

# UN MODELO LOGIT MULTINOMIAL MIXTO DE TENENCIA DE VIVIENDA

JAVIER A. BARRIOS GARCÍA  
JOSÉ E. RODRÍGUEZ HERNÁNDEZ  
*Universidad de La Laguna*

En este trabajo, basándonos en una muestra de sección cruzada procedente de la Encuesta Continua de Presupuestos Familiares (ECPF) base 1997, desarrollamos un modelo logit multinomial mixto que persigue aproximar las variables fundamentales que intervienen en las decisiones que adoptan los individuos a nivel nacional en materia de tenencia de vivienda durante el año 1999, considerándose cuatro supuestos en función del régimen de tenencia (propiedad o alquiler) y del tipo de edificio (unifamiliar o colectivo). Con este objetivo elaboramos previamente una medida de la renta permanente imputada a los individuos, así como un índice de precios hedónicos por comunidades autónomas. Procedemos a continuación a la estimación de un modelo logit multinomial mixto para la decisión de tenencia de vivienda que nos permitirá contemplar cierta heterogeneidad no observada en el comportamiento de los individuos, así como superar las hipótesis de independencia de alternativas irrelevantes y homoscedasticidad propias de los modelos logit multinomial y anidado.

*Palabras clave:* Mercado de vivienda, régimen de tenencia de la vivienda, logit mixto.

*Clasificación JEL:* C25, R21.

**E**l mercado de la vivienda se encuentra de rabiosa actualidad en España. En los últimos años hemos sido testigos de un amplio debate social en torno a este tema, alimentado fundamentalmente por una continua escalada de precios derivada de un amplio shock de demanda (en sus dos componentes, como bien de inversión y de consumo), cuyos principales efectos han sido el deterioro en la accesibilidad a una vivienda para los estratos de población más desfavorecidos, así como el creciente endeudamiento de las familias [García Montalvo (2003)].

El presente trabajo pretende colaborar en alguna medida a esta discusión al marcarse como meta estudiar los factores que determinan la elección del régimen de tenencia de vivienda dentro del ámbito nacional, considerándose ésta bajo cuatro tipos principales: propiedad de una vivienda unifamiliar, propiedad de una vivienda colectiva, alquiler de una vivienda unifamiliar y alquiler de una vivienda colectiva.

Con este objetivo proponemos un modelo logit multinomial mixto que persigue explicar el proceso de toma de decisiones de los individuos entre las cuatro

alternativas consideradas. Este tipo de especificación nos permitirá superar las hipótesis simplificadoras propias de la familia de modelos logit más ampliamente utilizados hasta el momento en la modelización de la elección de tenencia de vivienda, el logit multinomial y el logit anidado, y que se corresponden con los supuestos de independencia de alternativas irrelevantes y homocedasticidad.

Esta línea de investigación supone una continuación de la emprendida en Rodríguez y Barrios (2003) para analizar la elección entre alquilar o adquirir en propiedad una vivienda, distinguiéndose en cada caso el mercado libre y el de protección oficial, si bien ahora incorporamos diversas aportaciones relevantes. En primer lugar, ampliamos a todo el territorio nacional la cobertura geográfica de la muestra de sección cruzada empleada (Encuesta Continua de Presupuestos Familiares, ECPF, base 1997), a la vez que actualizamos la misma para el año 1999. A su vez, para evitar el problema de endogeneidad en la variable precio se construyen índices de precios hedónicos a nivel regional para cada uno de los tipos de tenencia de vivienda analizados partiendo de una “cesta” de características para una vivienda estándar determinada a nivel nacional<sup>1</sup>. Finalmente, incluimos una medida de la renta permanente de los individuos en lugar de la renta corriente.

Todo ello nos llevará a la estimación de un modelo logit multinomial mixto, novedoso en la metodología econométrica, que ajusta mejor los datos que el logit multinomial o incluso que un modelo logit multinomial mixto cuando se considera la renta corriente. Además, los efectos marginales obtenidos tanto para la renta permanente como para los precios son mayores en el modelo logit multinomial mixto propuesto que en el logit multinomial, acentuándose esta tendencia en el caso de la renta cuando se considera la renta permanente en lugar de la corriente.

La organización del trabajo es la siguiente. El apartado 1 recoge los principales antecedentes, tanto nacionales como internacionales, así como expone el modelo econométrico de elección de tenencia de vivienda que consideramos. En el apartado 2 se definen las variables que intervienen y las fuentes estadísticas empleadas. En el 3 se presenta el modelo de elección de tenencia estimado y se analizan los resultados obtenidos. Por último, en el epígrafe 4 se detallan las conclusiones del trabajo.

## 1. ANTECEDENTES Y MODELO ECONOMÉTRICO DE TENENCIA DE VIVIENDA

### 1.1. Antecedentes

Entre los principales antecedentes teóricos que analizan los determinantes de la elección entre formas de tenencia de la vivienda, cabe destacar el trabajo seminal ya clásico de Henderson y Ioannides (1983), así como las contribuciones posteriores de Goodman (1995, 2002) en las que se estudia simultáneamente, a lo largo de un cierto horizonte temporal, la decisión de mudarse o permanecer en la vivienda así como la demanda de vivienda derivada (que no la elección de tenen-

---

(1) En Rodríguez y Barrios (2003) se incluye la variable precio de cada alternativa a través de sendas regresiones hedónicas en las que se estima éste para cada individuo a partir de las características de la vivienda ya seleccionada bajo un determinado régimen de tenencia.

cia) en presencia de costes de transacción, y donde la utilidad del consumidor puede variar en el tiempo, particularmente con respecto a variables relativas al ciclo de vida (edad, número de hijos, etc.). Este modelo, planteado en términos de un problema de programación dinámica, permite deducir el plan óptimo de mudanza-consumo (de otros bienes y de vivienda) a lo largo de  $T$  períodos. Del trabajo de Goodman se deduce que una medida de la renta que obtiene el individuo en cada período a lo largo de su horizonte vital (renta permanente), el precio de la vivienda y de otros bienes de consumo, así como las preferencias a lo largo del tiempo del consumidor (que pueden venir determinadas por características socio-demográficas del mismo) influyen sobre la cantidad demandada de vivienda en cada estancia, incluso para los individuos que deciden no mudarse.

El modelo de Goodman (1995) es completado más tarde por Ioannides y Kan (1996) y Gobillon y Le Blanc (2002) para incorporar explícitamente la elección de tenencia y, en el último caso además, restricciones en el mercado hipotecario. En ambos trabajos se modeliza teóricamente el comportamiento dinámico de los individuos enfrentados simultáneamente a las decisiones de movilidad residencial, elección de tenencia de vivienda (entendiéndose éstas bajo tres supuestos: permanecer en la vivienda actual, mudarse y alquilar o mudarse y comprar) y cantidad de vivienda consumida (y de otros bienes de consumo), basándose en un marco de programación dinámica en el que los individuos proceden a lo largo de su ciclo vital realizando una secuencia de decisiones adoptadas a medida que poseen nueva información. Así, el problema del consumidor consistirá en determinar sus sendas de consumo de otros bienes, de vivienda y de decisiones de mudanza-tenencia de vivienda de cara a maximizar la utilidad esperada a lo largo de su vida.

Si bien la contrastación empírica más adecuada al marco teórico anterior se debe fundamentar en datos de panel, ante la carencia de datos de esta naturaleza que contengan información y longitud adecuadas, numerosos esfuerzos se han dirigido hacia su corroboración a través de datos de corte transversal, subsumiendo las ecuaciones de Bellman correspondientes al problema dinámico considerado en un marco estático basándose en los modelos de elección discreta inspirados en el paradigma de la maximización de la utilidad aleatoria [McFadden (2000)].

En este sentido cabe citar el trabajo de Lee y Trost (1978), que aborda un modelo probit que analiza simultáneamente la elección de tenencia y el gasto en vivienda, al igual que Jaén y Molina (1994) para el caso andaluz. Duce Tello (1995) especifica para España un modelo multinomial con racionamiento similar al desarrollado por King (1980), considerando cuatro alternativas: propiedad de vivienda de protección oficial, alquiler con renta antigua (previo al Decreto Boyer de 1985), propiedad de vivienda de renta libre y el alquiler libre (posterior a 1985). Börsch-Supan y Pitkin (1988), Bourassa (1995) o Walker *et al.* (2002) especifican sendos modelos logit multinomiales para la elección entre diferentes regímenes de tenencia de vivienda, mientras que Barrios y Rodríguez (2003) desarrollan un modelo logit multinomial para el estudio de la elección de tenencia, considerándose ésta bajo cuatro supuestos: propiedad o alquiler, bien sea en el mercado libre o en el de Protección Oficial. Por otra parte, diferentes trabajos emplean alternativamente el modelo logit multinomial anidado cuando el individuo se enfrenta a más de dos alternativas de decisión [Börsch-Supan y Pitkin (1988), Skaburskis (1999), o Colom

y Cruz (1997) para el caso español, este último considerando idénticas alternativas que en el presente trabajo], presentando éste mejores resultados empíricos en las comparaciones realizadas frente al logit multinomial.

No obstante, debido a las grandes limitaciones que presentan los modelos logit multinomial y anidado, y alentados por los avances tecnológicos en computación y métodos numéricos que posibilitan el uso de modelos más generales, se ha venido desarrollando en los últimos años el denominado modelo logit multinomial mixto, el cual da lugar a patrones de sustitución entre alternativas más complejos. En el campo particular de la demanda de vivienda, podemos citar como precursores en la aplicación del modelo logit multinomial mixto, los trabajos de Rouwendal y Meijer (2001), Börsch-Supan *et al.* (2001), así como Rodríguez y Barrios (2003), siendo este último el único precedente a nivel español.

### 1.2. Modelo econométrico de tenencia de vivienda

En este trabajo se pretende analizar conjuntamente el estudio del régimen de tenencia y tipo de edificio, estableciéndose para ello cuatro tipos de tenencia de vivienda: propiedad unifamiliar, propiedad colectiva, alquiler unifamiliar y alquiler colectivo. Para llevar a cabo esta labor consideraremos como modelo de elección entre estas alternativas no ordenadas el denominado logit multinomial mixto (*mixed logit*, también conocido como modelo logit con componentes de error o modelo logit con parámetros aleatorios), el cual se ha venido desarrollando de manera notable en los últimos años, siendo aplicado en un principio al estudio de la demanda de transporte [Boyd y Melman (1980), Cardell y Durban (1980), Brownstone y Train (1999), Train (2003)].

La principal ventaja del modelo logit multinomial mixto frente a los modelos de elección múltiple más utilizados hasta el momento, el modelo logit multinomial [McFadden (1974)] y el logit jerárquico o anidado [Williams (1977)], es su capacidad para afrontar situaciones en las que se vulneran las hipótesis de independencia de alternativas irrelevantes (IAI) y de homoscedasticidad propias de estos últimos (en el modelo logit anidado se considera una partición disjunta de las alternativas en grupos de tal forma que intra-grupos se satisface la hipótesis de IAI y de homoscedasticidad, pero no inter-grupos).

En los últimos tiempos han surgido numerosas aplicaciones del modelo logit multinomial mixto [Train (2003)] alentadas no sólo por su mayor flexibilidad para adaptarse a los patrones complejos de sustitución entre alternativas, sino también, y no menos importante, por el hecho de que permite aproximar cualquier modelo de maximización de la utilidad aleatoria mediante una selección adecuada de la especificación aleatoria [McFadden y Train (2000), McFadden (2000)].

El modelo logit multinomial mixto puede ser motivado desde dos puntos de vista diferentes pero formalmente equivalentes: bien como un modelo logit con componentes de error [Brownstone y Train (1999)], o bien mediante un modelo logit con parámetros aleatorios [Revelt y Train (1998)]. En nuestro trabajo seguiremos una especificación del tipo de parámetros aleatorios, cuya fundamentación se basa en permitir cierta heterogeneidad inobservada entre los individuos en cuanto a su respuesta frente a las variables exógenas observadas. Desde este punto

de vista, los individuos se enfrentan a la elección entre  $J$  alternativas, modelizándose la utilidad obtenida por el individuo  $i$  a causa de la alternativa  $j$  como:

$$U_{ij} = \beta_{ij}^t X_{ij} + \varepsilon_{ij} \quad i = 1, \dots, N; j = 1, \dots, J \quad [1]$$

donde  $X_{ij} \in \mathbb{R}^q$  es un vector de variables observables exógenas (relativas al individuo  $i$  y alternativa  $j$ ),  $\beta_i^t = (\beta_{i1}^t \dots \beta_{iJ}^t) \in \mathbb{R}^{q \times J}$  es un vector de parámetros que varía entre los individuos con densidad conjunta  $f(\beta|\Omega)$ , donde  $\Omega$  representa los parámetros fijos de esta distribución<sup>2</sup>, y,  $\varepsilon_{ij}$  constituye el término aleatorio que se supone independiente e idénticamente distribuido (iid) Gumbel sobre alternativas e individuos<sup>3</sup>.

Dado el valor de  $\beta_i$ , la probabilidad de que el individuo  $i$  escoja la alternativa  $j$  condicionada a este valor de  $\beta_i$  vendrá determinada como en el modelo logit multinomial:

$$L_{ij}(\beta_i) = \frac{e^{\beta_{ij}^t X_{ij}}}{\sum_{k=1}^J e^{\beta_{ik}^t X_{ik}}} \quad [2]$$

Y por tanto, la probabilidad de elección (no condicionada) vendrá dada por la integral múltiple:

$$P_{ij}(\Omega) = \int_D L_{ij}(\beta) f(\beta|\Omega) d\beta \quad D = \{\beta \in \mathbb{R}^{q \times J}\} \quad [3]$$

Como se puede apreciar, la probabilidad de elección de un modelo logit mixto no posee en general una expresión matemática cerrada, a diferencia del logit multinomial o del anidado, puesto que la integral anterior no podrá resolverse en general analíticamente. Una estrategia para facilitar el cálculo de estas probabilidades de elección consiste en evaluar la integral en [3] a través de métodos de simulación, esto es: para un valor dado de los parámetros en  $\Omega$  se obtiene un valor de  $\beta$  aleatoriamente de su distribución, y se calcula  $L_{ij}(\beta)$  a partir de [2]. Este proceso se repite para cierto número de valores de  $\beta$  (réplicas) obtenidos aleatoriamente de su distribución, y la media de las probabilidades de elección condicionadas  $L_{ij}(\beta)$  calculadas se toma como la probabilidad de elección aproximada o simulada (SP):

$$SP_{ij} = \frac{1}{R} \sum_{r=1}^R L_{ij}(\beta^r)$$

donde  $R$  es el número de replicaciones de  $\beta$ , y  $\beta^r$  es la replicación  $r$ -ésima.

(2) En la práctica se deben suponer distribuciones específicas para los parámetros aleatorios, siendo las más utilizadas la normal, la lognormal, la uniforme o la triangular, aunque puede ser considerada cualquier otra. El conjunto  $\Omega$  incluirá los parámetros que definen estas distribuciones. Por ejemplo, si la distribución conjunta es normal multivariante,  $\Omega$  contendrá el vector de medias y la matriz de varianzas-covarianzas de los parámetros.

(3) En general, la función de distribución Gumbel o de valor extremo tipo I vendrá definida por:  $F(x) = e^{-e^{-\mu(x-\eta)}}$ , para  $\mu > 0$ ,  $\eta \in \mathbb{R}$ .

Al mismo tiempo, se puede disminuir la varianza del estimador anterior si en lugar de emplear replicaciones aleatorias independientes de los parámetros utilizamos replicaciones “inteligentes” que se encuentren negativamente correlacionadas entre sí y que proporcionen a la vez una mejor cobertura del recinto de integración. El método de replicar a través de sucesiones Halton se ha revelado recientemente mucho más eficiente que las replicaciones aleatorias independientes, permitiendo reducir drásticamente el tiempo de computación [Train (2003), págs. 252-269].

De esta forma, para estimar los parámetros desconocidos del modelo, la función de verosimilitud logarítmica:  $\sum_{i=1}^N \text{Ln}(P_{ij})$  es aproximada por la función de verosimilitud logarítmica simulada:  $\sum_{i=1}^N \text{Ln}(SP_{ij})$  y los parámetros estimados son los que maximizan esta última. Se comprueba que, bajo condiciones de regularidad, el estimador así obtenido es consistente y asintóticamente normal. Además, cuando el número de replicaciones crece más rápido que la raíz cuadrada del número de observaciones, el estimador es asintóticamente equivalente al estimador máximo verosímil [Hajivassiliou y Ruud (1994)].

Obsérvese por otra parte que para el modelo logit multinomial mixto, el cociente entre las probabilidades de elección de dos alternativas diferentes determinadas por [3] depende de las características del resto de alternativas, y por tanto no se encuentra limitado por el supuesto de independencia de alternativas irrelevantes. Además, si  $\beta$  es una variable aleatoria múltiple con media  $b$  y desviación  $\mu$ , la expresión [1] se puede escribir también como:

$$U_{ij} = b_{ij}^t X_{ij} + [\mu_{ij}^t X_{ij} + \varepsilon_{ij}]$$

y, llamando  $\xi_{ij} = \mu_{ij}^t X_{ij} + \varepsilon_{ij}$ , obtenemos la expresión en forma de componentes de error del modelo logit mixto, donde  $\xi_{ij}$  representará un término aleatorio constituido por la suma de una componente Gumbel  $\varepsilon_{ij}$  independiente e idénticamente distribuida sobre alternativas e individuos, y por  $\eta_{ij} = \mu_{ij}^t X_{ij}$  que es un término aleatorio de media cero cuya distribución sobre los individuos y alternativas depende en general de ciertos parámetros subyacentes y de datos observables acerca del individuo y de la alternativa  $j$ , y que permite recoger la presencia de correlación y heteroscedasticidad entre los términos no observables de la utilidad de las alternativas ( $\xi_{ij}$ ) puesto que:

$$E[\xi_{ik} \cdot \xi_{ij}] = E[(\eta_{ik} + \varepsilon_{ik})(\eta_{ij} + \varepsilon_{ij})] = E[\eta_{ik} \cdot \eta_{ij}] \neq 0, \text{ en general}$$

## 2. DATOS Y VARIABLES UTILIZADOS

La fuente estadística para este trabajo ha sido la Encuesta Continua de Presupuestos Familiares (ECPF) base 1997. Esta contempla suficiente información tanto sobre los hogares como sobre los miembros que componen el mismo.

En este estudio sólo se han seleccionado los hogares de la ECPF que disfrutaban de la vivienda en propiedad o en alquiler, eliminándose aquellos hogares que

disfruten de la misma por cesión gratuita o semigratuita, los cuales constituyen un 1,4% del total. También se excluyen aquellos situados en Ceuta y Melilla en los que no se disponía de un tamaño de muestra adecuado para construir un índice de precios hedónico, así como todos aquellos de los que no se dispone información sobre alguna de las variables consideradas. Además, para no reducir aún más el tamaño de la muestra debido a los turnos de rotación trimestral propios del diseño de la ECPF, hemos adoptado los datos de un trimestre determinado del año 1999, concretamente el 2º trimestre, efectuándose las correspondientes anualizaciones (en proporción directa) para las variables de renta corriente y alquileres (reales e imputados) que se consideran en términos nominales.

En la línea de la literatura que aborda la elección de tenencia, se han incluido tres tipos de variables explicativas: características sociodemográficas del hogar, características económicas y otras características.

### 2.1. Características sociodemográficas del hogar

Incluimos en este grupo diversas características sociodemográficas del sustentador principal así como el número de miembros del hogar.

- Sexo: Valores: 0 = Mujer; 1 = Varón.
- Estudios: Recogido con tres variables *dummies* referidas al máximo nivel de estudios alcanzado por el sustentador principal: Estud1 = Sin estudios o estudios primarios (variable de referencia), Estud2 = Estudios secundarios, Estud3 = Estudios superiores o universitarios.
- Estado civil: Definido con dos variables *dummies*: casado = Casado, y otrasit = Otra situación: soltero, viudo, separado o divorciado (variable de referencia).
- Edad (Edad, Edad2): Se incluye la edad del sustentador principal y edad al cuadrado.
- Número de miembros (NMiemb): Refleja el número de miembros del hogar, incluyendo al sustentador principal.

### 2.2. Características económicas

Renta permanente (Lnrep): La noción relevante de renta a la hora de analizar la elección realizada por el consumidor sobre el régimen de tenencia de la vivienda debe incluir un concepto más amplio que el de renta corriente anual que permita recoger en su lugar la renta a largo plazo o permanente [Friedman (1957)].

Para obtener una medida de la renta permanente, y dada la disponibilidad de información contenida en la ECPF, hemos optado por la línea ya clásica de Goodman y Kawai (1982), en la que se estima ésta mediante el uso de un modelo de capital humano en donde la renta viene determinada por la inversión en capital humano y no humano. De esta manera, se podría escribir:

$$Y_C = Y_T \cdot Y_P = Y_T \cdot Y_P(H,N)$$

Esto es, la renta corriente ( $Y_C$ ) viene determinada por las desviaciones transitorias ( $Y_T$ ) de la renta permanente ( $Y_P$ ), donde  $Y_P$  constituye una función (posible-



mente no lineal) de los recursos de capital humano (H) (como son la educación, edad y formación) y los recursos de capital no humano (N) que posee el individuo. De este modo, la regresión del logaritmo de la renta corriente ( $\text{Ln}Y_C$ ) sobre las variables relativas al capital humano y no humano que posee el individuo nos proporcionaría una estimación del logaritmo de la renta permanente, constituyendo la parte residual de la regresión la componente transitoria (en logaritmo).

Las variables consideradas y los resultados de la estimación de la renta permanente imputada a los individuos se recogen en el apéndice 1.

**Precio:** La variable precio es un atributo de las alternativas que recoge los precios de distintos regímenes de tenencia para cada individuo (Precio compra de una vivienda unifamiliar o colectiva y alquiler de una vivienda unifamiliar o colectiva).

Es importante señalar que la variable precio para cada una de las alternativas no viene recogida en la ECPF. Sin embargo, frente a trabajos como el de González (1997) que opta sencillamente por omitirla por falta de información muestral, o por recogerla a través de índices de precios como es el caso de Lee y Trost (1978), Börsch-Supan y Pollakowski (1990), o Jaén y Molina (1994) (este último llega incluso al resultado antinatural de que el tipo de tenencia de vivienda en nuestro país no depende de forma significativa de los precios de compra y alquiler), o en una estimación de costes de uso anuales de la vivienda [King (1980) o Börsch-Supan y Pitkin (1988)], nosotros hemos preferido, al igual que Goodman y Kawai (1982), Thibodeau (1995) o Rapaport (1997), estimar índices de precios de vivienda hedónicos para cada comunidad autónoma que puedan servir como una aproximación de la valoración subjetiva que realizan los individuos en cada mercado regional en torno a la vivienda y sus diferentes regímenes de tenencia.

En el apéndice 2 se recoge la metodología empleada para obtener los índices de precios hedónicos por comunidades autónomas considerados.

### 2.3. Otras características

**Tamaño del municipio (Tmun1, Tmun2, Tmun3):** Indica el tamaño del municipio donde el individuo ha decidido vivir. Se recoge mediante tres variables *dummies*: Tmun1: municipio con más de 100.000 habitantes, Tmun2: municipio con más de 20.000 y menos de 100.000 habitantes, Tmun3 (variable de referencia): municipio con menos de 20.000 habitantes.

Adviértase que al igual que en Rodríguez y Barrios (2003), implícitamente realizamos la hipótesis de que los individuos deciden previamente el tamaño del municipio en el que quieren fijar su residencia y por lo tanto se entiende éste como una característica más de los mismos. Si admitiéramos que los individuos eligen simultáneamente el régimen de tenencia, el tipo de edificio y el municipio de residencia, estaríamos ante un modelo de elección discreta con 12 alternativas (bajo la desagregación del tamaño del municipio considerada), que de ser analizado como un logit multinomial mixto elevaría enormemente la complejidad y los tiempos de computación necesarios para su estimación, pudiendo incluso hacerse ésta inviable para determinadas alternativas (en alquiler fundamentalmente) donde quedaría reducida excesivamente la muestra.

**Comunidades por nivel de precios (Ccaa1, Ccaa2, Ccaa3):** Seguimos aquí la clasificación de las comunidades autónomas según el diferente nivel de precios resi-



denciales para el año 1998 establecido en Taltavull (2000) (pág. 212, basado en la estadística de precio medio de la vivienda nueva y usada publicada por el Ministerio de Fomento). Para ello se han creado tres variables *dummies*: Ccaa1: comunidades con precios superiores a la media nacional (Balears, Cataluña, Madrid, Navarra y País Vasco); Ccaa2: comunidades con precios entre el 80-100% de la media nacional (Aragón, Asturias, Canarias, Cantabria, Castilla y León, Galicia y Rioja); Ccaa3 (variable de referencia): comunidades con precios menores al 80% de la media nacional (Andalucía, Castilla la Mancha, C. Valenciana, Extremadura y Murcia).

Cuadro 1: ESTADÍSTICOS DESCRIPTIVOS DE LA MUESTRA

	Media	Desviación típica
Nmiemb	3,188	1,303
Edad	53,913	14,659
Edad2	3.121,390	1.636,876
Sexo	0,849	0,358
Estud2	0,299	0,458
Estud3	0,175	0,380
Ccaa1	0,350	0,477
Ccaa2	0,317	0,465
Tmun1	0,517	0,500
Tmun2	0,189	0,392
Casado	0,802	0,399
LnR (Y <sub>p</sub> )	14,557	0,437
LnR (Y <sub>c</sub> )	14,550	0,584
Precio PU (en Ln)	13,180	0,256
Precio PC (en Ln)	13,200	0,173
Precio AU (en Ln)	12,593	0,621
Precio AC (en Ln)	12,750	0,479
Número observaciones:	4.471	

### 3. MODELO DE ELECCIÓN DE TENENCIA DE VIVIENDA ESTIMADO

En este trabajo estimamos un modelo logit multinomial mixto en el que se ha considerado como aleatorios los coeficientes de la variable renta en la alternativa de propiedad unifamiliar, y de edad en las alternativas en propiedad colectiva y alquiler unifamiliar; siendo el resto de coeficientes constantes. Para estos tres coeficientes se especifica una distribución normal trivariante, admitiéndose la existencia de correlación entre ellos. La selección de los coeficientes aleatorios se ha realizado atendiendo a los resultados obtenidos con el test de especificación para

el modelo logit multinomial mixto propuesto por McFadden y Train (2000), del que se concluye que la conducta observada de los individuos presenta heterogeneidad en cuanto a las variables renta permanente y edad, frente al resto de variables, principalmente en las alternativas mencionadas. Adviértase además que entre los factores explicativos incluidos, el precio es la única característica intrínseca de las alternativas que barajan los individuos y, consecuentemente, figura con el mismo coeficiente en la expresión de la utilidad derivada de cada una de ellas.

Para llevar a cabo la estimación del modelo por el método de máxima verosimilitud simulada se ha utilizado el *software* Nlogit 3.0, empleándose 200 repeticiones Halton en nuestra muestra de estudio (4.471 observaciones), tras comprobarse que a partir de esa cifra los resultados no eran sensibles al número de repeticiones adoptado. En el cuadro 2 se detallan a modo comparativo los resultados de la estimación de los modelos logit multinomial y logit mixto que emplean la renta permanente ( $Y_P$ ) como medida de renta ( $\ln R$ ), y el modelo mixto que resulta del uso de la variable renta corriente ( $Y_C$ ) (el alquiler colectivo se toma como alternativa de referencia). En el cálculo de los errores estándar no se ha tenido en cuenta la introducción de variables explicativas predichas anteriormente (precios y renta permanente).

Los cuadros 3, 4 y 5 recogen los efectos marginales (calculados como promedio sobre el total de la muestra) de cada uno de los modelos. En el caso del modelo logit multinomial mixto, para calcular el efecto marginal del cambio en el atributo  $m$  del vector  $X_{ij}$  (con coeficiente  $\beta_{ij}^m$  en la expresión de  $U_{ij}$ ) sobre la probabilidad de que el individuo  $i$  escoja la alternativa  $j$  se tendrá que evaluar la integral:  $\int_D \beta_{ij}^m L_{ij}(\beta) (1 - L_{ij}(\beta)) f(\beta|\Omega) d\beta$ , mientras que su efecto sobre la probabilidad de la alternativa  $k \neq j$  vendrá dado por:  $\int_D \beta_{ij}^m L_{ij}(\beta) L_{ik}(\beta) f(\beta|\Omega) d\beta$ . Ambas integrales se calculan para cada individuo de la muestra empleando métodos de simulación. El cuadro 6 incluye la estimación obtenida de la matriz de varianzas-covarianzas para los parámetros considerados aleatorios en el modelo logit multinomial mixto con renta permanente<sup>4</sup>.

Los resultados contenidos en el cuadro 2 muestran claramente que las varianzas de los parámetros aleatorios considerados en el modelo logit mixto (bajo  $Y_P$  e  $Y_C$ ) son todas significativas, lo que indica que en realidad dichos parámetros no son fijos sino que varían entre los distintos individuos. En otras palabras, se evidencia que el peso que desempeñan las variables renta y edad en el modelo de elección de tenencia no es percibido de la misma manera por todos los individuos. En consecuencia, y a pesar de recoger diversas características de los individuos como factores explicativos de su proceso de toma de decisiones, se constata que aún persiste cierta heterogeneidad inobservada en la conducta de los mismos reflejada en el comportamiento aleatorio de los coeficientes de las variables renta y edad.

(4) El *software* Nlogit estima la matriz de Choleski  $\Gamma$  junto con las medias de los parámetros aleatorios y el resto de coeficientes fijos del modelo. El factor de Choleski representa la factorización de la matriz de varianzas-covarianzas  $V$  asociada a los parámetros aleatorios del modelo, esto es, aquella matriz cuadrada triangular inferior del mismo orden que  $V$  tal que  $\Gamma \cdot \Gamma' = V$ . Para obtener los errores estándar asociados a la matriz de varianzas-covarianzas aplicamos el método delta, tal y como se procede por ejemplo en Revelt y Train (1998). No recogemos la estimación obtenida para la matriz de Choleski por parecer de interpretación en sí misma.

Cuadro 2: ESTIMACIÓN DE LOS MODELOS LOGIT MULTINOMIAL Y MIXTO

Logit multinomial (Yp)		Logit mixto (Yp)		Logit mixto (Yc)		
Variables	Coefic.	Estad. t	Coefic.	Estad. t	Coefic.	Estad. t
Propiedad Unifamiliar (PU)						
Cte	-8,9638	-3,044	-15,761	-1,965	-14,144	-1,787
Nmiemb	0,1640	2,677	0,3141	1,959	0,3875	1,563
Edad	0,1174	3,997	0,3122	2,723	0,4392	2,178
Edad2	-0,0007	-2,675	-0,0016	-2,285	-0,0023	-2,059
Sexo	0,5498	2,538	1,1427	2,072	1,5763	1,773
Estud2	-0,2819	-1,671	-0,3742	-1,133	-0,3361	-0,793
Estud3	-0,7064	-3,039	-1,2071	-2,251	-1,2215	-1,672
Ccaa1	-2,0769	-11,763	-3,7629	-3,401	-4,8971	-2,324
Ccaa2	-0,9393	-5,722	-1,8046	-3,180	-2,2939	-2,302
Tmun1	-3,0231	-17,551	-5,7656	-3,118	-7,3073	-2,224
Tmun2	-1,4391	-7,090	-2,9928	-2,886	-3,7729	-2,132
Casado	0,1679	0,789	0,3865	0,873	0,4561	0,826
LnR: Media	0,5365	2,528	0,7196	1,412	0,4190	1,033
D. típica		0,3070	2,641	0,4206	2,018	
Propiedad Colectiva (PC)						
Cte	-17,566	-6,842	-29,223	-3,345	-23,845	-2,319
Nmiemb	-0,0370	-0,688	-0,0307	-0,336	-0,0571	-0,490
Edad: Media	0,0779	3,180	0,1980	2,5671	0,3130	2,244
D. típica			0,0596	4,07	0,0820	3,744
Edad2	-0,0005	-2,375	-0,0010	-1,965	-0,0018	-2,142
Sexo	0,2076	1,203	0,5474	1,490	0,7779	1,470
Estud2	-0,1564	-1,082	-0,1327	-0,530	0,1176	0,387
Estud3	-0,5423	-2,821	-0,8613	-2,161	-0,4297	-1,046
Ccaa1	-0,8337	-5,264	-1,6649	-3,350	-2,2119	-2,391
Ccaa2	-0,6260	-4,181	-1,2509	-3,070	-1,5753	-2,315
Tmun1	-0,3496	-2,195	-1,2717	-2,754	-1,5841	-2,171
Tmun2	-0,0594	-0,303	-0,6205	-1,644	-0,7351	-1,447
Casado	0,6818	3,859	1,1810	2,930	1,4208	2,354
LnR	1,1987	6,479	1,8833	3,394	1,3689	2,331

Cuadro 2: ESTIMACIÓN DE LOS MODELOS LOGIT MULTINOMIAL Y MIXTO (continuación)

Logit multinomial (Y <sub>p</sub> )		Logit mixto (Y <sub>p</sub> )		Logit mixto (Y <sub>c</sub> )		
Variables	Coefic.	Estad. t	Coefic.	Estad. t	Coefic.	Estad. t
Alquiler Unifamiliar (AU)						
Cte	6,3998	0,965	3,8899	0,303	-2,8140	-0,219
Nmiemb	0,3248	2,610	0,5736	2,585	0,6837	1,720
Edad: Media	-0,0231	-0,383	0,1330	0,751	0,0941	0,405
D. típica			0,1106	2,787	0,1487	2,158
Edad2	0,0002	0,466	-0,0016	-1,004	-0,0015	-0,798
Sexo	-0,2056	-0,438	0,1416	0,195	0,2003	0,210
Estud2	-0,8317	-1,929	-0,9121	-1,518	-1,3385	-1,484
Estud3	0,0705	0,136	-0,2113	-0,254	-0,5681	-0,583
Ccaa1	-0,5703	-1,566	-1,8481	-1,770	-2,2440	-1,705
Ccaa2	-0,3549	-0,978	-1,0426	-1,422	-1,4738	-1,474
Tmun1	-2,1911	-6,154	-4,9871	-2,823	-6,0097	-2,211
Tmun2	-0,8183	-2,184	-2,2365	-2,219	-2,5674	-1,819
Casado	-0,0218	-0,045	-0,0776	-0,109	-0,1516	-0,160
LnR	-0,4892	-1,009	-0,4340	-0,510	0,1074	0,137
Precio	-0,3669	-2,838	-0,5346	-2,233	-0,8132	-1,876
Log-verosim:	-3366,872		Log-verosim:	-3354,886	Log-verosim:	-3362,138
Log-verosim			Log-verosim		Log-verosim	
(sólo ctes):	-4236,431		(sólo ctes):	-4236,431	(sólo ctes):	-4236,431

Al mismo tiempo, la matriz de Choleski estimada (de la que se deriva la matriz de varianzas-covarianzas del cuadro 6) denota la existencia de correlación entre los parámetros aleatorios considerados, al resultar ser significativos a un 95% de confianza la mayor parte de sus elementos, como se confirma al realizar el test de razón de verosimilitud entre el mismo modelo logit mixto con y sin correlación entre los parámetros aleatorios incluidos.

Si se analiza el valor del Log-verosimilitud de los distintos modelos estimados, se observa que el modelo logit mixto que emplea la renta permanente (Y<sub>p</sub>) como medida de renta, es el que presenta un mayor valor, lo que muestra su mayor capacidad explicativa. Por otra parte, el test de razón de verosimilitud entre el logit multinomial y el logit mixto con Y<sub>p</sub> (al contemplar ambos las mismas variables) justifica el uso del modelo logit mixto frente al multinomial, al resultar ser rechazada la hipótesis que establece como nulas las varianzas de los parámetros aleatorios.

Cuadro 3: EFECTOS MARGINALES DEL LOGIT MULTINOMIAL ( $Y_p$ )

	Propiedad unifamiliar	Propiedad colectiva	Alquiler unifamiliar	Alquiler colectivo
LnR ( $Y_p$ )	-0,06417	0,16706	-0,01637	-0,08653
Nmiemb	0,02541	-0,02832	0,00354	-0,00062
Edad	0,00717	0,00113	-0,00135	-0,00695
Edad2	-0,00004	-0,00002	0,00000	0,00005
Sexo	0,05249	-0,02386	-0,00657	-0,02205
Estud2	-0,01652	0,00872	-0,00813	0,01592
Estud3	-0,03418	-0,02027	0,00775	0,04670
Ccaa1	-0,18590	0,08931	0,00815	0,08845
Ccaa2	-0,05441	-0,00612	0,00404	0,05651
Tmun1	-0,36529	0,30212	-0,01076	0,07394
Tmun2	-0,18752	0,16265	-0,00297	0,02786
Casado	-0,05666	0,10949	-0,00541	-0,04742
Precio Compra Unif.	-0,05060	0,04302	0,00178	0,00581
Precio Compra Colect.	0,04302	-0,06945	0,00233	0,02410
Precio Alquiler Unif.	0,00178	0,00233	-0,00462	0,00052
Precio Alquiler Colect.	0,00581	0,02410	0,00052	-0,03043

Nota: Los efectos marginales han sido calculados como promedio sobre el total de la muestra.

Cuadro 4: EFECTOS MARGINALES DEL LOGIT MIXTO ( $Y_p$ )

	Propiedad unifamiliar	Propiedad colectiva	Alquiler unifamiliar	Alquiler colectivo
LnR ( $Y_p$ )	-0,99606	2,35418	-0,25187	-1,10625
Nmiemb	0,08292	-0,0949	0,01313	-0,00115
Edad	0,51559	-0,1578	0,00084	-0,35865
Edad2	-0,18188	0,05464	-0,00726	0,13449
Sexo	0,04744	-0,02121	-0,00542	-0,02082
Estud2	-0,00463	0,00362	-0,00137	0,00238
Estud3	-0,00435	-0,00552	0,00134	0,00852
Ccaa1	-0,04629	0,00861	0,00239	0,03529
Ccaa2	-0,01888	-0,00276	0,00164	0,02000
Tmun1	-0,12256	0,08273	-0,00589	0,04573
Tmun2	-0,04895	0,04225	-0,00083	0,00751
Casado	-0,04584	0,08511	-0,00519	-0,03407
Precio Compra Unif.	-0,59428	0,52076	0,03846	0,03507
Precio Compra Colect.	0,52201	-0,82673	0,02731	0,27741
Precio Alquiler Unif.	0,03634	0,02595	-0,06412	0,00183
Precio Alquiler Colect.	0,03377	0,26902	0,00187	-0,30466

Nota: Los efectos marginales han sido calculados como promedio sobre el total de la muestra.

Cuadro 5: EFECTOS MARGINALES DEL LOGIT MIXTO ( $Y_C$ )

	Propiedad unifamiliar	Propiedad colectiva	Alquiler unifamiliar	Alquiler colectivo
LnR ( $Y_p$ )	-0,64585	1,37777	-0,10207	-0,62987
Nmiemb	0,08535	-0,09941	0,01271	0,00134
Edad	0,46913	0,00294	-0,03043	-0,44164
Edad2	-0,15010	-0,03761	0,00925	0,17846
Sexo	0,04924	-0,02071	-0,00534	-0,02320
Estud2	-0,00677	0,00985	-0,00197	-0,00113
Estud3	-0,00625	0,00252	0,00018	0,00357
Ccaa1	-0,04595	0,00692	0,00227	0,03677
Ccaa2	-0,01846	-0,00244	0,00097	0,01994
Tmun1	-0,12376	0,08513	-0,00566	0,04430
Tmun2	-0,04913	0,04303	-0,00093	0,00704
Casado	-0,04435	0,08287	-0,00548	-0,03305
Precio Compra Unif.	-0,70226	0,62634	0,03717	0,03875
Precio Compra Colect.	0,62787	-1,00498	0,03786	0,33925
Precio Alquiler Unif.	0,03515	0,03597	-0,07313	0,00201
Precio Alquiler Colect.	0,03733	0,32892	0,00205	-0,36830

Nota: Los efectos marginales han sido calculados como promedio sobre el total de la muestra.

Cuadro 6: MATRIZ DE VARIANZAS-COVARIANZAS PARA LOS PARÁMETROS ALEATORIOS DEL MODELO LOGIT MULTINOMIAL MIXTO ( $Y_p$ ) (ERRORES ESTÁNDAR ENTRE PARÉNTESIS)

	LnR en PU	Coeficientes de:	
		Edad en PC	Edad en AU
LnR en PU	0,0943 (0,0714)		
Edad en PC	0,0164 (0,0118)	0,0035 (0,0017)	
Edad en AU	0,0330 (0,0236)	0,0056 (0,0037)	0,0123 (0,0089)

Las principales pautas de comportamiento que en los valores promedio de la muestra se desprenden del modelo mixto estimado con la variable renta permanente ( $Y_p$ ), son los siguientes:

La renta permanente es el factor que en este modelo resulta ser significativamente más determinante en la elección de tenencia (principalmente sobre la pro-

propiedad), presentando incluso un efecto marginal notablemente superior al mostrado en el modelo con renta corriente. Respecto a su influencia sobre las distintas probabilidades, hay que decir que su signo y peso es el que cabría esperar en un principio (con la salvedad del exhibido en la propiedad unifamiliar), de modo que aumentos en la renta permanente suponen aumentos en la probabilidad de la propiedad colectiva y disminución en las probabilidades de alquiler. El impacto negativo que la variable renta permanente ejerce sobre la propiedad unifamiliar se debe con toda seguridad a la alta correlación positiva existente en la muestra de estudio entre las viviendas unifamiliares y las zonas rurales (más concretamente, el 58,46 % de las viviendas unifamiliares se encuentran ubicadas en zonas rurales, en donde los precios de las viviendas resultan ser inferiores). Por tanto, para tener una idea más ajustada del impacto de la renta sobre la decisión de tenencia se hace necesario estudiar de forma simultánea el régimen de tenencia y la localización de la vivienda en el entorno urbano.

La edad del sustentador principal es una de las variables, conjuntamente con el precio y la renta permanente, que posee una gran influencia sobre las probabilidades de las distintas alternativas, siendo ésta positiva sobre las viviendas unifamiliares en régimen de propiedad o alquiler, de modo que los hogares donde el sustentador presenta una mayor edad son los que más optan por viviendas unifamiliares (principalmente en propiedad). El nivel de estudios no constituye una característica relevante en la conducta de los hogares a la hora de decidir el tipo de tenencia, si bien un mayor nivel de estudios (secundarios o universitarios) genera un impacto negativo sobre la probabilidad de la propiedad unifamiliar, lo cual puede seguramente ser matizado si se considera de forma conjunta la variable estudios y renta, pudiéndose deducir en promedios que un incremento de ambas simultáneamente genera una mayor probabilidad de poseer la vivienda en propiedad.

En las comunidades con mayores precios en las viviendas, los individuos se decantan más por el alquiler colectivo y en menor medida por la propiedad colectiva, desincentivándose en este caso la propiedad unifamiliar (como era de esperar). También se observa de forma natural que aquellos individuos que han elegido un municipio de mayor tamaño para fijar su residencia optan en mayor medida por las viviendas colectivas (en alquiler y en propiedad).

En cuanto al sexo del sustentador principal, el hecho de ser mujer parece aumentar la probabilidad de optar por el régimen de alquiler y la propiedad colectiva. Por otra parte, el estado civil no resulta ser una variable muy relevante en la toma de decisiones. Cabe destacar la mayor preferencia que manifiesta el grupo de casados respecto a la propiedad colectiva. Al mismo tiempo, un mayor número de miembros del hogar da lugar a que los individuos se decanten más por las viviendas unifamiliares (en propiedad y alquiler).

Finalmente, la variable precio además de resultar ser significativa en todos los modelos estimados, en el modelo logit mixto pasa a convertirse en el elemento más relevante (junto con la renta permanente y la edad) para explicar las decisiones sobre elección de tenencia (a diferencia del escaso efecto marginal que desempeña en el modelo multinomial). El comportamiento del signo de su impacto se ajusta a lo esperado, esto es, subidas en el precio hacen disminuir la probabilidad de elección de la alternativa que se encarece, provocando un aumento en el



resto de las probabilidades. Se añade así evidencia empírica, al igual que ha sucedido en otras situaciones paralelas como la valoración del ahorro en el tiempo de viaje en modelos de transporte [Hensher (2001a), (2001b)], que apunta a que el modelo logit multinomial, al proponer un marco más restrictivo puede subestimar significativamente los efectos marginales en nuestro caso de las variables precio y renta, en comparación con el modelo logit multinomial mixto.

#### 4. CONCLUSIONES

En este trabajo se ha pretendido analizar las variables fundamentales que intervienen en el proceso de elección entre distintas alternativas de tenencia de la vivienda habitual, establecidas en función del régimen de tenencia y del tipo de edificio. Para ello, como modelo de probabilidad múltiple se ha especificado un modelo logit multinomial mixto por tratarse de un modelo más general, que permite modelar situaciones que presentan correlación y/o heteroscedasticidad, posibilitando que los coeficientes de las variables observadas, en nuestro caso la renta y la edad de los individuos, puedan variar aleatoriamente sobre los individuos en vez de ser fijos como en el caso multinomial.

Para la especificación del modelo mixto, se ha estimado la variable renta permanente y un índice de precios hedónicos por comunidades autónomas para cada una de las alternativas, dado que estas variables no figuran en la ECPF y a nuestro juicio se revelan como fundamentales en las decisiones de tenencia de vivienda.

Los resultados de la estimación justifican el uso de un modelo mixto, obteniéndose varianzas significativas para los distintos parámetros aleatorios barajados, lo que pone de manifiesto que el peso de la variable renta y edad en el modelo de elección de tenencia no es percibido de la misma manera por todos los individuos.

Por otra parte, se evidencia que precisamente las variables renta permanente y precio resultan ser las más relevantes a la hora de explicar la conducta de los hogares en la toma de decisiones sobre tenencia de vivienda. Por tanto, a pesar de la ausencia de información muestral sobre estas variables, se hace necesaria su inclusión, resultando una alternativa válida para ello el modelo de capital humano y el método de los precios hedónicos utilizados en este trabajo.

La renta permanente además de resultar ser significativamente más determinante en la elección de tenencia que la renta corriente (como muestran los efectos marginales), otorga una mayor capacidad explicativa al modelo mixto, al presentar éste un mayor valor en Log-verosimilitud.

En cuanto al resto de las variables analizadas, hay que destacar el importante papel que la edad del sustentador principal desempeña en las probabilidades de las distintas alternativas, al configurarse (conjuntamente con la variable precio y renta permanente) como una de las variables más influyente en la elección del tipo de tenencia.

#### APÉNDICE 1

Para la obtención de la renta permanente ( $Y_p$ ) se han considerado 5.259 hogares de la muestra entre propietarios e inquilinos para los cuales no existen ob-

servaciones perdidas en las variables incluidas, estimándose la renta permanente (en logaritmo) mediante la ecuación de regresión obtenida con el logaritmo de la renta corriente ( $Y_C$ ) como variable dependiente y diferentes características socioeconómicas del hogar como variables explicativas. Como medida de la renta corriente se toma la reportada por la ECPF y que se corresponde con la renta disponible del hogar o neta de impuestos en términos nominales.

En el cuadro 7 recogemos los resultados obtenidos junto con los t-ratios robustos a heteroscedasticidad (contraste de White). Cabe destacar el alto valor del  $R^2$  ajustado y el correcto signo y significatividad de las distintas variables incluidas.

## APÉNDICE 2

La variable precio recoge un indicador de los precios que poseen las distintas alternativas para cada individuo (precio de compra o de alquiler de una vivienda unifamiliar o colectiva).

Dada la ausencia de información muestral sobre la variable precio, hemos optado por estimar índices de precios de vivienda para cada comunidad autónoma, como aproximación a la valoración subjetiva que realizan los individuos sobre cada régimen de tenencia de la vivienda. Para ello estimamos en primer lugar un precio compra y alquiler de las viviendas en cada comunidad autónoma mediante dos regresiones lineales una vez separada la muestra de hogares entre propietarios e inquilinos.

Para llevar a cabo la estimación de la regresión se han considerado como variable dependiente, el alquiler anual imputado por el propietario a su vivienda (para la submuestra de propietarios) y el alquiler anual desembolsado por el inquilino (para la submuestra de inquilinos). Ambas variables son proporcionadas por la ECPF y se incluyen también en términos nominales.

Como variables independientes se han utilizado aquellas que recogen las características de la vivienda, del edificio donde está ubicada, al igual que características del entorno como son la densidad de población y el grado de urbanización de la zona.

En función a dichas variables independientes se han realizado dos regresiones separadas, cada una con la muestra de hogares correspondientes.

$$\begin{aligned} P_{pi} &= x_{pi} \beta_p + \mu_{pi} & i &= 1.2 \dots n_p \text{ (propietarios)} \\ P_{aj} &= x_{aj} \beta_a + \mu_{aj} & j &= 1.2 \dots n_a \text{ (inquilinos)} \end{aligned}$$

donde  $P_{pi}$  y  $P_{aj}$  son respectivamente, el alquiler anual imputado por el propietario (en logaritmo) y el alquiler anual real (en logaritmo) que paga el hogar  $j$  por su vivienda;  $x_{pi}$  y  $x_{aj}$  son los vectores de características de la vivienda y de su entorno del hogar  $i$  y  $j$ ;  $\beta_p$  y  $\beta_a$  son los vectores de parámetros desconocidos;  $\mu_{pi}$  y  $\mu_{aj}$  son las perturbaciones aleatorias.

Los precios de compra y alquiler se han considerado en logaritmos al comprobarse, dadas las características peculiares de este bien, que en los análisis sobre precios de vivienda dicha transformación resulta ser la más usual dadas las ventajas que presenta sobre la forma lineal [Malpezzi (2003)].

Cuadro 7: RENTA PERMANENTE (LNREP)

Variables	Coefficientes	Estad. t	Media
Cte	13,002	91,940	
Edad	0,022	7,594	53,819
Edad2	-0,0001	-7,083	3.107,530
Estud2	0,116	8,400	0,301
Estud3	0,312	18,448	0,175
Ahorro	0,126	10,882	0,372
Contrati	0,181	12,344	0,780
Mutpub	0,101	5,422	0,068
Mutpriv	0,109	5,620	0,077
Cualific	0,146	9,303	0,819
Nmieocup	0,260	33,332	1,116
Ftheadcap	0,112	4,828	0,502
Vivsecun	0,130	8,296	0,143
Fteing1	0,164	1,401	0,607
Fteing2	0,117	1,002	0,387
Fteing3	0,502	3,117	0,002
Finalmes	0,226	19,386	0,461
Nº observaciones: 5.259		R <sup>2</sup> Ajustado: 0,561	
F: 421,65 (Nivel signif.: 0,000)		Durbin-Watson: 1,779	

Nota: las variables que figuran en la regresión son las siguientes:

Lnrep: Renta permanente en logaritmo neperiano. ( $\text{Ln}Y_C = \text{Ln}Y_P + \text{Ln}Y_T$ ).

Cte: Constante o intercepto.

Edad, Edad2: Edad y edad al cuadrado del sustentador principal.

Nmieocup: Número de miembros ocupados en el último trimestre.

Las siguientes variables están codificadas con valor: 0 = No dispone/No posee, 1 = Sí dispone/Sí posee.

Estud2: Sustentador principal con estudios secundarios.

Estud3: Estudios superiores o universitarios.

Ahorro: Capacidad de ahorrar al final de mes.

Contrati: Contrato indefinido.

Cualific: Indica si posee cualificación según la Clasificación Nacional de Ocupaciones (CNO94).

El valor 0 se corresponde con el Grupo 9: Trabajadores no cualificados.

Mutpub: Cobertura sanitaria bajo una mutualidad pública.

Mutpriv: Cobertura sanitaria bajo una mutualidades privada.

Vivsecun: Poseer vivienda(s) secundaria(s).

Ftheadcap: Si además de la fuente principal de ingresos, recibe rentas de la propiedad y capital.

Fteing1: Principal fuente de ingresos derivada de trabajo por cuenta propia o ajena.

Fteing2: Principal fuente de ingresos derivada de pensiones, subsidios y prestaciones.

Fteing3: Principal fuente de ingresos derivada de rentas de la propiedad y capital.

Finalmes: Facilidad para llegar a final de mes.

Una vez estimadas sendas regresiones hedónicas para cada comunidad autónoma, definimos en la misma línea que los trabajos desarrollados por Goodman y Kawai (1982), Ermisch (1996), Rapaport (1997), Rouwendal y Meijer (2001) o Goodman (2002), una vivienda estándar en función de los valores medios que las distintas características presentan para el total de la muestra, y calculamos para dicha vivienda estándar, los índices de precios para cada una de las alternativas en las distintas regiones, a través de los parámetros estimados en las regresiones hedónicas.

Cuadro 8: ÍNDICES DE PRECIOS HEDÓNICOS

C. Autónoma	Compra		Alquiler	
	Unifamiliar	Colectiva	Unifamiliar	Colectivo
Andalucía	13,050	13,095	12,552	12,552
Aragón / Rioja / Navarra	13,179	13,179	12,788	12,788
Asturias	13,488	13,107	11,264	12,092
Baleares	13,118	13,256	12,850	12,850
Canarias	13,420	13,420	11,586	12,355
Cantabria	13,240	13,355	13,259	13,259
Castilla y León	12,820	13,097	12,401	12,401
C. La Mancha / Extremadura	12,896	13,058	11,493	11,968
Cataluña	13,495	13,369	13,181	13,181
C. Valenciana / Murcia	12,973	12,973	12,380	12,380
Galicia	13,003	13,155	13,041	13,041
Madrid	13,652	13,573	12,839	13,382
País Vasco	13,100	13,219	13,647	13,647

Para el cálculo de los índices de precios, hemos procedido a la agrupación de determinadas comunidades autónomas atendiendo a un criterio de cercanía territorial, debido al tamaño inadecuado de la muestra en dichas comunidades. Adviértase que en algunas comunidades coinciden los índices para viviendas unifamiliares y colectivas en ciertas formas de tenencia debido a que no resultaban significativas las variables relativas al tipo de edificio.

Según se desprende de los índices de precios estimados, el precio de las viviendas unifamiliares en determinadas comunidades (en contra de lo esperado) resulta ser inferior al precio de las viviendas colectivas. La explicación a este hecho se debe con seguridad a que la mayor parte de las viviendas unifamiliares (58,46% en nuestra muestra de estudio) se encuentran ubicadas en zonas rurales, en donde los precios de las viviendas resultan ser inferiores. Por tanto, los resultados obtenidos ponen de manifiesto la necesidad de estudiar de forma simultánea régimen de tenencia y ubicación de la vivienda.

Cuadro 9: CARACTERÍSTICAS DE LA VIVIENDA ESTÁNDAR

	Estadísticos descriptivos	
	Media	Desviación típica
Tedif = 0/1 (colectivo/unifamiliar)	0,34	0,48
Antiguo = 30 años	30	34,19
Antiguo2 = 900	2.073,85	43.361,2
Nhab = 5	5,13	1,32
Nhab2 = 25	28,11	17,72
m2 = 95	95,96	39,35
Durbalt = 0	0,083	0,28
Durbamed =1	0,67	0,47
Drural = 0	0,24	0,43
Caprov = 0	0,41	0,49
Dens1 = 1	0,51	0,50
Dens2 = 0	0,18	0,38
Dens3 = 0	0,32	0,47
Calefac = 0	0,43	0,50

Nota: las variables consideradas en las regresiones de precios hedónicos son:  
 Tedif: Tipo de edificio. Valor: 1 = Edificio de una sola vivienda (unifamiliar); 0 = Edificio de dos o más viviendas (colectiva).

Antiguo: Número de años de la vivienda.

Antiguo2: Número de años de la vivienda al cuadrado.

Nhab: Número total de habitaciones incluidos trasteros, sótanos y desvanes.

Nhab2: Variable "Nhab" al cuadrado.

M2: Metros cuadrados totales útiles de la vivienda.

Calefac: Calefacción en la vivienda (0 = No dispone; 1 = Sí dispone).

Dens: Variable que recoge la densidad de población de la zona y que viene expresada con tres variables *dummies*:

Dens1: Zona densamente poblada, es aquel conjunto de municipios contiguos que poseen todos una densidad de más de 500 habitantes por kilómetro cuadrado y cuya población conjunta es superior a 50.000 habitantes.

Dens2: Zona intermedia, es aquel conjunto de municipios contiguos, que no perteneciendo a una zona densamente poblada, cada uno de ellos tienen una densidad de más de 100 habitantes por kilómetro cuadrado y, o bien, la densidad del conjunto es de más de 50.000 habitantes, o bien, están situados al lado de una zona densamente poblada, independientemente del número de habitantes del conjunto.

Dens3: Zona diseminada (los que no constituyen dens1 ni dens2).

Zona de residencia: Recoge el tipo de urbanización en el que se encuentra ubicada la vivienda apreciada por el entrevistador de la ECPF. Se incluye a través de las *dummies*:

Durbalt: Zona catalogada como urbana alta. Pertenecen a ella las viviendas localizadas en municipios grandes (generalmente con más de 10.000 habitantes), dentro de barrios residenciales de clase media-alta con urbanización cuidada y viviendas confortables. Con comercio escaso o buen comercio.

Durbamed: Zona urbana media-inferior. Constituida por barrios populares habitados por trabajadores medios con viviendas en general de construcción antigua con comercio popular, o zonas deprimidas, de escasa urbanización con edificaciones baratas habitadas, en su mayoría, por obreros sin cualificar.

Drural: Zona rural. Formada por los hogares ubicados en municipios pequeños (generalmente con menos de 10.000 habitantes) o en zonas sin urbanizar.

Caprov: Variable *dummy* que recoge los municipios que son capital de provincia.

Los resultados de las regresiones hedónicas calculadas para cada comunidad autónoma podrán ser facilitados bajo petición.

### Vivienda estándar

Atendiendo a los valores medios que las distintas características de las viviendas consideradas presentan en la totalidad de la muestra, la vivienda estándar viene definida como una vivienda con treinta años de antigüedad, una superficie de noventa y cinco metros cuadrados útiles, sin calefacción y que dispone de cinco habitaciones, ubicada en una zona urbana media con una densidad de población superior a 500 habitantes por km<sup>2</sup>. Los valores que se corresponden con esta definición al igual que los estadísticos descriptivos son los que figuran en el cuadro 9.

Una vez estimados los vectores de parámetros  $\beta$  se puede obtener el precio de compra de la vivienda estándar si la hubiera comprado y un precio de alquiler si fuera inquilino, tanto para una vivienda unifamiliar como colectiva. Es importante subrayar que los índices de precios calculados no se corresponden con precios de mercado sino que representan precios hedónicos establecidos en función de las características de las viviendas e indicadores de la valoración subjetiva que realizan los individuos de éstas.



### REFERENCIAS BIBLIOGRÁFICAS

- Barrios, J.A. (2001): "El coste de uso del capital residencial en propiedad: revisión teórica y reciente evolución en España", en Calero F. *et al.* (coord.): "*Economía y Finanzas 2001*", Dir. Gral. Universidades e Investigación del Gobierno de Canarias, págs. 93-112.
- Barrios, J.A. y J.E. Rodríguez (2003): "Vivienda de protección oficial o libre: una caracterización de la elección de los individuos en el ámbito canario", *Estudios Económicos Regionales y Sectoriales*, vol. 3 n.º 1, págs. 79-103.
- Börsch-Supan, A., F. Heiss y M. Seko (2001): "Housing demand in Germany and Japan", *Journal of Housing Economics*, vol. 10 n.º 3, págs. 229-252.
- Börsch-Supan, A. y J. Pitkin (1988): "On discrete choice models of housing demand", *Journal of Urban Economics*, n.º 24, págs. 153-172.
- Börsch-Supan, A. y H.O. Pollakowski (1990): "Estimating housing consumption adjustments from panel data", *Journal of Urban Economics*, n.º 27, págs. 131-150.
- Bourassa, S.C. (1995): "A model of housing tenure choice in Australia", *Journal of Urban Economics*, n.º 37, págs. 161-175.
- Boyd, J. y R. Melman (1980): "The effect of fuel economy standards on the U.S. automotive market: an hedonic demand analysis", *Transportation Research*, 14A, n.º 5-6, págs. 367-378.
- Brownstone, D. y K. Train (1999): "Forecasting new product penetration with flexible substitution patterns", *Journal of Econometrics*, n.º 89, págs. 109-129.
- Cardell, N. y F. Dunbar (1980): "Measuring the societal impacts of automobile downsizing", *Transportation Research*, 14A, n.º 5-6, págs. 423-434.
- Colom, M. y M. Cruz (1997): "La demanda de vivienda en España", Documento de Trabajo del Instituto Valenciano de Investigaciones Económicas WP-EC 97-19.
- Duce Tello, R.M. (1995): "Un modelo de elección de tenencia de vivienda para España", *Moneda y Crédito*, n.º 201, págs. 127-152.

- Ermisch, J. (1996): "The demand of housing in Britain and population ageing: microeconomic evidence", *Economica*, n.º 63, págs. 383-404.
- Friedman, M. (1957): "A theory of the consumption function", Princeton University Press, Princeton.
- García Montalvo, J. (2003): "La vivienda en España: desgravaciones, burbujas y otras historias", *Perspectivas del Sistema Financiero*, n.º 78, págs. 1-43.
- Gobillon, L. y D. Le Blanc (2002): "The impact of borrowing constraints on mobility and tenure choice", Mimeo.
- González, L. (1997): "Estimación de la demanda de vivienda: tenencia y gasto en servicios. El mercado metropolitano de México", *El Trimestre Económico*, pág. 569-598.
- Goodman, A.C. (1995): "A dynamic equilibrium model of housing demand and mobility with transaction cost", *Journal of Housing Economics*, n.º 4, págs. 307-327.
- Goodman, A.C. (2002): "Estimating equilibrium housing demand for «stayers»", *Journal of Urban Economics*, n.º 51, págs. 1-24.
- Goodman, A.C. y M. Kawai (1982): "Permanent income, hedonic price, and demand for housing: new evidence", *Journal of Urban Economics*, n.º 12, págs. 214-237.
- Hajivassiliou, V. y P. Ruud (1994): "Classical estimation methods for LDV models using simulation", en Engle, R. y McFadden, D. (eds.), "Handbook of Econometrics Vol. IV", Elsevier, Nueva York, págs. 2383-2441.
- Haurin, D.R., P.H. Hendershott y D. Kin (1994): "Housing decisions of American youth", *Journal of Urban Economics*, n.º 35, págs. 28-45.
- Henderson, J.V. y Y.M. Ioannides (1983): "A model of housing tenure choice", *American Economic Review*, vol. 73 n.º 1, págs. 98-113.
- Hensher, D.A. (2001a): "The sensitivity of the valuation of travel time savings to the specification of unobserved effects", *Transportation Research Part E*, n.º 37, págs. 129-142.
- Hensher, D.A. (2001b): "The valuation of commuter travel time savings for car drivers: evaluating alternative model specifications", *Transportation*, n.º 28, págs. 101-118.
- Ioannides, Y.M. y K. Kan (1996): "Structural estimation of residential mobility and housing tenure choice", *Journal of Regional Science*, vol. 36 n.º 3, págs. 335-363.
- Jaén, M. y Molina, A. (1994): "Un análisis empírico de la tenencia y demanda de vivienda en Andalucía", *Investigaciones Económicas*, Vol. XVIII n.º 1, págs. 143-164.
- King, M.A. (1980): "An econometric model of tenure choice and demand for housing as a joint decision", *Journal of Public Economics*, n.º 14, págs. 137-159.
- Lee, L.F. y R.P. Trost (1978): "Estimation of some limits dependent variable models with application to housing demand", *Journal of Econometrics*, n.º 8, págs. 357-382.
- López García, M.A. (1996): "Precios de la vivienda e incentivos fiscales a la vivienda en propiedad en España", *Revista de Economía Aplicada*, vol. IV n.º 12, págs. 37-74.
- Malpezzi, S. (2003): "Hedonic pricing models: a selective and applied review", en O'Sullivan, A. y Gibb, K. (Eds): "Housing economics and public policy", Blackwell Publishers, Oxford, págs. 67-89.
- McFadden, D. (1974): "Conditional logit analysis of qualitative choice models", en Zarembka, P. (Ed.), "Frontiers in Econometrics", Academic Press, Nueva York, págs. 105-142.
- McFadden, D. (2000): "Disaggregate behavioral travel demand's RUM side: a 30-year retrospective", *Conference at the International Association of Travel Behavior Analyst*, Brisbane, Australia, July 2-7-2000.
- McFadden, D. y K. Train (2000): "Mixed MNL models for discrete response", *Journal of Applied Econometrics*, vol. 15 n.º 5, págs. 447-470.



- Rapaport, C. (1997): "Housing demand and community choice: an empirical analysis", *Journal of Urban Economics*, n.º 42, págs. 243-260.
- Revelt, D. y K. Train (1998): "Mixed logit with repeated choices: household's choices of appliance efficiency level", *Review of Economics and Statistics*, n.º 80, págs. 647-657.
- Rodríguez, J.E. y J.A. Barrios (2003): "Un modelo logit mixto de tenencia de vivienda en Canarias", *Estudios de Economía Aplicada*, vol. 21, n.º 1, págs. 175-193.
- Rosen, H.S. (1979): "Housing decisions and the U.S. income tax: an econometric analysis", *Journal of Public Economics*, n.º 11, págs. 1-23.
- Rouwendal, J. y E. Meijer (2001): "Preferences for housing, jobs, and commuting: a mixed logit analysis", *Journal of Regional Science*, vol. 41, n.º 3, págs. 475-505.
- Skaburskis, A. (1999): "Modelling the choice of tenure and building type", *Urban Studies*, vol. 36, n.º 13, págs. 2199-2215.
- Taltavull, P. (Coord.) (2000): "Vivienda y familia", Colección Economía Española, Vol. XIII, Fundación Argentaria, Madrid.
- Thibodeau, T.G. (1995): "House price indices from the 1984-1992 MSA American housing surveys", *Journal of Housing Research*, n.º 6, págs. 439-481.
- Train, K. (2003): "*Discrete choice methods with simulation*", Cambridge University Press, Cambridge.
- Walker, B., A. Marsh, M. Wardman y P. Niner (2002): "Modelling tenants' choices in the public rented sector: a stated preference approach", *Urban Studies*, vol. 39, n.º 4, págs. 665-688.
- Williams, H. (1977): "On the formation of travel demand models and economic evaluation measures of user benefit", *Environment and Planning*, 9A, págs. 285-344.

*Fecha de recepción del original: mayo, 2003*

*Versión final: febrero, 2004*

#### ABSTRACT

In this paper, we use cross sectional data from the Spanish Survey of Household Budgets (ECPF base 1997) to develop a multinomial mixed logit model that seeks to explain the fundamental variables influencing the individual decisions on housing tenure in Spain during 1999. We consider four types depending on the holding regime (property or rent) and the building type (single family or multi family). With this aim, we first elaborate a measure of the permanent income imputed to individuals, as well as a hedonic price index for regions in Spain. After that, we proceed to estimate a mixed logit model for the housing tenure decision that will allow us to contemplate a certain non-observed heterogeneity in the behavior of the individuals, as well as to overcome the hypothesis of the independence of irrelevant alternatives and homoscedasticity, both characteristic of multinomial and nested logit models.

*Key words:* housing market, housing tenure choice, mixed logit.

*JEL classification:* C25, R21.