

# Variabilidad semanal de los gastos en la EPF

por  
MARIA PAZOS MORAN  
Instituto de Estudios Fiscales  
Ministerio de Economía y Hacienda

## RESUMEN

El período muestral de la EPF es de una semana. Ello da lugar a distorsiones a la hora de utilizar los datos individuales para el análisis de la distribución de los gastos: de la variabilidad observada, una parte responde a desigualdades reales, y otra, espúrea, a la distinta semana de entrevista. En este trabajo se contrasta esta hipótesis mediante el análisis de la varianza. Para detectar qué comparaciones entre semanas resultan significativas se aplican los tests de Bonferroni, Tukey y Waller-Duncan. Los efectos de este problema sobre las estimaciones de los índices de desigualdad se evalúan a través de las descomposiciones del índice de Theil.

*Palabras clave:* gastos en la EPF, análisis de la varianza, test de Waller-Duncan.

*Clasificación AMS:* 62J10, 62P20.

## 1. INTRODUCCION

El problema que motiva este trabajo es la posible influencia de la estacionalidad de los gastos de los hogares en la estructura interna de los microdatos de la Encuesta de Presupuestos Familiares. El período muestral de la Encuesta es la semana, repartiéndose la muestra aleatoria y uniformemente a lo largo de todas las semanas del año, a fin de evitar posibles distorsiones en las estimaciones debidas a componentes estacionales. Considerando la distinta periodicidad con que las familias realizan los diferentes gastos, se han establecido distintos «períodos de referencia», o períodos de los que se recoge información: semanal (esencialmente para alimentos, bebidas y tabaco, exceptuando las grandes compras, que tienen referencia mensual); mensual, para bienes como vestido y calzado; etc. Los gastos anuales de cada familia se obtienen multiplicando los gastos obtenidos en la entrevista por 52, 12, 4 ó 1, según el gasto sea de referencia semanal, mensual, trimestral o anual (INE, 1990).

El efecto de la semana de la entrevista tiene dos vertientes: en cuanto a los gastos de referencia semanal, hay que señalar que, aunque su frecuencia sea efectivamente la semana, el nivel de consumo dentro de cada familia varía de unas semanas a otras. Los de referencia mensual reflejarán, asimismo, la variación del consumo por meses, haciendo, por ejemplo, que una familia entrevistada en una semana de un mes de gran consumo registre mayores gastos que otra entrevistada en otro período. A ello se une en este caso el efecto olvido: el encuestado tenderá a olvidar más fácilmente los gastos que realizó con anterioridad, y consignará con más probabilidad los realizados la semana de la entrevista, máxime cuando estos últimos debe anotarlos él mismo en las llamadas «libretas» según se produzcan.

De todo lo anterior puede derivarse que, aunque las estimaciones de las medias no se vean afectadas, las varianzas de los diferentes gastos, y del gasto total, estén sobreestimadas. En ese caso, todos los cálculos de los índices de desigualdad estarían recogiendo los dos efectos: la desigualdad real y el efecto de la variación estacional de los gastos, tanto semanal como mensual y trimestral.

Hay que señalar que este problema es presumiblemente menor con los ingresos, ya que, por un lado, el período de referencia de los ingresos es el año y, por otro, los ingresos se registran más y están más concentrados, por lo que el factor olvido es menor. Se ha señalado también la contradicción de que, mientras que la literatura de Economía del Bienestar y la coherencia con los datos macroeconómicos indicarían que la desigualdad es mayor en los ingresos, los estudios realizados con las EPFs de 1980 y 1990 ofrecen evidencia contraria. Ahí se han visto a veces indicios respecto a la distinta subestimación de los ingresos según grupos o niveles sociales. Sin embargo, podría ser debido en parte a la mayor estacionalidad de los gastos antes señalada.

El problema ha sido apuntado en distintas ocasiones (Sanz, 1992). El hecho de que en otros países el período de referencia sea mayor (dos semanas es bastante general) lo agrava, pues podría distorsionar la comparación.

En este trabajo se trata de calibrar la magnitud del problema y hasta qué punto puede cuestionar el uso de los gastos como datos individuales para algunos fines.

## 2. BASE DE DATOS Y VARIABLES UTILIZADAS

La EPF 1990-1991 se realizó entre abril de 1990 y marzo de 1991. Cada familia es encuestada durante una semana a lo largo de cuatro visitas. En la primera se entregan los cuestionarios «Libreta de cuentas del hogar» y «Libretas individuales de pagos», en los que los encuestados deben anotar todos los gastos a lo largo de esa semana. En el resto de las visitas se obtienen, mediante entrevista, los gastos de referencia superior a la semana que se hayan efectuado en el último período correspondiente (mes, trimestre o año) anterior a esa semana. En la tercera visita se preguntan los ingresos durante el último año.

Se dispone de una variable que indica la semana del año en la que se realizó la entrevista. Esta se indica mediante el mes en número seguido del número de orden de la semana dentro del mes (por ejemplo: 123 es la tercera semana de diciembre).

Se ha descompuesto el gasto total (GASTOT) según la referencia de los gastos en gastos de referencia semanal (REFSEM) y gastos de referencia superior a la semana (REFSUP). Por otro lado, se han considerado los gastos realizados en la semana de referencia (DENTRO) y los que se han producido fuera de dicha semana (FUERA). Como resultado, hemos creado las siguientes variables que podrían ser relevantes para nuestro análisis:

REFSEM: Gasto del hogar en todos los bienes y servicios de referencia semanal.

REFSUP: Idem de referencia superior a la semana.

.....

DENTRO: Gastos cuyos pagos fueron efectuados durante la semana de referencia.

FUERA: Idem fuera de la semana de referencia.

### 3. PROCEDIMIENTOS ESTADÍSTICOS

Para detectar si la semana de la entrevista tiene un efecto en la estimación de los gastos familiares o si, por el contrario, las medias en las submuestras entrevistadas durante las distintas semanas podrían ser distintas simplemente como consecuencia de un puro efecto aleatorio, se ha recurrido al análisis de la varianza clásico:

Sea  $J$  el número de grupos (semanas en este caso) e  $I$  el número de observaciones en cada grupo. Llamaremos  $SS_w$  y  $SS_b$  a las sumas de cuadrados dentro y entre grupos, respectivamente. El modelo es:

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

donde  $\mu$  es la media total,  $\alpha_i$  es el efecto diferencial de la semana  $i$  ( $\sum \alpha_i = 0$ ) y  $\varepsilon_{ij}$  es el error aleatorio. Los errores  $\varepsilon_{ij}$  se suponen independientes e idénticamente distribuidos con distribución normal de media cero y varianza  $\sigma^2$ , que se estima mediante:

$$s_p^2 = (SS_w) / [I(J-1)]$$

La hipótesis nula que contrastamos es:

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_{52} = 0$$

Para ello se utiliza el estadístico  $F = [SS_b / (I-1)] / [SS_w / I(J-1)]$ , que bajo  $H_0$  se distribuye según una distribución  $F$  con  $(I-1)$  y  $I(J-1)$  grados de libertad.

En el caso de que el resultado del test general fuera, como sucederá en nuestro caso, que la semana de la entrevista produce un efecto significativo, las dos preguntas que se plantean como consecuencia del análisis anterior son: en primer lugar, ¿qué semanas son diferentes?, y, en segundo lugar, ¿en qué medida afecta esta variabilidad debida a la semana de la entrevista a la utilización de los datos micro, y singularmente al cálculo de índices de desigualdad? La segunda pregunta se analizará en el apartado 5 mediante la descomposición del índice de Theil. La primera (apartado 4), con los tests que se describen a continuación.

Para discriminar entre grupos de semanas diferentes se han utilizado los tests de Tukey, de Bonferroni y de Waller-Duncan, los tres basados en la misma idea de encontrar una cifra mínima por encima de la cual las diferencias serían significativas (mínima diferencia significativa) para las medias de los grupos, aunque con estrategias algo diferentes (Rice, 1988).

El test de Tukey se basa en la distribución de la variable aleatoria:

$$\max_{1,2} ( |(Y_{.1} - \mu_1) - (Y_{.2} - \mu_2)| ) / (s_p / \sqrt{J})$$

donde  $Y_{.1}$ ,  $Y_{.2}$  son todos los posibles pares de medias muestrales de los grupos ( $\mu_1$ ,  $\mu_2$  sus correspondientes medias poblacionales). Bajo la hipótesis nula  $H_0$ , esta variable aleatoria se distribuye según la llamada «Distribución del recorrido estudentizado» (Mood y Graybill, 1972). Llamando  $q_{1,1(J-1)}(\alpha)$  al punto que señala el 100  $\alpha$  por ciento de la distribución, podemos construir intervalos de confianza simultáneos para las diferencias de medias al nivel 100 (1 -  $\alpha$ )%. Estos intervalos son:

$$(Y_{.1} - Y_{.2}) \pm q_{1,1(J-1)}(\alpha) (s_p / \sqrt{J})$$

Por la dualidad entre intervalos de confianza y test de hipótesis, si el intervalo de confianza para  $\mu_1 - \mu_2$  no incluye el cero, o sea, si la diferencia entre dos medias muestrales  $Y_{.1} - Y_{.2}$  es mayor que la mínima diferencia significativa:

$$|Y_{.1} - Y_{.2}| > q_{1,1(J-1)}(\alpha) (s_p / \sqrt{J})$$

entonces la hipótesis nula de que no hay diferencia entre  $\mu_1$  y  $\mu_2$  puede rechazarse al nivel de significación  $\alpha$ .

Bonferroni parte de la siguiente idea: si deben hacerse  $k$  comparaciones de pares de medias, un error de tipo 1 máximo de  $\alpha$  puede garantizarse si se hace cada comparación al nivel  $\alpha / k$ . Dicho de otra manera, si se tiene que obtener un intervalo de confianza global al nivel 100 (1 -  $\alpha$ )%, cada uno de los  $k$  intervalos debe ser al nivel de confianza de 100 (1 - ( $\alpha / k$ )) (mucho más grandes). Estos dos tests suelen arrojar resultados muy próximos.

El tercer test, el de Waller-Duncan, se basa en una aproximación bayesiana que consiste en minimizar el riesgo de Bayes bajo la hipótesis de aditividad de la pérdida (Waller y Duncan, 1969; Duncan, 1975). Se considera, para cada comparación de dos medias, la función pérdida por cada decisión de rechazar o no rechazar la hipótesis nula  $\mu_1 - \mu_2 \leq 0$ . Estas funciones se consideran aditivas, con lo que la pérdida total es la suma de las pérdidas de las comparaciones individuales. Finalmente, se trata de tomar la decisión que minimice el riesgo. La distribución *a priori* de las medias  $\mu_i$  se considera normal con varianza desconocida. Para cada par de medias muestrales  $Y_{.i}$ ,  $Y_{.j}$ , la hipótesis nula se rechaza si:

$$y_{.i} - y_{.j} \geq t_B s_p \sqrt{(2 / J)}$$

donde  $t_B$  es el valor de  $t$  bayesiano, que depende del estadístico  $F$  de ANOVA y de sus grados de libertad, así como de un factor  $k$  a especificar en la función de pérdida (Waller y Kemp, 1975) (ver también el manual de SAS: *Stat User's Guide*).

Como último contraste, si se comparasen las medias separadamente dos a dos usando la estimación de la varianza  $s_p$ , se obtendría que dos semanas son significativamente diferentes si sus medias difieren más de

$$t_{(J-1)}(\alpha/2) s_p \sqrt{(2/J)}$$

Este es el resultado más estricto, aunque en nuestro caso, para  $\alpha = 0,05$  prácticamente coincide con el resultado del test de Waller-Duncan.

#### 4. ANALISIS DE LOS DATOS

Se observa, sobre todo, que algunas semanas arrojan medias mucho más altas, estando estas semanas localizadas en los meses de diciembre y enero. Por lo demás, no parece que las oscilaciones sigan unas pautas estacionales claras (por ejemplo, dentro del mes, trimestres, verano, etc.). Es interesante el hecho de que esto sucede para todas las componentes del gasto que se han distinguido: no sólo para los gastos de referencia semanal, y no sólo para los gastos realizados dentro de la semana, aunque en éstos las diferencias son mayores.

Como se desarrolla en el apartado anterior, se ha utilizado el análisis de la varianza (ANOVA) para detectar si la semana de la entrevista tiene un efecto en la estimación de los gastos familiares o si, por el contrario, las medias de las distintas semanas podrían ser distintas simplemente como consecuencia de un puro efecto aleatorio. El test rechaza claramente esta última hipótesis, tanto para el total de gastos (GASTOT) como para las componentes consideradas ( $p = 0,0001$ ). Los resultados se ofrecen en el cuadro 1.

La variabilidad entre semanas es «demasiado grande» en comparación con la variabilidad dentro de cada semana. Debemos concluir, pues, que las semanas son diferentes, es decir, que el gasto medio familiar es significativamente diferente para los encuestados en diferentes semanas. No obstante, es necesario aludir aquí a un elemento de precaución: cuando las muestras son tan grandes como en nuestro caso, ANOVA (como en general todos los tests) tiende a detectar diferencias significativas aunque éstas sean muy moderadas, ya que el poder estadístico aumenta rápidamente con el tamaño de la muestra. Por ello, más que mirar simplemente el valor de  $p$  y concluir que las diferencias son significativas, merece la pena observar y comparar los valores del estadístico  $F$ , siendo su magnitud la que nos indicará el grado de los efectos.

**Cuadro 1**  
**ANOVA Y TEST DE WALLER-DUNCAN**

Variable	Estadístico F (gr. lib. 51, 21103)	Pr > F	Mínima dif. signific. según Waller	Semanas más diferentes
GASTOT	4,03	0,0001	233.914	11
REFSEM	4,75	0,0001	102.129	123 124 11
REFSUP	4,34	0,0001	170.139	11 12
DENTRO	5,39	0,0001	111.748	11 123 124
FUERA	3,78	0,0001	166.437	

En cuanto al análisis de qué semanas son diferentes, para el gasto total se han aplicado los tres tests, descritos en el apartado anterior, de Tukey, de Bonferroni y de Waller-Duncan.

El test de Tukey proporciona una mínima diferencia significativa de 467.146 pesetas. El test de Bonferroni, que, como se ha dicho, suele arrojar resultados muy próximos al de Tukey, indica en nuestro caso la mínima diferencia significativa de 479.817 ptas. Con el tercer test, el de Waller-Duncan, la mínima diferencia significativa es de 233.914 ptas. En el apéndice se presentan las medias de la variable gasto total (GASTOT) en cada semana, junto con el resultado detallado de este último test. Por último, comparando las medias separadamente dos a dos, usando para ello la estimación de la varianza  $s_p$ , se obtendría que dos semanas son significativamente diferentes, a un nivel de significación de 0,05, si sus medias difieren más de 228.263 ptas.

Una vez comparados los tests para el gasto total, para las componentes del gasto se ofrecen, en el cuadro 2, solamente los resultados del test de Waller-Duncan (para resultados más detallados, contactar con la autora).

El cuadro 2 compara el comportamiento de las distintas componentes del gasto a través de la relación entre el abanico de cada variable con la mínima diferencia significativa según el test de Waller-Duncan. Tanto los gastos de referencia semanal como los de referencia superior a la semana ofrecen una variación importante de unas semanas a otras. En el caso de los de referencia semanal (REFSEM), esto es explicable fácilmente, ya que se trata esencialmente del consumo en alimentación, y también lo es que la mayor variabilidad intersemanas esté en los gastos realizados dentro de la semana (DENTRO), pues ello muestra simplemente la estacionalidad del comportamiento de las economías domésticas. Los gastos de referencia superior a la semana (REFSUP) no tendrían en principio que mostrar una variabilidad intersemanal tan acusada. Dos

**Cuadro 2**  
**VARIABILIDAD DE LOS DISTINTOS GASTOS Y DEL INGRESO TOTAL**  
**(INGTOT)**

	GASTOT	REFSEM	REFSUP	DENTRO	FUERA	INGTOT
Mínima dif. signif. según Waller (MDS)	233.914	102.129	170.139	111.748	166.437	244.891
Valor máximo (MAX) (semana)	3.061.510 (11)	1.384.383 (123)	1.803.876 (11)	1.560.367 (11)	1.591.170 (11)	2.518.313 (122)
Valor mínimo (MIN) (semana)	2.119.794 (51)	968.079 (43)	1.135.968 (42)	1.054.582 (43)	1.054.298 (51)	1.877.791 (42)
MAX – MIN (DIF)	941.716	416.304	667.908	514.785	536.872	640.522
DIF / MDS	4,03	4,08	3,93	4,61	3,23	2,62

pueden ser las causas. La primera, la influencia de la variabilidad intermensual del comportamiento de las familias frente a los gastos de referencia mensual: si dos meses son diferentes, eso puede reflejarse en las semanas que pertenecen a ellos. La segunda es el factor olvido, que hace consignar más exhaustivamente los gastos realizados dentro de esa semana y en las semanas recientes, con lo que las semanas de más consumo arrojarían también mayores valores de esta variable. Esto queda confirmado por la mucho menor variabilidad de la variable FUERA (que resulta de restar a los gastos de referencia superior a la semana los realizados dentro de la semana de referencia). Si esta variable aún sigue mostrando variaciones intersemanales, ello puede ser debido a las mismas causas: estacionalidad de orden superior y olvido, esta vez reflejando las diferencias de las semanas inmediatamente anteriores frente a las de un pasado más remoto.

## 5. EFECTOS SOBRE LAS ESTIMACIONES DE LA DESIGUALDAD

Para observar la posible incidencia de la estacionalidad de los gastos en el cálculo de los índices de desigualdad, se ha descompuesto el índice de Theil por semanas. Se trata de saber cuánta de la desigualdad global medida está causada por diferencias reales (intra) y cuánta se debe a las diferencias en el consumo de unas semanas a otras (inter). Llamando  $i = 1, 2, \dots, l$  a las distintas semanas, con  $N_i$  familias entrevistadas en cada una, utilizaremos la ya clásica descomposición del índice de Theil (Borroach, 1991):

$$T(y; N) = \sum_{i=1}^I T(y^i; N_i) (\mu_i N_i) / (\mu N) + (1 / N) \sum_{i=1}^I N_i (\mu_i / \mu) \log (\mu_i / \mu)$$

El segundo sumando de la derecha ( $B$ ) representa el término de desigualdad atribuible al hecho de que los gastos medios de cada semana son distintos. Si, hipotéticamente, registráramos los gastos de todo el año para cada familia, este término sería cero, y estaríamos midiendo solamente la desigualdad real.

El resultado de la descomposición en nuestro caso es:

$$T(y; N) = 0,2010, \quad B = 0,0023$$

Es decir, un 1,14% de la desigualdad observada se debe a la estacionalidad. Se observa, pues, un efecto bastante moderado.

### 5.1. Análisis de determinadas semanas atípicas

Se han calculado los índices de Theil para submuestras temporales que resultaron ser significativamente diferentes en el análisis de la varianza. Los resultados se ofrecen en el cuadro 3.

**Cuadro 3**  
INDICES DE THEIL DE DIFERENTES SUBMUESTRAS SEMANALES  
SEGUN LA VARIABLE GASTO TOTAL (GASTOT)

Semanas..	11 (mayor gasto)	11 + 122 + 123	51 (menor gasto)	51 + 42 + 43	Total
I. THEIL de submuestra (% VARIAC. RESP. TOTAL)	0,211 (5%)	0,233 (15,9%)	0,1585 (-21,1%)	0,19344 (-3,76%)	0,2010 (0%)

Los gastos en períodos clave son no sólo mayores, sino más variables. Una muestra ideal que recogiera información sobre todo el año compensaría parte de esta variabilidad, aunque es difícil aventurar en qué medida, ya que podemos pensar que la variabilidad real está entre los límites de la sobrevaloración de los períodos de gasto fuerte y la infravaloración de los períodos de menos gasto.

## 5.2. Análisis de las componentes del gasto

Considerando las variables de descomposición del gasto según el período de referencia (REFSEM y REFSUP) y, por otro lado, según el momento de realización haya sido fuera o dentro de la semana de la entrevista (FUERA y DENTRO), se trata ahora de ver cómo influirían estas variables en la posible sobreestimación de la desigualdad.

La descomposición del índice de Theil de estas variables en sus componentes intra e intersemanas es la siguiente (cuadro 4):

**Cuadro 4**  
DESCOMPOSICION DE LOS INDICES DE THEIL

	DENTRO	FUERA	REFSEM	REFSUP	TOT
I. Theil Total	0,2101	0,3131	0,2022	0,3011	0,2010
Intersemanas (% del total)	0,0030 (1,4%)	0,0039 (1,2%)	0,0025 (1,2%)	0,0042 (1,4%)	0,0023 (1,14%)

Este cuadro indica varias cosas. En primer lugar, que la componente intersemanal es moderada. Comparando unas componentes del gasto con otras, se observa que los gastos realizados dentro de la semana de referencia son los más variables de unas semanas a otras, fenómeno que se observa tanto en este cuadro como en el análisis de la varianza, donde esta variable registra el mayor valor del estadístico  $F$  (5,39). Sin embargo, no son los gastos de referencia semanal los que causan todo este «exceso de variabilidad», ya que, de hecho (ver cuadro 1), éstos son menos variables entre semanas que los de referencia superior. Ello está conectado con el hecho de que los gastos de referencia superior a la semana presentan un abanico mucho mayor de unas familias a otras (ver índices de Theil totales en el cuadro 3). La desigualdad social se plasma especialmente en estos gastos.

## 6. COMPARACION CON INGRESOS

Como se ha señalado, los ingresos en la EPF se preguntan una sola vez, durante la tercera entrevista, y son de referencia anual.

Aunque las medias semanales resultan significativamente diferentes ( $p = 0,0001$ ), el análisis de la varianza para los ingresos totales muestra diferen-

cias entre semanas mucho menores que para los gastos de cualquier tipo ( $F = 2,94$ ). De hecho, para los ingresos, la diferencia entre la media semanal más alta y la más baja es igual a 2,6 veces la mínima diferencia significativa obtenida en el test de Waller-Duncan, mientras que para los gastos este abanico es de 4,03 veces la mínima diferencia significativa (ver cuadro 2).

## 7. CONCLUSIONES

Se ha analizado la variación estacional semanal del gasto, y en qué medida esta variación puede afectar a la utilización de los microdatos de la EPF para estudios de distribución. Aplicado el análisis de la varianza se ha encontrado que las muestras correspondientes a las diferentes semanas de entrevista producen estimaciones del gasto medio familiar significativamente diferentes entre sí. Si bien estas diferencias no afectarán al gasto medio agregado, añaden una fuente espúrea de variabilidad a los datos. Ello provocará que, al medir las diferencias entre las familias españolas, una parte de dichas diferencias responderán a que la EPF española tiene un período de entrevista de solamente una semana, provocando resultados diferentes según la semana en la que se entreviste a la familia.

El período que más provoca la estacionalidad es el de las últimas semanas de diciembre y las primeras de enero.

Es interesante observar que los gastos de referencia superior a la semana también registran variabilidad semanal. Una posible explicación de ello es el factor olvido, que hace a los encuestados declarar con más exactitud los gastos realizados más recientemente y olvidar los demás con mayor probabilidad. En efecto, considerando los gastos de referencia superior a la semana (vestido, calzado, etc.) que se han realizado durante la semana de la encuesta, se observa que éstos son, con diferencia, los más variables. Los gastos realizados fuera de la semana de la encuesta son los que muestran menor estacionalidad semanal, pero aun así la diferencia entre las muestras semanales es significativa. Se han señalado dos razones que pueden ser corresponsables de ello. En primer lugar, una vez más el factor olvido, pues las semanas de mayores gastos están cerca. Por ejemplo, los entrevistados en la primera semana de enero declararán más exhaustivamente sus gastos de la última semana de diciembre que los entrevistados tres meses más tarde. En segundo lugar, que los resultados hallados pueden estar recogiendo parte de la estacionalidad de los gastos de referencia mensual, y aun trimestral.

De todo lo anterior se deduce que una encuesta de período de referencia mayor sería más estable y robusta. De hecho, en la mayoría de los países de nuestro entorno el período es quincenal, y a veces mensual.

En cuanto a los posibles efectos del problema aquí analizado sobre las medidas de desigualdad, la descomposición del índice de Theil muestra que estos efectos son muy moderados para el gasto total. Sin embargo, podrían ser mayores para determinadas desagregaciones.

Por último, hay que señalar que se ha analizado solamente una de las fuentes de variabilidad espúrea en los datos de gastos. Otras han sido abordadas en la literatura, como la de la amplitud y exhaustividad del cuestionario, cuya correcta contestación ha mostrado estar correlacionada con el nivel de educación, o la del diferente grado de control sobre los gastos según características socioeconómicas, que correlacionaría el factor olvido con estas variables, y a la postre con la renta misma. Los ingresos son mucho más estables, lo que los hace imprescindibles de cara a su utilización para estudios de Economía del Bienestar. Recientemente se están utilizando los gastos de la EPF para estudios de pobreza y desigualdad. Aunque se han aducido en la literatura argumentos en favor de ello, lo cierto es que una razón que pesa especialmente es el problema de la subestimación de los ingresos. El origen de la EPF como instrumento para elaborar el IPC es, a su vez, una de las causas. Aunque hoy en día el objetivo es medir gastos e ingresos, la atención fundamental se sigue centrando en los gastos. De hecho, aún en algunas partes de la metodología de la EPF 1990-91 se muestra este sesgo cuando se señala: «el objetivo de la encuesta es dar estimaciones del gasto anual de los hogares» (p. 60). Si bien es cierto que es difícil eliminar la reticencia de los encuestados a declarar los ingresos, también lo es que el factor olvido es menor que en los gastos, y la recogida de datos más simple. Sirva este artículo para abundar en la repetidamente señalada necesidad de la mejora de la información sobre los ingresos de la EPF.

**APENDICE: MEDIAS DE LAS SUBMUESTRAS ENTREVISTADAS CADA SEMANA Y TEST DE WALLER-DUNCAN PARA LA VARIABLE GASTO TOTAL FAMILIAR (GASTOT)**

Test de W-D:  $K$ ratio = 100.  $DF$  = 21103.  $MSE$  = 2.753E12.  $F$  = 4.033131.  
Valor crítico de  $t$  = 2.00861. Mínima diferencia significativa = 233914

**MEDIAS SEMANALES**

Media	SEMANA	Media	SEMANA
3061510	11	2398798	111
2795469	122	2394301	115
2706812	123	2389754	92
2674361	12	2387436	14
2653709	124	2377365	54
2633147	101	2373460	63
2598200	13	2364614	102
2588445	121	2360070	93
2533773	74	2344724	72
2515596	23	2342612	35
2513498	52	2340463	41
2503083	34	2339164	53
2492903	94	2338965	61
2480640	112	2335929	85
2471737	21	2331985	103
2462417	81	2294507	55
2454836	44	2287116	32
2444613	62	2280978	82
2434908	104	2278984	91
2428651	113	2274362	24
2425474	84	2263840	83
2425178	71	2186545	64
2407982	114	2178405	43
2405125	33	2169657	73
2400670	31	2155541	42
2399079	22	2119794	51

**REFERENCIAS**

- BORROACH, V. K. (1991): «Measures of Inequality and Poverty», *Regional Income Inequality and Poverty in The U.K.*, Hants, Dartmonth.
- DUNCAN, D. B. (1975): «t-Tests and Intervals for Comparisons Suggested by the Data», *Biometrics* (31): 339.
- INE: «Encuesta de Presupuestos Familiares 1990-91. Metodología».
- MOOD, A. M., y GRAYBILL, F. A. (1972): *Introducción a la teoría de la Estadística*, Madrid, Aguilar.
- RICE, J. A. (1988): *Mathematical Statistics and Data Analysis*, California, Wadsworth and Brooks/Cole.
- SANZ, B. (1992): «La Encuesta de Presupuestos Familiares 1990-1991», *Situación*, BBV.
- SAS/STAT User's guide, Versión 6.03. Cary, NC: SAS Institute Inc., 1988.
- WALLER, R. A., y DUNCAN, D. B. (1969): «A Bayes Rule for the Symmetric Multiple Comparison Problem», *Journal of the American Statistical Association* (64): 1488.
- WALLER, R. A., y KEMP, K. E. (1975): «Computations of Bayesian t-Values for Multiple Comparisons», *Journal of Statistical Computation and Simulation* (75): 169.

**DISPERSION AMONG WEEKS OF THE EXPENDITURES  
AT THE SPANISH FAMILY BUDGET SURVEY  
(ENCUESTA DE PRESUPUESTOS FAMILIARES)**

**SUMMARY**

The period of the interview at the EPF is one week. This is a source of problems when using these data for the analysis of personal distribution of expenditure: some of the total variability observed would be due to real inequality, while some of it would be spurious, caused by the different week of the interview. In this paper, this hypothesis is tested by the Analysis of Variance. In order to find out which comparisons of weekly means are significant, the tests of Bonferroni, Tukey and Waller-Duncan are applied. The effect of this problem on the measures of inequality are assessed by decomposing the Theil index.

*Key Words:* expenditures an the EPF, ANOVA, Waller-Duncan test.

*AMS Classification:* 62J10, 62P20.