

Influencia de las sustituciones en la calidad de los datos en la encuesta de condiciones de vida y trabajo en España (ECVT)

por

SANTIAGO MURGUI

Universidad de Valencia

JUAN MURO

Universidad de Alcalá de Henares

EZEQUIEL URIEL

Universidad de Valencia

RESUMEN

Se analiza la influencia del elevado número de sustituciones en las viviendas inicialmente seleccionadas para la entrevista sobre la calidad de los datos de la encuesta de condiciones de vida y trabajo en España (ECVT). Dos aspectos fundamentales de la encuesta son tomados en consideración: el grado de aleatoriedad de la muestra efectivamente entrevistada y su representatividad respecto del conjunto de la población española en las edades contempladas.

Palabras clave: ECVT; calidad de datos; aleatoriedad de muestras; encuestas por muestreo.

Clasificación AMS: 62D05, 62P25.

AGRADECIMIENTOS

Los autores desean agradecer los comentarios del editor asociado y de un evaluador anónimo, que han contribuido a la mejora del presente artículo.

1. INTRODUCCION

La ECVT fue realizada durante los meses de octubre y noviembre de 1985 en todo el territorio del Estado, con la excepción de Ceuta y Melilla. Como se sabe, sus objetivos principales consistían en analizar el mercado de trabajo y completar la información obtenida periódicamente a través de los procedimientos institucionalizados-Encuesta de Población Activa (EPA) del INE y Registros en las Oficinas de Empleo del INEM (Paro registrado).

El Universo a investigar se refería a la población de catorce años o más residente en viviendas familiares. Para ello se seleccionó una muestra de 63.120 elementos, distribuidos en 3.156 secciones censales, efectuándose veinte entrevistas por seccion. El diseño muestral, de características similares al de la EPA, constaba de una doble estratificación de las secciones de cada provincia, atendiendo al número de habitantes del municipio al que pertenecían y a características socioeconómicas de la población. Dentro de cada estrato se seleccionaron las secciones con probabilidad proporcional al número de viviendas familiares que poseían. En cada sección se seleccionaron aleatoriamente veinte viviendas y en cada una de ellas se entrevistó a un miembro elegido al azar de entre los de catorce años o más. En cuanto al número de entrevistas a realizar por provincia, se eligió una solución de compromiso entre la uniforme y la proporcional, tendiendo a la proporcionalidad en la distribución entre los estratos de cada provincia. Para mayor detalle puede consultarse la publicación: *Condiciones de Vida y Trabajo en España (1986)*.

Es conocido que uno de los capítulos más importantes en el análisis de los resultados de una encuesta es el dedicado a investigar la calidad de los datos recogidos.

Uno de los factores determinantes en la calidad de los datos de la ECVT, es el elevado número de sustituciones que fueron llevadas a cabo, por diversas incidencias del trabajo de campo, en las viviendas inicialmente seleccionadas para la entrevista. En este artículo se analiza la influencia que tuvieron dichas sustituciones en dos aspectos fundamentales de la encuesta: el grado de aleatoriedad de la encuesta efectivamente entrevistada y su representatividad respecto del conjunto de la población española en esas edades.

2. ALEATORIEDAD EN LA SELECCION DEL MIEMBRO A ENTREVISTAR EN CADA VIVIENDA

La tercera etapa del diseño muestral de la ECVT, exigía la selección de un miembro de entre los que teniendo catorce años o más residían en las viviendas en las que se llevaron a cabo las entrevistas. Ordenados tales miembros por su edad en cada vivienda, con el fin de facilitar el proceso, se elaboró una tabla de selección, en la que según el número de miembros de catorce años o más que la habitaban, y partiendo de una fila aleatoria entre 1 y 14, la tabla indicaba el lugar que ocupaba en la lista el miembro a entrevistar. En el cuadro 1 se recoge la tabla, debida a Leslie Kish, utilizada en la ECVT para la selección de miembros de un hogar.

CUADRO 1

Tabla de seleccion aleatoria para determinar la persona a entrevistar

Fila	Numero total de individuos seleccionables						
	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1
3	1	1	1	1	1	2	2
4	1	1	1	1	2	2	2
5	1	1	2	2	2	3	3
6	1	1	2	2	2	3	3
7	1	2	2	3	3	4	4
8	1	2	2	3	3	4	4
9	1	1	2	2	3	5	5
10	1	2	3	3	4	5	5
11	1	2	3	4	4	5	6
12	1	2	3	4	4	6	6
13	1	2	3	4	5	6	7
14	1	2	3	4	5	6	7

Entrevistador: Una vez seleccionada la persona a entrevistar, proceda a realizar la entrevista, aplicando el cuestionario que figura a continuación.

Los resultados de la selección obtenidos por este procedimiento han sido recogidos en el cuadro 2. En dicho cuadro se han suprimido 73 viviendas que fueron tabuladas erróneamente, tal es el caso de viviendas en las que teniendo tres miembros con catorce o más años fue seleccionado el que ocupaba el lugar

seis de la lista. Se han despreciado 382 viviendas para las que no consta el número de orden en la lista del miembro entrevistado. Finalmente, no se han considerado en la tabla las viviendas con más de nueve miembros en la lista, puesto que su número no permite garantizar una mínima fiabilidad en su análisis.

CUADRO 2

N.º de miembros	N.º de orden									Total	
	1	2	3	4	5	6	7	8	9		
1	7.105										7.105
2	12.127	13.054									25.181
3	3.640	4.727	4.329								12.696
4	2.432	2.125	1.814	2.484							8.855
5	878	1.069	840	884	619						4.290
6	225	273	241	236	304	337					1.616
7	94	85	68	90	81	64	87				569
8	46	56	1	11	0	5	25	31			175
9	21	20	7	1	0	0	0	3	14		66
											60.553

Las causas por las que los resultados de la selección en cada tipo de vivienda no se ajustan a una distribución uniforme son dos. Por una parte está el hecho de que la utilización de la tabla de selección propuesta introducía ciertos sesgos, en la medida en que los miembros que ocupaban ciertos números de orden en la lista estaban sobrerrepresentados y otros en cambio infrarrepresentados. La otra causa, hay que atribuirle a la gran cantidad de sustituciones que debieron efectuarse sobre las viviendas originalmente seleccionadas, unido a las posibles manipulaciones intencionadas de los entrevistadores. A continuación se estudia la incidencia de cada una de estas causas por separado.

En el cuadro 3 se han recogido, expresadas en porcentajes la probabilidad de selección de cada miembro de la lista, para los distintos tipos de vivienda, derivadas de la aplicación de la tabla de selección utilizada.

CUADRO 3

N.º de miembros	N.º de orden								
	1	2	3	4	5	6	7	8	9
1	100								
2	50	50							
3	28,57	35,71	35,71						
4	28,57	21,43	21,43	28,57					
5	21,43	21,43	21,43	21,43	14,28				
6	14,28	14,28	14,28	14,28	21,43	21,43			
7	14,28	14,28	14,28	14,28	14,28	14,28	14,28		
8	14,28	10,71	10,71	14,28	14,28	10,71	10,71	14,28	
9	10,71	10,71	10,71	10,71	7,14	14,28	10,71	10,71	14,28

El cuadro 3 permite concluir que la equiprobabilidad fue alterada en todos los tipos de vivienda, excepto en las de dos y siete miembros de catorce o más años. La desviación máxima es de 7,14 % de probabilidad. Esto justificaría por ejemplo que en las 12.696 viviendas con tres miembros en la lista, el primero fuera seleccionado por término medio 906 veces menos que los otros dos, lo que significa una desviación de 604 respecto de las 4.232 veces en que por término medio sería seleccionado cada uno de los tres miembros si la distribución fuera la uniforme.

Puesto que las viviendas con un número de miembros de catorce o más años superior a cuatro únicamente representan un 11 % del total, una forma de reducir las desviaciones respecto de la equiprobabilidad consistiría en construir una tabla de selección con 12 ó 24 filas en lugar de las 14 que forman la tabla utilizada. De esta forma se garantizaría la equiprobabilidad entre los miembros de las viviendas con dos, tres, cuatro y seis miembros de catorce o más años. El error sería menor con 24 filas, por estar este número más próximo a un múltiplo de cinco. Sin embargo, si dicho número se prestara a una mayor confusión en el proceso de selección, sería preferible tomar 12 filas.

Utilizando el test de la X^2 y el número de viviendas entrevistadas que aparece en el cuadro 2, se ha contrastado estadísticamente si el modelo aleatorio del cuadro 3 difiere del modelo que supone la uniformidad para cada tipo de vivienda. Se concluye que no puede aceptarse la equivalencia entre ambos modelos, ni siquiera disminuyendo el nivel de significación hasta límites aceptables.

La segunda parte del análisis se orienta a comprobar si los resultados de la selección se ajustan al modelo proyectado que viene expresado en el cuadro 3. Con ello se pretende detectar las posibles anomalías cometidas en el proceso

de selección, bien por las sustituciones, bien por errores cometidos por parte de los entrevistadores.

Recurriendo al test de la X^2 para cada tipo de vivienda se han calculado los siguientes estadísticos.

CUADRO 4

N.º de miembros	2	3	4	5	6	7	8	9
χ^2	34,12	17,53	35,58	34,49	13,97	9,34	154,89	86,79

Considerando los niveles de significación usuales, los valores del cuadro 4 inducen a concluir que únicamente los resultados de la selección en las viviendas con tres, seis y siete miembros de catorce o más años, se ajustan al modelo propuesto. Para el resto de las viviendas, debe admitirse que ciertas causas ajenas al azar han intervenido en el proceso.

Con el fin de investigar las causas que han podido alterar los resultados de la selección, se han calculado los índices de variación relativa para cada tipo de vivienda y cada posible miembro a seleccionar. Tales índices se han determinado por cociente entre las diferencias de las frecuencias reales a las inducidas por el modelo y éstas últimas. Sus valores se recogen en el cuadro 5.

CUADRO 5

N.º de miembros	N.º de orden								
	1	2	3	4	5	6	7	8	9
1	0								
2	-0,0368	+0,0368							
3	+0,0035	+0,0425	-0,0453						
4	-0,0387	+0,1199	-0,0440	-0,0182					
5	-0,0450	+0,1629	-0,0862	-0,0384	+0,0100				
6	-0,0254	+0,1825	+0,0439	-0,0223	-0,1235	-0,0284			
7	+0,1565	+0,0458	-0,1634	+0,1073	-0,0034	-0,2126	+0,0704		
8	+0,8400	+1,9867	-0,9467	-0,5600	-1	-0,7333	+0,3333	+0,2400	
9	+1,9703	+1,8288	-0,0099	-0,8585	-1	-1	-1	-1	-0,5757

De la observación del cuadro 5 se deduce que en todo tipo de vivienda está sobrerrepresentado el segundo miembro de la lista, siendo mayor la sobrerrepre-

sentación en las viviendas con mayor número de miembros susceptibles de ser entrevistados.

Puesto que el número de viviendas no es el mismo para cada tipo, es conveniente la construcción de unos índices globales que expresen la variación relativa asociada a cada miembro de la lista. Para su cálculo se ha procedido a ponderar los índices del cuadro 5 con los números de viviendas que aparecen en el cuadro 2. Los índices que resultan son recogidos en el cuadro 6.

CUADRO 6

N.º orden	1	2	3	4	5	6	7	8	9
Índice	-0,0204	+0,0750	-0,0539	-0,0246	-0,0595	-0,1489	+0,0400	+0,0166	+0,4846

Cada índice global expresa la tendencia detectada a entrevistar por exceso o por defecto al respectivo miembro de la lista.

Teniendo en cuenta la poca significación de los índices asociados con los órdenes superiores, por haber sido calculados sobre un conjunto reducido de datos, puede concluirse que ha existido tendencia a entrevistar por exceso al segundo miembro de cada lista, y esto a costa de entrevistar menos a los demás.

Entre las razones que puede justificar tal conclusión debe destacarse la posible asociación entre el segundo miembro de cada lista y el ama de casa. Tal asociación induce a pensar que por estar esta persona más tiempo en la vivienda, puede haber sido objeto de un número mayor de entrevistas que los demás miembros. En esta tendencia puede haber influido fundamentalmente el elevado número de sustituciones de viviendas que se llevaron a cabo. Cabe tener en cuenta que si el miembro seleccionado era el primero de la lista y asociamos dicho miembro con el cabeza de familia, cuando estuviera ausente durante varias visitas a la vivienda, se sustituiría la misma por otra. En cuanto que la nueva vivienda podría tener un número de miembros distinto, el miembro seleccionado podría ser otro distinto al primero. En sucesivas sustituciones, el proceso tendería a detenerse cuando en una vivienda fuera seleccionada el ama de casa, con lo cual realmente se estaría sustituyendo un miembro cualquiera de la lista por el segundo.

Con el fin de apoyar esta interpretación de los resultados, se han calculado en el cuadro 7, los índices porcentuales de cada sexo para cada miembro de la lista.

CUADRO 7

N.º orden	Varones	Mujeres
1	68,71	31,29
2	23,31	76,69
3	50,57	49,43
4	49,11	50,89
5	51,96	48,04
6	51,02	48,98
7	54,08	45,92
8	50	50
9	44,66	53,34

En el cuadro 8 se ha calculado porcentualmente la distribución por edad y sexo de los segundos miembros de cada lista.

CUADRO 8

Edad	Varones	Mujeres
14 ó 15	50	50
15-19	40,4	59,6
20-24	30,8	69,2
25-29	24,3	75,7
30-34	20,8	79,2
35-39	21,9	78,1
40-44	21,5	78,5
45-49	23,2	76,8
50-54	25,3	74,7
55-59	24,9	75,1
60-64	23,5	76,5
65-69	19	81
70 o más	20,5	79,5

Se observa que el porcentaje del 76 % de mujeres, global para el segundo miembro se alcanza a partir de los veinticinco años y se mantiene en valores superiores en las edades más altas. Puede concluirse de los cuadros 7 y 8, que el segundo miembro de las listas es generalmente una mujer de edad superior a veinticinco años, por lo que resulta lógico asociar dicho miembro con el ama de casa.

3. REPRESENTATIVIDAD DE LA MUESTRA

En el apartado anterior, se concluye que debido al elevado número de sustituciones, existe sobre e infrarrepresentación de algunos miembros de la lista de cada vivienda. Si tenemos además en cuenta el cuadro 7 y los equivalentes al 8 para los demás miembros, se comprueba que la distribución por edad y sexo no es homogénea para todos los miembros. Debe esperarse por lo tanto que la distribución respecto a estas variables de todo el conjunto muestral, difiera de la que se hubiera obtenido en el caso de haber seleccionado la muestra según el diseño proyectado.

Tomando como referencia los datos del Censo de Población de 1981, se corrigieron las desviaciones anteriores modificando los factores de elevación de cada estrato, de manera que la inferencia deducida de la muestra en lo que respecta a las variables edad y sexo se ajustan a la población referenciada.

En síntesis, el proceso de equilibraje aplicado en la ECVT consistió en considerar cinco grupos de edad para los varones y otros cinco para las mujeres. A continuación y de manera iterativa se aplicó un factor de elevación para cada uno de los diez grupos, de manera que la muestra quedara equilibrada. Los símbolos con los que denotaremos a cada grupo, así como las edades que comprenden se recogen a continuación.

Varones		Mujeres	
Grupo	Edad en años	Grupo	Edad en años
V1	14-19	M1	14-19
V2	20-24	M2	20-24
V3	25-44	M3	25-44
V4	45-64	M4	45-64
V5	65 y más	M5	65 y más

Este procedimiento de corrección adolece de graves inconvenientes en la medida que no controla los efectos que puede producir en las variables distintas al sexo y edad. El problema así planteado no presenta posibilidades de solución, puesto que a priori es imposible reequilibrar la muestra respecto a las variables que precisamente se pretendía investigar.

El objetivo que se persigue a continuación es determinar la magnitud de las correcciones efectuadas sobre la muestra con el fin de equilibrarla. Son conocidos los factores de elevación que realmente se aplicaron, comparando tales factores con los que se debía haber aplicado en el supuesto de que la muestra

seleccionada se ajustara al criterio de aleatoriedad previsto en el diseño, obtendremos una aproximación de los cambios que fue necesario introducir.

De no haber tenido en cuenta el equilibraje, el factor de elevación debería ser único para cada estrato, y definido por el cociente entre el tamaño poblacional del estrato y el tamaño de la muestra seleccionada en el mismo. Las desviaciones entre tales factores teóricos y los realmente aplicados expresaron las correcciones que fue necesario introducir en cada estrato para equilibrarlo. La magnitud de tales correcciones permitirá asimismo cuantificar el grado de representatividad de la muestra.

Si expresamos por N_e la población de catorce o más años en el estrato e , y por n_e el tamaño de la muestra seleccionada en el mismo, el factor de elevación teórico asociado a dicho estrato vendrá determinado por $ft_e = N_e/n_e$.

Utilizando el paquete informático SPSS, se han calculado los factores de elevación aplicados a cada grupo de edad y sexo, y esto en cada uno de los estratos.

Denotaremos por f_{ie} el factor aplicado al grupo i del estrato e . Los productos $f_{ie} \times n_{ie}$, en el que n_{ie} expresa el tamaño de la muestra del grupo i del estrato e proporcionan una estimación N_{ie} a la población de catorce o más años del grupo i en el estrato e .

Los factores correctores para cada grupo de edad por sexo i del estrato e , pueden definirse por:

$$c_{ie} = f_{ie}/ft_e$$

y se han calculado para todos los estratos y grupos.

Los factores correctores de cada grupo pueden agregarse a nivel provincial, utilizando como ponderaciones las estimaciones N_{ie} . El factor asociado al grupo i y la provincia p , vendrá determinado por:

$$c_{ip} = \sum c_{ie} N_{ie} / \sum N_{ie}$$

Donde los sumatorios se extienden a todos los estratos presentes en la provincia p .

En los cuadros 9 y 10 se han calculado los factores correctores de cada provincia y grupo.

CUADRO 9

Provincia	14 - 19 años	20 - 24 años	25 - 44 años	45 - 64 años	65 y más años
Alava	1,38	1,00	1,24	1,25	0,73
Albacete	2,42	1,75	1,13	1,16	0,56
Alicante	1,44	2,04	1,09	1,31	0,80
Almería	1,41	2,19	1,32	0,93	0,89
Ávila	2,06	2,41	1,19	0,99	0,63
Badajoz	1,92	1,25	1,25	1,19	0,61
Baleares	1,80	1,76	1,35	1,09	0,65
Barcelona	1,37	1,33	1,14	1,24	0,72
Burgos	1,58	1,71	1,22	0,99	0,58
Cáceres	1,52	1,67	1,22	1,03	0,71
Cádiz	1,47	1,07	1,04	0,94	0,76
Castellón	1,50	1,41	1,24	1,18	0,65
Ciudad Real	1,74	1,54	1,21	1,15	1,05
Córdoba	1,47	1,67	1,12	1,14	0,70
Coruña	1,48	2,22	1,35	1,11	0,71
Cuenca	2,52	1,44	1,09	1,05	0,88
Gerona	1,73	1,67	1,06	1,08	0,69
Granada	1,45	2,05	1,40	1,16	0,65
Guadalajara	1,50	1,29	1,07	1,22	0,79
Guipúzcoa	1,53	1,89	1,17	1,16	0,79
Huelva	1,37	1,34	1,07	1,04	0,67
Huesca	1,96	2,12	1,22	1,59	0,76
Jaén	1,78	1,65	1,09	1,12	0,63
León	2,26	1,64	0,96	1,09	0,78
Lérida	1,63	2,05	1,15	1,04	0,85
La Rioja	1,50	2,66	1,18	1,13	0,64
Lugo	1,75	1,96	1,37	0,95	0,72
Madrid	1,56	1,52	1,26	1,16	0,63
Málaga	1,43	1,65	1,28	0,93	0,74
Murcia	1,45	1,73	1,10	1,10	0,74
Navarra	1,41	1,38	1,19	1,22	0,80
Orense	1,85	1,87	1,76	1,05	0,64
Asturias	1,31	1,56	1,19	1,01	0,80
Palencia	2,36	2,14	1,23	1,26	0,68
Las Palmas	1,40	1,54	1,50	1,15	0,59
Pontevedra	1,54	2,11	1,42	1,09	0,74
Salamanca	1,90	2,05	1,07	1,35	0,80
Tenerife	1,22	1,60	1,10	1,31	0,72
Cantabria	1,53	1,94	1,21	1,15	0,70
Segovia	1,90	1,23	1,31	0,94	0,72
Sevilla	1,52	1,25	1,33	1,05	0,73
Soria	1,33	1,42	1,50	1,15	0,76
Tarragona	1,10	2,03	1,19	1,17	0,77
Teruel	2,31	2,72	1,07	1,05	0,67
Toledo	1,61	1,65	1,25	1,13	0,67
Valencia	1,21	1,45	1,43	1,02	0,66
Valladolid	1,61	2,13	1,13	1,19	0,60
Vizcaya	1,40	1,14	1,43	1,21	0,65
Zamora	2,20	1,97	1,64	1,08	0,61
Zaragoza	1,67	1,24	1,08	1,11	0,65

CUADRO 10

Provincia	14 - 19 años	20 - 24 años	25 - 44 años	45 - 64 años	65 y más años
Alava	1,51	1,26	0,88	0,78	0,65
Albacete	2,16	1,38	0,83	0,86	0,68
Alicante	1,22	1,28	0,93	0,80	0,60
Almería	1,43	1,52	0,82	0,80	0,68
Ávila	1,85	2,44	0,88	0,90	0,66
Badajoz	1,52	1,69	0,87	0,79	0,71
Baleares	1,16	1,57	1,03	0,88	0,57
Barcelona	1,31	1,36	0,93	0,85	0,65
Burgos	1,77	1,58	1,01	0,88	0,66
Cáceres	1,85	1,42	0,89	0,92	0,62
Cádiz	1,90	1,70	0,92	0,88	0,60
Castellón	1,34	1,42	0,99	1,04	0,57
Ciudad Real	1,42	1,18	0,86	0,81	0,58
Córdoba	1,56	1,29	0,87	0,91	0,62
Coruña	1,20	1,10	0,85	0,95	0,65
Cuenca	1,28	1,90	0,91	0,93	0,62
Gerona	2,30	1,38	0,85	0,97	0,68
Granada	1,57	1,09	0,83	0,84	0,59
Guadalajara	1,47	1,30	0,95	0,85	0,74
Guipúzcoa	1,27	1,31	0,78	0,87	0,66
Huelva	1,09	1,14	1,01	1,03	0,74
Huesca	1,49	1,48	1,13	0,83	0,73
Jaén	2,07	1,09	0,91	0,83	0,66
León	1,78	1,40	0,90	0,85	0,64
Lérida	2,17	2,08	0,84	0,86	0,62
La Rioja	1,66	1,45	0,91	0,91	0,61
Lugo	2,08	1,29	1,14	0,92	0,64
Madrid	1,42	1,27	0,90	0,83	0,61
Málaga	1,35	1,22	0,93	0,84	0,65
Murcia	1,45	1,65	0,89	0,86	0,60
Navarra	1,27	1,33	0,94	0,81	0,64
Orense	1,39	1,75	1,05	0,97	0,52
Asturias	1,82	1,47	0,91	0,92	0,66
Palencia	1,74	1,13	0,87	0,84	0,59
Las Palmas	1,34	1,14	0,82	0,78	0,55
Pontevedra	1,34	0,98	0,84	0,90	0,62
Salamanca	1,58	1,70	0,79	0,86	0,63
Tenerife	1,67	1,12	0,85	0,81	0,68
Cantabria	1,59	1,27	0,83	0,96	0,60
Segovia	2,66	1,35	0,84	1,01	0,62
Sevilla	1,58	0,94	0,89	0,87	0,61
Soria	2,12	1,98	0,92	0,86	0,61
Tarragona	1,52	1,44	0,94	0,88	0,61
Teruel	2,73	1,38	0,77	1,02	0,74
Toledo	1,67	1,09	0,93	0,85	0,68
Valencia	1,27	1,27	1,01	0,91	0,58
Valladolid	1,56	1,28	0,83	0,90	0,65
Vizcaya	1,31	1,13	0,99	0,79	0,54
Zamora	1,88	1,48	1,06	1,08	0,51
Zaragoza	1,55	1,51	1,10	0,97	0,58

Los recorridos más amplios entre factores corresponden a las provincias de La Rioja, Segovia y Teruel, con valores respectivos 2,05, 2,04 y 2,06. Como medida del grado de dispersión de los factores de corrección para cada provincia puede utilizarse la raíz cuadrada del promedio de las desviaciones cuadráticas respecto de uno.

A nivel global, para el conjunto de toda la muestra, los factores, de corrección de cada uno de los diez grupos considerados se han calculado en el cuadro 11.

CUADRO 11

Grupo	V1	V2	V3	V4	V5	M1	M2	M3	M4	M5
F. C:	1,53	1,61	1,24	1,13	0,71	1,50	1,33	0,91	0,87	0,63

Se desprende de estos valores que en general las mujeres fueron sobrerrepresentadas frente a los hombres. El grupo más representado correspondió a las mujeres (M5) con 65 o más años y el menos representado al de los hombres (V2) entre 20 y 24 años. Resulta interesante observar el salto que se produce entre los grupos de mujeres M2 y M3, cuya frontera está en los 25 años. Esto realmente viene a confirmar la conclusión ya citada de que el miembro más representado de cada vivienda fue el ama de casa.

BIBLIOGRAFIA

Secretaría General de Economía y Planificación (1986). *Condiciones de Vida y Trabajo en España*. En colaboración con el Centro de Investigaciones Sociológicas. Madrid.

SUMMARY

ON THE EFFECTS OF THE SUBSTITUTION PROCESS UPON THE QUALITY OF DATA IN THE «WORKING AND LIVING CONDITIONS SURVEY» (ECVT) IN SPAIN

The effects of the high level of substitutions in the households initially chosen for the interview on the quality of data in the ECVT are analyzed. Two relevant features of the survey are considered: the degree of randomness of the sample effectively interviewed and its representativeness in relation to the spanish population in the ages taken into account.

Key words: ECVT; quality of data; randomness of sample; sampling surveys.

AMS Classification: 62D05, 62P25.

