

ESTADÍSTICA ESPAÑOLA
Vol. 49, Núm. 166, 2007, págs. 451 a 471

Modelos de elección discreta y especificaciones ordenadas: una reflexión metodológica

por

M^a CAROLINA RODRÍGUEZ DONATE
JOSÉ JUAN CÁCERES HERNÁNDEZ

Departamento de Economía de las Instituciones, Estadística Económica y Econometría.
Facultad de Ciencias Económicas y Empresariales. Universidad de La Laguna.

RESUMEN

Este trabajo pretende señalar que las especificaciones ordenadas pueden resultar demasiado rígidas para captar la naturaleza de un proceso de elección individual, en el que la decisión está orientada por la maximización de la utilidad, si el orden establecido entre las alternativas objeto de elección no expresa un orden subyacente de utilidades asociadas a aquéllas. Para ilustrar esta limitación, se comparan los resultados de un modelo *logit* y un *logit* ordenado que tratan de identificar el impacto de las características individuales en la decisión sobre la frecuencia de consumo de vino de los residentes en Tenerife.

Palabras clave: modelos de elección discreta, especificaciones multinomiales, alternativas ordenadas.

Clasificación AMS: 62P20

1. INTRODUCCIÓN

Los modelos de elección discreta resultan apropiados cuando el objetivo no es predecir el comportamiento medio de un agregado, sino analizar los factores determinantes de la probabilidad de que un agente económico individual elija un curso de acción dentro de un conjunto, generalmente finito, de opciones posibles. Si el objeto del análisis son las preferencias o utilidades individuales de los agentes económicos, puestas de manifiesto en una elección concreta, la variable explicada suele poseer naturaleza cualitativa y, a la hora de caracterizar el comportamiento probabilístico del atributo, resulta útil predecir la probabilidad asignada a cada una de las modalidades del mismo como función de aquellas características que, *a priori*, explican la decisión del individuo. Si el conjunto de alternativas es finito, los modelos multinomiales pueden constituir especificaciones apropiadas. En el caso particular de que los valores de la variable dependiente tengan un significado ordinal, puede ser conveniente especificar modelos multinomiales ordenados, que tienen en cuenta explícitamente este carácter.

En la literatura especializada reciente, existen múltiples aplicaciones de estos últimos modelos, que se han empleado profusamente en el ámbito de la economía(1). En muchos de estos trabajos se opta directamente por una especificación ordenada. Ahora bien, a pesar de que las alternativas estén ordenadas, el uso de estas especificaciones no es siempre la mejor opción. De hecho, se trata de especificaciones que resultan incapaces de recoger adecuadamente la naturaleza de un proceso de decisión en el que existen factores que actúan en el mismo sentido sobre la probabilidad de elección de las alternativas extremas. Esta rigidez de las especificaciones ordenadas puede entenderse si se analiza la justificación teórica que sustenta las especificaciones multinomiales, tal como se expone en el epígrafe siguiente. En el epígrafe tercero, se ilustra el problema mediante una aplicación de modelos *logit* multinomiales con objeto de analizar las decisiones sobre la frecuencia de consumo de vino de los residentes en Tenerife y, más en concreto, identificar las características que predominan entre aquéllos que presentan cada patrón de consumo.

(1) Véase, entre otros, Pudney y Shields (2000), Smith y Taylor (2001), Basile *et al.* (2003), Lauer (2003), Van Doorslaer y Jones (2003), Contoyannis *et al.* (2004), Jaggia y Thosar (2004), Lázaro *et al.* (2004), Boz y Akbay (2005), Choo *et al.* (2005), Ndejunga y Nelson (2005), Richards y Patterson (2005), Storchmann (2005), Fosgerau y Bjorner (2006).

2. ESPECIFICACIONES MULTINOMIALES Y ALTERNATIVAS ORDENADAS

Los ingredientes de un modelo multinomial son los objetos de la elección, el conjunto de alternativas disponibles, las características observadas de los agentes y de las alternativas y el modelo de comportamiento o elección individual. El proceso de decisión del individuo i puede representarse mediante una variable categórica Y_i tal que $Y_i = j$ si el individuo elige la alternativa j . Las alternativas deben ser mutuamente excluyentes y, además, exhaustivas, es decir, el conjunto de alternativas especificadas debe recoger todas las opciones posibles. Una vez especificada la variable dependiente, la probabilidad de que el individuo i elija la alternativa j , $P(Y_i = j)$, podrá expresarse como una función de un conjunto de factores, que pueden ser tanto características propias del agente que decide como características específicas de cada alternativa.

La especificación concreta de esta función puede deducirse a partir de la formulación de supuestos sobre los mecanismos de decisión. Y con objeto de analizar la racionalidad que subyace en el proceso de decisión, cabe considerar dos enfoques distintos. Por un lado, una elección puede interpretarse como resultado de la evaluación implícita, por parte del decisor, de una variable latente que es función de un conjunto de variables explicativas y cuyo valor determina que la elección sea una u otra (McKelvey y Zavoina, 1975). Desde otro punto de vista, puede suponerse que el agente decisor es perfectamente racional y escoge aquella alternativa que le proporciona máxima utilidad. En las aplicaciones económicas que hacen referencia a elecciones hechas por consumidores, productores o gobiernos, prevalece el enfoque de la teoría de la utilidad aleatoria, y las especificaciones obtenidas a partir de este enfoque se deducen a continuación.

2.1 Justificación teórica de los modelos de respuesta multinomial

La teoría del consumidor supone la existencia de consumidores racionales que toman decisiones maximizando su utilidad. Desde este punto de vista, McFadden (1973) formula un modelo de elección discreta en el que la probabilidad de que se elija una alternativa se define como la probabilidad de que dicha alternativa tenga la mayor utilidad entre el conjunto de alternativas posibles.

Supóngase que el proceso de decisión implica elegir entre $J+1$ alternativas, es decir, la variable dependiente Y_i toma valores desde 0 hasta J , de modo que interesa evaluar $P(Y_i = j)$, $j = 0, \dots, J$. La utilidad que obtiene el individuo i de la alternativa j puede representarse mediante $U_{i,j}$, $j = 0, \dots, J$. La utilidad de cada una de las alternativas no es observable por el investigador; sin embargo, depende de un conjunto de características del individuo, x_i , y un conjunto de atributos propios

de cada una de las alternativas, s_j , que sí son observables. Así, el componente determinístico o sistemático de la utilidad de la alternativa j para el individuo i puede especificarse como $V_{i,j} = v_{i,j}(x_i, s_j)$. La forma funcional que se utiliza generalmente para expresar el componente determinístico es lineal en los parámetros, de modo que $V_{i,j} = \beta_j^x x_i + \beta_j^s s_j$, donde β_j^x es un vector de parámetros que determina diferentes probabilidades para cada alternativa j en función de las características del individuo i , mientras que β^s es otro vector de parámetros que introduce diferencias en las probabilidades de elegir cada alternativa como función de los atributos propios de ésta.

Además, aparte de errores de medición, hay factores no observables por el analista que influyen en la utilidad y que no están incluidos en $V_{i,j}$; por ello, además del componente determinístico, debe considerarse un término aleatorio, cuya inclusión permite tener en cuenta que individuos aparentemente idénticos puedan escoger alternativas diferentes. Así, la utilidad del individuo para cada una de las alternativas queda recogida mediante la expresión $U_{i,j} = V_{i,j} + \varepsilon_{i,j}$. El individuo elegirá aquella alternativa que le proporcione máxima utilidad. Así, se elegirá la alternativa j si y sólo si $U_{i,j} > U_{i,k}$, $\forall k \neq j$. Entonces, la probabilidad de que el individuo i elija la alternativa j puede expresarse como

$$P(Y_i = j) = P(U_{i,j} > U_{i,k}, \forall k \neq j) = P(\varepsilon_{i,k} - \varepsilon_{i,j} < V_{i,j} - V_{i,k}, \forall k \neq j).$$

Así pues, el modelo finalmente especificado depende de la distribución del vector de términos de perturbación $(\varepsilon_{i,0}, \dots, \varepsilon_{i,J})$. Si $\varepsilon_i = (\varepsilon_{i,0}, \dots, \varepsilon_{i,J}) \approx N(0, \Sigma)$, se obtiene el modelo *probit* multinomial, originalmente propuesto por Thurstone (1927). Dado que la decisión depende de las diferencias de utilidad entre alternativas, la probabilidad $P(Y_i = j)$ puede expresarse en términos de estas diferencias. Si se definen las diferencias $\tilde{\varepsilon}_{i,j,k} = \varepsilon_{i,k} - \varepsilon_{i,j}$, $k = 0, \dots, J, k \neq j$, se obtiene un vector de dimensión J tal que $\tilde{\varepsilon}_{ij} = (\tilde{\varepsilon}_{i,j,0}, \dots, \tilde{\varepsilon}_{i,j,J}) \approx N(0, \tilde{\Sigma})$. Por tanto,

$$P(Y_i = j) = P(\tilde{\varepsilon}_{i,j,k} + \tilde{V}_{i,j,k} < 0, \forall j \neq k),$$

donde $\tilde{V}_{i,j,k} = \tilde{V}_{i,k} - \tilde{V}_{i,j}$, $k = 0, \dots, J, k \neq j$. La especificación final del modelo se construye expresando las diferencias en el componente determinístico de la utilidad como una función, habitualmente lineal, de un conjunto de variables explicativas, es decir,

$$\tilde{V}_{i,j,k} = V_{i,j} - V_{i,k} = \beta_j^x x_i + \beta^s s_j - \beta_k^x x_i - \beta^s s_k = (\beta_j^x - \beta_k^x)' x_i + \beta^s (s_j - s_k).$$

La especificación del modelo *probit* multinomial ofrece un amplio rango de posibilidades en función de la estructura de correlación especificada para los términos de perturbación del vector ϵ_i o del vector $\tilde{\cdot}$. En particular, la especificación apropiada de la matriz Σ permite considerar que el patrón de sustitución entre las alternativas j y k , definido como cociente de las probabilidades de elección respectivas, se modifique cuando cambia alguno de los atributos que define una tercera alternativa m . En muchas situaciones prácticas, es difícil admitir que dicho cociente de probabilidades se mantiene constante cuando cambia la utilidad que proporciona otra alternativa. Este supuesto rígido, poco frecuente en la práctica, se impone a priori en especificaciones menos flexibles como, por ejemplo, el modelo *logit* multinomial, que, sin embargo, presenta menos dificultades computacionales(2) y, además, ofrece posibilidades de interpretación más ricas de la compleja red de efectos a menudo presente en los procesos de decisión.

En el modelo *logit* multinomial se exige que los elementos del vector de perturbaciones $\epsilon_i = (\epsilon_{i,0}, \dots, \epsilon_{i,J})$ sigan distribuciones valor extremo tipo I (Gumbel) independientes, es decir, la función de densidad de cada término de perturbación $\epsilon_{i,j}$ es

$$f(\epsilon_{i,j}) = e^{-\epsilon_{i,j}} e^{-e^{-\epsilon_{i,j}}}$$

Por tanto,

$$P(Y_i = j) = P(\epsilon_{i,k} < \epsilon_{i,j} + V_{i,j} - V_{i,k}, \forall k \neq j) = \int \left(\prod_{k \neq j} e^{-e^{-(\epsilon_{i,j} + V_{i,j} - V_{i,k})}} \right) e^{-\epsilon_{i,j}} e^{-e^{-\epsilon_{i,j}}} d\epsilon_{i,j} .$$

Y, tras algunas operaciones algebraicas (Train, 2003: 85-86), se obtiene que

$$P(Y_i = j) = \frac{e^{V_{i,j}}}{\sum_{k=0}^J e^{V_{i,k}}} , j = 0, \dots, J .$$

Esta especificación permite advertir uno de sus principales inconvenientes. Y es que el patrón de sustitución entre las alternativas j y k , definido como cociente de

(2) De hecho, la aplicabilidad del modelo *probit* multinomial se ve restringida cuando el número de alternativas es elevado, dada la complejidad de las integrales que resultan y que deben ser evaluadas numéricamente. Esta complejidad ha limitado fuertemente el uso de estos modelos. Sin embargo, en los últimos años se han desarrollado métodos de aproximación basados en la simulación que amplían significativamente las posibilidades de aplicación de esta especificación en situaciones económicas reales (Bhat, 2003; Train, 2003).

las probabilidades respectivas, sólo depende de la diferencia de utilidades entre dichas alternativas, pero no de las utilidades que proporcionan las demás, dado que,

$$\frac{P(Y_i = j)}{P(Y_i = k)} = e^{V_{i,j} - V_{i,k}}, \quad j, k = 0, \dots, J, \quad j \neq k.$$

Desde el punto de vista del enfoque de la utilidad aleatoria, esta propiedad, conocida como independencia de alternativas irrelevantes (IIA), se deriva del supuesto distribucional sobre el vector $\tilde{\varepsilon} = (\varepsilon_{i,0}, \dots, \varepsilon_{i,J})$ y, más en concreto, del supuesto de igualdad de varianza e independencia para cada uno de sus elementos. Sin embargo, en muchas ocasiones la descripción adecuada del proceso de decisión exige considerar la presencia de estructuras de correlación entre los elementos del vector $\tilde{\varepsilon}$. Desde este punto de vista, el modelo *probit* multinomial supone, como ya se comentó, una especificación más flexible, pero también se han desarrollado otras especificaciones(3). Ahora bien, la propiedad de IIA puede proporcionar una representación adecuada de la realidad en determinados casos(4) y, entonces, la formulación del modelo *logit* supone ventajas prácticas.

Si el componente determinístico de la utilidad se expresa como una función lineal de características del individuo i y de la alternativa j , es decir, $V_{i,j} = \beta_j^x x_i + \beta_j^s s_j$, la especificación final del modelo depende de la naturaleza de las variables que se incorporen. El modelo propuesto por Theil (1969) se obtiene si

(3) Existen modelos *logit* específicos derivados a partir de hipótesis más generales sobre el vector $\tilde{\varepsilon}_i = (\varepsilon_{i,0}, \dots, \varepsilon_{i,J})$ que en el *logit* convencional. Así se obtiene el modelo *logit* anidado, que resulta apropiado cuando el conjunto de alternativas entre las que elige el tomador de la decisión puede ser particionado en subconjuntos o nidos de modo que la IIA se mantiene entre dos alternativas del mismo nido, pero no entre alternativas de nidos diferentes. Por otra parte, en procesos de decisión con respuesta policotómica, la naturaleza del proceso de decisión puede ser suficientemente compleja para que alguna de las especificaciones anteriores capte la multiplicidad de factores que explican la decisión y el modo en que actúa cada uno de ellos. En particular, en estas formulaciones se supone que una cualquiera de las variables explicativas contribuye al valor específico de la utilidad de una alternativa dada de acuerdo con una función lineal que implica que, ante un cambio dado de la variable explicativa, el efecto en la utilidad es siempre el mismo. Y, por tanto, la probabilidad de elegir cada alternativa reacciona demasiado rígidamente ante cambios en las variables explicativas. Considerar que las respuestas de los individuos no son tan homogéneas parece un supuesto más acorde con la realidad, pero entonces las probabilidades derivadas a partir de las especificaciones habituales recogen solamente una especie de respuesta media de los individuos. Una forma de captar la heterogeneidad en las respuestas individuales consiste en admitir la naturaleza estocástica de los parámetros que determinan dichas respuestas. Esta idea es la que da origen a los modelos mixtos. El modelo *logit* mixto ha recibido atención creciente en los últimos años y, aunque la formulación del *probit* mixto tampoco es novedosa, sólo los avances computacionales recientes han permitido su aplicación práctica.

(4) De hecho, Luce (1959) consideró la IIA como apropiada para describir ciertos procesos de decisión y derivó el modelo *logit* a partir de este supuesto.

$V_{i,j} = \beta_j^x \cdot x_i$. Entonces, el modelo *logit* derivado a partir del enfoque de utilidad aleatoria se transforma en

$$P(Y_i = j) = \frac{e^{\beta_j^x \cdot x_i}}{\sum_{k=0}^J \frac{e^{\beta_k^x \cdot x_i}}{e^{\beta_0^x \cdot x_i}}} = \frac{e^{(\beta_j^x - \beta_0^x) \cdot x_i}}{1 + \sum_{k=1}^J e^{(\beta_k^x - \beta_0^x) \cdot x_i}} = \frac{e^{\beta_j^x \cdot x_i}}{1 + \sum_{k=1}^J e^{\beta_k^x \cdot x_i}}, \quad j = 1, \dots, J.$$

Y la probabilidad de elegir la alternativa tomada de referencia se expresa como

$$P(Y_i = 0) = \frac{1}{1 + \sum_{k=1}^J e^{\beta_k^x \cdot x_i}}.$$

2.2 Justificación teórica de los modelos de respuesta multinomial con alternativas ordenadas

Si la variable dependiente de un modelo de elección discreta, Y_i , expresa preferencias u opiniones de los individuos sobre una determinada cuestión, las alternativas del proceso de decisión pueden expresar implícitamente un orden de utilidad y tener, por tanto, carácter ordinal(5). De manera que si se recurre a un modelo multinomial que no tenga en cuenta el carácter ordenado de los datos, tal vez no se describa adecuadamente el proceso de decisión del individuo. Una forma natural de interpretar este proceso de decisión consiste en considerar que el individuo tiene una valoración u opinión específica sobre la cuestión que determina su elección. En términos formales, la valoración u opinión del individuo puede definirse como una variable latente, Y_i^* , que es función de un conjunto de variables explicativas. De este modo, el rango de esta variable puede subdividirse en intervalos ordenados tales que si la valoración del individuo se sitúa en el intervalo j , el individuo elige la opción j , es decir, $Y_i = j$, $j = 0, \dots, J$.

La variable no observable o latente, Y_i^* , depende de un conjunto de variables explicativas que determinan la elección por parte del individuo i (6), es decir,

(5) Ejemplos de este tipo de variables aparecen frecuentemente en estudios a partir de encuestas, en las que se le pide al entrevistado que elija entre un rango de respuestas ordenadas para indicar sus preferencias o evaluar las características de un determinado producto.

(6) Aunque en los desarrollos posteriores se supondrá que el vector de variables explicativas contiene características de los individuos, en este conjunto de regresores podrían incluirse también, en algún caso, características específicas de la alternativa.

$$Y_i^* = \beta' x_i + \varepsilon_i,$$

donde ε_i es un término de perturbación aleatoria. La elección del individuo entre las alternativas $0, \dots, J$ es reflejo del valor de Y_i^* en relación con los umbrales μ_j , $j = 1, \dots, J$, es decir,

$$Y_i = \begin{cases} 0, & \text{si } Y_i^* \leq 0 \\ 1, & \text{si } 0 < Y_i^* \leq \mu_1 \\ 2, & \text{si } \mu_1 < Y_i^* \leq \mu_2 \\ \dots \\ J, & \text{si } \mu_{J-1} < Y_i^* \end{cases}$$

Estos umbrales, desconocidos a *priori*, deben ser tales que $0 < \mu_1 < \mu_2 < \dots < \mu_{J-1}$ (7). Si se adopta este enfoque, el modelo probabilístico que determina la elección queda definido de la siguiente manera

$$P(Y_i = 0) = F(-\beta' x_i),$$

$$P(Y_i = j) = F(\mu_j - \beta' x_i) - F(\mu_{j-1} - \beta' x_i), \quad j = 1, \dots, J-1,$$

$$P(Y_i = J) = 1 - F(\mu_{J-1} - \beta' x_i)$$

siendo F la función de distribución de la variable aleatoria ε_i . La distribución del término de error determina el tipo de modelo. De nuevo, las dos distribuciones más comúnmente usadas son la logística y la normal, dando lugar a los modelos *logit* ordenado y *probit* ordenado, respectivamente. Si se escoge el modelo *logit*,

$$F(\mu_j - \beta' x_i) = \Lambda(\mu_j - \beta' x_i) = \frac{e^{(\mu_j - \beta' x_i)}}{1 + e^{(\mu_j - \beta' x_i)}}, \quad j = 0, \dots, J-1.$$

Las formulaciones ordenadas pueden captar apropiadamente determinados procesos de decisión. Ahora bien, en problemas generales de decisión con alternativas ordenadas, dicha ordenación puede no implicar un orden implícito de utilidades. Parece entonces más apropiado considerar que el individuo elige la alternativa que le proporciona máxima utilidad. Desde este punto de vista, ligar la decisión al

(7) En esta formulación el primer umbral se considera nulo y se supone que la función índice $\beta' x_i$ contiene un término independiente. Otra formulación equivalente consiste en introducir un umbral μ_0 y eliminar el término independiente de la función índice.

valor de la variable latente implica suponer que existe una función lineal de las variables explicativas de la decisión tal que si dicha función toma valores entre μ_{j-1} y μ_j , la alternativa j es la que proporciona máxima utilidad. Nótese que en el planteamiento de los modelos con alternativas no ordenadas, las diferencias de utilidad entre cada par de alternativas quedan expresadas en términos de diferentes funciones lineales de las variables explicativas de la decisión. Ahora existe una única función lineal cuyo valor determina la alternativa con utilidad máxima. Este supuesto puede ser razonable si se analiza, por ejemplo, el efecto del número de horas de estudio de un alumno sobre la calificación obtenida por éste, expresada en escala ordinal, o el efecto de la renta familiar sobre el nivel educativo que elige un individuo. Pero no resulta tan evidente que este mismo supuesto sea capaz de captar la incidencia de la renta en la sensación de bienestar individual o el efecto del nivel educativo sobre la frecuencia de consumo de vino. Desde este punto de vista, a pesar de que las alternativas estén ordenadas, las especificaciones ordenadas introducen restricciones que pueden resultar excesivamente rígidas.

3. APLICACIÓN A LA DECISIÓN SOBRE LA FRECUENCIA DE CONSUMO DE VINO

En este apartado, y con objeto de mostrar la rigidez que puede incorporar la especificación ordenada, se comparan los resultados obtenidos mediante un modelo *logit* convencional y un modelo *logit* ordenado que pretenden describir el impacto de determinadas características socioeconómicas individuales en las decisiones sobre la frecuencia de consumo de vino de los residentes en Tenerife. Aunque podía haberse optado por otras especificaciones más complejas, el análisis se ha restringido a los modelos *logit* debido a su facilidad para interpretar el grado en que la especificación ordenada no es capaz de captar el efecto particular de determinados atributos.

Para desarrollar esta aplicación se han utilizado los datos de una encuesta dirigida al segmento de la población residente mayor de 18 años(8). Con objeto de recoger la frecuencia con que se realiza el acto de consumir vino, se han considerado 3 modalidades: no consumo, consumo ocasional o de mediana frecuencia y

(8) La encuesta fue realizada en el marco del proyecto de investigación titulado “Análisis de las preferencias de los consumidores y la demanda de vino en Tenerife”, financiado por el Excmo. Cabildo Insular de Tenerife (Proyecto nº 7-2001). Los individuos encuestados fueron elegidos por muestreo aleatorio estratificado con afijación proporcional en las tres grandes áreas geográficas en las que se dividió la isla (véase Guirao *et al.*, 2001). Los resultados que se muestran a continuación se refieren a los 1172 individuos que respondieron a cada una de las cuestiones consideradas en el estudio.

consumo de alta frecuencia. Si el objetivo último es diseñar estrategias comerciales que conduzcan a incrementar la demanda de vino en general o a captar clientes para la oferta insular, es necesario conocer los rasgos que identifican a los individuos que presentan cada patrón específico de consumo y, a partir de la identificación de conjuntos homogéneos de consumidores, diseñar estrategias de segmentación del mercado y, de un modo más general, elaborar distintos planes de marketing relativos a precios, producto, comunicación, distribución, etc... En la muestra considerada, los atributos utilizados para describir a los individuos con cada patrón de consumo son género, edad (18-29, 30-39, 40-49, 50-59, 60-60, 70 o más), área de residencia (zona norte, zona sur, área metropolitana), estado civil (casado, soltero, viudo-separado), ocupación (empleado, funcionario, estudiante, ama de casa, empresario, profesional, otros) y nivel de formación (sin estudios, estudios primarios, estudios secundarios, estudios universitarios).

Un análisis de la variabilidad del consumo atendiendo a ciertas características sociodemográficas de los encuestados puede ayudar a la hora de corroborar o refutar algunas de las ideas preestablecidas sobre los rasgos que identifican a los consumidores de vino e, incluso, puede servir de base para formular nuevas hipótesis. En particular, el consumo ocasional está más presente en el área metropolitana, mientras que el no consumo y el consumo de alta frecuencia tiene más peso en las zonas norte y sur. Por otro lado, la frecuencia de consumo parece estar inversamente relacionada con el nivel de estudios. Casi la mitad de los individuos sin estudios declararon no consumir vino, mientras que este porcentaje es mucho más bajo para los encuestados con estudios universitarios. Ahora bien, el consumo de alta frecuencia se presenta en mayor proporción de casos entre los individuos con bajo nivel de estudios, mientras que entre los entrevistados con niveles de estudios secundarios o universitarios el consumo es más ocasional (cuadro 1).

Cuadro 1
FRECUENCIA DE CONSUMO SEGÚN ÁREA DE RESIDENCIA Y NIVEL DE FORMACIÓN(a)

		<i>No consumo</i>	<i>Consumo Ocasional</i>	<i>Consumo de alta frecuencia</i>
Área de residencia	Norte	101 (25.9)	197 (50.5)	92 (23.6)
	Sur	99 (24.9)	206 (51.9)	92 (23.2)
	Metropolitana	82 (21.3)	245 (63.6)	58 (15.1)
Nivel de formación	Sin estudios	54 (45.8)	31 (26.3)	33 (28.0)
	Estudios primarios	87 (24.0)	176 (48.5)	100 (27.5)
	Estudios secundarios	90 (22.4)	246 (61.2)	66 (16.49)
	Estudios universitarios	51 (17.6)	195 (67.5)	43 (14.9)

(a) Entre paréntesis se indica el peso relativo de las alternativas para cada una de las modalidades de los dos atributos.

Sin embargo, la identificación de factores relevantes en la decisión y la cuantificación probabilística de los efectos particulares de cada uno de ellos en un modelo que los considere conjuntamente exige acudir a herramientas inferenciales como las proporcionadas por los modelos de elección discreta que, teniendo en cuenta el carácter discreto de la variable dependiente, permiten evaluar la probabilidad de que un individuo definido por un vector de características decida consumir vino con una frecuencia determinada. Y dado que puede considerarse que las modalidades del atributo que recoge dicha frecuencia están ordenadas, una aproximación inicial a la cuantificación de estas relaciones podría obtenerse a partir de un modelo *logit* ordenado. Ahora bien, la ordenación de categorías de acuerdo con la frecuencia de consumo no implica un orden subyacente de utilidades asociadas a dichas categorías y, si se acepta que la elección está, en última instancia, determinada por dichas utilidades, la especificación ordenada puede ser, como ya se comentó, una especificación demasiado rígida. Desde este punto de vista, un modelo *logit* multinomial estándar puede ser una especificación más adecuada.

Los resultados de la estimación de ambos modelos se muestran en los cuadros 2 y 3(9). Las alternativas que elige el individuo se han definido del modo siguiente:

(9) Las estimaciones de ambos modelos se han obtenido utilizando el algoritmo de Broyden, Fletcher, Goldfarb y Shanno (Train, 2003: 225), implementado en los programas Limdep 8.0 y Nlogit 3.0. (Greene 2002a, 2002b).

$Y_i = 0$, no consumo; $Y_i = 1$, consumo ocasional; $Y_i = 2$, consumo de alta frecuencia. Cada uno de los atributos explicativos de las probabilidades de elección se ha incorporado en términos de un conjunto de variables *dummies* que toman valor uno si el atributo en cuestión toma una modalidad determinada y cero en otro caso.

Cuadro 2
ESTIMACIÓN DE PARÁMETROS

	<i>Logit</i> ordenado	<i>Logit</i> multinomial	
		$P(Y_i = 1)$	$P(Y_i = 2)$
Constante	0.5816	-0.5122*	-1.7760
Mujer	-0.9714	-0.6589	-1.6199
30-39 años	0.0976*	-0.3381*	0.5099*
40-49 años	0.3227*	-0.6652	0.7453*
50-59 años	0.4060*	-0.6379	0.8570
60-69 años	0.3911*	-0.8735	0.8512*
Más de 70 años	0.1565*	-1.3250	0.7293*
Norte	0.2111*	-0.3121*	0.4497
Sur	0.2274*	-0.2702*	0.4307*
Casado	0.9006	0.9052	1.3481
Viudo/Separado	0.5384	0.5074*	1.0120
Empleado	0.1936*	0.7213	0.5295*
Funcionario	0.5321*	1.3070	1.5454
Ama de casa	-0.8781	0.4333*	-1.0014*
Empresario	0.4487*	0.5650*	0.8851*
Profesional	0.6008	1.3815	1.4139
Otra ocupación	0.2794*	0.5211*	0.2866*
Estudios primarios	0.5011	0.8469	0.3905*
Estudios secundarios	0.3361*	0.9920	0.1949*
Estudios universitarios	0.5527	1.4853	0.6382*
μ_1	2.9665		
N = 1172 ; LnL = -1047,591		N = 1172 ; LnL = -989,4755	

*Parámetros no significativos al 90% de confianza

Cuadro 3
EVALUACIÓN DE CAMBIOS DISCRETOS(a)

	<i>Logit ordenado</i>			<i>Logit multinomial</i>		
	$P(Y_i = 0)$	$P(Y_i = 1)$	$P(Y_i = 2)$	$P(Y_i = 0)$	$P(Y_i = 1)$	$P(Y_i = 2)$
Mujer	0.0724	0.0039	-0.0763	0.1542	0.0106	-0.1648
30-39 años	-0.0087	0.0013	0.0074	0.0336	-0.1190	0.0854
40-49 años	-0.0266	-0.0004	0.0270	0.0658	-0.2213	0.1556
50-59 años	-0.0324	-0.0027	0.0351	0.0568	-0.2276	0.1708
60-69 años	-0.0314	-0.0022	0.0336	0.0847	-0.2814	0.1966
Más de 70 años	-0.0137	0.0015	0.0122	0.1438	-0.3663	0.2225
Norte	-0.0179	0.0010	0.0169	0.0229	-0.1110	0.0881
Sur	-0.0192	0.0009	0.0183	0.0187	-0.0998	0.0811
Casado	-0.0748	0.0005	0.0743	-0.1668	0.0721	0.0947
Viudo/Separado	-0.0503	0.0145	0.0358	-0.1115	0.0236	0.0878
Empleado	-0.0136	-0.0017	0.0153	-0.1153	0.1079	0.0074
Funcionario	-0.0324	-0.0178	0.0502	-0.1109	-0.1620	0.2729
Ama de casa	0.0947	-0.0596	-0.0351	0.0456	0.0667	-0.1123
Empresario	-0.0283	-0.0123	0.0406	-0.1152	0.0342	0.0810
Profesional	-0.0355	-0.0230	0.0585	-0.2025	0.1377	0.0648
Otra ocupación	-0.0189	-0.0042	0.0232	-0.0823	0.0892	-0.0069
Estudios primarios	-0.0463	0.0094	0.0369	-0.1341	0.1534	-0.0193
Estudios secundarios	-0.0329	0.0096	0.0233	-0.1459	0.2062	-0.0603
Estudios universitarios	-0.0502	0.0087	0.0414	-0.2205	0.2726	-0.0521

(a) Los valores que aparecen en estas columnas representan la media de los cambios discretos en la probabilidad de elegir cada alternativa cuando cambia la modalidad de un atributo evaluados para cada uno de los individuos de la muestra.

Nótese que la rigidez del modelo *logit* ordenado impone que, cuando se compara un individuo del área metropolitana con uno residente en las zonas norte o sur de la isla, los cambios discretos en las probabilidades de elección de las alternativas $Y_i = 0$ y $Y_i = 2$ sean de signo contrario. Es decir, no puede considerarse que este cambio en las características del individuo produzca un efecto del mismo signo sobre las probabilidades de elección de las alternativas extremas. Pero este resultado es contrario a lo esperado de acuerdo con las frecuencias relativas observadas en la muestra y puede atribuirse a la rigidez que impone el modelo ordenado

frente a especificaciones más flexibles. El mismo problema se advierte si se analizan los cambios discretos cuando se modifica el nivel educativo del individuo. De acuerdo con este modelo, un cambio en la modalidad del atributo que identifica el nivel de estudios de la modalidad sin estudios a cualquier otra produce un efecto negativo en la probabilidad de no consumir vino, mientras que aumenta la probabilidad de consumir con alta frecuencia. Esta diferencia de signo en las probabilidades de elección de las alternativas extremas impide captar, como parece observarse en el análisis descriptivo, que el incremento del nivel educativo produce una tendencia hacia el consumo ocasional, mientras que el no consumo o el consumo de alta frecuencia pierden peso.

Las limitaciones del modelo ordenado se aprecian también a partir del cálculo de los cocientes de *odds-ratios*, definidos como cociente de los ratios $\Omega_j = \frac{P(Y_i \leq j)}{P(Y_i > j)}$ cuando cambian las modalidades de los atributos que definen al individuo, que permiten evaluar como se modifica el patrón de sustitución entre consumo de baja frecuencia y consumo de alta frecuencia. Si se denota por $\Omega_j(x_{i,m} = p)$ al valor del ratio Ω_j cuando el atributo $x_{i,m}$ toma la modalidad p , entonces

$$\frac{\Omega_j(x_{i,m} = p)}{\Omega_j(x_{i,m} = q)} = e^{\beta_m^q - \beta_m^p}$$
, siendo β_m^p y β_m^q los parámetros del modelo *logit* ordenado para las variables *dummies* que toman valor 1 cuando el atributo $x_{i,m}$ toma las modalidades p y q , respectivamente(10). Aunque es posible efectuar múltiples comparaciones de los patrones de sustitución entre alternativas, lo que proporciona una enorme cantidad de información útil para aproximar el conocimiento del proceso de decisión del individuo e identificar reacciones de individuos particulares cuando cambia alguno de los atributos que los definen, se ha optado por mostrar sólo aquellos cocientes de *odds-ratios* que ilustran algunos de los efectos más significativos. De acuerdo con estos ratios, el consumo de baja frecuencia es más probable entre los individuos residentes en el área metropolitana que entre los residentes en las zonas norte y sur y menos probable para los individuos con estudios que para los que declararon no tenerlos (cuadro 4).

(10) En el caso de que una de las dos modalidades consideradas sea la elegida como modalidad de referencia en la estimación del modelo, el parámetro correspondiente es nulo.

Cuadro 4
COCIENTES DE ODDS-RATIOS (LOGIT ORDENADO)

	$x_{i,m} = p ; x_{i,m} = q$		
$p \setminus q$	Norte		Sur
Metropolitana	1.2351		1.2553
$p \setminus q$	Primarios	Secundarios	Universitarios
Sin estudios	1.6505	1.3994	1.7380

Los comentarios anteriores muestran que la especificación del modelo *logit* ordenado es demasiado rígida para recoger efectos que, a veces, significan un incremento de las probabilidades de elección de alternativas extremas en detrimento de las intermedias o viceversa. En este sentido, el *logit* multinomial estándar parece una especificación más apropiada, dado que admite explícitamente cambios en los parámetros que intervienen en el componente determinístico de la utilidad asociada a cada alternativa. Las estimaciones de los parámetros de este último modelo deben interpretarse como diferencia entre los parámetros que intervienen en las funciones de utilidad de la alternativa j , $j > 0$, frente a la alternativa 0 (no consumo), pero reflejan también efectos diferenciados de los atributos considerados sobre las probabilidades de elección de alternativas que suponen consumir vino con distinta frecuencia.

En el caso particular del atributo que indica el área de residencia, los cambios discretos permiten apreciar que la probabilidad de no consumir es más alta entre los residentes de las zonas norte y sur que entre aquéllos que residen en el área metropolitana. Y las estimaciones de los parámetros sugieren que, con respecto a los individuos sin estudios, el incremento del nivel educativo tiene un efecto significativo sobre la probabilidad de que el individuo elija consumir ocasionalmente. Los cambios discretos muestran con más claridad que entre los individuos con algún tipo de estudios existe mayor probabilidad de elegir la opción anterior.

Los resultados comentados hasta ahora ponen ya de manifiesto que la especificación *logit* convencional es capaz de captar el proceso de decisión individual mejor que el *logit* ordenado. Pero una información más rica sobre los cambios en el patrón de sustitución entre alternativas puede obtenerse a partir de los cocientes de

odds-ratios, definidos ahora como cociente de los ratios $\Omega_{j/k} = \frac{P(Y_i = j)}{P(Y_i = k)}$, cuando uno de los atributos cambia de modalidad. Si se denota por $\Omega_{j/k}(x_{i,m} = p)$ al valor del ratio $\Omega_{j/k}$ cuando el atributo $x_{i,m}$ toma la modalidad p , entonces $\frac{\Omega_{j/k}(x_{i,m} = p)}{\Omega_{j/k}(x_{i,m} = q)} = e^{\beta_{m,j}^p - \beta_{m,k}^p - \beta_{m,j}^q + \beta_{m,k}^q}$, siendo $\beta_{m,j}^p$, $\beta_{m,j}^q$, $\beta_{m,k}^p$ y $\beta_{m,k}^q$ los parámetros en

el modelo *logit* para las variables *dummies* que toman valor 1 cuando el atributo $x_{i,m}$ toma las modalidades p o q en las probabilidades de elección de las alternativas j y k , respectivamente.

Estos coeficientes muestran que el modelo *logit* multinomial incorpora la flexibilidad necesaria con respecto a la especificación ordenada para recoger los efectos del área de residencia y de la educación sobre la frecuencia de consumo. Así, los ratios $\Omega_{1/0}$ y $\Omega_{1/2}$ muestran la mayor preferencia por el consumo ocasional entre los residentes en el área metropolitana. Del mismo modo, los dos ratios anteriores van creciendo conforme aumenta el nivel de estudios (cuadro 5).

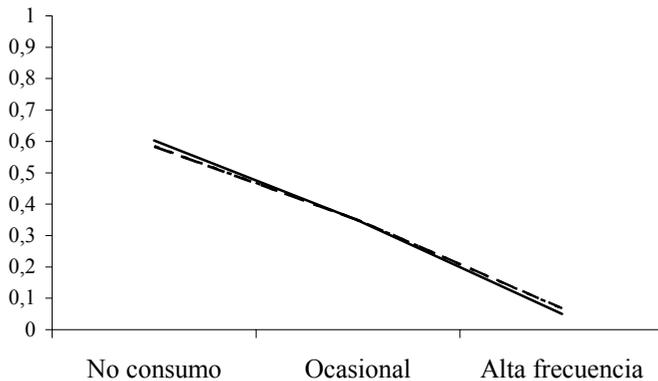
Cuadro 5

COCIENTES DE ODDS-RATIOS $\Omega_{1/k}$ (LOGIT MULTINOMIAL)		
p/q	$\frac{\Omega_{1/0}(x_{i,m} = p)}{\Omega_{1/0}(x_{i,m} = q)}$	$\frac{\Omega_{1/2}(x_{i,m} = p)}{\Omega_{1/2}(x_{i,m} = q)}$
Norte/Metropolitana	0.7319	0.4668
Sur/Metropolitana	0.7632	0.4917
Sin estudios/Estudios Primarios	0.4288	0.6336
Sin estudios/Estudios Secundarios	0.3708	0.4506
Sin estudios/Estudios Universitarios	0.2264	0.4287

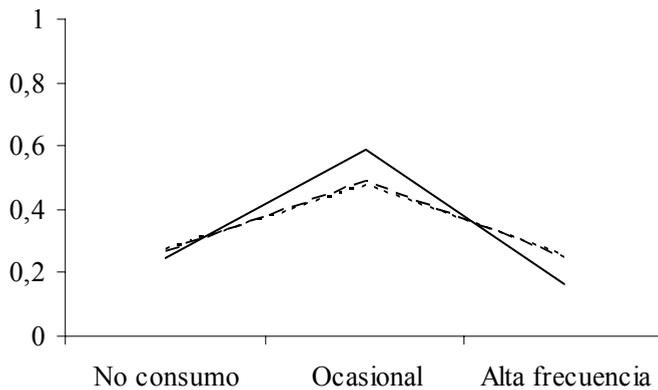
En general, las estimaciones del modelo captan una polarización en la frecuencia de consumo hacia las alternativas extremas entre los individuos residentes en las zonas norte y sur o con bajo nivel de estudios, mientras que el consumo de mediana frecuencia gana peso entre los individuos residentes en el área metropolitana o con mayor nivel educativo. Esta polarización se observa con claridad si se representan las medias de las probabilidades predichas para los individuos de la muestra cuando se modifica la modalidad de los atributos relativos al área de residencia y al nivel de formación (figuras 1 y 2).

Figura 1

PROBABILIDADES PREDICHAS SEGÚN ÁREA DE RESIDENCIA
LOGIT ORDENADO



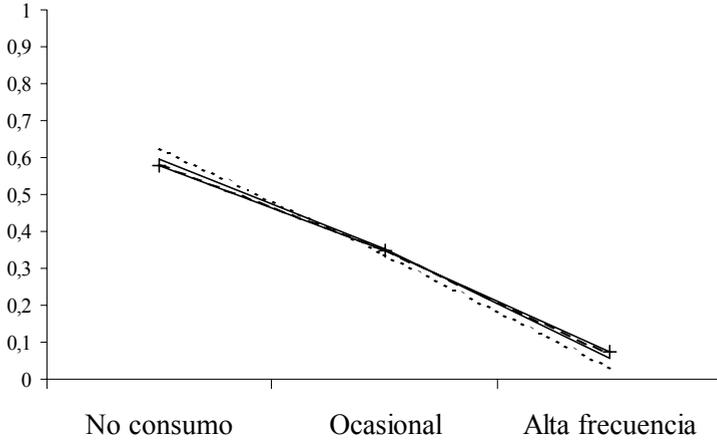
LOGIT MULTINOMIAL



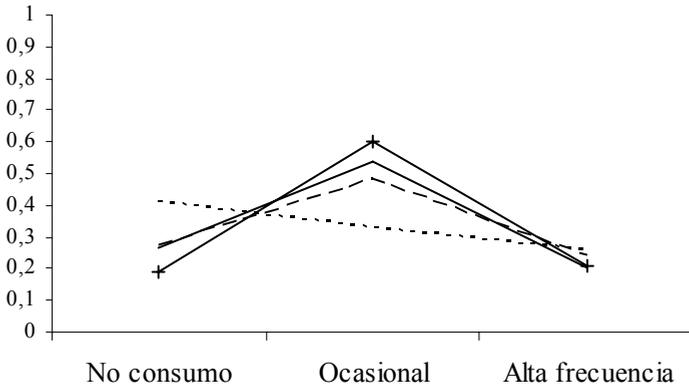
Área metropolitana: — ; Zona Norte: -----; Zona Sur: - · - · -

Figura 2

**PROBABILIDADES PREDICHAS SEGÚN NIVEL DE FORMACIÓN
LOGIT ORDENADO**



LOGIT MULTINOMIAL



Sin estudios: ; Primarios: - - - ; Secundarios: ——— ; Universitarios: +-+—

CONCLUSIÓN

Los modelos específicos para datos ordenados incorporan restricciones que facilitan la estimación y la interpretación. Ahora bien, a pesar de su uso generalizado en la literatura especializada, se trata de modelos que pueden ser demasiado rígidos para recoger adecuadamente los mecanismos subyacentes en los procesos de decisión individuales cuando el orden de las alternativas no implica un orden subyacente de utilidades asociadas a éstas. Esta característica impide captar el efecto de determinados atributos que influyen en el mismo sentido en las probabilidades de elección de las alternativas extremas. De ahí que, en general, resulte más apropiado acudir a especificaciones no ordenadas, que, en general, poseen mayor capacidad para ajustarse a los mecanismos que guían la decisión del individuo.

REFERENCIAS

- BASILE R, A GIUNTA Y JB NUGENT (2003) «Foreign expansion by Italian manufacturing firms in the nineties: An ordered probit analysis», *Review of Industrial Organization*, 23 (1): 1-24.
- BHAT CR (2003) «Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences» *Transportation Research Part B*, 37 (9): 837-855.
- BOZ I Y C AKBAY (2005) «Factors influencing the adoption of maize in Kahramanmaras province of Turkey», *Agricultural Economics*, 33 (3): 431-440.
- CHOO S, GO COLLANTES Y MOKHTARIAN PL (2005) «Wanting to travel, more or less: Exploring the determinants of the deficit and surfeit of personal travel», *Transportation*, 32 (2): 135-164.
- CONTOYANNIS P, AM JONES Y N RICE (2004) «The dynamics of health in the British household panel survey», *Journal of Applied Econometrics*, 19 (4): 473-503.
- FOSGERAU M Y TB BJORNER (2006) «Joint models for noise annoyance and willingness to pay for road noise reduction», *Transportation research Part B-Methodological*, 40 (2): 164-178.
- GREENE WH (2002a) *Limdep 8.0.: Econometric Modelling Guide*, Econometric Software, Inc. Vol 1 y 2.
- GREENE WH (2002b) *Nlogit 3.0.: Reference Guide*, Econometric Software, Inc.

- GUIRAO G, JJ CÁCERES, VJ CANO, M HERNÁNDEZ, MI LÓPEZ, FJ MARTÍN Y MC RODRÍGUEZ (2001) *El consumo de vino en Tenerife*, Servicio Técnico de Desarrollo Rural y Pesquero, Cabildo Insular de Tenerife.
- JAGGIA S Y S THOSAR (2004) «The medium-term aftermarket in high-tech IPOs: Patterns and implications», *Journal of Banking & Finance*, 28 (5): 931-950.
- LAUER C (2003) «Family background, cohort and education: A French-German comparison based on a multivariate ordered probit model of educational attainment», *Labour Economics*, 10 (2): 231-251.
- LÁZARO N, ML MOLTÓ Y R SÁNCHEZ (2004) «Paid employment and unpaid caring work in Spain», *Applied Economics*, 36 (9): 977-986.
- LUCE D (1959) *Individual Choice Behaviour*, New York, John Wiley and Sons.
- McFADDEN D (1973) «Conditional analysis of qualitative choice models» en P Zarembka, (ed.), *Frontiers in Econometrics*, New York, Academic Press.
- MCKELVEY RD Y W ZAVOINA (1975) «A statistical model for the analysis of ordinal level dependent variables», *Journal of Mathematical Sociology*, 4: 103-120.
- NDJEUNGA J Y CH NELSON (2005) «Toward understanding household preference for consumption characteristics of millet varieties: a case study from western Niger», *Agricultural Economics*, 32 (2): 151-165.
- PUDNEY S Y M SHIELDS (2000) «Gender, race, pay and promotion in the British nursing profession: Estimation of a generalized ordered probit model», *Journal of Applied Econometrics*, 15 (4): 367-399.
- RICHARDS TJ Y PM PATTERSON (2005) «Retail price fixity as a facilitating mechanism», *American Journal of Agricultural Economics*, 87 (1): 85-102.
- SMITH J Y R TAYLOR (2001) «Determinants of degree performance in UK universities: a statistical analysis of the 1993 student cohort», *Oxford Bulletin of Economics and Statistics*, 63 (1): 29-60.
- STORCHMANN K (2005) «English weather and Rhine wine quality: An ordered Probit model», *Journal of Wine Research*, 16, 2: 105-119.
- THEIL H (1969) «A multinomial extension of the linear logit model», *International Economic Review*, 10: 251-259.
- THURSTONE L (1927) «A law of comparative judgement», *Psychological Review*, 34: 273-286.
- TRAIN K (2003) *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge, UK.

VAN DOORSLAER E Y AM JONES (2003) «Inequalities in self-reported health: validation of a new approach to measurement» *Journal of Health Economics*, 22 (1): 61-87.

DISCRETE CHOICE MODELS AND ORDERED SPECIFICATIONS: A METHODOLOGICAL REFLECTION

ABSTRACT

In this paper, it is shown that ordered formulations may be too restrictive to capture the nature of an individual decision process guided by the maximization of utility if the order between alternatives does not match with the order of utilities related to them. To illustrate the disability of ordered formulations the results of estimating a standard logit and an ordered logit are compared. These models are used to identify the effect of individual features on the decisions about the frequency of wine consumption adopted by Tenerife residents.

Key words: discrete choice models, multinomial specifications, ordered alternatives.

AMS classification: 62P20