

La selección de los entrevistados últimos en encuestas presenciales: un análisis de la utilización conjunta del método de rutas y el método de cuotas¹

The selection of the last interviewees in face-to-face opinion polls: an analysis of the joint use of the route method and the quota method

Vidal Díaz de Rada

Universidad Pública de Navarra

vidal@unavarra.es

Palabras clave: Muestreo Aleatorio por Conglomerados, Método de Rutas, Método de Cuotas, Entrevistados, Centro de Investigaciones Sociológicas.

Keywords: Multistage Cluster Sampling, Route Method, Quota Sampling, Interviewee, Centro de Investigaciones Sociológicas.

RESUMEN

Los textos sobre muestreo presentan los métodos de selección muestral disponibles, así como las características que debe cumplir el universo objeto de estudio para la selección del método más adecuado. Todos ellos señalan la necesidad de tener un conocimiento previo del universo antes de llevar a cabo cualquier proceso de selección muestral; lo que implica que debe disponerse de alguna fuente que permita conocer en detalle el ámbito objeto de estudio. Este requerimiento presenta algunas dificultades cuando se trata de realizar muestras a la población

ABSTRACT

All texts about sampling present the available methods for sample selection and the characteristics that the universe which is the object of study must possess, in order for the most suitable method to be selected. All of them point to the need to have prior knowledge of the universe before carrying out any sampling. This implies that there has to be some source that provides detailed knowledge of the object of study domain. This requirement poses some difficulties when the samples are selected from the population in general. These difficulties are due

¹ Una primera versión de este trabajo ha sido presentada en el IX Congreso Español de Sociología, celebrado en Barcelona del 13 al 15 de septiembre de 2007. El autor desea agradecer la ayuda concedida por el CIS, dentro de su programa *Ayudas a la Investigación Sociológica* en el año 2005, para la explotación de los barómetros del año 2004. Deseamos expresar también nuestro agradecimiento a Adoración Núñez Villuendas, de la Unidad de Campo del CIS. Este artículo es una profundización de un capítulo incluido en un trabajo más amplio realizado junto con Adoración Núñez Villuendas y publicado en la colección *Monografías (Estudio de las incidencias en la investigación con encuesta: el caso de los barómetros del CIS)*. Gracias a Marina Sagaseta, del IEN, por la información sobre la Encuesta Navarra de Condiciones de Vida. Agradecemos también las sugerencias recibidas por dos revisores anónimos, así como por el Consejo Editorial de la REIS.

general. Estas dificultades se producen por el limitado acceso a los documentos que detallan las características de la población que reside en un determinado ámbito, así como por las inexactitudes de tales documentos. Por este motivo, las muestras realizadas a población general utilizan conjuntamente varios métodos de selección muestral. En este trabajo se realiza una *evaluación* del muestreo por cuotas aplicado a los barómetros realizados por el Centro de Investigaciones Sociológicas en el año 2004.

both to the limited access to documents reporting the characteristics of the population living in a specific area, and to the inaccuracies that these documents contain. This is why the samples drawn from the population in general use various methods of sampling jointly. In this paper, an evaluation of the quota sampling applied to the 2004 Barometers of the Sociological Research Centre is carried out.

Vidal Díaz de Rada

Doctor en Sociología por la Universidad de Deusto. Actualmente es Profesor Titular de Sociología en el Departamento de Sociología de la Universidad Pública de Navarra.

He gained his Ph.D. in sociology at the University of Deusto, and is currently Associate Professor of Sociology at the Department of Sociology of the Universidad Pública de Navarra.

Departamento de Sociología. Universidad Pública de Navarra. Campus de Arrosadía. 31006 Pamplona. Spain.

La práctica totalidad de los textos sobre muestreo, al presentar los métodos de selección muestral y las características que debe cumplir el universo objeto de estudio para el empleo del método más adecuado, coinciden en que se debe disponer de un conocimiento previo del universo antes de llevar a cabo cualquier proceso de selección muestral. Esto implica que, en numerosas ocasiones, sea necesario consultar algún *documento* (marco muestral) que permita identificar y conocer en detalle el ámbito objeto de estudio.

Este requerimiento, tan fácil de cumplir en la teoría, presenta algunas dificultades cuando se trata de realizar muestras a la *población general*, a la población que reside en viviendas. Estas dificultades se producen, por un lado, por el limitado acceso a los documentos que detallan las características de la población que reside en un determinado ámbito y, por otro, por las limitaciones e inexactitudes de tales documentos. Por este motivo, las muestras realizadas a población general utilizan conjuntamente varios métodos de selección muestral; algunos de los cuales recaen —en las últimas etapas— en los propios entrevistadores². En este trabajo se lleva a cabo una *evaluación* del método de cuotas —utilizado tras seleccionar las viviendas con el método de rutas— considerando su capacidad para reflejar adecuadamente el universo de donde se ha extraído la muestra.

1. INTRODUCCIÓN. EL PROCESO DE SELECCIÓN MUESTRAL EN LAS ENCUESTAS A POBLACIÓN GENERAL

Las muestras *probabilísticas* son las únicas que proporcionan resultados extrapolables, y se caracterizan porque se conoce la probabilidad que tiene cada uno de los miembros del universo para ser seleccionado. Para que una muestra sea probabilística es preciso que la aleatorización guíe todo el proceso de selección, y que se utilice un procedimiento que respete la probabilidad que tiene cada elemento de pertenecer a la muestra. La aleatoriedad debe respetarse no sólo en el diseño muestral, sino que debe guiar todo el proceso de recogida de información, especialmente durante el trabajo de campo. Es ahí donde aparecen más dificultades y pueden generarse mayores sesgos si no se respetan rigurosamente las instrucciones proporcionadas para la selección de los entrevistados (Rodríguez Osuna, 1991: 13).

De modo que el muestreo probabilístico se caracteriza porque se conoce la probabilidad que tiene cada miembro del universo para salir seleccionado en la muestra, condición imprescindible para proceder a la inferencia estadística y así estimar las características del

² Pese al *protagonismo* del entrevistador, el poder «decisorio» de éste es muy limitado, como tendremos ocasión de demostrar en las siguientes páginas.

universo a través de la muestra seleccionada. Ahora bien, para conocer esta probabilidad es necesario saber el tamaño del universo y su relación con el tamaño muestral³. Para poder efectuar este cálculo debe definirse con precisión el universo objeto de estudio, en la medida que limita el objeto de la investigación a un determinado entorno sociogeográfico. Para ello se procede a localizar y concretar de forma precisa el objeto del análisis, mencionando las características que lo ubican en un determinado espacio y en un tiempo concreto: los estudios de la extinta Fundación CIRES analizan la población de 18 y más años que reside en España; las «encuestas juventud», los jóvenes españoles entre 15 y 24 años (González Blasco *et al.*, 2006: 405); *British Social Attitudes*, la población de 18 y más años que reside en hogares privados (Park *et al.*, 2001: 263); la Encuesta General de Población del NORC de la Universidad de Chicago (*General Social Survey*), los adultos (mayores de 18 años) que viven en hogares en los Estados Unidos (Davis y Smith, 1992: 31)..., por citar algunas investigaciones de amplia difusión. De estos ejemplos se desprende que la población o universo es «un conjunto de unidades limitadas en el espacio, en el tiempo y en las características pertinentes, y que son el objeto de la investigación» (Azorín y Sánchez-Crespo, 1986: 17).

Delimitado el universo, a continuación se localiza un documento que permita identificar y conocer en detalle esta población, conocido como *marco muestral*. El marco muestral es un listado que contiene los elementos de la población objeto de estudio; acotando el universo y las unidades que lo componen (Czaja y Blair, 1996: 116). La estructura del marco, la información contenida y la calidad de esta información determinarán el tipo de muestreo a utilizar, así como los procedimientos de estimación utilizados en la investigación (Lessler y Kalsbeek, 1992: 40). En la medida que el conocimiento del universo sea más perfecto, la equivalencia entre el marco muestral y el universo será mayor, siendo posible reducir determinados errores susceptibles de cometerse en la selección muestral.

Cuando el universo objeto de estudio son colectivos especializados es usual la utilización de directorios profesionales, pero ¿cómo se opera cuando el objeto de estudio es la población general? En estos casos la información sobre el universo suele tomarse de los Censos y Padrones de Población, documentos que pueden presentar algunas deficiencias. Los Censos, por ejemplo, no reflejan adecuadamente la estructura de un territorio cuando se trata de poblaciones con gran movilidad, puesto que la revisión y la renovación del Censo se hacen cada diez años. La ausencia de un marco muestral actualizado y detallado lleva, en numerosas ocasiones, a utilizar muestreos por conglomerados.

³ Así, por ejemplo, la posibilidad de incluir a un español de 18 y más años en una muestra de 2.500 personas es de 0,00696%, cifra resultante de dividir el tamaño de la muestra entre el universo objeto de estudio; la totalidad de la población española mayor de 17 años, en este caso (2.500/35.895.962).

El muestreo por conglomerados se caracteriza porque el universo objeto de estudio no son los elementos individuales del universo, sino un conjunto de elementos que —bajo determinadas circunstancias— puede considerarse que forman una unidad; un conglomerado. Algunos ejemplos, extraídos del clásico libro de Kish, son los empleados de una gran empresa clasificados por «equipos» o zonas de trabajo, los estudiantes de una universidad considerando las clases a las que pertenecen, las viviendas de una ciudad considerando las manzanas, etc. (Kish, 1965). El marco muestral, en este caso, no está formado por una lista de elementos, sino por una lista de conglomerados, de «grupos de elementos» (grupos de trabajos, clases de la universidad, *manzanas* de viviendas, etc.). Esta forma de operar permite definir el universo sin necesidad de considerar todas las unidades que lo componen, ya que en vez de trabajadores considera «equipos de trabajadores», «clases» en vez de número de estudiantes, etc.

Existe una variante del muestreo por conglomerados, conocida como muestreo por conglomerados en varias etapas o muestreo polietápico, que se caracteriza porque la *unidad final* del muestreo no es el conglomerado, sino una subdivisión de éste. En este caso la selección se lleva a cabo en varias *etapas*. Considerando las entidades de población como conglomerados (una de las situaciones más frecuentes en encuestas a la población general), este muestreo selecciona —en una primera etapa— un número de municipios donde realizar las entrevistas, en la segunda unas determinadas *áreas* dentro de éstos (por ejemplo, una sección censal), en la tercera un grupo de viviendas y, posteriormente, a individuos (residentes en esas viviendas).

Esta forma de proceder, además de facilitar notablemente la recogida de información, supone importantes ahorros de costes a la hora de llevar a cabo el trabajo de campo⁴. Ahora bien, una vez que se han definido los conglomerados (municipios), y tras realizar el muestreo pertinente dentro de cada uno, ¿cómo se procede con la selección de los individuos si no se dispone de un listado con los lugares de residencia de las personas a entrevistar? Al no disponer de «listados de la población» pudiera parecer que no es posible localizar a unos individuos determinados.

La mayor parte de institutos de investigación llevan a cabo la selección muestral realizando una distribución de la muestra por regiones y hábitat⁵. Ante la imposibilidad de conocer con

⁴ En el *British Household Panel Survey*, por ejemplo, Rose *et al.* señalan que el trabajo de campo se incrementaría en un 40% si se utilizara un muestreo sin conglomerados (Rose, Buck y Corti, 1991: 22; Coxon, 1992: 6 y 8). Debe tenerse en cuenta que esta reducción de costes genera, por otro lado, una «pérdida de precisión»; esto es, un aumento del error muestral. Algunos expertos estiman que en el muestreo por conglomerados el error es aproximadamente el doble que en el muestreo aleatorio simple (Rodríguez Osuna, 2000: 469, 476 y 479). Una síntesis de las ventajas e inconvenientes del muestreo por conglomerados en varias etapas, comparado con el muestreo aleatorio simple, presenta Coxon (1992: 6-7).

⁵ Sólo los organismos pertenecientes a la Función Estadística Pública disponen de un marco detallado con los nombres y direcciones de todas las personas. Estos organismos tienen los datos personales de todos los residentes en su ámbito de competencia porque cooperan en la realización del Censo.

exactitud la distribución municipal, los muestreos con entrevistas personales utilizan —en los municipios seleccionados— un sistema de *rutas aleatorias* que *obliga* a los entrevistadores a recorrer varias calles del municipio para realizar las entrevistas asignadas. Definido como *método de rutas*, consiste en: una vez seleccionado un municipio se delimita un área geográfica donde llevar a cabo las entrevistas, y se elige un punto de partida donde comenzará la actividad del encuestador. A partir de aquí, éste deberá seguir una serie de normas fijadas de antemano sobre el recorrido a efectuar.

En España, concretamente, el Centro de Investigaciones Sociológicas realiza una delimitación geográfica de los municipios a nivel de distrito y sección censal, y las *normas de ruta* indican a los encuestadores que deben seleccionar un portal de cada tres, o uno de cada cinco cuando se trata de viviendas unifamiliares. Dentro de un portal hay que realizar una entrevista cada 12 viviendas o fracción, limitando a tres el número máximo de entrevistas en un portal. La vivienda dentro del portal donde realizar la entrevista es la situada en el piso inmediatamente superior al que se ha realizado la entrevista anterior, o en el piso superior que aparece en el «punto de partida».

En el cuadro 1 se presenta una de las *hojas de selección muestral* empleadas por el CIS, concretamente del estudio 2554, correspondiente al Barómetro de enero de 2004. En la parte inferior se presentan las calles que deben ser recorridas por el encuestador. Junto a esta información, el encuestador recibe instrucciones sobre el punto de partida de la ruta. En el caso del ejemplo mostrado en el cuadro 1 esta instrucción decía: «calle del listado de la sección electoral que figura en primer lugar, primer número de los indicados, primer piso, primera letra».

Considerando las calles mostradas en la parte inferior del cuadro 1, la ruta comenzará en la calle Villamanin, número 1, primer piso, primera letra. Tras realizar la entrevista en este portal se pasará al siguiente, efectuando el primer contacto en el piso inmediatamente superior al del portal anterior. Cuando no se consigue una entrevista, el portal seleccionado se sustituye por el contiguo. Cuando esto sucede, para el recuento de la próxima entrevista se considera como si la entrevista se hubiera realizado en el primer contacto. Estos sistemas de rutas, con más o menos variación, se aplican en la mayor parte de los sondeos con entrevistas personales (por ejemplo, los realizados por la Fundación CIRES, los estudios de juventud, *British Social Attitudes*, NORC de la Universidad de Chicago, etc.).

El resto de investigadores únicamente pueden disponer de esta información a nivel *agregado*. Entiéndase por «cifras agregadas» el número de personas que viven en un tamaño de hábitat dentro de una provincia; por ejemplo, el número de personas de la provincia de Zaragoza que residen en municipios menores de 101 habitantes, de 101 a 500, etc.

CUADRO 1

Hoja de selección muestral

N.º DE ESTUDIO: 2554

DESCRIPCIÓN DEL ESTUDIO: Muestra nacional

COMUNIDAD AUTÓNOMA: 13 MADRID (COMUNIDAD DE) - ÁREAS
 PROVINCIA: 28 MADRID - RESTO
 MUNICIPIO: 079 MADRID
 TAMAÑO DE HÁBITAT: 70 >1.000.001
 DISTRITO: 10
 SECCIÓN: 071
 N.º DE ENTREVISTAS: 10
 DISTANCIA A LA CAPITAL: 0 km

SEXO	EDADES					
	18-24	25-34	35-44	45-54	55-64	>65
Varones	0	1	1	1	1	1
Mujeres	0	1	1	1	1	1

PROVINCIA: 28 MADRID
 MUNICIPIO: 079 MADRID
 DISTRITO: 10
 SECCIÓN: 071

Entidad	Denominación de la vía	Vía	Desde	Hasta	C. Postal
Madrid	Villamanin	Calle	0001	0007	28011
Madrid	Villasandino	Calle	0001	0007	28011
Madrid	Villasandino	Calle	0004	0006	28011
Madrid	Villavaliante	Calle	0004	0008	28011
Madrid	Villavieja	Calle	0003	0003	28011
Madrid	Extremadura	Paseo	0210	0226	28011

FUENTE:

Centro de Investigaciones Sociológicas (2004a).

Seleccionada una vivienda, a continuación se procede a la elección de la persona a entrevistar. Cuando reside una persona se procede a su elección, algo que sucede —según el Censo del año 2001— en el 21,3% de los hogares españoles (Instituto Nacional de Estadística, 2004: 2). Considerando que en la mayoría de los hogares españoles residen varias personas, será preciso utilizar un método de selección dentro de la vivienda.

De todos los métodos disponibles para la selección de la persona a entrevistar dentro de la vivienda, el más utilizado en España es la selección por cuotas. Este método fue diseñado en el primer tercio de siglo xx para asegurar que determinados individuos difíciles de localizar estuvieran bien representados en las muestras, y entre sus grandes ventajas destaca su bajo coste en relación a los otros métodos de selección (Kish, 1965: 565). Sudman estima que el muestreo por cuotas es tres veces más barato que un muestreo probabilístico con más de una visita al hogar (Sudman, 1976: 199).

El muestreo por cuotas trata de elaborar una muestra que sea similar, en una serie de características, al universo objeto de estudio. Para ello se seleccionan determinadas características de las unidades a entrevistar —normalmente referidas a rasgos sociodemográficos como edad, sexo, nivel de estudios, profesión, etc.— y posteriormente se elaboran una serie de *fichas de selección* con las características de las personas a entrevistar en cada zona. En la parte central del cuadro 1 se muestran las cuotas de sexo y edad utilizadas en esta investigación⁶. Estas cuotas —elaboradas por los técnicos de muestreo— se han fijado considerando la distribución de la población en la zona correspondiente (Kish, 1965: 564). En las viviendas visitadas los encuestadores únicamente podrán entrevistar a las personas que cumplan las características fijadas en las cuotas, eligiendo a la persona más joven cuando varios miembros de la vivienda cumplen tales características.

Otro rasgo que caracteriza al muestro por cuotas es que no precisa de varias visitas cuando no se obtiene respuesta en la vivienda seleccionada. En el caso del CIS, cuando sus encuestadores se encuentran con esta situación, esto es, cuando no se consigue una entrevista en el portal seleccionado, éste se sustituye por el contiguo: «cuando una entrevista no se consiga en el primer contacto se puede seguir intentando en la puerta contigua»; «si no se consigue ninguna entrevista en el portal seleccionado, se sustituye por el contiguo» (Centro de Investigaciones Sociológicas, 2004c: 2).

Existe un relativo acuerdo a nivel internacional de la idoneidad de los métodos de selección de las viviendas (método de rutas). Este grado de acuerdo desaparece cuando se consideran los métodos de selección dentro del hogar, puesto que un gran número de investigadores (entre otros, Bradburn, 1992: 393; Butcher, 1995: 5-8; Kish, 1965: 563; Marsh y Scarbrough, 1990: 493-494; Moser, 1952: 413; Moser y Kalton, 1971: 129; Morton-Williams, 1993: 33; Stephenson, 1979: 488; Sudman, 1976: 199-200; Worcester y Downham, 1986: 105-107) han destacado la escasa idoneidad del método de cuotas, lo que les lleva a recomendar no utilizarlo.

⁶ La dificultad para encontrar distribuciones de la profesión en el universo, unida a la complejidad del trabajo de campo en rutas con más de dos variables, llevan a utilizar —la mayoría de las veces— cuotas de sexo y edad.

Esta escasa idoneidad lleva a que algunos expertos aconsejen utilizar este método en investigaciones exploratorias o en aquellas que no necesitan resultados muy precisos (Kish, 1965: 562-563; Sudman, 1976: 199; Azorín y Sánchez-Crespo, 1986: 21; O'Muirheartaigh, 1995: 13-15; Lynn, 1995: 20-23), aunque otros afirman que una investigación basada en un método de cuotas puede proporcionar resultados tan precisos como otra basada en un muestreo probabilístico⁷ (Moon, 1995: 9-12; Orton, 1995: 16-19). A este respecto son célebres las investigaciones con cuotas realizadas por Gallup para las elecciones presidenciales de 1940 y 1944; aunque también es conocido el fracaso en la prospección electoral de 1948⁸. Así, Morton-Williams (1993: 35) señala que cuando se trata de investigaciones grandes y complejas es mejor utilizar un muestreo probabilístico, aunque sea más caro, puesto que la no respuesta es menor y se reduce la capacidad de elección del entrevistador (Morton-Williams, 1993: 35).

La profusa utilización de este método por la investigación privada de opinión y mercado (Sánchez-Crespo, 1973: 15; Azorín y Sánchez-Crespo, 1986: 20-22; Rodríguez Osuna, 1991: 43; Camacho *et al.*, 2000: 128), y el debate aún sin cerrar sobre los efectos de su utilización (SCPR, 1995; Kish, 1998), nos han llevado a elaborar este trabajo. Su objetivo es realizar una *evaluación* del método de cuotas —utilizado tras seleccionar las viviendas con el método de rutas— considerando su capacidad para reflejar adecuadamente el universo de donde se ha extraído la muestra.

2. PLANTEAMIENTO DE LA INVESTIGACIÓN

Buscando resolver el *objetivo general* señalado al final del epígrafe anterior consideramos tres aspectos; tres objetivos concretos: el primero busca identificar (o cuantificar) el número de intentos que deben realizarse antes de conseguir una entrevista, y los problemas que esto genera a la hora de representar adecuadamente el universo objeto de estudio. Esto implica, necesariamente, considerar las dificultades producidas al utilizar el muestreo por cuotas con rutas con pocas entrevistas.

Como segundo aspecto, reconstruiremos el proceso de selección de las personas entrevistadas a lo largo de la ruta. El tercer objetivo considera hasta qué punto las personas selec-

⁷ En otro trabajo (Díaz de Rada, 2001: 146-154) se exponen sintéticamente las ventajas y desventajas del método de cuotas.

⁸ Algunos expertos explican este fracaso por la utilización de un Censo no actualizado (Cea D'Ancona, 2004), mientras que otros (Jaime Castillo, 2004: 145) señalan que los errores de tal prospección estaban originados por la utilización del método de cuotas en la selección de los entrevistados.

cionadas presentan similitud —en cuanto a sus características sociodemográficas— con los rasgos del universo. Con el fin de proporcionar información sobre estos tres aspectos proponemos las siguientes hipótesis:

H1 Uno de los grandes problemas del método de cuotas es la cantidad de *intentos* (visitas) realizados que no culminan en una entrevista. Estos intentos se producen por la propia filosofía del método de cuotas (maximizar el número de contactos por unidad de tiempo; según Groves y Couper, 1998: 38-39), así como por el elevado número de contactos con personas que no pueden ser seleccionadas como consecuencia de tener ya cubiertas las características de sexo y edad buscadas en las cuotas.

H1a El número de intentos varía en función del número de orden de la entrevista en la ruta⁹ y del tamaño de la ruta¹⁰:

- A medida que avanza la ruta, el entrevistador va *cubriendo* determinadas cuotas de edad, de modo que —al final— está buscando una persona con unos rasgos determinados (pertenecer a un determinado grupo de edad según sea hombre o mujer).
- Respecto al *tamaño* de la ruta, éste contribuye al aumento del número de visitas infructuosas por el incremento del número de contactos en las rutas más pequeñas.

H1b Esta selección «con condiciones» supone una importante alteración de la equiprobabilidad; dista mucho de producir una elección aleatoria, que recordemos es el principio fundamental del muestreo probabilístico.

H2 Las primeras entrevistas realizadas recogen —fundamentalmente— jubilados y personas que realizan tareas domésticas no remuneradas, mientras que en las últimas entrevistas de la ruta se localizan las personas más difíciles de ser localizadas: varones, jóvenes (menores de 35 años) y ocupados.

H2a La localización de jóvenes (menores de 35 años) y de ocupados requiere visitar un gran número de viviendas, hasta encontrar el prototipo de persona ideal a entrevistar.

⁹ Con esta terminología nos referimos a la «posición» de cada entrevista dentro de la ruta; esto es, si se trata de la primera entrevista realizada, la segunda, la tercera..., la novena o la décima.

¹⁰ Considerando el número de entrevistas que se realizan en cada ruta.

H2b Normalmente, las últimas entrevistas de la ruta se realizan en horario nocturno (a partir de las ocho de la tarde).

H3 El método de cuotas consigue una adecuada representatividad en relación al sexo, la edad y la actividad laboral.

H3a En relación a esta última, la no utilización de «varias visitas» cuando nadie responde a la llamada del entrevistador produce una gran selección de las personas que más tiempo pasan en la vivienda. Esto genera algunos problemas para representar adecuadamente la población activa.

Para llevar a cabo los objetivos propuestos utilizaremos información extraída de los *barómetros* del Centro de Investigaciones Sociológicas realizados durante el año 2004. El Centro de Investigaciones Sociológicas es un organismo autónomo, adscrito al Ministerio de la Presidencia, cuya labor principal es realizar investigaciones científicas sobre la realidad española (CIS, 2006). El CIS cuenta con una red de campo de más de 400 entrevistadores, y se trata —sin duda— del organismo de máxima referencia de la investigación social en España. En el año 2004 entrevistaron a 106.000 personas mediante encuestas (Díaz de Rada y Núñez Villuendas, 2008: Introducción).

En este trabajo utilizaremos información extraída de los *barómetros* realizados durante el año 2004. Los *barómetros* son investigaciones con periodicidad fija que utilizan un cuestionario que combina una serie de indicadores fijos con temas o eventos de interés en un momento concreto. El CIS realiza cada año once *barómetros*, uno al mes excepto el mes de agosto, y todos utilizan el mismo diseño de investigación. El universo objeto de estudio son las personas españolas de 18 y más años (excepto las residentes en Ceuta y Melilla), que son entrevistadas en su domicilio en presencia del encuestador (entrevista personal o cara a cara). El trabajo de campo se lleva a cabo en una semana, normalmente en la segunda o tercera de cada mes, realizándose 2.500 entrevistas en 168 municipios y 49 provincias. Para un nivel de confianza del 95,5% (dos sigmas), y $P = Q$, el error real en cada *barómetro* es de $\pm 2,0\%$ para el conjunto de la muestra y en el supuesto de muestreo aleatorio simple. Se utiliza un diseño muestral en varias etapas, seleccionando las unidades primarias de muestreo (municipios) y las unidades secundarias (secciones) de forma aleatoria proporcional, y las unidades últimas (individuos) por rutas aleatorias y cuotas de sexo y edad.

Durante el año 2005 se llevaron a cabo un total de 27.359 entrevistas, que constituyen la base de datos de esta investigación. Con este tamaño el error muestral se reduce al $\pm 0,604\%$, considerando los mismos parámetros de nivel de confianza, P y Q , y el supuesto de muestreo aleatorio simple. En el anexo 1 se detallan los aspectos técnicos de cada *barómetro*.

Todos los cuestionarios terminan con una serie de preguntas de control del trabajo de campo donde el encuestador debe anotar las incidencias en el proceso de localización del entrevistado. Se muestra un ejemplo en el cuadro 2. Para completarlo, el encuestador debe anotar el «número de veces» que ha ocurrido cada una de estas situaciones hasta efectuar la entrevista. Es decir, que cada cuestionario lleva unidos, junto a la información propia de su temática, los problemas que han tenido lugar para localizar a cada entrevistado. En este trabajo nuestro interés se centrará en el relativo a los contactos fallidos por no cumplir cuotas (I.7)¹¹. Los encuestadores realizan anotaciones en esta incidencia cuando no se llega a realizar la entrevista, porque en la vivienda donde se contacta no hay nadie que cumpla las condiciones para ser elegido (sexo o edad) o bien porque —aun viviendo esta persona en la vivienda seleccionada— no se encuentra en su domicilio durante la visita del encuestador.

CUADRO 2

Parte final de los cuestionarios del CIS:

Preguntas para los entrevistadores. Incidencias en la entrevista

- I.1. Número de orden de entrevista en la ruta
- I.2. Dificultad de acceso al edificio, casa, urbanización, etc.
- I.3. Viviendas en las que no hay nadie
- I.4. Viviendas en las que se niegan a recibir ninguna explicación
- I.5. Negativas de varones a realizar la entrevista
- I.6. Negativas de mujeres a realizar la entrevista
- I.7. Contactos fallidos por no cumplir cuotas
- I.8. Contactos fallidos por no ser una vivienda (oficinas, consultas médicas, etc.)
- I.9. Viviendas de inmigrantes

FUENTE:

Centro de Investigaciones Sociológicas (2004a).

¹¹ Una perspectiva global del resto de incidencias se realiza en Díaz de Rada y Núñez Villuendas (2008).

3. CONTACTOS FALLIDOS POR NO CUMPLIR CUOTAS: PRINCIPALES RESULTADOS. IMPORTANCIA DEL NÚMERO DE ORDEN DE LA ENTREVISTA EN LA RUTA¹²

Recordemos que una vez que el entrevistador ha accedido a una vivienda y ha localizado a personas comprendidas en las cuotas de sexo y edad indicadas en la *hoja de muestra* (cuadro 1) procede con su elección. Cuando varias personas cumplen tales condiciones se recomienda a los encuestadores que entrevisten a la más joven. Ahora bien, ¿qué sucede cuando el entrevistador ha completado la mayor parte de su ficha de cuotas y va buscando una persona con unos rasgos determinados? O, mejor aún, ¿qué ocurre si nadie está dentro de las cuotas?

Cuando se produce esta situación, el entrevistador deberá abandonar la vivienda agradeciendo la atención prestada. Los protocolos utilizados en estos casos indican que «cuando una entrevista no se consiga en el primer contacto se puede seguir intentando en la puerta contigua» (Centro de Investigaciones Sociológicas, 2004c: 2); de modo que el entrevistador anotará la incidencia y llamará de nuevo en la vivienda más cercana. En el estudio realizado por el CIS a sus encuestadores en 1989, éstos señalaron que al comienzo de la ruta visitaban —como media— entre 3,35 y 3,07 viviendas, y al final de la ruta entre 5,17 y 8 (Núñez Villuendas, 2005: 227). En los barómetros del año 2004 esto sucede en dos de cada tres *intentos* realizados por los encuestadores, siendo necesario visitar cinco viviendas para efectuar la entrevista (tabla 1). Dicho de otro modo, cuando el encuestador no realiza la entrevista en la primera vivienda suele visitar un promedio de cinco viviendas en las que no puede realizarse la entrevista por «no cumplir cuotas».

Otras investigaciones realizadas en el mismo contexto —y con una metodología similar— incrementan el número de intentos infructuosos para contactar con una vivienda hasta el 85% (González Blasco *et al.*, 2006: 407). Así, en el estudio *Jóvenes Españoles 2005* se visitaron 42.414 viviendas para conseguir el tamaño muestral fijado, establecido en 4.000 entrevistas. Eliminando los «no hogares» (2.754) porque no forman parte del universo, el número de visitas se queda en 39.660 (ver datos en anexo 2). Teniendo en cuenta que en 21.751 viviendas no se pudo hacer la entrevista por no haber personas con las características de la cuota, esto supone que el 54,8% (21.751/39.660) de los intentos no culminaron con la realización de la entrevista por tener cumplida la ficha de cuotas.

De hecho, mientras que en los barómetros del CIS esta incidencia supone el 26,2% del total de incidencias (Díaz de Rada y Núñez Villuendas, 2008: tabla 2.2 o cuadro 2.1), en el estudio de juventud citado esta cifra llega hasta el 51% cuando se consideran todos los

¹² Indicador I.1 de la ficha de incidencias (cuadro 2).

TABLA 1

Contactos fallidos por no cumplir cuota

	Distribución porcentual	
	n	Porcentaje
Ninguno	8.886	33,5
Uno	3.635	13,7
Dos	3.149	11,9
Tres	2.386	9,0
Cuatro	1.711	6,4
Cinco	1.251	4,7
Seis	1.075	4,0
Siete y más	4.459	16,8
NÚMERO DE CASOS	26.552	100,0

Promedio de contactos fallidos por no cumplir cuota: 4,94.

FUENTE:

Elaboración propia con datos del Centro de Investigaciones Sociológicas (2004b).

contactos efectuados, y hasta el 54% cuando se tienen en cuenta únicamente los contactos a personas que forman parte del universo. Téngase en cuenta que esta situación puede explicarse por la *menor amplitud* de las cuotas utilizadas y por la limitación del universo a los jóvenes entre 15 y 24 años; los más difíciles para cooperar (Pasadas del Amo *et al.*, 2004: 11, y 2006a). Investigaciones realizadas con encuestas telefónicas señalan que uno de cada tres contactos (33%) se encuentra con «cuotas cubiertas»; pese a intentar el contacto con el hogar en cinco ocasiones durante siete días (Pasadas del Amo *et al.*, 2006b: 5 y 7).

Volviendo a los datos de los barómetros del CIS, en la tabla 2 puede observarse que el número de incidencias presenta una gran variación según el «número de orden de la entrevista en la ruta», disminuyendo a partir de la cuarta: en la primera entrevista de la ruta, dos de cada tres entrevistas no están afectadas por esta incidencia, cifra que desciende paulatinamente hasta la cuarta. A partir de ésta el número de *intentos* para conseguir una entrevista experimenta un importante aumento, desde 4,4 intentos hasta 7,40 en la décima entrevista de la ruta.

Considerando que el universo objeto de estudio es la población española de 18 y más años (anexo 1), y teniendo en cuenta que en la primera visita es posible entrevistar a cualquier persona dentro de su ficha de cuotas, resulta sorprendente que —para realizar la primera entrevista de la ruta— un 40% de los intentos hayan derivado en una entrevista in-

TABLA 2

Contactos fallidos por no cumplir cuotas según número de orden de la entrevista en la ruta

	Distribución porcentual incidencias (porcentajes horizontales)		Número de casos	Promedio de contactos fallidos por no cumplir cuota	
	Ninguna	Una y varias			
TOTAL	33,5	66,5	26.552	4,94 ¹³	(17.666)
<i>Número orden de la entrevista</i>					
Uno	+59,9	-40,1	3.443	<u>2,76</u>	(1.379)
Dos	+47,9	-52,1	3.447	2,92	(1.795)
Tres	+37,3	-62,7	3.415	3,65	(2.142)
Cuatro	-30,5	+69,5	3.385	4,39	(2.352)
Cinco	-25,2	+74,8	3.220	4,96	(2.407)
Seis	-23,6	+76,4	3.012	5,80	(2.301)
Siete	-20,4	+79,6	2.572	6,11	(2.047)
Ocho	-20,0	+80,0	1.967	6,50	(1.573)
Nueve	-16,8	+83,2	1.303	7,16	(1.084)
Diez	-23,6	+76,4	758	7,40	(579)

(+) Residuos estandarizados corregidos (valor positivo) con un nivel de significación igual o menor a 0,01.

(-) Residuos estandarizados corregidos con valor negativo.

El subrayado indica que es el valor promedio más bajo de esta variable.

La negrilla indica que es el promedio valor más alto de esta variable.

FUENTE:

Centro de Investigaciones Sociológicas (2004b). Reproducido de Díaz de Rada y Núñez Villuendas (2008: tabla 3.5).

fructuosa al no localizar a nadie dentro de la cuota. Esto implica preguntarse por la tipología de personas que se han encontrado en las viviendas visitadas: ¿sólo menores de 18 años?, ¿o quizás personas no incluidas en la ficha de cuotas del entrevistador?¹⁴. Las incidencias por no cumplir cuota pueden estar producidas por contactar con viviendas donde sólo hay menores de esa edad en el momento de la entrevista, o por ausencias de determinadas características en la ficha de cuotas del entrevistador. Esta segunda razón será muy frecuente —planteamos como hipótesis— en las rutas menores de 10 entrevistas (tal y como se propuso en la hipótesis H1b). La explicación viene del elevado número de rutas con menos de 10 entrevistas: al tener pocas entrevistas por rutas, la ficha de características está bastante vacía, por lo que aumentará el número de intentos en la primera visita.

¹³ Esta cifra se refiere a la media de contactos fallidos por no cumplir cuota, y se ha calculado considerando únicamente las entrevistas realizadas tras incidencias. Es decir, se han eliminado aquellas entrevistas realizadas «a la primera».

¹⁴ En la ficha de cuotas mostrada en el cuadro 1, por ejemplo, no había ninguna entrevista para menores de 24 años.

Lo explicaremos mejor utilizando un ejemplo, concretamente la hoja de muestra que fue mostrada en el cuadro 1. Como pudimos apreciar, se trata de una ruta de 10 entrevistas donde se incluyen personas de ambos sexos mayores de 25 años. Dicho de otro modo, no debe hacerse ninguna entrevista a personas entre 18 y 24 años¹⁵. En una ruta de ocho entrevistas, por ejemplo, será necesario dejar más cuotas vacías (con 0 entrevistas), y más aún en una ruta de siete entrevistas. De modo que la ficha de cuotas del cuadro 1 se transformará, para una ruta de seis entrevistas, en una tabla con seis ceros; con seis «tipos de personas» que no pueden ser seleccionadas. La situación se muestra en el cuadro 3. Nuestra hipótesis plantea que, lógicamente, resultará más difícil completar las cuotas del cuadro 3 que la mostrada en el cuadro 1.

CUADRO 3

Ficha de cuotas de una ruta con seis entrevistas

Sexo	Edades					
	18-24	25-34	35-44	45-54	55-64	>65
Varones	0	0	1	1	1	0
Mujeres	0	1	0	1	0	1

FUENTE:

Elaboración propia con datos del Centro de Investigaciones Sociológicas (2004b).

En la tabla 3 se muestran los tamaños de las rutas empleadas en los barómetros realizados por el CIS en el año 2004: un 28% de las entrevistas se han realizado en rutas de diez entrevistas, un 22% en rutas de nueve entrevistas, un 20% en rutas de ocho, un 15% en rutas de siete, y otro 14,5% en rutas menores de seis entrevistas. En la segunda parte de la tabla puede comprobarse cómo a medida que se consideran rutas mayores descende el número de incidencias (aumentan los porcentajes de la columna «ninguna incidencia»); lo que permite aceptar el planteamiento propuesto en el párrafo anterior.

La misma conclusión se desprende al analizar la tabla 4, donde el promedio de contactos fallidos es siempre más alto en las rutas pequeñas que en las grandes. Realizar un trabajo de campo con varios tamaños de ruta permite recoger mejor la heterogeneidad del univer-

¹⁵ Tal y como puede apreciarse en el cuadro 1, cada *hoja de muestra* dispone de 12 posibles cuotas; seis grupos de edad para hombres y otros seis para mujeres.

TABLA 3

Tamaño de las rutas. Contactos fallidos por no cumplir cuotas considerando el tamaño de la ruta

	Tamaño de las rutas ¹⁶		
	Frecuencia	Porcentaje (vertical)	Porcentaje acumulado
Ruta de 4 entrevistas y menos	463	1,7	1,7
Ruta de 5 entrevistas	821	3,0	4,7
Ruta de 6 entrevistas	2.678	9,8	14,5
Ruta de 7 entrevistas	4.161	15,2	
Ruta de 8 entrevistas	5.474	20,0	
Ruta de 9 entrevistas	6.025	22,1	
Ruta de 10 y 11 entrevistas	7.702	28,2	
TOTAL	27.324		
Contratos fallidos por no cumplir cuotas considerando el tamaño de la ruta			
	Porcentajes horizontales		
	Ninguna	Una y varias	Número de casos
TOTAL	59,9	40,1	26.552
Ruta de 6 entrevistas y menos	44,3	55,7	3.820
Ruta de 7 entrevistas	49,9	50,1	4.161
Ruta de 8 entrevistas	58,1	41,9	5.474
Ruta de 9 entrevistas	70,4	29,6	6.025
Ruta de 10-11 entrevistas	74,8	25,2	7.702

FUENTE:

Elaboración propia con datos del Centro de Investigaciones Sociológicas (2004b).

so al utilizar varios puntos de partida en cada municipio, pero implica también problemas para la organización del campo al aumentar notablemente el número de contactos fallidos por no cumplir cuota. El Centro de Investigaciones Sociológicas, por ejemplo, suele partir de un número de rutas que oscila entre 6 y 10 entrevistas, con lo que se asegura una gran dispersión muestral (Rodríguez Osuna, 2000: 472).

En la tabla 5 se presenta una información que justifica la utilización de rutas de diferente tamaño. La relación entre el hábitat y el tamaño de la ruta desvela un mayor empleo de rutas menores de 7 entrevistas en municipios «intermedios» (entre 2.000 y 50.000 habitantes). Otros análisis, que no desvelamos aquí por motivos de espacio, desvelan que en este hábitat estas rutas «pequeñas» se han combinado con rutas de 10 entrevistas. En España,

¹⁶ Es preciso indicar que esta variable no está incluida en los barómetros del CIS y que ha sido elaborada por el autor.

TABLA 4

Promedio de contactos fallidos por no cumplir cuotas: según tamaño de la ruta y según el número de orden de la entrevista en la ruta

	Según tamaño de la ruta				
	Ruta de 6	Ruta de 7	Ruta de 8	Ruta de 9	Ruta de 10
TOTAL	5,33	5,37	4,86	4,77	4,75
	Según el número de orden de la entrevista en la ruta				
	Ruta de 6	Ruta de 7	Ruta de 8	Ruta de 9	Ruta de 10
<i>Número orden de la entrevista</i>					
Uno	3,25	2,85	2,51	2,17	2,45
Dos	3,35	3,07	2,72	2,48	2,43
Tres	4,58	4,09	3,43	2,88	2,85
Cuatro	5,92	5,01	3,90	3,49	3,42
Cinco	6,56	5,70	4,67	4,21	3,74
Seis	7,36	7,10	5,75	4,96	4,43
Siete		8,23	5,95	5,23	5,08
Ocho			7,74	6,52	5,37
Nueve				7,60	6,78
Diez					7,41

FUENTE:

Elaboración propia con datos del Centro de Investigaciones Sociológicas (2004b).

sin embargo, la gran mayoría de las empresas de investigación de opinión y mercados utilizan rutas del mismo tamaño, algo que simplifica tremendamente el diseño muestral y la organización del trabajo de campo.

Concluiremos este apartado reflexionando sobre las implicaciones para la teoría muestral de la información presentada en las tablas anteriores; siempre considerando que el muestreo probabilístico se fundamenta en la idea que «todos los miembros del universo tienen la misma probabilidad de ser elegidos». Téngase en cuenta, como vimos en la tabla 1, que en dos de cada tres viviendas no se realiza una entrevista porque la cuota ya está cumplida. Es decir, en un 66% de los intentos realizados se produce una probabilidad de selección «0» porque el entrevistador ya entrevistó a una persona de ese grupo de edad y sexo. El número de personas (entrevistados potenciales) que estará afectado por esta situación varía en función del tamaño de la ruta (tabla 3), algo que está condicionado —a su vez— por el tamaño del municipio.

TABLA 5

Tamaño de las rutas en relación con el hábitat

	Tamaño del municipio (en porcentajes)						Número de casos
	Menos de 2.000 habitantes	2.001 a 10.000 habitantes	10.001 a 50.000 habitantes	50.001 a 100.000 habitantes	100.001 a 400.000 habitantes	Más de 400.000 habitantes	
TOTAL	7,9	16,2	23,9	9,7	23,9	18,3	27.324
<i>Tamaño de las rutas</i>							
2-5	+10,7	+45,1	+28,1	-4,7	-10,7	-0,8	1.284
6-7	-0,6	+31,4	+34,0	-8,5	+25,1	-0,4	6.839
8	+21,6	-14,6	+25,7	+10,9	+26,1	-1,0	5.474
9	+11,6	-4,6	23,5	+11,4	-21,5	+27,5	6.025
10-11	-1,3	-8,3	-13,3	9,2	+25,5	+42,3	7.702

FUENTE:

Elaboración propia con datos del Centro de Investigaciones Sociológicas (2004b).

Téngase en cuenta que esta situación supone una gran alteración de la equiprobabilidad, hipótesis H1b, lo que implica una contradicción con el planteamiento básico del muestreo probabilístico («todos los miembros del universo tienen la misma probabilidad de ser elegidos»).

4. RASGOS SOCIODEMOGRÁFICOS DE LAS PERSONAS ENTREVISTADAS EN CADA MOMENTO DE LA RUTA

La enorme importancia de la variable «número de orden de la entrevista» en la ruta para explicar los contactos fallidos por no cumplir ruta precisa llevar a cabo un análisis detallado de la misma, cuyos resultados se muestran en la tabla 6. En esta tabla se relacionan el número de orden de la entrevista en la ruta —considerando los diferentes tamaños de las rutas— y los rasgos sociodemográficos de las personas entrevistadas en cada momento; desvelando así los rasgos que caracterizan a los entrevistados en la primera entrevista de la ruta, en la segunda, etc.

Con el fin de presentar esta información con mayor claridad se ha realizado una transformación de la variable «número de orden de la entrevista en la ruta» agrupando varias categorías con el fin de reducir su amplitud. Para llevar a cabo tal agrupación se ha utilizado la información proporcionada por el promedio de contactos fallidos por no cumplir cuota según el número de orden de la entrevista (mostrada en la tabla 4), considerando cinco nive-

les: 1.º) menos de tres contactos fallidos; 2.º) contactos fallidos entre tres y cuatro; 3.º) media de contactos fallidos entre cuatro y cinco; 4.º) en torno a cinco contactos fallidos, y 5.º) número de contactos fallidos superior a cinco. En el cuadro 4 se presentan los detalles de la agrupación efectuada.

CUADRO 4

Reducción de categorías de la variable «número de orden de la entrevista en la ruta» considerando el «promedio de contactos fallidos por no cumplir cuota»

Nivel	Contactos fallidos por no cumplir cuota (medias)	Categorías agrupadas
<i>Agrupación de la ruta de 6 entrevistas</i>		
1.º	Inferior a 3	Ninguna
2.º	Entre 3 y 4	Primera y segunda entrevistas
3.º	Entre 4 y 5	Tercera
4.º	Rondando el 5 (4,75-5,25)	Ninguna
5.º	Medias superiores a 5,26	Cuarta, quinta y sexta
<i>Agrupación de la ruta de 7 entrevistas</i>		
1.º	Inferior a 3	Primera entrevista
2.º	Entre 3 y 4	Segunda
3.º	Entre 4 y 5	Tercera
4.º	Rondando el 5 (4,75-5,25)	Cuarta
5.º	Medias superiores a 5,26	Quinta, sexta y séptima
<i>Agrupación de la ruta de 8 entrevistas</i>		
1.º	Inferior a 3	Primera y segunda entrevistas
2.º	Entre 3 y 4	Tercera y cuarta
3.º	Entre 4 y 5	Quinta
4.º	Rondando el 5 (4,75-5,25)	Ninguna
5.º	Medias superiores a 5,26	Sexta, séptima y octava
<i>Agrupación de la ruta de 9 entrevistas</i>		
1.º	Inferior a 3	Primera, segunda y tercera entrevistas
2.º	Entre 3 y 4	Cuarta
3.º	Entre 4 y 5	Quinta
4.º	Rondando el 5 (4,75-5,25)	Sexta
5.º	Medias superiores a 5,26	Séptima, octava y novena
<i>Agrupación de la ruta de 10 entrevistas</i>		
1.º	Inferior a 3	Primera, segunda y tercera entrevistas
2.º	Entre 3 y 4	Cuarta y quinta
3.º	Entre 4 y 5	Sexta
4.º	Rondando el 5 (4,75-5,25)	Séptima
5.º	Medias superiores a 5,26	Octava, novena y décima

La información mostrada por la tabla 6 —donde se relacionan el número de orden de la entrevista (considerando los diferentes tamaños de las rutas) y los rasgos sociodemográficos de las personas entrevistadas— indica que las primeras entrevistas de la ruta se hacen a mujeres, con una diferencia de 5 puntos porcentuales a favor de éstas. La distribución por sexos se iguala en el «nivel 4», y a partir de este momento se realiza una mayor selección de hombres con el fin de «cuadrar» la sobrerrepresentación de las mujeres realizada en las primeras entrevistas.

Respecto a la *edad*, las primeras entrevistas de la ruta se realizan a las personas mayores de 55 años, mientras que el «nivel 3» se caracteriza por una gran elección de entrevistados entre 35 y 44 años. En el «nivel 4» se corrige la infrarrepresentación de los grupos de edad más jóvenes. Las últimas entrevistas de la ruta (nivel 5) se caracterizan por realizar más entrevistas a los menores de 24 años y a aquellos entre 45 y 54 años. Es decir, aquellas cuotas más difíciles; en línea con lo que han revelado otras investigaciones, aun cuando utilizan procedimientos distintos: «... en los últimos años se ha experimentado un incremento notable para cerrar el campo de los estudios, sobre todo en lo que se refiere a las cuotas de los más jóvenes» (Pasadas del Amo *et al.*, 2004: 11). Los más jóvenes precisan de más contactos y son localizados al final del trabajo de campo. En la explicación de este fenómeno, Pasadas del Amo *et al.* (2006a) analizan la influencia de la disponibilidad horaria, la predisposición a participar en encuestas (tasa cooperación) y la localización de los jóvenes en sus residencias habituales. Desde su punto de vista, este último factor explica la dificultad en la cooperación del colectivo menor de 30 años, y lo interpretan aludiendo a que se trata de un «segmento afectado por circunstancias vitales de transición como son la formación y la incorporación al trabajo». Estas circunstancias generan movilidad geográfica que produce diferencias entre la residencia real y el domicilio censal. Un análisis del Censo en la provincia de Córdoba les permite detectar que el 12% de los jóvenes entre 18 y 29 años residen en un lugar diferente al declarado en el Censo; cifra que aumenta al 18% en los municipios entre 5.000 y 20.000 habitantes, y al 20% en los municipios menores de 5.000 habitantes (Pasadas del Amo *et al.*, 2006a: 10).

El perfil del colectivo que reside en un lugar diferente al recogido en el Censo se caracteriza por su masculinidad (58% hombres), juventud (67% entre 18 y 29 años, edad media 27,5) y residir en municipios menores de 20.000 habitantes (61,2%). Se trata de un problema muy importante en la medida que las cuotas (sexo, edad, hábitat, etc.) se realizan considerando datos censales o padronales, datos que —en ocasiones— tienen poco que ver con la realidad. Por otro lado, debe tenerse en cuenta que es un análisis del Censo del año 2001 realizado en el año 2006; problema que aumentará sustancialmente a medida que envejece el Censo.

TABLA 6

Rasgos sociodemográficos de los entrevistados según el número de orden de la entrevista en la ruta (categorizada en cinco valores)

	Número de orden de la entrevista en la ruta categorizada en cinco grupos (porcentajes horizontales)					Número de casos
	Nivel 1 (a)	Nivel 2 (b)	Nivel 3 (c)	Nivel 4 (d)	Nivel 5 (e)	
TOTAL	24,1	19,4	12,1	12,9	31,4	25.996 ¹⁷
<i>Sexo</i>						
Hombre	-21,3	-18,0	11,8	12,9	+36,0	12.536
Mujer	+26,8	+20,8	12,4	12,9	-27,1	13.460
<i>Edad en cuotas</i>						
18-24 años	-18,4	-16,1	11,9	+15,0	+38,6	3.319
25-34 años	-22,5	18,6	12,7	13,1	+33,1	5.412
35-44 años	-23,0	19,2	+13,0	13,1	31,8	4.771
45-54 años	-22,6	20,0	11,4	13,0	-33,0	4.002
55-64 años	+26,8	-18,2	11,6	12,0	32,2	3.263
65 y más años	+30,5	+22,8	11,8	11,8	23,2	5.229
<i>Actividad principal</i>						
Trabaja	-21,6	-18,0	12,0	+13,5	+34,9	12.464
Parado	22,9	19,8	+13,2	13,6	30,6	2.254
Jubilado (ha trabajado)	+27,1	+21,2	12,1	12,3	-27,3	5.166
Jubilado (no ha trabajado)	+29,6	+24,1	-10,7	13,0	-22,7	1.272
Tareas domésticas no remuneradas	+28,6	+20,9	12,4	-11,7	-26,4	3.430
Estudiante	-21,6	-16,3	12,2	12,9	+37,0	1.337
<i>Situación profesional</i>						
Asalariado fijo	+24,9	19,4	-11,7	13,0	31,0	14.348
Asalariado eventual	-22,4	19,1	+12,9	+13,6	32,0	6.333
Empresarios, profesionales con asalariados	23,0	19,2	10,9	11,8	+35,1	1.201
Profesionales y autónomos	24,1	19,9	12,6	12,1	31,2	3.669
<i>Persona que aporta más ingresos al hogar</i>						
Entrevistado	-23,5	19,1	12,2	12,6	+32,6	11.509
Otra persona	+24,7	+19,8	12,2	13,1	-30,2	12.332
Entrevistado y otro a partes iguales	24,6	18,7	11,5	13,4	31,8	2.030

(a) Estas agrupaciones se han realizando considerando los valores medios de la variable «promedio de contactos fallidos por no cumplir cuota» que han sido mostrados en la tabla 4. En el cuadro 4 se han presentado las agrupaciones y sus características. Así, el nivel 1 se caracteriza por tener menos de tres contactos fallidos por no cumplir cuota.

(b) Nivel 2: media contactos fallidos por no cumplir cuota entre tres y cuatro.

(c) Nivel 3: media contactos fallidos por no cumplir cuota entre cuatro y cinco.

(d) Nivel 4: cinco contactos fallidos por no cumplir cuota.

(e) Nivel 5: media contactos fallidos por no cumplir cuota superior al cinco.

FUENTE:

Elaboración propia con datos del Centro de Investigaciones Sociológicas (2004b).

¹⁷ El número de casos es inferior al mostrado en otras tablas porque únicamente se han considerado las rutas de 6 y más entrevistas, lo que implica dejar de considerar 1.284 rutas (ver tabla 3).

Volviendo de nuevo a la información proporcionada por la tabla 6, la *actividad principal* revela que las primeras entrevistas se realizan a jubilados y a personas que realizan tareas domésticas no remuneradas, al tiempo que los trabajadores y los estudiantes son los menos elegidos en las primeras entrevistas. A partir del nivel 4 se realiza una mayor selección de trabajadores y estudiantes, compensando así las bajas elecciones realizadas en las primeras entrevistas. El análisis de la *situación profesional* muestra que los asalariados fijos son entrevistados en las primeras entrevistas de la ruta, mientras que los asalariados eventuales son entrevistados —fundamentalmente— en el tercer y cuarto niveles. Los empresarios y profesionales con asalariados apenas son seleccionados en las primeras entrevistas, siendo elegidos a partir del quinto nivel. Se aprecia también una similitud de los profesionales y autónomos con la distribución total, así como un elevado porcentaje de asalariados eventuales que son entrevistados en el tercer y cuatro niveles.

Respecto a la persona que aporta más ingresos al hogar, en las primeras entrevistas son seleccionadas «otras personas». A partir del quinto nivel es cuando suele entrevistarse a los entrevistados. Es importante recordar cómo en el quinto nivel aumentaba el número de hombres seleccionados, personas entre 18 y 35 años, y trabajadores.

La información mostrada en los párrafos anteriores permite aceptar la hipótesis H2 cuando planteaba que *las primeras entrevistas realizadas recogen —fundamentalmente— jubilados y personas que realizan tareas domésticas no remuneradas, mientras que en las últimas entrevistas de la ruta se localizan las personas más difíciles de ser localizadas: hombres, jóvenes (menos de 35 años) y trabajadores*. Esta tabla proporciona información suficiente para aceptar la primera de las subhipótesis planteadas, que señalaba que *la localización de varones jóvenes (menores de 35) y ocupados requiere visitar un gran número de viviendas, hasta encontrar el prototipo de persona ideal a entrevistar* (H2a). Las personas de mediana edad son entrevistadas a partir del nivel 3, los trabajadores a partir del 4, y los menores de 35 años a partir del 5. Puede observarse cómo el colectivo al que se refiere la subhipótesis (varones trabajadores menores de 35 años) no es entrevistado hasta el quinto nivel, hasta las últimas entrevistas de la ruta.

Considerando que los entrevistados en las últimas entrevistas de la ruta (nivel 5) son los estratos más difíciles (tabla 6), será interesante conocer los horarios en los que se lleva a cabo la entrevista. Así, la relación entre el horario de realización de la entrevista y el «número de contactos fallidos por no cumplir cuota» (tabla 7) muestra que las entrevistas más difíciles (los estratos que precisan de un mayor número de intentos, los más complicados de localizar) son realizadas en las horas finales del día; en horario vespertino (16-20 horas) y nocturno (20-22 horas). No obstante, este *desplazamiento hacia las horas finales del día* no permite aceptar la segunda de las subhipótesis apuntadas dentro de la hipótesis dos:

normalmente las últimas entrevistas de la ruta se realizan en horario nocturno, a partir de las 8 de la tarde (H2b). Como puede apreciarse en la tabla 7, tan sólo el 11% de las últimas entrevistas de la ruta (nivel 5) se lleva a cabo en horario nocturno, mientras que el 45% se realiza en horario de tarde (de 16 a 20 horas).

TABLA 7

Hora de realización de la entrevista y contactos fallidos por no cumplir cuota según número de orden (cuadro 4)

	Hora de realización de la entrevista (porcentajes horizontales)				Número de casos
	Mañana (9-12)	Mediodía (12-16)	Tarde (16-20)	Noche (20-22)	
TOTAL	17,4	40,4	37,2	5,0	25.855
<i>Contactos fallidos por no cumplir...</i>					
Nivel 1	+37,7	-33,5	-27,8	-1,1	6.244
Nivel 2	+23,0	+42,2	-33,1	-1,7	5.011
Nivel 3	-13,5	+46,6	37,4	-2,6	3.140
Nivel 4	-8,1	+44,2	+42,4	5,3	3.332
Nivel 5	-3,6	40,6	+44,9	+10,9	8.128

FUENTE:

Elaboración propia con datos del Centro de Investigaciones Sociológicas (2004b).

5. COMPARACIÓN CON EL UNIVERSO AL QUE SE DESEAN EXTRAPOLAR LOS RESULTADOS

En los apartados anteriores se ha mostrado el proceso de selección de los entrevistados con el método de cuotas, donde se han evidenciado algunas limitaciones del método, pero todavía queda la «prueba» definitiva: hasta qué punto el método consigue representar adecuadamente el universo objeto de estudio. Daremos respuesta a este interrogante comparando varios aspectos (variables) del universo y la muestra obtenida; concretamente, la distribución por edades y sexos, la relación con la actividad económica y la situación profesional del entrevistado (o del cabeza de familia cuando el entrevistado no trabaja).

La distribución de la población española de 18 años según grupos de edad y sexo se muestra en la tabla 8, y ha sido obtenida de la revisión del Padrón Municipal a 1 de enero

TABLA 8

Comparación de la muestra con datos padronales considerando la distribución según edad y sexo (en porcentajes)

Edad	Universo (Padrón 2004)	
	Hombres	Mujeres
18-24 años	12,0	10,9
25-34 años	22,0	19,8
35-44 años	20,0	18,6
45-54 años	15,5	14,9
55-64 años	12,4	12,4
64 y más años	17,6	23,0
TOTAL	17.534.300	18.361.662

Edad	Muestra			
	Hombres		Mujeres	
18-24 años	13,6	+1,6 (a)	11,9	+1,0
25-34 años	21,7	-0,3 (b)	19,6	-0,2
35-44 años	18,8	-1,6	18,0	-0,6
45-54 años	15,4	-0,1	15,1	+0,2
55-64 años	12,6	-0,2	12,5	-0,1
64 y más años	18,0	+0,4	20,7	-2,3
TOTAL	13.214		14.145	

(a) Valores positivos: Categorías sobrerrepresentadas por la muestra por cuotas.

(b) Valores negativos: Categorías infrarrepresentadas por la muestra por cuotas.

FUENTE:

Instituto Nacional de Estadística (2005a) y Centro de Investigaciones Sociológicas (2004b).

del año 2004. Como puede observarse, la desviación entre el universo y la distribución muestral es muy pequeña. Destaca, tan sólo, la sobrerrepresentación muestral de los menores de 24 años (+1,6), así como la infrarrepresentación de los hombres entre 35 y 44 años (-1,6) y de las mujeres mayores de 64 años (-2,3).

Esta adecuada representatividad tiene su explicación en el método de selección de los entrevistados últimos, utilizando cuotas de sexo y edad que son similares a la distribución del universo. Recordemos que los entrevistadores *buscan* personas con unos rasgos determinados (recordar cuadro 1); de modo que la correcta definición de las cuotas trae

como consecuencia la correspondiente adecuación de la muestra al universo objeto de estudio. Esta adecuada representatividad se consigue pese a que el Centro, consciente de la dificultad de localizar a una persona de determinadas características, permite sustituirla por otra de las cuotas contiguas: «En caso de ser imposible conseguir una determinada cuota de edad, se podrá sustituir por una de las cuotas adyacentes», se señala en las instrucciones para la aplicación de la muestra (CIS, 2004c: 4). Es importante precisar que se permite una alteración en la cuota de edad, nunca en el sexo, y sólo puede realizarse una vez en cada ruta. En tal caso deberá anotarse en el cuestionario —y en la hoja de ruta— la edad real (o la cuota) de la persona entrevistada y el cambio realizado. Esta información, junto con los motivos del cambio, se anotarán también en el informe de trabajo de campo.

Verificada la adecuada representatividad muestral del sexo y la edad, continuaremos con información laboral de los entrevistados, concretamente la relación con la actividad económica y la situación profesional. En este caso la muestra será comparada con la Encuesta de Población Activa, una estadística cuyo objetivo primordial es «conocer el grado de actividad económica de la población y otras características relacionadas con dicha actividad» (Instituto Nacional de Estadística, 2002: 6). Numerosos trabajos han señalado que esta investigación es la fuente más adecuada para conocer los datos en relación a la actividad económica (Ballano Fernández y Martínez Vidal, 2000: 270, entre otros); mucho más cuando tenemos en cuenta que se trata de una de las investigaciones más importantes que se realizan en España. Entre las razones de esta afirmación está su elevado tamaño muestral (200.000 personas cada trimestre, según el Instituto Nacional de Estadística, 2002: 16), así como los recursos utilizados para la actualización de los *marcos*. A juicio de Losilla (2005: 3), es la principal encuesta dirigida a los hogares, considerando el tamaño muestral, el coste y el personal empleado.

El análisis de la tabla 9 desvela una gran similitud existente entre ambas tasas de actividad cuando se compara todo el agregado poblacional, así como los varones y las mujeres por separado. Ahora bien, en relación a la tasa de actividad por grupos de edad la muestra por cuotas infraestima el colectivo entre 20 y 54 años, sobreestimándola en los mayores de 55 años. Téngase en cuenta que se trata de diferencias pequeñas, de dos y tres puntos porcentuales, respectivamente¹⁸. La tasa de paro, sin embargo, presenta grandes diferencias: la muestra por cuotas produce una sobrerrepresentación del 4,36 en el caso del universo total, y del 7,08 en el colectivo de mujeres.

¹⁸ No se presta atención a la tasa de actividad del grupo con menos edad porque se comparan dos poblaciones diferentes: una muestra donde se ha entrevistado a personas de 18 y más años, y otra que considera aquellas que tienen 16 años y más.

TABLA 9

Comparación de la muestra con otras fuentes considerando la relación con la actividad económica

	Encuesta Población Activa (a)	Muestra por cuotas (b)	Diferencias
Relación con la actividad. Datos totales			
<i>Ambos sexos</i>			
Población + de 16 años	143.243	27.305 (c)	
Activos	80.738	15.371	
— Ocupados	71.883	13.025	
— Parados	8.855	2.346	
Inactivos	62.505	11.934	
Tasa de actividad (%)	56,36	56,29	-0,07 (e)
Tasa de paro (%)	10,97	15,26	4,36 (f)
Relación con la actividad por sexos			
<i>Varones</i>			
Población + de 16 años	69.953	13.191 (c)	
Activos	47.620	9.107	
— Ocupados	43.737	8.145	
— Parados	3.883	962	
Inactivos	22.332	4.084	
Tasa de actividad (%)	68,09	69,04	0,95 (f)
Tasa de paro (%)	8,15	10,56	2,41 (f)
<i>Mujeres</i>			
Población + de 16 años	73.290	14.114 (c)	
Activas	33.118	6.264	
— Ocupadas	28.146	4.880	
— Paradas	4.972	1.384	
Inactivas	40.173	7.850	
Tasa de actividad (%)	45,19	44,38	-0,81 (e)
Tasa de paro (%)	15,01	22,09	7,08 (f)
Tasas de actividad por grupos de edad (%)			
De 16 a 19 años	25,46	38,70 (d)	13,24 (f)
De 20 a 24 años	64,30	62,10	2,20
De 25 a 54 años	80,56	77,90	2,66
De 55 a más años	18,06	21,40	3,34

(a) Datos correspondientes al año 2004. Se han elaborado agregando las cuatro oleadas de la encuesta de Población Activa del año.

(b) En el caso de la muestra por cuotas se trata de la población entrevistada, aquellos con 18 años y más.

(c) El apartado «mayores de 16 años» corresponde al número de personas que han respondido esa pregunta en el caso de la Encuesta de Población Activa, y en el caso de la muestra por cuotas los que tienen 18 y más años.

(d) Téngase en cuenta que en la muestra por cuotas se ha entrevistado a los mayores de 17 años, y no a los mayores de 15.

(e) Valores negativos: Categorías infrarrepresentadas por la muestra por cuotas.

(f) Valores positivos: Categorías sobrerrepresentadas por la muestra por cuotas.

FUENTE:

Instituto Nacional de Estadística (2005b) y Centro de Investigaciones Sociológicas (2004b). Reproducido de Díaz de Rada y Núñez Villuendas (2008: tabla 2.1).

Tras la tasa de actividad procederemos con la situación profesional del entrevistado, cuando trabaja, o el cabeza de familia, cuando no sea así. En la tabla 10 puede comprobarse que las diferencias son —de nuevo— escasas, aunque la muestra por cuotas sobrerrepresenta a los trabajadores por cuenta propia e infrarrepresenta a los asalariados. Cuando se realiza un análisis en detalle se aprecia que esta sobrerrepresentación se produce por el gran número de autónomos¹⁹ entrevistados en la muestra por cuotas, en perjuicio de los empleadores y los miembros de cooperativas. Respecto a los asalariados, apenas hay diferencia entre ambas fuentes cuando se considera el tipo de contrato —atendiendo a la dimensión indefinido/temporal—, si bien la muestra por cuotas consigue más trabajadores del sector privado que los declarados por la Encuesta de Población Activa.

La información proporcionada por las tablas 8, 9 y 10 ha demostrado que el método de cuotas representa adecuadamente el universo en relación al sexo, edad, tasa de actividad y situación profesional. Hemos constatado también algunas ligeras deficiencias en la representatividad conseguida en la tasa de actividad por grupos de edad²⁰, y grandes problemas de representatividad en el caso de la tasa de paro femenino.

Consideramos que estos desajustes en las *estimaciones* pueden estar provocados por las diferencias metodológicas entre ambos estudios. La Encuesta de Población Activa es una encuesta probabilística nominal que utiliza como marco muestral la información recogida por el Censo de Población²¹, mientras que la muestra por cuotas del CIS se realiza con un sistema de rutas aleatorias con selección de las unidades últimas por cuotas de sexo y edad (ver anexo 1). Recordemos, como ya hemos señalado anteriormente, que a los encuestadores se les indica que «cuando encuentren una vivienda en la que no hay nadie, procedan a visitar la siguiente» (CIS, 2004c: 2). Esta forma de proceder explica que la probabilidad que una persona tiene de ser elegida está en relación con el tiempo pasado en el hogar. Consideramos que este motivo explica la sobrerrepresentación de las personas que están en paro; aquellas que pasan más tiempo en la vivienda al no disponer de trabajo. Otra consecuencia derivada de esta situación es el aumento del número de intentos (viviendas visitadas). Así, en la muestra por cuotas los entrevistadores contactan con más de 15 viviendas para lograr una entrevista (Díaz de Rada y Núñez Villuendas, 2008: gráfico 2.1), mientras que la Encuesta de Población Activa visita 1,19 viviendas²² para conseguir un *contacto exitoso* (Instituto Nacional de Estadística, 2005c: 64-65).

¹⁹ Empresarios sin asalariados y trabajadores independientes.

²⁰ La muestra infraestima ligeramente el colectivo entre 20-54 años, al tiempo que sobreestima a los mayores de 55 años.

²¹ Nos referimos, concretamente, a los nombres, apellidos y dirección de las personas que viven en un determinado ámbito.

²² Se han considerado únicamente las «primeras entrevistas», no las «entrevistas sucesivas» del panel.

TABLA 10

Comparación de la muestra con otras fuentes considerando la situación profesional
(en porcentajes)

	Encuesta Población Activa (a)	Muestra por cuotas (b)	Diferencias
Ocupados por situación profesional			
<i>Ambos sexos</i>			
Trabajadores cuenta propia	18,4	20,1	1,4 (c)
Asalariados	81,5	79,6	-1,9 (d)
Otra situación	0,1	0,3	0,2
TOTAL OCUPADOS	71.883	27.114	
<i>Ambos sexos</i>			
Trabajadores cuenta propia			
— Empleadores	5,6	4,7	-0,9 (d)
— Empresarios sin asalariados y trabajadores independientes	10,9	14,8	3,9 (c)
— Miembros de cooperativa y ayuda familiar	2,0	0,6	-1,4
Asalariados			
— Contrato indefinido	56,5	55,2	-1,3
— Contrato temporal	24,9	24,4	-0,5
Otra situación	0,1	0,3	0,2
TOTAL OCUPADOS	71.883	27.114	
Asalariados según tipo de empresa (sector público o privado)			
<i>Asalariados</i>			
— Sector público	19,0	14,4	-4,6
— Sector privado	81,0	85,6	4,6
TOTAL ASALARIADOS	58.883	21.508	

(a) Población ocupada de 16 y más años.

(b) En el caso de la muestra por cuotas, se trata de la población entrevistada, los mayores de 18 años.

(c) Valores positivos: Categorías sobrerrepresentadas por la muestra por cuotas.

(d) Valores negativos: Categorías infrarrepresentadas por la muestra por cuotas.

FUENTE:

Instituto Nacional de Estadística (2005b) y Centro de Investigaciones Sociológicas (2004b).

La estrategia seguida por la muestra por cuotas está determinada por el tiempo disponible para la recogida de información. Así, mientras que el trabajo de campo de los sondeos del CIS se realiza en siete días (ver anexo 1), en la Encuesta de Población Activa se lleva a cabo en seis semanas. Esto permite que los encuestadores de la Encuesta de

Población Activa vuelvan a aquellas viviendas que no han podido ser contactadas en la primera visita (Losilla, 2005: 25). En ocasiones también se utilizan encuestas telefónicas para reducir el número de hogares no contactados (Ballano Fernández y Martínez Vidal, 2000: 266). Además, cuando estas estrategias no consiguen realizar un contacto, el hogar será visitado el siguiente trimestre (no olvidemos que es una encuesta panel); y no es eliminado de la muestra hasta que niega la cooperación en tres momentos distintos (Losilla, 2005: 27).

Otras diferencias están referidas a que la Encuesta de Población Activa utiliza la mayor parte de las estrategias conocidas para animar a la cooperación²³; así como a la diferente carga de trabajo de los entrevistadores. Además, en la EPA los inspectores hacen una revisión periódica de las zonas con el fin de incluir dentro del marco muestral las nuevas construcciones, las viviendas desocupadas o aquellos edificios que eran destinados a otros fines (Díaz de Rada, 2005: 49-56).

Los ajustes mostrados en las tablas anteriores pueden resultar «sospechosos» a algunos lectores, en la medida que se comparan dos encuestas (EPA/Barómetros) que pudieran estar *alterando* la realidad de la misma forma. Aunque ya hemos señalado que se trata de instrumentos muy diferentes²⁴, pudiera ser que la EPA cometiera errores para la medición de la actividad. No consideramos que sea así y, de hecho, la mayor parte de países desarrollados realizan la medición del empleo y paro mediante encuestas (Ballano Fernández y Martínez Vidal, 2000: 270), si bien para evitar suspicacias hemos procedido con una comparación entre los barómetros y la totalidad de la población. El Censo del año 2001 recoge la totalidad de la población, de modo que para que el análisis resulte adecuado utilizaremos los datos relativos a los barómetros de ese mismo año²⁵.

Como puede comprobarse en la tabla 11, existen grandes similitudes entre ambas fuentes en relación a las tasas totales de actividad y paro, produciéndose —por parte de los barómetros— una ligera infraestimación de la tasa de actividad (en un punto) y una sobreesti-

²³ Aumento del plazo para recoger los datos de una sección, empleo de cartas de presentación para anunciar la visita del entrevistador y explicar la importancia de su cooperación en la encuesta; estudio previo de la sección con el fin de fijar los mejores horarios para localizar a los hogares en función de sus características sociodemográficas; utilización de «avisos de visita» que se dejan en la vivienda cuando —tras varios intentos— no se contacta con ningún miembro de la vivienda; visitas sucesivas; volver —en las siguientes oleadas— a viviendas no seleccionadas; incrementos salariales a los entrevistadores según su disponibilidad horaria; entrevistas *proxy*; utilización de diversos procedimientos de recogida de información; estrategias de «conversión» de rechazos para tratar de mitigar las negativas a cooperar, etc.

²⁴ Se ha expuesto con más detalle en Díaz de Rada y Núñez Villuendas (2008).

²⁵ Agradecemos a Adoración Núñez Villuendas, de la Unidad de Campo del CIS, su generosidad al proporcionarnos esta información.

mación de la tasa de paro en algo más de un punto y medio (exactamente 1,68). En el caso de las tasas de actividad por sexos las diferencias son similares: los barómetros presentan una ligera infraestimación de la actividad de los hombres y las mujeres (diferencias de 1,67 y 1,79, respectivamente), aunque las mayores diferencias se encuentran en la sobreestimación de la tasa de paro femenina (4,18 puntos). Sintetizando, poca diferencia entre el Censo y los barómetros en relación a la actividad, tanto considerando la población en general como diferenciada por sexos; lo que nos llevar a concluir que los barómetros reproducen adecuadamente la actividad de la población española.

TABLA 11

Comparación del Censo 2001 y los barómetros (año 2001) considerando la relación con la actividad económica (en porcentajes)

	Censo 2001 (a)	Barómetros 2001 (b)
Relación con la actividad. Datos totales		
<i>Ambos sexos</i>		
Tasa de actividad	55,58	54,57
Tasa de paro	14,16	15,84
Relación con la actividad por sexos		
<i>Varones</i>		
Tasa de actividad	68,17	67,90
Tasa de paro	11,09	11,17
<i>Mujeres</i>		
Tasa de actividad	43,66	42,13
Tasa de paro	18,69	22,87

(a) En el caso del Censo se trata de la población de 16 años y más.

(b) En los barómetros se refiere a los mayores de 17 años.

FUENTE:

Instituto Nacional de Estadística (2005d) y Centro de Investigaciones Sociológicas (2001).

6. CONCLUSIONES

En las páginas anteriores hemos demostrado que el método de cuotas —utilizado tras un sistema de rutas aleatorias que identifica las viviendas donde realizar las entrevistas— proporciona una «adecuada» representatividad del universo en relación a la edad, el sexo, la actividad económica y la situación profesional. No deja de resultar sorprendente este hecho, mucho más cuando hemos demostrado la gran relación existente entre los rasgos sociodemográficos de los entrevistados y el número de orden de la entrevista en la ruta (tabla 6). Si tenemos en cuenta, por otro lado, que las primeras entrevistas de la ruta se hacen por la mañana, y las últimas por la noche (tabla 7), apreciaremos que a medida que avanza el día se seleccionan más activos —y ocupados (laboralmente hablando)—. Dicho de otro modo, por la mañana se entrevista fundamentalmente a no activos (principalmente mujeres y jubilados), y al final del día a población activa y ocupada.

Consideramos que esta situación se explica por la «particularidad» de las tasas de actividad (y paro) por sexos y grupos de edad en España. Como se aprecia en la tabla 12, las tasas de actividad de las mujeres son inferiores a los hombres en todas las edades, diferencia que llega a 25 puntos en el grupo de edad entre 35 a 44 años, y a 35 puntos en el grupo de edad de 55 a 64 años. Una situación similar, pero de signo contrario, se presenta al analizar las tasas de paro. El paro es notablemente superior en el colectivo de mujeres, doblando la tasa masculina en los grupos de edad entre 25 y 54 años.

Las bajas tasas de actividad de las mujeres, unidas a las elevadas tasas de paro de las mujeres que forman parte de la población activa, generan que este colectivo pase mucho tiempo en su vivienda, lo que incrementa su probabilidad de ser entrevistado. En el caso de los hombres se produce la situación opuesta, lo que garantiza el éxito del método de cuotas. La práctica totalidad de los hombres mayores de 25 años forman parte de la población activa²⁶, al tiempo que presentan bajas tasas de paro, de modo que tratar de localizar un varón perteneciente a ese grupo de edad implica —casi siempre— entrevistar a una persona que esté trabajando. Éste es, a nuestro juicio, el gran éxito del método de cuotas en nuestro país, y lo que le lleva a seleccionar un elevado número de trabajadores (recordar tabla 9).

La información proporcionada por las tablas 6 y 7 explica algunas de las ligeras *desviaciones* obtenidas en el momento de comparar los barómetros con el universo del que son extraídos. El hecho que las primeras entrevistas (considerando el número de orden de la entrevista en

²⁶ El 90% de los hombres mayores de 25 años son activos, porcentaje que aumenta hasta el 95% en los que superan los 30 años.

TABLA 12

Tasas de actividad y paro por sexos y grupos de edad (año 2004)

	Varones	Mujeres	Total
<i>Tasa de actividad</i>			
De 18 a 24	50,48	39,01	44,90
De 25 a 34	92,33	77,24	85,00
De 35 a 44	94,61	68,78	81,83
De 45 a 54	90,01	30,84	38,35
De 55 a 64	61,39	26,63	43,55
Más de 65	3,29	1,59	2,36
TOTAL	68,05	45,175	56,35
<i>Tasa de paro</i>			
De 18 a 24	21,53	30,75	25,29
De 25 a 34	8,98	15,33	11,76
De 35 a 44	5,96	13,86	9,25
De 45 a 54	2,70	5,66	3,91
De 55 a 64	6,09	9,25	7,09
Más de 65	0,86	2,03	1,26
TOTAL	8,17	15,02	10,97

FUENTE:

Instituto Nacional de Estadística (2005b).

la ruta, así como la hora de realización) se realicen a mujeres y a «no trabajadores fuera del hogar» explica la sobreestimación de la tasa de paro en las mujeres entrevistadas (seleccionadas con el método de cuotas)²⁷: esto se produce porque —una vez cumplimentadas las cuotas de mujeres durante las primeras horas del día— se dejan las entrevistas de tarde para entrevistar a los hombres. La *búsqueda* de hombres implica —según la información mostrada en la tabla 12— la localización de hombres trabajadores. Proponer a los entrevistadores que realicen por la tarde un determinado número de entrevistas a mujeres podría —desde nuestro punto de vista— disminuir la elevada tasa de paro de este colectivo²⁸.

Lo expuesto hasta el momento explica los ACTUALES procesos de selección muestral empleando el método de cuotas. Ahora bien, pensar sobre el FUTURO implica tener en cuen-

²⁷ Nos referimos, concretamente, a los datos mostrados en la tabla 9. Allí indicamos que la EPA sitúa la tasa de paro femenino en un 15,01%, mientras que el 22,09% de las mujeres entrevistadas por la muestra por cuotas se encuentran en paro.

²⁸ Realizamos esta recomendación siendo plenamente conscientes de la dificultad de la labor del entrevistador, y la tremenda complicación de solicitar «nuevas tareas» a este colectivo.

ta el progresivo aumento en las tasas de actividad de las mujeres más jóvenes (obsérvese que en el grupo entre 25 y 34 años ésta llega hasta el 77,24%), y el consiguiente descenso de mujeres que realizan tareas domésticas no remuneradas. Esta situación va a generar —en los próximos años— algunos *déficits* a la hora de representar adecuadamente la población objeto de estudio; en la medida que no será posible realizar las entrevistas matutinas a los «no activos», y las vespertinas a los «activos» (tal y como ocurre actualmente). Existen, desde nuestro punto de vista, tres soluciones para esta situación: 1) utilizar cuotas que consideren la relación con la actividad, lo que complica —y encarece— enormemente el proceso de selección; 2) realizar un mayor número de entrevistas por la tarde, como se procede actualmente con las encuestas telefónicas; y 3) cambiar el método de selección de los entrevistados últimos por otro método que respete la aleatoriedad (números aleatorios, último cumpleaños, etc.). La última solución es, desde nuestro punto de vista, la mejor, en la medida que sustituye un método de selección no aleatorio por otro aleatorio. Ésta es la fórmula que han adoptado, en nuestro país, la Encuesta Social Europea (Cuxart y Riba, 2005: 21-40), el Eurobarómetro (Clemente, 2006: 9), etc. Se trata, por otro lado, de una tendencia dominante en la investigación internacional, y como muestra basta considerar el hecho que la Encuesta Mundial de Valores —en su oleada del año 2008— abandonará la muestra por cuotas sustituyéndola por un muestreo aleatorio.

Ahora bien, es un cambio que tiene importantes implicaciones en términos económicos y de tiempo. Estimaciones realizadas en el Reino Unido consideran que el trabajo de campo de un muestreo aleatorio con una duración de siete días se reduce a la mitad con un muestreo con cuotas. Respecto al importe económico, un muestreo con cuotas supone una reducción del 30-40% en el coste total de la investigación (Worcester y Downham, 1986: 105). Se trata de unas cifras que no se mantienen en este momento por el gran incremento en el número de «contactos infructuosos» (del método de cuotas) consecuencia de contactar con personas con *cuota cubierta*. De hecho, el análisis de la Encuesta Navarra de Condiciones de Vida —que en el año 2001 se realizó mediante rutas y cuotas y en el año 2003 con rutas y el método de Kish para la selección de los entrevistados— permite realizar una somera valoración de tal cambio. En términos económicos, la muestra probabilística tiene un coste ligeramente superior, con una diferencia que no llega al euro por encuesta respondida en esta investigación. Considerando que «es más importante la representatividad que el tamaño muestral» (Alderete, 1996: 32), este aumento de costes podría sufragarse con una ligera reducción del tamaño muestral.

Las implicaciones en términos de calendario son más elevadas, puesto que el muestreo probabilístico implicó casi duplicar la duración del trabajo de campo, consecuencia de la realización de tres contactos cuando no se localiza a la persona seleccionada. Esto genera que sería tremendamente complicado utilizar este método en investigaciones con cortos

trabajos de campo, como ocurre con los barómetros (ocho días) o con las encuestas electorales, que, por motivos legales (prohibición de publicación de sondeos cinco días antes de los comicios), deben realizar el *campo* en pocos días.

REFERENCIAS UTILIZADAS

- ALDERETE, J. (1996): «¿Utilizamos muestras representativas?», *Investigación y Marketing*, vol. 50, pp. 32-35.
- AZORÍN, F., y SÁNCHEZ-CRESPO, J. L. (1986): *Métodos y aplicaciones del muestreo*, Madrid: Alianza.
- BALLANO FERNÁNDEZ, C., y MARTÍNEZ VIDAL, M. A. (2000): «Características de la no respuesta en la Encuesta de Población Activa», *Estadística Española*, vol. 42, n.º 146, pp. 263-277.
- BRADBURN, N. M. (1992): «A Response to the Nonresponse Problem», *Public Opinion Quarterly*, vol. 56, pp. 391-397.
- BUTCHER, B. (1995): «Sampling methods: an overview and review», *Survey Methods Centre Newsletter*, vol. 15, n.º 1, pp. 4-8.
- CAMACHO, C.; PRADO, C.; ROMERO, B., y VALERA, S. (2000): «¿Cómo hacen encuestas los encuestadores?», *Metodología de Encuestas*, vol. 2, pp. 117-133.
- CEA D'ANCONA, M. A. (2004): *Métodos de encuesta. Teoría y práctica, errores y mejora*, Madrid: Síntesis.
- CENTRO DE INVESTIGACIONES SOCIOLOGICAS (2001): *Barómetros del año 2001*.
- (2004a): *Barómetro de enero*, Madrid: Centro de Investigaciones Sociológicas, estudio 2554.
- (2004b): *Barómetros del año 2004*. Fichero agregado uniendo los datos comunes de los estudios 2554, 2556, 2558, 2561, 2565, 2568, 2570, 2573, 2577, 2581 y 2584.
- (2004c): *Instrucciones para la correcta aplicación de la muestra*.
- (2006): *El CIS*. http://www.cis.es/cis/opencms/ES/6_cis/.
- CLEMENTE, J. (2006): «Eurobarómetro, el pulso demoscópico de Europa», en *XIV Seminario de Investigaciones Políticas y Sociológicas*, organizado por AEDEMO y celebrado el 25 de mayo de 2006, pp. 5-24.
- COXON, A. P. M. (ed.) (1992): *Sample Design Issues in a Panel Survey: The Case of the British Household Panel Study*, Working Papers of the ESRC Research Centre on Micro-social Change, Paper 3, Colchester: Universidad de Essex.
- CUXART, A., y RIBA, C. (2005): «Aspectos metodológicos de la Encuesta Social Europea», en M. Torcal, L. Morales y S. Pérez-Nievas, *España: Sociedad y política en perspectiva comparada*, Valencia: Tirant lo Blanch, pp. 21-40.
- CZAJA, R., y BLAIR, J. (1996): *Designing Surveys*, California: Pine Forge.
- DAVIS, J. A., y SMITH, T. W. (1992): *The NORC General Social Survey. A User's Guide*, Newbury Park, CA: Sage.
- DEVILLE, J. C. (1991): «A theory of Quota Surveys», *Survey Methodology*, vol. 17, n.º 2, pp. 163-181.
- DÍAZ DE RADA, V. (2001): *Organización y gestión de los trabajos de campo con encuestas personales y telefónicas*, Barcelona: Ariel.

DÍAZ DE RADA, V. (2005): *Manual del Trabajo de Campo en la encuesta*, Madrid: Centro de Investigaciones Sociológicas.

DÍAZ DE RADA, V., y NÚÑEZ VILLUENDAS, A. (2005): *La participación en encuestas. Un indicador de las transformaciones sociales del siglo XXI*. Investigación patrocinada por el Centro de Investigaciones Sociológicas dentro de la «Convocatoria de Ayudas a la Investigación Sociológica para el año 2005».

— (2008): *Estudio de las incidencias en la investigación con encuesta. El caso de los barómetros del CIS*, Madrid: Centro de Investigaciones Sociológicas, colección Monografías.

GONZÁLEZ BLASCO, P.; GONZÁLEZ-ANLEO, J.; ELZO IMAZ, J.; GONZÁLEZ-ANLEO SÁNCHEZ, J. M.; LÓPEZ RUIZ, J. A., y VALLS IPARAGUIRRE, M. (2006): *Jóvenes Españoles 2005*, Madrid: Fundación SM.

GROVES, R. M., y COUPER, M. P. (1998): *Nonresponse in Household Interview Surveys*, Nueva York: Wiley.

INSTITUTO NACIONAL DE ESTADÍSTICA (2002): *Encuesta de Población Activa. Informe Técnico*, Madrid: INE.

— (2004): «¿Cuántos somos en casa?», en *Cifras INE. Boletín Informativo del Instituto Nacional de Estadística*, n.º 6. Disponible en <http://www.ine.es/revistas/cifraine/0604.pdf>.

— (2005a): *Revisión del Padrón municipal 2004, explotación a 1 de enero de 2004*, Madrid: INE. Disponible en www.ine.es/inebase/cgi/axi.

— (2005b): *Encuesta de Población Activa 2004*, Madrid: INE. Disponible en www.ine.es.

— (2005c): *Evaluación de la calidad de la EPA 2004*, Madrid: INE. Disponible en www.ine.es/docutrab/eval_epa/evaluacion_epa04.pdf.

— (2005d): *Censo de Población y Viviendas 2001*, Madrid: INE. Disponible en www.ine.es.

JAIME CASTILLO, A. J. (2004): «George Gallup y los orígenes de la investigación social por encuesta en los Estados Unidos», en J. Andreu Abela (ed.), *Investigación Aplicada de clásicos de la Sociología: viejos y nuevos problemas de investigación social*, Granada: Fundación Centra, pp. 133-164.

KISH, L. (1965): *Survey Sampling*, Nueva York: Wiley.

— (1998): *Quota Sampling: Old Plus New Thought*, Working Paper, Universidad de Michigan. <http://www.websm.org/uploadi/editor/1132389572Kish.doc>.

LESSLER, J. T., y KALSBECK, W. D. (1992): *Nonsampling Error in Surveys*, Nueva York: Wiley.

LOSILLA, J. (2005): «Recogida de datos de una encuesta continua dirigida a los hogares: la Encuesta de Población Activa», en *Los Trabajos de Campo en las encuestas del INE*, Curso impartido en la Escuela de Estadística de las Administraciones Públicas, del Instituto Nacional de Estadística, 21-23 de junio.

LYNN, P. (1995): «Evidence of the inaccuracy of quota samples», *Survey Methods Centre Newsletter*, vol. 15, n.º 1, pp. 20-23.

MARSH, C., y SCARBROUGH, E. (1990): «Testing nine hypotheses about quota sampling», *Journal of the Market Research Society*, vol. 32, n.º 4.

MOON, N. (1995): «The case of quota sampling», *Survey Methods Centre Newsletter*, vol. 15, n.º 1, pp. 9-12.

MORTON-WILLIAMS, J. (1993): *Interviewer Approaches*, Vermont: Dartmouth.

MOSER, C. A. (1952): «Quota Sampling», *Journal of the Royal Statistical Society, Series A*, vol. 115, pp. 411-423.

MOSER, C. A., y KALTON, G. (1971). *Survey Methods in Social Investigation*, Aldershot: Gower Publishing Company Limited.

- NÚÑEZ VILLUENDAS, A. (2005): «Incidencias en la entrevista personal en la investigación por encuesta», *REIS*, n.º 109, pp. 219-236.
- O'MUIRCHEARTAIGH, C. (1995): «The case of probability sampling (against quota sampling)», *Survey Methods Centre Newsletter*, vol. 15, n.º 1, pp. 13-15.
- ORTON, S. (1995): «Evidence of the efficiency of quota samples», *Survey Methods Centre Newsletter*, vol. 15, n.º 1, pp. 16-19.
- PARK, A.; CURTICE, J.; THOMSON, K.; JARVIS, L., y BROMLEY, C. (2001): *British Social Attitudes, the 18th Report. Public Policy, Social Ties*, Londres: Sage.
- PASADAS DEL AMO, S.; SORIA ZAMBRANO, M., y ZARCO URIBE-ECHEVARRÍA, M. (2006a): «¿Dónde están los jóvenes? Un análisis de las dificultades de localización y acceso al segmento juvenil en las encuestas telefónicas», IV Congreso de Metodología de Encuestas, Pamplona, 20-22 de septiembre.
- (2006b): «¿Importa el prefijo? Una aproximación a las diferencias territoriales en las pautas de respuesta a las encuestas telefónicas», en *Metodología de Encuestas*, vol. 7, n.º 3, pp. 3-12.
- PASADAS DEL AMO, S.; ZARCO URIBE-ECHEVARRÍA, M., y SORIA ZAMBRANO, M. (2004): *La telefonía móvil: un nuevo reto para las encuestas telefónicas*, Córdoba: Documento de Trabajo del Instituto de Estudios Sociales de Andalucía, Consejo Superior de Investigaciones Científicas, IESA 15-04.
- RODRÍGUEZ OSUNA, J. (1991): *Métodos de Muestreo*, Madrid: Centro de Investigaciones Sociológicas, colección Cuadernos Metodológicos, n.º 1.
- (2000): «La muestra: teoría y aplicación», en M. García Ferrando, J. Ibáñez y F. Alvira (eds.), *El Análisis de la Realidad Social*, Madrid: Alianza, 3.ª ed., pp. 445-482 (edición original 1986).
- ROSE, D.; BUCK, N., y CORTI, L. (1991): *Design Issues in the British Household Panel Study*, Working Papers of the ESRC Research Centre on Micro-social Change, Paper 2, Colchester: Universidad de Essex.
- ROSSI, P. H.; WRIGHT, J. D., y ANDERSON, A. B. (1983): «Sample surveys: History, current practice, and future prospects», en P. H. Rossi, J. D. Wright y A. B. Anderson, *Handbook of survey research*, San Diego: Academic Press, pp. 1-20.
- SÁNCHEZ-CRESPO, J. L. (1973): *Muestreo de poblaciones finitas aplicado al diseño de encuestas*, Madrid: INE.
- SOCIAL AND COMMUNITY PLANNING RESEARCH - SCPR (1995): «Quota versus probability sampling», monográfico del *Survey Methods Centre Newsletter*, vol. 15, n.º 1.
- SMITH, T. W. (1981): *The Hidden 25 percent: An Analysis of Nonresponse of the 1980 General Social Survey*, GSS Technical Report No. 25, Chicago: National Opinion Research Center.
- STEPHENSON, C. B. (1979): «Probability Sampling With Quotas: An Experiment», en *Public Opinion Quarterly*, vol. 43, pp. 477-496.
- SUDMAN, S. (1976): *Applied Sampling*, Nueva York: Academic Press.
- WORCESTER, R., y DOWNHAM, J. (1986): *Consumer Market Research Handbook*, Holland: Elsevier.

ANEXO 1: FICHA TÉCNICA DE LOS BARÓMETROS REALIZADOS POR EL CIS

Ámbito: Nacional.

Universo: Población española de ambos sexos de 18 años y más, excepto residentes en Ceuta y Melilla.

Tamaño de la muestra:

Enero (estudio 2554): diseñada, 2.500 entrevistas; realizada, 2.489 entrevistas.
 Febrero (estudio 2556): diseñada, 2.500 entrevistas; realizada, 2.490 entrevistas.
 Marzo (estudio 2558): diseñada, 2.500 entrevistas; realizada, 2.494 entrevistas.
 Abril (estudio 2561): diseñada, 2.500 entrevistas; realizada, 2.493 entrevistas.
 Mayo (estudio 2565): diseñada, 2.500 entrevistas; realizada, 2.496 entrevistas.
 Junio (estudio 2568): diseñada, 2.500 entrevistas; realizada, 2.479 entrevistas.
 Julio (estudio 2570): diseñada, 2.500 entrevistas; realizada, 2.487 entrevistas.
 Septiembre (estudio 2573): diseñada, 2.500 entrevistas; realizada, 2.483 entrevistas.
 Octubre (estudio 2577): diseñada, 2.500 entrevistas; realizada, 2.494 entrevistas.
 Noviembre (estudio 2581): diseñada, 2.500 entrevistas; realizada, 2.496 entrevistas.
 Diciembre (estudio 2584): diseñada, 2.500 entrevistas; realizada, 2.458 entrevistas.
 Muestra agregada: 27.359

Afijación: Proporcional.

Ponderación: No procede.

Puntos de muestreo: 168 municipios y 49 provincias.

Procedimiento de muestreo: Polietápico, estratificado por conglomerados, seleccionando las unidades primarias de muestreo (municipios) y las unidades secundarias (secciones) de forma aleatoria proporcional, y las unidades últimas (individuos) por rutas aleatorias y cuotas de sexo y edad.

Los estratos se han formado por el cruce de las 17 Comunidades Autónomas con el tamaño del hábitat, dividido en siete categorías: menor o igual a 2.000 habitantes; de 2.001 a 10.000; de 10.001 a 50.000; de 50.001 a 100.000; de 100.001 a 400.000; de 400.001 a 1.000.000, y más de 1.000.001 habitantes.

Los cuestionarios se han aplicado mediante entrevista personal en los domicilios.

Error muestral: Para un nivel de confianza del 95,5% (dos sigmas), y $P = Q$, el error real en cada barómetro es de $\pm 2,0\%$ para el conjunto de la muestra y en el supuesto de muestreo aleatorio simple. Considerando la muestra agregada, el error muestral se reduce al 0,60%, considerando los mismos parámetros de nivel de confianza, P y Q, y en el supuesto de muestreo aleatorio simple.

Fecha de realización:

Enero (estudio 2554): del sábado 10 al jueves 15.
 Febrero (estudio 2556): del viernes 20 al viernes 27.
 Marzo (estudio 2558): del martes 16 al domingo 21.
 Abril (estudio 2561): del jueves 22 al martes 27.
 Mayo (estudio 2565): del jueves 27 al lunes 31.
 Junio (estudio 2568): del viernes 25 al miércoles 30.
 Julio (estudio 2570): del viernes 9 al viernes 16.
 Septiembre (estudio 2573): del viernes 24 al jueves 30.
 Octubre (estudio 2577): del jueves 21 al martes 26.
 Noviembre (estudio 2581): del viernes 19 al miércoles 24.
 Diciembre (estudio 2584): del miércoles 15 al lunes 20.

ANEXO 2: INCIDENCIAS EN EL ESTUDIO *JÓVENES ESPAÑOLES 2005*

	TOTAL
No hogar	2.754
Negativa total	9.137
No cuota	21.751
Negativa (Sí cuota)	4.772
Realización	4.000
Total contactos	42.414
Media	10.604
No cuota/total (21.751/42.414)	51,28
No cuota/contactos (21.751/39.660)	54,80

FUENTE:

González Blasco *et al.* (2006).