

## CONSTRUCCIÓN DE UN ÍNDICE DE CALIDAD DE LA VIVIENDA

RÁQUEL ARÉVALO-TOMÉ

*Universidad Complutense de Madrid*

*Universidad de Vigo*

*En este trabajo se construye un Índice de Calidad de la Vivienda que sintetiza las características de la vivienda principal ocupada en España. Se comprueba que en la década de 1980 a 1990 la calidad de la vivienda ha mejorado en todas las Comunidades Autónomas. (JEL C43, D12, D39)*

### 1. Introducción

La intervención por parte del Estado en el mercado de la vivienda en propiedad y/o alquiler es una práctica habitual en Europa desde la Primera Guerra Mundial y en Estados Unidos desde la Segunda. Existen numerosos trabajos que tratan de estimar su impacto en el precio a través de modelos de precios hedónicos, utilizando datos de corte transversal. Ejemplos concretos pueden ser Börsch-Supan (1994) para el caso de Alemania o Peña y Ruiz-Castillo (1984) y Arévalo (1998) para España. Una recopilación crítica de un gran número de este tipo de trabajos realizados en Estados Unidos se recoge en Olsen (1990). En todos ellos se reconoce la importancia de tratar la vivienda como un bien heterogéneo, la hipótesis que subyace es que la calidad de las viviendas está relacionada con la divergencia de precios observada.

En España, los datos disponibles más detallados sobre las características de la vivienda los proporcionan las Encuestas de Presupuestos Familiares (EPF) realizadas por el Instituto Nacional de Estadística (INE) en 1980-81 y 1990-91. Esta información será utilizada para construir una medida sintética que refleje de la manera más fiel posible las particularidades relevantes que pueden caracterizar a una vivienda.

Deseo agradecer los comentarios y sugerencias de Gragorio Serrano y Javier Ruiz-Castillo al realizar este trabajo. No obstante, la autora es la única responsable de las deficiencias que en él puedan encontrarse.

La técnica multivariante utilizada es el Análisis de Correspondencias Múltiples (ACM), que hace posible adjudicar a cada vivienda una puntuación, combinación lineal de sus características físicas, que puede identificarse con un índice de calidad de la vivienda. Adicionalmente, se proporciona una medida de la pérdida de información que esta agregación trae consigo. Este índice procede de la consideración de las dos muestras conjuntas y se trata, por tanto, de una medida homogénea para las viviendas de ambas encuestas. Se aplica para analizar en qué sentido ha variado la calidad de las viviendas principales ocupadas de 1980-81 a 1990-91. Se pretende responder a preguntas como: ¿puede decirse que la vivienda en 1990-91 tiene mejor calidad que la de diez años antes?, ¿cuál es la ordenación de Comunidades Autónomas respecto a la calidad de sus viviendas en ambos períodos?

La exposición del trabajo se divide en las siguientes partes. En el segundo apartado se expone de forma resumida la metodología del ACM. El tercero contiene la aplicación de esta técnica multivariante al sector de la vivienda, así como los resultados que permiten construir e interpretar el índice. En el cuarto apartado se compara la calidad de las viviendas desglosada por Comunidades Autónomas. Se reserva el último apartado para comentar los resultados más relevantes.

## 2. Metodología del análisis de correspondencias múltiples

El ACM se aplica a  $Q$  variables de tipo cualitativo, y posiblemente correladas, referidas a un conjunto de sujetos con el objeto de obtener  $K$  variables incorreladas ( $K < Q$ ) que, siendo combinación lineal de las anteriores, pueden ayudar a interpretar los datos observados. El ACM es la generalización más simple del llamado Análisis de Correspondencias Simples, para el que se utiliza una tabla de contingencia de dos variables (Lebart *et al.*, 1984 y Greenacre, 1984). Ambos métodos permiten, además, analizar posibles asociaciones entre variables. Existe una gran variedad de técnicas multivariantes que tienen objetivos similares en el sentido de reducir el número de variables que proporcionan información de un determinado conjunto de observaciones. En concreto, el Análisis de Componentes Principales, de aplicación a variables continuas, presenta gran similitud con el ACM (Tenenhaus y Young, 1985).

### 2.1 Notación

Se dispone de información sobre  $Q$  variables cualitativas referentes a un total de  $N$  sujetos objeto de estudio. Sea  $Z$  la matriz de datos ( $N \times J$ ) donde  $J$  hace referencia al número total de clases o modalidades de las  $Q$  variables. Se denota por  $J_q$  al número de modalidades de cada variable  $q$  ( $q = 1, 2, \dots, Q$ ). El término general  $z_{ij}$  de la matriz  $Z$  toma valor uno si el sujeto  $i$  tiene la característica  $j$  y cero en caso contrario,  $i = 1, 2, \dots, N$ ,  $j = 1, 2, \dots, J$ , siendo  $N \gg J$ . A cada sujeto le corresponde una, y sólo una, de las  $J_q$  modalidades de cada variable, de forma que  $Q = \sum_{j=1}^J z_{ij}$ . La suma de los elementos de la columna  $j$  de  $Z$  es la frecuencia absoluta de la  $j$ -ésima modalidad denotada por  $N_j$  y por construcción  $\text{rango}(Z) \leq J - Q$ . Para cada  $q$ ,  $N = \sum_{j \in I_q} \sum_{i=1}^N z_{ij}$ , donde  $I_q$  es el conjunto de modalidades de la variable  $q$ . Por tanto,  $NQ = \sum_{j=1}^J \sum_{i=1}^N z_{ij}$ .

La matriz de frecuencias relativas es  $F = (1/NQ)Z$ , y a partir de ella se define el *perfil columna*  $j$  como la  $j$ -ésima columna de  $F$  expresada en relación al total de sus elementos; esto es, los perfiles columna son las frecuencias relativas condicionadas de las columnas. De forma análoga quedaría definido el *perfil fila*  $i$ . Los respectivos perfiles medios son  $r = F^T \mathbf{1}_N$  para las columnas donde  $r_i = 1/N$ , y  $c = F \mathbf{1}_J$  para las filas, donde cada elemento es  $c_j = N_j/NQ$ . Las matrices diagonales que contienen estos perfiles medios se denotan por  $D_r$  y  $D_c$  respectivamente.

En este contexto, las similitudes entre dos perfiles se miden a través de su distancia chi-cuadrado, siendo ésta una distancia euclídea ponderada con la métrica definida por  $D_r^{-1}$  para los perfiles columna y  $D_c^{-1}$  para los perfiles fila. Se define la matriz:

$$E = D_r^{-1/2}(F - rc^T)D_c^{-1/2} \quad [1]$$

de forma que el elemento  $e_j^T e_j$  de la diagonal principal de la matriz  $E^T E$  es la distancia chi-cuadrado entre el perfil columna  $j$  y el perfil medio ( $r$ ), ponderada por su frecuencia relativa ( $c_j$ ). De forma análoga puede interpretarse el elemento  $e_j e_j^T$  de  $EE^T$  para las filas. La suma de estos elementos, la traza de  $E^T E$ , se denomina *inerencia total* ( $IT$ ), concepto usado en la literatura de análisis de correspondencias, y se

asocia con la noción de varianza ponderada. En ACM su valor está determinado por  $IT = tr(E^T E) = tr(EE^T) = J/Q - 1$ .

## 2.2 Procedimiento del ACM

El ACM realiza la descomposición en valores singulares de la matriz  $E$  de la ecuación [1], es decir,  $E = UD_\alpha V^T$ , siendo  $U$  y  $V$  ortogonales. En la práctica, como  $N \gg J$  se realiza la descomposición en valores propios (DVP) de la matriz simétrica  $E^T E$ ,  $E^T E = \Gamma D_\lambda \Gamma^T$ , donde  $D_\lambda = D_\alpha^2$ . Los valores propios contenidos en la matriz diagonal  $D_\lambda$ , indican la inercia proyectada sobre cada uno de los vectores propios asociados (columnas de  $\Gamma$ ), los cuales representan direcciones ortogonales de proyección de los perfiles columna centrados. La dirección del primer vector propio (el asociado al mayor de los valores propios) es un eje de proyección óptimo desde el punto de vista de representar la orientación lineal que recoge la máxima disparidad entre los sujetos en lo que a las  $Q$  variables se refiere. El segundo vector propio es ortogonal al primero y representa la orientación lineal que captura la máxima disparidad residual, es decir, la no recogida en el primer eje de proyección, y así sucesivamente hasta que la inercia total está contenida en un total de  $K$  ejes ortogonales con  $K \leq J - Q$ , de forma que  $IT = \sum_{k=1}^K \lambda_k$ . En términos geométricos, esto supone un cambio de base del espacio original de los perfiles, a otro de dimensión reducida  $K$ .

Las coordenadas de los  $J$  puntos perfiles columna en el  $k$ -ésimo eje están contenidos en la columna  $k$  de la matriz:

$$M = D_c^{-1/2} \Gamma D_\lambda^{1/2} \quad [2]$$

con elemento genérico  $m_{ij}$ . Las  $K$  variables indicador  $w_k$  ( $k = 1, 2, \dots, K$ ) se definen como combinaciones lineales de las modalidades de las  $Q$  variables observadas, y son las columnas de la matriz:

$$W = ZM \quad [3]$$

El valor  $m_{jk}$  es el peso de la  $j$ -ésima modalidad en la nueva variable  $w_k$  y  $z_{ij}$  toma valor uno (cero) si el  $i$ -ésimo sujeto presenta (no presenta) el atributo  $j$ .

Teniendo en cuenta que a cada eje le corresponde una inercia proyectada  $\lambda_k$ , y que la suma de todos ellos es la inercia total, es posible

hallar el porcentaje de inercia recogido en cada eje o variable indicador, disponiendo así de una medida de su capacidad para resumir la información contenida en  $E$ . En el ACM estos porcentajes son siempre muy pequeños y muestran una idea pesimista de la proporción de inercia proyectada (Greenacre, 1990). Esto se debe a la disposición de la matriz de datos  $Z$  que, a costa de aumentar artificialmente la dimensión del espacio original, permite una generalización del ACS cuando  $Q > 2$ , sin afectar a los resultados óptimos buscados: máxima discriminación de los sujetos respecto a las variables que los definen (véase Greenacre, 1984, pp. 131-132). Para conocer la verdadera representatividad de los ejes, Benzécri (1979) propone considerar sólo los  $P$  ejes relevantes, esto es, los asociados a los valores propios  $\lambda_p > 1/Q$ ,  $p = 1, 2, \dots, P$  y  $P \leq K$ . Igualmente propone corregir los valores propios con la transformación  $\lambda_p^c = [Q/(Q - 1)]^2 [\lambda_p - 1/Q]^2$  y hallar la proporción de inercia explicada en relación a  $\sum_{p=1}^P \lambda_p^c$ . En la medida en que se consideren  $P$  variables indicador en lugar de  $J - Q$  modalidades, se reduce la dimensión de la matriz original con la mínima pérdida de información.

La interpretación del  $p$ -ésimo indicador debe hacerse en función de: a) la importancia del eje  $p$  en relación al porcentaje de inercia corregida proyectada en él, b) la correlación del indicador con las  $Q$  variables originales que intervienen en el análisis, y c) las posiciones relativas de las modalidades proyectadas en cada eje de proyección, es decir, las ponderaciones de las modalidades en la nueva variable.

Dado que la matriz sobre la que se aplica la DVP es centrada, el valor  $w_{ip}$  corresponde a una medida de la discrepancia entre el  $i$ -ésimo sujeto y el sujeto de perfil medio para el que, por construcción,  $\bar{w}_p = 0, \forall p$ . Es en relación a este valor medio de referencia respecto al que se pueden realizar comparaciones, lo que resulta especialmente importante en este análisis puesto que las variables indicador son ordinales.

### 3. Aplicación del ACM a la vivienda española 1980-1990

En la aplicación del ACM realizada en este trabajo, los sujetos que definen la población de estudio son las viviendas principales ocupadas en España que se investigan en las EPF de 1980-81 y 1990-91. Dichas encuestas se realizan por el INE a nivel nacional y proporcionan información detallada sobre las viviendas en ambos períodos. El tamaño muestral es de 23.971 en 1980-81 y 21.155 en 1990-91. Estas muestras

representan, estadísticamente, un universo de 10.024.166 y 11.298.509 hogares respectivamente y sus correspondientes viviendas principales

La selección de variables a incluir en el análisis se realiza en función de su relevancia a priori para determinar la calidad de la vivienda. Con el fin de hacer posible su comparación, se renuncia a información que, referida a 1990-91 no está disponible en la EPF 1980-81. Por el mismo motivo se codifican las variables de forma homogénea en ambos períodos. La definición de las  $Q = 8$  variables disponibles y su división en  $J = 32$  modalidades se muestra en el Cuadro 1.

CUADRO 1  
Definición de variables y su distribución en % para cada período muestral

VARIABLES. modalidades	EPF 80-81	EPF 90-91
ACON. Años de antigüedad del edificio desde su construcción		
antig0. 10 entre o menos	25.8	17.9
antig1. 10 entre 11 y 20	27.4	32.9
antig2. 10 entre 21 y 30	11.4	23.8
antig3. 10 entre 31 y 50	10.9	12.9
antig4. más de 20	24.5	12.5
HIG Servicios higiénicos		
hig0. no tiene o compartido con otras viviendas	7.6	1.9
hig1. un baño completo o menos (aseo(s) y/o retrete(s))	76.8	73.4
hig2. más de un baño completo	15.6	24.7
M2V Número de metros cuadrados de superficie construida		
metros0. menos de 61	23.1	9.4
metros1. entre 61 y 90	48.4	37.7
metros2. entre 91 y 130	21.5	37.2
metros3. más de 130	7.0	15.7
AGUA Agua corriente		
agua0. no tiene	3.8	0.4
agua1. sólo fría	19.5	5.2
agua2. fría y caliente	76.7	94.4
TEL Teléfono		
telef0. no tiene	49.3	23.1
telef1. sí tiene	50.7	76.9
CALF Calefacción		
calef0. no tiene	61.7	11.1
calef1. sí tiene, individual, colectiva y/o aparatos móviles	38.3	88.9
GAR Garaje		
garaj0. no tiene	83.7	72.5
garaj1. sí tiene	16.3	27.5
LUZ Luz eléctrica		
luz0. no tiene	0.9	0.2
luz1. sí tiene	99.1	99.8

La aplicación del ACM se realiza sobre las viviendas que componen la EPF de 1980-81 y 1990-91 consideradas conjuntamente con el fin de disponer de un indicador que permita comparar las viviendas de ambos períodos. En el Cuadro 2 se muestra la descomposición de la inercia total después de haber realizado la corrección de Benzécri (1979), por lo que el porcentaje de inercia recogido en cada eje de proyección muestra la verdadera representatividad del eje.

CUADRO 2  
Descomposición de la inercia proyectada

Factores relevantes	EPF 80-90 $\lambda_p^c$	%
1º	0.0561	87.92
2º	0.0076	11.85
3º	0.0002	0.23
Total	0.0639	100

Como se puede apreciar, la simplificación que puede conseguirse con este cambio de base ortogonal es muy importante, al pasar de 32 modalidades a tan sólo una o dos variables indicador.

Respecto a las modalidades proyectadas sobre el plano formado por los dos primeros ejes, se observó una disposición parabólica de las modalidades de cada variable, conocida como *efecto Guttman*. Este efecto aparece si existe una estructura de orden entre las modalidades con una asociación respecto a las viviendas, e indica que la relación entre las modalidades se puede resumir en la estructura de orden manifestada por el primer eje, el cual se interpreta globalmente. El segundo eje representa una oposición entre las modalidades extremas (de frecuencias pequeñas) y las modalidades medias (de mayor frecuencia), de una misma variable. Dada esta interpretación del segundo eje, el análisis se limita al primero y, por tanto, a la primera variable indicador de la vivienda, ya que contiene la información esencial de la variación entre viviendas.

### 3.1 Interpretación de la variable indicador

El indicador obtenido sintetiza las particularidades relevantes que distinguen a una vivienda respecto a las demás en lo que a las características físicas se refiere. Los pesos obtenidos mediante [2], permiten definir un índice, que se denotará por I8090, a través de la ecuación [3]. Este índice recoge el 86.7% de la variabilidad total de las ocho variables (32 modalidades) que caracterizan las viviendas analizadas.

### 3.2 Correlaciones entre variables originales e índice

Los pesos de las modalidades proporcionan la máxima correlación lineal posible entre las características físicas y el índice y alcanza el valor medio de 0.58. Esta correlación se halla utilizando la contribución absoluta ( $CA$ ) de cada modalidad  $j$  al eje  $p$ , las cuales vienen dadas por la expresión  $CA_{jp} = (N_j/NQ)m_{jp}^2$ ,  $p = 1, 2, \dots, P$  (véase Greenacre, 1984). Matemáticamente, dicha correlación al cuadrado coincide con la media de las correlaciones al cuadrado de las variables, y se puede descomponer en tantas partes como variables se consideran. Cada componente informa del papel jugado por cada variable en la determinación del índice en términos de correlación lineal, por lo que es una información relevante a la hora de interpretar el índice. En el Cuadro 3 se muestran estas correlaciones.

CUADRO 3  
Correlación Lineal entre variables originales e índice

	18090
Años de construcción del edificio	0.53
Servicios higiénicos	0.74
Metros cuadrados construidos	0.49
Agua corriente	0.75
Teléfono	0.63
Calefacción	0.61
Garaje	0.43
Luz eléctrica	0.26
Correlación media	0.58

Las variables HIG (servicios higiénicos) y AGUA son las más influyentes a la hora de determinar el índice, mientras que LUZ es la que muestra menor correlación. A pesar de las pocas viviendas que no disponen de este atributo (0.9% en 1980-81 y 0.2% en 1990-91), se ha decidido su inclusión en el ACM considerando la relevancia de esta información a la hora de discriminar viviendas respecto a los servicios instalados.

### 3.3 Pesos de las modalidades sobre el índice

Se debe analizar el peso relativo de las modalidades, sin olvidar la contribución de cada variable a la construcción del índice dada por su correlación. Los pesos obtenidos son valores estandarizados y la escala de medida depende directamente de la inercia total explicada por el índice.

En el Cuadro 4 se presenta la ordenación de modalidades según su peso tras haber aplicado una normalización que conserva tanto el signo de cada peso de la modalidad, como el valor cero referido al índice medio en cada conjunto de hogares. Se asigna el valor arbitrario -10 a la modalidad de menor peso y el valor 10 al equivalente valor extremo de signo positivo, conservando los pesos relativos de los servicios que componen la vivienda. El valor máximo no aparece, lo cual indica que no hay ninguna modalidad que compense positivamente el valor extremo negativo tomado por la modalidad de peso -10.

CUADRO 4  
Modalidades ordenadas según su peso

luz0	-10.0
agua0	-8.72
hig0	-7.89
agua1	-4.38
antig4	-2.89
metros0	-2.75
teléf0	-2.38
calef0	-2.36
antig3	-0.79
garaj0	-0.66
hig1	-0.24
metros1	-0.19
luz1	0.05
antig2	0.35
agua2	0.81
antig1	0.83
metros2	1.11
calef1	1.26
teléf1	1.31
antig0	1.39
metros3	1.58
garaj1	2.32
hig2	2.67

Teniendo en cuenta cada variable individualmente, la ordenación de sus modalidades presenta un peso negativo para el caso de su peor característica, signo positivo en su mejor clase, y una gradación acorde con su significado en el caso de existir más de dos modalidades. Este resultado es sistemático para todas las variables involucradas. La no posesión luz, agua y servicios higiénicos, resultan ser las modalidades extremas con influencia negativa. Además, las dos últimas son las más correladas con el índice. Por otro lado, la disponibilidad de garaje y servicios higiénicos con más de un baño completo son las características que se relacionan con los valores positivos extremos.

La correlación entre índice y variables, junto con los pesos de las modalidades, permiten interpretar la nueva variable como un *índice de calidad de la vivienda*. Con este indicador se pueden ordenar la totalidad de viviendas en relación a su calidad, ya que representa un valor indicativo de la desviación (positiva o negativa) respecto a una calidad media que toma valor cero. Las modalidades asociadas con la vivienda media son: una antigüedad del edificio entre 20 y 30 años, vivienda con luz, un baño (tal vez no completo), agua caliente, entre 61 y 90 metros cuadrados, sin garaje, con teléfono y calefacción.

### *3.4 Limitaciones del indicador de calidad de la vivienda*

Es importante hacer notar que el índice es interpretable en términos de un concepto de calidad restringido a las ocho variables que intervienen en el ACM, con unos únicos pesos de las características físicas para toda vivienda a comparar. En este punto, deben tenerse en cuenta dos limitaciones del índice hallado: a) La no disponibilidad de variables como. grado de contaminación, seguridad ciudadana, etc., que son relevantes para un concepto más amplio de calidad, limita la información contenida en el índice. b) No se recoge la posible heterogeneidad en el peso de las modalidades sobre la calidad para distintos segmentos de viviendas. En concreto, es de suponer que la calefacción en Canarias debería tener menor peso sobre la calidad que en cualquier zona del norte de España. La consideración apropiada de esta circunstancia podría resolverse hallando un índice por cada agrupación de interés, aunque impediría la comparación entre zonas. Dados los objetivos de este trabajo, se ha optado por calcular un único índice.

## **4. Comparación de la calidad de la vivienda por CCAA**

El análisis de correspondencias conjunto de las EPF 1980-81 y 1990-91 permite la comparación de ambos períodos. Para ello, una vez construido el índice I8090, se ordenan los hogares de menor a mayor calidad y se dividen en cuartiles. De esta forma, utilizando la distribución de frecuencias de las características en cada cuartil de calidad, es posible identificar una vivienda tipo de dicho cuartil. La distribución de frecuencias de las modalidades que caracterizan cada cuartil se presenta en el Cuadro 5. Como se puede observar, en el primero de ellos se encuentra la totalidad de hogares sin servicios higiénicos (4.8% del total de hogares y el 18.3% de los hogares que componen este 25% de viviendas de peor calidad); la totalidad de los que no tienen luz, y

prácticamente el 100% de los que no poseen agua corriente. De forma progresiva, en los siguientes cuartiles, se van registrando mayores porcentajes en las modalidades asociadas a un mayor nivel de calidad. Así, en el cuarto cuartil, el 97% de las viviendas tienen calefacción y/o teléfono, el 100% tiene agua caliente; el 67.5% tiene menos de 20 años de antigüedad; el 76.3% tiene más de 90 metros cuadrados y el 63.1% tiene más de un baño. Resulta revelador comprobar la correspondencia de la medida ordinal "índice de calidad", con las variables que representa y determinan su significado.

CUADRO 5  
Distribución en % de los cuartiles según las características de cada variable

EPF 1980-90	Frecuencia relativa	Cuartil 1	Cuartil 2	Cuartil 3	Cuartil 4	
ACON	antig0	21.6	9.6	19.1	19.6	38.3
	antig1	30.3	16.3	32.1	33.6	39.2
	antig2	18.0	12.6	17.8	27.4	14.0
	antig3	12.0	16.3	10.7	13.6	7.3
	antig4	18.1	45.2	20.3	5.8	1.2
HIG	hig0	4.6	18.3	-	-	-
	hig1	75.0	80.3	95.4	87.3	36.9
	hig2	20.4	1.3	4.5	12.7	63.1
M2V	metros0	15.8	35.7	25.3	2.3	0.1
	metros1	42.7	45.7	40.5	61.5	23.6
	metros2	29.8	14.1	25.4	28.7	51.2
	metros3	11.6	4.8	8.8	7.6	25.1
AGUA	agua0	2.0	8.0	0.1	-	-
	agua1	11.9	45.0	2.4	0.3	-
	agua2	86.1	47.0	97.5	99.7	100
TEL	teléf0	35.4	85.6	41.8	11.3	2.9
	teléf1	64.6	14.4	58.2	88.7	97.1
CALF	calef0	34.9	82.8	42.3	11.5	2.9
	calef1	65.1	17.2	57.7	88.5	97.1
GAR	garaj0	77.8	96.3	92.5	84.8	37.5
	garaj1	22.2	3.7	7.5	15.2	62.5
LUZ	luz0	0.5	2.0	-	-	-
	luz1	99.5	98.0	100	100	100

En el Cuadro 6 se muestra la media de I8090 y su desviación típica, así como la distribución de las viviendas por cuartiles para el total nacional, las 18 CCAA y cada período de análisis. En términos medios, la calidad de la vivienda en 1990-91 ha mejorado de forma significativa respecto a la de 1980-81 para el conjunto total de hogares. La media nacional de la variable I8090 pasa de un valor -0.24 a 0.22. Aunque a distintos niveles, este resultado se obtiene en todas y cada una de las CCAA. Las CCAA que, en media, han experimentado una mejora

CUADRO 6  
Media y distribución en % del Índice I8090 por

	EPF	Media (D T.)	Cuartil 1	Cuartil 2	Cuartil 3	Cuartil 4
País Vasco	80-81	0.073 (0.013)	17.4	29.1	30.3	23.2
	90-91	0.266 (0.008)	4.6	18.5	41.8	35.1
Madrid	80-81	0.027 (0.013)	18.4	37.7	25.0	18.9
	90-91	0.248 (0.015)	2.8	21.8	47.0	28.4
Navarra	80-81	0.018 (0.025)	29.0	21.6	24.3	25.0
	90-91	0.400 (0.015)	2.3	11.0	32.1	54.6
Aragón	80-81	-0.091 (0.016)	30.8	27.9	21.8	19.4
	90-91	0.281 (0.011)	5.8	16.3	36.5	41.4
Cataluña	80-81	-0.100 (0.010)	35.3	32.9	15.7	16.1
	90-91	0.321 (0.008)	3.6	16.9	39.4	40.1
Murcia	80-81	-0.120 (0.023)	33.2	37.5	14.4	14.9
	90-91	0.235 (0.020)	10.4	23.2	22.3	44.1
Valencia	80-81	-0.134 (0.012)	39.0	27.5	18.9	14.5
	90-91	0.294 (0.008)	5.7	20.5	32.3	41.5
La Rioja	80-81	-0.182 (0.028)	46.3	19.3	20.8	13.7
	90-91	0.329 (0.019)	6.4	18.4	27.5	47.7
Cantabria	80-81	-0.236 (0.027)	38.6	33.5	14.9	13.0
	90-91	0.181 (0.022)	8.4	24.8	33.2	33.7
Asturias	80-81	-0.253 (0.024)	46.3	23.4	14.7	15.6
	90-91	0.153 (0.024)	17.7	18.7	26.7	36.9
Balears	80-81	-0.267 (0.030)	46.0	23.4	15.3	15.2
	90-91	0.326 (0.019)	7.1	14.8	30.0	48.1
Castilla-León	80-81	-0.355 (0.012)	48.5	21.4	15.3	14.7
	90-91	0.237 (0.008)	10.4	19.2	28.7	41.7
Canarias	80-81	-0.414 (0.021)	55.0	28.5	9.3	7.2
	90-91	-0.093 (0.015)	31.0	37.2	20.1	11.7
Andalucía	80-81	-0.449 (0.010)	56.9	27.2	10.1	5.8
	90-91	0.130 (0.007)	13.9	30.7	28.1	27.3
Ceuta y Melilla	80-81	-0.483 (0.031)	66.6	22.0	7.6	3.8
	90-91	-0.151 (0.029)	32.6	41.2	17.8	8.3
Castilla-La Mancha	80-81	-0.511 (0.018)	58.2	21.7	10.6	9.5
	90-91	0.256 (0.010)	9.0	21.9	27.9	41.2
Galicia	80-81	-0.578 (0.020)	62.4	16.7	7.9	10.0
	90-91	0.090 (0.014)	20.5	22.7	22.7	34.1
Extremadura	80-81	-0.740 (0.023)	74.7	14.0	5.9	5.4
	90-91	0.085 (0.017)	18.1	27.6	27.1	27.2

más significativa son las que presentaban una calidad más baja en 1980-81: Andalucía, Ceuta-Melilla, Galicia, Castilla-La Mancha y Extremadura. En este sentido se puede afirmar que la desigualdad entre comunidades, en lo referente al acondicionamiento de la vivienda, ha disminuido en la década de estudio. Madrid es la CA que experimenta un menor cambio medio en su índice de calidad pasando, del segundo puesto en 1980-81, al octavo en 1990-91. El fenómeno contrario

ocurre en Castilla-La Mancha, que pasa de ocupar la antepenúltima posición en 1980-81, precediendo a Galicia y Extremadura, a situarse en el séptimo lugar.

La distribución por cuartiles de los valores del índice de calidad en el total nacional se invierte de 1980-81 a 1990-91, ganando peso los dos cuartiles superiores en detrimento de los inferiores. En 1980-81, el 70.3% de los hogares se corresponden con la calidad representada por el primer y segundo cuartil mientras que, una década después, la mayoría de los hogares (65%) se concentran en los dos últimos cuartiles, lo que permite afirmar que la calidad de la vivienda, en conjunto, ha experimentado una mejoría apreciable. Un comportamiento similar se observa en la mayoría de las CCAA, entre las que destaca Navarra por ser la que tiene mayor porcentaje de viviendas en los dos últimos cuartiles, 32.1% en 1980-81 y 54.6% en 1990-91. En Andalucía y Extremadura también se produce un cambio en el mismo sentido que en el total nacional, pero en menor medida que en la mayoría de las CCAA. Canarias y Ceuta-Melilla difieren del resto de comunidades por acumular tanto en 1980-81 como en 1990-91 más del 50% de viviendas en el primer y segundo cuartil.

## 5. Conclusiones

En este trabajo se aplica la técnica multivariante de Análisis de Correspondencias Múltiples para construir un indicador del nivel de acondicionamiento físico de la vivienda en España. El indicador es una variable ordinal que refleja un nivel de calidad relativo a una vivienda media y es aplicado para comparar la calidad de la vivienda en las distintas Comunidades Autónomas en los períodos 1980-81 y 1990-91. Las conclusiones más importantes son:

La correlación lineal media entre el índice de calidad hallado y las características que lo determinan es de 0.58. Las variables más correladas con el índice son los servicios higiénicos y el agua corriente. La pérdida de información que supone considerar la variable índice de calidad en lugar de las variables que representa, es del orden del 12%. Dicha información está recogida en dos variables incorreladas con la calidad, por construcción.

La vivienda principal ocupada en 1990-91 posee mayor calidad que en 1980-81, tanto a nivel nacional como por Comunidades Autónomas. Aunque este resultado se mantiene para todas las Comunidades, se

observa un comportamiento heterogéneo entre ellas. Navarra destaca por tener la mayor concentración de viviendas de mejor calidad en 1990-91 mientras que, Ceuta-Melilla ocupa la peor situación.

## References

- Arévalo, R. (1998): "Caracterización de la vivienda en España y determinación de su valor corriente. 1980-1990", Universidad Complutense de Madrid, Documento de Trabajo 9810.
- Benzécri, J.-P. (1979): "Sur le calcul des taux d'inertie dans l'analyse d'un questionnaire. Addendum et erratum", *Cahiers de l'Analyse des Données* 4, pp. 377-378.
- Bösch-Supan, A. (1994): "On the west german tenants protection legislation", *Journal of Institutional and Theoretical Economics* Jun-1994, pp. 411-418.
- Greenacre, M.J. (1990): "Some limitations of multiple correspondence analysis", *Computational Statistics Quarterly* 3, pp. 249-256.
- Greenacre, M.J. (1984), *Theory and Applications of Correspondence Analysis*. London: Academic Press.
- Lebart, L., A. Morineau y K.M. Warwick (1984), *Multivariate Descriptive Statistical Analysis*. Chichester, UK: Wiley.
- Olsen, E O. (1990): "What is know about the effects of rent controls?", Consulting Report for U.S. Department of Housing and Urban Development
- Peña, D. y J. Ruiz-Castillo (1984): "Distributional aspects of público rentar housing and rent control polacos in sopón", *Journal of Urban Economics* 15, pp. 350-370.
- Tenenhaus, M. y F.W. Young (1985): "Analysis and synthesis of multiple correspondence analysis, optimal scaling, dual scaling, homogeneity analysis and other methods for quantifying categorical multivariate data", *Psychometrika* 50, 1, pp. 91-119.

## Abstract

*We construct a Housing Quality Index using Spanish macroeconomic data. Housing quality seems to have improved in Spain from 1980 to 1990.*

*Recepción del original, marzo de 1998  
Versión final, enero de 1999*