

EFFECTO DEL PRECIO Y LA LEALTAD EN EL COMPORTAMIENTO DE ELECCIÓN DEL CONSUMIDOR: MODELIZACIÓN LOGIT MULTINOMIAL CON DATOS DE ESCÁNER

Juan Carlos Gázquez Abad
Manuel Sánchez Pérez

RESUMEN

Los directivos de los establecimientos comerciales pueden utilizar los modelos de elección para la toma de sus decisiones comerciales, siendo muy útiles para aspectos como la planificación promocional, la selección de marca o el rendimiento de las categorías. En este trabajo, analizamos la influencia que tienen algunos factores sobre el comportamiento de elección, utilizando el modelo logit multinomial, a partir de datos de escáner de aceite de oliva. Incluimos factores como el precio de compra regular de cada alternativa, y la lealtad del consumidor hacia cada una de ellas, comparando los resultados para determinar cuál de las dos variables analizadas es más relevante.

PALABRAS CLAVES: Comportamiento del consumidor, modelos de elección, MNL, datos de escáner

ABSTRACT

Store managers may make decisions regarding product categories with the help of a choice model, and may find it very useful for pricing promotion planning, brand selection and category assortment. In this paper we analyse some factors influencing consumer choice behavior using multinomial logit models based on scanner data for oliva oil. We include factors like regular price or loyalty. We compare the results for both factors in order to stablish which is the most relevant one.

KEY WORDS: Consumer behavior, choice models, MNL, scanner data

1. INTRODUCCIÓN

La introducción del escáner en los puntos de venta en régimen de autoservicio apoyada por la adopción general de los códigos de barras que identifican cada referencia concreta de producto (código *EAN*³⁶), ha revolucionado la facilidad con la que se pueden realizar diferentes análisis de marketing, con objeto de comprender mejor el comportamiento del consumidor que acude a estos establecimientos (Christen, Gupta, Porter, Staelin y Wittink, 1997). Ya en el año 1989, más del 50% de los supermercados de Estados Unidos utilizaban los datos de escáner como sistema de gestión de sus productos (Capps Jr., 1989:750). A esta información procedente de los equipos de escáner, se le une la elevada disponibilidad de datos proporcionados por el comercio electrónico y por las actividades de marketing directo (Rossi y Allenby, 2000).

³⁶ European Article Numbering.

Los datos que proporciona el escáner constituyen una valiosa fuente de información para el minorista, teniendo múltiples aplicaciones en el plano económico. Se trata de datos cuya cantidad, precisión y coste resulta extremadamente ventajoso, a la vez que es posible obtener información fiable y válida de los productos que oferta el distribuidor sobre una base temporal (Capps Jr., 1989). Ha sido, precisamente, esta creciente disponibilidad (tanto electrónica como no) de datos de compra del consumidor, la que ha propiciado el predominio que en la actualidad está teniendo la investigación cuantitativa en la literatura de marketing (Hildebrandt y Klapper, 2001:139). Así, el creciente número de estudios que buscan determinar el impacto de las variables de *marketing-mix* (sobre todo el precio y/o las promociones) en las ventas de un producto, viene motivado por esta mayor disponibilidad de la secuencia histórica de las compras del consumidor a través de la utilización de sistemas de escáner (Bentz y Merunka, 2000).

2. FACTORES QUE INFLUYEN EN EL COMPORTAMIENTO DE ELECCIÓN DEL CONSUMIDOR

Desde la literatura sobre elección de marca (p.e. Bucklin, Gupta y Han, 1995; Guadagni y Little, 1983; Kamakura y Rusell, 1989; Papatla y Krishnamurthi, 1996), se le ha dedicado un esfuerzo importante a su modelización en función de diferentes variables de marketing. Así, existe cierto consenso a la hora de considerar que éstas deben ser incluidas en los modelos de utilidad del consumidor, además de constantes específicas de las alternativas de elección para recoger las preferencias no incorporadas por las variables explicativas.

Existen, de un lado, modelos que emplean únicamente variables referentes a las alternativas de elección, denominadas como “variables de marketing” y asumiendo, por tanto, la homogeneidad en el comportamiento de elección de todos los individuos, y de otro lado, existe un conjunto de modelos que incluyen de modo adicional variables referentes al individuo (Sánchez, Gómez y Frasset, 2001).

La consideración en el modelo de variables referentes al individuo como la fidelidad a la marca o la preferencia hacia un determinado tamaño y envase, (véase Guadagni y Little, 1983), permite incorporar la heterogeneidad en las probabilidades de elección considerando una única modelización para toda la muestra. En contraposición a esta opción, se puede recoger la heterogeneidad en las preferencias de los consumidores a nivel de muestra o incluso de individuo a través de una segmentación de la misma. Tradicionalmente se distinguen dos maneras de segmentar un mercado (Kamakura y Wedel, 2000):

- ✓ A priori, donde se divide a los consumidores en función de una serie de características (llamadas normalmente variables descriptoras) y se verifica si los segmentos resultantes son heterogéneos en su comportamiento.
- ✓ A posteriori, que consiste en buscar los segmentos latentes que forman el mercado y que a veces no son fácilmente reconocibles utilizando variables descriptoras previamente definidas.

La segmentación a posteriori, aunque de más difícil estimación, garantiza el cumplimiento del principio fundamental de homogeneidad/heterogeneidad, si bien en los últimos años con el desarrollo del escáner se han desarrollado nuevas metodologías que pueden utilizarse para segmentar el mercado de forma óptima.

Las variables de marketing incluidas en los modelos de elección de marca hacen referencia a aspectos fundamentales, tanto en la toma de decisiones de los fabricantes, como en la toma de decisiones de los intermediarios. Estas variables son fundamentalmente el precio de venta del producto, las promociones de venta, la publicidad o a la disposición en el lineal. La generalización en la utilización de estas variables no es caprichosa, sino que, como posteriormente analizaremos, responde a la tipología informativa ofrecida por el escáner de los terminales de venta de los establecimientos.

De entre estas variables de marketing, la más utilizada y a la que más importancia se le atribuye es el precio de venta del producto, tanto que en muchos modelos es la única variable explicativa indicada (Cavero y Cebollada, 1999; Kamakura y Rusell, 1989; Kim, Blattberg y Rossi, 1995), lo que pone de manifiesto la importancia que se le concede para la explicación de las preferencias de los consumidores.

Sin embargo, además de estas variables es necesario incorporar en los modelos los aspectos dinámicos del comportamiento de elección diseñados para determinar los efectos de las variables explicativas (Gupta, Chintagunta y Wittink, 1997). La incorporación en el modelo de determinados aspectos referentes al comportamiento del consumidor está supeditada a las características de disponibilidad de la base de datos con la que estemos trabajando, lo que viene derivado, fundamentalmente, de la existencia de tarjetas de pago o fidelidad emitidas por el establecimiento de venta.

Se identifican tres dimensiones claves para la especificación de los modelos dinámicos de elección (McAlister, Srivastava, Horwitz, Jones, Kamakura, Kulchitsky, Ratchford, Russell, Sultan, Yai, Weiss y Winer, 1991): la dependencia de estado, que existe cuando las elecciones actuales del individuo están influidas por las elecciones realizadas anteriormente, la heterogeneidad, debida a las diferencias en los individuos en términos de sus preferencias hacia las diferentes alternativas, su sensibilidad a las variables explicativas de la utilidad o a diferencias en características de los individuos y la no estacionariedad del entorno, que consiste en los cambios sistemáticos en la naturaleza de las actividades de marketing de las empresas a lo largo del tiempo. Se toma por supuesto que estos efectos suelen ser constantes en el tiempo (Gupta *et al.*, 1997), si bien este planteamiento puede ser generalizado de modo que pueda especificarse la posición estable o dinámica del consumidor en la distribución de la heterogeneidad (Allenby y Lenk, 1995). Sea como fuere, el interés por la incorporación de este tipo de variables es creciente, derivado de la capacidad que tienen para capturar la heterogeneidad en el comportamiento de elección del consumidor (Sánchez *et al.*, 2001).

3. LA TEORÍA DE LA ELECCIÓN ALEATORIA. EL MODELO LOGIT MULTINOMIAL (MNL): FORMULACIÓN E INTERPRETACIÓN

3.1. Formulación del modelo logit

Los modelos logit pertenecen al tipo de modelos de elección discreta (Greene, 1993; Novales, 1993) o de regresión discreta (Maddala, 1977), y son aquellos en los que la variable dependiente toma valores discretos. La elección de marca, cuando existen más de dos alternativas competidoras, corresponde a una variable policotómica, categórica y no ordenada, siendo en esta situación donde se aplica el modelo logit multinomial (MNL). No obstante, el modelo se puede ampliar para la consideración de observaciones repetidas, es decir para la agregación de las selecciones de un conjunto de consumidores.

El modelo MNL permite estimar la probabilidad de elección de una alternativa determinada en función de una serie de factores explicativos y conocer las cuotas de mercado de las diferentes alternativas competidoras (Ahn y Ghosh, 1989; Currim, 1982). Esto es posible, según Basuroy y Nguyen (1998), porque este tipo de modelo cumple las condiciones establecidas por Naert y Bultez (1973) de consistencia lógica: la cuota de mercado de cada alternativa toma un valor entre cero y uno, y la suma de las cuotas de mercado es igual a la unidad. La

derivación del modelo MNL por McFadden (1974)³⁷ parte del supuesto de que el objetivo del consumidor a la hora de la elección es maximizar su utilidad. Se asume que la aleatoriedad de la elección del individuo entre alternativas discretas puede ser debida al desconocimiento del investigador respecto a las características de las alternativas o del individuo, o ambos. De este modo aunque los consumidores fueran capaces de elegir la alternativa que maximiza la utilidad, los investigadores pueden desconocer los factores que determinan la función de utilidad y por ello hacer predicciones erróneas respecto a la elección del individuo.

De acuerdo con este planteamiento, McFadden (1974) entiende que la utilidad de una alternativa es una función aleatoria, que se puede descomponer en un componente determinista y uno aleatorio:

$$U_{ij} = V_{ij} + \varepsilon_{ij} \quad [1]$$

donde:

U_{ij} = utilidad aleatoria percibida de la alternativa j por el consumidor i

V_{ij} = utilidad asignada por el consumidor i a la alternativa j

ε_{ij} = componente aleatorio de la función de utilidad

McFadden (1974) demuestra como si, y sólo si, se asume que el componente aleatorio ε_{ij} está independiente e idénticamente distribuido y sigue una distribución doble exponencial tipo Weibull, la probabilidad de que un determinado consumidor i elija la alternativa j entre un conjunto de elección compuesto por J alternativas, viene dada por la conocida formulación del modelo logit multinomial:

$$P_{ij} = \frac{\exp(V_{ij})}{\sum_{j=1}^J \exp(V_{ij})} \quad \text{donde } 0 \leq P_{ij} \leq 1 \text{ y } \sum_{j=1}^J P_{ij} = 1 \quad [2]$$

Como hemos mencionado, la particularidad del modelo logit de McFadden (1974) reside en que en la ecuación anterior, la componente determinista de la utilidad, V_{ij} , es una función lineal de los atributos de la alternativa j , siendo una especificación del mismo muy frecuente en sociometría la que incluye interacciones de los atributos y de variables dummy específicas de las alternativas:

$$V_{ij} = \beta_{0j} x_j + \sum_{v=1}^V \beta_v x_{vj} \quad [3]$$

donde:

β_{0j} = Coeficiente específico de la alternativa j

x_j = Variable dummy específica de la alternativa, que toma el valor 1 para la alternativa j y 0 para el resto

β_v = Parámetro que indica la sensibilidad del consumidor hacia el atributo v -ésimo

x_{vj} = Valor del atributo v -ésimo que describe la alternativa j

Por tanto, la formulación del modelo incorporando los atributos determinantes de la utilidad, vendrá dada por la ecuación:

³⁷ El modelo MNL propuesto por McFadden (1974), que el autor denomina modelo logit condicional, difiere del modelo MNL general en que el primero considera como variables explicativas, características de las alternativas de elección, mientras que el modelo MNL general considera únicamente características de los individuos (Greene, 1993; Maddala, 1977).

$$P_j = \frac{e^{\beta_{0j}x_j + \sum_{v=1}^V \beta_v x_{vj}}}{\sum_{j=1}^J e^{\beta_{0j}x_j + \sum_{v=1}^V \beta_v x_{vj}}} \quad [4]$$

Así pues, la especificación del modelo según la ecuación anterior, consta de dos tipos de parámetros: aquellos específicos para cada alternativa (β_{0j}) y aquellos específicos para cada atributo (β_v). Los primeros acompañan a las variables *dummy*, o constantes específicas de las alternativas, y recogen el efecto medio de la omisión de variables relevantes en la función de utilidad (Ben-Akiva y Lerman, 1985; Liao, 1994). Se recomienda su inclusión en el modelo para evitar sesgos en las estimaciones de los parámetros β_v (Currim, 1982; Gautschi, 1981), interpretándose su valor como la diferencia entre la utilidad de dos alternativas cuando todos los demás atributos se mantienen constantes. Los segundos, que acompañan a los V atributos, recogen, según Maddala (1977), la valoración o importancia relativa de los atributos considerados influyentes en la elección, siendo, de acuerdo con Finn y Louviere (1990), los que realmente tienen importancia estratégica en la modelización. Por tanto, el número de parámetros a estimar (B) es igual al número de atributos más el número de alternativas menos uno: B= (V+J-1).

Respecto al método de estimación de los parámetros del modelo MNL, el de máxima verosimilitud resulta ser el que da lugar a estimadores más adecuados (Cadwallader, 1996; Gracia, 1988; Greene, 1993; Maddala, 1977; Malhotra, 1984). La aplicación de dicho método no permite incluir una variable *dummy* para todas las alternativas, por lo que se debe imponer una condición de normalización, haciendo $\beta_{0k}=0$, para $j=k$. De este modo, la alternativa cuyo término constante se suprime actúa como referencia para el resto de alternativas. Se obtienen así (J-1) ecuaciones que determinan de forma única las probabilidades de elección y garantizan que su suma sea igual a la unidad.

En el modelo logit multinomial, el valor de los parámetros no tiene una interpretación directa, ya que recoge el efecto marginal de la variable en la transformación logística de la función de utilidad, pero no directamente en la probabilidad de selección de una alternativa (Greene, 1993; Liao, 1994). El efecto marginal en dicha probabilidad, se obtiene derivando P_j respecto a x_{vj} (variable asociada a la alternativa j) y a x_{vk} (variable asociada a la alternativa $k \neq j$):

$$\frac{\partial P_j}{\partial x_{vj}} = P_j(1 - P_j)\beta_v \quad [5]$$

$$\frac{\partial P_j}{\partial x_{vk}} = -P_j \cdot P_k \cdot \beta_v \quad [6]$$

Lo cual se puede expresar en términos de elasticidades del siguiente modo:

$$\varepsilon_{P_j, x_{vj}} = (1 - P_j)x_{vj}\beta_v \quad \forall j = 1 \dots J \quad [7]$$

$$\varepsilon_{P_j, x_{vk}} = -P_j x_{vk} \beta_v \quad \forall k \neq j \quad [8]$$

Correspondiendo [7] a la elasticidad directa y [8] a la elasticidad cruzada. Por consiguiente, el efecto marginal o la elasticidad de las variables explicativas en la probabilidad de selección de una determinada alternativa, depende tanto de los coeficientes estimados como de los valores que toman las variables explicativas.

3.2. Ventajas e inconvenientes del MNL

A partir del análisis de la formulación del modelo MNL, diversos autores han puesto de manifiesto su eficacia para analizar la selección de alternativas competidoras. Así, a partir del trabajo de Gensch y Recker (1979), que compara el modelo logit frente al modelo lineal para la selección de establecimiento comercial, otros autores como Agrawal y Schorling (1996), Arnold, Roth y Tigert (1980), Currim (1982), Fotheringham (1988), Jain, Acito, Malhotra y Mahajan (1979), Malhotra (1984), y Van Ophem, Stam y Van Praag (1999), han evaluado empíricamente el modelo MNL frente a otras alternativas metodológicas. Las conclusiones de estos trabajos subrayan los beneficios del modelo MNL, justificando así su aplicación a nuestro objeto de estudio:

- ✓ Se deriva de una teoría sobre el comportamiento del consumidor, lo cual da solidez y consistencia a las conclusiones derivadas de su estimación.
- ✓ El conjunto de elección puede ser diferente para cada individuo de la muestra.
- ✓ La estimación de los parámetros considera las diferencias en los atributos de las alternativas de elección, aspecto especialmente relevante en la evaluación de la posición competitiva.
- ✓ Las elasticidades consideran la probabilidad de selección de una alternativa y el nivel de presencia del atributo, siendo así consecuentes con el principio de utilidad marginal decreciente y de saturación.
- ✓ Es eficiente incluso con muestras pequeñas ($N > 50$).

3.3. Criterios de bondad del ajuste en los modelos MNL

Al no ser lineales, los modelos logit requieren procedimientos distintos a los modelos lineales para analizar la bondad del ajuste. Malhotra (1984) recomienda utilizar más de un criterio para evaluar el poder predictivo del modelo, así como evaluar la bondad del ajuste en términos relativos y no absolutos, es decir examinando distintas especificaciones del modelo. Estos criterios de comparación de los modelos deben considerar tanto la capacidad explicativa del modelo como el grado de parsimonia o sencillez del mismo (Cavero y Cebollada, 1999). Entre los métodos y criterios disponibles, destacamos los siguientes (Elorz, 1997; Guadagni y Little, 1983; Malhotra, 1984; Novales, 1993):

1.- *Cuota de mercado real frente a cuota de mercado predicha.* La agregación de las probabilidades de elección para la muestra representa la cuota de mercado esperada de dicha alternativa. La comparación, como sugiere Vázquez (1990), a través del análisis de la correlación entre dichas cuotas y las reales a lo largo del tiempo proporciona una representación visual de la bondad del ajuste.

2.- *Significatividad y signo de los parámetros estimados.* Es requisito previo antes de discutir el signo de un parámetro, comprobar si es significativamente distinto de cero, para lo cual es útil el test de la t de Student. Este test se emplea pues, según Arnold *et al.* (1980), para contrastar la determinancia de un atributo, a través de la hipótesis nula de que su coeficiente es cero. Si se rechaza la hipótesis nula, se discute a continuación si el signo de los parámetros es el esperado de acuerdo con la literatura o con los resultados de investigaciones previas.

3.- *Valor de la función de verosimilitud.* La estimación de los modelos logit busca maximizar el valor de la función de verosimilitud $LL(\beta)$, por ello, en la elección de especificaciones alternativas, un primer criterio es el de máxima función de verosimilitud del modelo estimado.

4.- *Indice de la razón de verosimilitud.* Indica el grado de variabilidad del modelo que es explicado por los coeficientes del mismo, teniendo así un significado similar al coeficiente de determinación R² en el modelo lineal. Sin embargo, como indica Novales (1993:547), este coeficiente es difícil de interpretar, porque “ni llega a valer 1 ni es fácil interpretar los valores entre 0 y 1”. En cuanto al valor del estadístico que se considera aceptable, McFadden (1978) indica que un coeficiente entre 0,2 y 0,4 es indicador de un ajuste satisfactorio del modelo. La expresión del estadístico es:

$$\rho^2 = 1 - \frac{LL(\beta)}{LL(C)} \quad [9]$$

Tal y como indican Ahn y Ghosh (1989), el valor del estadístico aumenta con la inclusión de parámetros adicionales, por lo que se recomienda para el contraste del modelo, la corrección del coeficiente con el número de parámetros a estimar:

$$\bar{\rho}^2 = 1 - \frac{LL(\beta) - k}{LL(C)} \quad [10]$$

5.- *Contraste de la razón de verosimilitud.* Permite comprobar que el ajuste del modelo estimado es mejor que el de un modelo restringido en los parámetros, es decir, que incluya un número menor de variables explicativas. Se hace operativo a través de un estadístico que compara el valor de ambas funciones de verosimilitud, y que se distribuye como una chi-cuadrado con BNR-BR grados de libertad, siendo BNR y BR, el número de parámetros a estimar en el modelo no restringido y restringido, respectivamente:

$$RV = -2 [LL(\beta_{NR}) - LL(\beta_R)] \quad [11]$$

El contraste de la razón de verosimilitud puede emplearse para contrastar diferentes hipótesis, siendo fundamental el contraste de la significación global de modelo (Novales, 1993), es decir, la hipótesis nula H₀: $\beta=0$, calculando el valor LL(0) que se obtiene cuando todos los parámetros simultáneamente son cero. Si se rechaza la hipótesis nula, significa que el modelo estimado aporta información adicional respecto al modelo nulo, o de cuotas de mercado idénticas.

6.- *Criterios de información teórica.* Este tipo de estadísticos consideran que es mejor aquel modelo que combina un bajo número de parámetros con un alto valor de la función de verosimilitud (Elorz, 1997). Por lo tanto, según indica Malhotra (1984), son especialmente indicados para contrastar modelos con diferentes grados de libertad. Entre estos criterios teóricos, destacan el Criterio de Información de Akaike (AIC) y el Criterio de Información Bayesiana de Schwarz (SBIC), definidos respectivamente como:

$$\begin{aligned} AIC &= -2 [LL(\beta) - B] \\ SBIC &= -2 [LL(\beta) - B/2 \ln(n)] \end{aligned} \quad [12]$$

donde:

n= tamaño muestral

4. METODOLOGÍA DE LA INVESTIGACIÓN

4.1. Selección de la categoría de producto y del tipo de datos

Los datos utilizados para esta investigación, han sido recogidos en un hipermercado de Almería capital. Se trata de datos de escáner de elección de marcas de aceite de oliva 0,4° en el formato de botellas de 1 litro. La elección de la categoría del aceite de oliva se fundamenta, primero en la importancia que esta categoría de producto tiene en el contexto económico y social de Andalucía, y, posteriormente, en la elevada frecuencia y utilización que este producto tiene en los hogares españoles, lo que permite contar con un número de datos suficiente como para realizar un estudio de este tipo. La selección del aceite 0,4° se realizó por ser el más utilizado, normalmente, en la preparación de comidas por parte del consumidor, si bien es cierto, que existen otras graduaciones también importantes. La elección del formato de 1 litro se debe a que, junto a la garrafa de 5 litros, se trata del envase más habitual. Estos datos de compra abarcan un período total de 53 semanas. El conjunto de elección está constituido por las diez marcas con las que el hipermercado contaba en aquel período. En la tabla 1, podemos observar cuál es la cuota de mercado de las mismas durante el año 2002.

<i>CARBONELL</i>	26,67
<i>MARCA DE DISTRIBUIDOR</i>	22,38
<i>COOSUR</i>	16,34
<i>LA MASÍA</i>	11,87
<i>LA ESPAÑOLA</i>	9,83
<i>ELOSUA</i>	4,01
<i>LA GIRALDA</i>	3,87
<i>YBARRA</i>	2,45
<i>MARCA BLANCA</i>	2,17
<i>MUELOLIVA</i>	0,4

Tabla 1. Cuotas de mercado de las marcas utilizadas

Una vez escogidas las marcas, hemos de seleccionar a los clientes que vamos a incluir en nuestra investigación. Para ello, y debido a la necesidad de seguir el comportamiento del consumidor a lo largo de las 53 semanas analizadas, sólo pudimos considerar a aquellos consumidores que realizaron su pago con tarjeta del establecimiento, no considerando aquellos consumidores que pagaron al contado o los que, teniendo tarjeta del establecimiento, hicieron alguna de sus compras igualmente al contado.

Por otra parte, y ya que queremos analizar la importancia del precio de venta en relación al nivel de lealtad en la elección del consumidor, fue necesario dividir el período total considerado. Esta división es un aspecto complejo en el que no existe unanimidad en la forma de selección de los hogares a utilizar. En la diferente y abundante literatura existente, que utiliza datos de escáner (p.e. Bellizzi, Hamilton, Krueckeberg y Martin, 1981; Bucklin *et al.*, 1995; Chintagunta, 1994; Guadagni y Little, 1983; Kamakura y Russell, 1989) no existe una metodología clara a la hora de realizar esta selección, el tipo de modelo a utilizar, las variables a emplear, los períodos en los que dividir los datos, etc.

Nosotros, siguiendo a Sivakumar y Raj (1997), hemos dividido las 53 semanas en dos períodos: un primer período de inicialización, que nos va a servir para calcular la variable de lealtad, y que abarca de la semana 1 a la 20 (ambas inclusive), y un período de estimación, en el que se estiman los parámetros del modelo, y que comprendería el número restante de semanas. Una vez realizada esta división, contemplaremos, únicamente, aquellos hogares que hayan realizado, al menos, dos compras en cada uno de los períodos. De este modo, el conjunto definitivo de datos de escáner que se recogieron está constituido por 389 hogares, que realizan un total de 3241 compras (8,33 compras por hogar), de las que 1427 compras fueron utilizadas para la inicialización de las variables de lealtad, mientras que el resto, 1814, fueron utilizadas para la estimación de los parámetros del modelo.

4.2. Hipótesis

Como hemos comentado con anterioridad, pretendemos conocer cuáles son las variables que influyen en el comportamiento de compra del consumidor. Dentro de esas variables, hemos introducido variables de precio, tanto precio de compra como precio de referencia, así como variables de lealtad. En base a estas variables, formulamos las siguientes hipótesis:

H1: Existe una relación inversa y significativa entre el precio de venta de una alternativa y su probabilidad de compra

H2: Existe una relación positiva y significativa entre la lealtad del individuo hacia una alternativa y la probabilidad de compra de la misma

H3: La lealtad a la marca tiene una mayor influencia que el precio en la decisión de compra de los consumidores

4.3. Variables analizadas

Las variables utilizadas en nuestro análisis han sido las siguientes:

- **Precio de compra (PCOM_ACT_{jt}):** Precio de compra de la alternativa j en el período (semana) t
- **Lealtad histórica (LEAL_HIS_{ijt}):** Fidelidad del individuo i a la alternativa j en el período t , considerada como la lealtad histórica.
- **Lealtad anterior (LEAL_INM_{ijt}):** Fidelidad del individuo i a la alternativa j en el momento t , considerada como fidelidad inmediata.

Con objeto de contrastar las hipótesis planteadas anteriormente, se han analizado dos modelos alternativos:

- **Modelo 1:** En el que únicamente se contrasta el efecto del precio sobre la probabilidad de elección de las diferentes alternativas
- **Modelo 2:** En el que se incorporan, las variables de lealtad, con objeto de conocer en qué medida son más explicativas del comportamiento de elección.

Para ambos modelos se han estimado los coeficientes de los atributos analizados, así como sus correspondientes parámetros de significación. A partir de estos análisis se aceptan o rechazan las diferentes hipótesis planteadas.

CITIES IN COMPETITION

En ambos casos, se han estimado, igualmente, los coeficientes de cada una de las alternativas que componen el conjunto de elección.

Modelo 1: $U_{ijt} = \beta_0 X_j + \beta_j PCOM_ACT_{jt}$

Modelo 2: $U_{ijt} = \beta_0 X_j + \beta_j PCOM_ACT_{jt} + \beta_2 LEAL_HIS_{ijt} + \beta_3 LEAL_ANT_{ijt}$

5. RESULTADOS

En la tabla 2, podemos observar los resultados obtenidos en la estimación de cada uno de los modelos.

	<i>Modelo 1</i>	<i>Modelo 2</i>
<i>Constantes específicas de las alternativas</i>		
Carbonell	6,404*	4,407*
Marca de distribuidor	3,063***	0,687(n.s.)
Coosur	4,316*	2,042*
La Masía	3,073*	1,589*
La Española	3,549*	2,131*
Elosúa	2,711*	1,074**
Giralda	1,15**	-1,247***
Ybarra	2,797*	1,502**
Marca Blanca	0,266 (n.s.)	-1,695***
Mueloliva	fijo	fijo
<i>Coefficientes de los atributos de las alternativas</i>		
Precio de compra	-5,981*	-8,319*
Fidelidad histórica	-	1,273*
Fidelidad anterior	-	4,893*
Parámetros de bondad del ajuste		
Nº de parámetros	10	12
<i>LL (β)</i>	-6637,51	-2280,005
χ^2	178,41	8893,429
ρ^2	0,013	0,661
ρ^2 ajust.	0,011	0,659

*p<0,001 **p<0,05 ***p<0,1 n.s. no significativo

Tabla 2. Parámetros estimados de los modelos logit

Las constantes específicas de las alternativas, de acuerdo con Guadagni y Little (1983), representan cualidades únicas de la alternativa, o bien recogen errores de especificación, reflejando la utilidad relativa de cada marca respecto a las otras, siempre que el resto de variables o factores sean iguales. Su valor es relativo respecto a una

alternativa que se pone de referencia, y que suele ser aquella de menor cuota de mercado, por ser la menos representativa dentro del conjunto de elección. De acuerdo con la interpretación de Elorz (1997), un valor positivo de β_{0j} indica que, en igualdad de condiciones del resto de variables, la alternativa j aportaría a los consumidores mayor valor que la alternativa tomada como base, que en este caso corresponde a la alternativa número 10 (*Mueloliva*), que es la que tiene menor cuota de mercado. Como podemos observar en la tabla 2, vemos como la alternativa *Carbonell* sería la de mayor valor intrínseco en el conjunto de alternativas contempladas (siendo esto cierto para todos los modelos). Esto indica que la imagen de marca que posee *Carbonell* es la mejor posicionada en la mente del consumidor, por lo que, en igualdad de condiciones del resto de variables, esta será la marca que el consumidor escoja ya que es hacia la que presenta una mayor preferencia. Igualmente, en la tabla 2 podemos observar que el precio de compra presenta coeficientes negativos, que indica una relación inversa entre el precio de una alternativa y su probabilidad de elección, siendo, por lo general, valores elevados, lo que indica una fuerte e intensa relación inversa. Por otra parte, esta relación es significativa, por lo que se puede **aceptar** la hipótesis **H1**.

Ahora bien, las variables que toman una importancia fundamental son las variables de lealtad. En este sentido, y como se deriva de los resultados que se muestran en la tabla 2., tanto la fidelidad histórica como la anterior, presentan valores positivos, siendo en el caso de la fidelidad anterior, valores más elevados (4,893). Esto es coincidente con trabajos como el de Cavero y Cebollada (1997), donde se le otorga una importancia elevada a la lealtad del consumidor hacia las diferentes marcas existentes, frente a otro tipo de variables que puedan influir en su comportamiento de elección. Nuestros resultados indican que la elección de aceite de oliva está fuertemente influenciada por comportamientos de lealtad y que el consumidor es poco proclive al cambio de marca, siendo incluso, mal vista la reducción del precio de venta y afectando a su imagen de marca (así por ejemplo, la marca con mayor nivel de preferencia, *Carbonell*, no realiza, prácticamente, ninguna promoción en precio durante las 53 semanas contempladas). Esta lealtad es más elevada a corto plazo que en el medio y largo plazo (lo que viene indicado, además de por la diferencia de valor entre los coeficientes de ambas variables porque la lealtad anterior tiene un valor del estadístico $t = 41,974$ frente al valor de la variable de lealtad ($t = 13,213$), por lo que el consumidor se ve muy influenciado por la elección que ha realizado en las ocasiones más recientes en las que ha visitado el establecimiento comercial y ha seleccionado la categoría de producto en cuestión. Estos resultados otorgan una elevada importancia a los programas de fidelización que realizan estos establecimientos, y justifican, en gran medida, la notable importancia que está tomando la utilización de tarjetas de fidelización por parte del individuo y el desarrollo, por parte del establecimiento, de condiciones de compra y de uso más favorables para los poseedores de la misma. Es por todo ello, que se puede **aceptar** la hipótesis **H2**.

Si analizamos los parámetros de bondad del ajuste para cada uno de los modelos, podemos observar como, en el modelo 1, la capacidad de predicción de la variable precio es muy reducida, lo que nos lleva a pensar que la variable precio no tiene capacidad por sí sola, para explicar fielmente el comportamiento de elección del consumidor. Sin embargo, y cuando se incorporan las dos variables de lealtad en el modelo, ρ^2 presenta un valor de 0,661, lo que supone un incremento sustancial en dicho valor respecto al primer modelo. Esto viene a indicar la gran capacidad de explicación que tienen estas variables sobre el comportamiento de elección, en detrimento de la variable precio, por lo que se puede **aceptar** la hipótesis **H3**.

6. CONCLUSIONES

Una de las etapas fundamentales en el proceso de compra del consumidor, y a la que mayor atención se le ha prestado en la literatura de marketing, es la elección. El conocimiento de los factores que influyen en este proceso, va a ayudar, tanto a las empresas minoristas como a las empresas mayoristas, a optimizar su presupuesto comercial, conociendo qué variables del marketing-mix son las más importantes, y a las que mayor

atención deben prestarle. En este sentido, en este trabajo se pone de manifiesto como, si bien el precio de compra sigue siendo importante para el consumidor, los niveles de lealtad del individuo son cada vez más intensos, lo que indica la necesidad de las empresas minoristas de incrementar y afianzar esa lealtad mediante programas efectivos de marketing relacional.

BIBLIOGRAFÍA

- AGRAWAL, D. Y SCHORLING, C. (1996): "Market share forecasting: an empirical comparison of artificial neural networks and multinomial logit model", *Journal of Retailing*, Vol. 72, nº 4, pp. 383-407.
- AHN, K. Y GHOSH, A. (1989). "Hierarchical models of store choice", *International Journal of Retailing*, Vol. 4, nº 5, pp. 39-52.
- ALLENBY, G. Y LENK, P. (1995): "Reassessing brand loyalty, price sensitivity and merchandising effects on consumer brand choice", *Journal of Business & Economics Statistics*, 13, (3), pp. 281-289.
- ARNOLD, S., ROTH, V. Y TIGERT, D. (1980): "Conditional logit versus MDA in the prediction of store choice", *Advances in Consumer Research*, Vol. 20, pp. 665-670.
- BASUROY, S. Y NGUYEN, D. (1998): "Multinomial logit market share models: equilibrium characteristics and strategic implications", *Management Science*, Vol. 44, nº 10, pp. 1396-1408.
- BELLIZZI, J., HAMILTON, J., KRUECKEBERG, H. Y MARTIN, W. (1981): "Consumer perceptions of national, private and generic brands", *Journal of Retailing*, Vol. 57, nº 4, pp. 56-70.
- BEN AKIVA, M. Y LERMAN, S. (1985): *Discrete Choice Analysis. Theory and Application to Travel Demand*. Cambridge, MA.: MIT Press.
- BENTZ, Y. Y MERUNKA, D. (2000): "Neural networks and the multinomial logit for brand choice modelling: a hybrid approach", *Journal of Forecasting*, Vol. 19, pp. 177-200.
- BUCKLIN, R., GUPTA, S. Y HAN, S. (1995): "A brand's eye view of response segmentation in consumer brand choice behavior", *Journal of Marketing Research*, Vol. XXXII, nº Febrero, pp. 66-74.
- CADWALLADER, M. (1996): *Urban geography. An analytical approach*, Prentice Hall.
- CAPPS Jr., O. (1989): "Utilizing scanner data to estimate retail demand functions for meat products", *American Agricultural Economics Association*, Vol. Agosto, pp. 750-760.
- CAVERO, S. Y CEBOLLADA, J.(1997): "Las estrategias de captación y retención de clientes y la lealtad. Un análisis de segmentación", *Revista Española de Investigación de Marketing Esic*, Vol.1, nº Septiembre, pp. 55-71.
- CAVERO, S. Y CEBOLLADA, J. (1999): "Análisis de la competencia entre tipos de marca. Una aplicación empírica" *Información Comercial Española*, Vol.79, nº Julio-Agosto, pp. 55-70.
- CHINTAGUNTA, P. (1994): "Heterogeneous logit model implications for brand positioning", *Journal of Marketing Research*, Vol. XXXI, nº Mayo, pp. 304-311.
- CHRISTEN, M., GUPTA, S., PORTER, J., STAELIN, R. Y WITTINK, D. (1997): "Using market-level data to understand promotion effects in a nonlinear model", *Journal of Marketing Research*, Vol. XXXIV, nº Agosto, pp. 322-334.
- CURRIM, I. (1982): "Predictive testing of consumer choice models not subject to independence of irrelevant alternatives", *Journal of Marketing Research*, Vol. XIX, nº Mayo, pp. 208-222.
- ELORZ, M. (1997): *El Valor de las Marcas y la Gestión Minorista de la Categoría de Producto*. Tesis Doctoral no publicada. Universidad Pública de Navarra.
- FINN, A. Y LOUVIERE, J. (1990): "Shopping center patronage models: fashioning a consideration set segmentation solution", *Journal of Business Research*, Vol. 21, nº Noviembre, pp. 259-275.
- FOTHERINGHAM, A. (1988): "Consumer store choice and choice set definition", *Marketing Science*, Vol. 7, nº 3, pp. 299-310.
- GAUTSCHI, D. (1981): "Specification of patronage models for retail center choice", *Journal of Marketing Research*, Vol. XVIII, nº Mayo, pp. 162-174.
- GENSCH, D. Y RECKER, W. (1979): "The multinomial, multiattribute logit choice model", *Journal of Marketing Research*, Vol. XVI, nº Febrero, pp. 124-32.
- GRACIA, M. (1988): "Modelos con variable dependiente cualitativa y de variación limitada", *Información Comercial Española*, Vol. 39, nº 2, pp. 7-49.
- GREENE, W.H. (1993): *Econometric Analysis*. Prentice Hall, Londres.
- GUADAGNI, P. Y LITTLE, J. (1983): "A logit model of brand choice calibrated on scanner data", *Marketing Science*, Vol. 2, nº 3, pp. 203-238.
- GUPTA, S., CHINTAGUNTA, P. Y WITTINK, D. (1997): "Household heterogeneity and state dependence in a model of purchase strings: empirical results and managerial implications", *International Journal of Research in Marketing*, Vol. 14, pp. 341-57.
- HILDEBRANDT, L. Y KLAPPER, D. (2001): "The analysis of price competition between corporate brands", *International Journal of Research in Marketing*, Vol. 18, pp. 139-159.

NEW TRENDS IN MARKETING MANAGEMENT

- JAIN, A., ACITO, F., MALHOTRA, N. Y MAHAJAN, V. (1979): "A comparison of the internal validity of alternative parameter estimation methods in descompositional multiattribute preference models", *Journal of Marketing Research*, Vol. XVI, n° Agosto, pp. 312-322.
- KAMAKURA, W. Y RUSSELL, G. (1989): "A probabilistic choice model for market segmentation and elasticity structure", *Journal of Marketing Research*, Vol. XXVI, n° Noviembre, pp. 379-390.
- KAMAKURA, W. Y WEDEL, M. (2000): "Factor analysis and missing data", *Journal of Marketing Research*, Vol. 37, n° Noviembre, pp. 490-498.
- KIM, B., BLATTBERG, R.C. Y ROSSI, P.E. (1995): "Modeling the distribution of price sensitivity and implications for optimal retail pricing", *Journal of Business & Economic Statistics*, Vol. 13, n° 3, pp. 291-303.
- LIAO, T. (1994): *Interpreting probability models. Logit, probit and other generalized linear models*. Sage Publications, U.S.A.
- MADDALA, G.S. (1977): *Econometrics*, McGraw-Hill.
- MALHOTRA, N. (1984): "The use of linear logit models in marketing research", *Journal of Marketing Research*, Vol. XXI, n° Febrero, pp. 20-31.
- MCALISTER, L., SRIVASTAVA, R., HORWITZ, J., JONES, M., KAMAKURA, W., KULCHITSKY, J., RATCHFORD, B., RUSSELL, G., SULTAN, F., YAI, T., WEISS, D. Y WINER, R. (1991): "Incorporating choice dynamics in models of consumer behavior", *Marketing Letters*, Vol. 2, pp. 241-252.
- McFADDEN, D. (1974): "Conditional logit analysis of qualitative choice behavior", en ZAREMBKA, P. (ed.), *Frontiers in econometrics*, Academic Press, New York, pp. 105-142.
- McFADDEN, D. (1978): "Modelling the Choice of Residential Location", en KARLQUIST, A., LUNDQVIST, L., SNICKARS, F. Y WEIBULL, J.W. (eds.): *Spatial Interaction Theory and Residential Location, Studies in Regional Science and Urban Economics*, Vol. 3, North Holland, Amsterdam, pp. 75-96.
- NAERT, PH. Y BULTEZ, A. (1973): "Logically consistent market share models", *Journal of Marketing Research*, Vol. X, n° Agosto, pp. 334-340.
- NOVALES, A. (1993): *Econometría*. McGraw-Hill, Madrid.
- PAPATLA, P. Y KRISHNAMURTHI, L. (1996): "Measuring the dynamic effects of promotions on brand choice", *Journal of Marketing Research*, Vol. XXXIII, n° Febrero, pp. 20-35.
- ROSSI, P. Y ALLENBY, G. (2000): "Statistics and marketing", *Working Paper*, Graduate School of Business, Universidad de Chicago.
- SÁNCHEZ, M., GÓMEZ, M.A. Y FRASQUET, M. (2001): "Modelización de la elección del consumidor con datos de escáner: logit multinomial y logit anidado", Seminario "El uso de datos de escáner en empresas minoristas", Albacete, 4 y 5 de Junio.
- SIVAKUMAR, K. Y RAJ, S. (1997): "Quality tier competition: how price change influences brand choice and category choice", *Journal of Marketing*, Vol. 61, n° Julio, pp. 71-84.
- VAN OPHEM, H., STAM, P. Y VAN PRAAG, B. (1999): "Multichoice logit. Modeling incomplete preference rankings of classical concerts", *Journal of Business & Economics Statistics*, Vol. 17, n° 1, pp. 117-128.
- VÁZQUEZ, R. (1990): "Aplicaciones potenciales de los modelos de cuota de mercado a partir de la información generada por un scanner", *II Encuentro de Profesores de Marketing*, Universidad de Valencia: Valencia, 4-5 Octubre