Estudios de Economía Aplicada
Nº 10, 1998. Págs. 131-146

Comparación de tres métodos de estimación del análisis conjunto: diferencias en las preferencias en el consumo de vino y en la segmentación del mercado

SÁNCHEZ, M.[1]
Universidad Pública de Navarra
GIL, JOSÉ M.[2]
SIA-DGA. Unidad de Economía Agraria

Esta versión incluye todas las correcciones sugeridas por el evaluador, las cuales nos han parecido oportunas y por las que le quedamos muy agradecidos.

RESUMEN

El Análisis Conjunto se ha convertido en una técnica habitual en la estimación de las utilidades asignadas por los consumidores a los distintos atributos que conforman un producto. Las etapas requeridas en el desarrollo de este instrumento incluyen el diseño, la administración del experimento, la especificación y estimación del modelo conjunto y la evaluación de los resultados hallados. En este trabajo, se ha diseñado un experimento con el fin de analizar las diferencias, tanto en las preferencias como en la segmentación del mercado del vino, obtenidas dependiendo del método de estimación seleccionado que a su vez depende de la consideración que se haya dado a la variable dependiente (de tipo numérico o de tipo ordinal). Las diferencias en los resultados de las distintas elecciones se han estudiado mediante tres métodos de estimación: 1) estimación por Mínimos Cuadrados, 2) estimación por Máxima Verosimilitud de un modelo Tobit Doblemente Censurado y 3) modelo Logit Ordenado. La información se ha obtenido mediante una encuesta personal dirigida a los compradores habituales de vino en Navarra. Los atributos seleccionados como más relevantes en la adquisición

1. Profesora titular de Universidad. Universidad Pública de Navarra. Departamento de Gestión de Empresas. Campus de Arrosadía, 31006 Pamplona. Telf.: 948 16 93 96. Fax: 948 16 94 04. e-mail: mersan@si.upna.es.
2. Investigador científico. SIA-DGA Unidad de Economía Agraria. Apdo. 727, 50080 Zaragoza. Telf.: 976 57 63 61. Fax: 976 57 55 01. e-mail: jmgil@posta.unizar.es.

de estos productos han sido el precio, el origen y el tipo de vino. Los resultados observados ofrecen similares estructuras de preferencias en la compra de los vinos, así como clasificaciones del mercado muy homogéneas entre sí.

Palabras clave: mercado vitivinícola, preferencias del consumidor, análisis conjunto, tobit doblemente censurado, logit ordenado.

SUMMARY

Conjoint analysis has become an increseasingly popular approach to estimate the benefits received from the attributes of a product. The specific steps in conjoint analysis include conjoint desing, administration of the conjoint experiment, conjoint model specification and estimation, and result's evaluation. In this paper, we have concentrated on the consequences of choosing different methods of estimating the conjoint model. Three alternative methods of estimation have been considerend which differs in the way the dependent variable is considered: 1) least squeares estimation of a linear regression, 2) maximum likelihood estimation of a two-limit Tobit model; and 3) maximum likelihood estimation of an Ordered Logit model. Estimed parameters from these approaches are compared and consequences on wine consumer preferences and wine market segmentation are discussed. Data are obtained from a survey camet out in Navarra. Respondents were the main purchasers of wine in a particular household. Three wine attributes have been used in the conjoint experiment desing: price, origin and grape vintage. Results shown that more less the three methods provide similar structures of consumer preferences and market segments.

Key words: wine market, consumer preferences, conjoint analysis, two-limit tobit model, ordered logit model.

Artículo recibido en diciembre de 1997. Revisado en mayo de 1998.

1. Introducción

La saturación de los mercados vitivinícolas ha originado una eficiencia menor de las políticas comerciales indiferenciadas. La mejora de las calidades ofertadas y la presencia de marcas, individuales o colectivas, se han convertido en elementos estratégicos para competir en estos mercados. Entre otras razones, la reducida dimensión de las empresas que conforman este subsector en España ha motivado un importante desarrollo de las marcas colectivas, diferenciando las producciones según sus regiones de origen. Esta situación ha justificado el incremento de la relevancia de la etapa de la evaluación de las alternativas en la decisión de compra de estos productos. A la hora de evaluar las diferentes estrategias, las preferencias del consumidor juegan un papel destacado. Entre los métodos existentes para analizar la estructura de preferencias del consumidor el Análisis Conjunto, ocupa un lugar importante sobre todo en los últimos años. Este método descomposicional parte de dos supuestos básicos: 1) una definición multiatributo del producto y 2) una importancia real de estos atributos en la decisión de adquisición realizada por el comprador (Múgica, 1989).

Las etapas incluidas en la aplicación de la técnica del análisis conjunto abarcan: diseño y administración del experimento, especificación y estimación del modelo conjunto, e interpretación de los resultados. El diseño incluye fundamentalmente la definición de los atributos y de los niveles de cada atributo más relevantes en la elección analizada. La administración se ocupa de la selección de los hipotéticos productos que van a ser evaluados por los entrevistados. La especificación y estimación del modelo trata de seleccionar los métodos más adecuados para la obtención de las preferencias. Esta elección depende, fundamentalmente, de la consideración por parte del investigador de la variable dependiente o preferencia manifestada por el consumidor. La principal clasificación divide las opciones entre numéricas y categóricas o no numéricas. Este trabajo se va a centrar en la contrastación de las diferencias existentes en los resultados, tanto en las estructuras de preferencias como en la clasificación o segmentación del mercado, en base a la consideración que se realice sobre la escala de medición empleada (numérica o no numérica). Las implicaciones de está decisión radican fundamentalmente en el método que se emplea para estimar el modelo conjunto.

La información utilizada se refiere a las preferencias manifestadas por los compradores navarros de vino en sus elecciones de compra. Se ha obtenido mediante una entrevista personal realizada a 282 compradores, lo que ha supuesto incurrir en un error máximo del 5 por cien[1]. Los atributos seleccionados como más relevantes

1. De forma previa se ha realizado una encuesta piloto, dónde se observó un 80 por cien de compradores de vino en el territorio analizado.

en la adquisición de este producto fueron el precio, el origen geográfico de su producción y el tipo o año de elaboración. Estos aspectos fueron seleccionados mediante la realización de un estudio cualitativo previo. La escala utilizada en la valoración de las combinaciones de dichos atributos ha tomado valores entre 1 y 10, siendo 10 la máxima preferencia. Se han considerado los siguientes métodos de estimación: 1) Mínimos Cuadrados, 2) Tobit Doblemente Censurado y 3) Logit Ordenado. Los dos primeros se basan en la consideración de las preferencias como numéricas, mientras que en el tercero se consideran categóricas.

El trabajo ha sido estructurado en tres apartados adicionales. En el apartado 2 se describe la metodología utilizada en el estudio, prestando especial atención al diseño del análisis conjunto y a una breve exposición de las distintas técnicas de estimación seleccionadas. El apartado tercero recoge los resultados obtenidos mediante las distintas alternativas de estimación, así como las consecuencias de la utilización de cada método en el análisis de la estructura de las preferencias de los consumidores y en las segmentaciones del mercado obtenidas a partir de dichas preferencias. Finalmente, se incluyen las principales conclusiones del estudio.

2. Metodología

El método del Análisis Conjunto se ha convertido en una importante herramienta en la evaluación de las preferencias asignadas por un comprador a los distintos atributos que lo componen (Ruiz de Maya y Munuera, 1993). Las utilidades asignadas a cada uno de los aspectos relevantes en la compra del producto, se obtienen mediante la descomposición de las valoraciones globales realizadas por los sujetos entrevistados sobre los diferentes productos hipotéticos que se les ofrecen (Green y Srinivasan, 1978). El interés de esta técnica radica en que sus resultados pueden utilizarse, al menos potencialmente, en la toma de decisiones sobre la política de cartera de productos (renovación, retirada, cuotas de mercado, etc.).

La aplicación del análisis conjunto parte de una etapa destinada a su diseño. Se deben elegir los atributos y los niveles considerados como claves en el proceso de adquisición de los productos estudiados. En este análisis se han seleccionado como atributos y niveles más relevantes en la compra de vinos: el precio (oscilando entre 250, 400 y 600 pts.), el origen geográfico de la producción (local, Rioja u otros) y el tipo o año de elaboración del vino (del año y crianza). Las elecciones anteriores se han realizado en base a un estudio cualitativo realizado previamente y a otros estudios cuantitativos desarrollados en la literatura (véase por ejemplo Johnson et al., 1991). Sobre las múltiples combinaciones posibles entre los atributos y niveles anteriores se ha extraído un diseño ortogonal compuesto por nueve tarjetas o vinos hipotéticos.

La escala de valoración de los distintos tipos de vino ofrecidos toma valores entre 1 y 10, indicando 10 la máxima preferencia. Tal y como se ha indicado con anterioridad, el principal objetivo del estudio consiste en contrastar las diferencias en los resultados obtenidos dependiendo de los métodos de estimación elegidos. A pesar de las variaciones existentes en relación a los métodos de estimación empleados por cada una de las alternativas anteriores, que van a ser expuestos a continuación, en todos ellos se va a partir de un modelo de tipo aditivo. Es decir, se considera la valoración global efectuada por el comprador como la suma de las puntuaciones individuales de cada atributo (Steenkamp, 1987). En base a los tres atributos seleccionados, el modelo conjunto se formula como sigue (1):

$$Valoracion = \beta_0 + \sum_{i=1}^{n} \beta_{1i} D_{1i} + \sum_{j=1}^{m} \beta_{2j} D_{2j} + \sum_{k=1}^{p} \beta_{3k} D_{3K} \qquad (1)$$

dónde β_{1i}, β_{2j}, y β_{3k} son los part-worths asociados a los niveles i (i=1,2,....., n); j (j=1,2,.....,m) y k (k=1,2,....., p), de los atributos precio (1), origen (2) y tipo de vino (3) respectivamente. Las variables dummy D_{1i}, D_{2j}, y D_{3K} toman el valor 1 si el nivel correspondiente del atributo está presente y el valor 0 en el resto de los casos.

Se incluye a continuación, de forma breve, las principales diferencias existentes entre los tres métodos de estimación seleccionados: 1) Mínimos Cuadrados, 2) Tobit Doblemente Censurado y 3) Logit Ordenado.

Estimación por Mínimos Cuadrados

El modelo a estimar mediante la clásica metodología de los Mínimos Cuadrados de acuerdo con los atributos y niveles elegidos en este diseño es el siguiente (2):

$$Valoración = \beta_0 + \beta_1 \Pr ecio + \beta_2 Orig1 + \beta_3 Orig2 + \beta_4 DA + \varepsilon \qquad (2)$$

Dónde:

Valoración = Preferencia asignada a cada uno de los vinos hipotéticos. Los valores asignados oscilan entre 1 y 10, siendo 10 la máxima preferencia.

Precio = Precio

Orig1 = Variable dummy que toma el valor 1 si el vino se ha producido en la propia región y 0 en el resto de casos

Orig2 = Variable dummy que toma el valor 1 si el vino se ha producido en la Comunidad Riojana y 0 en el resto de casos

DA = Variable dummy que toma el valor 1 si el vino es del año y 0 en el resto de los casos

ε = Término de error

La estimación de este modelo se ha realizado por Mínimos Cuadrados Ponderados ya que la estimación Mínimo Cuadrática presentaba problemas de heteroscedasticidad[2].

Estimación mediante el modelo Tobit Doblemente Censurado

El segundo método de estimación considerado ha sido un modelo Tobit Doblemente Censurado. Una aplicación del mismo en el análisis conjunto puede encontrarse en Lin et al. (1996). En este caso, se parte de la clásica regresión lineal:

$$Valoración^* = \beta_0 + \beta_1 \Pr ecio + \beta_2 Orig1 + \beta_3 Orig2 + \beta_4 DA + \varepsilon \qquad (3)$$

en la que la variable Valoración* no es directamente observable y las variables exógenas son las mismas que las incluidas en (2). La variable que sí es observable es Valoración que está doblemente censurada con respecto a la variable Valoración* de la siguiente forma:

$$Valoración = \begin{cases} 1 & \text{si valoración}^* \leq 1 \\ \text{valoracion}^* = \beta_0 + \beta_1 \Pr ecio + \beta_2 Orig1 + \beta_3 Orig2 + \beta_4 DA + \varepsilon & \text{si } 1 < \text{valoración}^* < 10 \\ 10 & \text{si valoración}^* \geq 10 \end{cases}$$

En este caso, suponemos que las perturbaciones aleatorias se distribuyen para cada observación según una normal de media cero y varianza constante (σ^2), que coincide con la varianza de la variable artificial latente la cual, a su vez, se distribuye según una normal. En definitiva, esta variable latente es una mezcla de dos componentes, uno continuo y otro discreto asociado a los dos niveles.

A la vista de (3) podría parecer que el modelo Tobit es un modelo de regresión lineal y que , por tanto, podemos estimarlo por Mínimos Cuadrados (MC). Sin embargo, Maddala (1983) demuestra que los estimadores MC son sesgados e inconsistentes, debiendo recurrirse a una estimación Máximo-Verosimil (una exposición detallada del método de estimación se encuentra muy alejada del objetivo de este trabajo, aunque el lector interesado puede acudir al Capítulo 6 del manual de Maddala citado anteriormente).

Un supuesto básico en la estimación de (3) es el de homoscedasticidad, ya que si la perturbación aleatoria es heteroscedástica, las estimaciones de los parámetros

2. Se ha realizado el contraste de Breusch y Pagan (Novales, 1993), utilizándose el precio como variable explicativa. El valor obtenido es 42,72, por lo que para una χ^2 de 4 grados de libertad se puede rechazar la hipótesis nula de homoscedasticidad (el valor crítico al 5% es 3,84).

β_i (y=0,1,2,3 y 4) no son ni siquiera consistentes. Dado que el modelo Tobit se ha estimado mediante Máxima Verosimilitud se ha utilizado un estadístico basado en el ratio de verosimilitud. El valor del citado estadístico ascendió a 44, muy superior al valor crítico para un nivel de significación del 5%. Finalmente, en este trabajo, el modelo (3) se ha estimado considerando que la varianza de la perturbación es proporcional al cuadrado de la variable precio, con el fin de ser consistente con la solución adoptada en la regresión lineal.

Estimación mediante modelos de probabilidad ordenados

Este tipo de modelos parte de la hipótesis de que los valores de una variable cualquiera puede clasificarse en un conjunto de categorías ordenadas. Por ejemplo, en el caso del nivel de educación podríamos distinguir tres categorías: no haber finalizado el BUP, poseer el título de bachillerato, y tener educación universitaria. En el caso que nos ocupa, y teniendo en cuenta la distribución de frecuencias de los valores recogidos en la variable "valoración", hemos procedido a su clasificación en tres categorías: 1) valor inferior o igual a 3; 2) valor entre 4 y 7 y 3) valor superior o igual a 8.

El modelo de probabilidad ordenado fue introducido por McKelvey y Zavoina (1975) quienes, a su vez, desarrollaron un algoritmo para su estimación, de tal forma que el output resultante se asemejase bastante al de la típica regresión lineal. El modelo subyacente puede escribirse como:

$$Y_i = \beta_0 + \beta_1 \Pr ecio_i + \beta_2 Orig1_i + \beta_3 Orig2_i + \beta_4 DA_i + \varepsilon_i \qquad (4)$$

donde las variables explicativas han sido definidas anteriormente; es una secuencia de perturbaciones aleatorias; e Y_i es la valoración subyacente de un determinado vino. La variable Y_i no es observable, pero sí podemos saber a qué categoría pertenece dependiendo de la valoración que se haya otorgado al vino (en nuestro caso a la variable Y_i se le ha asignado los valores 0, 1 y 2 para las categorías 1, 2 y 3, respectivamente). Dado que la variable endógena sólo es observable de forma ordinal, podemos utilizar la regla de normalización de que $Var(\varepsilon) = 1$ (tal como se hace en el modelo probit). Por tanto, podemos suponer que las perturbaciones aleatorias se distribuyen de forma idéntica, según una normal de medio cero y varianza unitaria.

En el presente trabajo, se ha estimado una versión distinta, asumiendo que la perturbación aleatoria sigue una distribución logística en vez de una distribución normal. En este caso, el modelo puede expresarse como:

$$Y = 1/\left[1 + \exp^{-(X\beta)}\right]$$

donde

$$X\beta = \alpha_0 + \alpha_1 + \alpha_2 \, \mathrm{Pr}ecio + \alpha_3 Orig1 + \alpha_4 Orig2 + \alpha_5 DA + \varepsilon$$

siendo α_0 una variable ficticia de intervalo que toma el valor unitario si la valoración asignada pertenece a la categoría 1y valor nulo en el resto; α_1 es otra variable ficticia que toma el valor unitario si la valoración asignada pertenece a la categoría 2 y valor nulo en el resto. El resto de las variables ya han sido definidas.

Como puede apreciarse en la especificación anterior, se estima una constante separada para cada categoría de la variable dependiente. En nuestro caso, sólo aparecen dos ya que la constante del modelo base había sido normalizada.

En base a los resultados obtenidos en cada una de estos modelos de estimación se han evaluado las preferencias y se han segmentado los mercados. El siguiente epígrafe resume los resultados hallados.

3. Resultados

Estimaciones

Los resultados de las estimaciones obtenidos por los distintos métodos se recogen en el Cuadro 1. En general, los modelos ofrecen una aceptable valoración. La bondad del ajuste en los tres métodos (medidos por el R^2 y los estadísticos F y χ^2, según el caso) puede considerarse apropiada, sobre todo si se tiene en cuenta el elevado número de observaciones disponibles.

Cuadro 1. Coeficientes obtenidos en el modelo conjunto para cada uno de los tres métodos de estimación [a]

Variable	REGRESIÓN LINEAL	TOBIT DOBLEMENTE	Variable	LOGIT ORDENADO[b]
Constante	4,841733***	4,5050***	Constante 1	-10,540
			Constante 2	-2386,7***
Precio	-0,002913***	0,00081752**	Precio	7,5272
Orig1[d]	2,006374	0,10848*	Orig1[d]	2168***
Orig2[e]	2,385819***	0,091122	Orig2[e]	2505***
DA[f]	-0,768841***	-0,04088	DA[f]	825,58***
R^2 Ajustado	0,12606			
Estadístico F[c]	75,60727***	71,06***		
Log-likelihood		-5166,16		-2.191,9
Chi-Cuadradoc				3.782,6***
Tamaño muestra	2.048	2.048		2.048

a ***,**,* implica que existen diferencias significativas para unos niveles máximos de error del 1, 5 y 10 por cien respectivamente.
b Los modelos han sido estimados considerando que la varianza de la perturbación es proporcional al cuadrado del precio.
c En cada caso este estadístico contrasta la significatividad conjunta de los parámetros del modelo.
d Variable ficticia que toma el valor 1 si el vino se ha producido en Navarra y 0 en el resto.
e Variable ficticia que toma el valor 1 si el vino se ha producido en Rioja y 0 en el resto.
f Variable ficticia que toma el valor 1 si el vino es del año y 0 en el resto

En términos comparativos, los valores manifestados para cada uno de los niveles de los atributos son similares, independientemente del método de estimación empleado. En todos los modelos, la mayoría de las variables son significativas. El origen preferido para los vinos es el local o el riojano, a diferencia de los producidos en otros territorios. En cuanto al precio, excepto en la estimación por Mínimos Cuadrados, se asigna una mayor utilidad a aquellos caldos con precio superior. Finalmente, las preferencias manifestadas en relación con la elaboración, posicionan al vino más elaborado con puntuaciones superiores, excepto en el caso del modelo Logit que asigna una preferencia superior a los caldos del año.

En base a las estimaciones del Cuadro 1 es posible determinar las utilidades de los distintos niveles de cada atributo. Los valores son similares a los recogidos en dicho cuadro para las variables consideradas de tipo categórico (origen y tipo de vino), pero se deben transformar en el atributo numérico (precio). El cálculo de la utilidad asignada a cada nivel de precios se realiza multiplicando cada uno de los precios (250, 400 y 600 pts.) por el coeficiente obtenido para este atributo en cada uno de los métodos de estimación. En el caso de las variables categóricas, se puede observar que los valores obtenidos en las estimaciones corresponden a la preferencia por cada uno de los niveles sobre aquel nivel elegido como base. En el caso del origen del vino, las utilidades se refieren a la preferencia de vinos regionales sobre los nacionales o de vinos Riojanos sobre los nacionales. Para el tipo de vino, se indica la valoración otorgada a los caldos del año sobre los obtenidos mediante métodos de crianza.

Los resultados se presentan en el Cuadro 2. De nuevo, se puede observar que, en general, los resultados son similares, asignando una mayor preferencia a los vinos navarros o riojanos, de crianza (excepto en el modelo Logit Ordenado) y de precio superior (salvo en el caso de la estimación mínimo cuadrática).

Teniendo en cuenta las utilidades obtenidas en el cuadro anterior se puede calcular la importancia relativa de cada uno de los atributos dependiendo del método de estimación. Esta importancia se determina como la proporción de rango asignada a cada atributo sobre la variación de rangos total (Halbrendt *et al.*, 1991):

$$\text{Importancia Relativa} = \frac{Rango(i)}{\sum Rangos(i)} * 100 \qquad (6)$$

En el caso de la variable continua, el precio, su rango corresponde con la diferencia entre el mayor y menor valor en términos absolutos. Para las variables categóricas, teniendo en cuenta que las utilidades mostradas se refieren a preferencias relativas, esto es, sobre el nivel considerado como base, el cálculo del rango puede realizarse de dos formas distintas, teniendo en cuenta el número de niveles en cada atributo. Por ejemplo, en el caso del tipo de vino, en el que solo existen dos atributos, el rango viene determinado por el valor absoluto de la utilidad relativa (0.768 en la regresión lineal). En el caso del origen, en el que existen tres niveles, el rango viene determinado por el máximo valor de las utilidades relativas en valor absoluto, si todos son del mismo signo. En caso contrario, el rango viene determinado por la diferencia entre las utilidades máxima y mínima. En nuestro caso todas las utilidades relativas asignadas son del mismo signo, por lo que, por ejemplo en la regresión lineal el rango para el origen sería 2.385. A modo de ejemplo, si dividimos el valor anterior (2.385) entre la suma de rangos de dicho método de estimación (4.903) obtendremos para el origen geográfico una importancia relativa del 49% (Gráfico 1).

Cuadro 2. Utilidades de los distintos niveles de cada atributo para cada uno de los tres métodos de estimación

	REGRESIÓN LINEAL	TOBIT DOBLEMENTE CENSURADO	LOGIT ORDENADO
PRECIO			
250 Pts.	-0,72825	0,204380	1.881,80
400 Pts.	-1,16520	0,327008	3.010,88
600 Pts.	-1,74780	0,490512	4.516,32
ORIGEN			
Regional sobre nacional	2,006374	0,108480	2.168,00
Rioja sobre nacional	2,385819	0,091122	2.505,00
TIPO			
Del año sobre crianza	-0,768841	-0,04088	825,58

El Gráfico 1 resume las importancias relativas obtenidas para cada atributo en los tres métodos de estimación seleccionados en el estudio. El precio de los vinos, en primer lugar, seguido de su origen geográfico, son los aspectos más valorados en los modelos Tobit Doblemente Censurado y en el modelo Logit Ordenado. El tipo de vino es el atributo incluido en tercer lugar en estos dos modelos. En la regresión lineal

el origen geográfico de los caldos es el aspecto más valorado, seguido de los precios y su tipo de elaboración. Por lo tanto, los resultados obtenidos por los tres métodos presentan ligeras diferencias.

Gráfico 1. Importancia relativa de cada atributo para cada uno de los tres métodos de estimación

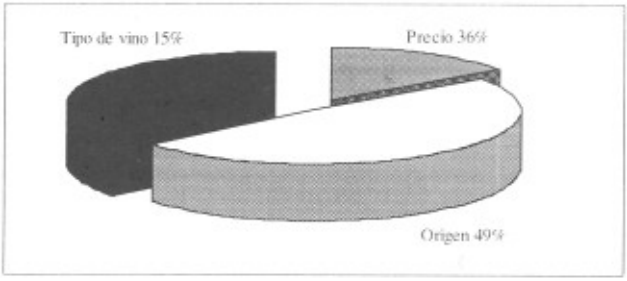

Fig. 1.—Regresión lineal

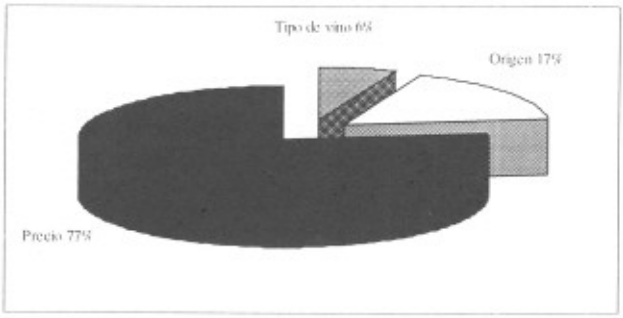

Fig. 2.—Tobit doblemente censurado

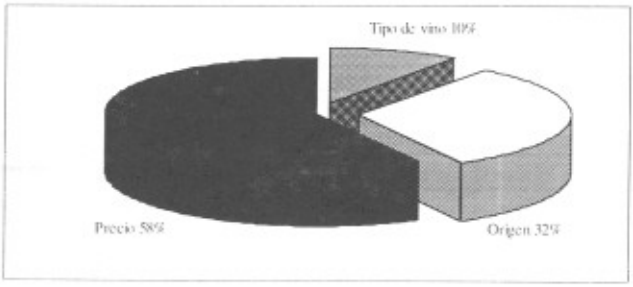

Fig. 3.—Logit ordenado

El siguiente caso en estudio ha consistido en calcular las preferencias a nivel individual, para cada uno de los métodos de estimación. En base a dichas preferencias se procederá a segmentar el mercado analizándose las características de cada segmento y se determinarán las diferencias existentes entre los resultados obtenidos para cada modelo.

Segmentación

En la segmentación se ha elegido la modalidad a posteriori (Green y Krieger, 1991 y Vriers et al, 1996). En una primera etapa, se han determinado los grupos existentes en base sus estructuras de preferencias y, en una segunda fase, se ha caracterizado a cada uno de los segmentos hallados. El proceso de determinación de los grupos se ha basado en el método del Cluster no Jerárquico con el procedimiento de K-Means, tomando como observaciones la importancia relativa asignada por cada individuo a cada uno de los tres atributos seleccionados (origen, precio y tipo de vino), en cada uno de los tres métodos de estimación aplicados. La aplicación de esta metodología ha permitido dividir al mercado en tres segmentos de composición similar según los tres tipos de estimación considerados. La caracterización de estos grupos se ha realizado en base a aspectos demográficos, socioeconómicos y de comportamiento en el mercado. En cuanto a las características demográficas y socioeconómicas se han incluido el sexo, la edad, la residencia, el tamaño de la familia y el nivel de renta. Las variables de comportamiento seleccionadas han sido el nivel de consumo de los vinos calificados y la cantidad de litros adquirida. Los grupos obtenidos junto a su caracterización se incluyen en el Cuadro 3.

Cuadro 3. Segmentación del mercado de acuerdo con las preferencias de los consumidores obtenidas mediante los tres métodos de estimación

	REGRESIÓN LINEAL			TOBIT DOBLEMENTE CENSURADO			LOGIT ORDENADO		
	Seg. (35%)*	Seg.2 (19%)	Seg.3 (46%)	Seg. 1 (25,7%)	Seg.2 (16%)	Seg. 3 (58,3%)	Seg. 1 (8%)	Seg. 2 (64%)	Seg. 3 (28%)
Precio	27,88	27,03	64,63	22,26	23,24	59,04	22,86	85,09	44,87
Origen	54,05	25,24	19,24	59,51	27,70	24,49	61,53	7,24	17,11
Tipo	18,06	47,42	16,12	18,22	49,45	16,46	15,59	7,66	38,01
Menores de 40 años	nd32,5%	30,95%	30,75%	nd29,3%	36,1%	31,05%	55,6%	34%	26,6%
R.Urbana	nd45%	33,3%	46,2%	37,9%	33,3%	*48,5%	55,6%	37,5%	53,1%
Menos 4 miembros	nd87,2%	84,2%	90,2%	85,7%	94,1%	87,5%	87,5%	84,1%	96,9%
Renta alta	nd15%	23,8%	13,5%	20,7%	22,2%	12,1%	nd-	18,1%	15,5%
Consumo Medio D.O.	nd6,5%	13,3%	7%	14,3%	5,4%	11,5%	4,5%	15,4%	7,4%
Cantidad D.O. (litros)nd	6,74	6,26	5,49	8,94	3,99	5,38	nd6,64	6,24	5,48

* Indica el tamaño del segmento

m Indica los valores medios obtenidos en cada segmento.

p Indica la proporción de individuos de cada grupo distinguidos por cada característica.

nd Indica la no existencia de diferencias significativas entre los grupos obtenidos.

Los resultados obtenidos a partir de los diferentes métodos de estimación son interesantes. La segmentación proporcionada por los métodos que consideran la variable dependiente como numérica es similar. Un primer segmento estaría formado por individuos que otorgan mayor importancia al origen del vino, siendo el precio el segundo aspecto más valorado. Un segundo grupo está formado por compradores que, en su decisión de compra otorgan al origen y al tipo de vino mayor importancia. Finalmente, un tercer segmento valoraría primordialmente el precio del producto adquirido. En términos porcentuales, el número de individuos del grupo que valora el precio es superior al resto, seguido del segmento centrado en el origen del vino.

La segmentación resultante utilizando el método logit (variable dependiente categórica) es, en cierto modo, diferente a la anterior. En este caso, existe un grupo un poco mayor que el resultante con los métodos comentados anteriormente que atribuyen una mayor importancia al precio (segundo segmento). El resto de los individuos encuestados asignaría mayor utilidad al tipo de vino, aunque al precio también se la da una elevada importancia. Finalmente, existirá un grupo marginal (3% de la población) para quienes el origen del vino constituye el factor más importante de los considerados a la hora de adquirir un vino.

A pesar de estos rasgos diferenciadores, se podrían equiparar los primeros segmentos de los métodos de variable cuantitativa con los grupos primero y tercero del modelo logit. Finalmente, los terceros segmentos de los modelos cuantitativos pueden asimilarse al segundo grupo del modelo categórico.

Si ahora introducimos las características sociodemográficas y de comportamiento, podemos caracterizar a los segmentos que hemos equiparado en el párrafo anterior. Si analizamos los segmentos proporcionados por los métodos de estimación cuantitativos los resultados son bastante similares. Los compradores que adquieren vinos por su precio se caracterizan por residir en mayor proporción en Pamplona, ser de familias menos numerosas y consumir menor cantidad de vino calificado que aquellos que adquieren los caldos por su origen o por su método de elaboración. De forma particular, entre los compradores que valoran más el tipo de vino se incluye una proporción superior de personas de renta elevada.

En el modelo categórico, el grupo mayoritario de precio presenta características similares a las consideradas anteriormente, mientras que en el segmento que valora tanto el precio como el tipo de vino se incluye una proporción superior de personas mayores de cuarenta años. Finalmente, el segmento preocupado por el origen ofrece similares características a los obtenidos en los modelos con variable dependiente numérica.

4. Conclusiones

En el presente trabajo se ha tratado de identificar las diferencias existentes en el Análisis Conjunto según se considere que las preferencias manifestadas por el con-

sumidor sea una variable numérica u ordinal. No se pretende discutir si es más adecuado una u otra consideración. Por el contrario, dado que en la literatura abundan los trabajos en ambos sentidos, nos ha parecido interesante determinar, en un ejemplo concreto, las consecuencias de la elección sobre una de las aplicaciones más utilizada en el Análisis Conjunto: la segmentación de mercados.

Se parte de las valoraciones realizadas por una muestra de consumidores navarros hacia nueve vinos tintos hipotéticos obtenidos mediante un diseño ortogonal a partir de tres atributos: precio, origen y tipo de elaboración. En el trabajo se han considerado tres tipos de modelos conjuntos, cada uno de los cuales se ha estimado asumiendo que la perturbación del modelo era proporcional al cuadrado del precio. El primer modelo ha considerado las preferencias como una variable continua que depende de una combinación lineal de los atributos. La estimación se realizó por Mínimos Cuadrados Ponderados. En segundo lugar, se ha considerado que las preferencias estaban acotadas por un valor inferior y superior para los cuales la variable se comportaba como discreta y continua para el resto de valores (modelo Tobit doblemente censurado). Este modelo se estimó al igual que en el siguiente por Máxima Verosimilitud. Finalmente, se sustituyó las preferencias por una escala ordinal ordenada de menor a mayor importancia (modelo de probabilidad ordenado).

Los resultados de las estimaciones de los tres modelos han resultado ser ligeramente diferentes, pudiéndose considerar en los tres casos como aceptables. Las principales diferencias se han localizado en las valoraciones del origen y del precio. Mientras el modelo Tobit Doblemente Censurado y el modelo Logit Ordenado asignan una preferencia superior al precio sobre el origen, la estimación por Mínimos Cuadrados Ponderados manifiesta una valoración superior por el origen geográfico del vino.

De forma adicional, otro resultado interesante lo constituye la segmentación obtenida con los diferentes métodos. El modelo de regresión lineal y el Tobit segmentan de forma idéntica, con pequeñas variaciones en cuanto al número de individuos que componen cada segmento. El modelo de probabilidad ordenado identifica tres segmentos, todos ellos bastante equiparables a los obtenidos mediante los otros dos métodos, a pesar de las diferencias existentes en sus tamaños, siendo superior aquél que valora especialmente los precios, e inferior el que se centra en el origen de los caldos.

Finalmente, el intento de caracterización de los segmentos obtenidos en base a características sociodemográficas y a variables de comportamiento ha dado como resultado una composición muy similar entre los grupos hallados según los distintos métodos de estimación. Si bien, la principal conclusión que se puede obtener es que no existen diferencias significativas importantes entre los segmentos obtenidos.

En cualquier caso, estos resultados únicamente deben circunscribirse a los datos utilizados y al criterio de segmentación elegido. Como se ha podido comprobar, úni-

camente hemos segmentado en base a las preferencias individuales para, a continuación, caracterizar estos segmentos. Con ello se trataba de evitar que las variables sociodemográficas discriminaran más que las preferencias. Extensiones de este trabajo hacia otros criterios de segmentación, o hacia otras aplicaciones del Análisis Conjunto, como puede ser la simulación, están siendo objeto de atención por parte de los autores.

Bibliografía

ECONOMETRIC SOFTWARE (1991). Limdep. User's manual and reference guide. Version 6.0. Australia.

GAN, Ch.; LUZAR, E.J. (1993): A conjoint analysis in waterfowl hunting in Louisiana. Journal of Agricultural and Applied Economics, 25 (2): págs. 36-45.

GINEO, W.M. (1990): A conjoint/logit analysis of nursery stock purchases. Northeastern Journal of Agricultural and Resources Economics, 19 (1): págs. 49-58.

GREEN, P.E.; KRIEGER, A.M. (1991): Segmenting markets with conjoint analysis. Journal of Marketing, 55: págs. 20-31.

GREEN, P.E.; SRINIVASAN, V. (1978): Conjoint analysis in consumer research: issues and outlook. Journal of Consumer Research, 5: págs. 103-123.

HALBRENDT, C.K.; BACON, J.R.; PESEK, J. (1992): Weigthed least squares analysis for conjoint studies: the case of hybrid striped bass. Agribusiness, 8 (2): págs. 187-198.

HALBRENDT, C.K.; WIRTH, E.F.; VAUGHN, G.F. (1991): Conjoint analysis of the Mid-Atlantic food-fish market for farm-raised hybrid stripted bass. Southtern Journal of Agricultural Economics, July, págs. 155-163.

JOHNSON, L.W. ; RINGHAM, K.; JURD,K. (1991): Behavioural segmentation in the Australian wine market using conjoint choice analysis. International Marketing Review, 8 (2), págs.26-31.

LIN, B.H.; PAYSON, S.; WERTZ, J. (1996): Opinions of professional buyers toward organic produce: a case study of Mid-Atlantic market for fresh tomatoes. Agribusiness, 12 (1): págs. 89-97.

MADDALA, G.S. (1983): Limited-dependent and qualitative variables in econometrics. Cambridge University Press.

McKELVEY, R.; ZAVOINA, W. (1975): A statistical model for the analysis of ordinal level dependent variables. Journal of Mathematical Sociology, 4: págs. 103-120.

MÚGICA, J.M. (1989): El análisis conjunto. Alternativas, problemas y limitaciones. Ipmark, 326: págs. 45-54.

NOVALES, A (1993): Econometría. Segunda Edición. McGraw Hill, Madrid.

RUIZ DE MAYA, S.; MUNUERA, J.L. (1993): Las preferencias del consumidor: estudio de su composición a través del análisis conjunto. Estudios sobre consumo, 28: págs. 27-43.

STEENKAMP, J-E-.B. (1987): Conjoint measurement in ham quality evaluation. Journal of Agricultural Economics, 38 (3): págs. 473-480.

VRIENS, M.; WEDEL, M.; WILMS, T. (1996): Metric conjoint segmentation methods: a Monte Carlo comparison. Journal of Marketing Research, 33: págs. 73-85.