

Datos de panel y demanda de alimentos en España

B. DHEHIBI (*)

J. M. GIL (**)

A. M. ANGULO (***)

1. INTRODUCCIÓN

La demanda de alimentos en España ha sido objeto de numerosos trabajos de investigación en los últimos años, utilizando tanto datos de corte transversal como de series temporales. Entre los primeros, se pueden destacar los de Laajimi y Albisu (1997), Gracia y Albisu (1998) y Manrique y Jensen (1997, 1998), por mencionar sólo los más recientes. Entre los segundos, se pueden destacar los de Molina (1994) y Gracia *et al.* (1998). En los últimos años, en la modelización de la demanda de alimentos en España se han tratado de incorporar otros determinantes además de los tradicionales factores económicos (renta y precios) y de las variables sociodemográficas, en el caso de utilizar datos de corte transversal. Así, Ben Kaabia *et al.* (2001) analizan la posible influencia de la creciente información relativa a la relación existente entre nutrición y salud sobre la demanda de carnes. Finalmente, destacar el trabajo de Angulo *et al.* (2001), en el que se considera el contenido nutricional de los alimentos, en términos de calorías ingeridas, y se analiza si las elasticidades renta y calorías de ciertos productos en los países de la Unión Europea convergen en el largo plazo.

(*) INRAT. Túnez.

(**) DEAB. Universidad Politécnica de Cataluña.

(***) Departamento de Análisis Económico. Facultad de Ciencias Económicas y Empresariales. Universidad de Zaragoza.

La mayor parte de los trabajos mencionados adolecen de cierta rigidez en un doble sentido. Por un lado, imponen, *a priori*, una determinada forma funcional sobre las ecuaciones de demanda (Sistema Casi Ideal de Demanda, Rotterdam, CBS, GADS, etc.), sin contrastar si otro modelo alternativo pudiera ajustarse mejor a los datos. Por otro lado, estos trabajos únicamente consideran una única dimensión de los datos (trasversal o temporal). La consideración conjunta de ambas dimensiones (datos de panel) permitiría mejorar sustancialmente la precisión de las estimaciones de los parámetros de los modelos, si bien hay que reconocer la dificultad de llevar a cabo este tipo de estudios en España debido a la limitada disponibilidad de información para construir un auténtico panel (1).

Recientemente, algunos trabajos han tratado de superar algunas de las dos limitaciones que acabamos de mencionar. En relación con la forma funcional, Chung (1994) analiza diversas especificaciones utilizando datos de corte transversal provinciales. Lombán (1996) adopta una forma funcional envolvente, a partir de la cual se pueden obtener especificaciones alternativas, para estimar un sistema de demanda inverso para las carnes. En relación con la dimensión de los datos, sólo el trabajo realizado por Angulo *et al.* (2002) analiza la demanda de alimentos en España sobre la base empírica de un panel de datos. Sin embargo, en dicho trabajo se considera directamente un sistema de demanda Rotterdam.

En este sentido, el objetivo del presente trabajo se centra en analizar la demanda de alimentos en España utilizando un panel de datos y prestando especial atención a la selección de la forma funcional. Así, por un lado, se analizarán cuatro de los modelos más frecuentemente utilizados en la literatura: los modelos Rotterdam, SCID, CBS y NBR. La selección entre ellos se llevará a cabo a partir de la especificación de un modelo sintético o envolvente que, a su vez, constituye un nuevo sistema general de demanda. Sobre dicho modelo sintético se realizarán una serie de contrastes paramétricos con el fin de determinar cuál de las cuatro especificaciones se ajusta mejor a los datos o, en última instancia, si es más coherente trabajar con el modelo sintético. Con respecto a la base de datos utilizada, en este trabajo se hará uso de un panel de consumidores diseñado a partir de la Encuesta Continua de Presupuestos Familiares (ECPF). Es

(1) Entre las fuentes de datos disponibles: Encuesta de Presupuestos Familiares, Encuesta Continua de Presupuestos familiares (ambas elaboradas por el Instituto Nacional de Estadística) y el Panel de Consumo Alimentario (elaborado por el Ministerio de Agricultura, Pesca y Alimentación) sólo es posible disponer de información a nivel familiar y para un periodo de tiempo razonable a partir de la Encuesta Continua de Presupuestos Familiares.

decir, utilizaremos la información sobre la demanda de alimentos efectuada por un conjunto de consumidores a lo largo de un período continuado de tiempo.

La estructura del trabajo es la siguiente. A continuación, se presenta una breve descripción de la estructura de la demanda de alimentos en España. En el apartado 3 se describen brevemente los cuatro sistemas de demanda considerados y el procedimiento utilizado para su selección. A continuación, se analiza la fuente de datos utilizada y cómo se ha construido el panel de datos. El apartado 5 describe la metodología utilizada para la estimación de sistemas de demanda con datos de panel. Los resultados obtenidos en la aplicación empírica se presentan en el apartado 6. El trabajo finaliza ofreciendo las conclusiones más relevantes.

2. ESTRUCTURA DE LA DEMANDA EN ESPAÑA

En este apartado se analizan las principales características del consumo de alimentos en el hogar, lo que ayudará a comprender mejor los resultados obtenidos en la aplicación empírica. En el cuadro 1 se recoge la evolución de la estructura del gasto en alimentación entre diferentes grupos de productos para el período 1990-1997. Las principales tendencias que se pueden observar son un aumento considerable del gasto en cereales y patatas, pescados, aceites, azúcar, refres-

Cuadro 1

EVOLUCIÓN DE LA ESTRUCTURA DEL GASTO EN ALIMENTACIÓN DE LOS HOGARES (%)

Prod. alimentarios	1990	1991	1992	1993	1994	1995	1996	1997
Cereales y patatas	13,83	14,67	14,53	15,36	15,82	16,01	16,41	15,22
Carne roja	16,87	16,64	17,83	17,04	16,00	15,50	15,12	16,16
Carne blanca	7,59	7,01	6,47	6,97	6,98	6,47	6,29	5,72
Pescados	10,86	11,38	12,00	11,94	11,52	10,42	11,02	12,64
Lácteos	12,33	11,69	10,73	11,60	11,46	10,98	10,07	9,74
Huevos	4,20	3,27	3,12	3,09	3,07	2,72	3,73	2,63
Aceites	4,55	5,13	4,84	4,31	5,19	5,82	6,65	7,80
Frutas	12,76	13,40	13,37	11,86	12,61	13,59	13,06	12,49
Hortalizas	8,35	7,77	7,94	7,62	7,97	8,11	7,69	5,84
Azúcar y otros (a)	4,27	4,35	4,51	5,12	4,82	4,99	4,35	5,08
Otros productos (b)	3,79	3,98	4,42	4,35	4,30	4,68	5,05	6,29 (a)

(a) Incluye preparados de azúcar y refrescos y bebidas gaseosas.

(b) Incluye bebidas alcohólicas y otros productos alimenticios (café, infusiones,...).

Fuente: Elaboración propia a partir de la ECPF (varios años).

Cuadro 2

**ESTRUCTURA PORCENTUAL DEL GASTO MEDIO DE ALIMENTACIÓN ATENDIENDO
A LAS CARACTERÍSTICAS DEMOGRÁFICAS Y SOCIOECONÓMICAS DEL HOGAR
PARA EL PERÍODO 1990-1997 (%)**

	Cereales y patatas	Carne roja	Carne blanca	Pesca- dos	Lácteos	Huevos	Aceites	Frutas	Hortal.	Azúcar ^a	Otros ^b product.
TAMAÑO DE MUNICIPIO											
< de 10.000	19,84	13,87	7,52	9,25	11,08	3,03	6,14	12,66	8,17	4,49	3,97
10.001 - 50.000	14,96	18,48	5,91	10,72	11,46	3,34	5,62	12,81	6,96	4,63	4,87
50.001 - 500.000	14,08	18,18	6,60	12,80	10,52	2,74	5,68	13,54	6,8	4,98	4,06
> 500.000	13,53	15,98	6,76	11,80	10,85	2,82	4,68	14,08	9,32	4,93	5,23
NIVEL DE ESTUDIOS DEL SUSTENTADOR PRINCIPAL											
Sin Estudios	19,18	12,08	7,80	9,90	11,20	3,14	6,26	13,28	8,06	4,60	4,06
Estudios Primarios	13,94	16,82	6,64	11,54	10,60	2,89	5,89	12,86	7,83	4,18	4,06
Estudios Medios	14,20	19,38	6,64	11,30	10,58	2,92	5,24	12,40	7,36	4,88	4,30
Estudios Superiores	9,84	17,06	5,38	13,30	11,36	2,44	4,74	14,82	8,32	3,76	5,86
EDAD DEL SUSTENTADOR PRINCIPAL DE LA FAMILIA											
Entre 18 y 25 Años	16,42	17,46	7,48	5,48	14,00	3,78	5,26	11,08	7,12	7,30	4,16
De 26 a 45 Años	14,18	17,52	6,78	10,48	11,70	2,90	4,58	12,62	7,36	5,06	4,32
De 46 a 65 Años	16,62	16,46	6,72	11,68	9,92	2,98	5,90	12,76	8,22	3,94	4,34
Más de 65 Años	15,60	14,18	6,82	11,90	10,64	2,64	7,18	14,26	7,90	4,18	3,48
COMPOSICIÓN FAMILIAR: % DE MIEMBROS MAYORES DE 13 AÑOS											
25% - 50%	15,22	15,12	6,86	10,52	11,70	2,80	5,96	12,84	7,98	4,32	4,34
50% - 75%	13,82	15,42	7,88	11,63	11,37	3,08	6,92	13,18	8,06	4,32	4,20
Más del 75%	12,84	15,92	7,06	11,76	11,51	3,25	7,13	13,38	7,86	4,44	4,20

(a) Incluye preparados de azúcar y refrescos y bebidas gaseosas.

(b) Incluye bebidas alcohólicas y otros productos alimenticios (café, infusiones,...).

Fuente: Elaboración propia a partir de la ECPF (varios años).

cos, bebidas gaseosas y otros productos. Por el contrario, en el resto de productos (carnes, lácteos, huevos, frutas y hortalizas) se observa una clara tendencia a la baja.

Una vez analizada la evolución de la estructura del gasto en términos generales, es conveniente, a su vez, analizar las diferencias existentes en dicha estructura atendiendo a las principales características demográficas y económicas del hogar: 1) el tamaño de municipio donde habita la familia; 2) el nivel de educación del sustentador principal; 3) su edad; y 4) la composición de la unidad familiar, medida a partir del porcentaje de miembros mayores de 13 años. Los resultados obtenidos, utilizando los valores medios correspondientes al período 1990-1997, se recogen en el cuadro 2.

Con respecto a la variable tamaño de municipio, se aprecia como en los municipios de gran tamaño (más de 500.000 habitantes) disminuye considerablemente el gasto destinado a cereales y patatas, lácteos, huevos y aceites. Los municipios con un número de habitantes

entre 50.001 y 500.000, que son los que representan un mayor porcentaje de la población total, presentan una estructura del gasto en alimentación similar a la media nacional. Por el contrario, las participaciones del gasto en alimentación en los municipios de menor tamaño son, en general, inferiores a dicha media.

En relación con el efecto del nivel de educación del sustentador principal, las familias cuyo cabeza posee estudios superiores gastan, en términos relativos, un mayor porcentaje en carne roja, pescados, lácteos y frutas. En cambio, las familias cuyo cabeza no posee estudios destinan un mayor porcentaje de su gasto a adquirir cereales y patatas, carne blanca, huevos y aceite. En el resto de segmentos (estudios primarios y estudios medios), se observa que la estructura del gasto en alimentación no difiere sustancialmente con respecto a la media nacional.

En cuanto a la edad del sustentador principal, el porcentaje de gasto en alimentación destinado a adquirir pescados, aceites, frutas y hortalizas aumenta de forma espectacular en los hogares cuyo cabeza de familia supera los 46 años. Por el contrario, en los hogares cuyo cabeza de familia es más joven, se aprecian mayores participaciones de gasto destinadas a adquirir carne roja, carne blanca, y, sobre todo, productos lácteos, debido a la posible existencia de hijos pequeños.

Finalmente, en relación con la composición del hogar, medida a través del porcentaje de miembros que superan los trece años, se puede concluir que conforme aumenta la media de edad de los miembros de la unidad familiar el porcentaje de gasto en alimentación destinado a adquirir carne roja, pescados, aceites, frutas y hortalizas tiende a ser mayor. Sin embargo, este porcentaje disminuye cuando se adquieren productos como carne blanca y lácteos. En el resto de productos (azúcar, refrescos y bebidas gaseosas y bebidas alcohólicas) no se aprecian diferencias significativas.

3. SISTEMAS DE DEMANDA. SELECCIÓN DE LA FORMA FUNCIONAL

La teoría tradicional de la demanda establece que el consumidor racional trata de maximizar la función de utilidad, que depende de las cantidades de un conjunto de n bienes, sujeto a su restricción presupuestaria. Como resultado del anterior proceso de optimización se obtienen las siguientes funciones de demanda marshallianas:

$$q_i = f_i(x, p_1, \dots, p_n) \quad i=1, 2, \dots, n \quad [1]$$

que indican que la demanda de un bien (q_i) depende del nivel de renta x y de los precios de los n bienes (p_i).

A la hora de estimar [1] es necesario especificar una forma funcional. Como se ha comentado en la introducción, la mayor parte de los trabajos han elegido *a priori* una determinada forma funcional a partir de la cual se contrastaban las restricciones sugeridas por la teoría y se calculaban las correspondientes elasticidades. Sin duda, los sistemas a los que se ha prestado mayor atención han sido el modelo Rotterdam (a partir del cual se puede obtener el modelo CBS [Keller y van Driel (2), 1985]) y el Sistema Casi Ideal de Demanda (SCID) (a partir del cual Neves (1987) obtuvo el sistema NBR). En este trabajo, en vez de seleccionar *a priori* uno de estos modelos, lo que hemos hecho ha sido especificar una familia paramétrica de formas funcionales, de tal forma que, dando valores a unos determinados parámetros, se pudiesen obtener los cuatro modelos anteriormente definidos.

Antes de definir el modelo sintético, revisaremos brevemente las expresiones de los cuatro modelos considerados.

3.1. El modelo de Rotterdam

El modelo Rotterdam, propuesto inicialmente por Theil (1965), se obtiene a partir de la diferenciación logarítmica de la función de demanda expresada en [1] y viene dado por:

$$w_i d \ln q_i = \alpha_i + \theta_i d \ln Q + \sum_{j=1}^n \pi_{ij} d \ln p_j \quad [2]$$

donde: w_i representa la participación en el gasto total en alimentación del producto i .

α_i recoge el cambio de gustos relativo al producto i .

$\theta_i = w_i \frac{\partial \ln q_i}{\partial \ln x}$ es la propensión marginal a consumir, considerada constante.

π_{ij} son los coeficientes de Slutsky que también se suponen constantes.

$d \ln Q$ es el índice divisa de volumen, definido por:

$$d \ln Q = \sum_{j=1}^n w_j d \ln q_j = d \ln x - \sum_{j=1}^n w_j d \ln p_j \quad [3]$$

(2) El nombre de CBS proviene de que los dos autores reseñados trabajaban en el Central Bureau of Statistics en Holanda.

3.2. El Sistema Casi Ideal de Demanda (SCID)

El modelo SCID, propuesto por Deaton y Muellbauer (1980), se obtiene a partir de la función de costes derivada a partir del planteamiento dual del problema de partida (es decir, el consumidor minimiza el gasto para alcanzar un determinado nivel de utilidad). Aunque el modelo de partida es no lineal, la aplicación más frecuente del mismo se formula a partir de la siguiente linealización:

$$w_i = c_i + \rho_i(\ln x - \sum_{j=1}^n w_j \ln p_j) + \sum_{j=1}^n \Psi_{ij} \ln p_j \quad [4]$$

Como se puede apreciar, el modelo Rotterdam y el SCID no se pueden relacionar directamente comparando las expresiones [2] y [4]. Para poder relacionar ambos modelos es necesario obtener una expresión diferencial del modelo SCID. Diferenciando la expresión anterior, y teniendo en cuenta la ecuación [3], llegamos a la siguiente expresión de un sistema SCID diferencial:

$$dw_i = \alpha_i + \rho_i d \ln Q + \sum_{j=1}^n \Psi_{ij} d \ln p_j \quad [5]$$

que, como puede observarse, sólo se diferencia de la versión del modelo Rotterdam recogida en [2] en la variable dependiente. Además, Barten (1993) obtuvo la relación existente entre los parámetros de ambos modelos a partir del siguiente procedimiento. Si se diferencia logarítmicamente la participación presupuestaria

$$w_i = \frac{P_i q_i}{x}$$

y se multiplica por w_i , se obtiene:

$$w_i d \ln w_i = dw_i = w_i d \ln p_i + w_i d \ln q_i - w_i d \ln x \quad [6]$$

Si, a continuación, en dicha expresión se sustituye $w_i d \ln q_i$ por la parte derecha del modelo de Rotterdam y se tiene en cuenta la expresión [3], se obtiene una expresión alternativa para el modelo SCID diferenciado:

$$dw_i = \alpha_i + (\theta_i - w_i) d \ln Q + \sum_{j=1}^n (\pi_{ij} + w_i \delta_{ij} - w_i w_j) d \ln p_j \quad [7]$$

o bien,

$$w_i d \ln q_i = \alpha_i + (\rho_i + w_i) d \ln Q + \sum_{j=1}^n (\psi_{ij} - w_i \delta_{ij} + w_i w_j) d \ln p_j \quad [8]$$

donde δ_{ij} es el delta de Kroneker que toma el valor unitario si $i = j$, y cero, en caso contrario.

De esta forma, se obtiene la relación existente entre los parámetros de los modelos SCID y Rotterdam, que viene dada por:

$$\rho_i = \theta_i - w_i \quad \text{y} \quad \psi_{ij} = \pi_{ij} + w_i \delta_{ij} - w_i w_j$$

3.3. El Modelo CBS

En este modelo, desarrollado por Keller y Van Driel (1985), se trata de eliminar la restricción de parámetros constantes del modelo Rotterdam. Se obtiene mediante la sustitución de la propensión marginal al consumo (θ_i) del modelo Rotterdam [2], por la expresión $(\rho_i + w_i)$ (3):

$$w_i d \ln q_i = \alpha_i + (\rho_i + w_i) d \ln Q + \sum_{j=1}^n \pi_{ij} d \ln p_j \quad [9]$$

El modelo CBS se considera un modelo híbrido entre el modelo SCID (ya que posee los mismos coeficientes renta que éste) y el modelo Rotterdam (ya que ambos comparten los mismos coeficientes precio y la misma variable dependiente).

3.4. El Modelo NBR

El modelo NBR fue propuesto inicialmente por Neves (1987). Al igual que el modelo anterior, se trata de un sistema híbrido, pero, en este caso, incorpora los coeficientes renta del modelo Rotterdam y los coeficientes precio del modelo SCID. En efecto, si en [7] sustituimos el parámetro correspondiente al precio por el parámetro ψ_{ij} obtenemos:

$$dw_i = \alpha_i + (\theta_i - w_i) d \ln Q + \sum_{j=1}^n \Psi_{ij} d \ln p_j \quad [10]$$

o bien,

(3) Esta expresión corresponde a las participaciones de gasto marginales obtenidas a partir del modelo de Working (1943): $w_{ij} = \mu_i + \rho_i \log x$.

$$w_i d \ln q_i = \alpha_i + \theta_i d \ln Q + \sum_{j=1}^n (\psi_{ij} - w_i \delta_{ij} + w_i w_j) d \ln p_j \quad [11]$$

Como se puede apreciar, los modelos Rotterdam, SCID, CBS y NBR representados por [2], [8], [9] y [11] comparten las mismas variables, pero no se encuentran anidados. Para poder seleccionar el modelo que mejor se ajuste a los datos disponibles, Barten (1993) definió un sistema de demanda sintético o envolvente que anidaba a los cuatro modelos mencionados, de tal forma que para la selección del mejor modelo bastaba con realizar una serie de contrastes sobre dicho sistema envolvente. Este sistema general viene definido por la siguiente expresión:

$$w_i d \ln q_i = \alpha_i + (d_i + \delta_1 w_i) d \ln Q + \sum_{j=1}^n (r_{ij} - \delta_2 w_i (\delta_{ij} - w_j)) d \ln p_j \quad [12]$$

donde:

$$d_i = \delta_1 \rho_i + (1 - \delta_1) \theta_i$$

$$r_{ij} = \delta_2 \psi_{ij} + (1 - \delta_2) \pi_{ij}$$

δ_1 y δ_2 son parámetros a estimar.

Nótese que si en [12] se impone $\delta_1 = \delta_2 = 0$, se obtiene el modelo Róterdam definido en [2]. El modelo SCID recogido en [8] es la especificación preferida si la hipótesis nula de que no puede ser rechazada. Si se cumple la restricción $\delta_1 = 1, \delta_2 = 0$, entonces se obtiene el modelo CBS definido en [9]. Finalmente, el modelo NBR [11] sería el modelo preferido si no es posible rechazar las hipótesis $\delta_1 = 0$ y $\delta_2 = 1$. En caso de no cumplirse ninguna de las restricciones expuestas sobre los parámetros, la solución propuesta consistiría en estimar el modelo envolvente definido en [12], ya que dicho modelo constituye en sí mismo un sistema de demanda.

En dicho modelo, las restricciones de agregación, homogeneidad y simetría vienen dadas por:

Agregación:
$$\sum_{i=1}^n \alpha_i = 0, \sum_{i=1}^n d_i = 1 - \delta_1 \text{ y } \sum_{i=1}^n r_{ij} = 0$$

Homogeneidad:
$$\sum_{j=1}^n r_{ij} = 0$$

Simetría:
$$r_{ij} = r_{ji}$$

La propiedad de negatividad será satisfecha si la matriz formada a partir de los elementos r_{ij} es semidefinida negativa. Finalmente, las elasticidades vendrán dadas por las siguientes expresiones:

$$\text{Elasticidad renta: } \eta_i^{\text{SINT}} = \frac{(d_i + \delta_1 w_i)}{w_i}$$

Elasticidad precio-directa no compensada:

$$\mu_{ii}^{\text{SINT}} = \frac{(r_{ii} - \delta_2 w_i (\delta_{ii} - w_i)) - (d_i + \delta_1 w_i) w_i}{w_i}$$

Elasticidad precio-cruzada no compensada:

$$\mu_{ij}^{\text{SINT}} = \frac{(r_{ij} - \delta_2 w_i (\delta_{ij} - w_j)) - (d_i + \delta_1 w_i) w_j}{w_i}$$

Elasticidad precio-directa compensada:

$$e_{ii}^{\text{SINT}} = \frac{r_{ii} - \delta_2 w_i (\delta_{ii} - w_i)}{w_i}$$

Elasticidad precio-cruzada compensada:

$$e_{ij}^{\text{SINT}} = \frac{r_{ij} - \delta_2 w_i (\delta_{ij} - w_j)}{w_i}$$

A partir de estas expresiones, si se introducen las restricciones sobre δ_1 y δ_2 mencionadas anteriormente, se obtienen las elasticidades correspondientes a cada modelo.

4. DATOS UTILIZADOS Y TRANSFORMACIONES PREVIAS

La información disponible proviene de la Encuesta Continua de Presupuestos Familiares (ECPF), elaborada por el Instituto Nacional de Estadística (INE). La información recogida, para cada hogar y trimestre, se refiere al gasto y cantidades consumidas de diversos productos durante una semana de dicho trimestre. El tamaño muestral de cada uno de los trimestres oscila en torno a 3.200 hogares, seleccionados con los condicionantes de estratificación y aleatoriedad requeridos estadísticamente. Esta muestra se renueva cada trimestre en una octava parte y, por tanto, ofrece la posibilidad de seguir el comportamiento de un hogar durante 8 trimestres (2 años).

En este trabajo, se ha optado por trabajar con un panel de datos que abarca los trimestres comprendidos entre el primero de 1996 y el cuarto de 1997, ambos incluidos. No obstante, se han eliminado de la muestra aquellas observaciones atípicas que pudieran distorsionar los resultados obtenidos, concretamente aquellas familias que pre-

sentaban gastos nulos en todos los grupos de productos, o en todos menos en uno. El porcentaje de familias eliminadas ha sido del 13,88 por ciento. El resultado de esta labor es un panel real de consumidores con un tamaño muestral de 1.736 observaciones.

La información así obtenida contempla dos tipos de aspectos. Por un lado, tiene en cuenta la evolución temporal en el comportamiento de una serie de familias (de forma análoga a los análisis de series temporales), a la vez que, por otro lado, se dispone de una muestra considerable de familias para cada uno de los períodos de tiempo analizados (de modo similar a un corte transversal). La consideración conjunta de estas dos dimensiones (temporal y transversal) dota al análisis de importantes ventajas. Por un lado, el gran número de observaciones disponibles permite aumentar los grados de libertad del modelo mejorando, por tanto, la eficiencia de las estimaciones. Además, también se consigue mejorar la precisión de las mismas, gracias a la menor colinealidad entre las variables explicativas del modelo. Por último, los datos de panel permiten reducir o evitar sesgos en la estimación del modelo al controlar tanto los efectos individuales como los temporales.

La información disponible se refiere al gasto y a las cantidades consumidas de distintos productos alimenticios, así como a toda una serie de características sociodemográficas referidas a las distintas familias. Los grupos de productos considerados en este trabajo son los siguientes: 1) cereales y patatas; 2) carne roja; 3) carne blanca y huevos; 4) pescados; 5) lácteos; 6) aceites; 7) frutas; 8) hortalizas; 9) azúcar (incluye todos los preparados de azúcar, así como los refrescos y las bebidas gaseosas); y 10) otros productos (se incluyen, fundamentalmente, las bebidas alcohólicas, así como infusiones, café y otros productos no contemplados en los epígrafes anteriores, con el fin de considerar la totalidad del gasto en alimentación). En relación con las variables explicativas, se han considerado, además del gasto total en alimentación, como aproximación de la renta (4), y de los precios de todos los productos analizados, las siguientes variables sociodemográficas: 1) el tamaño de municipio en el que habita la familia; 2) el nivel de educación del sustentador principal; 3) su edad; y 4) la composición del hogar, medida a través del porcentaje de miembros mayores de trece años.

Por último, la fuente utilizada no ofrece información directa acerca de los precios de los respectivos productos. En estos casos, una alter-

(4) Se ha supuesto separabilidad débil de las preferencias.

nativa usada tradicionalmente ha consistido en calcular los llamados «índices de valor unitario» obtenidos al dividir el gasto entre las cantidades consumidas (Deaton, 1989). No obstante, la utilización de estos índices plantea dos importantes problemas. En primer lugar, dado que el valor unitario está afectado por las decisiones del consumidor relativas tanto a la cantidad como a la calidad de producto, aquél no puede igualarse al precio del producto. Por otra parte, es de esperar que las familias de mayor renta tiendan a adquirir productos de mayor calidad y, por tanto, aparezcan problemas de correlación entre los valores unitarios y la renta. Teniendo en cuenta estos problemas, en este trabajo se han calculado los precios ajustando los valores unitarios mediante las diferencias en calidad, en la línea llevada a cabo por autores como Cowling y Raynor (1970), Gao *et al.* (1995) o Gao *et al.* (1997). Este ajuste se realiza a partir de la estimación de una ecuación de precios hedónicos, que relaciona el precio con ciertas características sociodemográficas de la familia, con objeto de aproximar la selección de la calidad de cada una de las familias.

5. ESTIMACIÓN DE SISTEMAS DE DEMANDA CON DATOS DE PANEL

Una vez contextualizado el marco teórico del análisis y definido el tipo de datos que se van a utilizar, quedaría por concretar el marco econométrico en el que se va a desenvolver este trabajo. En este sentido, es preciso tener en cuenta que la complejidad de los análisis relativos a sistemas con datos de panel ha supuesto que prácticamente toda la literatura existente relativa al tratamiento de datos de panel se haya llevado a cabo en el marco uniecuacional (Baltagi, 1999). Por este motivo, en este apartado trataremos de describir los rasgos básicos de la estimación en sistemas multiecuacionales (5).

Dentro del contexto de datos de panel, nos ceñiremos a la especificación más general conocida como especificación de efectos aleatorios, cuya expresión, para el caso del modelo sintético definido en [12], viene dada por:

$$\begin{aligned}
 w_{imt} d \ln q_{imt} &= \alpha_i + (d_i + \delta_1 w_{imt}) d \ln Q + \\
 &\sum_{j=1}^n (r_{ij} - \delta_2 w_{imt} (\delta_{ij} - w_{jmt})) d \ln p_{jt} + u_{imt} \\
 u_{imt} &= \mu_{im} + v_{imt}
 \end{aligned}
 \tag{13}$$

(5) Una descripción más extensa puede encontrarse en Arellano y Bover (1990) y Arellano (2003).

donde w_{imt} y q_{imt} son, respectivamente, la participación presupuestaria y la cantidad del bien genérico i ($i=1, \dots, n$) en el período t ($t=1, \dots, T$), para la familia m ($m=1, \dots, M$); p_{jt} es el precio del bien j en el período t ; y, finalmente, el término de error u_{imt} está compuesto por los efectos específicos individuales μ_{im} , diferente para cada individuo pero constante en el tiempo, y el término de error convencional v_{imt} . En forma matricial el sistema [13] puede expresarse como:

$$\begin{aligned} Y_j &= Z_j \delta_j + u_j & j=1, 2, \dots, n & \quad [14] \\ u_j &= Z_\mu \mu_j + v_j \end{aligned}$$

donde: Y_j es un vector de orden de $MT \times 1$

Z_j es una matriz $MT \times k_j$

$\delta'_j = (\alpha'_j, \beta'_j)$ es un vector de k_j parámetros

$Z_\mu = (I_M \otimes I_T)$

$\mu'_j = (\mu_{1j}, \mu_{2j}, \dots, \mu_{Mj})$ y $v'_j = (v_{11j}, v_{12j}, \dots, v_{1Tj}, \dots, v_{M1j}, \dots, v_{MTj})$ son vectores de perturbaciones aleatorias con media cero y matriz de varianzas y covarianzas.

$$E \begin{pmatrix} \mu_j \\ v_j \end{pmatrix} \begin{pmatrix} \mu_i & v_i \end{pmatrix} = \begin{bmatrix} \sigma_{\mu_{ji}}^2 I_M & 0 \\ 0 & \sigma_{v_{ji}}^2 I_{MT} \end{bmatrix} \quad i, j = 1, 2, \dots, n$$

dado que $\mu \sim (0, \sum_\mu \otimes I_M)$ y $v \sim (0, \sum_v \otimes I_{MT})$, siendo $\mu = (\mu_1, \mu_2, \dots, \mu_n)$; $v = (v_1, v_2, \dots, v_n)$; $\sum_\mu = [\sigma_{\mu_{ij}}^2]$; $\sum_v = [\sigma_{v_{ij}}^2]$; I_M , la matriz identidad de dimensión M ; e I_{MT} , la matriz identidad de dimensión MT .

A partir de [14] la submatriz de varianzas y covarianzas para los productos j e i puede definirse como:

$$\Omega_{ij} = E (\hat{u}_j \hat{u}'_i) = \sigma_{\mu_{ji}}^2 (I_M \otimes J_T) + \sigma_{v_{ji}}^2 (I_M \otimes I_T)$$

donde J_T es una matriz de unos de dimension T . La matriz de varianzas y covarianzas del conjunto de las n ecuaciones adoptará, por tanto, la siguiente expresión:

$$\Omega = \sum (\hat{u} \hat{u}') = \sum_\mu \otimes (I_M \otimes J_T) + \sum_v \otimes (I_M \otimes I_T) \quad [15]$$

Por último, definiendo $\bar{J}_T = \frac{J_T}{T}$, reemplazando J_T por $T\bar{J}_T$ en la expresión anterior, y agrupando términos, se obtiene la expresión:

$$\begin{aligned} \Omega &= (T \sum_\mu + \sum_v) \otimes (I_M \otimes \bar{J}_T) + \sum_v \otimes (I_M \otimes E_T) = & [16] \\ &= \sum_\mu \otimes P + \sum_v \otimes Q \end{aligned}$$

donde: $\Sigma_l = T\Sigma_\mu + \Sigma_v$

$P = I_M \otimes \bar{J}_T$, matriz que recoge las medias de las observaciones de cada individuo en el tiempo.

$Q = I_M \otimes E_T = I_M \otimes (I_T - \bar{J}_T) = I_{MT} - P$, matriz que recoge las desviaciones de las familias con respecto a la media en el tiempo.

De esta forma, la matriz definida en [16] no es sino la descomposición espectral de Ω derivada por Baltagi (1980), donde:

$$\Omega^r = \Sigma_l^r \otimes P + \Sigma_v^r \otimes Q$$

para cualquier potencia r . Para $r = -1$, se obtiene la matriz inversa Ω^{-1} y para $r = -1/2$, obtenemos:

$$\Omega^{-1/2} = \Sigma_l^{-1/2} \otimes P + \Sigma_v^{-1/2} \otimes Q$$

Finalmente, el estimador mínimo cuadrado generalizado de δ vendrá dado por:

$$\hat{\sigma}_{MCG} = (Z' \Omega^{-1} Z)^{-1} Z' \Omega^{-1} Y \quad [17]$$

De cara a obtener la estimación mínimo cuadrática generalizada factible (MCGF) de δ , se necesita la estimación de Σ_l y Σ_v . Prucha (1984) muestra que si el estimador de Σ_v es consistente y el estimador Σ_μ tiene un límite positivo finito, el estimador factible [17] es asintóticamente eficiente. Según Baltagi (1999), Σ_v se puede estimar por $\hat{U}' Q \hat{U} / M(T-1)$ y Σ_l por $\hat{U}' P \hat{U} / M$ donde $\hat{U} = [\hat{u}_1, \dots, \hat{u}_n]$ es la matriz $MT \times n$ de residuos estimados por Mínimos Cuadrados Ordinarios (MCO) (Wallace y Hussain, 1969).

6. RESULTADOS Y DISCUSIÓN

Tal y como se ha descrito con anterioridad, en este trabajo se pretende analizar la demanda de alimentos en España a partir del modelo que mejor se ajuste a los datos. En este sentido, dado que la selección entre los modelos alternativos se lleva a cabo mediante los contrastes indicados sobre el modelo general o sintético, el primer paso del trabajo ha consistido en estimar el modelo definido en [13], incluyendo las correspondientes variables sociodemográficas, mediante el procedimiento descrito en el apartado anterior. Las restricciones de agregación, homogeneidad y simetría se han impuesto

en el modelo (6). En el caso de la negatividad, se ha comprobado que todos los valores propios de la matriz constituida a partir de los elementos r_{ij} fueran negativos. En el modelo estimado, los valores propios estimados han sido: $-0,1083$; $-0,0784$; $-0,0564$; $-0,0395$; $-0,0213$; $-0,0214$; $-0,0141$; $-0,0100$; $-0,0036$, respectivamente, mientras que el último valor propio ha sido nulo. Por lo tanto, el modelo estimado cumplía la condición de negatividad.

Finalmente, se han sometido al modelo estimado diversos contrastes de especificación. En primer lugar, se ha contrastado la posible presencia de heterocedasticidad utilizando el estadístico de White (1980) para cada una de las ecuaciones. Los resultados obtenidos indicaban que sólo en dos de las ecuaciones estimadas (carne roja y carne blanca y huevos) se podían apreciar problemas de heterocedasticidad. Teniendo en cuenta estos resultados, en este trabajo se ha considerado que el modelo estimado no presentaba problemas graves de heterocedasticidad, optándose por calcular las desviaciones típicas robustas de los parámetros estimados utilizando las fórmulas propuestas por White (1980) en las ecuaciones con problemas de heteroscedasticidad.

En segundo lugar, se ha contrastado la posible presencia de autocorrelación, si bien a priori no parecía que éste fuese un problema relevante ya que el número de períodos de tiempo considerados era muy reducido. Se han llevado a cabo dos contrastes propuestos por Bera *et al.* (2001). El primero (RS_{μ}) contrasta exclusivamente la existencia de correlación serial bajo la hipótesis mantenida de un modelo de efectos aleatorios, mientras que el segundo ($RS_{\mu\rho}$) contrasta conjuntamente la especificación conjunta de un modelo de efectos aleatorios y ausencia de autocorrelación. Ambos estadísticos se distribuyen como una χ^2 con 1 y 2 grados de libertad, respectivamente. Los resultados obtenidos de ambos contrastes indican que, en ningún caso se rechaza la hipótesis nula de ausencia de autocorrelación (7). Tras obtener evidencia empírica a favor de la correcta especificación del sistema general, se ha procedido a seleccionar la forma funcional

(6) La restricción de agregación ha requerido eliminar una ecuación del sistema para evitar el problema de singularidad de la matriz de varianzas y covarianzas de las perturbaciones. En nuestro caso, la ecuación eliminada ha sido la correspondiente a los otros productos alimentarios, cuyos parámetros se han obtenido a partir de las condiciones de agregación expuestas anteriormente.

(7) Los valores obtenidos del estadístico RS_{μ} para las 10 ecuaciones fueron: cereales (1,56), carne roja (1,13), carne blanca (1,12), pescados (1,65), lácteos (2,14), aceites (1,99), frutas (1,78), hortalizas (2,86), azúcar (2,56), y otros (2,17), todos ellos inferiores al valor crítico al 5 por ciento que es de 3,84. Los valores obtenidos del estadístico $RS_{\mu\rho}$ para las 10 ecuaciones fueron: cereales (2,45), carne roja (3,34), carne blanca (2,32), pescados (2,98), lácteos (2,10), aceites (3,49), frutas (4,81), hortalizas (3,68), azúcar (4,05), y otros (3,97), todos ellos inferiores al valor crítico al 5 por ciento que es de 5,99.

que mejor se ajusta a los datos. Como se mencionó en el apartado 3, a partir del modelo sintético estimado se pueden obtener como casos particulares los modelos Rotterdam, SCID, CBS y NBR introduciendo ciertas restricciones sobre los parámetros estimados. En el cuadro 3 se recogen los resultados de dichos contrastes. Como se puede apreciar, los cuatro modelos planteados son rechazados por los datos y, por tanto, se obtiene evidencia empírica a favor del modelo general o sintético.

Cuadro 3

CONTRASTE DE WALD PARA SELECCIONAR LA FORMA FUNCIONAL ADECUADA

Sistemas de demanda	Contraste	Test de Wald	Grados de libertad	Valor crítico (5%)
Róterdam	$\delta_1 = \delta_2 = 0$	96,24	2	5,99
SCID	$\delta_1 = \delta_2 = 1$	58,43	2	5,99
CBS	$\delta_1 = 1$ y $\delta_2 = 0$	27,66	2	5,99
NBR	$\delta_1 = 0$ y $\delta_2 = 1$	80,60	2	5,99

En cualquier caso, y para finalizar con el diagnóstico del modelo elegido, se ha calculado la bondad del ajuste del sistema calculando el coeficiente de determinación conjunto propuesto por Bewley y Young (1987) y se ha comparado con el que se obtendría de estimar cada uno de los cuatro modelos anidados (Rotterdam, SCID, CBS y NBR). Los resultados obtenidos se recogen en el cuadro 4. Como puede apreciarse, el coeficiente de determinación del modelo sintético seleccionado es igual a 0,51, el cual es relativamente alto teniendo en cuenta que la variable dependiente está definida en primeras diferencias. Asimismo, se observa como los coeficientes de determinación de los cuatro modelos alternativos son considerablemente más bajos.

Una vez comprobada la correcta especificación del modelo seleccionado, procederemos a describir los principales resultados obtenidos.

Cuadro 4

COEFICIENTE DE DETERMINACIÓN CONJUNTO DE LOS DIFERENTES SISTEMAS DE DEMANDA

Sistemas de demanda	Sintético	Rotterdam	SCID	CBS	NBR
R^2	0,51	0,41	0,44	0,46	0,26

En primer lugar analizaremos los parámetros estimados para, posteriormente, pasar al análisis de las correspondientes elasticidades y efectos sociodemográficos. En relación con las estimaciones de los parámetros del modelo (cuadro 5) (8), puede observarse como la mayor parte de los mismos son reveladores al nivel de significación del 5 por ciento. Todos los coeficientes relativos a la variable gasto total en alimentación son positivos e inferiores a la unidad, mientras que todos los coeficientes propio precio han resultado ser negativos. Finalmente, la mayoría de los coeficientes precio cruzados estimados son positivos.

Cuadro 5

PARÁMETROS GASTO Y PRECIOS ESTIMADOS DEL MODELO SELECCIONADO

	Cereales y patatas	Carne roja	Carne blanca y huevos	Pescados	Lácteos	Aceites	Frutas	Hortalizas	Azúcar*	Otros* product.
Constante	0,66**	-0,21	0,84**	0,075	-0,44	-0,36	-0,81**	-0,11	0,52	-0,165
Gasto	0,031**	0,052**	0,036**	0,025**	0,021**	0,015**	0,062**	0,032**	0,002**	0,74
Precios cereales y patatas	-0,10**	-	-	-	-	-	-	-	-	-
Carne roja	0,013**	-0,014	-	-	-	-	-	-	-	-
C. blanca y huevos	0,005	0,012**	-0,06**	-	-	-	-	-	-	-
Pescados	0,002	-0,006	0,011**	0,019**	-	-	-	-	-	-
Lácteos	0,032**	0,0002	0,004	-0,012**	-0,036**	-	-	-	-	-
Aceites	0,004	-0,003	0,058**	-0,003	-0,003	0,012**	-	-	-	-
Frutas	0,032**	0,0001	0,015**	-0,007	0,012**	-0,006*	-0,054**	-	-	-
Hortalizas	0,012**	0,004	0,004	-0,002	-0,006	-0,004	-0,003	-0,004	-	-
Azúcar (a)	0,003	0,0001	-0,002	0,003	0,003	-0,005**	0,009**	-0,003	-0,009**	-
Otros productos (b)	-0,003	-0,006	-0,046**	-0,002	-0,049**	-0,05	0,002	0,003	0,001	0,146**

(a) Incluye preparados de azúcar y refrescos y bebidas gaseosas.

(b) Incluye bebidas alcohólicas y otros productos alimenticios (café, infusiones,...).

Nota:

** Indica la significatividad de los parámetros al 5%.

* Indica la significatividad de los parámetros al 10%.

A partir de dichos parámetros estimados, se ha procedido al cálculo de las correspondientes elasticidades gasto y precio marshallianas y hicksianas considerando los valores medios. Las elasticidades con respecto al gasto total en alimentación, recogidas en la primera

(8) Únicamente recogemos en el cuadro 5 los parámetros correspondientes al gasto y a los precios. Los parámetros correspondientes a las variables sociodemográficas serán analizados más adelante.

columna del cuadro 6, son positivas e individualmente significativas al 5 por ciento. Sin embargo, mientras que los grupos referidos a carne roja, carne blanca y huevos, frutas, hortalizas y otros productos (bebidas alcohólicas) pueden considerarse como bienes de lujo, el resto de productos son bienes de primera necesidad. Los resultados obtenidos son bastante consistentes con los esperados. Quizás en el caso de las carnes rojas, los pescados y los lácteos se podía haber esperado una elasticidad gasto superior debido a que los precios de los componentes de estos grupos suelen ser altos. Sin embargo, en la Encuesta Continua de Presupuestos Familiares no se dispone de información sobre cantidades consumidas para todos los productos alimenticios, aunque los productos incluidos dentro de cada uno de los grupos de alimentos considerados en este trabajo aproximadamente representan entre el 60 y el 70 por ciento del gasto efectuado en cada grupo. A pesar de ello, en determinados grupos de alimentos no han podido ser considerados productos con precios elevados. Así, en el caso de la carne roja, no se ha incluido la carne de ovino y, en el caso de los lácteos, se ha incluido solamente leche fresca y leche esterilizada, prescindiéndose de productos de alto valor añadido como el queso y algunos derivados lácteos. En lo que respecta a los pescados, en dicho grupo únicamente se han incluido la merluza y la pescadilla (frescas y congeladas), productos con un precio relativamente bajo.

La segunda columna del cuadro 6 muestra las elasticidades precio directas marshallianas de todos los grupos de productos. Como puede observarse, son todas negativas y significativas. En términos generales, la demanda de los diferentes productos alimenticios es bastante inelástica. Los mayores valores se obtienen para cereales y patatas, carne blanca y huevos, frutas y lácteos, que son aquellos en los que se ha apreciado una mayor dispersión de precios y consumos para las diferentes familias y períodos.

Las elasticidades hicksianas de los distintos productos se recogen en el cuadro 7. Por un lado, el signo negativo de todas las elasticidades directas reafirma el cumplimiento de la condición de negatividad. Por otro, las elasticidades cruzadas nos indican el grado de complementariedad y/o de sustitución existente entre los diferentes grupos de productos. Del total de las elasticidades cruzadas, 91 han resultado positivas y 9 negativas, siendo la mayor parte de ellas significativas. De esta forma, puede concluirse acerca del carácter sustitutivo de la mayor parte de los productos.

Finalmente, los efectos de las distintas variables sociodemográficas incluidas en el análisis se recogen en el cuadro 8. Comenzaremos con

Cuadro 6

ELASTICIDADES GASTO Y PROPIO PRECIO, CALCULADAS EN LOS VALORES MEDIOS

Productos alimentarios	Elasticidad gasto	Elasticidad propio precio
Cereales y patatas	0,882**	-1,05**
Carne roja	1,028**	-0,53**
Carne blanca y huevos	1,086**	-1,05**
Pescados	0,936**	-0,22**
Lácteos	0,910**	-0,75**
Aceites	0,938**	-0,18**
Frutas	1,166**	-0,85**
Hortalizas	1,093**	-0,44**
Azúcar (a)	0,654**	-0,56**
Otros productos (b)	1,286**	-0,39**

(a) Incluye preparados de azúcar y refrescos y bebidas gaseosas.

(b) Incluye bebidas alcohólicas y otros productos alimenticios (café, infusiones,...).

Nota:

** Indica la significatividad de los parámetros al 5%.

* Indica la significatividad de los parámetros al 10%.

Cuadro 7

ELASTICIDADES PRECIO HICKSIANAS DEL MODELO SINTÉTICO,
CALCULADAS EN LOS VALORES MEDIOS

	Cereales y patatas	Carne roja	Carne blanca y huevos	Pescados	Lácteos	Aceites	Frutas	Hortalizas	Azúcar *	Otros ^b product.
Cereales y patatas	-0,90**	0,131**	0,061**	0,049**	0,22**	0,042**	0,23**	0,101**	0,033*	0,025
Carne roja	0,141**	-0,37**	0,107**	-0,0045	0,035	0,0033	0,045	0,049*	0,016	0,061
C. blanca y huevos	0,109**	0,178**	-0,95**	0,157**	0,075**	0,082**	0,203**	0,064**	-0,0058	0,26**
Pescados	0,078**	-0,006	0,139**	-0,12**	-0,08**	-0,0075	-0,022	0,0024	0,043	0,043
Lácteos	0,37**	0,055	0,069**	-0,085**	-0,06**	-0,004	0,169**	-0,03	0,045	0,30**
Aceites	0,114**	0,0084	0,123**	-0,012	-0,0076	-0,12**	-0,058	-0,032	-0,060	0,06
Frutas	0,29**	0,054	0,145**	-0,017	0,129**	-0,027**	-0,70**	0,0049	0,08**	0,38**
Hortalizas	0,211**	0,096*	0,074**	0,003	-0,038	-0,025	0,008	-0,35**	-0,26	0,102
Azúcar (a)	0,126	0,055	-0,01	0,102	0,101	-0,08	0,25**	-0,047	-0,53**	0,23**
Otros productos (b)	0,039*	0,068	0,08**	-0,055**	0,031	-0,026**	0,023	0,016*	0,043	-0,25**

(a) Incluye preparados de azúcar y refrescos y bebidas gaseosas.

(b) Incluye bebidas alcohólicas y otros productos alimenticios (café, infusiones,...).

Nota:

** Indica la significatividad de los parámetros al 5%.

* Indica la significatividad de los parámetros al 10%.

los resultados relativos al tamaño de municipio, los cuales deben interpretarse teniendo en cuenta que el segmento de referencia es aquel que reúne a los municipios con menos de 10.000 habitantes (hábitat rural, en su mayor parte). En términos generales, se observa como los municipios de menor tamaño, categoría de referencia, son los que destinan el mayor porcentaje de gasto sobre el total, a las carnes rojas y a otros productos, mientras que la participación del azúcar es la menor entre los segmentos considerados. Por otra parte, los municipios más grandes, con más de 500.000 habitantes, son los que presentan el mayor porcentaje de gasto en carne blanca y huevos, lácteos y hortalizas, mientras que su porcentaje de gasto destinado a pescados y a frutas es el menor. En términos generales, podemos afirmar que, desde el punto de vista nutricional, son los hogares localizados en las ciudades de menor tamaño (entre 10.000 y 50.000 habitantes) los que ingieren una dieta más equilibrada. En efecto, en estos hogares se destina un mayor porcentaje sobre el total a pescados, aceites, frutas y hortalizas, mientras que la participación de las carnes es la menor entre los diferentes tamaños de municipio considerados. En cualquier caso, como se puede apreciar en el cuadro 8, las diferencias existentes entre los diferentes segmentos no son estadísticamente significativas.

Con respecto al efecto del nivel de estudios del sustentador principal, se observa como aquellos hogares sustentados por una persona que no posee ningún tipo de estudios (grupo de referencia) presentan el mayor porcentaje de gasto destinado a cereales y patatas y a azúcar. No obstante, en términos generales, puede decirse que la educación del sustentador principal afecta de forma negativa al gasto efectuado en cereales y patatas, carne blanca y huevos, mientras que ejerce un efecto positivo en pescados, carnes rojas, lácteos, aceites y frutas, si bien sólo en el caso del pescado el efecto es estadísticamente significativo. Finalmente, se puede decir que el nivel de estudios no está relacionado con la presencia de una dieta más equilibrada, desde el punto de vista de los grupos de alimentos considerados.

Al analizar el efecto provocado por la edad del sustentador principal es necesario tener en cuenta que la variable de referencia se refiere a aquellos hogares cuyo cabeza de familia se encuentra entre 18 y 25 años. Como puede observarse, en general, esta variable es la que permite, en mayor medida, encontrar diferencias significativas. Así, el aumento de la edad del sustentador principal ejerce una influencia positiva sobre la participación de lácteos, aceites, frutas, hortalizas y otros productos, si bien, sólo en el caso de los lácteos el efecto es sig-

Cuadro 8

EFFECTOS DE LAS VARIABLES SOCIO-DEMOGRÁFICAS DE LOS HOGARES

	Cereales y patata	Carne roja	Carne blanca y huevos	Pescados	Lácteos	Aceites	Frutas	Hortalizas	Azúcar*	Otros* product.
EFFECTO DEL TAMAÑO DE MUNICIPIO										
10.001 - 50.000	-0,123	-0,061	-0,068	0,099	-0,007	0,123	0,086	0,068	0,208	-0,325
50.001 - 500.000	0,055	-0,044	-0,061	0,014	-0,044	-0,045	0,101	-0,024	0,136	-0,088
> 500.000	0,045	-0,121	0,029	-0,109	0,030	0,016	-0,04	0,102	0,053	-0,005
EFFECTO DEL NIVEL DE ESTUDIOS DEL SUSTENTADOR PRINCIPAL										
Estudios Primarios	-0,02	-0,05	0,21*	-0,037	-0,02	-0,05	-0,057	0,05	-0,134	0,108
Estudios Medios	-0,29	0,101	0,25	-0,068	-0,013	0,065	-0,062	-0,033	-0,094	0,144
Estudios Superiores	-0,56**	0,114	-0,38**	0,51**	0,22	0,127	0,25	-0,137	-0,11	-0,034
EFFECTO DE LA EDAD DEL SUSTENTADOR PRINCIPAL DE LA FAMILIA										
De 26 a 45 años	-0,02	0,018	-0,56**	-0,26**	0,74**	0,11	0,38	-0,01	-0,84**	0,44**
De 46 a 65 años	0,013	-0,09	-0,66**	-0,08**	0,60	0,36	0,55	-0,09	-0,68**	0,077
Más de 65 años	-0,017	-0,19	-0,71**	-0,13	0,61	0,19	0,67	0,23	-0,81**	0,157
EFFECTO DEL PORCENTAJE DE MIEMBROS DE MÁS DE TRECE AÑOS										
25% - 50%	-0,44**	0,29**	0,007	0,073	0,17	0,22	-0,16	-0,27	0,39**	-0,28
50% - 75%	0,056	-0,05	-0,194	0,131	-0,07	0,001	0,13	-0,016	-0,26	0,272
Más del 75%	-0,16	0,15	-0,147	-0,11	-0,32**	-0,08	0,24	0,31	0,15	-0,033

(a) Incluye preparados de azúcar y refrescos y bebidas gaseosas.

(b) Incluye bebidas alcohólicas y otros productos alimenticios (café, infusiones,...).

Nota:

** Indica la significatividad de los parámetros al 5%.

* Indica la significatividad de los parámetros al 10%.

nificativo. En sentido contrario, conforme aumenta la edad disminuye significativamente la participación de productos que están relacionados con problemas de salud, como son las carnes blancas y los huevos (principales responsables del colesterol) y el azúcar (asociado a determinadas enfermedades cardiovasculares).

Por último, en relación con el efecto de la edad de los miembros que componen la unidad familiar, se observa como, aunque las diferencias no son estadísticamente significativas en la mayor parte de los casos, los hogares constituidos por un menor porcentaje de miembros mayores de trece años tienden a destinar un porcentaje significativamente mayor de su gasto al consumo de carne roja, lácteos y azúcar, mientras que el porcentaje destinado a cereales disminuye. Como acabamos de mencionar, para el resto de productos, las diferencias no son significativas, no pudiéndose obtener una conclusión clara, desde el punto de vista de la dieta, sobre la incidencia de la composición del hogar sobre la misma.

7. CONCLUSIONES

En este trabajo se han analizado las reacciones de los consumidores de alimentos en España, desde el punto de vista de las cantidades de producto adquiridas, ante cambios en las variables económicas tradicionales, renta y precios, y en ciertas variables demográficas y socioeconómicas. Desde el punto de vista metodológico, el trabajo se ha orientado a la estimación de sistemas de demanda diferenciados utilizando datos de panel. En este sentido, este trabajo constituye una de las primeras aplicaciones en España con relación a la demanda de alimentos que cubre simultáneamente dos importantes aspectos: la no-imposición *a priori* de una forma funcional, por un lado, y la utilización de un panel completo de datos, por otro.

Con el fin de seleccionar la mejor forma funcional, se ha definido un sistema envolvente de tal forma que, introduciendo ciertas restricciones sobre los parámetros del mismo, se pudiesen obtener cuatro de los modelos más utilizados en la literatura. Desde este punto de vista, uno de los resultados más interesantes obtenidos es que ninguno de los cuatro modelos se ajusta mejor que el modelo sintético, que también puede ser considerado como un sistema de demanda. Dicho en otras palabras, en este trabajo, la elección de una forma funcional *a priori* hubiese supuesto la pérdida de un importante volumen de información. En relación con la utilización de datos de panel, se ha partido de la especificación de un modelo de efectos aleatorios (los efectos individuales no observables afectan al componente estocástico del modelo). En cualquier caso, al modelo estimado se le ha sometido a una batería de contrastes, obteniéndose que dicho modelo había sido especificado correctamente.

Los resultados obtenidos en términos de elasticidades son comparables con los obtenidos en otros trabajos, fundamentalmente con los que han utilizado datos de corte trasversal, si bien se aprecian ciertas diferencias en la magnitud de ciertas elasticidades gasto y precio. La agrupación que se ha hecho de los alimentos y las propias características de la fuente de información utilizada han sido, sin duda, los principales condicionantes de este trabajo, por lo que es difícil comparar resultados.

En cualquier caso, creemos que la estrategia utilizada en nuestro trabajo es adecuada y proporciona mayor flexibilidad y consistencia a la estimación de sistemas de demanda, si bien no está exenta de cierta complejidad. Hasta qué punto esa mayor complejidad permite «mejorar» los resultados obtenidos anteriormente es una cuestión abierta, que se aleja del objetivo de este trabajo y que cada lector

debería valorar. En cualquier caso, lo que sí debería ser una constante en el trabajo aplicado es que cualquier modelo que se estime debería ser validado tanto desde el punto de vista económico como estadístico, tal como hemos realizado en este estudio.

El trabajo únicamente ha pretendido iniciar una cierta línea de investigación complementando dos aspectos que ya habían sido tratados en la literatura. Evidentemente, este trabajo puede mejorarse en un futuro, tanto desde el punto de vista metodológico como aplicado. En el primer caso, la estimación y contrastación en sistemas multiecuacionales utilizando datos de panel es un campo que está siendo objeto de investigación teórica, sobre todo en lo que se refiere al desarrollo de contrastes de especificación en versión multivariante. En cuanto estas investigaciones se divulguen podrían ser objeto de utilización en trabajos aplicados futuros. Desde el punto de vista aplicado, creemos que el futuro pasa por la introducción de variables diferentes de las tradicionales renta, precios y características sociodemográficas. Aspectos como la calidad de la dieta, la información sobre la relación dieta-salud, la información sobre escándalos o crisis alimentarias son factores que condicionan la demanda de, al menos, determinados tipos de alimentos y que, por tanto, deberían ser susceptibles de introducirse en futuras investigaciones en línea con algunos trabajos que han ido apareciendo recientemente.

BIBLIOGRAFÍA

- ANGULO, A. M.; GIL, J. M.; DHEHIBI, B. y MUR, J. (2002). «Town Size and the Consumer Behaviour of Spanish Households: a Panel Data Approach». *Applied Economics*, 34: pp. 503-507.
- ANGULO, A. M.; GIL, J. M. y GRACIA, A. (2001). «Calorie Intake and Income Elasticities in EU countries: A Convergence Analysis Using Cointegration». *Papers in Regional Science*, 80(2): pp. 165-187.
- ARELLANO, M. y BOVER, O. (1990): «La econometría de datos de panel». *Investigaciones Económicas*, 14: pp. 3-45.
- ARELLANO, M. (2003): *Panel data econometrics*. Oxford University Press.: pp. 231.
- BALTAGI, B. H. (1980): «On Seemingly Unrelated Regressions With Error Components». *Econometrica*, 48: pp. 1.547-1.551.
- BALTAGI, B. H. (1999): *Econometric Analysis of Panel Data*. John Wiley & Sons Press.
- BARTEN, A. P. (1993): «Consumer Allocation Models: Choice of Functional Form». *Empirical Economics*, 18: pp. 129-158.
- BEN KAABIA, M.; ANGULO, A. M. y GIL, J. M. (2001): «Health Information and the Demand for Meat in Spain». *European Review of Agricultural Economics*, 28(4): pp. 499-517.

- BERA, A. K.; SOSA-ESCUADERO, W. y YOON, M. (2001): «Tests for the Error Component Model in the Presence of Local Misspecification». *Journal of Econometrics*, 101: pp. 1-23.
- BEWLEY, R. y YOUNG, T. (1987): «Applying Multinomial Extension of the Linear Logit Model to Meat Expenditure Data». *American Journal of Agricultural Economics*, 69: pp. 151-157.
- CHUNG, C. F. (1994): «A Cross-Section Demand Analysis of Spanish Provincial Food Consumption». *American Journal of Agricultural Economics*, 76: pp. 513-521.
- COWLING, R. G. D. y RAYNOR, A. J. (1970): «Price, Quality and Market Shares». *Journal of Political Economy*, 78: pp. 1.292-1.309.
- DEATON, A. (1989): «Quality, Quantity and Spatial Variation of Price». *The American Economic Review*, 78: pp. 418-430.
- DEATON, A. y MUELLBAUER, J. (1980): «An Almost Ideal Demand System». *The American Economic Review*, 70: pp. 312-326.
- GAO, X. M.; WAILES, E. J. y CRAMER, G. L. (1995): «Double Hurdle Model With Bivariate Normal Errors: An Application to U.S. Rice Demand». *Journal of Agricultural and Applied Economics*, 72(2): pp. 363-376.
- GAO, X. M.; RICHARDS, T. J. y KAGAN, A. (1997): «A Latent Variable Model of Consumer Taste Determination and Taste Change for Complex Carbohydrates». *Applied Economics*, 29: pp. 1.643-1.654.
- GRACIA, A. y ALBISU, L. M. (1998): «The Demand for Meat and Fish in Spain: Urban and Rural Areas». *Agricultural Economics*, 19 (3): pp. 359-366.
- GRAVELLE, H. y REES, R. (1992): *Microeconomics*. 2ª Edición. Editorial Longman, New York.
- GRACIA, A.; GIL, J. M. y ANGULO, A. M. (1998): «Spanish Food Demand: a Dynamic Approach». *Applied Economics*, 30: pp. 1.399-1.405.
- KELLER, W. J. y VAN DRIEL, J. (1985): «Differential Consumer Demand Systems». *European Economic Review*, 27: pp. 375-390.
- LAAJIMI, A. y ALBISU, L. M. (1997): «La Demande de Viandes et de Poissons en Espagne: Une Analyse Micro-Economique». *Cahiers d'Economie et Sociologie Rurales*, 42-43: pp. 71-91.
- LAMBÁN, J. (1996): *Sistemas diferenciales de demanda inversa. Aplicación a la demanda de carne y pescado en España*. Tesis Doctoral. Universidad Politécnica de Madrid.
- MANRIQUE, J. y JENSEN, H. H. (1997): «Spanish Household Demand for Convenience Meat Products». *Agribusiness*, 13(6): pp. 579-586.
- MANRIQUE, J. y JENSEN, H. H. (1998): «Working Women and Expenditures on Food Away-From-Home and At-Home in Spain». *Journal of Agricultural Economics*, 49(3): pp.321-333.
- MOLINA, J. A. (1994): «Food Demand in Spain: An application of the Almost Ideal System». *Journal of Agricultural Economics*, 45(2): pp. 252-258.

- NEVES, P. (1987): *An Analysis of Consumer Demand in Portugal, 1958-1981. Memoire de Maitrise en Sciences Economiques*. Université Catholique de Louvain, Louvain-la-Neuve, Bélgica.
- PRUCHA, I. R. (1984): «On the Asymptotic Efficiency of Feasible Aitken Estimators for Seemingly Unrelated Regression Models With Error Components». *Econometrica*, 52: pp. 203-207.
- THEIL, H. (1965): «The Information Approach to Demand Analysis». *Econometrica*, 33: pp. 67-87.
- WALLACE, T. D. y HUSSAIN, A. (1969): «The Use of Error Components Models in Combining Cross-section and Time-series Data». *Econometrica*, 37: pp. 55-72.
- WHITE, H. (1980): «A Heteroskedasticity Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity». *Econometrica*, 48: pp. 817-838.
- WORKING, H. (1943): «Statistical Laws of Family Expenditure». *Journal of American Statistical Association*, 38: pp. 43-56.

RESUMEN

Datos de panel y demanda de alimentos en España

Hasta la fecha, la mayor parte de los trabajos que han estudiado la demanda de alimentos en España han considerado únicamente datos de series temporales o de sección cruzada. El principal objetivo de este estudio consiste en considerar conjuntamente las dos dimensiones para abordar este problema, tratando de ofrecer al lector unas pautas no sólo para estimar sistemas de demanda con un panel de datos sino también para contrastar su correcta especificación. Una segunda característica de este trabajo consiste en definir un esquema general a partir del cual se pueda especificar una forma funcional del sistema de demanda que mejor se ajuste a la información disponible, en vez de seleccionar a priori una forma funcional determinada. Los resultados empíricos obtenidos están en la línea de trabajos anteriores, si bien las elasticidades gasto y precio directas calculadas son ligeramente superiores, sobre todo en el caso de los productos perecederos. Al analizar los efectos de las características sociodemográficas de los individuos, la edad y el porcentaje de miembros de la familia menores de 13 años son los principales determinantes del consumo de carnes, pescados y azúcar, experimentándose un descenso en las cantidades consumidas de estos productos conformen aumentan dichas variables.

PALABRAS CLAVES: Datos de panel, selección de la forma funcional, sistemas de demanda de alimentos, España.

SUMMARY

Panel data and Spanish demand for food

The demand for food in Spain has been traditionally studied only considering time series or cross section data. This paper aims to analyse food demand jointly considering both data dimensions. It also provides some guidance on how estimate demand systems using a panel data set and on how to carry out misspecification tests. The second main characteristics of this paper is that instead of selecting an a priori functional form, we provide a flexible approach from which the functional form which fits better the data can be selected. Empirical results are consistent with those found in previous literature although expenditure and price elasticities are slightly higher for perishable products. As far as the socio-demographic characteristics of households, the age and the percentage of household members below 13 years are main determinants of meat, fish and sugar consumption. Consumed quantities of these products decrease as the two considered variables increase.

KEYWORDS: Panel data, selection of the functional form, food demand systems, Spain.