo es un recurso escaso y no se puede producir. A esto se le suma no sólo que el espacio es limitado, sino además que los atributos del suelo (accesibilidad, calidad del vecindario, etc.) son resultado de externalidades no pecuniarias urbanas sin costo para los dueños (a excepción de impuestos) no modificables en el corto plazo. Estas características hacen a cada fracción de terreno altamente peculiar, cuasi-único (Martínez, 1992). Bajo este enfoque, la localización es el resultado de un proceso de remate, donde el mejor postor usa o consume cada bien inmueble del mercado y su postura define la renta o valor de uso del bien inmueble (Ellickson, 1981). A diferencia de Rosen, que estima directamente la renta, el modelo de Ellickson sólo permite estimar parámetros relativos, por lo que las rentas se definen sólo hasta fijar una constante exógena. Por otra parte, Lerman y Kern (1983) proponen un enfoque para estimar el modelo basado en posturas de remate, esto es estimar la probabilidad conjunta de que un hogar sea el mejor postor y que la postura de ese hogar sea igual al precio observado de venta.

Si bien, los modelos mencionados, en su deducción consideran que existe una restricción presupuestaria, nada garantiza que, a la hora de estimar las posturas que éstas no violen esta restricción. La función de disposición a pagar no se conoce, dado que la función de utilidad no es conocida. Aun así se puede usar una función no lineal que garantiza el cumplimiento de la restricción presupuestaria (Martínez et al, 2009).

2.4 Modelo probabilístico de elección discreta Logit Multinomial basado en disposición a pagar

En esta sección se presenta un modelo probabilístico de elección donde la función de valor es una variable aleatoria i.i.d *Gumbel*($0, \mu$). Como se ha mencionado antes, este modelo aplica a la probabilidad de localización de agentes bajo el enfoque (*bid auction*), donde éstos compiten por una localización realizando posturas por esta. Las posturas por un bien raíz se pueden escribir siguiendo (11), como una función de los atributos de la localización $B(w_i)$ (Hurtubia y Bierlaire, 2012).

Sean H es el conjunto de agentes (o categorías de agentes, si se los representa de manera agregada) y $B_h(w_i)$ la postura del agente (o tipo de agente) h por el bien raíz i . Sea $\bar{B}_h(w_i)$ la componente determinística de la postura y ε_h un término aleatorio que recoge una variabilidad o heterogeneidad no incluida en la componente determinística, entonces la postura $B_h(w_i)$ se puede expresar como:

$$B_{hi} = \bar{B}_h(w_i) + \varepsilon_h = \bar{B}_{hi} + \varepsilon_h \quad (15)$$

Como se aprecia, se asume que solamente la parte sistemática o determinística de la postura, \bar{B}_{hi} se ve afectada por los atributos de la localización.

La probabilidad de que una localización i sea ocupada por un agente h , es la probabilidad de que el agente sea el mejor postor de entre todos los demás agentes interesados en el bien raíz, por lo tanto, se expresa como:

$$P_{hi} = (\bar{B}_{hi} + \varepsilon_h > \bar{B}_{h'i} + \varepsilon_{h'}; \forall \neq h', i) \quad (16)$$

Si los términos del error son independientes e idénticamente distribuidos (i.i.d.) *Gumbel* $(0, \mu)$, la probabilidad de que h sea mejor postor puede ser expresada como un modelo Logit Multinomial (McFadden, 1978):

$$P_{hi} = \frac{\exp(\mu \bar{B}_{hi})}{\sum_{g \in H} \exp(\mu \bar{B}_{gi})}, \forall h \in H, i \quad (17)$$

2.5 Estimación por máxima verosimilitud

Por sus buenas propiedades, el método más ampliamente usado para estimar modelos probabilísticos es el de la Maximización de la Verosimilitud. A continuación se explica brevemente este método, en consideración a que es utilizado en esta investigación, para la estimación de funciones de disposición a pagar (Ellickson, 1981).

Los estimadores máximo-verosímiles (resultantes de la aplicación de este método) son aquellos que maximizan las probabilidades de elección observadas en una muestra, es decir, que el modelo probabilístico propuesto se ajuste a las probabilidades observadas (Ortúzar y Willumsen, 2011).

Para ilustrar esta idea, considérese una muestra de n observaciones de una determinada variable, que pueden ser las características de una localización, $w = \{w_1, \dots, w_n\}$, obtenida de una población caracterizada mediante un parámetro θ (media, varianza, etc.). Al ser w una variable aleatoria, tiene asociada una función de densidad $f(w/\theta)$ que depende de los valores de θ . Por tanto, siendo los valores de la muestra independientes, se puede escribir la función de densidad conjunta como:

$$f(w_1, w_2, \dots, w_n/\theta) = f(w_1/\theta)f(w_2/\theta) \cdots f(w_n/\theta) \quad (18)$$

La interpretación estadística más corriente de esta función es con w como variables y θ fijo. Invirtiendo el proceso, la ecuación anterior puede interpretarse como una función de verosimilitud $L(\theta)$. Maximizando la función de verosimilitud con respecto a θ , el resultado se denomina estimador de máxima verosimilitud porque corresponde al valor del parámetro que tiene la probabilidad mayor de haber generado la muestra observada.

Sea Q número de observaciones de elección disponibles de una muestra y A el conjunto de alternativas de elección disponibles en cada observación. Se supone que para cada uno de las observaciones de la muestra se observa:

$$\delta_{ij} = \begin{cases} 1 & \text{si } i \text{ es elegido en la observación } j \\ 0 & \text{si } i \text{ no es elegido en la observación } j \end{cases} \quad (19)$$

Sea ahora P_{ij} la probabilidad de que la opción i sea elegida en la observación j , determinada por el modelo usado por el investigador. Si se asume que las distintas observaciones de la muestra son independientes, la probabilidad conjunta de la muestra, denominada función de verosimilitud, es igual al producto de las probabilidades de que cada individuo elija la opción que realmente seleccionó. Una expresión general para la verosimilitud es:

$$L(\theta) = \prod_{i=1}^Q \prod_{j \in A} (P_{ij}(\theta))^{\delta_{ij}} \quad (20)$$

Los estimadores θ de máxima verosimilitud son aquellos que maximizan esta función. Por simplicidad, habitualmente no se maximiza esta función directamente, sino que su logaritmo $l(\theta)$, lo que conduce al mismo óptimo, dado que el logaritmo es una función creciente. La función a maximizar es, por lo tanto:

$$l(\theta) = \log L(\theta) = \sum_{i=1}^Q \sum_{j \in A} \delta_{ij} \log P_{ij}(\theta) \quad (21)$$

El resultado de la maximización de $l(\theta)$ son estimadores θ^* de los parámetros poblacionales y se distribuyen asintóticamente $N(\theta, s^2)$, donde:

$$s^2 = \frac{-1}{E \left(\frac{\partial^2 l(\theta)}{\partial \theta^2} \right)} \quad (22)$$

Los estimadores θ^* pueden ser sesgados para pequeñas muestras, pero el sesgo es pequeño en muestras suficientemente grandes. Ortúzar y Willumsen (2011) sugieren que normalmente tamaños de muestras de 500 a 1000 observaciones son adecuadas.

2.6 *Constrained Multinomial Logit*

En la literatura especializada de modelos probabilísticos de elecciones discretas que abordan la existencia de restricciones en las funciones de valor (utilidad o disposición a pagar) existe un modelo interesante de revisar con detalle por su aporte a la modelación de restricciones: el modelo Logit Multinomial Restringido (*Constrained Multinomial Logit* de Martínez *et al*, 2009).

Este modelo introduce una nueva componente a función de utilidad, denominada *cutoff*, la cual tiene el propósito eliminar la probabilidad de elección de aquellas alternativas que violan determinadas restricciones que actúan sobre éstas. Ciertamente, la restricción que es relevante en este trabajo es la presupuestaria. Estas componentes de penalización no actúan dicotómicamente eliminando las alternativas cuando no cumplen una restricción, sino que lo hacen de un modo probabilístico reduciendo las probabilidades de elección en la medida de que las restricciones, si bien siendo cumplidas, se aproximen a ser violadas. Castro *et al* (2013) estudian su calibración.

A continuación, se mostrará el contexto en el cual el comportamiento de los consumidores se ve afectado por restricciones y cómo el modelo las trata a través de la utilización de funciones de *cutoff*.

Comportamientos no compensatorios

Uno de los supuestos microeconómicos principales en la modelación de la elección que realizan los consumidores, es que utilizan una estrategia compensatoria, esto es que siempre existe un *trade off* entre los atributos que componen las funciones de utilidad. Este supuesto ha sido criticado, estableciéndose incluso que un comportamiento no compensatorio sería más realista para modelar sus decisiones. En la formulación del modelo Logit Multinomial Restringido, los autores revisaron algunos comportamientos no compensatorios que se discuten a continuación.

Una estrategia no compensatoria es la *Eliminación por aspectos*, propuesto por Tversky (1972) donde se eliminan aquellas alternativas del *choice set* o conjunto de elecciones posibles por el agente, que tienen algún atributo de la elección que no cumple cierto criterio o condición.

Otra forma para atender a comportamientos no compensatorios consiste en desacoplar la consideración de alternativas de elección, del proceso de elección dentro de este espacio de opciones consideradas. Este enfoque fue propuesto originalmente por Manski y Lerman (1977), y luego continuado por Swaitt y Ben-Akiva (1987), Ben-Akiva y Boccara (1995) y Cantillo y Ortuzar (2005), entre otros. Se modela el proceso de elección en dos etapas. Primero, se supone que el individuo define el conjunto factible de alternativas de elección, para luego, en una segunda etapa, con un modelo compensatorio, se calculan las probabilidades de elección condicionales al *choice set* que fue generado en la primera etapa. Este método, si bien es riguroso es muy costoso computacionalmente, dado que el número posible de *choice sets* es muy grande, igual a 2^{m-1} , siendo m el número de alternativas.

Cascetta y Papola (2001) presentaron un nuevo modelo para tratar este problema, el denominado *implicit availability perception model* (IAP). En esta formulación el *choice*

set es difuso, es decir, cada opción de elección tiene un grado de pertenencia a cada *choice set*.

Swait (2001) por su parte introduce restricciones a los valores de los atributos y precios por medio de cutoffs que se introducen en la función de utilidad. Similar al proceso de *Eliminación por aspectos* las alternativas que sobrepasan los valores de las restricciones, escapan del dominio factible. La forma en que actúan las funciones de *cutoff* es progresiva porque en la medida que los valores de los atributos restringidos se aproximan a su cota, la función de *cutoff* empieza a penalizar la función de utilidad. Sin embargo, como la penalización es lineal no es diferenciable en el punto en que es activa la restricción.

Problema de elección discreta restringido

Considérese el siguiente tipo de problemas de optimización utilizados ampliamente para describir el comportamiento de un agente frente a bienes discretos. Cada agente n se comporta de acuerdo a su función de utilidad indirecta U_n cuando decide la mejor elección dentro de un grupo de I alternativas discretas que pertenecen al conjunto C de alternativas disponibles. Se asume que esa función de utilidad depende de $K - 1$ atributos, denotadas por el vector $X \in R^{(K-1) \times I}$ y del vector de precios $p \in R^I$, tal como aparece en Martínez *et al* (2009).

$$\begin{aligned} & \max_{\delta_{ni}} \sum_{i \in C} \delta_{ni} U_n(X_i, p_i) \\ & \text{s. a. } \sum_{i \in C} \delta_{ni} = 1, \delta_{ni} \in \{0,1\} \forall i \in \{1,2, \dots, I\} \in C \end{aligned} \tag{23}$$

En este problema, δ_{ni} representa la elección del individuo X_i , los atributos exógenos de la alternativa i y p_i el precio exógeno de la alternativa i y $U(X, p)$ es la función de utilidad indirecta.

Es crucial notar que en el problema (23) la función de utilidad indirecta está recogiendo, por definición, todas las restricciones a las que se enfrenta el consumidor (por ejemplo, de tiempo y presupuesto, sujeta a la elección discreta). La violación de estos supuestos puede llevar a grandes errores al predecir demanda (Martínez *et al*, 2009). El problema, tal como está, presenta una dificultad para el modelador a la hora de especificar una función de utilidad indirecta que recoja la complejidad del comportamiento de los consumidores bajo restricciones. En vez de buscar una función que represente el complejo comportamiento de los usuarios, se modifica el problema de los usuarios haciendo explícitas las restricciones.

$$\begin{aligned}
& \max_{\delta_{ni}} \sum_{i \in C} \delta_{ni} U_n(X_i, p_i) \\
& \text{s. a. } \sum_{i \in C} \delta_{ni} = 1 \quad \delta_{ni} \in \{0,1\} \quad \forall i \in C \\
& a_{nk} \leq X_{ik} \leq b_{ik} \quad \forall i \in C, k = \{1, \dots, K-1\} \\
& a_{nk} \leq p_i \leq b_{ik} \quad \forall i \in C
\end{aligned} \tag{24}$$

En el interior del dominio el comportamiento del individuo está gobernado por U y en el borde por las restricciones, esta descomposición permite una forma funcional de U más sencilla que una que incluyera el comportamiento completo de los individuos. Esto motiva el enfoque de Manski de dos etapas, donde primero se encuentra un subconjunto donde las restricciones se cumplen, seguidas de una segunda etapa que modela la elección de máxima utilidad dentro del subconjunto usando una utilidad compensatoria.

Si por simplicidad se define el vector $Z_i = (X_i, p_i) \in R^K$ y se reescribe el problema (24), queda un problema de maximización irrestricto.

$$\begin{aligned}
\max_{\delta_{ni}} L = & \left(\sum_{i \in C} \delta_{ni} U_n(X_i, p_i) + \sum_k \lambda_{nki} (b_{nk} - Z_{ik}) - \sum_k \kappa_{nik} (a_{nk} - Z_{ik}) \right) \\
& + \gamma_n \left(\sum_{i \in C} \delta_{ni} - 1 \right)
\end{aligned} \tag{25}$$

Donde λ, κ y γ representan los multiplicadores de Lagrange asociados a las restricciones. El hecho de que el problema (25) sea irrestricto motiva el enfoque implícito de Swait (2001) y Cascetta y Papola (2001).

La utilidad aleatoria restringida

Consideremos, ahora el modelo clásico donde la función de utilidad es una variable aleatoria, eso es $U_n = V_n + \varepsilon_n$, con V_n como el término compensatorio sistemático de la utilidad y ε_n un término de error. El modelo Logit, mostrado en 2.5. asume que los términos aleatorios distribuyen Gumbel, lo que implica que $\varepsilon \in [-\infty, \infty]$, y las utilidades son irrestrictas. Por definición, entonces, los modelos logit, asumen utilidades irrestrictas y por consiguiente, el comportamiento de los consumidores es también irrestricto a menos que se especifique su *choice set*, eliminando alternativas de elección cuyo vector de atributos está fuera del dominio de elección del elector. Este es enfoque en dos etapas anteriormente mencionado.

Alternativamente, se puede utilizar un método implícito, definiendo una función de utilidad restringida, una que induce al individuo a elegir dentro de su dominio factible con una probabilidad, además de ser definida como compensatoria al interior del dominio del individuo.

Para obtener una función de utilidad restringida al dominio, como en el modelo IAP, se incorpora a la típica función de utilidad compensatoria un nuevo término de *cutoff* llamado penalización de la utilidad (*utility penalty*). Entonces, la utilidad restringida se define como la suma del término compensatorio V^c y el *cutoff* aditivo, como indica la siguiente expresión:

$$V_n(Z_i) = V_n^c + \frac{1}{\mu} \ln \phi_{ni}(Z_i) + \varepsilon_{ni} \quad (26)$$

Donde Z_i representa el atributo de la alternativa i bajo restricción, ε el error aleatorio que se distribuye Gumbel $(0, \mu)$ al igual que en el modelo Logit Multinomial.

Como las restricciones que afectan la elección de una alternativa dada pueden ser más que una, el término de penalidad se define como un factor multidimensional de *cutoff* $\phi_{ni} = \prod_{k=1}^K \phi_{nki}^L \cdot \phi_{nki}^U$ que está compuesto por un set de K *cutoffs* superiores e inferiores. Para cada n_{ki} se define el factor de cutoff elemental superior e inferior como:

$$\phi_{nki}^L = \frac{1}{1 + \exp(\omega_k(a_{nk} - Z_{ki} + \rho_k))} = \begin{cases} 1 & \text{si } (a_{nk} - Z_{ki}) \rightarrow -\infty \\ \eta_k & \text{si } a_{nk} = Z_{ki} \end{cases} \quad (27)$$

$$\phi_{nki}^U = \frac{1}{1 + \exp(\omega_k(Z_{ki} - b_{nk} + \rho_k))} = \begin{cases} 1 & \text{si } (b_{nk} - Z_{ki}) \rightarrow \infty \\ \eta_k & \text{si } b_{nk} = Z_{ki} \end{cases} \quad (28)$$

Se define $\omega_j > 0, \forall j$, y adicionalmente se define:

$$\rho_k = \frac{1}{\omega_k} \cdot \ln\left(\frac{1 - \eta_k}{\eta_k}\right) \quad (29)$$

Con η el parámetro de tolerancia del cutoff.

El factor de *cutoff* es una penalización a la función de utilidad actuando cuando algún valor de los atributos Z_i se acerca a la restricción o la supera y si los valores permanecen en el interior del dominio D_n , alejado de los bordes, entonces el cutoff baja su influencia en la función de utilidad. Si cualquier elemento tiende a violar su restricción, es decir, $Z_{ki} \notin D_n$ entonces $\phi_{ni} \rightarrow 0$, la utilidad de la alternativa tiende a menos infinito y la probabilidad de elección tiende a cero. Por el contrario, si el elemento se aleja de violar la restricción $Z_i \in D_n$, y $\phi_{ni} \rightarrow 1$. La pendiente de la función de cutoff está controlada por el parámetro ω , definiendo la influencia de la restricción.

El modelo *Constrained Multinomial Logit*

Siguiendo con lo descrito en Martínez *et al* (2009), la alternativa que un consumidor escoge es aquella que le provee la máxima utilidad, y en el caso de considerar penalidades, es la solución del siguiente problema:

$$\max_{i \in C} \left\{ V_n(Z_i) = V_n^c + \frac{1}{\mu} \ln \phi_{ni} + \varepsilon_{ni} \right\} \quad (30)$$

Con C el conjunto universal de alternativas y $\phi_{ni} = \phi_{ni}(Z_i)$. La expresión (30) es la función objetivo estocástica reducida que representa una versión estocástica del problema de elección (24), con V la función de utilidad indirecta. Bajo este enfoque, la probabilidad de elección de una alternativa i es la siguiente:

$$P_{ni} = \text{Prob} \left[V_n^c + \frac{1}{\mu} \ln \phi_{ni} + \varepsilon_{ni} \geq \max_{j \in C} \left(V_n^c + \frac{1}{\mu} \ln \phi_{nj} + \varepsilon_{nj} \right) \right] \quad (31)$$

Bajo el supuesto que la utilidad restringida distribuye i.i.d. Gumbel, la probabilidad de elección es la siguiente, dando origen al *Constrained Multinomial Logit*:

$$P_{ni} = \frac{\phi_{ni} \cdot \exp(\mu V_{ni}^c)}{\sum_{j \in C} \phi_{nj} \cdot \exp(\mu V_{nj}^c)} \quad (32)$$

Esta expresión representa una probabilidad de elección que cumple con el dominio factible D_n , haciendo que se descarten, en probabilidad, opciones que violan determinadas restricciones.

2.7 Síntesis del *Constrained Multinomial Logit*

En la sección anterior se ha hecho una revisión exhaustiva del Modelo Logit Multinomial Restringido este modelo corresponde a una heurística para incorporar las restricciones que aplican sobre las alternativas de elección en un modelo de elección discreta, y que reducen el conjunto disponible de éstas. En particular, este modelo se desarrolla en el caso en que las funciones de valor son variables aleatorias i.i.d Gumbel. Este modelo fue aplicado a un caso particular de restricciones: cuando los valores de ciertos atributos de las alternativas escapan de un rango de admisibilidad para el usuario y por tanto, eliminan del conjunto de elección a las alternativas afectadas. Este enfoque también puede aplicarse a otro caso: cuando los valores de las funciones de valor (utilidades o disposiciones a pagar) caen fuera del espacio de admisibilidad y esto conduce a la eliminación de alternativas. Cuando se trata de la disposición a pagar en modelos de elección de remate (por ejemplo, en el caso del modelo *bid* de localización de actividades) este enfoque puede aplicarse cuando existen cotas mínimas en estas funciones, representando la existencia de remates con posturas mínimas. Sin embargo, las cotas máximas de las disposiciones a pagar que provienen de restricciones

presupuestarias de los agentes, no calza con este enfoque, porque no conduce a la eliminación de postores (alternativas de elección en un modelo de remate), sino a la consideración de que las posturas de cada uno de los postores presentes en un remate no pueden ser superiores a estas cotas. Bajo el enfoque del CMNL, en la medida de que los agentes hagan posturas que se aproximen a sus cotas presupuestarias serían “eliminados” (en probabilidad), pero eso no debería ocurrir, sino que estimar posturas que respeten estas restricciones.

3 MODELO PROPUESTO

En esta sección se desarrolla un modelo general de probabilidad de adjudicación de bienes discretos a agentes postores bajo la regla de remate usando la condición que sus posturas no pueden sobrepasar sus respectivas disponibilidades presupuestarias. Por simplicidad, se deduce este modelo para el caso binomial, que considera solamente dos agentes competidores, pero el enfoque puede extenderse al caso de múltiples agentes. El modelo se deduce para posturas aleatorias con leyes de probabilidad general, pero se aplica al caso en que las posturas son variables aleatorias i.i.d. Gumbel obteniéndose expresiones explícitas de estas probabilidades al igual que en el caso del Logit Binomial que no considera cotas superiores en las posturas. Ciertamente, este modelo es usado como un nuevo modelo de probabilidad de localización de agentes en una ciudad bajo el enfoque bid o de remate, donde todos los bienes raíces son adjudicados a los mejores postores. Para fijar ideas, en toda la deducción del modelo que se presenta a continuación, se lo considerará aplicado al mercado inmobiliario, pero puede ser aplicado directamente a otros problemas donde los bienes son únicos, y por tanto, es aplicable la regla de remate.

Para cada bien inmueble del mercado se supone que existen dos tipos de agentes postores interesados en localizarse allí y sus disposiciones a pagar representan el precio, o el gasto que cada agente está dispuesto a incurrir para adquirir el bien, considerando que han fijado sus gastos en otros bienes según su ingreso para maximizar sus niveles de utilidad (paradigma de comportamiento racional). Por lo tanto, la propia definición de una disposición a pagar depende del presupuesto disponible para adquirir el bien. Es interesante observar que la disposición a pagar que representa preferencias no restringidas corresponde a aquella en que el ingreso es suficientemente alto. No obstante, debe tenerse presente, que los valores de la disposición a pagar que un usuario realiza efectivamente en el mercado deben encontrarse siempre dentro de su real disponibilidad presupuestaria..

Se supone que las posturas de los agentes (B_1 y B_2) están compuestas por una componente determinística (\bar{B}_1 y \bar{B}_2), que corresponden a valoraciones pecuniarias de estos agentes por los atributos de los bienes raíces (x), y otra estocástica que recoge la heterogeneidad no observada por el modelador, que se comporta como una variable aleatoria (ε_1 y ε_2). Si estos errores se distribuyeran *Gumbel* ($0, \mu$), las distribuciones estocásticas de las posturas serían las siguientes:

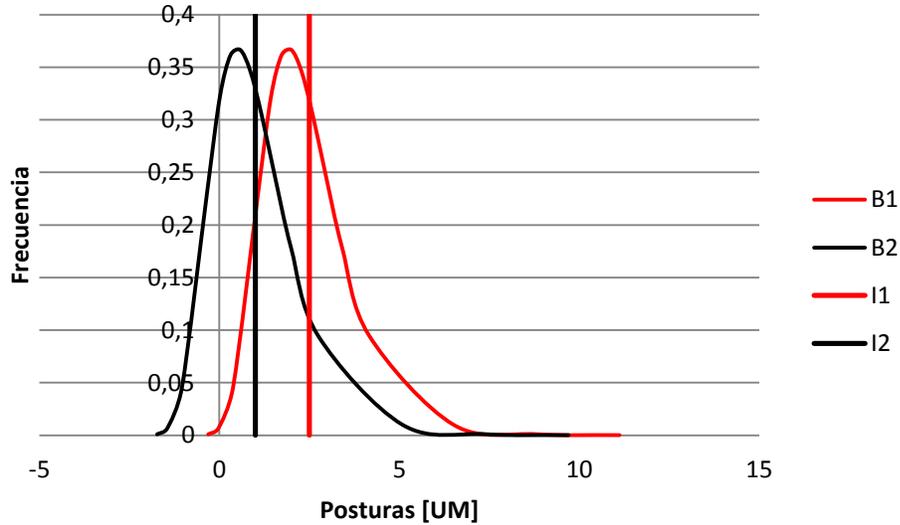


Figura 1: Frecuencia de posturas $\mu=1$

Las líneas verticales representan la restricción presupuestaria de los agentes (I_1 e I_2). Como se observa, las posturas no restringidas (con un ingreso suficientemente alto) tienen un recorrido desde $-\infty$ a $+\infty$ con una concentración de valores en un rango bastante menor dependiendo del factor de escala μ usado, pero las posturas observadas siempre serán menores o iguales a su respectiva cota (I_1 o I_2).

A continuación, se derivará la expresión analítica de las probabilidades de que un agente gane el remate bajo restricción de ingreso.

3.1 Formulación matemática de las probabilidades Logit binomiales con cota superior

Un agente (i) se localiza cuando éste realiza una postura o bid (B_i) mayor que la del agente competidor. Ahora, sea P_i a probabilidad de que i se localice sujeta a condiciones de remate al mejor postor, condicional en la restricciones presupuestarias que actúan sobre ambos agentes (D).

$$P_i = \frac{P_{i,D}}{P_D} \quad (33)$$

Donde $P_{i,D}$ es la probabilidad conjunta de elección y de cumplimiento de la restricción y P_D es la probabilidad del cumplimiento de la restricción.

Escrito de otro modo:

$$P_{i,D} = P(B_i \geq B_j \quad j \neq i; B_j \leq I_j; B_i \leq I_i) \quad (34)$$

Y se cumple que:

$$P_{i,D} + P_{j,D} = P_D \quad (35)$$

Demostrado en una propiedad análoga en Anexo A.

Ahora, sean:

$$B_k = \bar{B}_k + \varepsilon_k \quad (36)$$

$$\Delta_{ij} = \bar{B}_i - \bar{B}_j \quad (37)$$

Luego:

$$P_{i,D} = P(\varepsilon_j \leq \varepsilon_i + \Delta_{ij}; \varepsilon_j \leq I_j - \bar{B}_i + \Delta_{ij} \quad j \neq i; \varepsilon_i \leq I_i - \bar{B}_i) \quad (38)$$

$$P_{i,D} = P(\varepsilon_j \leq \text{Mín}(\varepsilon_i, I_j - \bar{B}_i) + \Delta_{ij}; \quad j \neq i; \varepsilon_i \leq I_i - \bar{B}_i)$$

Al integrar sobre el recorrido de la variable ε_i se obtiene:

$$P_{i,D} = \int_{-\infty}^{I_i - \bar{B}_i} F_j(\text{Mín}(z, I_j - \bar{B}_i) + \Delta_{ij}) f_i(z) dz \quad (39)$$

Para calcular esta expresión se consideran dos casos: $I_j \leq I_i$ y $I_i < I_j$

Caso N°1: $I_j \leq I_i$

Si $I_j \leq I_i$ entonces:

$$I_j - \bar{B}_i \leq I_i - \bar{B}_i \quad (40)$$

Luego:

$$\text{Mín}(z, I_j - \bar{B}_i) = \begin{cases} z & \text{si } z \leq I_j - \bar{B}_i \\ I_j - \bar{B}_i & \text{si } I_j - \bar{B}_i \leq z \leq I_i - \bar{B}_i \end{cases} \quad (41)$$

Por lo tanto, este resultado sugiere dividir la integral anterior en dos tramos:

$$\begin{aligned}
P_{i,D} = & \int_{-\infty}^{I_j - \bar{B}_i} F_j(\text{Mín}(z, I_j - \bar{B}_i) + \Delta_{ij}) f_i(z) dz \\
& + \int_{I_j - \bar{B}_i}^{I_i - \bar{B}_i} F_j(\text{Mín}(z, I_j - \bar{B}_i) + \Delta_{ij}) f_i(z) dz
\end{aligned} \tag{42}$$

Al evaluar cada integral se obtiene:

$$P_{i,D} = \int_{-\infty}^{I_j - \bar{B}_i} F_j(z + \Delta_{ij}) f_i(z) dz + \int_{I_j - \bar{B}_i}^{I_i - \bar{B}_i} F_j(I_j - \bar{B}_j) f_i(z) dz \tag{43}$$

Al observar que el término $F_j(I_j - \bar{B}_j)$ de la segunda integral no depende de z , se deduce que:

$$P_{i,D} = \int_{-\infty}^{I_j - \bar{B}_i} F_j(z + \Delta_{ij}) f_i(z) dz + F_j(I_j - \bar{B}_j) \left(F_i(I_i - \bar{B}_i) - F_i(I_j - \bar{B}_i) \right) \tag{44}$$

Caso N°2: $I_i < I_j$

Si $I_i < I_j$, entonces:

$$I_i - \bar{B}_i < I_j - \bar{B}_i \tag{45}$$

Luego, si $z \leq I_i - \bar{B}_i$, entonces:

$$z \leq I_j - \bar{B}_i \tag{46}$$

Por lo tanto, la expresión (39):

$$P_{i,D} = \int_{-\infty}^{I_i - \bar{B}_i} F_j(z + \Delta_{ij}) f_i(z) dz \tag{47}$$

Se ha deducido un modelo válido para errores independientes con leyes generales de probabilidad. Para obtener expresiones explícitas se supondrá en lo que sigue que los errores ε_i son i.i.d. *Gumbel*(0, μ), con $\mu > 0$, cuya función de densidad es f y función de distribución es:

$$F(x) = \exp(-\exp(-\mu x)) \quad (48)$$

Por lo tanto, la expresión para el

Caso N°1: $I_j \leq I_i$

$$P_{i,D} = \int_{-\infty}^{I_j - \bar{B}_i} F_j(z + \Delta_{ij}) f_i(z) dz + F_j(I_j - \bar{B}_j) \left(F_i(I_i - \bar{B}_i) - F_i(I_j - \bar{B}_i) \right) \quad (49)$$

Al utilizar el Anexo B:

$$P_{i,D} = F(I_j - \bar{B}_j) F(I_j - \bar{B}_i) \frac{\exp(\mu \bar{B}_i)}{\exp(\mu \bar{B}_i) + \exp(\mu \bar{B}_j)} + F_j(I_j - \bar{B}_j) \left(F_i(I_i - \bar{B}_i) - F_i(I_j - \bar{B}_i) \right) \quad (50)$$

O equivalentemente:

$$P_{i,D} = F(I_j - \bar{B}_j) \left(F(I_j - \bar{B}_i) \frac{\exp(\mu \bar{B}_i)}{\exp(\mu \bar{B}_i) + \exp(\mu \bar{B}_j)} + F(I_i - \bar{B}_i) - F(I_j - \bar{B}_i) \right) \quad (51)$$

Revisando el otro caso:

Caso N°2: $I_i < I_j$

$$P_{i,D} = \int_{-\infty}^{I_i - \bar{B}_i} F_j(z + \Delta_{ij}) f_i(z) dz \quad (52)$$

Utilizando el Anexo B:

$$P_{i,D} = F(I_i - \bar{B}_j)F(I_i - \bar{B}_i) \frac{\exp(\mu\bar{B}_i)}{\exp(\mu\bar{B}_i) + \exp(\mu\bar{B}_j)} \quad (53)$$

Si $I_j \leq I_i$

$$P_{i,D} = F(I_j - \bar{B}_j) \left(F(I_i - \bar{B}_i) - F(I_j - \bar{B}_i) \frac{\exp(\mu\bar{B}_j)}{\exp(\mu\bar{B}_i) + \exp(\mu\bar{B}_j)} \right) \quad (54)$$

$$P_{j,D} = F(I_j - \bar{B}_i)F(I_j - \bar{B}_j) \frac{\exp(\mu\bar{B}_j)}{\exp(\mu\bar{B}_i) + \exp(\mu\bar{B}_j)} \quad (55)$$

En ambos casos se cumple que:

$$P_{i,D} + P_{j,D} = F(I_i - \bar{B}_i)F(I_j - \bar{B}_j) = P_D \quad (56)$$

Notación:

$$q_i \equiv \frac{\exp(\mu\bar{B}_i)}{\exp(\mu\bar{B}_i) + \exp(\mu\bar{B}_j)} \text{ y } q_j \equiv \frac{\exp(\mu\bar{B}_j)}{\exp(\mu\bar{B}_i) + \exp(\mu\bar{B}_j)} \quad (57)$$

Notar que q_j y q_i se puede expresar en función de $\Delta\bar{B}_{ij}$:

$$\begin{aligned} q_j &= \frac{\exp(\mu\bar{B}_j)}{\exp(\mu\bar{B}_i) + \exp(\mu\bar{B}_j)} \\ &= \frac{\exp(\mu\bar{B}_i)}{\exp(\mu\bar{B}_i) + \exp(\mu\bar{B}_j)} * \frac{\exp(-\mu\bar{B}_j)}{\exp(-\mu\bar{B}_j)} \\ &= \frac{1}{\exp(\mu\bar{B}_i - \mu\bar{B}_j) + 1} \end{aligned}$$

$$q_j(\Delta\bar{B}_{ij}) = \frac{1}{\exp(\mu\Delta\bar{B}_{ij}) + 1} \quad (58)$$

Se demuestra de manera homóloga para q_i .

Luego, para el caso general:

$$P_i = \frac{P_{i,D}}{P_D} = 1 - \frac{F(I_j - \bar{B}_i)}{F(I_i - \bar{B}_i)} \frac{\exp(\mu\bar{B}_j)}{\exp(\mu\bar{B}_i) + \exp(\mu\bar{B}_j)} = 1 - \frac{F(I_j - \bar{B}_i)}{F(I_i - \bar{B}_i)} q_j \quad (59)$$

$$P_j = \frac{P_{j,D}}{P_D} = \frac{F(I_j - \bar{B}_i)}{F(I_i - \bar{B}_i)} \frac{\exp(\mu\bar{B}_j)}{\exp(\mu\bar{B}_i) + \exp(\mu\bar{B}_j)} = \frac{F(I_j - \bar{B}_i)}{F(I_i - \bar{B}_i)} q_j \quad (60)$$

Es interesante notar que la probabilidad restringida coincide con la irrestricta, salvo por un factor que pondera la probabilidad Logit Binomial (irrestricta) q_j . Este factor depende de la postura determinística del agente con mayor ingreso (\bar{B}_i) y las cotas (I_1 e I_2). A (59) y (60) se les llamará versión uno (v_1).

Utilizando las propiedades de F se pueden expresar las probabilidades P_i y P_j en función de $(I_i - \bar{B}_i)$ y $(I_j - \bar{B}_j)$ respectivamente.

Notación:

$$\Delta I_{ij} \equiv I_i - I_j \quad \text{y} \quad \Delta\bar{B}_{ij} \equiv \bar{B}_i - \bar{B}_j \quad (61)$$

Entonces, P_i queda expresado como:

$$\begin{aligned} P_i &= 1 - \frac{F(I_j - \bar{B}_i)}{F(I_i - \bar{B}_i)} q_j \\ &= 1 - \frac{F(I_j - \bar{B}_i + I_i - I_i)}{F(I_i - \bar{B}_i)} q_j \\ &= 1 - \frac{F(I_i - \bar{B}_i - \Delta I_{ij})}{F(I_i - \bar{B}_i)} q_j \end{aligned}$$

$$\begin{aligned}
&= 1 - \frac{F(I_i - \bar{B}_i) \exp(\mu \Delta I_{ij})}{F(I_i - \bar{B}_i)} q_j \\
&= 1 - F(I_i - \bar{B}_i) \exp(\mu \Delta I_{ij})^{-1} q_j
\end{aligned} \tag{62}$$

Esta expresión junto a su respectiva P_i se les llamará versión 2 ($v2$).

Procediendo de manera similar, P_j queda determinado de la siguiente manera:

$$\begin{aligned}
P_j &= \frac{F(I_j - \bar{B}_i)}{F(I_i - \bar{B}_i)} q_j \\
&= \frac{F(I_j - \bar{B}_i + \bar{B}_j - \bar{B}_j)}{F(I_i + I_j - I_j - \bar{B}_i)} q_j \\
&= \frac{F(I_j - \bar{B}_i + \bar{B}_j - \bar{B}_j)}{F(I_i + I_j - I_j - \bar{B}_i)} q_j \\
&= \frac{F(I_j - \bar{B}_j - \Delta \bar{B}_{ij})}{F(I_j - \bar{B}_i + \Delta I_{ij} + \bar{B}_j - \bar{B}_j)} q_j \\
&= \frac{F(I_j - \bar{B}_j - \Delta \bar{B}_{ij})}{F(I_j - \bar{B}_j + \Delta I_{ij} - \Delta \bar{B}_{ij})} q_j \\
&= \frac{F(I_j - \bar{B}_j) \exp(\mu \Delta \bar{B}_{ij})}{F(I_j - \bar{B}_j) \exp(\mu \Delta \bar{B}_{ij}) \exp(-\mu \Delta I_{ij})} q_j \\
&= F(I_j - \bar{B}_j) \exp(\mu \Delta \bar{B}_{ij})^{-1} \exp(-\mu \Delta I_{ij}) q_j \\
&= F(I_j - \bar{B}_j) \exp(\mu \Delta \bar{B}_{ij}) (1 - \exp(-\mu \Delta I_{ij})) q_j
\end{aligned} \tag{63}$$

Junto a su respectiva P_i , esta reformulación de la probabilidades se denominará versión tres ($v3$).

3.2 Condición en que el modelo restringido recupera el irrestricto.

El modelo de logit binomial con restricciones recién presentado es un caso general del modelo logit irrestricto para dos agentes, por lo que es interesante encontrar bajo qué condiciones el modelo nuevo se comporta del mismo modo que el modelo irrestricto, es decir $P_j = q_j$ y $P_i = q_i$.

Evidentemente, ambos modelos son idénticos si las cotas tienden a infinito, pero existe otro caso menos intuitivo donde esto también ocurre. Esto ocurre cuando las cotas son iguales y finitas ($I_i = I_j$). Se puede ver en las tres expresiones equivalentes de las probabilidades.

En v_1 :

$$P_j = \frac{F(I_j - \bar{B}_i)}{F(I_i - \bar{B}_i)} q_j = \frac{F(I_i - \bar{B}_i)}{F(I_i - \bar{B}_i)} q_j = q_j \quad (64)$$

En v_2 :

$$P_j = F(I_i - \bar{B}_i)^{\exp(\mu\Delta I_{ij})-1} q_j = F(I_i - \bar{B}_i)^{\exp(0)-1} q_j = F(I_i - \bar{B}_i)^0 q_j = q_j \quad (65)$$

En v_3 :

$$\begin{aligned} P_j &= F(I_j - \bar{B}_j)^{\exp(\mu\Delta\bar{B}_{ij})(1-\exp(-\mu\Delta I_{ij}))} q_j = F(I_i - \bar{B}_j)^{\exp(\mu\Delta\bar{B}_{ij})(1-1)} q_j \\ &= F(I_i - \bar{B}_j)^0 q_j = q_j \end{aligned} \quad (66)$$

3.3 Las cotas pueden inducir un cambio en el orden de las probabilidades.

El orden de las probabilidades puede ser alterado por la presencia de las cotas, un agente que tiene una probabilidad dada de ganar el remate puede ver esa probabilidad favorecida o mermada por la restricción.

Se toma el caso en que el modelo irrestricto es equiprobable, i.e. $q_i = q_j = 0,5$. Ahora, si las cotas son iguales el factor que acompaña a q_j en P_j es igual a uno ocurriendo que $P_j = q_j$ y $P_i = q_i$.

Si las cotas son distintas, por muy pequeña que sea esta diferencia, el factor se hace menor estricto que uno, por lo que, en el caso $q_i = q_j = 0,5$ se rompe el empate en las probabilidades.

Demostración.

Sean:

$$\begin{aligned}
 I_i &= I_j + \varepsilon \\
 \varepsilon &> 0 \\
 q_i &= q_j = 0,5
 \end{aligned}
 \tag{67}$$

Reemplazando en v_1 :

$$\begin{aligned}
 P_j &= \frac{F(I_j - \bar{B}_i)}{F(I_i - \bar{B}_i)} q_j = \frac{F(I_j - \bar{B}_i)}{F(I_j + \varepsilon - \bar{B}_i)} 0,5 = \\
 &= \frac{F(I_j - \bar{B}_i)}{F(I_j - \bar{B}_i)^{\exp(-\mu\varepsilon)}} 0,5 = F(I_j - \bar{B}_i)^{1-\exp(-\mu\varepsilon)} 0,5
 \end{aligned}
 \tag{68}$$

Como F es una función de probabilidad acumulada, está acotada superiormente por uno, i.e. $I_j - \bar{B}_i \rightarrow \infty \Rightarrow F(I_j - \bar{B}_i) \rightarrow 1$.

A su vez, si se toma un ε mayor que cero, $1 - \exp(-\mu\varepsilon)$ siempre será menor estricto que uno, que multiplica a $q_j = 0,5$ por lo que $P_j < 0,5$.

3.4 Caracterización del modelo restringido.

Una forma de que quede más claro cómo se comporta el modelo restringido bajo variaciones de sus variables es gráficamente.

Cuando se incluye en la modelación el hecho de que las posturas están acotadas superiormente, las probabilidades de elección cambian entre el modelo irrestricto y el restringido. A este cambio se le denomina “efecto restricción”.

3.4.1 Sensibilidad del modelo de elección respecto al valor de las cotas superiores.

Primero se mostrará la variación de las probabilidades restringidas e irrestrictas en la situación en que $\bar{B}_1 = \bar{B}_2 = 1$ y $\mu = 1$, haciendo variar la diferencia de las cotas dejado fija I_2 en uno y moviendo I_1 desde uno hasta 11, variando ΔI de cero a diez. Recordar que las posturas tienen una componente determinística (\bar{B}_1 y \bar{B}_2) y una aleatoria (ε_1 y ε_2) y es esta componente la que puede tomar valores que pertenecen a $(-\infty, \infty)$.

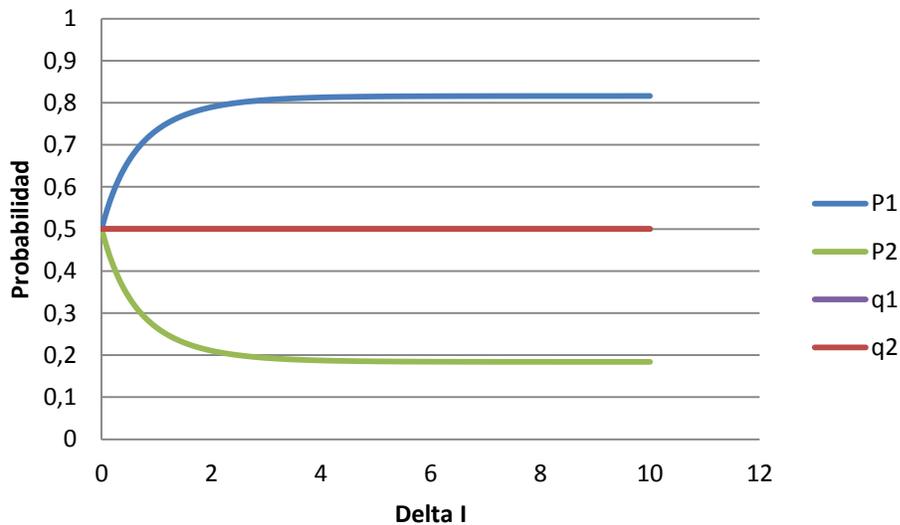


Figura 2: Probabilidades en función de ΔI , con $I_2 = 1$, $\mu=1$ y $B_1=B_2=1$

En el gráfico se observa que el valor de las probabilidades irrestrictas, q_1 y q_2 , evidentemente se mantienen iguales a un medio, dado que las posturas son iguales y la expresiones de las probabilidades irrestrictas no dependen de las cotas. Sin embargo los valores de las probabilidades restrictas, P_1 y P_2 se empiezan a alejar a medida que la diferencia en las cotas se hace mayor. Es interesante notar que llega un punto en que la diferencia de las cotas sigue aumentando pero las probabilidades de localización se mantienen iguales. Observar que en este caso la distribución de \bar{B}_1 está casi en su totalidad dentro del espacio permitido $(-\infty, I_1]$, mientras que la distribución de \bar{B}_2 tiene un menor espacio para quedar admisible $(-\infty, I_2]$.

Observar también, que cuando se cambia el valor de \bar{B}_1 y \bar{B}_2 a uno menor, el valor asintótico al que las probabilidades se estabilizan cambia.

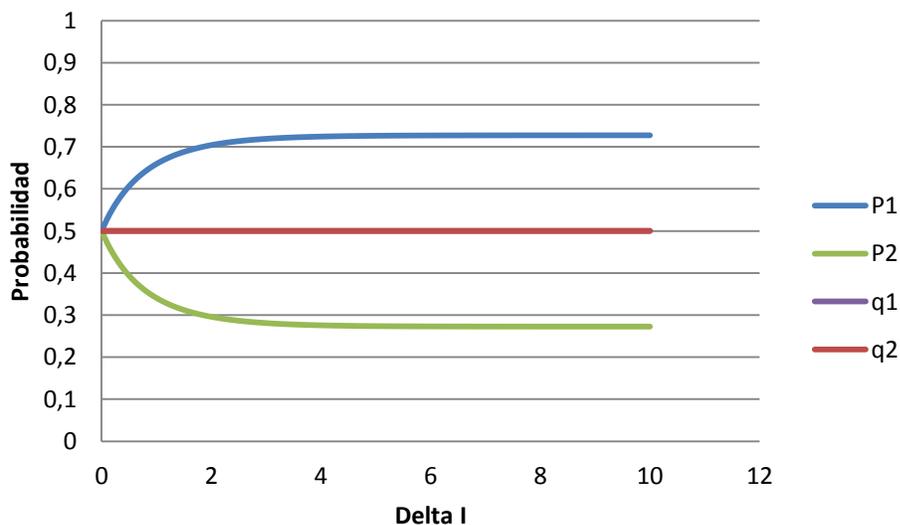


Figura 3: Probabilidades en función de ΔI , con $I_2 = 1$, $\mu=1$ y $B_1=B_2=0,5$

Este comportamiento motiva a indagar si existiera una expresión paramétrica en las posturas e I_2 a la que tienden los valores de las probabilidades restrictas cuando I_1 tiende a infinito y se deja fijo I_2 .

Tomando el límite, en efecto, se llega a la expresión:

$$\lim_{I_1 \rightarrow \infty} P_1 = \exp\left(-\exp(-\mu(I_2 - \bar{B}_1))\right) \quad (69)$$

Es interesante revisar lo que ocurre cuando las posturas son distintas entre sí. A continuación se grafican los casos en que $\bar{B}_1 > \bar{B}_2$ y luego $\bar{B}_2 > \bar{B}_1$.

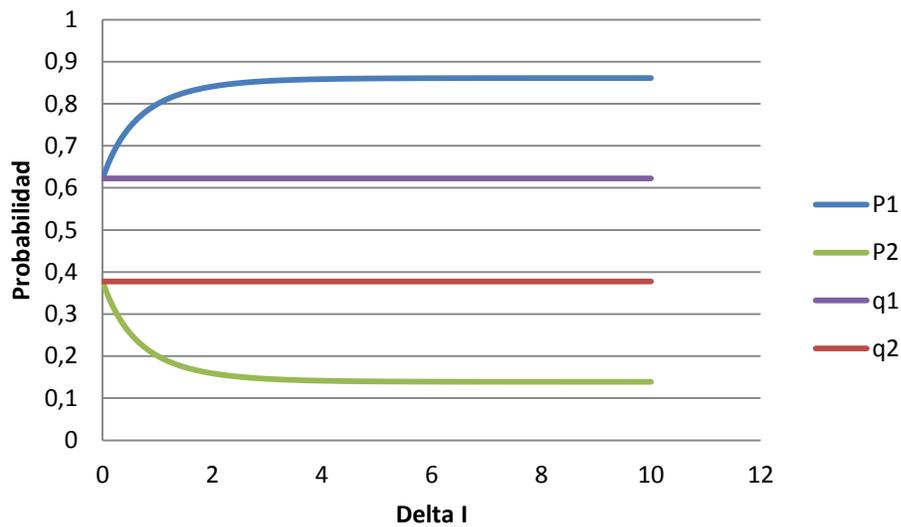


Figura 4: Probabilidades en función de ΔI , con $I_2 = 1$, $\mu=0$ y $B_1=1$ $B_2=0,5$

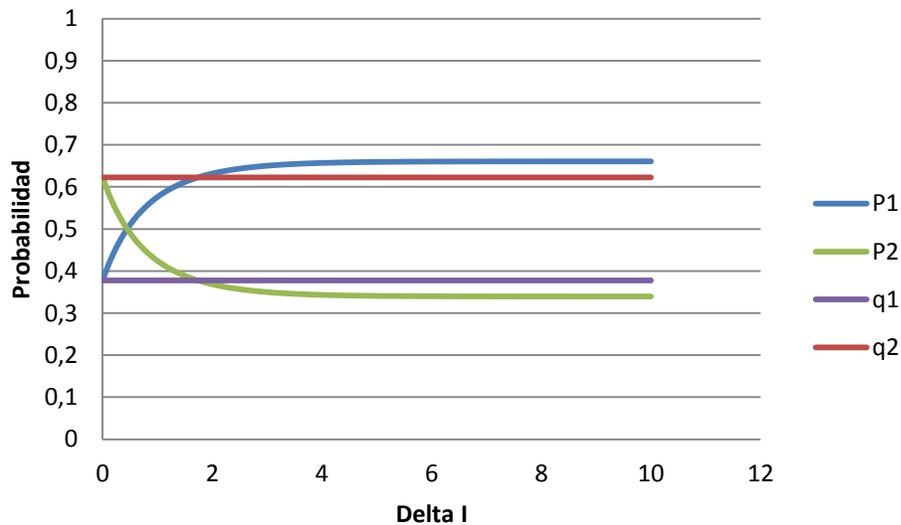


Figura 5: Probabilidades en función de ΔI , con $I_2 = 1$, $\mu=1$ y $B_1=0,5$ $B_2=1$

En la figura 4 se aprecia que, partiendo con un $\Delta I = 0$, las probabilidades restringidas coinciden con las irrestringidas. Sin embargo a medida que ΔI aumenta, el efecto restricción extrema el valor de P_1 y P_2 .

En la figura 5, como $\bar{B}_2 > \bar{B}_1$ entonces $P_2 > P_1$ cuando las cotas son iguales. Pero a medida que ΔI aumenta, el efecto restricción hace aumentar P_1 y disminuir P_2 al punto en que hay un ΔI que compensa la diferencia entre posturas, dando a ambos agentes la misma probabilidad de localización. A medida que aumenta ΔI se invierte el orden de las probabilidades y $P_1 > P_2$.

3.4.2 Sensibilidad del modelo respecto a las posturas.

A continuación se muestran dos ejemplos en que varían las probabilidades variando el $\Delta \bar{B}_{12}$ de maneras distintas. En ambos casos, ΔI es fijo igual a uno, $I_1 = 2,5$ e $I_2 = 1,5$ y $\mu = 1$.

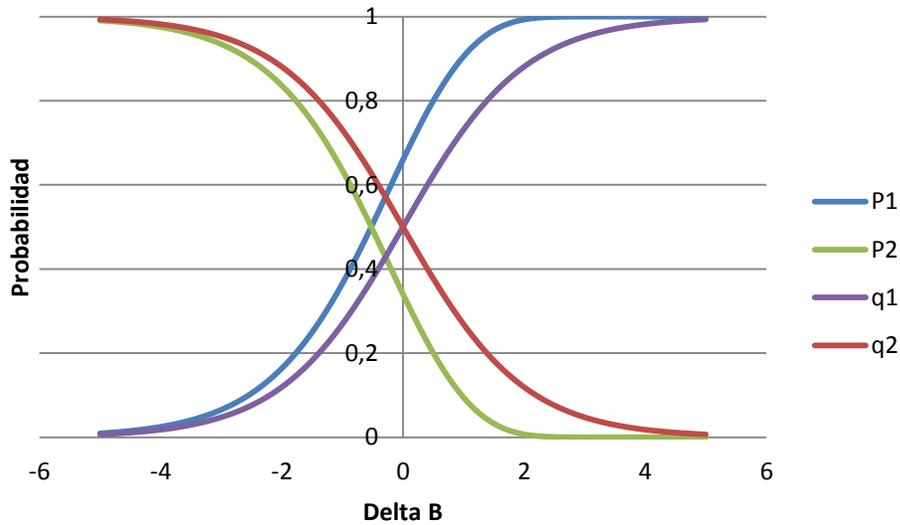


Figura 6: Probabilidades en función de ΔB , con $I_1=2,5$ $I_2= 1,5$ $\mu=1$ y $B_2=1$

En este caso la postura que varía es \bar{B}_1 y es interesante observar dos fenómenos. En primer lugar, se observa que restricto en un punto en que las posturas son distintas indica que hay una diferencia en las posturas que compensa la diferencia en presupuesto de los agentes. Y segundo, el que a pesar de que \bar{B}_1 sea mayor a I_1 y que 2 tenga una postura siempre admisible, esto no merma las posibilidades de localización de 1. Esto tiene sentido desde el punto de vista en que el tener posturas superiores a la cota propia no inhabilita al postor y que su contrincante tiene menos espacio para realizar una postura.

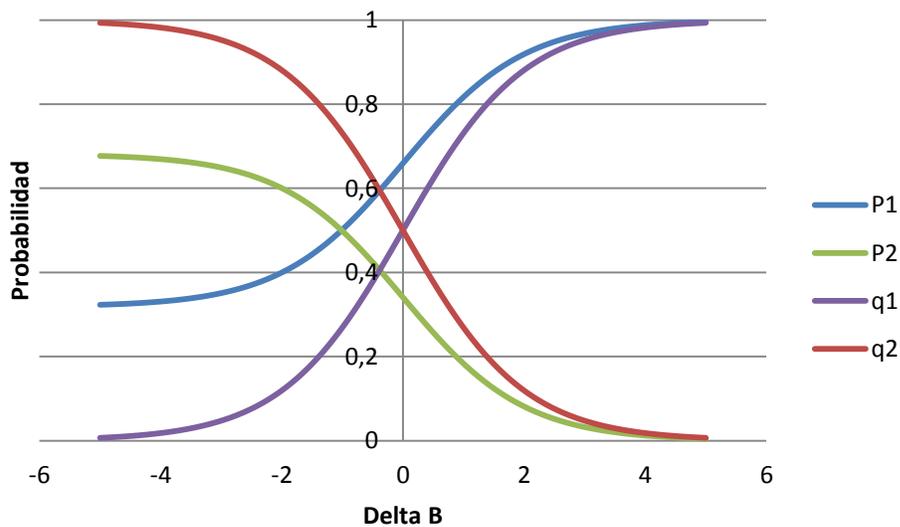


Figura 7: Probabilidades en función de ΔB , con $I_1=2,5$ $I_2= 1,5$ $\mu=1$ y $B_1=1$

En la figura 7 se observa un efecto levemente distinto al anterior. A la parte izquierda del gráfico, es intuitivo que, al tener 2 una postura inferior, 1 tenga una mayor probabilidad de ganar el remate, siendo éste doblemente favorecido por tener una restricción presupuestaria mayor y postura mayor. Luego, cuando \bar{B}_2 empieza a crecer y sobrepasa a \bar{B}_1 , hay dos tramos, uno donde $P_1 > P_2$ con $\bar{B}_1 < \bar{B}_2$, que se justifica porque $I_1 > I_2$ y otro tramo donde $P_1 < P_2$. En este último tramo hay dos efectos que van en direcciones opuestas, el tener una postura mayor favorece la probabilidad de ganar el remate y la diferencia de cotas, que desfavorece al que tiene una cota menor, haciendo que el efecto ingreso aumente para 2 y disminuya para 1. Esto hace que, a pesar de que \bar{B}_2 sea significativamente mayor que \bar{B}_1 , jamás se alcance un dominio total del agente dos sobre el agente uno, lo que vuelve sumamente importante la diferencia presupuestaria a la hora de que los agentes entren a pujar por la localización.

Tomando el límite de P_2 con B_2 tendiendo a infinito, se llega a que el límite existe y es igual al factor que acompaña a la expresión de la probabilidad irrestricta q_2 .

$$\lim_{B_2 \rightarrow \infty} P_2 = \frac{\exp\left(-\exp(-\mu(I_2 - \bar{B}_1))\right)}{\exp\left(-\exp(-\mu(I_1 - \bar{B}_1))\right)} \quad (70)$$

Este valor, como se ha mencionado antes (3.3.), es estrictamente menor que uno, cuando el valor de las cotas es distinto.

Es de importancia mencionar que, si bien, existe una diferencia en las posturas que compensa la diferencia en restricciones presupuestarias, esta no es calculable de forma analítica. De igual modo ocurre para buscar una expresión analítica para la diferencia en restricciones presupuestarias que compense las probabilidades de localización producto al $\Delta\bar{B}_{12}$.

3.4.3 Sensibilidad del modelo respecto al parámetro μ .

El parámetro μ es inversamente proporcional a la varianza de ε_1 y ε_2 . Un μ bajo, hace que los errores ocupen una fracción importante de las posturas, en cambio uno más bien alto hace que los errores tomen menor protagonismo dentro de la postura final de los agentes. A continuación se muestran gráficos que permiten comprender el efecto que tiene μ sobre las probabilidades restringidas e irrestrictas, distintos valores de B_1 y B_2 .

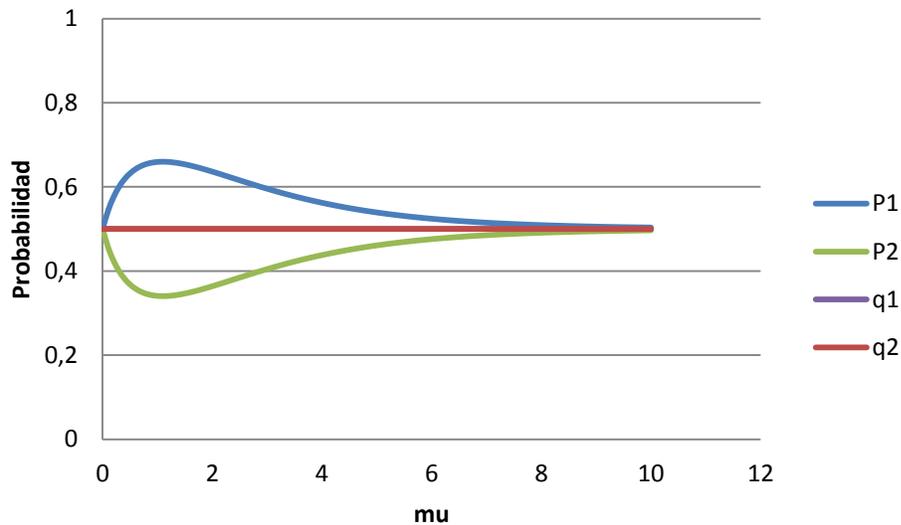


Figura 8: Probabilidades en función de μ , con $I_1=2,5$ $I_2= 1,5$ y $B_1=B_2=1$

Para el caso de posturas iguales, hay dos efectos interesantes de notar. En primer lugar, si μ es cero, las probabilidades de localización para ambos agentes son iguales. Esto se debe a que si μ es cero, las posturas no tienen relevancia para las probabilidades y el modelo irrestricto se vuelve un modelo de elección equiprobable, esto ocurre de la misma forma al modelo con restricciones. El segundo punto que es interesante de notar, es que dado que las posturas son iguales, lo único que diferencia a los agentes son las restricciones presupuestales que producen el efecto restricción.

Ahora se presentan los valores de las probabilidades en función de μ en un caso similar al anteriormente presentado, con la diferencia de que las posturas \bar{B}_1 y \bar{B}_2 ya no son iguales.

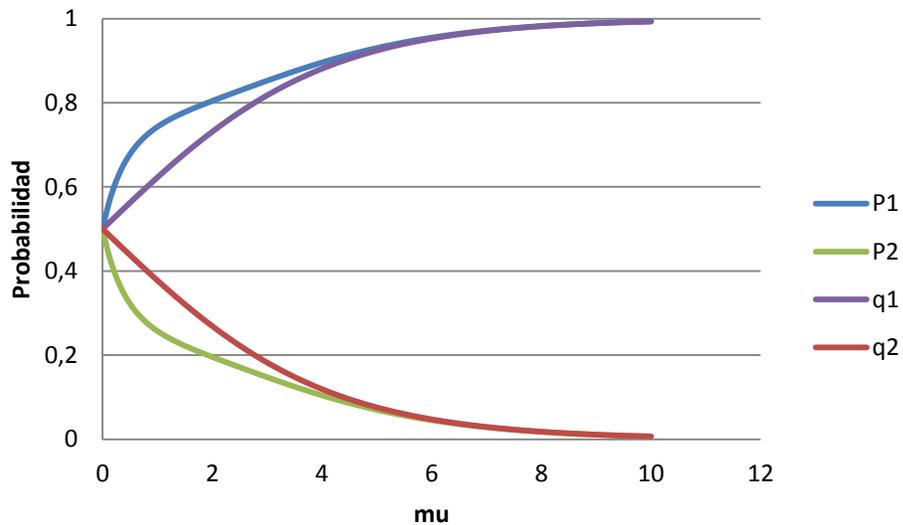


Figura 9: Probabilidades en función de μ , con $I_1=2,5$ $I_2= 1,5$, $B_1=1$ y $B_2=0,5$

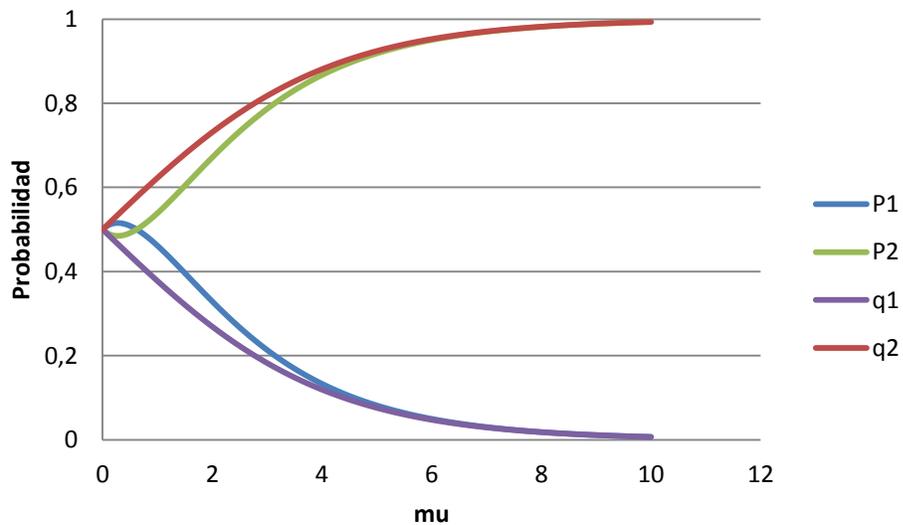


Figura 10: Probabilidades en función de μ , con $I_1=2,5$ $I_2= 1,5$, $B_1=0,5$ y $B_2=1$

En las figuras 9 y 10, es interesante notar que hay dos componentes en el efecto restricción, por un lado, la diferencia en las posturas, y la proporción del error que le da un valor a la postura dentro del espacio admisible.

En la figura 9, a 1 lo favorecen ambas componentes, primero, el tener una postura mayor y luego tener una mayor cota, sin embargo este segundo efecto tiene la particularidad de que a medida de que aumenta μ su peso va disminuyendo, dado que los errores que se van generando son más pequeños y por ende una mayor fracción de los errores generados están creando posturas admisibles para el agente dos,

disminuyendo el efecto restricción. Finalmente cuando ya μ pasa a ser muy grande, los errores generados son tan pequeños que su peso es prácticamente nulo y la diferencia en posturas determina que las probabilidades de ambos modelos coincidan en que la probabilidad del agente uno tiende a uno como ocurre en q_1 .

En el caso $B_2 > B_1$, los efectos se contraponen, las posturas favorecen al agente dos, sin embargo, cuando μ se empieza a alejar del cero, el agente que tiene mayor probabilidad de localizarse es dos, dado que las posturas están formadas en su mayor grado como parte aleatoria que determinística en este tramo y dado que el agente uno tiene mayor espacio para que sus errores generen posturas admisibles. A medida que crece μ , el error se empieza a hacer más insignificante, por lo que se invierte la jerarquía en un punto y el agente dos pasa a tener mayor probabilidad de ganar el remate. En el límite, ocurre el mismo fenómeno que en los otros dos casos, el error es tan pequeño que las probabilidades terminan siendo gobernadas por la postura determinística.

3.5 Identificabilidad de parámetros.

Uno de los objetivos de este trabajo es revisar la capacidad de estimar la capacidad de estimar los parámetros de la disposición a pagar utilizando el nuevo modelo. Para entender los parámetros identificables en el modelo con restricción, se propone primero estudiar los parámetros identificables con el modelo irrestricto.

3.5.1 Identificabilidad en el modelo logit irrestricto.

En el modelo Logit Multinomial las posturas son de la forma $B_i = \bar{B}_i + \varepsilon_i$ donde \bar{B} es la componente determinística y ε es una variable aleatoria que distribuye Gumbel con parámetro μ . Como fue mostrado en el capítulo 2.4, la probabilidad de que un postor haga una postura mayor a la de sus competidores es, de forma simplificada, la siguiente:

$$q_i = \frac{\exp(\mu\bar{B}_i)}{\sum_j \exp(\mu\bar{B}_j)} \quad \forall j \quad (71)$$

Ahora bien, si se reduce el problema a dos postores, la probabilidad de que cada uno de ellos gane el remate queda de la siguiente forma:

$$q_1 = \frac{\exp(\mu\bar{B}_1)}{\exp(\mu\bar{B}_1) + \exp(\mu\bar{B}_2)} \quad y \quad q_2 = \frac{\exp(\mu\bar{B}_2)}{\exp(\mu\bar{B}_1) + \exp(\mu\bar{B}_2)} \quad (72)$$

Estas expresiones pueden ser reescritas de modo de que queden en función de la diferencia de las posturas.

$$q_1 = \frac{1}{1 + \exp(-\mu(\bar{B}_1 - \bar{B}_2))} \quad (73)$$

$$P_2 = \frac{1}{1 + \exp(\mu(\bar{B}_1 - \bar{B}_2))} \quad (74)$$

Si se consideran las funciones de postura lineales:

$$B_1 = \bar{B}_1 + \varepsilon_1 = \beta_{01} + \beta_{11} x + \varepsilon_1 \quad (75)$$

$$B_2 = \bar{B}_2 + \varepsilon_2 = \beta_{02} + \beta_{12} x + \varepsilon_2 \quad (76)$$

Reemplazando en las expresiones de probabilidad en función de la diferencia de las posturas:

$$\begin{aligned} q_1 &= \frac{1}{1 + \exp(-\mu(\beta_{01} + \beta_{11} x - (\beta_{02} + \beta_{12} x)))} \\ &= \frac{1}{1 + \exp(-\mu((\beta_{01} - \beta_{02}) + (\beta_{11} - \beta_{12})x))} \end{aligned} \quad (77)$$

Del mismo modo para q_2 :

$$q_2 = \frac{1}{1 + \exp(\mu((\beta_{01} - \beta_{02}) + (\beta_{11} - \beta_{12})x))} \quad (78)$$

Los parámetros que se pueden estimar de esta manera son tres, por lo que se propone es fijar en cero el valor de β_{02} .

$$B_1 = \beta_{01} + \beta_{11} x + \varepsilon_1 \quad (79)$$

$$B_2 = \beta_{12} x + \varepsilon_2 \quad (80)$$

Reemplazando en la expresión de las probabilidades:

$$q_1 = \frac{1}{1 + \exp(-\mu(\beta_{01} + (\beta_{11} - \beta_{12})x))} \quad (81)$$

$$q_2 = \frac{1}{1 + \exp(\mu(\beta_{01} + (\beta_{11} - \beta_{12})x))} \quad (82)$$

De aquí se desprende que sólo son identificables $\mu\beta_{01}$, $\mu\beta_{11}$ y $\mu\beta_{12}$.

3.5.2 Identificabilidad en el modelo logit restringido.

En v_1 , si se consideran i y j como 1 y 2 respectivamente:

$$P_1 = 1 - \frac{F(I_2 - \bar{B}_1)}{F(I_1 - \bar{B}_1)} * q_2 \quad (83)$$

Donde q_2 es la probabilidad logit de que el agente dos gane el remate, en el modelo irrestricto.

En extensión, P_1 se puede escribir de la siguiente manera:

$$\begin{aligned} P_1 &= 1 - \frac{\exp(-\exp(\mu I_2 - \mu \bar{B}_1))}{\exp(-\exp(\mu I_1 - \mu \bar{B}_1))} * \frac{1}{1 + \exp(\mu \Delta \bar{B}_{12})} \\ &= 1 - \frac{\exp(-\exp(\mu(I_2 - \bar{B}_1)))}{\exp(-\exp(\mu(I_1 - \bar{B}_1)))} * \frac{1}{1 + \exp(\mu \Delta \bar{B}_{12})} \end{aligned} \quad (84)$$

Utilizando v_2 para P_1 es:

$$\begin{aligned} P_1 &= 1 - \exp(-\exp(\mu \bar{B}_1 - \mu I_1))^{\exp(\mu \Delta I_{12})-1} * \frac{1}{1 + \exp(\mu \Delta \bar{B}_{12})} \\ &= 1 - \exp(-\exp(\mu(\bar{B}_1 - I_1)))^{\exp(\mu \Delta I_{12})-1} * \frac{1}{1 + \exp(\mu \Delta \bar{B}_{12})} \end{aligned} \quad (85)$$

Usando la forma lineal para las posturas (75) (76) y reemplazando en las tres versiones de la probabilidad:

$$v_1: P_1 = 1 - \frac{\exp(-\exp(\mu(\beta_{01} - I_2) + \mu\beta_{11}x))}{\exp(-\exp(\mu(\beta_{01} - I_1) + \mu\beta_{11}x))} * \frac{1}{1 + \exp(\mu\Delta\beta_0 + \mu\Delta\beta_1x)} \quad (86)$$

$$v_2: P_1 = 1 - \frac{\exp(-\exp(\mu(\beta_{01} - I_1) + \mu\beta_{11}x))^{\exp(\mu\Delta I_{12})-1}}{1 + \exp(\mu\Delta\beta_0 + \mu\Delta\beta_1x)} \quad (87)$$

$$v_3: P_1 = 1 - \frac{\exp(-\exp(\mu(\beta_{02} - I_2) + \mu\beta_{12}x))^{\exp(\mu\Delta\beta_0 + \mu\Delta\beta_1x)(1-\exp(-\mu\Delta I_{12}))}}{1 + \exp(\mu\Delta\beta_0 + \mu\Delta\beta_1x)} \quad (88)$$

De estas tres expresiones, es posible ver que hay términos identificables así como lo había en el modelo irrestricto, sin embargo se puede llegar a mayor resolución de estos si se escriben por extensión los términos que aparecen como diferencias:

$$v_1: P_1 = 1 - \frac{\exp(-\exp(\mu(\beta_{01} - I_2) + \mu\beta_{11}x))}{\exp(-\exp(\mu(\beta_{01} - I_1) + \mu\beta_{11}x))} * \frac{1}{1 + \exp(\mu(\beta_{01} - \beta_{02}) + \mu(\beta_{11} - \beta_{12})x)} \quad (89)$$

$$v_2: P_1 = 1 - \frac{\exp(-\exp(\mu(\beta_{01} - I_1) + \mu\beta_{11}x))^{\exp(\mu(I_1 - I_2)) - 1}}{1 + \exp(\mu(\beta_{01} - \beta_{02}) + \mu(\beta_{11} - \beta_{12})x)} \quad (90)$$

$$v_3: P_1 = 1 - \frac{\exp(-\exp(\mu(\beta_{02} - I_2) + \mu\beta_{12}x))^{\exp(\mu(\beta_{01} - \beta_{02}) + \mu(\beta_{11} - \beta_{12})x)(1 - \exp(-\mu(I_1 - I_2)))}}{1 + \exp(\mu(\beta_{01} - \beta_{02}) + \mu(\beta_{11} - \beta_{12})x)} \quad (91)$$

De acá se puede desprender que hay cinco factores estimables por medio de la máxima verosimilitud para cada una de las versiones. Estos son:

Para v_1 :

1. $\mu(\beta_{01} - I_2)$
2. $\mu\beta_{11}$
3. $\mu(\beta_{01} - I_1)$
4. $\mu(\beta_{01} - \beta_{02})$
5. $\mu(\beta_{11} - \beta_{12})$

Para v_2 :

1. $\mu(\beta_{01} - I_1)$
2. $\mu\beta_{11}$
3. $\mu(I_1 - I_2)$
4. $\mu(\beta_{01} - \beta_{02})$
5. $\mu(\beta_{11} - \beta_{12})$

Para v_3 :

1. $\mu(\beta_{02} - I_2)$
2. $\mu\beta_{12}$
3. $\mu(\beta_{01} - \beta_{02})$
4. $\mu(\beta_{11} - \beta_{12})$
5. $\mu(I_1 - I_2)$

En las tres formulaciones aparentemente se estiman parámetros distintos, lo cual sugeriría que si se estimaran por separado se podría en conjunto estimar más de los cinco parámetros de cada caso. Sin embargo, esto no es posible, por ejemplo, se toma el primer y segundo grupo. Se observa que el parámetro 1. 2. 4. Y 5. Ya están estimados en el primero y el tercer parámetro se deduce por la resta entre 1. y 3. del primer grupo.

En este modelo sería interesante poder estimar siete parámetros ($\beta_{01}, \beta_{11}, \beta_{02}, \beta_{12}, I_1, I_2$ y μ). Recordar que I_1 e I_2 son conocidos en el modelo teórico, sin embargo, en la práctica podrían ser desconocidos debido a la dificultad de observarlos y sería útil que con el modelo se puedan estimar. En cada una de las formas de escribir la probabilidad P_1 , se tienen cinco ecuaciones. Lo que implica que no se pueden obtener todos los valores de manera simultánea. De modo que, si se quieren realizar estimaciones, hay que tener en consideración la información con la que se cuenta y, según eso, estimar.

Tomando las ecuaciones derivadas de las probabilidades P_1 y P_2 , en sus tres versiones, es interesante ver qué se puede estimar dependiendo el escenario al cuál uno se enfrenta. He aquí dos escenarios interesantes.

I_1 e I_2 conocidos

En el caso en que I_1 e I_2 conocidos, es posible identificar los cinco parámetros restantes ($\beta_{01}, \beta_{11}, \beta_{02}, \beta_{12}$ y μ). Esto proviene de despejar cada uno de los parámetros a partir de las ecuaciones. Notar que con el MNL sólo son identificables $\mu\beta_{01}, \mu\beta_{11}$ y $\mu\beta_{12}$.

$$\beta_{02} = 0$$

Si se toma una especificación en que β_{02} es igual a cero, como con el MNL, es posible identificar $\mu\beta_{01}$, $\mu\beta_{11}$, $\mu\beta_{12}$, μI_1 y μI_2 .

Tabla 1: Parámetros identificables con el modelo restringido.

Caso	Parámetros identificables
I1 e I2 conocidos	β_{01} , β_{02} , β_{11} , β_{12} y μ
$\beta_{02} = 0$, conocido	$\mu\beta_{01}$, $\mu\beta_{11}$, $\mu\beta_{12}$, μI_1 y μI_2

4 SIMULACIÓN.

Se propone un nuevo modelo de probabilidades de elección discreta que incorpora la restricción presupuestaria. Estructuralmente este modelo es más complejo que el Multinomial Logit (MNL) porque incorpora nuevas componentes no lineales en su formulación y nuevos parámetros que pueden ser estimados.

En el capítulo anterior se analizó este modelo en sus características matemáticas y se requiere conocer distintos aspectos del modelo en relación a su estimación econométrica. Por ejemplo, el método de estimación basado en el criterio de máxima verosimilitud, la factibilidad de ser estimado, los parámetros que se pueden estimar, los datos requeridos para la estimación y, ciertamente, casos reales donde pueda ser aplicado.

Como paso inicial, en esta sección se analiza, se propone y se implementa un método de estimación de este modelo y con bases de datos sintéticas o simuladas, se realizan distintas pruebas con el objetivo de analizar la factibilidad de estimar este modelo con distintos escenarios de parámetros conocidos o desconocidos y características de los atributos.

Este primer análisis empírico permitirá entender la factibilidad de utilizar este modelo en casos reales y comparar este proceso de estimación con el tradicionalmente ocupado MNL sin restricciones.

4.1 El simulador

Con el fin de comprobar empíricamente la estimación de funciones de disposición a pagar utilizando el modelo propuesto se utiliza un simulador que consiste en un programa escrito para correr en el software estadístico Gauss u otro software estadístico, que permite utilizar el método de máxima verosimilitud que permite estimar econométricamente parámetros de disposiciones a pagar.

Los datos necesarios para utilizar el método de la máxima verosimilitud son: una base de datos con atributos y una base de datos con observaciones de localización. En este estudio no se cuenta con dichas bases de datos, por lo que se ha dispuesto crear bases sintéticas para poner a prueba el modelo.

4.1.1 Bases de datos sintéticas.

Utilizar bases de datos sintéticas para probar el modelo tiene como beneficio la variedad de escenarios que se permite proponer y la posibilidad de saber de antemano los valores de los parámetros que uno pretende estimar, de modo de hacer una validación del modelo.

La primera base de datos que se construye es aquella con los atributos. Como las funciones de disposición a pagar que se estimarán tienen un atributo, basta sólo simular uno. Para esto es necesario ingresar un archivo con las características del atributo que se busca generar, estas son media \bar{x} y desviación estándar σ . Cada una de las observaciones del atributo x_i , con i el número de observación, se genera utilizando una distribución

Normal (\bar{x}, σ). Gauss no permite generar variables aleatorias que distribuyen *Normal* (\bar{x}, σ), por este motivo es necesario generar realizaciones de una variable aleatoria que distribuye *Normal* (0,1) (Gauss cuenta con la función que genera valores distribuidos de esta forma) y modificarla de la siguiente manera:

$$x_i = \bar{x} + \sigma z_i \text{ con } z_i \sim N(0,1) \quad (92)$$

La segunda base que se necesita para poder estimar con los modelos, es la de elecciones. Ésta se genera a partir de la base de datos de atributos, los parámetros poblacionales de las funciones de disposición a pagar de los agentes ($\beta_{01}, \beta_{11}, \beta_{02}$ y β_{12}), las cotas de los agentes (I_1 e I_2) y parámetro μ de la distribución de los errores Gumbel. Estos parámetros son definidos por el modelador.

Para generar las elecciones simuladas, el programa lee los parámetros un archivo de entrada y simula los errores Gumbel para cada postor en cada una de las observaciones. En vista que Gauss no cuenta con una función que entregue realizaciones de números distribuidos como una Gumbel, se utilizó un método que emplea la distribución *Uniforme* (0,1) para poder generar distribuciones de variables aleatorias. Para esto, es necesario despejar la función de probabilidad acumulada de la Gumbel ($F(x) = \exp(-\exp(-\mu x))$).

$$F(x) = \exp(-\exp(-\mu x)) \Leftrightarrow \frac{-\ln(-\ln(F(x)))}{\mu} = x \quad (93)$$

Luego, si se reemplaza $F(x)$ por una variable aleatoria y_i que distribuye *Uniforme* (0,1), la función entrega una variable aleatoria que distribuye *Gumbel* (μ):

$$x_i = \frac{-\ln(-\ln(y_i))}{\mu}, \text{ si } y_i \sim \text{Uniforme}(0,1) \text{ entonces } x_i \sim \text{Gumbel}(\mu) \quad (94)$$

Así, entonces, se crean los errores ε_{i1} y ε_{i2} para cada agente uno y dos respectivamente en la observación i .

Luego el programa calcula la parte determinística \bar{B}_{i1} y \bar{B}_{i2} utilizando los datos de los parámetros de las funciones de postura de entrada y la base de datos de los atributos generada anteriormente:

$$\begin{aligned} \bar{B}_{i1} &= \beta_{01} + \beta_{11}x_i \\ \bar{B}_{i2} &= \beta_{02} + \beta_{12}x_i \end{aligned} \quad (95)$$

Tras realizar esta operación se suma el error correspondiente a cada postura para armar las posturas finales:

$$B_{i1} = \bar{B}_{i1} + \varepsilon_{i1} = \beta_{01} + \beta_{11}x_i + \varepsilon_{i1} \quad (96)$$

$$B_{i2} = \bar{B}_{i2} + \varepsilon_{i2} = \beta_{02} + \beta_{12}x_i + \varepsilon_{i2}$$

En seguida se hace un filtro de las observaciones generadas, dejando dentro sólo aquellas que no violan su cota, es decir, se revisa una a una las posturas de los agentes y se las compara con su respectiva cota. Si alguna de las dos posturas viola la restricción presupuestaria, entonces, se elimina dicha fila de las observaciones. El supuesto detrás de este proceso es que en el mercado los agentes sólo pueden realizar posturas que sean menores o iguales a su restricción presupuestaria.

Luego de eliminar posturas inadmisibles, se arma el vector de elecciones comparando cada postura B_{i1} con su respectiva contraparte B_{i2} .

En la figura 11 se muestra esquemáticamente como distribuyen las frecuencias de las posturas generadas de manera sintética.

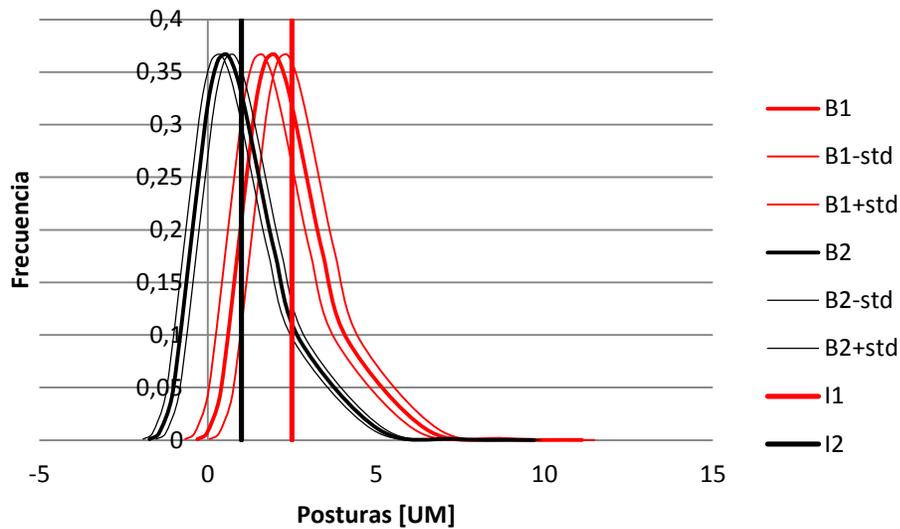


Figura 11: Distribución de frecuencias de posturas simuladas con $\mu=1$.

Las posturas que son consideradas válidas son aquellas que se encuentran a la izquierda de su propia cota. Notar que hay tres curvas para cada tipo de agente. Esto se debe a que las frecuencias de las posturas estarán dentro de la franja que se dibuja entre $B_i - std$ y $B_i + std$ para un agente del tipo i , con una probabilidad de 0,99 si se considera que los atributos distribuyen *Normal* (\bar{x}, σ) a una distancia de menos cuatro (-4σ) y cuatro desviaciones estándar de la media (4σ) .

4.1.2 Estimación.

Para la estimación de los parámetros se utiliza la rutina MaxLik programada en Gauss, que ejecuta el método de máxima verosimilitud. Como se explica en 2.5, el método utiliza elecciones (en el caso de esta memoria son aquellas generadas con el simulador) y probabilidades de elección. Dependiendo del modelo que se quiere utilizar, se le entrega la expresión de las probabilidades restringidas o irrestringidas en función de los parámetros a estimar.

De la rutina MaxLik se recupera el tiempo que demoró en procesar el computador, el número de iteraciones en que se llegó a los valores y para cada parámetro estimado, su valor, su error estándar y su t-valor.

4.2 Escenarios estimados.

En esta sección se muestran los resultados de utilizar el simulador y la rutina Maxlik en distintos escenarios exploratorios de estimación de funciones de disposición a pagar utilizando el modelo propuesto. Se comparará el desempeño de utilizar dicho modelo en contraste al modelo Logit Binomial irrestringido y se muestra los resultados de estimar distintos juegos de parámetros de ambos modelos de probabilidad. Será interesante revisar la estimación de ciertos parámetros del nuevo modelo que no pueden identificarse en el modelo Logit Binomial tradicional.

Las estimaciones se llevaron a cabo en un computador portátil que cuenta con un procesador Intel Core i5-4200U @1.6GHz 2.30 GHz, 4GB de RAM y sistema operativo Windows 8.1 de 64 bit.

4.2.1 Escenario 1.

Este escenario fue diseñado para estimar funciones de disposición a pagar utilizando los modelos Logit Binomial irrestringido y el modelo propuesto y luego comparar los resultados obtenidos. Recordar que el modelo propuesto tiene la capacidad de estimar más parámetros, sin embargo lo que se busca con este escenario es mostrar que con el modelo propuesto también es posible estimar los parámetros que con el modelo irrestringido. Las funciones que se estiman son las siguientes:

$$B_1 = \beta_{01} + \beta_{11} x \quad (97)$$

$$B_2 = \beta_{02} + \beta_{12} x \quad (98)$$

Sin embargo, utilizando el modelo Logit irrestringido no es posible estimar estos cuatro parámetros, sólo tres (ver 3.5.1) y multiplicados por el parámetro μ . Para poder comparar el desempeño de las estimaciones utilizando ambos modelos, se estimarán los mismos parámetros con cada uno de ellos.

Los parámetros que se estimarán, entonces, son $\mu\beta_{01}$, $\mu\beta_{11}$ y $\mu\beta_{12}$. Se tomarán como conocidas las cotas de los agentes (I_1 e I_2) y el valor $\beta_{02} = 0$.

Las características del experimento son las siguientes: se generan 100.000 valores para el atributo x que ingresan en las funciones de disposición a pagar, luego utilizando el método de selección descrito en 4.1.1 se consolida una base de datos con posturas de tipo uno y tipo dos, utilizando los atributos antes generados. De esa población se obtienen los datos para estimar las funciones. Se toman muestras de 500, 1.000, 2.000, 5.000, 10.000 y 20.000 observaciones. Con cada una de dichas muestras se estiman las funciones utilizando el modelo irrestricto y restricto.

A continuación, en la tabla 2, se muestran los parámetros considerados en el escenario 1 para construir las posturas.

Tabla 2: Parámetros del escenario 1.

Escenario 1										
Parámetro	\bar{x}	σ	c_v	μ	β_{01}	β_{11}	β_{02}	β_{12}	I_1	I_2
Valor	0,7	0,07	0,1	1	0,1	1,3	0	0,7	1,3	0,8

En la tabla 2, \bar{x} es el valor de la media de los datos generados, σ su desviación estándar. El valor c_v corresponde al coeficiente de variación, calculado como σ/\bar{x} . A continuación se listan los parámetros poblacionales de las funciones de disposición a pagar (β_{01} , β_{11} , β_{02} y β_{12}) y las cotas superiores I_1 e I_2 asociadas a las restricciones presupuestarias de cada uno de los agentes. Se utilizó el valor de $\mu = 1$ para simplificar el análisis de los resultados. La elección de los valores de los parámetros del escenario 1 no es trivial, dado que se busca construir posturas que no sean ni muy cercanas ni muy lejanas.

En la tabla 3 se presentan los resultados de las estimaciones bajo el escenario 1, cambiando el número de observaciones de 500 a 20.000. En ella se presentan los valores de los parámetros que se pueden estimar con el modelo Logit Multinomial irrestricto, estimados con dicho modelo y el modelo restricto. Se entregan los parámetros poblacionales (Pob.), valores estimados por cada modelo, t-valor (con respecto al parámetro poblacional), número de iteraciones que tomó el método para converger a los parámetros y tiempo en minutos que se demoró en procesar el conjunto de iteraciones.

Tabla 3: Resultados de estimaciones bajo el escenario 1 variando el tamaño de la muestra.

Número de observaciones	500					1000			
Modelo	Irrestringido			Restringido		Irrestringido		Restringido	
Parámetros	Pob.	Estimado	t-valor	Estimado	t-valor	Estimado	t-valor	Estimado	t-valor
$\mu\beta_{01}$	0,1	-0,10	0,26	-0,23	0,41	0,69	0,80	1,18	0,38
$\mu\beta_{11}$	1,3	1,73	0,55	1,59	0,30	1,13	0,33	1,10	0,02
$\mu\beta_{12}$	0,7	-0,23	1,20	0,29	1,22	0,37	0,62	4,80	0,74
Iteraciones		14		9		13		49	
Tiempo de CPU [min]		0,00157		0,00135		0,00167		0,00665	
Número de observaciones	2000					5000			
Modelo	Irrestringido			Restringido		Irrestringido		Restringido	
Parámetros	Pob.	Estimado	t-valor	Estimado	t-valor	Estimado	t-valor	Estimado	t-valor
$\mu\beta_{01}$	0,1	0,55	13,91	0,26	0,41	0,53	21,63	0,23	0,56
$\mu\beta_{11}$	1,3	1,24	2,15	0,99	0,70	1,23	4,32	0,99	1,36
$\mu\beta_{12}$	0,7	0,26	16,71	0,50	1,23	0,27	25,10	0,51	0,98
Iteraciones		5		10		4		8	
Tiempo de CPU [min]		0,0008		0,00183		0,00083		0,00228	
Número de observaciones	10000					20000			
Modelo	Irrestringido			Restringido		Irrestringido		Restringido	
Parámetros	Pob.	Estimado	t-valor	Estimado	t-valor	Estimado	t-valor	Estimado	t-valor
$\mu\beta_{01}$	0,1	0,32	0,76	0,09	0,07	0,24	0,94	0,02	0,45
$\mu\beta_{11}$	1,3	1,40	0,44	1,17	0,68	1,47	1,66	1,26	0,20
$\mu\beta_{12}$	0,7	0,10	2,77	0,43	3,82	0,03	6,55	0,40	4,04
Iteraciones		13		7		13		7	
Tiempo de CPU [min]		0,00373		0,00302		0,00592		0,00532	

Sobre la reproducción de los parámetros, se observa que utilizando ambos modelos, a medida que aumenta el número de observaciones, la diferencia entre las estimaciones y los valores poblacionales se reducen. En los casos de muestras de entre 500 y 5000 observaciones todos los parámetros estimados con el modelo restringido son estadísticamente similares a los valores poblacionales, presentando t-valores menores a 1,96 a diferencia del modelo irrestringido. Recordar que cuando el t-valor es mayor que 1,96, hay evidencia estadística de que el valor del parámetro estimado es distinto al poblacional con un 95% de certeza y cuando es menor que 1,96, no se rechaza la hipótesis de que son iguales. En los casos en que la muestra tiene 10.000 y 20.000 observaciones, en ambos modelos se aprecia que $\mu\beta_{01}$ y $\mu\beta_{11}$ presentan t-valores inferiores a 1,96. En el parámetro $\mu\beta_{12}$ si bien para ambos modelos se rechaza la

hipótesis de que los valores son iguales a los parámetros poblacionales, los valores de las estimaciones con el modelo restricto con más cercanos a los parámetros poblacionales.

En la figura 12 se ilustra que, a medida que se aumentan las observaciones de la muestra el número de iteraciones que toma llegar a los parámetros estimados se mantiene relativamente constante bajo el enfoque de los dos modelos. El tiempo de convergencia a los parámetros tiene una tendencia a aumentar en la medida que la muestra tiene más observaciones para ambos modelos (tabla 13), presentando, una pequeña superioridad en velocidad el modelo restricto que el modelo irrestricto. Vale la pena mencionar la estimación de tamaño de muestra igual a 1.000, caso en que el modelo restricto presenta una convergencia significativamente más lenta y en un mayor número de iteraciones. Debiese tenerse presente para investigaciones futuras la causalidad tras este fenómeno que quiebra el patrón que seguían las estimaciones.

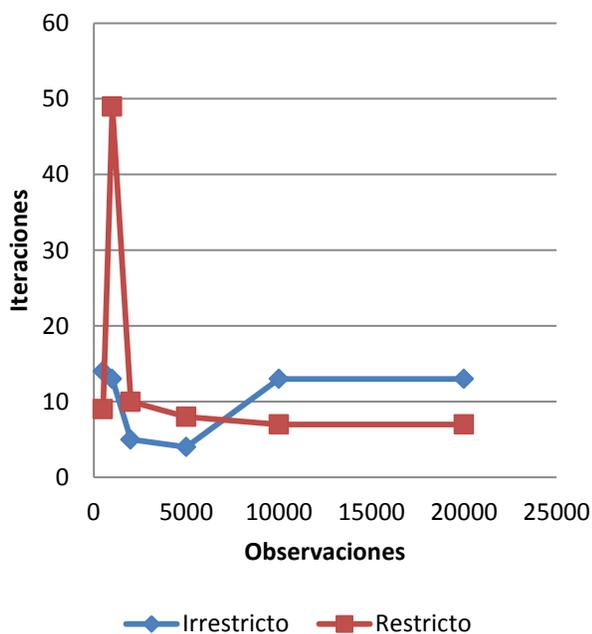


Figura 12: Numero de iteraciones versus número de observaciones.

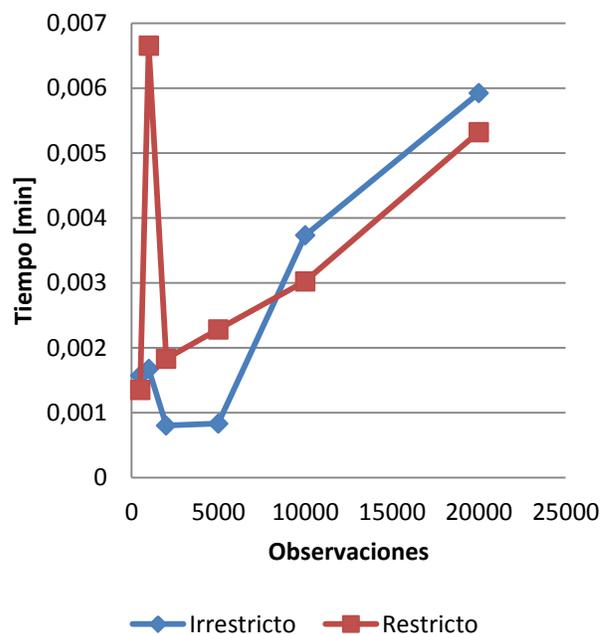


Figura 13: Tiempo de convergencia versus número de observaciones.

Este primer resultado muestra que es posible la estimación de funciones de disposición a pagar de manera similar que utilizando el modelo Logit Binomial.

4.2.2 Escenario 2.

Este escenario fue diseñado para estimar funciones de disposición a pagar y el parámetro μ de la distribución Gumbel utilizando el modelo restricto propuesto en esta memoria.

Se estiman los cuatro parámetros de las funciones de disposición a pagar ($\beta_{01}, \beta_{11}, \beta_{02}$ y β_{12}) y el parámetro μ , dejando exógeno valor de las cotas de los agentes. Notar que la

estimación de estos parámetros por separado no es posible hacerla en el caso del modelo Binomial clásico.

El experimento se diseñó, distinto al escenario 1, para estimar los parámetros y tres valores distintos de μ , generándose por tanto tres conjuntos distintos de elecciones, uno con cada valor de μ . En la tabla 4 se encuentran los parámetros del escenario 2.

Tabla 4: Parámetros del escenario 2.

Escenario 2													
Parámetro	\bar{x}	σ	c_v	Número de observaciones	β_{01}	β_{11}	β_{02}	β_{12}	l_1	l_2	μ_1	μ_2	μ_3
Valor	0,7	0,07	0,1	20000	0,2	0,8	0,1	0,6	1,3	1	0,7	0,9	1,1

Los valores de μ elegidos varían solamente en dos decimales entre sí, sin embargo, esta variación tiene efectos significativos en las bases de elección que se usan porque la varianza del error aumenta cuando disminuye el valor del parámetro μ . En cada una de las estimaciones se utilizó una muestra de 20.000 observaciones de elección, usando este gran tamaño muestral debido a que favorece la reproducción de los parámetros originales. Hay un cambio en el valor de una de las cotas y de los parámetros poblacionales con respecto al escenario 1; esto se debe a la búsqueda de una combinación de parámetros que dejara al menos 20.000 observaciones admisibles al aplicar la selección.

En la parte superior de la tabla 5 se presentan los valores de μ y parámetros poblacionales para cada caso, acompañados de sendos parámetros estimados y su t-valor. En la parte media se presenta el valor de μ multiplicando por los parámetros poblacionales.

Tabla 5: Resultados de estimaciones bajo el escenario 2 variando el parámetro μ

Parámetros	Pob.	Estimado	t-valor	Pob.	Estimado	t-valor	Pob.	Estimado	t-valor
μ	0,7	0,7890	0,9366	0,9	0,9088	0,14549	1,1	1,0150	1,40189
β_{01}	0,2	0,3706	1,7587	0,2	0,3229	1,75117	0,2	0,2420	0,70161
β_{11}	0,8	0,7283	0,5704	0,8	0,8821	0,91071	0,8	1,0300	2,77402
β_{02}	0,1	0,2811	1,8705	0,1	0,3434	3,57486	0,1	0,4307	5,42961
β_{12}	0,6	0,6614	0,9125	0,6	0,5790	0,3626	0,6	0,4959	2,432
Iteraciones	13			13			11		
Tiempo de CPU [min]	0,01383			0,01382			0,0109		

Se observa que casi todos los parámetros tienen t-valores inferiores al valor crítico de 1,96. Es importante notar que las estimaciones del parámetro μ fueron muy cercanas a los valores poblacionales y con t-valores menores a 1,96.

4.2.3 Escenario 3

El diseño de este escenario pretende explorar la capacidad del nuevo modelo de estimar las cotas de cada uno de los agentes.

Como se presenta en 3.5.2 para estimar las cotas es necesario fijar uno de los parámetros de las funciones de disposición a pagar. Se fija el valor de β_{02} en cero. Estas condiciones permiten identificar tres parámetros y las cotas, pero no el parámetro μ , por lo que los parámetros a estimar son: $\mu\beta_{01}$, $\mu\beta_{11}$, $\mu\beta_{12}$, μI_1 y μI_2 .

Tabla 6: Parámetros del escenario 3.

Escenario 3											
Media	σ	c_v	μ	Número de observaciones	β_{01}	β_{11}	β_{02}	β_{12}	I_1	I_2	
0,7	0,07	0,1	1	20000	0,1	0,9	0	0,4	1	0,7	

Por simplicidad del análisis de los resultados se eligió como valor de μ el número uno. Notar que los parámetros poblacionales son distintos que en los escenarios anteriores, como así lo son las cotas.

Tabla 7: Resultados de la estimación exploratoria bajo el escenario 3.

Parámetros	Pob.	Estimado	t-valor
μ		1	
$\mu\beta_{01}$	0,1	0,33	1,35
$\mu\beta_{11}$	0,9	0,99	0,92
$\mu\beta_{12}$	0,4	0,51	2,36
μI_1	1	0,75	10,84
μI_2	0,7	0,66	1,61
Iteraciones		13	
Tiempo de CPU [min]		0,01058	

El valor de los parámetros estimados es relativamente cercano al de los parámetros poblacionales, presentando la diferencia mayor en el parámetro μI_1 . A pesar de que se rechaza la hipótesis de que la estimación de los parámetros μI_1 y $\mu\beta_{12}$ es similar al valor poblacional, estas diferencias son relativamente pequeñas.

El número de iteraciones se mantiene similar al de los escenarios anteriores, de igual modo que el tiempo de procesamiento.

A la luz de los resultados se observa que el modelo propuesto permite estimar, sin mucha diferencia con el modelo irrestricto, parámetros a través del método de máxima verosimilitud. Además permite identificar parámetros que con el modelo Logit Binomial clásico era imposible de hacer, cómo el parámetro μ .

5 Conclusiones

En esta memoria se dedujo un modelo probabilístico y se realizaron estimaciones exploratorias para probar su validez.

Con este modelo se define una expresión explícita y rigurosa que representa la elección considerando la restricción presupuestaria.

Para hacer frente a este problema, se realizó una deducción general de las probabilidades de adjudicarse un remate al mejor postor con dos agentes, sujeto a que cada uno de los dos agentes tiene una restricción presupuestaria que acota superiormente su postura, bajo un enfoque en que las posturas presentan un error i. i. d. Gumbel. Es decir, se modeló la probabilidad de realizar una postura superior dado que la postura de cada agente no puede superar su restricción presupuestaria.

El nuevo modelo muestra que los postores que tienen una restricción presupuestaria más fuerte, es decir, que cuentan con un presupuesto menor para realizar posturas en el remate, ven afectadas negativamente sus probabilidades de ganar dicho remate. El modelo Logit Binomial irrestricto, en este caso, sobreestimaría la probabilidad de que el agente con menos presupuesto se adjudique el remate.

El modelo propuesto, es una extensión del modelo que no presenta restricciones (Logit Binomial irrestricto). Y cuando las cotas son iguales o éstas tienden a infinito se recuperan los resultados del modelo de probabilidades irrestrictas.

Se estudió el efecto sobre las probabilidades de modelar las posturas acotadas superiormente y se caracteriza el efecto restricción.

También se mostró que, en general, existe una diferencia en las posturas puede compensar la diferencia en presupuestos de los agentes al momento de estimar las elecciones.

Con el modelo restricto se pueden estimar los mismos parámetros que se pueden estimar con el modelo Logit Binomial irrestricto. Los tiempos que toma estimar funciones de disposición a pagar utilizando este modelo son similares que con el modelo clásico.

El nuevo modelo puede estimar más parámetros que el modelo Logit Binomial irrestricto. Cuando se conocen las cotas, es capaz de estimar todos los parámetros de las funciones de disposición a pagar y es posible estimar, a diferencia del modelo clásico, el parámetro μ de la distribución Gumbel de los errores de las posturas. Si se fija un parámetro en cero de las disposiciones a pagar, es posible estimar las cotas además del resto de los parámetros de las funciones de disposición a pagar, eso sí todas ellas multiplicadas por el parámetro μ .

Hay una serie de ventajas que provienen de poder estimar μ y los parámetros poblacionales por separado. Con respecto a μ , indica que es posible caracterizar totalmente a los errores, en particular su varianza. Al conocer los parámetros poblacionales es podrán estimar las valoraciones marginales de los atributos.

Se generaron estimaciones de los parámetros para ciertos valores de parámetros de escenario. Como línea de trabajo se puede desarrollar estimaciones para una mayor variación de parámetros de escenario, motivar la estimación de modelos con data real y desarrollo del modelo multinomial.

Bibliografía

Alonso, W. (1964). Location and land use. Toward a general theory of land rent. *Location and land use. Toward a general theory of land rent*.

Anas, A. (1982). Residential location markets and urban transportation: Economic theory, econometrics, and policy analysis with discrete choice models. Nueva York: Academic Press.

Ben-Akiva, M. y Boccara, B. (1995). Discrete choice models with latent choice sets. *International Journal of Research in Marketing*, 12(1), 9-24.

Birch, S., Gafni, A., & O'Brien, B. (1999). Willingness to pay and the valuation of programmes for the prevention and control of influenza. *Pharmacoeconomics*, 16(1), 55-61.

Cascetta, E. y Papola, A. (2001). Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand. *Transportation Research Part C: Emerging Technologies*, 9(4), 249-263.

Castro, M., Martínez, F., & Munizaga, M. A. (2013). Estimation of a constrained multinomial logit model. *Transportation*, 40(3), 563-581.

Cantillo, V. y Ortúzar, J. D. D. (2005). A semi-compensatory discrete choice model with explicit attribute thresholds of perception. *Transportation Research Part B: Methodological*, 39(7), 641-657.

Ellickson, B. (1981). An alternative test of the hedonic theory of housing markets. *Journal of Urban Economics*, 9(1), 56-79.

Follain, J. R., y Jimenez, E. (1985). Estimating the demand for housing characteristics: a survey and critique. *Regional science and urban economics*, 15(1), 77-107.

Gil, J. M., Gracia, A., y Sanchez, M. (2000). Market segmentation and willingness to pay for organic products in Spain. *The International Food and Agribusiness Management Review*, 3(2), 207-226.

Gross, D. J. (1988). Estimating willingness to pay for housing characteristics: an application of the Ellickson bid-rent model. *Journal of Urban Economics*, 24(1), 95-112.

Hurtubia, R., y Bierlaire, M. (2012). Estimation of bid functions for location choice and price modeling with a latent variable approach. *Networks and Spatial Economics*, 1-19.

Jara-Díaz, S. R. y Martínez, F. J. (1999). On the specification of indirect utility and willingness to pay for discrete residential location models. *Journal of Regional Science*, 39(4), 675-688.

Lerman, S. R. y Kern, C. R. (1983). Hedonic theory, bid rents, and willingness-to-pay: Some extensions of Ellickson's results. *Journal of Urban Economics*, 13(3), 358-363.

- Manski, C. F., y Lerman, S. R. (1977). The estimation of choice probabilities from choice based samples. *Econometrica: Journal of the Econometric Society*, 1977-1988.
- Martinez, F. J. (1992). The bid-choice land-use model: an integrated economic framework. *Environment and Planning A*, 24(6), 871-885.
- Martinez, F. (1996). MUSSA: land use model for Santiago city. *Transportation Research Record: Journal of the Transportation Research Board*, 1552(1), 126-134.
- Martínez, F. J., Aguila, F. y Hurtubia, R. (2009). The constrained multinomial logit: A semi-compensatory choice model. *Transportation Research Part B: Methodological*, 43(3), 365-377.
- McFadden, D. (1978). *Modelling the choice of residential location* (pp. 75-96). Institute of Transportation Studies, University of California.
- McFadden, D. (1980). Econometric models for probabilistic choice among products. *Journal of Business*, S13-S29.
- McFadden, D. (1998). Measuring willingness-to-pay for transportation improvements. *Theoretical Foundations of Travel Choice Modeling*, 339, 364.
- Ortúzar, J. de D. y Willumsen, L.G. (2011), *Modelling Transport*, 4th Edition, Wiley.
- Roe, B., Teisl, M. F., Levy, A., y Russell, M. (2001). US consumers' willingness to pay for green electricity. *Energy policy*, 29(11), 917-925.
- Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *The journal of political economy*, 34-55.
- Sillano, M., y Ortúzar, J. de D. (2005). Willingness-to-pay estimation with mixed logit models: some new evidence. *Environment and Planning A*, 37(3), 525-550.
- Swait, J., y Ben-Akiva, M. (1987). Incorporating random constraints in discrete models of choice set generation. *Transportation Research Part B: Methodological*, 21(2), 91-102.
- Swait, J. (2001). Choice set generation within the generalized extreme value family of discrete choice models. *Transportation Research Part B: Methodological*, 35(7), 643-666.
- Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological review*, 79(4), 281.

Anexos

Anexo A: Proposición 1

$$\sum_i P(A_i/B) = 1$$

Demostración

$$\sum_i P(A_i/B) = \sum_i \frac{P(A_i/B)}{P(B)} = \frac{P(\cup (A_i \cap B))}{P(B)} = \frac{P((\cup A_i) \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

Anexo B: Proposición 2

Sean ε_i errores i.i.d Gumbel($0, \mu$), con $\mu > 0$, cuya función de densidad es f y función de distribución es:

$$F(x) = \exp(-\exp(-\mu x))$$

Esta función cumple con la siguiente propiedad:

$$F(x + a) = F(x) \exp(-\mu a)$$

Entonces:

$$\int_{-\infty}^b F(z + \Delta_{ij}) f(z) dz = F(b + \Delta_{ij}) F(b) \frac{\exp(\mu \bar{B}_i)}{\exp(\mu \bar{B}_i) + \exp(\mu \bar{B}_j)}$$

Demostración

Si se aplica el cambio de variable $t = F(z)$ se deduce que:

$$\int_{-\infty}^b F(z + \Delta_{ij}) f(z) dz = \int_0^{F(b)} F(F^{-1}(t) + \Delta_{ij}) dt$$

La función F tiene la siguiente propiedad:

$$F(x + a) = F(x)^{\exp(-\mu a)}$$

Al aplicar esta propiedad se deduce:

$$\int_{-\infty}^b F(z + \Delta_{ij})f(z)dz = \int_0^{F(b)} t^{\exp(-\mu\Delta_{ij})} dt = \frac{F(b)^{\exp(-\mu\Delta_{ij})+1}}{\exp(-\mu\Delta_{ij}) + 1}$$

Al multiplicar y dividir el cociente por $\exp(\mu\bar{B}_i)$ y aplicar la propiedad se deduce:

$$\int_{-\infty}^b F(z + \Delta_{ij})f(z)dz = F(b + \Delta_{ij})F(b) \frac{\exp(\mu\bar{B}_i)}{\exp(\mu\bar{B}_i) + \exp(\mu\bar{B}_j)}$$