

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

ESCUELA DE GRADUADOS



ESTIMACIÓN BAYESIANA DE EFECTOS DE RED: EL  
MODELO LOGIT MIXTO

TESIS PARA OPTAR POR EL GRADO DE MAGISTER EN  
ESTADÍSTICA

Presentado por:

Paulo Roberto Chahuara Vargas

Asesor: Cristian Luis Bayes Rodriguez

Miembros del jurado:

Dr. Oscar Edgardo Millones Destefano

Dr. Cristian Luis Bayes Rodriguez

Dr. Luis Hilmer Valdivieso Serrano

Lima, Julio 2017

## Dedicatoria

A mi familia y a la leve piedad del tiempo.



## Agradecimientos

A mi asesor, el Dr. Cristian Bayes, por su valiosa guía y consejos durante el proceso de este trabajo.

Al Dr. Valdivieso y el Dr. Millones por sus imprescindibles sugerencias.

También quisiera expresar mi agradecimiento a los profesores de la Maestría de Estadística de la Pontificia Universidad Católica del Perú (PUCP): Valdivieso, Bayes, Flores, Tarazona, Veliz, Giancarlo, Millones y Camiz, quienes con sus enseñanzas y su buena disposición han contribuido en mi proceso continuo de formación, no solo a nivel del aprendizaje técnico, sino también como persona. Mi más profundo respeto y admiración hacia ustedes.



## Resumen

Los efectos o externalidades de red son factores que pueden condicionar las decisiones de contratación de los consumidores en favor de empresas ya establecidas y en contra de los nuevos competidores, pudiendo limitar la competencia efectiva y potencial de los mercados, en especial, en aquellas industrias donde el número de empresas es bajo y la entrada de nuevos competidores es poco frecuente. Por ello, es importante verificar su existencia y la magnitud de sus efectos sobre las decisiones de compra de los consumidores con el objetivo de justificar o establecer medidas que impulsen una competencia más equilibrada entre las empresas. Además, teniendo en consideración que los consumidores pueden tener cierto grado de heterogeneidad en sus comportamientos de adquisición, también resulta relevante estudiar el grado de diferenciación de los efectos de red entre los consumidores a fin de mejorar las políticas que fomenten la competencia.

Este trabajo tiene por objetivo estimar un modelo logit mixto bajo el enfoque de la inferencia bayesiana, para estudiar empíricamente la existencia y heterogeneidad de los efectos de red sobre las decisiones de contratación de los consumidores en la industria de telefonía móvil peruana. El análisis se hace con base a una muestra que combina información de la Encuesta Residencial de Servicios de Telecomunicaciones (ERESTEL) del año 2015 e información de las empresas operadoras del servicio de telefonía móvil.

Los resultados de las estimaciones realizadas sugieren que los efectos de red tendrían un condicionamiento importante sobre las decisiones de contratación del servicio de telefonía móvil, además de presentar un grado de heterogeneidad estadísticamente significativo en la magnitud de sus efectos.

**Palabras clave:** modelo logit mixto, inferencia bayesiana, efectos de red.

# Índice general

<b>Índice de figuras</b>	<b>VI</b>
<b>Índice de cuadros</b>	<b>VII</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Consideraciones Preliminares	1
1.2. Objetivos	3
1.3. Organización del Trabajo	4
<b>2. Modelos de Elección Discreta</b>	<b>5</b>
2.1. Utilidad y regla de comportamiento	5
2.2. Derivación de las probabilidades de elección	6
<b>3. Modelo Logit Estándar</b>	<b>7</b>
3.1. Especificando una función densidad para el término de error	7
3.2. Probabilidades de elección	8
<b>4. Modelo Logit Mixto (MLX)</b>	<b>10</b>
4.1. Especificación del Comportamiento	10
4.2. Probabilidades de elección	10
4.3. Aproximación Bayesiana para el MLX	11
<b>5. Aplicación</b>	<b>17</b>
5.1. Base de datos y medición de las variables de investigación	17
5.2. Resultados de la investigación	19
<b>6. Conclusiones</b>	<b>25</b>
6.1. Conclusiones	25
6.2. Sugerencias para investigaciones futuras	26
<b>A. Código</b>	<b>28</b>
<b>Bibliografía</b>	<b>31</b>

## Índice de figuras

5.1. Distribución de $\beta_n^{CN}$ en la muestra . . . . .	21
5.2. Distribución de $\beta_n^P$ en la muestra . . . . .	22
5.3. Distribución de las correlaciones del vector de medias $b$ en la muestra simulada . . . . .	22
5.4. Distribución de las correlaciones de las desviaciones estándar de la matriz de covarianzas $W$ en la muestra simulada . . . . .	23
5.5. Distribución de la cadena del vector de medias $b$ en la muestra simulada . . . . .	23
5.6. Distribución de la cadena de las desviaciones estándar de la matriz de covarianzas $W$ en la muestra simulada . . . . .	24



## Índice de cuadros

5.1. Medición de las Variables introducidas en el Modelo . . . . .	18
5.2. Resultados de la Estimación . . . . .	20



# Capítulo 1

## Introducción

### 1.1. Consideraciones Preliminares

Los efectos o externalidades de red pueden definirse como aquellos efectos que hacen que el valor que obtiene una persona por consumir un producto, o hacer uso de un servicio, dependa no solo de la utilidad del producto o servicio en sí mismo (valor intrínscico), sino del número de individuos que lo consumen o utilizan (valor de sincronización). Alternativamente, se dice que las preferencias de los consumidores muestran externalidades de red si la utilidad de cada consumidor se eleva cuando el número total de consumidores se incrementa (Farrell y Klemperer, 2007; Shy, 2001). Por ejemplo, el servicio de telefonía móvil ofrecido por una empresa operadora resultará más útil o beneficioso para una persona a medida que esta compañía tenga más clientes, ya que de este modo, el individuo no solo podrá comunicarse con un número de usuarios mayor, sino que esta comunicación puede resultar menos costosa para el consumidor: una mayor base de suscriptores en una compañía móvil disminuye los costos unitarios de la empresa, además de permitirle menores costos relacionados a la interconexión de llamadas que terminan en las redes de otras compañías móviles, lo cual se traduce en ofertas comerciales más atractivas.

Los efectos de red pueden condicionar de forma importante la decisión del consumidor para elegir al proveedor de un bien o servicio, pues la presencia de externalidades de red implica que en los compradores emerjan deseos de afinidad con el resto de consumidores, y, en consecuencia, las personas tomarán sus decisiones de elección de proveedor condicionados por la aspiración de ser compatibles con lo que ya han elegido la mayoría de usuarios en el mercado, o en algunos casos con las elecciones ya realizadas por los integrantes de su red social (familiares, amigos, compañeros o colegas). Esto genera una situación eficaz para que las empresas establezcan relaciones con sus consumidores encaminadas a una muy larga duración (e.g. Czajkowski y Sobolewski, 2015; Maicas et al., 2009, Ocoña et al., 2009 y 2010, Birke y Swann, 2006, Corrocher y Zirulia, 2009, Maicas y Sese, 2011).

Si bien los efectos de red pueden resultar beneficiosos para las firmas, su presencia puede representar un desequilibrio para la competencia entre empresas. Esto es particularmente sustancial, en las industrias donde el número de competidores es limitado y la oportunidad que aparezcan nuevas empresas que reten a las compañías establecidas es poco frecuente. Bajo este contexto, la oportunidad de que el consumidor cuente con más opciones de provee-



dores en el mercado puede perderse ya que las externalidades de red se vuelven una ventaja sostenible para las compañías con más tiempo o participación en el mercado a la hora competir por atraer o retener clientes frente a las nuevas empresas retadoras o de menor cuota en la industria (Fu, 2004). Como resultado, el mercado no puede desconcentrarse de las empresas ya establecidas y termina o continua en manos de pocos proveedores, degradándose la competencia (Farrell y Klemperer, 2007; Katz y Shapiro, 1985 y 1994). Ello les permite a las compañías fijar precios que resultan superiores a los precios de un mercado sin externalidades de red, limitando -por los altos precios- que más personas puedan hacer uso del bien o servicio y disminuyendo el excedente del consumidor (la diferencia entre lo que está dispuesto a pagar como máximo un cliente y lo que realmente termina pagando). Además, dada la fragilidad o ausencia de competencia, estas pocas empresas tienen el incentivo de fijar la calidad del producto en un nivel que no es óptimo para la sociedad e inclusive darse el lujo de operar con costos altos, generando una industria ineficiente. En suma, una reducción del bienestar social.

Por lo expuesto, existe un importante interés y preocupación sobre el rol de las externalidades de red de parte de las instituciones públicas encargadas de diseñar y establecer las políticas de competencia y regulación que arbitrarán el juego competitivo de los mercados. En esta línea, el desarrollo e implementación eficaz de tales políticas requiere no solo un análisis que defina e identifique para cada industria la manera en que los efectos de red se manifiestan y operan, sino que también cuantifique su magnitud y distribución entre los agentes económicos. Sin embargo, no son muchos los estudios de naturaleza empírica que analizan la relevancia de las externalidades de red sobre la elección de proveedor en los consumidores. Particularmente, se han encontrado estudios de esta índole a nivel de Estados Unidos, Europa o Asia. Lo que es más, dentro de este grupo de investigaciones, parte de la perspectiva adoptada ha sido cuantificar los efectos de red a nivel promedio o agregado (Karakuka et al., 2013; Maicas et al., 2010; Doganoglu y Grzybowski, 2007; Birke y Swann, 2006; Kim y Kwon, 2003), obviando que puede existir un grado de diferenciación en el comportamiento de cada consumidor y que estas heterogeneidades podrían influir sobre la medición de las externalidades de red, lo cual convierte al estudio de estos elementos restrictivos de la competencia en un tema importante para los hacederos de políticas públicas y privadas, los agentes económicos y la comunidad científica.

La presente investigación tiene como finalidad utilizar un modelo estadístico para estimar y analizar la relevancia y heterogeneidad de los efectos de red sobre la elección de proveedor que realizan los consumidores. El trabajo toma como ámbito de aplicación la industria de telefonía móvil peruana debido a que, en primer lugar, reúne las características de ser una industria con pocas empresas compitiendo y en donde la mayoría de clientes está concentrada en una operadora, lo que incrementa las posibilidades del daño que pueden tener las externalidades de red sobre la competencia. En segundo lugar, es un sector que tiene una importancia fundamental en el bienestar social del país (e.g. Beuermann et. al., 2012). En tercer lugar, tiene instituciones públicas tales como el Ministerio de Transporte y Comunicaciones (MTC) y el Organismo Supervisor de la Inversión Privada en Telecomunicaciones

(OSIPTEL) especialmente interesados en el desarrollo del sector. Por ejemplo, el OSIPTEL es la autoridad regulatoria y en los últimos años ha estado aplicando diversas políticas que han buscado mejorar la competencia efectiva entre las empresas (e.g., cargos de terminación móvil diferenciados). En esta línea, una cuestión que siempre ha estado presente es si unos cargos (precios) de terminación de llamadas elevados pueden ser utilizados por los operadores ya establecidos para expulsar del mercado a nuevos retadores, situación que se ve reforzada por las externalidades de red. Así, cuando las llamadas que terminan en la red de otro operador resultan más caras que las llamadas que terminan en la misma red del operador, la red con más suscriptores se vuelve más atractiva para los usuarios, ya que si se suscriben a esta red una mayor proporción de sus llamadas serán de la misma red y, por tanto, incurrirán en un menor costo de comunicación. Por el contrario, ser cliente de una red pequeña supone un costo para los usuarios porque reciben un menor número de llamadas, además de incurrir en mayores gastos por comunicación, ya que la mayoría de sus llamadas terminarán en la red del operador más grande. Por último, a diferencia de otras industrias reguladas en el Perú (saneamiento, electricidad, gas natural o transporte), los usuarios de telefonía móvil tienen la facilidad técnica o legal de elegir con qué compañía contratar su servicio.

Para contrastar lo señalado en el párrafo anterior, se dispone de una base de datos que combina información sobre los atributos de las diferentes empresas operadoras de telefonía móvil con información sobre las decisiones de contratación de usuarios del servicio de telefonía móvil. La metodología empleada en la presente tesis sigue la adoptada por el grupo pequeño de investigaciones que estudiaron la magnitud y heterogeneidad de los efectos de red a nivel individual (Czajkowski y Sobolewski, 2015; Sobolewski y Czajkowski, 2012; Maicas et al., 2009; Polo y Sesé, 2009). Dicha metodología, se basa en la utilización del modelo logit mixto estimado bajo inferencia bayesiana. Esto debido a las ventajas del procedimiento bayesiano frente a la estimación clásica. Por ejemplo, el procedimiento bayesiano no requiere la maximización de una función de verosimilitud, a diferencia de la inferencia clásica, lo que expone a la estimación clásica a problemas de convergencia por valores iniciales pobres, óptimos locales o problemas de curvaturas de la función de verosimilitud que pueden derivar en resultados engañosos (Train, 2009; Balcombe et al., 2009).

## 1.2. Objetivos

A partir de lo señalado en la sección anterior, el presente trabajo tiene el objetivo general de estimar el modelo logit mixto bajo el enfoque de la inferencia bayesiana, para evaluar la existencia y heterogeneidad de los efectos o externalidades de red sobre la elección de proveedor que realizan los consumidores en la industria peruana de servicio de telefonía móvil.

Los objetivos específicos de la tesis son los siguientes:

- Estudiar el modelo de elección discreta logit mixto bajo inferencia bayesiana.
- Estimar la influencia de los efectos de red sobre las decisiones de los consumidores para elegir un operador de telefonía móvil.

- Evaluar si esta influencia presenta una heterogeneidad importante entre los consumidores.
- Evaluar si estos efectos son estadísticamente significativos.

### 1.3. Organización del Trabajo

En el Capítulo 2, se presenta el marco general de los modelos de elección bajo el enfoque de la teoría de la utilidad aleatoria. Luego, en el Capítulo 3 se muestra la versión básica del modelo logit mixto, el modelo logit estándar, para seguidamente en el Capítulo 4 centrarse concretamente en el modelo logit mixto y su estimación bajo el enfoque de la inferencia bayesiana. Después, en el Capítulo 5 se utiliza el modelo logit mixto para obtener el efecto estimado de las externalidades de red sobre el proceso de elección de compañía de telefonía móvil que el consumidor peruano realiza, y se estudia la heterogeneidad de dichos efectos. Finalmente, en el Capítulo 6 se expone las conclusiones obtenidas del trabajo. Adicionalmente, en el anexo de la tesis, se presenta el código que fue programado en el *software* estadístico R para realizar la estimación (Apéndice A).



## Capítulo 2

# Modelos de Elección Discreta

De acuerdo a los trabajos previos sobre la estimación de efectos de red, el punto de partida de la estimación son los modelos de elección discreta (Thurstone, 1927; Marschak, 1960; McFadden, 1974). Dichos modelos se derivan generalmente bajo el enfoque de la utilidad aleatoria, cuyo supuesto angular es la maximización de la utilidad en el comportamiento del tomador de decisiones. Bajo este enfoque cada alternativa u opción del conjunto de elección le reporta al decisor un nivel de utilidad. La utilidad es el beneficio o bienestar percibido de elegir una alternativa del conjunto de elección, y el decisor se comporta de forma tal que escoje la alternativa que le reporta el mayor nivel de utilidad.

A continuación se describe el modelo de elección discreta desde la optica de la teoría de la utilidad aleatoria. Es importante señalar que esta sección se basa en gran medida en Train (2009).

### 2.1. Utilidad y regla de comportamiento

Se asume que una persona  $n = 1, \dots, N$ , enfrenta en cada período  $t = 1, \dots, T$ , la decisión de elegir un operador de telefonía móvil  $j = 1, \dots, J$ . En cada situación de elección  $t$ , la persona  $n$  asocia a cada alternativa  $j$  cierto nivel de utilidad denotado como  $U_{njt} \forall j, t$ . La regla de comportamiento, en base al enfoque de la utilidad aleatoria, es que el decisor eligirá la alternativa que le brinda el máximo nivel de bienestar o utilidad en el período de decisión  $t$ . Esto es, elegir la alternativa  $y_{nt}$ , si y sólo si:  $U_{ny_{nt}} > U_{njt}, \forall j \neq y_{nt}$ .

Sin embargo, desde el punto de vista del investigador, no es posible observar la utilidad del decisor en cada situación de elección. El investigador solo puede observar un vector  $1 \times K$  no estocástico de atributos que describen a la empresa operadora  $j$  tal y como se presentan al decisor  $n$  en el período  $t$ . Este vector se denota como  $x_{njt} \forall j, t$  y es importante señalar que también puede incluir variables que caracterizan al decisor. Asimismo, el investigador puede especificar en cada período una función que relacione las variables observadas con la utilidad que percibe el decisor. Esta función se denota como  $V_{njt} = V(x_{njt}), \forall j, t$ , y suele llamarse utilidad representativa (Train, 2009).

## 2.2. Derivación de las probabilidades de elección

Dado que existen aspectos de la utilidad que el investigador no observa o no puede observar,  $V_{njt} \neq U_{njt}$ , se puede descomponer la utilidad como  $U_{njt} = V_{njt} + \varepsilon_{njt}$ , donde  $\varepsilon_{njt}$  captura los factores que afectan a la utilidad, pero no están incluidos en la parte determinística  $V_{njt}$ . Según Train (2009), esta descomposición es totalmente general, ya que  $\varepsilon_{njt}$  se define simplemente como la diferencia entre la verdadera utilidad  $U_{njt}$  y la parte de la utilidad que el investigador captura en  $V_{njt}$ . El investigador no conoce  $\varepsilon_{njt} \forall j, t$ , y por lo tanto trata estos términos como variables aleatorias. Dichas variables suelen ser llamados términos de error. La densidad conjunta del vector de errores en el período  $t$ ,  $\varepsilon_{nt} = (\varepsilon_{n1t}, \dots, \varepsilon_{nJt})$ , es etiquetada como  $f(\varepsilon_{nt})$ . Conociendo esta densidad, el investigador puede hacer afirmaciones probabilísticas acerca de la elección de proveedor de servicio de telefonía móvil que toma el decisor. Así, la probabilidad que el decisor  $n$  elija en el periodo de decisión  $t$  a la alternativa  $y_{nt}$  es

$$\begin{aligned} P_{nt}(y_{nt}) &= P(U_{ny_{nt}} > U_{njt}, \forall j \neq y_{nt}) \\ &= P(V_{ny_{nt}} + \varepsilon_{ny_{nt}} > V_{njt} + \varepsilon_{njt}, \forall j \neq y_{nt}) \\ &= P(\varepsilon_{njt} - \varepsilon_{ny_{nt}} < V_{ny_{nt}} - V_{njt}, \forall j \neq y_{nt}) \end{aligned}$$

Esta probabilidad es una distribución acumulativa, es decir, es la probabilidad de que cada término aleatorio  $\varepsilon_{njt} - \varepsilon_{ny_{nt}}$  esté por debajo de la cantidad observada  $V_{ny_{nt}} - V_{njt}$ . Usando la densidad  $f(\varepsilon_{nt})$  la probabilidad acumulada de elegir  $y_{nt}$  en la situación de elección  $t$  puede ser escrita como

$$\begin{aligned} P_{nt}(y_{nt}) &= P(\varepsilon_{njt} - \varepsilon_{ny_{nt}} < V_{ny_{nt}} - V_{njt}, \forall j \neq y_{nt}) \\ &= \int_{\varepsilon_{nt}} I(\varepsilon_{njt} - \varepsilon_{ny_{nt}} < V_{ny_{nt}} - V_{njt}, \forall j \neq y_{nt}) f(\varepsilon_{nt}) d\varepsilon_{nt} \end{aligned}$$

donde  $I(\cdot)$  es una función indicadora, igual a 1 cuando la expresión entre paréntesis es verdadera y cero en caso contrario. Esta expresión es una integral multidimensional, sobre la densidad de probabilidad de la parte no observada de la utilidad,  $f(\varepsilon_{nt})$ . Diferentes modelos de elección discreta se obtienen mediante especificaciones diferentes de esta densidad, es decir, a partir de diferentes supuestos acerca de cómo se distribuye la densidad de probabilidad de la parte no observada de la utilidad. La integral tiene una forma cerrada sólo para ciertas especificaciones de  $f(\cdot)$ . En particular, el modelo logit estándar tienen una expresión cerrada para esta integral.



## Capítulo 3

### Modelo Logit Estándar

La presente sección describe el modelo de logit estándar siguiendo en su mayor parte la exposición que realiza McFadden (1974) y Train (2009) sobre este modelo.

#### 3.1. Especificando una función densidad para el término de error

El modelo logit estándar se obtiene suponiendo que cada  $\varepsilon_{njt}$  se distribuye independientemente y de forma idénticamente distribuida respecto a  $n$ ,  $j$  y sobre todo,  $t$ , de acuerdo a una densidad de probabilidad valor extremo tipo I con parámetro de localización igual a cero y parámetro de escala igual a 1. Dicha densidad, también es conocida como la distribución estándar de Gumbel.

Así, se define la densidad de cada componente no observado de la utilidad en la situación de elección  $t$  como

$$f(\varepsilon_{njt}) = \exp(-\varepsilon_{njt})\exp(-\exp(-\varepsilon_{njt})),$$

cuya distribución acumulada es

$$F(\varepsilon_{njt}) = \exp(-\exp(-\varepsilon_{njt}))$$

Considerar la distribución de valor extremo para los errores es similar a asumir que los errores se distribuyen normalmente y de forma independiente. La distribución de valor extremo tiene colas ligeramente más gruesas que una distribución normal, lo que implica que permite un comportamiento ligeramente más aberrante que la normal. Por lo general, sin embargo, la diferencia entre errores distribuidos según el valor extremo y según distribuciones normales independientes es indistinguible empíricamente (Train, 2009). De hecho, el supuesto clave del modelo no es tanto la forma de la distribución como que los errores sean independientes entre sí. La suposición de que los errores sean independientes unos de otros es muy importante y podría ser visto como restrictiva, ya que implica que el investigador tiene especificado la utilidad representativa ( $V_{njt}$ ) con tal grado de precisión que la parte restante de la utilidad ( $\varepsilon_{njt}$ ) es esencialmente ruido: toda la información necesaria en el proceso de decisión es capturada en la forma analítica de  $V_{njt}$ , haciendo que el error de una alternativa no proporciona al investigador ninguna información sobre el error de otra alternativa diferente.

### 3.2. Probabilidades de elección

Siguiendo la aproximación de McFadden (1974), a continuación se derivan las probabilidades de elección logit. La probabilidad que el usuario  $n$  elija en el periodo de decisión  $t$  la empresa de telefonía móvil  $y_{nt}$  es

$$\begin{aligned} P_{nt}(y_{nt}) &= P(V_{ny_{nt}t} + \varepsilon_{ny_{nt}t} > V_{njt} + \varepsilon_{njt}, \forall j \neq y_{nt}) \\ &= P(\varepsilon_{njt} < \varepsilon_{ny_{nt}t} + V_{ny_{nt}t} - V_{njt}, \forall j \neq y_{nt}) \end{aligned}$$

Si  $\varepsilon_{ny_{nt}t}$  esta dado, la distribución acumulada para cada  $\varepsilon_{njt}$  evaluada en  $\varepsilon_{ny_{nt}t} + V_{ny_{nt}t} - V_{njt}$  es

$$F_{\varepsilon_{njt}}(\varepsilon_{ny_{nt}t} + V_{ny_{nt}t} - V_{njt}) = \exp(-\exp(-(\varepsilon_{ny_{nt}t} + V_{ny_{nt}t} - V_{njt})))$$

Se etiqueta a  $P_{nt}(y_{nt}|\varepsilon_{ny_{nt}t})$  como el valor de la probabilidad de elección  $P_{nt}(y_{nt})$  dado el valor de  $\varepsilon_{ny_{nt}t}$ . Ya que los  $\varepsilon$ s son independientes entre alternativas, la probabilidad sobre todo  $j \neq y_{nt}$  es el producto de todas las distribuciones acumuladas individuales:

$$P_{nt}(y_{nt}|\varepsilon_{ny_{nt}t}) = \prod_{j \neq y_{nt}} \exp(-\exp(-(\varepsilon_{ny_{nt}t} + V_{ny_{nt}t} - V_{njt})))$$

En vista de que  $\varepsilon_{ny_{nt}t}$  en realidad no está dado, la probabilidad de elección  $P_{nt}(y_{nt})$  es la integral de  $P_{nt}(y_{nt}|\varepsilon_{ny_{nt}t})$  sobre todos los valores de  $\varepsilon_{ny_{nt}t}$  ponderados por la densidad descrita para  $f(\varepsilon_{ny_{nt}t})$ :

$$P_{nt}(y_{nt}) = \int \left( \prod_{j \neq y_{nt}} e^{-e^{-(\varepsilon_{ny_{nt}t} + V_{ny_{nt}t} - V_{njt})}} \right) e^{-\varepsilon_{ny_{nt}t}} e^{-e^{-\varepsilon_{ny_{nt}t}}} d\varepsilon_{ny_{nt}t}$$

que puede ser reescrito como

$$P_{nt}(y_{nt}) = \int_{s=-\infty}^{+\infty} \left( \prod_{j \neq y_{nt}} e^{-e^{-(s+V_{ny_{nt}t}-V_{njt})}} \right) e^{-s} e^{-e^{-s}} ds$$

donde  $s = \varepsilon_{ny_{nt}t}$ . Además, se debe notar que  $V_{ny_{nt}t} - V_{ny_{nt}t} = 0$ . Por lo que, agrupando términos en el exponente de  $e$ , se tiene

$$\begin{aligned} P_{nt}(y_{nt}) &= \int_{s=-\infty}^{+\infty} \left( \prod_j e^{-e^{-(s+V_{ny_{nt}t}-V_{njt})}} \right) e^{-s} ds \\ &= \int_{s=-\infty}^{+\infty} \exp \left( - \sum_j e^{-(s+V_{ny_{nt}t}-V_{njt})} \right) e^{-s} ds \\ &= \int_{s=-\infty}^{+\infty} \exp \left( -e^{-s} \sum_j e^{-(V_{ny_{nt}t}-V_{njt})} \right) e^{-s} ds \end{aligned}$$

Reescribiendo  $\exp(-s) = z$ , con  $-\exp(-s)ds = dz$ :

$$\begin{aligned}
P_{nt}(y_{nt}) &= \int_{-\infty}^0 \exp\left(-z \sum_j e^{-(V_{ny_{nt}t} - V_{njt})}\right) (-dz) \\
&= \int_0^{\infty} \exp\left(-z \sum_j e^{-(V_{ny_{nt}t} - V_{njt})}\right) dz \\
&= \frac{\exp\left(-z \sum_j e^{-(V_{ny_{nt}t} - V_{njt})}\right)}{-\sum_j e^{-(V_{ny_{nt}t} - V_{njt})}} \Big|_0^{\infty} \\
&= \frac{1}{\sum_j e^{-(V_{ny_{nt}t} - V_{njt})}} \\
&= \frac{e^{V_{ny_{nt}t}}}{\sum_j e^{V_{njt}}}
\end{aligned}$$

La utilidad representativa suele especificarse de forma que sea lineal en relación a los parámetros:  $V_{njt} = \beta' x_{njt}$ , donde  $x_{njt}$  es el vector no estocástico observable mencionado en el Capítulo 2 y, que, dado el planteamiento de la presente tesis, incluirá la variable independiente proxy de los efectos de red. Con esta especificación, la probabilidad de que el consumidor  $n$  elija a la empresa operadora  $y_{nt}$  se convierte en

$$P_{nt}(y_{nt}) = \frac{e^{\beta' x_{ny_{nt}t}}}{\sum_j e^{\beta' x_{njt}}}$$

que es la probabilidad de elección del modelo logit estándar para el periodo de decisión  $t$ . El hecho de que las probabilidades de elección se expresan en una forma cerrada es una de las mayores ventajas de logit frente a otros modelos de elección discreta, como el probit. En este sentido, las probabilidades de elección logit son rápidamente calculadas y esto es beneficioso cuando se realiza estimaciones basadas en simulación. Además, McFadden (1974) demostró que la función logaritmo de la verosimilitud (log-verosimilitud o *log-likelihood*) con estas probabilidades de elección, es globalmente cóncava respecto a los parámetros  $\beta$ .

Ahora bien, la última probabilidad también se puede reexpresar como

$$\prod_{y_{nt}=1}^J (P_{nt}(y_{nt}))^{d_{y_{nt}}}$$

donde  $d_{y_{nt}} = 1$  si el consumidor  $n$  eligió la opción  $y_{nt}$  en el periodo  $t$  y cero en caso contrario. Luego, si se tiene una muestra de  $N$  usuarios cuyas elecciones de proveedor de telefonía móvil fueron observadas durante  $T$  períodos, la función de verosimilitud de la muestra o - dicho de otro modo - la probabilidad de que cada persona de la muestra haya elegido la alternativa que realmente se ha observado que eligió, dado que los  $\varepsilon$ 's también son independientes entre decisores y períodos de elección, es

$$\prod_{n=1}^N \prod_{t=1}^T \prod_{y_{nt}=1}^J [P_{nt}(y_{nt})]^{d_{y_{nt}}}$$



## Capítulo 4

### Modelo Logit Mixto (MLX)

La siguiente sección describe el modelo de comportamiento MLX y el enfoque bayesiano para estimar el MLX. Es importante señalar que esta sección se basa en gran parte en McFadden y Train (2000), Train (2001 y 2009), Huber y Train (2001), y Train y Sonnier (2005).

#### 4.1. Especificación del Comportamiento

El modelo logit mixto parte de reespecificar la utilidad que recibe el consumidor  $n$  de la empresa operadora  $j$  en el período de elección  $t$ , de la siguiente forma:

$$U_{njt} = \beta_n' x_{njt} + \varepsilon_{njt},$$

donde nuevamente  $x_{njt}$  es un vector  $1 \times K$  no estocástico de variables observables que incluyen la covariable de los efectos de red, o, en términos generales,  $x_{njt}$  contiene variables independientes relacionadas al proveedor de servicio de telefonía móvil  $j$  o al decisor  $n$  en el tiempo  $t$ . Por su parte,  $\varepsilon_{njt}$  es el término de error cuya función de densidad es de valor extremo tipo I los cuales se asumen son independientes e idénticamente distribuidos. Sin embargo, ahora  $\beta_n$  es un vector de coeficientes que representa la influencia específica de  $x_{njt}$  sobre las preferencias del usuario  $n$ . Así, estos coeficientes varían entre decisores y el investigador puede especificar una densidad para  $\beta_n$  que describa la distribución de estos parámetros en la población, con un vector de medias  $b$  y matriz de covarianzas  $W$ . Con ello, esta especificación es igual a la de un modelo logit estándar, excepto que los  $\beta$ s varían entre consumidores en lugar de ser fijos, y hace que el MLX se caracterize como un modelo de coeficientes aleatorios. Seguidamente, se plantea que el decisor conoce el valor de su propia  $\beta_n$  y de las  $\varepsilon_{njt}$  para cualquier opción  $j$ , y elige la alternativa  $y_{nt}$  siempre que  $U_{ny_{nt}} > U_{njt}, \forall j \neq y_{nt}$ .

#### 4.2. Probabilidades de elección

El investigador observa las  $x_{njt}$  pero no los  $\beta_n$ 's o los  $\varepsilon_{njt}$ 's. Si el investigador observase las  $\beta_n$ 's, entonces la probabilidad de elección en el periodo  $t$  sería la del modelo logit estándar, ya que los  $\varepsilon_{njt}$ s son valor extremo independientes e idénticamente distribuidos. Es decir, la probabilidad de elegir a  $y_{nt}$  condicionada sobre  $\beta_n$  es

$$L(y_{nt}|\beta_n) = \frac{e^{\beta_n' x_{ny_{nt}}t}}{\sum_{j=1}^J e^{\beta_n' x_{njt}}}$$

Si definimos la secuencia de elecciones concretas de la persona  $n$  como  $y_n = \{y_{n1}, \dots, y_{nT}\}$ , la probabilidad de las alternativas elegidas por el consumidor  $n$ , condicionado a  $\beta_n$ , es

$$L(y_n|\beta_n) = \prod_{t=1}^T \frac{e^{\beta_n' x_{nynt}}}{\sum_{j=1}^J e^{\beta_n' x_{njt}}}$$

No obstante, el investigador desconoce los  $\beta_n$ s y por ende no puede condicionar sobre  $\beta_n$ . En este sentido, la probabilidad de elección incondicionada es la integral de  $L(y_n|\beta_n)$  sobre todos los posibles valores de  $\beta_n$ :

$$P_n(y_n|b, W) = \int L(y_n|\beta_n) f(\beta_n|b, W) d\beta_n$$

que es la probabilidad de elección del modelo logit mixto.

En la literatura económica o de marketing, la distribución de los parámetros  $f(\beta_n|b, W)$  se suele especificar como Normal:  $\beta \sim N(b, W)$  (Hall et al., 2006; Johnson et al., 2000; King et al., 2007; Lancsar et al., 2007; Negrín et al., 2008). Empero, otros trabajos han utilizado la distribución log normal,  $\ln\beta \sim N(b, W)$  (Train, 1998), la distribución uniforme (Revelt y Train, 2000), distribución triangular (Revelt y Train, 2000), distribución SB de Johnson (Train y Sonnier, 2005), y una distribución normal censurado por debajo en cero (Bhat, 2000). Los coeficientes también se pueden especificar como fijos; es decir, que no varían en la población.

Por su parte, la función de verosimilitud para la muestra de consumidores de tamaño  $N$ , dada la independencia de los errores entre decisores, es

$$\prod_{n=1}^N P_n(y_n|b, W)$$

### 4.3. Aproximación Bayesiana para el MLX

Dado los fines del presente trabajo, más que la estimación de las probabilidad de elección  $P_n$ , se busca la estimación de la distribución posteriori de  $b$  y  $W$ , ya que la cuantificación de  $b$  y  $W$  es sustancial para analizar la importancia y la heterogeneidad, respectivamente, de los efectos de red sobre la elección de compañía de telefonía móvil.

La distribución posteriori de  $b$  y  $W$ , se denota como  $K(\cdot)$  y es por definición proporcional a la distribución a priori de  $b$  y  $W$ ,  $k(b, W)$ , multiplicado por la la función de verosimilitud del modelo logit mixto:

$$K(b, W|Y) \propto \prod_{n=1}^N P_n(y_n|b, W) k(b, W)$$

donde  $Y = (y_1, \dots, y_N)$  representa las elecciones de los decisores de toda la muestra.

Resulta posible simular valores de  $K(b, W|Y)$  utilizando el algoritmo Metropolis-Hastings

(MH). Sin embargo, por cada iteración de este algoritmo, sería necesario calcular  $P_n(y_n|b, W)$  que contiene una integral no cerrada que también debe ser simulada. Para sortear esta restricción, el algoritmo MH debe combinarse con el algoritmo de muestreo de Gibbs. El muestreo de Gibbs implica la estimación de una secuencia de simulaciones en donde cada simulación de un parámetro es estimada condicionada sobre los demás parámetros en el modelo (en una forma jerárquica). Para ello, se debe tener en cuenta que  $\beta_n$  es un parámetro más junto a  $b$  y  $W$ . Así, la distribución posteriori de  $b$ ,  $W$  y  $\beta = (\beta_1, \dots, \beta_n)^T$  es

$$K(b, W, \beta|Y) \propto \prod_{n=1}^N L_n(y_n|\beta_n) f(\beta_n|b, W) k(b, W)$$

Antes de describir los pasos en detalle, se enunciarán dos lemas útiles para este desarrollo<sup>1</sup>:

**Lema 1** (Priori conjugada para una normal multivariada con media desconocida, pero varianza conocida). *Si se tiene un vector aleatorio  $\beta$  de dimensión  $K \times 1$  que sigue una distribución normal, con vector de medias desconocido  $b$  y matriz de covarianzas conocida  $W$ ,  $\beta \sim N(b, W)$ , y se tiene una muestra aleatoria  $\beta_n (n = 1, \dots, N)$  de  $N(b, W)$ , donde  $\bar{\beta} = (1/N) \sum_{n=1}^N \beta_n$  es la media muestral. Si la distribución a priori de  $b$  es normal, con vector de medias  $b_0$  y matriz de covarianzas  $W_0$ ,  $b \sim N(b_0, W_0)$ , cuya varianza tiende a infinito, entonces la distribución a posteriori de  $b$  es normal, con vector de medias  $\bar{\beta}$  y matriz de covarianzas  $W/N$ ,  $b \sim N(\bar{\beta}, W/N)$ .*

**Lema 2:** (Priori conjugada para una normal multivariada con media conocida, pero varianza desconocida). *Si se tiene un vector aleatorio  $\beta$  de dimensión  $K \times 1$  que sigue una distribución normal, con vector de medias conocido  $b$  y matriz de covarianzas desconocida  $W$ ,  $\beta \sim N(b, W)$ , y se tiene una muestra aleatoria  $\beta_n (n = 1, \dots, N)$  de  $N(b, W)$ , donde  $\bar{S} = (1/N) \sum_{n=1}^N (\beta_n - b)(\beta_n - b)'$  es la varianza alrededor de la media conocida  $b$ . Si la distribución a priori de  $W$  es wishart invertida con  $K$  grados de libertad y matriz escalar  $I$  ( $I$  es la matriz identidad),  $W \sim IW(K, KI)$ , entonces la distribución a posteriori de  $W$  es wishart invertida, con  $K + N$  grados de libertad y matriz escalar  $\frac{(KI + N\bar{S})}{K + N}$ ,  $W \sim IW\left(K + N, \frac{(KI + N\bar{S})}{K + N}\right)$ .*

Luego, asumiendo que la distribución a priori de  $k(b, w) = k(b)k(W)$ , donde  $k(b) \sim N(b_0, W_0)$  con una varianza que tiende a infinito,  $k(W) \sim IW(K, KI)$  y  $\beta_n \sim N(b, W)$ . A continuación se detalla el muestreo de Gibbs para los tres conjuntos de parámetros  $b$ ,  $W$  y  $\beta$ .

(1)  $b|W, \beta$ .

En este paso se condiciona respecto a  $W$  y a los  $\beta_n$ s de cada persona, lo que significa que se trata a estos parámetros como si se conocieran. Así, los  $\beta_n$ s constituyen una muestra de  $N$  realizaciones de una distribución normal con media desconocida  $b$  y varianza  $W$

<sup>1</sup>Una prueba del cumplimiento de estos lemas se muestra en las páginas 13 y 14.

conocida. La probabilidad de observar la muestra  $\boldsymbol{\beta}$  dado  $b$  y  $W$  es

$$\begin{aligned} p(\boldsymbol{\beta}|b, W) &= \prod_{n=1}^N (2\pi)^{-K/2} |W|^{-1/2} \exp \left[ -\frac{1}{2} (\beta_n - b)' W^{-1} (\beta_n - b) \right] \\ &= (2\pi)^{-NK/2} |W|^{-N/2} \exp \left[ -\frac{1}{2} \sum_{n=1}^N (\beta_n - b)' W^{-1} (\beta_n - b) \right] \\ &= (2\pi)^{-NK/2} |W|^{-N/2} \exp \left[ -\frac{1}{2} \sum_{n=1}^N \text{tr} \left[ W^{-1} (\beta_n - b) (\beta_n - b)' \right] \right] \\ &= (2\pi)^{-NK/2} |W|^{-N/2} \exp \left[ -\frac{1}{2} \text{tr} \left( W^{-1} \sum_{n=1}^N (\beta_n - b) (\beta_n - b)' \right) \right], \end{aligned}$$

donde la expresión “ $tr$ ” que se encuentra dentro de los corchetes indica el operador traza. Asimismo, cabe señalar que la expresión dentro de los paréntesis puede ser reexpresada de la siguiente forma

$$\sum_{n=1}^N (\beta_n - b) (\beta_n - b)' = S + N(\bar{\beta} - b)(\bar{\beta} - b)', \text{ donde } S = \sum_{n=1}^N (\beta_n - \bar{\beta})(\beta_n - \bar{\beta})'$$

Considerando esta reexpresión, se puede escribir  $p(\boldsymbol{\beta}|b, W)$  como

$$p(\boldsymbol{\beta}|b, W) = (2\pi)^{-NK/2} |W|^{-N/2} \exp \left[ -\frac{1}{2} \text{tr}(W^{-1}[S + N(\bar{\beta} - b)(\bar{\beta} - b)']) \right]$$

Ahora, dado que  $W$  es conocido, y que la distribución a priori de  $b$  es  $N(b_0, W_0)$  se obtiene

$$\begin{aligned} p(b|W, \boldsymbol{\beta}) &\propto \exp \left[ -\frac{1}{2} (b - b_0)' W_0^{-1} (b - b_0) \right] \\ &\quad \times \exp \left[ -\frac{1}{2} \text{tr}(W^{-1}[S + N(\bar{\beta} - b)(\bar{\beta} - b)']) \right] \\ &\propto \exp \left[ -\frac{1}{2} ((b - b_0)' W_0^{-1} (b - b_0) + (b - \bar{\beta})' (NW^{-1})(b - \bar{\beta})) \right] \end{aligned}$$

Dando la forma cuadrática en  $b$

$$\begin{aligned} &(b - b_0)' W_0^{-1} (b - b_0) + (b - \bar{\beta})' (NW^{-1})(b - \bar{\beta}) \\ &= (b - \bar{b})' (W_0^{-1} + NW^{-1})(b - \bar{b}) + (b_0 - \bar{\beta})' [W_0^{-1} (W_0^{-1} + NW^{-1})^{-1} NW^{-1}] (b_0 - \bar{\beta}), \end{aligned}$$

donde  $\bar{b} = (W_0^{-1} + NW^{-1})^{-1} (W_0^{-1} b_0 + NW^{-1} \bar{\beta})$ . Luego, si se simplifica la distribución posteriori de  $p(b|W, \boldsymbol{\beta})$  de la siguiente forma

$$p(b|W, \boldsymbol{\beta}) \propto \exp \left[ -\frac{1}{2} (b - \bar{b})' (W_0^{-1} + NW^{-1})(b - \bar{b}) \right]$$

se puede reconocer inmediatamente que esta última expresión es la función de densidad

de una distribución normal multivariada. Por lo tanto,

$$b|W, \boldsymbol{\beta} \sim N(\bar{b}, [W_0^{-1} + NW^{-1}]^{-1})$$

Ahora, ya que se consideró que la distribución a priori de  $b$  es  $N(b_0, W_0)$  con una varianza que tiende a infinito. Entonces,  $W_0 \rightarrow \infty \Rightarrow W_0^{-1} \rightarrow 0$ , lo que nos lleva de inmediato al cumplimiento del lema 1:

$$b|W, \boldsymbol{\beta} \sim N\left(\bar{\beta}, \frac{W}{N}\right)$$

(2)  $W|b, \boldsymbol{\beta}$ .

En este caso, los  $\beta_n$ s constituyen una muestra de una distribución normal con media  $b$  conocida y varianza  $W$  desconocida. Entonces, si se considera la probabilidad  $p(b|W, \boldsymbol{\beta})$  y la distribución a priori de  $W$ ,  $IW(v_0, v_0 S_0)$ , la distribución a posteriori de  $W$  es:

$$\begin{aligned} p(W|b, \boldsymbol{\beta}) &\propto |W|^{-(v_0+K+1)/2} \exp\left[-\frac{1}{2}\text{tr}(v_0 S_0 W^{-1})\right] \\ &\times \left[ |W|^{-N/2} \exp\left(-\frac{1}{2}\text{tr}([S + N(\bar{\beta} - b)(\bar{\beta} - b)']W^{-1})\right) \right] \\ &\propto |W|^{-(v_0+N+K+1)/2} \exp\left[-\frac{1}{2}\text{tr}([v_0 S_0 + S + N(\bar{\beta} - b)(\bar{\beta} - b)']W^{-1})\right] \\ &\propto |W|^{-(v_0+N+K+1)/2} \exp\left[-\frac{1}{2}\text{tr}([v_0 S_0 + N\bar{S}]W^{-1})\right] \\ &\propto |W|^{-(v_1+K+1)/2} \exp\left[-\frac{1}{2}\text{tr}(v_1 S_1 W^{-1})\right]. \end{aligned}$$

Esta última expresión representa la función de densidad de una distribución Wishart invertida  $IW(v_1, v_1 S_1)$ , donde

$$v_1 = v_0 + N \quad S_1 = \frac{v_0 S_0 + N\bar{S}}{v_0 + N}$$

No obstante, la distribución a priori de  $W$  se definió como  $IW(K, KI)$  y no como  $IW(v_0, v_0 S_0)$ , por lo que procediendo con la equivalencia de parametros, se tiene el inmediato cumplimiento del lema 2: la distribución a posteriori de  $W$  es una Wishart invertida con  $K + N$  grados de libertad y matriz de escala  $(KI + N\bar{S})/(K + N)$ , donde  $\bar{S} = (1/N) \sum_{n=1}^N (\beta_n - b)(\beta_n - b)'$  es la varianza muestral de las  $\beta_n$ s alrededor de la media conocida  $b$ .

(3)  $\beta_n|b, W$ .

La distribución a posteriori de los  $\beta_n$  de cada persona, condicionada respecto a sus elecciones y a los parámetros de la población, es

$$K(\beta_n|b, W, y_n) \propto L_n(y_n|\beta_n)\phi(\beta_n|b, W)$$



Para extraer valores al azar de esta distribución posteriori se utilizó el algoritmo MH. Así, para cada  $\beta_n^i$  el algoritmo MH opera de la siguiente forma:

- (a) Se comienza con un valor inicial  $\beta_n^i$ .
- (b) Se extrae  $K$  valores independientes de una densidad normal estándar, y se agrupan los valores en un vector etiquetado como  $\eta^i$ .
- (c) Se crea un valor de prueba de  $\beta_n^{i+1}$  como  $\tilde{\beta}_n^{i+1} = \beta_n^i + \rho L \eta^i$ , donde  $\rho$  es un escalar especificado por el investigador y  $L$  es el factor Choleski de  $W$ . Se debe tener en cuenta que la distribución propuesta del algoritmo MH se especifica como normal con media cero y varianza  $\rho^2 W$ .
- (d) Se extrae un valor de una variable uniforme estándar  $\mu^{i+1}$ .
- (e) Se calcula el ratio:

$$F = \frac{L(y_n | \tilde{\beta}_n^{i+1}) \phi(\tilde{\beta}_n^{i+1} | b, W)}{L(y_n | \beta_n^i) \phi(\beta_n^i | b, W)}$$

- (f) Si  $\mu^{i+1} \leq F$ , se acepta  $\tilde{\beta}_n^{i+1}$  y se define  $\beta_n^{i+1} = \tilde{\beta}_n^{i+1}$ . Si  $\mu^{i+1} > F$ , se rechaza  $\tilde{\beta}_n^{i+1}$  y se deja  $\beta_n^{i+1} = \beta_n^i$ .
- (g) Se repite el proceso varias veces. Para un  $i$  suficientemente alto,  $\beta_n^i$  es un valor extraído al azar de la distribución a posteriori.

En el algoritmo MH, el escalar  $\rho$  es especificado por el investigador. Este escalar determina el tamaño de cada salto dentro de la distribución. Usualmente, saltos más pequeños se trasladan en más aceptaciones, y saltos más grandes resultan en pocas aceptaciones. Sin embargo, usar saltos pequeños implica que el algoritmo MH necesitará más iteraciones para converger e involucra más correlación serial en la muestra después de la convergencia.

La tasa de aceptación óptima para el algoritmo MH es de aproximadamente 0.44 cuando  $K = 1$  y cae hasta 0.23 cuando  $K$  se eleva (vease Gelman et al. (1995, p. 335) para mayor detalle). En este sentido, el valor de  $\rho$  puede ser ajustado por el investigador para lograr una tasa de aceptación entorno a estos valores, bajando  $\rho$  para obtener una tasa de aceptación mayor y elevándolo para obtener una tasa de aceptación menor.

De hecho,  $\rho$  se puede ajustar como parte del proceso de iteración. El investigador establece el valor inicial de  $\rho$ . En cada iteración, un valor de prueba de  $\beta_n$  es aceptado o rechazado para cada muestra  $n$ . Si en una iteración, la tasa de aceptación entre las  $N$  observaciones es superior a un valor dado (por ejemplo, 0.33), entonces,  $\rho$  se eleva. Si la tasa de aceptación está por debajo de este valor,  $\rho$  se baja. Por tal, el valor de  $\rho$  puede moverse durante el proceso de iteración para alcanzar el nivel de aceptación especificado (Train, 2009).

Enunciando nuevamente pero de forma más concisa el procedimiento de estimación, el algoritmo comienza con valores iniciales de  $b^0$ ,  $W^0$ , y  $\beta_n^0$ . Luego, la  $i$ -ésima iteración del muestreo de Gibbs puede ser estimada en 3 pasos:

- Extraer  $b^i$  de  $N(\bar{\beta}^{i-1}, W^{i-1}/N)$  donde  $\bar{\beta}^{i-1}$  es la media de  $\beta_n^{i-1}$
- Extraer  $W^i$  de  $IW(K+N, (KI+N\bar{S}^{i-1})/(K+N))$ , donde  $\bar{S}^{i-1} = \sum_n (\beta_n^{i-1} - b^i)(\beta_n^{i-1} - b^i)' / N$
- Para cada muestra  $n$ , extraer  $\beta_n^i$  usando una iteración del algoritmo MH descrito anteriormente, empezando por  $\beta_n^{i-1}$  y usando la densidad  $\phi(\beta_n | b^i, W^i)$ .

Estos tres pasos se repiten para muchas iteraciones. Los valores resultantes convergen a valores extraídos de la distribución posteriori conjunta de  $b$ ,  $W$  y  $\beta$ . Una vez se obtienen los valores convergentes de la distribución posteriori, se puede calcular la media y la desviación estándar de los valores extraídos para obtener estimaciones y errores estándar de los parámetros.



## Capítulo 5

### Aplicación

#### 5.1. Base de datos y medición de las variables de investigación

De acuerdo al planteamiento de la presente tesis, se necesita incluir las externalidades de red como parte del vector de variables observables,  $x_{njt}$ , y poder así obtener la media ( $b$ ) y la desviación estándar ( $\sigma$ ) asociada a la distribución del coeficiente. La aproximación de los efectos de red dentro del logit mixto, se realizó utilizando la cuota de mercado de la empresa operadora  $j$  en el período  $t$ . Asimismo, los trabajos anteriores han incluido al precio como otro factor coadyuvante en la contratación del servicio de telefonía móvil y lo han aproximado por el ARPU (ingreso promedio reportado por las empresas operadoras entre el número de líneas en servicio). Sin embargo, dado que en el presente caso la información del ARPU no es pública, se utilizó el gasto mensual que declararon tener los consumidores e una encuesta.

Así, la base de datos que se dispone combina información de las empresas de telefonía móvil y de los clientes. La información sobre el número de líneas en servicio (total de suscriptores) de cada empresa operadora procede de los reportes administrativos publicados en la página institucional del OSIPTEL, mientras que los datos sobre la elección de los usuarios del servicio de telefonía móvil y el gasto asociado al servicio provienen de la Encuesta Residencial de Servicios de Telecomunicaciones (ERESTEL) del año 2015, y que también se encuentra disponible en la web del OSIPTEL.

La ERESTEL permite recoger información de los servicios de telecomunicaciones (telefonía fija y móvil, internet fijo y móvil, televisión de paga y telefonía de uso público) en términos de la demanda de acceso y uso de dichos servicios en el sector residencial (hogares y sus miembros). Asimismo, contiene información sobre las características demográficas y socioeconómicas de los encuestados. Por otra parte, es importante mencionar, que si bien en sus inicios la ERESTEL fue concebida como una base del tipo corte transversal, es a partir del año 2013 que se determinó que parte de la muestra global sea elegida para un seguimiento longitudinal cuyo cierre está programado para el año 2017. Por lo que actualmente, se cuentan con 3 olas panel correspondiente al periodo 2013 - 2015, que aún no son de acceso público.

La ERESTEL 2015 entrevistó a 55323 miembros del hogar, cuya edad van desde infantes hasta adultos mayores, y donde naturalmente no todas las personas cuentan con un celular o tienen el servicio de telefonía móvil. En este sentido, fue necesario aplicar una serie de filtros



a la muestra potencial de la ERESTEL y así tener una muestra de interés para el presente trabajo. Al respecto, es conveniente señalar que en general la estimación bayesiana presenta limitaciones de carácter computacional para trabajar con muestras bastantes grandes, por lo que se hace necesario en algunos casos acotar las bases de datos. Así pues, los criterios para la elección de los individuos que conformaron la muestra de análisis fueron los siguientes:

- La persona debe ubicarse en Lima Metropolitana. Esto a fin de reducir los problemas de contratación con una empresa operadora por falta de cobertura o disponibilidad del servicio al interior del país.
- La persona debe poseer un teléfono móvil.
- La persona debe tener una edad de 18 a 64 años.
- La persona debe ser el único responsable del pago de su servicio de telefonía móvil.
- La modalidad de contratación del servicio de telefonía móvil es postpago o control.
- El hogar de la persona debe pertenecer al nivel socioeconómico C.
- La persona declaró el último gasto mensual en su servicio de telefonía móvil.

Con estos filtros, la muestra utilizada para los fines del presente trabajo es de 615 individuos. Ahora bien, un aspecto importante en la construcción de la base de datos para la estimación del logit mixto es aproximar o construir los escenarios contrafactuales que dan respuesta a la siguiente pregunta: ¿Cuánto hubiera sido el precio que pagaba un consumidor si contrataba con el operador  $s$  en lugar del  $j$  en el período  $t$ ? Para ello, se imputó el gasto medio que declararían tener los consumidores en cada empresa operadora del servicio de telefonía móvil. Este promedio se calculó considerando al total de usuarios de Lima Metropolitana que declararían tener el servicio de telefonía móvil bajo la modalidad postpago o control. En el caso de la covariable asociada a los efectos de red, la cuota de mercado de las empresas operadoras se calculó teniendo en cuenta el promedio de líneas móviles en servicio de cada empresa operadora durante el año 2015. El Cuadro 5.1 resume la estrategia de aproximación de las variables que intervienen en el modelamiento.

Cuadro 5.1: Medición de las Variables introducidas en el Modelo

Variable		Definición	Efecto Esperado
Utilidad	$U_{njt}$	Variable dummy que representa la elección efectuada por el individuo: 1 =el consumidor $n$ elige la empresa operadora $j$ en el período $t$ . 0 =en caso contrario.	
Precio	$P_{njt}$	Logaritmo natural del gasto del usuario en la empresa $j$ en el período $t$ .	Negativo
Efectos de red clásicos	$CN_{njt}$	Participación de mercado del número de líneas en servicio de la empresa operadora $j$ en el período $t$ , expresado en porcentaje.	Positivo

## 5.2. Resultados de la investigación

El análisis de regresión se realizó con el modelo logit mixto bajo la aproximación bayesiana, el cual fue descrito en el Capítulo 4, y cuya función de utilidad considerando la base de datos y las variables utilizadas, toma la siguiente forma:

$$U_{njt} = \beta_n^P P_{njt} + \beta_n^{CN} CN_{njt} + \varepsilon_{njt}; \quad t = 1, j = 1, 2, 3, 4.$$

Las estimaciones se obtuvieron utilizando un código de simulación propio, que fue desarrollado en el *software* para el análisis estadístico R. Dicho código se basó en el procedimiento presentado en Train (2009, pág. 301-302) y que fue expuesto en el Capítulo 4. En la estimación, se especificó un total de 160.000 simulaciones, de las cuales las 10.000 primeras iteraciones se destinaron para la fase de *burn-in* (quemado), y de las 150.000 restantes se conservaron uno de cada trecientos (*thin*). Esto, para obtener los resúmenes estadísticos de los parámetros de interés <sup>1</sup>. Los valores iniciales para el vector de medias  $b$  se obtuvieron a partir de la regresión logit condicional (logit estándar), mientras que para la matriz de covarianzas  $W$  se partió de una matriz identidad ( $I$ ). Luego, esta matriz  $I$  y el vector de medias mencionado, fueron empleados para generar los valores iniciales de los  $\beta$ s individuales por medio de una distribución normal multivariada. Asimismo, el parámetro  $\rho$  se fijó en 2.608 (redondeado a 3 decimales) y se creó un contador para registrar la tasa de aceptación obtenida luego de terminada la simulación. Mayores detalles sobre el código de estimación pueden encontrarse en el Apéndice A.

Los resultados del procedimiento bayesiano del modelo logit mixto se presentan en el Cuadro 5.2. Dicho cuadro muestra la siguiente información: (1) En la primera columna se encuentran los coeficientes asociados a las covariables consideradas en la función de utilidad, (2) en la segunda columna se tiene el parámetro de la media y la desviación estándar correspondiente a cada coeficiente, (3) en la tercera columna se muestra la estimación de la media por parámetro, en tanto que en la (4) cuarta columna se encuentra el valor obtenido de su desviación estándar. Por último, (5) en la quinta y (6) sexta columna se presenta los límites del intervalo de credibilidad (ICr) formado por los cuantiles 0.025 (límite inferior o LI) y 0.975 (límite superior o LS) de la muestra simulada (Albert, 2009; pág. 64-65). Cabe señalar, que el ICr permite evaluar la relevancia estadística de los parámetros. Así, cuando dentro del ICr de la media simulada se encuentra el cero (0), este parámetro se considerará como estadísticamente no significativo y el análisis de su dispersión pierde relevancia.

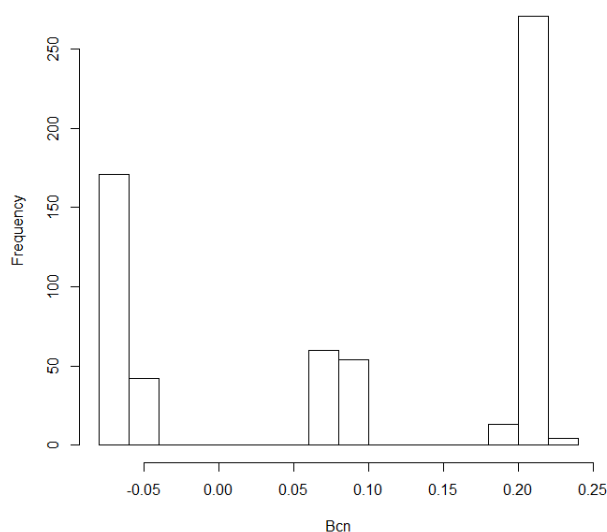
<sup>1</sup>Salto de 300 permitieron asegurar que las cadenas simuladas no tuvieran problemas de autocorrelación de orden mayor a 1 o 2, aunque es importante mencionar que saltos de 100 solo generarían autocorrelaciones que van descendiendo hasta el orden 5, para luego no ser significativas

Cuadro 5.2: Resultados de la Estimación

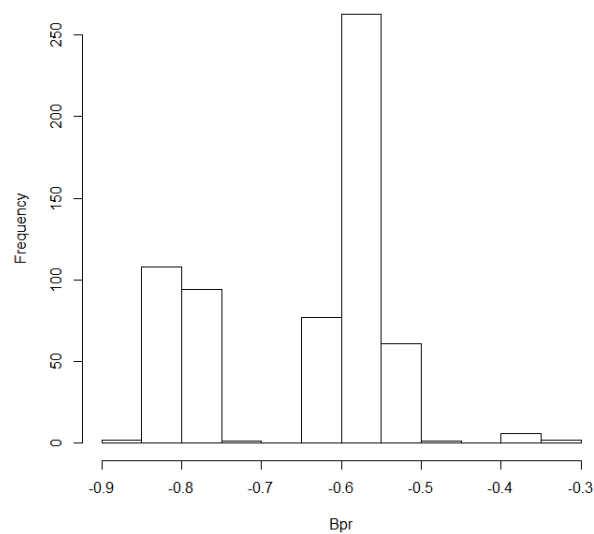
Coefficiente	Parámetros de la distribución	Media	Error Estándar	LI	LS
$\beta_n^P$	Media de la distribución ( $b^P$ )	-0,648	0,090	-0,819	-0,476
	Desv. est. de la distribución ( $\sigma^P$ )	0,413	0,077	0,288	0,590
$\beta_n^{CN}$	Media de la distribución ( $b^{CN}$ )	0,090	0,010	0,071	0,110
	Desv. est. de la distribución ( $\sigma^{CN}$ )	0,170	0,014	0,148	0,198

La tasa de aceptación obtenida fue de 0,249 y se encuentra dentro del rango óptimo señalado por Gelman et al. (1995, p. 335). Los resultados muestran que el impacto sobre la utilidad asociado a los efectos de red ( $\beta_n^{CN}$ ) presenta una media positiva ( $b^{CN} = 0,090$ ) y estadísticamente relevante (su ICr no contiene al 0). Ello significa que el tamaño de la empresa operadora es un factor importante para el consumidor cuando decide elegir una compañía de telefonía móvil ya que su nivel de bienestar o utilidad tiene un incremento estadísticamente importante. Así, los operadores de mayor cuota de mercado -que usualmente son los que llevan más tiempo operando en la industria- pueden direccionar por defecto las decisiones de contratación de los consumidores hacia sus redes, afectando la captación de clientes de nuevos entrantes o competidores, mermando la desconcentración del mercado. En cuanto a la heterogeneidad de los impactos de las externalidades de red, se registró que el incremento de utilidad que supone la elección de un operador de mayor tamaño no es idéntica para todos los individuos, sino que existen diferencias estadísticamente relevantes que ponen de manifiesto cierta diversidad en la influencia de las externalidades de red sobre los consumidores; lo cual queda reflejado a través de la desviación estándar de  $\beta_n^{CN}$  ( $\sigma^{CN} = 0,170$ ).

Así pues, la Figura 5.1 presenta la distribución estimada de los  $\beta_n^{CN}$  para los individuos de la muestra bajo estudio, donde se puede apreciar que para la gran mayoría de consumidores el tamaño del proveedor de telefonía móvil implica un aumento de su utilidad en alrededor de 0,23. Sin embargo, existen grupos minoritarios de consumidores donde este efecto puede llegar a cambiar, siendo negativa en algunos casos (entorno a -0,05) o con una influencia positiva leve (entorno a 0,07). Esta heterogeneidad se manifiesta a pesar de tener una muestra acotada a un grupo específico de consumidores.

Figura 5.1: Distribución de  $\beta_n^{CN}$  en la muestra

Naturalmente, otro factor que también se encuentre como determinante en la elección de empresa operadora que realicen los consumidores estaría representado por el precio del servicio de telefonía móvil. Esto, de acuerdo a la estimación de la media ( $b^P = -0,648$ ) correspondiente al coeficiente  $\beta_n^P$  y su respectivo ICr. Además, este efecto del precio presenta una importante heterogeneidad de acuerdo a la desviación estándar estimada ( $\sigma^P = 0,413$ ) y su distribución en la muestra de estudio, según se puede observar en la Figura 5.2. Así, en algunos consumidores, la influencia del precio puede llegar a ubicarse en  $-0,9$  o reducirse a  $-0,3$ . Esta heterogeneidad se encuentra, nuevamente, pese a tener una muestra bajo análisis correspondiente a un grupo particular de usuarios del servicio de telefonía móvil.

Figura 5.2: Distribución de  $\beta_n^P$  en la muestra

Por último, las Figuras 5.3 y 5.4 presentan la función de autocorrelación de la media y la desviación estándar de cada coeficiente estimado, mientras que las Figuras 5.5 y 5.6 muestran la evolución de las cadenas de los valores estimados. Así, estos gráficos dan indicios que los valores simulados lograron alcanzar la convergencia y formar una muestra estadísticamente independiente.

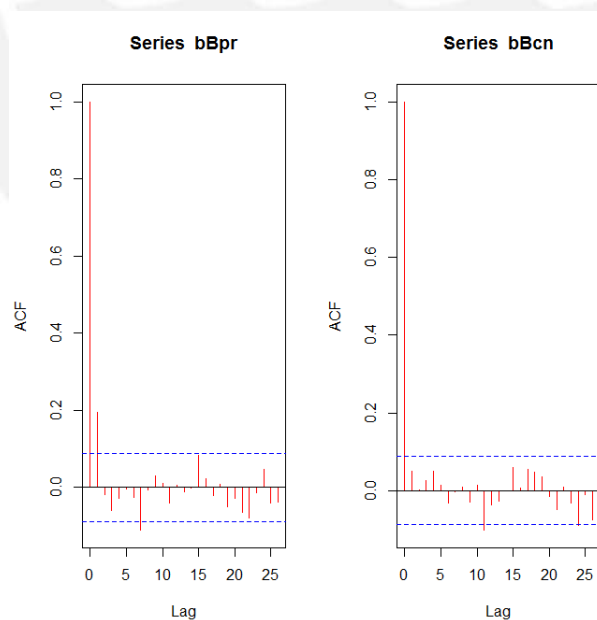
Figura 5.3: Distribución de las correlaciones del vector de medias  $b$  en la muestra simulada

Figura 5.4: Distribución de las correlaciones de las desviaciones estándar de la matriz de covarianzas  $W$  en la muestra simulada

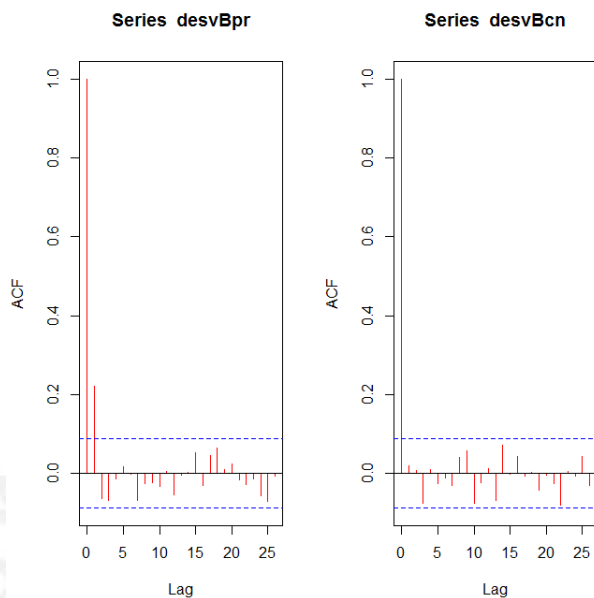


Figura 5.5: Distribución de la cadena del vector de medias  $b$  en la muestra simulada

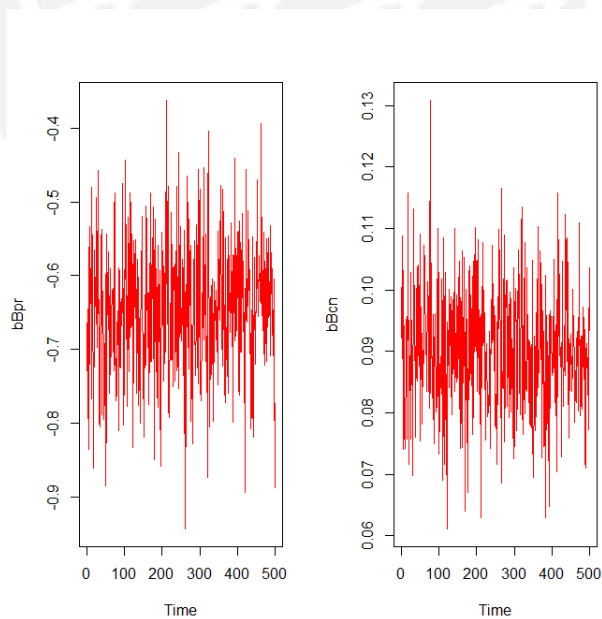
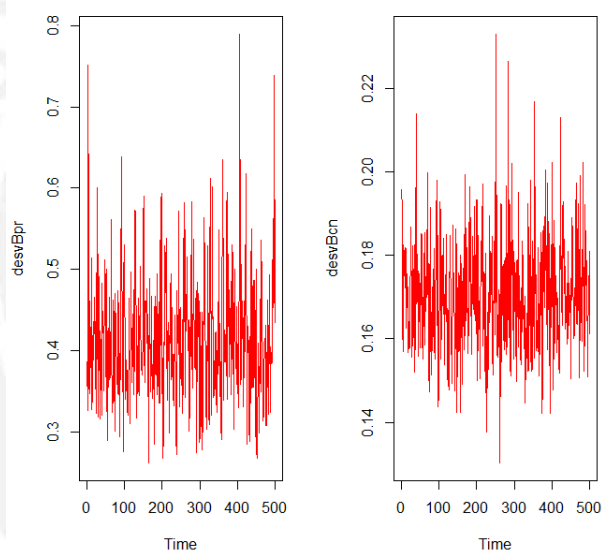


Figura 5.6: Distribución de la cadena de las desviaciones estándar de la matriz de covarianzas  $W$  en la muestra simulada





## Capítulo 6

# Conclusiones

### 6.1. Conclusiones

Como se señaló en el Capítulo 1, la existencia de efectos de red, al generar una atracción o dependencia en los consumidores por proveedores de mayor cuota de mercado, puede degradar la intensidad competitiva de un mercado, protegiendo la base de clientes de las empresas operadoras de mayor tamaño frente a la competencia de nuevos retadores. Así, se generaría una situación donde se consolida el dominio de las empresas ya establecidas y se elimina la competencia efectiva y potencial de nuevos competidores, llevando al mercado a una situación oligopólica donde es difícil la desconcentración.

La presente investigación tenía por objetivo principal estudiar la magnitud y heterogeneidad de los efectos de red en el mercado de telefonía móvil peruano usando un modelo logit mixto bayesiano. Así, dado la base de datos utilizada, el período de análisis, y la metodología y estrategia de estimación empleada, la evidencia empírica obtenida permitiría sostener que los efectos de red no solo se encontrarían presentes en el mercado de telefonía móvil peruano, a través de su influencia en la utilidad de los usuarios, sino que también esta influencia resultaría tener cierto grado de heterogeneidad entre los consumidores. Esto, pese a que en el presente estudio se utilizó una muestra específica de clientes: usuarios postpago o control del nivel socioeconómico C.

Dado estos hallazgos, se puede plantear algunas implicaciones de política para los diferentes agentes económicos, y, en especial, para el organismo regulador. En este sentido, las medidas que ha venido implementando el OSIPTEL en los últimos años a fin de potenciar la competencia en el mercado de telefonía móvil, tales como el establecimiento de cargos de terminación móvil diferenciados según el tamaño o participación de mercado de las empresas operadoras, se encontraría justificado y debe seguir siendo mejorado y complementado por otras medidas que ayuden a relativizar la influencia de los efectos o externalidades de red. Para ello, sería importante tener en cuenta que la magnitud en que afectan dichas externalidades no sería homogénea en todos los usuarios del servicio de telefonía móvil, sino que existirían diferencias significativas entre ellos.

Por ejemplo, se podría argumentar que la propia dinámica competitiva del mercado móvil peruano estaría relativizando la importancia de realizar políticas que fomenten un juego com-



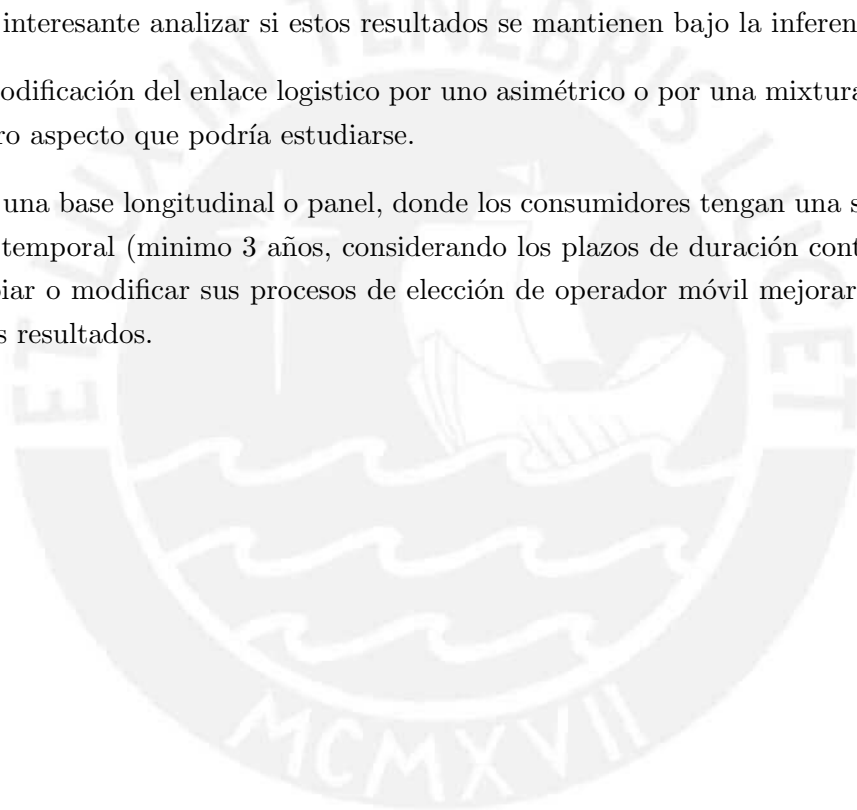
petitivo más equilibrado, como los cargos móviles diferenciados, ya que se está observando que cada vez más consumidores están optando por elegir operadores de menor cuota de mercado, hecho que se reflejaría en la estimación realizada, donde para un grupo de abonados el efecto de las externalidades de red resultó negativo. No obstante, la misma estimación llevada a cabo muestra que para la mayor parte del grupo de análisis aún se prefiere elegir los operadores de tamaño de red grandes: el efecto de las externalidades de red en este grupo es positivo, además de que su magnitud es mayor al grupo donde el efecto es negativo, lo que en neto da una influencia positiva de las externalidades de red sobre la utilidad del consumidor. En este sentido, la política del establecimiento de cargos no podría descartarse en el corto plazo. Sin embargo, dicha afirmación no necesariamente se puede mantener en el largo plazo, particularmente, si se considera la creciente inversión que están realizando los operadores en redes de datos (internet) y el direccionamiento de las preferencias de la demanda hacia servicios de voz o mensajería que se realizan por internet (por ejemplo, el mayor uso de *WhatsApp* o *Skype*).

## 6.2. Sugerencias para investigaciones futuras

Este trabajo no está libre de limitaciones, algunas de las cuales pueden servir de punto de partida o referencia para futuros análisis. A continuación, se listan posibles alternativas de mejora o extensiones:

- Para realizar la estimación se tuvo que recurrir a un procedimiento de imputación sobre el precio que pagaría cada decisor si estuviera en otro operador diferente a su actual proveedor. Asimismo, los efectos de red solo tenían la variabilidad disponible de la cuota de mercado de las empresas operadoras al 2015. Futuros trabajos, podrían construir un experimento de elección considerando la técnica de *Conjoint*, donde el precio y el número de clientes de cada opción o alternativa de proveedor se defina previamente y varíen en una serie de preguntas hipotéticas donde el consumidor tiene que elegir, dado el precio y la cuota de mercado de la compañía, su operador de telefonía móvil.
- No se puede descartar que la estimación de los coeficientes pueda tener un grado de sobre o subestimación como consecuencia de la posible endogeneidad de las variables explicativas en la especificación mostrada, particularmente en el caso del precio. Por ende, trabajos posteriores podrían ampliar el modelo logit mixto para incorporar más covariables, reduciendo las posibilidades de sesgo por variables omitidas relevantes, o un método que permita estimar el modelo de elección en presencia de variables explicativas endógenas, como el enfoque basado en la función de control.
- El modelamiento de la media de los  $\beta$ s ha sido incondicional. Empero, se podría incluir un segundo nivel de análisis, donde la media dependa de ciertas covariables como el nivel socioeconómico, la edad, el género, la modalidad de contratación del servicio de telefonía móvil, la duración del contrato, la empresa a la que se pertenece, el tiempo con el servicio móvil, etc.

- Las estimaciones obtenidas de la media y la varianza correspondientes a los  $\beta$ s se obtuvieron considerando una distribución normal. Nuevas investigaciones, podrían considerar la modificación de este supuesto y analizar su influencia sobre los resultados finales.
- El uso del algoritmo de Gibbs para simular la distribución posteriori también podría modificarse por algoritmos que tengan otras ventajas comparativas. Por ejemplo, en la reducción del tiempo de estimación o frente a la necesidad de reducir los problemas de autocorrelación o convergencia de las cadenas generadas. En esta línea, futuros trabajos podrían realizar la implementación del logit mixto (o sus mejoras) en el *software* Stan, una herramienta desarrollada recientemente para la inferencia bayesiana en los modelos estadísticos y que mejora la calidad de las estimaciones por el algoritmo que utiliza.
- Los resultados hallados han sido bajo el enfoque de la inferencia bayesiana. No obstante, sería interesante analizar si estos resultados se mantienen bajo la inferencia clásica.
- La modificación del enlace logístico por uno asimétrico o por una mixtura de funciones es otro aspecto que podría estudiarse.
- Usar una base longitudinal o panel, donde los consumidores tengan una suficiente ventana temporal (mínimo 3 años, considerando los plazos de duración contractual) para cambiar o modificar sus procesos de elección de operador móvil mejoraría la robustez de los resultados.



## Apéndice A

### Código

```
set.seed(475446) # fija la semilla del random number generator.

##install.packages("mvtnorm")
library(mvtnorm) # rmvnorm function
##install.packages("MCMCpack")
library(MCMCpack) # rwish and riwish function
##install.packages("mcmcse")
library(mcmcse)
library(bayesm)

#####
# Datos #
#####
library("foreign")
base1=read.dta("C:/Users/Paulo_Roberto/Desktop/Tesis/base_y_do_siset/base.dta")
head(base1, n=15) # las 15 primeras líneas
lapply(base1, class) # Toda la base esta en formato número
save(base1, file="base1.Rda")
head(base1, n=6)

choice1<-(base1$choice==1)
choice2<-(base1$choice==2)
choice3<-(base1$choice==3)
choice4<-(base1$choice==4)

#Frecuencia de las elecciones
fe=table(base1$choice)

N<-nrow(base1)

#####
# Repositorios y Valores Iniciales #
#####

memory.limit(size = 40000)
nsim<-160000
K<-2

be<-matrix(0,nsim,K); be[1,]<-c(-1.800999,.030279)
bedoble<-array(0,dim=c(K,K,nsim)); bedoble[, ,1]<-diag(K)
beta<-betanew<-array(0,dim=c(N,K,nsim)); beta[, ,1]<-rmvnorm(N,be[1,],bedoble[, ,1])

rho<-2.60831426
counter<-0

#####
# MCMC #
#####
for (t in 2:nsim){

  # Extrayendo be
  v<-bedoble[, ,t-1]/N
  m<-apply(beta[, ,t-1],2,mean)
  be[t,]<-rmvnorm(1,m,v)

  # Extrayendo bedoble
  tmp<-t(beta[, ,t-1])-be[t,]
  bedoble[, ,t]<-rwishart( K+N, solve(K*diag(K) + tmp%*(tmp)))$IW
```

```

for (n in 1:N){
# Calcula la verosimilitud basada en el antiguo valor de beta
v1 <- log(base1[n,4])*beta[n,1,t-1] + base1[n,12]*beta[n,2,t-1]
v2 <- log(base1[n,5])*beta[n,1,t-1] + base1[n,13]*beta[n,2,t-1]
v3 <- log(base1[n,6])*beta[n,1,t-1] + base1[n,14]*beta[n,2,t-1]
v4 <- log(base1[n,7])*beta[n,1,t-1] + base1[n,15]*beta[n,2,t-1]

lold <- log ( (exp(v1) * choice1[n] + exp(v2) * choice2[n] + exp(v3) * choice3[n]
+ exp(v4) * choice4[n]) / ( exp(v1) + exp(v2) + exp(v3) + exp(v4) ) )

# Extrayendo candidatos
betanew[n,,t]<- beta[n,,t-1] + t(chol(rho*bedoble[, ,t])) %norm(K)

# Calcula la verosimilitud basada en el nuevo valor de beta
v1 <- log(base1[n,4])*betanew[n,1,t] + base1[n,12]*betanew[n,2,t]
v2 <- log(base1[n,5])*betanew[n,1,t] + base1[n,13]*betanew[n,2,t]
v3 <- log(base1[n,6])*betanew[n,1,t] + base1[n,14]*betanew[n,2,t]
v4 <- log(base1[n,7])*betanew[n,1,t] + base1[n,15]*betanew[n,2,t]

lnew <- log ( (exp(v1) * choice1[n] + exp(v2) * choice2[n] + exp(v3) * choice3[n]
+ exp(v4) * choice4[n]) / ( exp(v1) + exp(v2) + exp(v3) + exp(v4) ) )

# Probabilidad de Aceptación (en escala logaritmica)
r <- lnew + dmvnorm(betanew[n,,t], be[t,], bedoble[, ,t], log=T) -
( lold + dmvnorm(beta[n,,t-1], be[t,], bedoble[, ,t], log=T) )

if ( log ( runif(1) ) < r ) {
beta[n,,t]<-betanew[n,,t]
counter=counter + 1
} else {
beta[n,,t]<-beta[n,,t-1]
}
}

if (t%30==0) print(t)
} # Fin MCMC

accept.rate=counter/(nsim*N)

#CADENAS SIMULADAS DEL VECTOR b

##Media de los elementos del vector b
nsim<-160000
burn<-10001
thin<-300
index<-seq(1,nsim-burn+1,thin)
str(index)

bBpr<-be[burn:nsim,1][index]
bBcn<-be[burn:nsim,2][index]

length(bBpr)
length(bBcn)

mean_b=c(mean(bBpr), mean(bBcn))

##Desviación estándar de los elementos del vector b
sd_b=c(sd(bBpr), sd(bBcn))

##Intervalos de credibilidad para la media de los elementos del vector b, basados en cuantiles
qbBpr=quantile(bBpr,c(.025,.975))
qbBcn=quantile(bBcn,c(.025,.975))

#CADENAS SIMULADAS DE LA MATRIZ W
varBpr=bedoble[1,1,burn:nsim][index]
varBcn=bedoble[2,2,burn:nsim][index]

desvBpr=sqrt(varBpr)
desvBcn=sqrt(varBcn)

length(desvBpr)
length(desvBcn)

##Media de las Varianzas (y Desviación Estándar) de la matriz W
mean_W=c(mean(varBpr), mean(varBcn))
mean_sdW=c(mean(desvBpr), mean(desvBcn))

##Desviación estándar de las Varianzas (y Desviación Estándar) de la matriz W
sd_W=c(sd(varBpr), sd(varBcn))
sd_sdW=c(sd(desvBpr), sd(desvBcn))

```

```

##Intervalos de credibilidad para la desviación estándar de las desviaciones estándar
##de la matriz W
qdvBpr=quantile(desvBpr,c(.025,.975))
qdvBcn=quantile(desvBcn,c(.025,.975))

#Bs individuales

Bpr=apply(beta[,1,seq(burn,nsim,by=300)],1,mean)
Bcn=apply(beta[,2,seq(burn,nsim,by=300)],1,mean)

hist(Bpr,main=NULL)
hist(Bcn,main=NULL)

#Mostrando los resultados de la simulación: Tasa de aceptación, medias, desviaciones estándar
#e intervalos de credibilidad

accept.rate

mean_b
sd_b
qbBpr
qbBcn

mean_sdW
sd_sdW
qdvBpr
qdvBcn

#Mostrando los resultados de la simulación: La función de autocorrelación

par(mfrow=c(1,2))
acf(bBpr,col=2)
acf(bBcn,col=2)

par(mfrow=c(1,2))
acf(desvBpr,col=2)
acf(desvBcn,col=2)

#Mostrando los resultados de la simulación: La cadena

par(mfrow=c(1,2))
ts.plot(bBpr,col=2)
ts.plot(bBcn,col=2)

par(mfrow=c(1,2))
ts.plot(desvBpr,col=2)
ts.plot(desvBcn,col=2)

save(accept.rate, be, bedoble, beta, fe, file=
"C:/Users/Paulo_Roberto/Desktop/Tesis/base_y_do_siset/Especificación_final/myscript11_160k.rda")
load(file=
"C:/Users/Paulo_Roberto/Desktop/Tesis/base_y_do_siset/Especificación_final/myscript11_160k.rda")

```

## Bibliografía

- Albert, J. (2009). *Bayesian Computation with R*. Springer.
- Baker, M. (2014). *Adaptive Markov chain Monte Carlo sampling and estimation in Mata*. The Stata Journal, 14(3), 623-661.
- Balcombe, K., Chalak, A., y Fraser, I. (2009). *Model selection for the mixed logit with Bayesian estimation*. Journal of Environmental Economics and Management, 57, 226-237.
- Bhat, C. (2000). *Incorporating observed and unobserved heterogeneity in urbanwork travel mode choice modeling*. Transportation Science, 34 (2), 228-238.
- Birke, D., y Swann, G.P.(2006). *Network effects and the choice of mobile phone operator*. Journal of Evolutionary Economics, 16(1-2), 65-84.
- Beuermann, D., McKelvey, C., y Sotelo C. (2012). *The Effects of Mobile Phone Infrastructure: Evidence from Rural Peru*. Documento de Trabajo N° 2012-12. BCRP.
- Czajkowski, M., y Sobolewski, M. (2015). *How much do switching costs and local network effects contribute to consumer lock-in in mobile telephony?*. Telecommunications Policy.
- Corrocher, N., y Zirulia, L.(2009). *Me and you and everyone we know: an empirical analysis of local network effects in mobile communications*. Telecommunications Policy, 33(1), 68-79.
- Doganoglu, T. y Grzybowski, L. (2007). *Estimating network effects in mobile telephony in Germany*. Information Economics and Policy, 19(1), 65-79.
- Farrell, J., y Klemperer, P. (2007). *Coordination and lock-in: competition with switching costs and network effects*. En M. Armstrong, y R.Porter (Eds.), Handbook of industrial organisation (pp. 1967-2072). Elsevier: Elsevier, 2007.
- Fu, W.W. (2004). *Termination-discriminatory pricing, subscriber bandwagons, and network traffic patterns: the Taiwanese mobile phone market*. Telecommunications Policy, 28(1), 5-22.
- Hall, J., Fiebig, D.G., King, M.T., Hossain, I., y Louviere, J.J. (2006). *What influences participation in genetic carrier testing? Results from a discrete choice experiment*. Journal of Health Economics, 25(3), 520-537.
- Huber, J., y Train, K. (2001). *On the Similarity of Classical and Bayesian Estimates of Individual Mean Partworths*. Marketing Letters, 12(3), 259-269.



- Johnson, F.R., Banzhaf, M.R., y Desvousges, W.H. (2000). *Willingness to pay for improved respiratory and cardiovascular health: a multiple-format, stated-preference approach*. Health Economics, 9(4), 295-317.
- Karaçuka, M., Çatik, A.N., y Haucap, J.(2013). *Consumer choice and local network effects in mobile telecommunications in Turkey*. Telecommunications Policy, 37(4), 334-344.
- Katz, M., y Shapiro, C. (1985). *Networks externalities, competition, and compatibility*. The American Economic Review, 75(3), 424-440.
- Katz, M., y Shapiro, C. (1994). *Systems competition and network effects*. Journal of Economic Perspectives, 8(2), 93-115.
- Klemperer, P. (1995). *Competition when consumers have switching costs: an overview with applications to industrial organization, macroeconomics, and international trade*. The Review of Economic Studies, 4, 515.
- Kim, H.S., y Kwon, N. (2003). *The advantage of network size in acquiring new subscribers: a conditional logit analysis of the Korean mobile telephony market*. Information Economics and Policy, 15(1), 17-33.
- King, M.T., Hall, J., Lancsar, E., Fiebig, D. Hossain, I., Louviere, J., Reddel, H.K., y Jenkins, C.R. (2007). *Patient preferences for managing asthma: results from a discrete choice experiment*. Health Economics, 16 (7), 703-717.
- Lancsar, E., Hall, J.P., Kennny, P., Louviere, J.J., Fiebig, D.G., Hossain, I., Thien, F.C., Reddel, H.K., y Jenkins, C.R. (2007). *Using discrete choice experiments to investigate subject preferences for preventive asthma medication*. Respiriology, 12(1), 127-136.
- Maicas, J.P., Polo, Y., y Sese, F.J. (2009). *The role of (personal) network effects and switching costs in determining mobile users' choice*. Journal of Information Technology, 24(2), 160-171.
- Maicas, J.P., Polo, Y., y Sese, F.J. (2010). *El Papel de los efectos de red y los costos de cambio en la elección de la compañía de telefonía móvil*. Revista Europea de Dirección y Economía de la Empresa, vol. 19, núm 3, pp. 161-172.
- Maicas, J.P., y Sese, F.J. (2011). *Network effects in the mobile communications industry: an overview*. In Maicas Juan P.(Ed.), Recent Developments in the Mobile Communications Industry: A Multidisciplinary Approach (pp. 131-140). Croatia: In Tech.
- Marschak, J. (1960). *Binary choice constraints on random utility indications*. in K. Arrow, ed., Stanford Symposium on Mathematical Methods in the Social Sciences, Stanford University Press, Stanford, CA, pp. 312-329.
- McFadden, D. (1974). *Conditional logit analysis of qualitative choice behavior*. in P. Zarembka, ed., Frontiers in Econometrics, Academic Press, New York, pp. 105-142.

- McFadden, D., y Train, K. (2000). *Mixed MNL models for discrete response*. Journal of Applied Econometrics, 15(5), 447-470.
- Negrin, M.A., Pinilla, J., y León, C.J. (2008). *Willingness to pay for alternative policies for patients with Alzheimer's Disease*. Health Economics, Policy and Law 3 (Pt 3), 257-275.
- Regiera, D., Ryanb, M., Phimister, E., y Marrad, C. (2009). *Bayesian and classical estimation of mixed logit: An application to genetic testing*. Journal of Health Economics, 598-610.
- Revelt, D., y Train, K. (2000). *Specific taste parameters and mixed logit*. Working paper, Department of Economics, University of California, Berkeley.
- Rust, R., y Verhoef, P. (2005). *Optimizing the Marketing Interventions Mix in Intermediate-Term CRM*. Marketing Science, 24(3), 477-489.
- Shy, O. (2001). *The Economics of Network Industries*. Cambridge University Press.
- Sobolewski, M., y Czajkowski, M. (2012). *Network effects and preference heterogeneity in the case of mobile telecommunications markets*. Telecommunications Policy, 36(3), 197-211.
- Ocaña, C., Polo, F., y Sesé, F. (2009). *Existencia y heterogeneidad de los costes de cambio en la industria de la telefonía móvil*. Hacienda Pública Española / Revista de Economía Pública, 191(4/2009): 926.
- Polo, Y., y Sesé, J. (2009). *La Retención de los Clientes. Un Estudio Empírico de sus Determinantes*. Revista Española de Investigación de Marketing ESIC. Vol. 13, n° 1, pp. 117-137.
- Thurstone, L. (1927). *A law of comparative judgement*. Psychological Review 34, 273-286.
- Train, K. (1998). *Recreation demand models with taste differences over people*. Land Economics, 74 (2), 230-239.
- Train, K. (2001). *A Comparison of Hierarchical Bayes and Maximum Simulated Likelihood for Mixed Logit*. Department of Economics, University of California, Berkeley.
- Train, K. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press.
- Train, K., y Sonnier, G. (2005). *Mixed logit with bounded distributions of correlated part-worths*. In: Alberini, A., Scarpa, R. (Eds.), Applications of Simulation Methods in Environmental and Resource Economics. Springer, Dordrecht, pp. 117-134.