

Documento de Trabajo 91-04  
Julio 1991

Departamento de Economía  
Universidad Carlos III de Madrid  
Calle Madrid, 126  
28903 Getafe (España)

EFFECTOS BLOQUE EN LA ESTIMACION DE CURVAS DE  
ENGEL CON DATOS DE LA ENCUESTA DE  
PRESUPUESTOS FAMILIARES 1980-81

María Jesús San Segundo\*

Resumen

---

En este trabajo se analizan los efectos que el diseño en bloques de las Encuestas de Presupuestos Familiares puede tener sobre la estimación de Curvas de Engel. Al realizar tests de efectos fijos y de efectos aleatorios se detecta la presencia de efectos bloque. Sin embargo, se comprueba que los parámetros obtenidos por mínimos cuadrados proporcionan una aproximación aceptable a las estimaciones consistentes (de efectos fijos), especialmente si se incluye en la regresión variables que miden características comunes a los hogares de un bloque.

---

Palabras clave: Efectos fijos, Curvas de Engel.

\* María Jesús San Segundo, Universidad del País Vasco y Universidad Carlos III de Madrid. El trabajo original en el que se basa este artículo fue realizado mientras la autora era miembro del Instituto de Economía Pública (U.P.V.). En primer lugar, se agradece la financiación del Instituto Nacional de Estadística para la realización de este proyecto. La colaboración de José Antonio Bartolomé en las estimaciones se agradece muy sinceramente. Los valiosos comentarios y sugerencias de Javier Ruiz-Castillo han mejorado notablemente este trabajo. Por último, se agradece la generosa ayuda de Javier Prieto en la edición de este artículo.

# 1 INTRODUCCION

En este trabajo se estudian los efectos que algunas características del diseño muestral de la Encuesta de Presupuestos Familiares 1980/81 (EPF) pueden tener sobre la estimación de modelos de regresión lineal que utilizan datos de esta Encuesta.

El diseño estratificado de una muestra como la EPF presenta ventajas indudables de economía, y para garantizar la representatividad de la muestra por provincias (INE (1983, p.48)); pero, Deaton and Irish (1983) han puesto de manifiesto los problemas que el muestreo en bloques de observaciones "próximas geográficamente" puede crear para la estimación de modelos de regresión lineal. En la próxima sección de este trabajo se describen estos problemas y se argumenta que se pueden presentar en la estimación de Curvas de Engel (curvas de demanda con precios fijos) con datos de la EPF. En la sección 3 se describe la muestra utilizada en este trabajo. En las secciones 4 y 5 se contrasta la existencia de los efectos bloque, como efectos fijos y como efectos aleatorios, en las Curvas de Engel estimadas con datos de la EPF 80-81. Estos contrastes ponen de manifiesto que los problemas detectados por Deaton y Irish pueden desaparecer al incluir en el modelo variables que miden características comunes a las observaciones de un bloque. Aún en el caso de no estar interesados en la relación de estas variables "bloque" con la composición del consumo, parece recomendable su inclusión para mitigar el impacto del diseño muestral sobre nuestras estimaciones.

## 2 EFECTOS BLOQUE EN CURVAS DE ENGEL

### 2.1 Efectos Bloque en Modelos de Regresión Aplicados a la EPF 80/81

La Encuesta de Presupuestos Familiares no se diseñó como una muestra aleatoria simple de los hogares españoles en 1980. "El diseño muestral de la EPF/80 se enmarca en el diseño general realizado por el INE para encuestas dirigidas a la población y que recibe el nombre de Encuesta General de Población: (INE (1983), p.48)".

La EPF es una muestra aleatoria estratificada en dos etapas. En la primera etapa se eligieron "secciones censales" dentro de cada estrato (por provincias, por tipo de municipio) con una probabilidad proporcional a su peso demográfico en el estrato.

En la segunda etapa se seleccionaron viviendas familiares dentro de cada sección,

mediante un muestreo sistemático con arranque aleatorio (INE (1983, p.49)).

El tamaño muestral fue de 3.412 secciones y la muestra efectiva se refiere a 24.000 hogares. Veamos a continuación los efectos que esta selección de hogares por secciones o bloques puede tener en la estimación de modelos de regresión con datos de la EPF.

Si se denota con el índice  $i$  a los bloques de observaciones ( $i = 1, 2, \dots, B$ ) y se ordenan las observaciones por bloques de forma que todas las pertenecientes a un mismo bloque sean consecutivas, con  $n_i$  observaciones en el bloque  $i$  se tiene que  $N = \sum_{i=1}^B n_i$ .

Supongamos que el modelo de regresión lineal tiene la siguiente forma

$$Y_{ji} = X_{ji}\beta + U_{ji} \quad j = 1, \dots, n_i \quad i = 1, \dots, B, \quad (1)$$

donde  $Y$  representa a la variable dependiente,  $X$  es una matriz de  $K$  variables independientes y  $U$  es una perturbación aleatoria que se supone normal, de media cero.

Una hipótesis habitual en los modelos que utilizan datos transversales (de sección cruzada) es la independencia de las perturbaciones  $U_{ji}$ . En nuestro caso, se sabe que las observaciones están agrupadas por bloques geográficos, y se puede pensar que aunque sean independientes entre los diferentes bloques, pudieran no serlo dentro de los bloques. En el caso más sencillo se puede suponer que la correlación es idéntica en los distintos bloques, es decir:

$$\begin{aligned} E(U_{ji}) &= 0, & E(U_{ji}U_{ji}) &= \sigma^2 \\ E(U_{ji}U_{ki}) &= \sigma^2\rho, & j \neq k, \quad \forall i & \\ E(U_{ji}U_{ks}) &= 0, & i \neq s & \end{aligned} \quad (2)$$

o en notación matricial.  $E(uu') = \sigma^2V$ , donde  $V = \bigotimes_{i=1}^B V_i$  con  $V_i = (1-\rho)I_{n_i} + \rho E_{n_i}$ , donde  $I_{n_i}$  es la matriz identidad  $n_i \times n_i$ , y  $E_{n_i}$  es una matriz  $n_i \times n_i$  con todos los elementos iguales a la unidad.

Si el diseño muestral en bloques hace que  $\rho$  sea distinto de cero, se tiene que la estimación del modelo de regresión (1) por mínimos cuadrados ordinarios (MCO) produce estimaciones consistentes, pero no eficientes, de los parámetros  $\beta$  del modelo. En las próximas secciones se discute cómo contrastar este modelo y qué métodos de estimación utilizar cuando  $\rho \neq 0$ .

La relación del modelo presentado con el de "efectos aleatorios" en los trabajos con datos longitudinales es clara (véase Hsiao (198)). Si se denota por  $\alpha$  los efectos bloque se tiene que

$$Y_{ji} = X_{ji}\beta + U_{ji} = X_{ji}\beta + \alpha_i + \varepsilon_{ji} \quad (3)$$

$$\text{con } \begin{cases} E(U_{ji}U_{ji}) = \sigma_{\alpha}^2 + \sigma_{\epsilon}^2 = \sigma^2 \\ E(U_{ji}U_{ki}) = \sigma_{\alpha}^2 = \rho\sigma^2, & j \neq k. \\ E(U_{ji}U_{ks}) = 0, & i \neq s. \end{cases}$$

Ahora se puede interpretar  $\sigma^2 = \sigma_{\alpha}^2 + \sigma_{\epsilon}^2$ , y  $\rho = 1 - \frac{\sigma_{\epsilon}^2}{\sigma^2} = \frac{\sigma_{\alpha}^2}{\sigma^2}$ .

Esta analogía de nuestro problema con los que aparecen en los modelos de datos de panel lleva inmediatamente a plantear si los efectos comunes a los bloques ( $\alpha$ ) pueden ser "efectos fijos" (en la terminología de Hausman (1978)), no independientes de las variables explicativas. Es bien conocido que, si  $E[\alpha_i/x_{ji}] \neq 0$ , entonces los estimadores de mínimos cuadrados (ordinarios o generalizados) de (1) no son consistentes<sup>1</sup>. En la próxima subsección se discute la relevancia de este problema para la estimación de Curvas de Engel.

## 2.2 Estimaciones de Curvas de Engel con Efectos Bloque

Una de las aplicaciones fundamentales de las muestras transversales de gasto familiar como la EPF es la estimación de curvas de demanda con precios fijos, o curvas de Engel. Estas curvas relacionan el gasto de los hogares en diversos grupos de bienes con la renta familiar, el número y la edad de los miembros del hogar, y otras características socioeconómicas de la unidad de gasto.

Siguiendo a Deaton y Muellbauer (1986) especificamos una forma funcional flexible:

$$W_{ji} = \beta_0 + \beta_1 \ln(X/m)_{ji} + \beta_2 [\ln(X/m)_{ji}]^2 + \sum_k \gamma_k M_{kji} + Z_{1ji} \delta_1 + Z_{2ji} \delta_2 + U_{ji} \quad (3)$$

donde  $W$  representa la proporción del gasto total ( $X$ ) destinado a algún bien (por ej. a alimentación),  $m$  es el número de miembros del hogar  $j$ ,  $M_k$  es el n° de miembros del hogar del grupo demográfico  $k$ ,  $Z_1$  y  $Z_2$  recogen otras características observadas de la unidad de consumo, y  $U$  es la perturbación aleatoria<sup>2</sup>.

<sup>1</sup>Ecuaciones como (2) podrían contener efectos fijos (o aleatorios) individuales  $\phi_{ji}$  que generarían problemas similares a los introducidos por  $\alpha_i$  para estimar estas ecuaciones. Si se dispusiera de datos de panel ( $Y_{jit}, X_{jit}, t = 1, \dots, T$ ) se podrían tener en cuenta estos efectos; con datos de corte transversal hay que suponer que no hay efectos individuales  $\phi_{ji}$  en el término de error  $U_{ji}$ .

<sup>2</sup>Deaton, Ruiz Castillo y Thomas (1986) consideran diferentes especificaciones de las curvas de Engel, para los datos de esta encuesta, y encuentran evidencia en favor de la especificación de Working-Leser que se utiliza en este trabajo. Una discusión de especificación similar, con datos Sri Lanka, aparece en Deaton (1981).

La estimación de curvas de Engel es un ejercicio de economía aplicada que tiene interés si se desea predecir la evolución del consumo. Así, la estimación de curvas de demanda como (3) permite conocer la relación entre la renta de los hogares y su consumo de distintos bienes; permite, por ejemplo, estimar los posibles efectos sobre el consumo de diferentes políticas distributivas (Abadía (1985)), a falta de datos experimentales.

La estimación de curvas de Engel también tiene aplicaciones en el análisis de bienestar. El trabajo de Deaton, Ruiz Castillo y Thomas (1985) ilustra la utilización de curvas de Engel, obtenidas a partir de los datos de la EPF 80-81, para estimar el coste de los niños en España. Los artículos de Seneca y Tausing (1971) y Bennett (1979) ilustran como utilizar estas estimaciones para evaluar las compensaciones familiares del sistema impositivo y de la Seguridad Social, respectivamente.

Si tiene interés estimar consistente y eficientemente los parámetros de estas curvas de demanda (3), según lo expuesto en la sección anterior, hay que analizar la posible presencia de efectos bloque, ya que las estimaciones de M.C.O. pueden ser ineficientes, si hay efectos aleatorios, e incluso inconsistentes, si los efectos son fijos.

Se espera que efectos bloque puedan estar presentes en curvas de Engel como (3) porque parece lógico suponer que nuestra especificación de estas ecuaciones no incluye todas las variables que afectan al consumo de diversos bienes, y las variables no incluidas explícitamente en cada ecuación pasan a formar parte del término de error  $U_{ji}$ . Si las observaciones que pertenecen a un mismo bloque tienen en común alguna de estas características de los hogares que no se observan y que afectan al consumo, aparece el problema de la existencia de efectos bloque.

Nos proponemos por lo tanto estudiar si el diseño de la EPF crea problema de efectos bloque para la estimación de curvas de Engel, y creemos que los resultados de nuestro trabajo serán de interés para los usuarios de la Encuesta de Presupuestos Familiares.

### **3 DESCRIPCION DE LA MUESTRA. ALGUNOS PROBLEMAS ECONOMETRICOS**

En este trabajo se utilizan datos de la Encuesta de Presupuestos Familiares 1980-81 referentes a 21.705 hogares. Se han excluido las observaciones correspondientes a Ceuta y Melilla (300 aproximadamente), y las correspondientes a hogares de un sólo miembro (1.921). Es bien conocido que el comportamiento de estas unidades de consumo es muy diferente al del resto de la población, y no parece adecuado utilizar todos los datos

conjuntamente<sup>3</sup>. Por último, se excluyen los bloques que cuentan con un solo hogar en la muestra (82) y no pueden ser utilizados en el análisis.

En la TABLA 1 aparecen las estadísticas descriptivas correspondientes a las variables que se emplean en este estudio. Para tener en cuenta la estratificación de la muestra las observaciones en cada estrato son ponderadas por el cociente de la proporción poblacional del estrato entre la proporción muestral. Aunque esta práctica es habitual al presentar estadísticas descriptivas, es objeto de controversia cuando se aplica a la estimación de un modelo de regresión. Si la estratificación se basa en variables exógenas del modelo entonces los estimadores de mínimos cuadrados ordinarios (M.C.O.) son los mejores estimadores lineales insesgados, bajo los supuestos habituales de perturbaciones homocedásticas (véase DuMouchel y Duncan (1983)).

Deaton, Ruiz Castillo y Thomas (1985) argumentan que, en caso de duda sobre la validez de un modelo uniforme para los diferentes estratos, merece la pena arriesgar una pérdida de eficiencia utilizando los estimadores ponderados, para evitar que los parámetros estén determinados por el diseño muestral. En este trabajo se presentan estimaciones de M.C.O. y de mínimos cuadrados ponderados para obviar esta controversia.

En todos los casos se utilizan los estimadores de la matriz de varianzas y covarianzas de White (1980), que son consistentes bajo formas generales de heteroscedasticidad. Como es habitual en estos estudios se encuentra evidencia de heterocedasticidad (véanse los estadísticos de Breusch-Pagan (1979) en las tablas), y los errores standard crecen entre un 20 y un 30% al aplicar el método de White. Aunque Mackinnon y White (1984) muestran que el estimador "jackknife" es superior al de White (1980), Deaton, Ruiz Castillo y Thomas (1986) comprueban que los dos estimadores proporcionan la misma matriz de varianzas y covarianzas, para las estimaciones de mínimos cuadrados de nuestras curvas de Engel.

Otra característica de los datos que causa problemas econométricos es la presencia de observaciones (hogares) con gastos cero en algunos grupos de bienes. Si estos "ceros" se deben a que los bienes no se compran frecuentemente (por ej.: bienes duraderos), la estimación por mínimos cuadrados es consistente. Sin embargo, si se trata de bienes que algunas familias deciden no comprar (por ej.: tabaco, sanidad privada, ...) se de-

---

<sup>3</sup>Los datos de la Encuesta Permanente de Consumo (INE, 1977-1983) permiten comprobar esta diferencia entre el comportamiento del consumo de los hogares de 1 miembro y los de dos o más miembros. En el cálculo de escalas de equivalencia también es práctica habitual eliminar estas observaciones de la muestra.

**TABLA 1**  
**ESTADISTICAS DESCRIPTIVAS<sup>4</sup>**

GASTOS ( $W_i$ )	MEDIA	DESVIACION STANDARD
Participaciones en el GASTO TOTAL de		
Alimentación	0.351	0.15
Alcohol y Tabaco	0.028	0.03
Vestido y Calzado	0.083	0.06
Bienes del Hogar y Servicios	0.069	0.06
Vivienda	0.187	0.12
Medicina y Farmacia	0.021	0.04
Transportes y Comunicaciones	0.108	0.12
Educación	0.016	0.03
Gasto per cápita (ln)		
$\ln(x/n)$	12.24	0.61
Número de miembros del hogar		
$n_1$ : entre 0 y 4 años	0.231	0.50
$n_2$ : entre 5 y 8 años	0.374	0.66
$n_3$ : entre 9 y 13 años	0.394	0.70
$n_4$ : entre 14 y 17 años	0.301	0.59
$na_1$ : entre 18 y 23 años	0.396	0.70
$na_2$ : entre 24 y 60 años	1.714	0.83
$n_0$ : Mayores de 60 años	0.520	0.79
Nivel educativo del cabeza de familia		
$ED_1$ : Analfabeto	0.06	0.23
$ED_2$ : Sin estudios	0.251	0.43
$ED_4$ : Ed. Secundaria	0.134	0.34
$ED_5$ : Ed. Superior	0.068	0.25

<sup>4</sup> NOTA: Las regresiones incluirán las siguientes variables adicionales:

- 5 variables ficticias sobre la situación laboral del cabeza de familia
- 4 variables ficticias de tamaño de municipio
- 16 variables ficticias para las Comunidades Autónomas
- Variables que recogen las edades del cabeza de familia y de su esposa
- variables ficticias sobre el trimestre y la semana en que se realiza la entrevista.

bieran utilizar métodos de estimación no lineal que requieran fuertes hipótesis sobre la distribución de las variables (véanse los tests de Newey (1987)), son muy costosos para nuestro tamaño muestral, y no están suficientemente explorados para varios grupos de bienes que pueden tener gasto cero (véase Wales y Woodland (1983) para un ejemplo con 3 bienes y una muestra más reducida que la nuestra). En este trabajo se sigue la opción habitual de utilizar las estimaciones de mínimos cuadrados y se confía que sean consistentes.

## 4 CONTRASTES DE LA PRESENCIA DE EFECTOS BLOQUES

### 4.1 Efectos fijos. Descripción del problema

Reformulando las ecuaciones de demanda se incluyen explícitamente los efectos bloque ( $\alpha$ ). Así, la proporción del gasto destinada a un bien ( $W$ ) es una función

$$\begin{aligned} W_{ji} &= \beta_0 + \beta_1 \ln\left[\frac{X}{m_{ji}}\right] + \beta_2 \left[\ln\left[\frac{X}{m_{ji}}\right]\right]^2 + \sum_k \gamma_k M_{kji} + Z_{1ji}\delta_1 + Z_{2i}\delta_2 + \alpha_i + \varepsilon_{ji} \\ &= S_{ji}B + Z_{2i}\delta_2 + \alpha_i + \varepsilon_{ji} \end{aligned} \quad (4)$$

donde  $Z_{2i}$  recoge las características comunes a los hogares de un mismo bloque ( $i$ ),  $S_{ji}$  contiene las demás características observables del hogar  $j$ ; y  $\varepsilon_{ji}$  son perturbaciones aleatorias independientes pero no idénticamente distribuidas.

Para contrastar la hipótesis nula,  $H_0 : E(\alpha_i/S_{ji}, Z_{2i}) = 0$ , de no existencia de efectos fijos, se utiliza un test de especificación de Hausman (1978). En este test se compara el estimador de efectos fijos (o "within groups" o dentro de los bloques), que es consistente se satisfaga o no la hipótesis nula, y el estimador entre grupos ("between groups" o entre bloques), que sólo es consistente si la hipótesis nula es cierta.

Para derivar estos estimadores se pueden tomar medias de las variables, para cada bloque, en la ecuación (4), y se obtiene:

$$W_{.i} = S_{.i}B + Z_{2i}\gamma_2 + \alpha_i + \varepsilon_{.i} \quad (5)$$

Si se considera la diferencia (4) - (5), se obtiene la ecuación en desviaciones sobre la media del bloque:

$$W_{ji} - W_{.i} = (S_{ji} - S_{.i})B + \varepsilon_{ji} - \varepsilon_{.i} \quad (6)$$

Aplicando MCO a esta ecuación se obtienen los estimadores “within groups” ( $\hat{B}_W$ ) cuya consistencia no se ve afectada por la presencia de efectos fijos en (4). Aplicando MCO a (5) se obtienen los estimadores “between groups” ( $\hat{B}_B$ ) que son consistentes si la hipótesis nula es cierta.

Aunque el test de Hausman de efectos fijos debiera comparar el estimador de mínimos cuadrados generalizados (el más eficiente bajo  $H_0$ ) con el estimador de efectos fijos (consistente sea  $H_0$  cierta o no), Hausman-Taylor (1981) demuestran que bajo la hipótesis  $H_0$  el estadístico  $Q = (\hat{B}_W - \hat{B}_B)' [cov(\hat{B}_W) + cov(\hat{B}_B)]^{-1} (\hat{B}_W - \hat{B}_B)$  es una  $\chi^2$ , idéntica a la que compara  $\hat{B}_W$  y el estimador de mínimos cuadrados de (4).

Por comodidad, se comparan  $\hat{B}_W$  y  $\hat{B}_B$ , lo que no requiere estimar  $\rho$  (resolver el problema planteado en 2.a). Y si este test rechaza la hipótesis nula de ortogonalidad, hay que utilizar los estimadores de efectos fijos en (6) en lugar de los de MCO en (4), para garantizar la consistencia de las estimaciones de  $\beta$ . Si se acepta la hipótesis nula, aún hay que contrastar la importancia de los efectos bloque como efectos aleatorios.

## 4.2 Estimaciones de efectos fijos de curvas de Engel

En las TABLAS 2 y 3 se presentan las estimaciones de efectos fijos (ecuaciones (6)) de las Curvas de Engel correspondientes a ocho grupos de bienes<sup>4</sup>. De la comparación de las dos TABLAS se deduce que la estratificación no tiene una influencia significativa en la estimación de los parámetros.

Los efectos de cambios en el gastos total ( $x$ ) sobre la composición del consumo se resumen en la TABLA 4 que recoge las ELASTICIDADES-RENTA de los diferente bienes (evaluadas en las medias muestrales). Se observa que si crece el gasto total disminuye la proporción destinada a alimentación, alcohol y tabaco, y a bienes del hogar; mientras que aumenta la participación en el gasto de los demás grupos de bienes.

---

<sup>4</sup>El noveno grupo (OTROS BIENES) no aparece en las TABLAS ya que se cumple la condición de agregación; los gastos en los nueve grupos de bienes agregan al gasto total ( $x$ ): es decir  $\sum_{i=1}^9 W_i = 1$ . Por lo tanto,  $\sum_{i=1}^9 \beta_{0i} = 1$ ,  $\sum_{i=1}^9 \beta_{1i} = \sum_{i=1}^9 \beta_{2i} = \sum_{i=1}^9 \delta_i = 0$ .

**TABLA 2**  
**CURVAS DE ENGEL**  
**ESTIMACIONES DE EFECTOS FIJOS PONDERADOS<sup>1</sup>**

	Alimen.	Alcohol y Tabaco	Vestido y Calzado	Hogar	Vivienda	Medicina y Farmacia	Transport. y Comunicac.	Educación
$\ln(x/n)$	13.93 (2.1)	1.33 (0.8)	22.02 (10.6)	- 55.73 (- 9.2)	2.61 (0.9)	0.56 (0.2)	-7.75 (-1.7)	9.89 (9.7)
$(\ln(x/n))^2$	-1.09 (-4.1)	-0.07 (-1.1)	- 0.86 (-10.1)	2.19 (8.8)	-0.008 (- 0.0)	-0.01 (-0.1)	0.62 (3.4)	-0.40 (-9.8)
$n_1$	-0.52 (-2.6)	-0.23 (-4.6)	0.18 (1.6)	-0.89 (-4.5)	0.82 (6.2)	0.31 (4.)	0.74 (3.5)	-0.15 (-2.7)
$n_2$	-0.49 (-3.2)	-0.22 (-5.7)	0.32 (3.8)	-1.41 (-9.5)	0.13 (1.4)	-0.01 (-0.3)	1.10 (6.8)	0.66 (13.9)
$n_3$	-0.03 (-0.2)	-0.31 (-9.)	0.44 (5.4)	-1.20 (-9.)	0.29 (3.6)	0.008 (0.1)	0.33 (2.4)	0.70 (15.2)
$n_4$	-1.10 (-6.9)	-0.13 (-3.5)	0.52 (5.7)	-1.38 (-8.7)	0.02 (0.2)	-0.03 (-0.6)	0.10 (0.6)	0.87 (15.5)
$na_1$	-2.35 (-16.7)	0.09 (2.9)	0.29 (3.5)	-1.70 (-12.)	-0.09 (-1.0)	-0.11 (-2.2)	1.26 (8.8)	0.34 (7.)
$na_2$	-1.43 (-8.3)	0.14 (3.)	0.03 (0.3)	-1.95 (-11.1)	-0.62 (-6.1)	-0.007 (-0.1)	1.86 (11.2)	0.19 (4.2)
$n_0$	-0.84 (-3.9)	-0.003 (-0.0)	-0.18 (-1.5)	-0.36 (-1.6)	-0.36 (-3.1)	0.06 (0.7)	0.98 (4.8)	0.04 (0.6)
$ED_1$	1.51 (3.1)	0.48 (3.2)	-0.46 (-1.9)	-0.78 (-1.7)	0.89 (3.6)	-0.04 (-0.2)	-0.85 (-2.1)	-0.14 (-1.4)
$ED_2$	0.80 (2.8)	0.21 (2.7)	-0.33 (-2.1)	0.14 (0.5)	0.30 (1.8)	0.18 (1.6)	-0.59 (-2.1)	-0.07 (-1.)
$ED_4$	-1.95 (-6.9)	-0.09 (-1.4)	0.34 (2.)	0.05 (0.1)	0.02 (0.1)	0.06 (0.5)	0.33 (1.)	0.48 (5.2)
$ED_5$	-2.16 (-5.95)	-0.15 (-1.8)	0.60 (2.22)	0.38 (0.8)	0.99 (3.7)	-0.20 (-0.8)	-0.19 (-0.4)	0.89 (6.2)
$R^2$	0.35	0.03	0.03	0.11	0.04	0.007	0.18	0.15
$F$	499.7	34.2	34.16	112.67	43.65	7.11	119.7	166.9
BP $\chi^2( )$	10,089.8	9,169.	10,722.1	8,098.2	9,416.8	8,794.9	8,250.3	7,519.6

<sup>1</sup> NOTAS: Ecuaciones (6) estimadas por mínimos cuadrados.

Estimaciones de la matriz de varianzas y covarianzas de WHITE. T. estadísticos en paréntesis.

Todos los coeficientes (excepto la constante) aparecen multiplicados por 100.

BP = Estadístico de Breusch-Pagan.

**TABLA 3**  
**CURVAS DE ENGEL**  
**ESTIMACIONES DE EFECTOS FIJOS<sup>1</sup>**

	Alimen.	Alcohol y Tabaco	Vestido y Calzado	Hogar	Vivienda	Medicina y Farmacia	Transport. y Comunicac.	Educación
$\ln(x/n)$	13.13 (2.3)	0.94 (0.7)	21.52 (11.)	-50.18 (-9.8)	0.91 (0.3)	0.81 (0.5)	-7.93 (-2.)	8.74 (10.2)
$(\ln(x/n))^2$	-1.05 (-4.7)	-0.06 (-1.2)	-0.83 (-10.5)	1.97 (9.4)	0.06 (0.7)	-0.02 (-0.3)	0.61 (3.8)	-0.36 (-10.4)
$n_1$	-0.69 (-4.)	-0.20 (-4.8)	0.22 (2.2)	-0.88 (-5.3)	0.91 (7.9)	0.32 (5.3)	0.72 (4.)	-0.19 (-4.6)
$n_2$	-0.45 (-3.4)	-0.25 (-7.7)	0.34 (4.6)	-1.19 (-9.4)	0.07 (1.0)	-0.06 (-1.3)	0.93 (7.1)	0.56 (14.9)
$n_3$	-0.06 (-0.5)	-0.29 (-10.2)	0.44 (6.4)	-1.05 (-9.1)	0.28 (3.9)	0.01 (0.4)	0.34 (3.0)	0.62 (17.1)
$n_4$	-1.01 (-7.5)	-0.14 (-4.)	0.56 (7.1)	-1.30 (-9.8)	-0.002 (-0.)	-0.04 (-0.8)	-0.03 (-0.2)	0.87 (19.1)
$na_1$	-2.34 (-19.5)	0.12 (4.3)	0.23 (3.2)	-1.76 (-15.3)	-0.11 (-1.5)	-0.11 (-2.7)	1.11 (9.5)	0.36 (8.9)
$na_2$	-1.53 (-10.3)	0.11 (2.9)	0.04 (0.5)	-1.89 (-12.4)	-0.62 (-7.1)	0.03 (0.5)	1.92 (13.5)	0.17 (4.5)
$n_0$	-0.80 (-4.4)	-0.04 (-0.9)	-0.19 (-1.9)	-0.21 (-1.1)	-0.19 (-1.7)	0.09 (1.5)	0.81 (4.8)	0.008 (0.1)
$ED_1$	1.77 (4.)	0.52 (4.1)	-0.33 (-1.4)	-0.86 (-2.1)	0.69 (3.2)	-0.20 (-1.4)	-0.56 (-1.6)	-0.30 (-3.5)
$ED_2$	0.89 (3.6)	0.20 (3.2)	-0.29 (-2.1)	-0.08 (-0.3)	0.29 (2.1)	0.14 (1.6)	-0.23 (-1.)	-0.17 (-2.8)
$ED_4$	-1.88 (-7.9)	-0.11 (-2.)	0.28 (1.9)	0.25 (0.9)	0.06 (0.4)	0.11 (1.1)	0.02 (0.)	0.40 (5.6)
$ED_5$	-2.29 (-7.4)	-0.22 (-3.1)	0.14 (0.6)	0.72 (1.9)	1.09 (4.4)	-0.04 (-0.3)	-0.17 (-0.4)	0.76 (7.)
$R^2$	0.35	0.04	0.03	0.10	0.05	0.007	0.17	0.14
$F$	488.6	38.7	32.22	109.2	49.86	6.84	189.6	152.11
BP $\chi^2( )$	9,904.2	8,364.1	10,624.8	7,356.9	9,339.7	7,126.9	7,474.2	7,388.5

<sup>1</sup> NOTAS: Ecuaciones (6) estimadas por mínimos cuadrados.

Estimaciones de la matriz de varianzas y covarianzas de WHITE. T. estadísticos en paréntesis.

Todos los coeficientes (excepto la constante) aparecen multiplicados por 100.

BP = Estadístico de Breusch-Pagan.

**TABLA 4**  
**ELASTICIDADES RENTA**

Regresión	Alimen.	Alcohol y Tabaco	Vestido y Calzado	Hogar	Vivienda	Medicina y Farmacia	Transport. y Comunicac.	Educación
Efectos fijos	0.641	0.814	1.143	0.895	1.345	1.147	1.648	0.956
Efectos fijos ponderados	0.636	0.865	1.003	0.887	1.35	1.145	1.687	1.059
Mínimos cuadrados	0.627	0.830	1.137	0.948	1.370	1.269	1.611	0.949
Mínimos cuadrados ponderados	0.624	0.843	1.143	0.952	1.33	1.22	1.63	0.98

En cuanto a los EFECTOS DEMOGRAFICOS SOBRE LA COMPOSICION DEL CONSUMO se observa que todas las variables demográficas ( $n_1$  a  $n_o$ ) tienen coeficientes negativos en las demandas de Alimentación y Gastos de Hogar, y coeficientes positivos en las demandas de Transportes y Comunicaciones, y Educación. Es decir, si se mantuviese el gasto total per cápita ( $\frac{x}{m}$ ) constante, al aumentar el tamaño familiar disminuiría el gasto en alimentos y bienes del hogar y crecería la participación de los bienes educativos y los gastos de transporte en el gasto total. Si el nuevo miembro del hogar fuese un niño ( $n_1$  a  $n_4$ ) también crecerían las participaciones en el gasto de la vivienda y del vestido, mientras que se reduciría la proporción del gasto destinada a alcohol y tabaco.

Las estimaciones de las TABLAS 2 y 3 también permiten observar que el nivel educativo del cabeza de familia muestra una relación inversa con la proporción del gasto destinada a alimentación, y a alcohol y tabaco; y una relación directa (significativa) con las participaciones de los gastos educativos en el gasto total.

En cuanto al ajuste total de estas ecuaciones, se obtienen estadísticos  $R^2$  que no superan el valor de 0.35 (alimentación), como es habitual en regresiones con datos de sección cruzada. Los estadísticos  $F$  indican que todas las ecuaciones son significativas, superando ampliamente los criterios de significación habituales [ $F_{(1\%)} \geq 1.47$ ]; aunque se debe señalar que el modelo no parece muy adecuado para explicar el gasto privado en medicina y farmacia. Ni el gasto total per capita, ni la composición familiar, ni las otras características socioeconómicas (educación, situación laboral, etc.) muestran una relación estadística significativa con los gastos médicos recogidos en la E.P.F.

### 4.3 Tests de efectos fijos

En las filas 2 y 4 de la TABLA 5 se contrastan la hipótesis  $H_0 : E(\alpha_i/S_{ji}) = 0$  comparando  $\hat{B}_W$  con el estimador  $\hat{B}_B$  que se obtiene en ecuaciones que no incluyen las variables  $Z_{2i}$ , que recogen características comunes a los hogares de un bloque. Los tests rechazan la hipótesis nula para todas las curvas de demanda; es decir, la presencia de efectos bloque hace que las estimaciones de mínimos cuadrados de las curvas de Engel sean inconsistentes.

**TABLA 5**  
**TESTS DE EFECTOS FIJOS**

$\chi^2$		Alimen.	Alcohol y Tabaco	Vestido y Calzado	Hogar	Vivienda	Medicina y Farmacia	Transport. y Comunicac.	Educación
Regresión sin Ponderar									
1	$S_{ji}, Z_i$	96.6	78.3	34.8*	901.1	65.4	37.7*	91.8	112.4
2	$S_{ji}$	157.7	98.8	67.2	392.9	54.3	45.8	96.0	197.2
Regresión Ponderada									
3	$S_{ji}, Z_i$	81.9	72.7	44.1*	190.3	40.1*	45.5*	98.3	121.0
4	$S_{ji}$	106.1	91.5	103.4	322.5	49.3	45.4*	98.6	201.8

$$\chi_{24}^2(0,5\%) = 45,6$$

$$\chi_{24}^2(1\%) = 43$$

\*: No se rechaza  $H_0$ .

En las filas 1 y 3 de la TABLA 5 se contrasta la hipótesis  $H_0 : E(\alpha_i/S_{ji}, Z_{2i}) = 0$ , comparando  $\hat{B}_W$  con los parámetros  $\hat{B}_B$  estimados en ecuaciones que incluyen las variables explicativas comunes a los hogares de un bloque ( $Z_{2i}$ ) (por ej: tamaño del municipio, comunidad autónoma, trimestre de la encuesta, etc.).

Los tests rechazan de nuevo la hipótesis nula en la mayor parte de los casos; las excepciones aparecen para los bienes: vestido y calzado, vivienda, y medicina y farmacia. Para estos 3 grupos de bienes la inclusión de las variables  $Z_{2i}$  en las curvas de Engel elimina el problema de inconsistencia que crean los efectos bloque. Merece la pena resaltar este resultado ya que, aún en los casos en que no exista un interés genuino en los coeficientes de estas variables,  $Z_{2i}$ , se observa que su inclusión puede eliminar los problemas de inconsistencia en la estimación de  $B$  que generan los efectos bloque. A priori no se podía predecir que la inclusión de  $Z_{2i}$  haría más o menos probable la consistencia de  $\hat{B}$

de Mínimos Cuadrados.

Dado el tamaño de la muestra utilizada ( $N=21.705$ ) se puede argumentar que los criterios de significación habituales no son aplicables. La utilización del procedimiento clásico de contraste de hipótesis, comparando los valores del estadístico  $Q$  en la TABLA 5 con un valor crítico prefijado  $\chi^2_\alpha$ , tal que

$$Pr(Q \geq \chi^2_\alpha) \leq 1 - \alpha$$

no tiene en cuenta el tamaño de la muestra que se está utilizando y es fuente de polémica entre estadística clásica y bayesiana. Así, en lugar del procedimiento clásico, se puede utilizar la regla de Schwartz (1975) para elegir la dimensión de un modelo con muestras grandes; pero esta regla no parece directamente aplicable a nuestro contraste de efectos fijos.

En cualquier caso, los valores del estadístico  $Q$  para algunos grupos de bienes son muy elevados por ej: para gastos del hogar, educación, alimentación, y transporte, y difícilmente se puede interpretar que estos valores de  $Q$  no aportan evidencia de la existencia de efectos fijos.

Si se interpreta que los resultados de la TABLA 5 aportan evidencia sobre la existencia de efectos bloque, entonces se pueden obtener estimaciones consistentes (aunque pueden no ser las más eficientes) de  $B$ , a partir de las ecuaciones en desviaciones (6). Pero no se pueden recuperar los parámetros ( $\delta_2$ ) de las variables comunes a los hogares de un bloque ( $Z_{2i}$ )<sup>5</sup>.

Por el contrario, si se pudiera descartar la presencia de efectos fijos en las ecuaciones, se podrían obviar todos estos problemas ya que las estimaciones de M.C.O. de  $B$  y  $\delta_2$  serían consistentes (aunque no necesariamente las más eficientes). En las TABLAS 6 y 7 se presentan las estimaciones de los parámetros de las Curvas de Engel por mínimos cuadrados, con y sin ponderaciones de las observaciones en diferentes estratos.

Se observa que no hay "grandes" diferencias entre estas estimaciones y las obtenidas en las TABLAS 2 y 3 (de efectos fijos). Aunque se debe señalar que los coeficientes de las variables educativas (por ej.) son aproximadamente un 20% más elevados al emplear Mínimos Cuadrados, mientras que los coeficientes del gasto per cápita crecen entre un 10 y un 15% al aplicar el método de efectos fijos. En la TABLA 4 se puede comprobar que

---

<sup>5</sup>Si hay interés en obtener estimaciones consistentes de estos coeficientes se pueden utilizar procedimientos en dos etapas.(Véase Hausman y Taylor (1981)), y San Segundo (1985)).

las elasticidades-renta apenas varían con el método de estimación utilizado<sup>6</sup>.

En resumen, aunque al disponer de 21.705 observaciones los tests de especificación habituales permiten detectar efectos bloque (TABLA 5); para este problema concreto de estimación de Curvas de Engel con datos de la EPF no parece que al utilizar los parámetros estimados por mínimos cuadrados se incurra en grandes sesgos; cuidando de que la especificación de las curvas de demanda incluya variables "bloque" ( $Z_{2i}$ ).

En la próxima sección del trabajo se considera la importancia en la estimación de Curvas de Engel de los efectos bloque como efectos aleatorios.

## 5 CONTRASTES DE EFECTOS ALEATORIOS

### 5.1 Derivación de los tests

Si, a pesar de los resultados obtenidos en la TABLA 5, se decide utilizar las estimaciones de mínimos cuadrados, no tratando a los posibles efectos bloque como efectos fijos, queda por considerar su importancia como efectos aleatorios, de acuerdo con el modelo propuesto en las ecuaciones (2).

El test del multiplicador de Lagrange de la hipótesis  $H_0 : \rho = 0$  utiliza los estimadores máximo verosímiles que satisfacen  $\rho = 0$ , y se basa en el siguiente estadístico:

$$r^2 = \frac{\partial \log L}{\partial \rho}(\rho = 0) V^{-1} \frac{\partial \log L}{\partial \rho}(\rho = 0) = \frac{2}{\sum_i n_i (n_i - 1)} \left[ \frac{\sum_i \sum_j \sum_{h < j} \hat{U}_{ji} \hat{U}_{hi}}{\sum_i \sum_j \hat{U}_{ji}^2} \right]^2$$

Bajo  $H_0$  este estadístico se distribuye como una  $\chi^2_1$ , siendo  $\hat{U}$  los residuos de M.C.O. Deaton y Irish (1983) derivan este estadístico aplicando los resultados de Breusch y Pagan (1980) a nuestro modelo.

Si se rechaza la hipótesis nula de no correlación, se pueden estimar  $\rho$  y  $\sigma^2$  consistentes de la siguiente forma:

$$\hat{\rho} = \frac{\sum_{i=1}^B \sum_{j=1}^{n_i} \sum_{h < j} \hat{U}_{ji} \hat{U}_{hi}}{\frac{1}{2} \sum_i n_i (n_i - 1) \hat{\sigma}^2} \quad (7)$$

$$\hat{\sigma}^2 = \frac{1}{N(1 - \hat{\rho})} \sum_i \left[ \hat{U}_i \hat{U}'_i - \frac{\hat{\rho}}{1 + (n_i - 1)\hat{\rho}} (\hat{U}'_i L)^2 \right]$$

donde  $\hat{\sigma}^2$  y  $\hat{U}$  son estimadores de M.C.O.

<sup>6</sup>En la versión original de este trabajo se comprueba que las diferencias mínimas entre  $\hat{B}_W$  y  $\hat{B}_{M.C.O.}$  tampoco se traducen en diferentes escalas de equivalencia (medidas del coste de los niños).

**TABLA 6**  
**CURVAS DE ENGEL**  
**MINIMOS CUADRADOS PONDERADOS**

	Alimen.	Alcohol y Tabaco	Vestido y Calzado	Hogar	Vivienda	Medicina y Farmacia	Transport. y Comunicac.	Educación
$\hat{\alpha}$	0.45 (1.)	0.071 (0.8)	-1.25 (-9.2)	3.23 (8.)	-0.08 (-0.4)	-0.05 (-0.3)	-0.007 (-0.003)	-0.55 (-8.9)
$\ln(x/n)$	10.3 (1.5)	-0.08 (-0.0)	20.3 (9.1)	-48.6 (-7.2)	0.56 (0.1)	0.68 (0.2)	-3.95 (-0.9)	9.05 (8.9)
$(\ln(x/n))^2$	-0.96 (-3.4)	-0.02 (-0.3)	-0.78 (-8.7)	1.95 (7.)	0.07 (0.6)	-0.008 (-0.07)	0.44 (2.4)	-0.37 (-8.9)
$n_1$	-0.84 (-4.)	-0.17 (-3.3)	0.12 (1.1)	-0.72 (-3.5)	0.70 (5.2)	0.40 (5.9)	0.91 (4.2)	-0.15 (-2.8)
$n_2$	-0.59 (-3.7)	-0.22 (-5.9)	0.28 (3.3)	-1.16 (-7.4)	0.12 (1.3)	0.03 (0.6)	1.01 (6.1)	0.68 (12.9)
$n_3$	-0.19 (-1.3)	-0.28 (-8.2)	0.49 (5.9)	-1.07 (-7.7)	0.28 (3.6)	0.07 (1.5)	0.22 (1.6)	0.68 (13.6)
$n_4$	-1.11 (-6.8)	-0.12 (-3.1)	0.52 (5.4)	-1.40 (-8.7)	0.011 (0.1)	0.003 (0.05)	0.18 (1.2)	0.85 (14.)
$na_1$	-2.28 (-15.7)	0.11 (3.4)	0.29 (3.5)	-1.65 (-11.1)	-0.08 (-0.9)	-0.07 (-1.5)	1.18 (8.3)	0.30 (6.1)
$na_2$	-1.32 (-7.6)	0.14 (3.)	0.10 (1.1)	-1.95 (-10.7)	-0.73 (-7.3)	0.01 (0.2)	1.98 (12.1)	0.19 (4.4)
$n_0$	-0.71 (-3.3)	-0.02 (-0.3)	-0.17 (-1.5)	-0.26 (-1.1)	-0.42 (-3.6)	0.18 (2.6)	0.93 (4.6)	0.04 (0.7)
$ED_1$	1.84 (3.8)	0.51 (3.5)	-0.45 (-1.9)	-0.48 (-0.9)	0.43 (1.7)	-0.08 (-0.5)	-0.62 (-1.8)	-0.23 (-2.3)
$ED_2$	0.56 (2.2)	0.33 (4.9)	-0.30 (-2.1)	0.26 (1.)	0.11 (0.8)	0.10 (1.)	-0.32 (-1.4)	-0.14 (-2.4)
$ED_4$	-2.21 (-8.2)	-0.20 (-3.1)	0.23 (1.3)	0.47 (1.5)	0.06 (0.3)	0.14 (1.2)	-0.20 (-0.6)	0.76 (8.2)
$ED_5$	-3.05 (-9.1)	-0.28 (-3.4)	0.35 (1.3)	1.62 (3.5)	1.17 (4.3)	0.05 (0.2)	-0.98 (-2.0)	1.26 (8.6)
$R^2$	0.63	0.16	0.14	0.29	0.14	0.05	0.28	0.23
$F$	738.	82.4	72.5	177.5	69.8	23.9	164.9	129.6
BP $X_{51}^2$	6,037.4	12,610.2	3,841.5	9,762.7	6,643.9	18,635.1	10,552.8	13,742.4

<sup>1</sup> NOTAS: Ecuaciones ( ) estimadas por mínimos cuadrados ponderados.

Estimaciones de la matriz de varianzas y covarianzas de WHITE. T estadísticos en paréntesis.

Las regresiones incluyen como variables explicativas: 16 variables ficticias para indicar la Comunidad Autónoma de residencia y 3 variables ficticias que indican la semana del mes en la que se realizó la entrevista.

**TABLA 7**  
**CURVAS DE ENGEL**  
**MINIMOS CUADRADOS ORDINARIOS**

	Alimen.	Alcohol y Tabaco	Vestido y Calzado	Hogar	Vivienda	Medicina y Farmacia	Transport. y Comunicac.	Educación
$\hat{\alpha}$	0.50 (1.5)	0.08 (1.1)	-1.22 (-9.8)	2.83 (8.2)	-0.08 (-0.5)	-0.05 (-0.5)	0.06 (0.2)	-0.47 (-9.)
$\ln(x/n)$	7.98 (1.3)	-0.24 (-0.2)	20.0 (9.8)	-42.1 (-7.4)	0.35 (0.1)	0.83 (0.4)	-4.90 (-1.2)	7.75 (9.)
$(\ln(x/n))^2$	-0.86 (-3.5)	-0.01 (-0.3)	-0.77 (-9.3)	1.68 (7.1)	0.09 (0.9)	-0.01 (-0.2)	0.47 (2.8)	-0.32 (-9.1)
$n_1$	-0.91 (-5.2)	-0.12 (-2.9)	0.15 (1.4)	-0.73 (-4.3)	0.82 (7.0)	0.35 (6.2)	0.90 (4.9)	-0.22 (-4.8)
$n_2$	-0.50 (-3.7)	-0.24 (-7.5)	0.27 (3.7)	-0.99 (-7.5)	0.07 (1.0)	-0.005 (-0.1)	0.85 (6.4)	0.56 (16.)
$n_3$	-0.18 (-1.5)	-0.26 (-9.5)	0.47 (6.7)	-0.95 (-7.9)	0.27 (3.8)	0.07 (1.7)	0.25 (2.2)	0.60 (18.8)
$n_4$	-0.99 (-7.2)	-0.11 (-3.3)	0.56 (6.8)	-1.29 (-9.5)	-0.04 (-0.5)	-0.01 (-0.2)	-0.01 (-0.1)	0.85 (22.5)
$na_1$	-2.29 (-18.7)	0.13 (4.5)	0.22 (3.1)	-1.73 (-14.6)	-0.09 (-1.2)	-0.07 (-1.9)	1.04 (8.9)	0.33 (10.1)
$na_2$	-1.37 (-9.)	0.13 (3.2)	0.10 (1.1)	-1.94 (-12.1)	-0.68 (-7.9)	0.06 (1.2)	1.97 (13.8)	0.18 (4.7)
$n_0$	-0.64 (-3.4)	-0.03 (-0.7)	-0.14 (-1.4)	-0.23 (-1.1)	-0.28 (-2.7)	0.20 (3.2)	0.79 (4.7)	0.02 (0.4)
$ED_1$	1.82 (4.1)	0.52 (4.1)	-0.25 (-1.1)	-0.70 (-1.6)	0.44 (2.1)	-0.17 (-1.3)	-0.45 (-1.5)	-0.34 (-3.5)
$ED_2$	0.47 (2.1)	0.30 (5.3)	-0.28 (-2.3)	0.06 (0.3)	0.17 (1.5)	0.15 (1.9)	0.02 (0.1)	-0.20 (-3.8)
$ED_4$	-2.08 (-9.)	-0.21 (-3.9)	0.22 (1.4)	0.61 (2.4)	0.09 (0.5)	0.15 (1.6)	-0.49 (-1.8)	0.64 (9.5)
$ED_5$	-3.04 (-10.4)	-0.31 (-4.7)	0.03 (0.1)	1.54 (4.0)	1.25 (5.1)	0.02 (0.2)	-0.73 (-1.9)	1.09 (11.5)
$R^2$	0.44	0.08	0.04	0.10	0.06	0.01	0.19	0.17
$F$	342.1	38.8	20.1	49.89	30.2	8.34	100.3	92.04
BP $\chi^2_{51}$	2,439.1	5,776.2	496.1	4,596.7	2,671.1	7,023.1	4,411.5	8,157.7

<sup>1</sup> NOTAS: Estimaciones de la matriz de varianzas y covarianzas de WHITE. T estadísticos en paréntesis.

Todos los coeficientes (excepto la constante) aparecen multiplicados por 100.

Las regresiones incluyen como variables explicativas: 16 variables ficticias para indicar la Comunidad Autónoma de residencia y 3 variables ficticias que indican la semana del mes en la que se realizó la entrevista.

**TABLA 9**  
**ESTIMACIONES Y TESTS DE  $\sigma^2$**

	Alimen.	Alcohol y Tabaco	Vestido y Calzado	Hogar	Vivienda	Medicina y Farmacia	Transport. y Comunicac.	Educación
Regresión ponderada								
$\hat{\sigma}^2$	0.0135	0.0008	0.0043	0.0151	0.0044	0.0019	0.0120	0.0009
$\hat{\sigma}^2$	0.0124	0.0008	0.0041	0.0133	0.0043	0.0018	0.0116	0.0008
Regresión sin ponderar								
$\hat{\sigma}^2$	0.0140	0.0008	0.0046	0.0153	0.0046	0.0017	0.0121	0.0009
$\hat{\sigma}^2$	0.0128	0.0008	0.0043	0.0134	0.0045	0.0017	0.0116	0.0008

A partir de los resultados obtenidos en las TABLAS 8 y 9 se decide no corregir las estimaciones de la matriz de varianzas y covarianzas por la presencia de efectos bloque, y se opta por mantener las estimaciones consistentes bajo heteroscedasticidad.<sup>8</sup>

Para interpretar los resultados, es útil compararlos con los obtenidos por Deaton y Irish (1983) con datos de consumo de alimentos de Sri-Lanka. (resumidos en la TABLA 10). Se observa que ellos obtienen evidencia de efectos aleatorios, con estimaciones de  $\rho$  entre 0.26 y 0.41; y, para la muestra de la zona Tamil encuentran evidencia de efectos fijos.

Las curvas de Engel estimadas por Deaton y Irish para los gastos de alimentación tienen la siguiente forma:

$$W_{ji} = \alpha + \beta \log\left[\frac{x}{m}\right]_{ji} + \gamma_a M_{a,ji} + \gamma_c M_{c,ji} + \varepsilon_{ji}$$

Es decir, sólo se utilizan 3 variables explicativas (gasto per cápita, número de adultos y número de niños en el hogar  $j$ , más el término constante), en contraste con las 50 variables que aquí empleadas con la muestra de la EPF. Al realizar los contrastes de efectos fijos, en la TABLA 5 se observ que los problemas que puede crear el muestreo en bloques se pueden eliminar o al menos mitigar si las ecuaciones incluyen suficientes variables explicativas. Para enfatizar este resultado, en la TABLA 11 se comprueba que

<sup>8</sup>En las TABLAS 8 y 9 se contrasta un modelo muy particular de correlación en los bloques: se supone que la correlación ( $\rho$ ) es idéntica en todos los bloques de la muestra. Se podrían haber considerado modelos alternativos, por ej. si hubiera razones para suponer que los efectos bloque son comunes a algunas de las demandas (para diferentes bienes)<sub>j</sub> se utilizaría un modelo similar al de Chamberlain-Griliches (1975), que en este caso requeriría un proceso de estimación muy costoso.

**TABLA 10**  
**CURVAS DE ENGEL DE ALIMENTACION PARA SRI LANKA**

	Zona URBANA	Zona RURAL	Zona III
$\hat{\rho}$	0.2661	0.2893	0.4099
$\hat{\rho}_b$	0.2756	0.2938	0.3474
TEST $r^2$ de $\rho = 0$	18,8	18,8	21,6
TEST de ef. fijos $\chi^2_3$	3,48*	2,43*	15,1
n° observaciones	1102	1032	585
n° bloques	119	122	60

Fuente: Deaton y Irish (1983)

la estimación de una Curva de Engel para el gasto en alimentos con únicamente 9 variables explicativas:

$$W_{ji} = \alpha + \beta_1 \ln\left[\frac{x}{m}\right]_{ji} + \beta_2 \left[\ln\left[\frac{x}{m}\right]_{ji}\right]^2 + \sum_k \gamma_k M_{kji} + \varepsilon_{ji}$$

produce estimaciones de  $\rho$  (la correlación dentro de los bloques) entre 0.08 y 0.11, significativamente más altas que las obtenidas en la Tabla 9; y los tests de efectos fijos también son más favorables a los efectos bloque.

**TABLA 11**  
**CURVA DE ENGEL DE ALIMENTACION**  
**EPF 80-81**

Regresión Ponderada	Estimaciones y Tests de $\hat{\rho}$		Tests de Efectos Fijos		
	50 var.expl.	9 var.expl.		50 var.expl.	9 var.expl.
$\hat{\rho}$	0.087	0.1141			
$\hat{\rho}_b$	0.078	0.1007	$\chi^2$	81.9	112.7
$r^2$ : Test de $\rho = 0$	485,3	834,8			
Regresión sin Ponderar					
$\hat{\rho}$	0.089				
$\hat{\rho}_b$	0.079	0.1047	$\chi^2$	96.6	163.2
$r^2$ : Test de $\rho = 0$	507,9				

## 6 CONCLUSIONES

En este trabajo se analizan los efectos que el diseño en "bloques" de la Encuesta de Presupuestos Familiares 1980-81 puede tener sobre la estimación de Curvas de Engel con datos de la EPF.

Al utilizar una muestra de 21.705 observaciones un test de efectos fijos de Hausman permite detectar la existencia de efectos bloque que hacen que los estimadores de mínimos cuadrados sean inconsistentes. Aunque, a efectos prácticos se comprueba que las estimaciones de mínimos cuadrados no son muy diferentes a las de efectos fijos, y ambos conjuntos de parámetros generan elasticidades - renta (TABLA 4) prácticamente idénticas.

Los contrastes de efectos aleatorios de nuevo detectan efectos bloque, a los niveles de significación habituales, pero no se obtiene evidencia que haga necesaria la corrección de la matriz de varianzas y covarianzas (TABLAS 8 y 9) para tener en cuenta estos efectos bloque.

Conviene resaltar el papel crucial que, en estos tests de especificación, juegan las variables que miden características comunes a los hogares de un bloque. Aún en los casos en que no se esté interesado en la relación de estas variables con la composición del consumo, su inclusión parece recomendable para mitigar el impacto del diseño muestral sobre los parámetros estimados.

En resumen, el diseño muestral en bloques de observaciones sugiere que un test de efectos fijos sería recomendable en modelos de regresión que utilicen datos de la EPF. Sin embargo, en este trabajo se comprueba, para la estimación de Curvas de Engel, que los parámetros obtenidos por mínimos cuadrados proporcionan una aproximación aceptable a las estimaciones consistentes (de efectos fijos), especialmente si se incluyen variables "bloque" en la regresión.

## REFERENCIAS

- ABADIA, A. (1985): "Income distribution and composition of consumer demand in the Spanish Economy". *European Economic Review* 29 (pp. 1-13).
- BALESTRA, P. and NERLOVE, M. (1966): "Pooling cross-section and time series data in the estimation of a dynamic model". *Econometrica* (pp. 585-612).
- BENNET, C.T. (1979): "The social security benefit structure. Equity considerations of the family as its basis". *American Economic Review* 69 (pp. 227-231).
- BREUSCH, T.S. and PAGAN, R. (1979). "A simple test for heteroskedasticity and random coefficient variation". *Econometrica* (pp. 1287-1294).
- CHAMBERLAIN, G. and GRILICHES, Z. (1975): "Unobservables with a variance - components structure: ability, schooling, and the economic success of brothers". *International Economic Review* 16 (2) (pp. 422-449).
- DEATON, A.S. (1981): "Three essays on a Sri Lankan Household survey". World Bank, Washington D.C.
- DEATON, A.S. and IRISH, M. (1983): "Block effects in regression analysis using survey data". Unpublished, Princeton D.P. n° 65.
- DEATON, A.S. and MUELLBAUER, J. (1980): "Economics and Consumer Behavior". Cambridge Univ. Press.
- (1986): "On measuring child costs in poor countries". *Journal of Political Economy* 94 (pp. 720-745).
- DEATON, A.S., RUIZ-CASTILLO, J. and THOMAS, D. (1989): "The influence of household composition on household expenditures patterns. Theory and Spanish evidence". *Journal of Political Economy* 97 (pp. 179-200).
- DuMOUCHEL, W. and DUNCAN, G. (1983): "Using sample survey weights in multiple regression analyses of stratified samples". *JASA* (pp. 535-543).
- HAUSMAN, J. (1978): "Specification tests in Econometrics". *Econometrica* 46 (pp. 1251-1272).
- HAUSMAN, J. and TAYLOR, W. (1981): "Panel data and unobservable individual effects". *Econometrica* 49 (pp. 1377-1398).
- HSIAO, C. (19): "Panel Data". *Econometric Society Monographs*.
- INE (1977-1983): "Encuesta Permanente de Consumo", números 1, 2 y 3.

- INE (1983): "Encuesta de Presupuestos Familiares 1980-81. El gasto y el ingreso de los hogares". Tomo 1°. Madrid.
- MACKINNON, J.G. and WHITE, H. (1985): "Some heteroskedasticity consistent covariance matrix estimators with improved finite sample properties". *Journal of Econometrics* 29 (pp. 305-325).
- NEWKEY, W.K. (1987): "Specification tests for distributional assumptions in the Tobit model". *Journal of Econometrics* 34 (pp. 125-145).
- POLLACK, R.A. and WALES, T.J. (1979): "Welfare comparisons and equivalence scales". *American Economic Review* 69 (pp. 216-221).
- SAN SEGUNDO, M. (1985): "Empirical studies of quality of schooling". Unpublished Ph.D. Thesis, Princeton University.
- SENECA, J. and TAUSING, M. (1971): "Family equivalence scales and personal income tax exemptions for children". *Review of Economics and Statistics* 53 (pp. 253-262).
- SCHWARTZ, G. (1978): "Estimating the dimension of a model". *Annals of Statistics* 6 (pp. 461-464).
- WALES, T.J. and WOODLAND, A. (1983): "Estimation of consumer demand systems with binding non-negativity constraints". *Journal of Econometrics* 21 (pp. 263-285).
- WHITE, H. (1980): "A heteroskedasticity consistent covariance matrix and a direct test for heteroskedasticity". *Econometrica* 48 (pp.817-838).