

CALIDAD EN LA RECOGIDA DE INFORMACIÓN DE LA ENCUESTA DE CONDICIONES DE VIDA 2004 DE LA C.A. DE EUSKADI

Javier San Vicente Alfambra
J_SanVicente@eustat.es
Instituto Vasco de Estadística (EUSTAT)

RESUMEN

La investigación por encuestas tradicionalmente ha centrado su atención en los errores de muestreo y, en menor medida en alguno de los denominados errores de no muestreo. Entre estos últimos, los más destacados por los esfuerzos dedicados han sido los errores de especificación, los errores de medida, relacionados con el cuestionario, y los errores de proceso.

Sin embargo, en pocas ocasiones los errores de cobertura y de no respuesta han sido el centro del análisis. Baste para comprobarlo la escasa literatura que al respecto hay en castellano.

En este artículo se analizan las incidencias recogidas durante la fase de campo con ocasión de la realización de la Encuesta de Condiciones de Vida 2004 de Eustat, sus consecuencias en la cobertura muestral y en la no respuesta, caracterizando ambos fenómenos según las variables incluidas en la muestra.

Palabras clave: incidencias, no respuesta, errores de cobertura

Introducción

Todas las encuestas basadas en la utilización de muestras en lugar de toda la población están sometidas al error de muestreo. El error de muestreo es la diferencia entre el valor del parámetro en la población y el valor estimado mediante la muestra. Todas las encuestas, estén o no basadas en la utilización de muestras, como por ejemplo los censos, están sometidas al error de no muestreo.

Biemer y Liberg (2003) identifican cinco componentes del error de no muestreo. A saber,

- 1.- los errores de especificación: cuando los conceptos utilizados para definir el fenómeno que se quiere medir no están correctamente formulados;
- 2.- los errores de cobertura o marco originados por las omisiones: duplicados o especificaciones imprecisas de las unidades sobre las que se realiza el muestreo;
- 3.- la no respuesta: o fallo en obtener la respuesta de alguno de los elementos de la muestra;
- 4.- los errores de medida: producidos en la transformación de los objetivos de la encuesta en preguntas utilizables para obtener respuestas, recogerlas, grabarlas y procesarlas, y
- 5.- los errores de proceso: que comprenden los errores de edición, codificación, de entrada de datos y de programación.

Durante la fase de campo la atención se centra en los errores de cobertura y en los errores de no respuesta, ya sea total o parcial, que aparecen recogidos en las llamadas incidencias de campo. Las incidencias son los apuntes sobre el estado provisional, mientras dura la investigación, o definitivo, una vez ésta concluye, en que se encuentra la una unidad objeto de estudio, así como sobre la causa que motiva tal situación.

Nunca la información que se pretende recoger mediante una encuesta está completa. Incluso ni en aquellos casos en que la obligatoriedad viene prescrita por ley, como sucede con las encuestas promovidas por organismos oficiales, se consigue el cien por cien de los cuestionarios correspondientes a la muestra tal y como fue diseñada originalmente.

Una cantidad variable de hogares de los inicialmente llamados a responder, han de ser sustituidos por negativa a colaborar por parte del individuo seleccionado para informar, por la imposibilidad de establecer contacto en los horarios estipulados, por ausencia prolongada, por incapacidad física o mental para hacerlo o por cualquier otro motivo. Igualmente, algunas de las viviendas enumeradas en el marco muestral no están habitadas en el período de referencia, han cambiado de uso, han desaparecido o, sencillamente, resulta imposible localizarlas. Estas situaciones en relación a las viviendas, las familias o los individuos que la componen resultan determinantes en la calidad final de las estimaciones derivadas de la muestra.

Este artículo pretende dar cuenta de cómo estos aspectos han influido en la calidad final de los datos recogidos en la Encuesta de Condiciones de Vida 2004 de la Comunidad Autónoma de Euskadi (ECV04).

Diseño de la EC04

El universo poblacional de la encuesta está formado por las viviendas ocupadas y la población residente en esas viviendas de la Comunidad Autónoma del País Vasco en el año 2004. Quedan excluidas las viviendas colectivas y las personas o familias que puedan estar en ellas.

Siguiendo el tamaño de muestra utilizado habitualmente en las encuestas de familias realizadas por EUSTAT, en ECV04 se partió de una muestra teórica de 5.000 hogares. Para su distribución se ha utilizado un diseño de muestreo probabilístico en dos etapas, con selección primera de las viviendas y dentro de cada vivienda, segunda etapa, selección de un individuo de forma aleatoria.

El método de muestreo de las viviendas es el aleatorio estratificado y dentro de las viviendas el muestreo para seleccionar las personas es aleatorio. En la primera etapa los estratos los componen los tres Territorios Históricos, el año de construcción del edificio, con tres grupos (hasta 1940, de 1941 a 1970 y 1971 y posterior) y el tamaño de la familia, con cuatro (1, 2, 3 - 4 y 5 y más personas). Por combinación de las modalidades anteriores se obtienen 36 estratos diferentes.

El reparto de la muestra se hizo proporcional a la raíz cuadrada de la población de cada Territorio aplicando, además, el criterio de similar error para los tres Territorios. Se llegó al reparto final siguiente: Araba 1.500, en Bizkaia 1.900 y Gipuzkoa 1.600 encuestas. El reparto posterior por estratos dentro de cada Territorio es proporcional al número de viviendas.

La base muestral de las viviendas fue el Directorio de Viviendas con información de las viviendas y de la población a 31 de diciembre de 2001. Adicionalmente se llevó a cabo un contraste con los ficheros de defunciones

posteriores para eliminarlas de la muestra. De esta base se extraen las viviendas familiares con los criterios de estratificación mencionados.

En lo que se refiere a los individuos, se cruzó la información del Directorio con la del Registro de Población para verificar las personas que viven en cada vivienda y hacer, de forma aleatoria, el sorteo del individuo que tiene que contestar dentro de cada familia. Dentro de esa misma familia se selecciona un individuo sustituto para casos de ausencia prolongada, enfermedad etc.

Además del fichero de viviendas titulares se extrajo un fichero de 5.000 viviendas sustitutas, también con una persona titular y otra suplente designadas para completar la encuesta. La sustitución de una vivienda se hace por otra del mismo estrato (Territorio, año de construcción y tamaño de la vivienda).

La encuesta se compone de dos cuestionarios: uno familiar, cumplimentable por cualquier miembro idóneo para contestar a las preguntas acerca de los miembros del colectivo familiar, y otro individual, aplicado únicamente a la persona designada en el sorteo.

Se combinaron varios métodos de encuestación. Las entrevistas eran telefónicas cuando la vivienda disponía de teléfono fijo conocido. En otro caso, o cuando expresamente así lo solicitaba el encuestado, la información se recogía mediante visita personal. La proporción inicial para las viviendas titulares era de un 68,6% telefónico un 31,4% personal. Por primera vez para una encuesta a hogares se ofrecía a los encuestados la posibilidad de completar los dos cuestionarios a través de la Web.

El marco de la ECV04. Errores de cobertura.

El marco de una población es el conjunto de unidades (personas, familias...) para las cuales puede ser hecha algún tipo de enumeración previa a la selección de la muestra de la encuesta. La función del marco de muestreo es proveer un medio para escoger los miembros del objeto de población que han de ser entrevistados en la encuesta.

El marco de muestreo tiene implicaciones significativas sobre el coste y la calidad de cualquier encuesta. En las encuestas a hogares los marcos de muestreo defectuosos son fuente común de error de no muestreo, particularmente la falta de cobertura de importantes subgrupos de población. Los errores de cobertura se refieren a la discrepancia entre los estadísticos calculados sobre el marco de población y los que resultarían de calcularlos sobre la población objeto de estudio.

En un sentido estadístico, el marco de muestreo debe de capturar la población objeto de estudio en un momento dado. Mas allá de esto, un perfecto marco es aquel que es completo, preciso y actualizado. En la medida en que exista algún fallo en alcanzar cada una de estas propiedades produce encuestas que están sesgadas en varias direcciones, pero frecuentemente en una subestimación de la población objetivo (Groves R.M. 1989). La importancia de la falta de cobertura depende de las características de las unidades omitidas. La población marco debe constituir, por lo tanto, una colección actualizada y exhaustiva de las unidades de muestreo, sin solapamientos, con límites bien definidos y fácilmente identificables, sin duplicidades, sin omisiones, sin unidades extrañas ni vacías.

La falta de cobertura impide la cooperación de un número de unidades muestrales puesto que determinados individuos no pueden ser seleccionados en la muestra, pero no es considerada como falta de respuesta total, sino que es computada dentro de otros errores de no muestreo presentes en la investigación social (Groves R.M. 1989).

Durante el trabajo del entrevistador en campo inevitablemente aparecen una serie de situaciones que afectan a las viviendas y que impiden que finalmente se aplique el instrumento de medida previsto, el cuestionario. Las más usuales son la imposibilidad de identificar sobre el terreno la dirección que figura en su relación de viviendas, la vivienda utilizada para otros fines distintos de la residencia y un variado elenco de situaciones menores bajo el epígrafe de otras incidencias, entre las que destacarían las viviendas ruinosas, en obras y los establecimientos colectivos.

La ECV04 tiene como población objetivo dos poblaciones relacionadas: los individuos y las familias residentes en la C.A. de Euskadi. El marco muestral, la relación física de unidades de muestreo, fue, como en el caso de todas las encuestas a hogares de Eustat, el directorio de viviendas. Se seleccionaron aquellas viviendas que tuvieran empadronado como mínimo un individuo mayor de 16 años.

Tabla 1. Incidencias de viviendas titulares por tipo según modo de entrevista y Territorio Histórico (%)

	C.A. de Euskadi	Personal	Telefónica	Araba	Bizkaia	Gipuzkoa
Total	100	100	100	100	100	100
Localizada	82,5	84,0	81,8	84,8	83,2	79,6
No localizada	13,5	4,7	17,5	10,9	12,8	16,7
Temporada	0,9	2,5	0,2	1,3	0,6	0,9
Deshabitada	2,1	6,1	0,2	2,1	2,2	1,9
Otros fines	0,2	0,3	0,1	0,2	0,2	0,1
Desaparecida	0,1	0,3	0,1	0,0	0,3	0,1

Reforma	0,4	1,0	0,1	0,4	0,5	0,3
Ruinosas	0,2	0,6	0,0	0,3	0,3	0,1
Inaccesible	0,1	0,2	0,0	0,0	0,0	0,2
Establecimiento	0,1	0,3	0,0	0,1	0,1	0,1
Otros	0,0	0,1	0,0	0,0	0,0	0,1

La utilización en la EC04 de métodos mixtos de entrevista, entre los que se incluye el teléfono, dificulta discernir la elegibilidad de las viviendas muestreadas. Un número de teléfono del que nunca se obtiene respuesta puede corresponder a un teléfono averiado o a un hogar en el que no hay nadie en casa, segunda residencia, vivienda en obras etc... La cuestión no es trivial si se considera que las viviendas de las que no se obtiene respuesta vía telefónica pueden considerarse no contactadas, de elegibilidad dudosa o no elegibles. Solo en el primer caso, esas viviendas (o individuos) forman parte directa del denominador para el cálculo de tasas de respuesta, cooperación y rechazo. En el caso de elegibilidad dudosa pueden formar parte del denominador vía coeficiente de ponderación.

En la fase de campo se dieron instrucciones para que las viviendas de cuyo teléfono no se obtenía respuesta una vez de cumplido el calendario de llamadas, se anotaran como no localizadas. De esta forma, tienen la consideración de elegibilidad desconocida y son incluidos en el cálculo de tasas. (The American Association for Public Opinion Research, 2006) Por otra parte, todo teléfono que no se correspondía con la vivienda listada en el marco no era objeto de entrevista al no formar parte de la muestra, calificándose, igualmente, como vivienda no localizada.

Aparentemente y de cara a la calidad de la cobertura, todos estos impedimentos aconsejarían utilizar preferentemente marcos basados únicamente en direcciones de viviendas, con el consiguiente desplazamiento a campo (tabla 1). Así sería si las incidencias de viviendas se limitaran a su localización, ya que hay 13% puntos porcentuales más de viviendas no localizadas mediante entrevista telefónica que mediante entrevista personal. Sin embargo, la frecuencia de todos los tipos de incidencias en el marco es muy similar en uno y otro caso, 16,1% en entrevista personal frente a 18,2% en telefónica. La entrevista personal, lógicamente, permite un mayor detalle y precisión en la categorización de la incidencia.

La repercusión de las incidencias varía sustancialmente entre las distintas categorías de cada una de las variables de estratificación. El tamaño familiar y el año de construcción de la vivienda discriminan nítidamente la probabilidad de encontrar una incidencia (tabla 2).

Las diferencias territoriales vendrían explicadas en el caso de Araba por concentrarse el 76% de sus viviendas principales en su capital, municipio de 225.631 habitantes, donde como en el resto de municipios grandes son menos abundantes los errores de marco. La misma razón de concentración de población en municipios de gran tamaño sería válida para Bizkaia y justificaría las diferencias de los dos Territorios con Guipúzcoa.

Como se ha comentado anteriormente, la ECV04 está basada en un muestreo aleatorio estratificado por Territorio Histórico, tamaño familiar y año de construcción de la vivienda. Se escogieron éstas como variables de estratificación en base al análisis de los resultados de la anterior encuesta, en 1999.

La relación entre antigüedad de la vivienda y frecuencia de las incidencias es directa. Cuanto más antigua es la vivienda más se incrementa la posibilidad de aparición de una incidencia. Aunque cuantitativamente es el grupo menos numeroso, menos de una de cada diez viviendas del total, se ha detectado alguna incidencia en casi el 13% de ellas. Como en el resto de estratos, la incidencia más frecuente es la de vivienda no localizada, pero también es típica la de vivienda deshabitada o vacía y, comparativamente con los estratos de viviendas más recientes, la vivienda desaparecida o en reforma.

De igual modo parece existir cierta asociación entre imperfecciones en el marco muestral y tamaño familiar, más evidente para las viviendas de un único residente y en las de tres y cuatro personas, aunque de signo contrario. En las viviendas de una única persona las incidencias abundan. Además de la no localización, la más frecuente es la de residencia de temporada y vivienda deshabitada. Mientras en las de tres y cuatro personas, casi el 46% del total de las viviendas de la muestra, en el total de incidencias su participación no alcanza el 40%.

Tabla 2. Viviendas titulares por tamaño familiar, año de construcción y Territorio según incidencia (%).

	Total	Localizada	Incidencia
Total	100	82,5	17,5
Tamaño familiar			
1 persona	18,1	16,9	24,0
2 personas	24	24,1	23,6
3-4 personas	45,6	47	39,2
5 y más personas	12,2	12	13,2

Año construcción de la vivienda			
Hasta 1940	8,9	8,9	12,8
De 1941 a 1970	43,1	43,1	44,5
1971 y posterior	47,9	47,9	42,7
Territorio Histórico			
Bizkaia	38,0	38,3	36,5
Gipuzkoa	32,0	30,9	37,3
Araba	30,0	30,8	26,2

En resumen, de los treinta y seis estratos establecidos en el diseño, media docena superan el 20% de incidencias y básicamente se corresponden a viviendas construidas antes de 1940, con un único residente y preferentemente en el Territorio de Gipuzkoa.

Con respecto a las diferencias entre viviendas en muestra inicial y de titulares finalmente elegibles para la realización de la encuesta, se puede concluir que las incidencias de viviendas, a pesar de constituir una proporción elevada, no han alterado sustancialmente la composición del marco tal y como se diseñó en un principio, al no existir grandes diferencias sea cual sea el estrato considerado.

Por último, hay que señalar que los errores de cobertura son difíciles de estimar y requieren investigaciones especiales o la utilización de fuentes externas a la encuesta. El primer caso se conoce como método de reenumeración y consiste en volver a enumerar las unidades de una submuestra de la muestra original. Su costo y duración reduce su empleo a operaciones de gran volumen, como el censo, o de carácter repetitivo, como en las encuestas tipo panel. Un método menos gravoso, aunque menos preciso, consiste en comparar los resultados de la muestra para ciertas variables de las que se conoce su distribución en la población total. En este caso, lo que no se puede precisar es en qué conjunto de población se haya el error, si en la muestra o en la población.

La no respuesta total

Se distingue entre dos tipos de no respuesta: no respuesta total y no respuesta parcial. La primera se produce cuando no se consigue realizar una entrevista con la persona seleccionada. La segunda cuando se realiza la entrevista pero falta la información para una o varias de las preguntas del cuestionario (Sánchez Carrión J.J., 2000).

En la ECV04 se registran ambos tipos de no respuesta. La forma más habitual de no respuesta parcial es el abandono o la negativa a continuar antes de finalizar la entrevista. La falta de información en alguna cuestión quedó circunscrita a la pregunta sobre ingresos económicos, de carácter voluntario y situada al final del cuestionario. En el resto del cuestionario los entrevistados eran instados por el encuestador a dar una respuesta.

No es objeto de estas notas este tipo de no respuesta. Su incidencia fue escasa, aunque cabe hacer alguna matización. Un número de encuestas no precisable por haberse contemplado como incidencia temporal y no definitiva en la hoja de ruta, fueron comenzadas por Internet, abandonadas y completadas posteriormente vía entrevista personal o telefónica a solicitud de agente, a quien se pasaba periódicamente la relación de viviendas con esa incidencia. Otras en la misma situación engrosaron finalmente las negativas.

La no respuesta total deriva de la incapacidad de medir sobre todas las unidades de la muestra todas las variables de interés; resulta inevitable en las encuestas basadas en muestreos probabilísticos a pesar de los esfuerzos económicos y de recursos humanos tanto en la fase de diseño como durante la recogida de datos. Las causas de la no respuesta total son diversas pudiéndose descomponer en: imposibilidad de contactar con un respondiente o con el respondiente seleccionado, incapacidad para responder por parte del interesado y el rechazo a prestar su colaboración.

Estrategias en campo para mitigar la falta de respuesta

Existe una abundante literatura sobre los factores que intervienen en la no respuesta, estrategias de prevención y posibles actuaciones durante el trabajo de campo (Díaz de Rada, Vidal 2000). Con el objetivo de minimizar el sesgo derivado de la no respuesta total durante la fase de campo se tomaron las siguientes medidas:

1.- Se utilizaron técnicas mixtas de encuestación de cara a aligerar la carga del encuestado y facilitar el requerimiento de información. Los modos mixtos de encuestación proporcionan una oportunidad de compensar las debilidades de cada uno de los métodos. Es ampliamente reconocido que algunos respondientes prefieren ser entrevistados por un método mientras otros prefieren otro (Dillman, 2001).

Pese a que la entrevista debía ser preferentemente telefónica se ofertaban cuatro métodos alternativos de respuesta: entrevista personal, entrevista telefónica, cuestionario Web y correo postal. Este último método solo era

sugerido por los encuestadores como última posibilidad, con el objetivo de recuperar las negativas débiles. La encuestación por correo entraña dificultades de calendario y control por lo que se configura como un método poco óptimo para operaciones con duración de campo ajustada.

2.- Formación sobre técnicas de encuestación telefónica previa para los agentes que participaron en la operación mediante un curso de veinte horas de duración. El cuerpo de entrevistadores de Eustat cuenta con amplia experiencia en la técnica de encuestación cara a cara. Sin embargo con la EC04 Eustat ensayó por primera vez el cambio de encuestación personal a entrevista telefónica como método preferente de recogida de información para las encuestas a los hogares (en las encuestas a establecimientos económicos se viene utilizando desde hace tiempo).

3.- Contacto previo con los hogares seleccionados mediante envío de carta notificativa personalizada en la que se anticipaba su inclusión en la muestra, contenido de la encuesta, obligatoriedad legal de responder y garantía legal del secreto de la información proporcionada.

4.- Uso de incentivos en forma de obsequio a todos aquéllos que colaboraran en la encuesta. El regalo tenía un valor de mercado de tres euros.

5.- Monitorización del progreso del trabajo de los encuestadores, con especial control sobre las incidencias tanto de viviendas como de personas, que posibilita la identificación rápida de dificultades. Se habilitó una consulta on-line sobre la base de datos donde quedaban registradas las incidencias de la Hoja de Ruta, consultable en cualquier momento por los miembros del equipo.

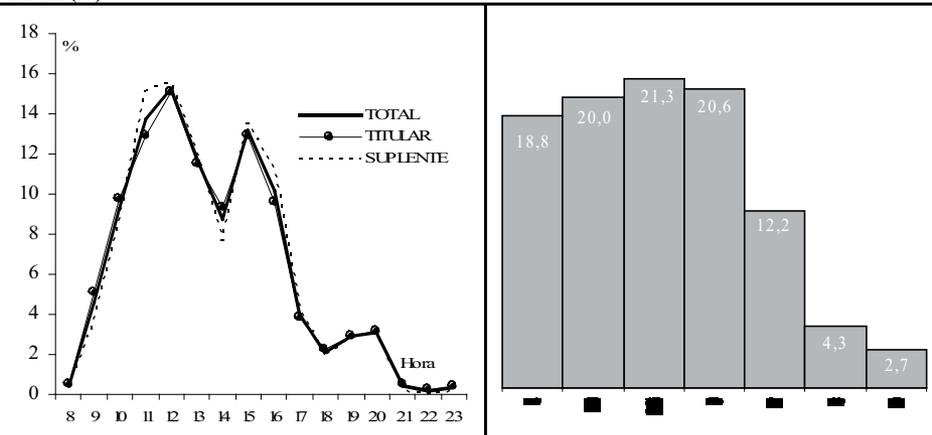
6.- Determinación y extensión del número de contactos. No había una experiencia previa documentada sobre el número y horario de contactos idóneo que rentabilizara los esfuerzos para reducir la no respuesta. Ésto era especialmente cierto en lo referente a la encuestación telefónica, primera vez que se intentaba con los hogares.

Sí que había la certeza de que los horarios de trabajo debían de adaptarse para facilitar la localización en casa de los encuestados. Siguiendo la literatura al respecto en las encuestas telefónicas un entrevistador debe intentar entre 8 y 10 rellamadas antes de ceder sobre una unidad, se estableció un máximo de ocho llamadas para el cuestionario familiar y otras tantas para el individual. Las cuatro primeras llamadas, en horario continuo hasta las 18:00, debían de hacerse en horas y días diferentes. Las cuatro siguientes llamadas eran realizadas por un grupo reducido de agentes cuyo horario de encuestación abarcaba de lunes a viernes de 15: a 21:00 y los sábados de 10:00 a 15:00, con la obligación de que al menos las dos últimas llamadas fueran realizada en viernes y sábado.

Para las encuestas mediante visita personal se marcó un máximo de tres visitas antes de la sustitución. En la práctica, tanto en la entrevista personal como en la telefónica era frecuente superar los contactos estipulados.

Como muestra el gráfico siguiente (gráfico 1), el horario de encuestación marca dos máximos, uno hacia mediodía y el otro un poco más tarde, entre las 15 y las 16 horas, momentos previos y posteriores a la comida en muchos hogares. En el horario especial, se recogió más del 10% del total de encuestas, como se ve, algunas de ellas más allá incluso de la hora límite, correspondiéndose a citas previamente establecidas.

Gráfico 1. Encuestas telefónicas realizadas por hora del día (%) **Gráfico 2. Encuestas realizadas por día (%)**



El día más productivo de la semana es el miércoles, aunque el ritmo resulta ser bastante equilibrado de lunes a jueves (gráfico 2). En cambio los viernes no parecen ser buen día para hacer encuestas. En fin de semana, incluido el domingo, se consiguió el 7% de los cuestionarios.

7.- Consulta a la base de datos de territorio para localización de direcciones postales y, edición, en su caso, de mapas y planos mediante el SIG que facilitaron la configuración de rutas adecuadas a los agentes de campo.

Más recientemente, con ocasión de la formación del nuevo panel para la Encuesta de Población en Relación con la Actividad (PRA) de Eustat, se ha puesto en marcha con cierto éxito un procedimiento recuperación de negativas. Una vez producido el rechazo a colaborar, personal especializado se pone en contacto con el encuestado siguiendo las pautas marcadas en un protocolo de actuación para recuperación de negativas. Por esta vía se llega a recuperar cerca de un 30% de las negativas iniciales. La extensión de este tipo de medidas en encuestas no panel requerirá ciertas adaptaciones, dado los diferentes ritmos de trabajo y la necesidad de completar la muestra en un período breve de tiempo, pero no cabe duda que habrá que implementarlas.

Tasas de respuesta

Son muchas ya, las oficinas que están realizando estudios sobre no respuesta, reflejados la más de las veces en informes de evaluación de la calidad de sus estadísticas (a nivel nacional por ejemplo, Informe de calidad de la Población en Relación con la Actividad de Eustat (2002) o el informe de No Respuesta en la Encuesta de Condiciones de Vida (2004). Otros institutos han puesto ya en marcha mecanismos para el tratamiento de la no respuesta. Así, por ejemplo, los Países Bajos han centrado sus esfuerzos en la reducción de las tasas de no respuesta, mientras las oficinas sueca y finlandesa se proponen realizar las mejores estimaciones posibles basadas en los datos recogidos de aquéllos que responden a la encuesta, y en la información auxiliar relevante de que se dispone sobre la población y sus elementos, sean respondientes o no.

Existe, por lo tanto, conciencia generalizada de que la falta de respuesta es un problema al que las oficinas estadísticas han de enfrentarse. Sin embargo, desafortunadamente no existe el mismo consenso sobre cómo medir la falta de respuesta. Siendo consideradas las tasas de respuesta el indicador más importante de la calidad de las encuestas a hogares, las comparaciones se ven ampliamente obstaculizadas entre encuestas, años y organizaciones por las definiciones de qué contabilizar como no respuesta y por el modo de calcular las tasas (Lynn et al., 2002).

Para el cálculo de las tasas que figuran a continuación se han seguido las recomendaciones sobre definición, categorización y cálculo de tasas efectuadas por el Institute for Social and Economic Research (ISER) (Lynn, Beerten, Laiho y Martin, 2001). Existen otras organizaciones que promueven igualmente la estandarización en el uso de definiciones y cálculo de tasas, como el Council of American Survey Research Organizations (CASRO, 1982), o, más conocida, la American Association for Public Opinion Research (AAPOR, 2006). Se ha escogido la metodología propuesta por el ISER, aunque las diferencias son mínimas, por encajar mejor en la naturaleza de los marcos y métodos de muestreo empleados en las encuestas a hogares europeas (Lynn et al., 2001).

Se han excluido de los cálculos las viviendas sustituidas en campo ya que, en sentido estricto, la sustitución es una forma de imputación y, por lo tanto, de tratamiento de la no respuesta (Vehovar, 1999).

La herramienta de partida es la Hoja de Ruta. La Hoja de Ruta, tradicionalmente en papel, es hoy en día una aplicación informática donde se recoge información acerca del propio proceso de recogida de información. Existen tres bloques diferentes de datos.

El primero detalla la situación de la unidad muestral con respecto al marco (resultado de la localización y elegibilidad), el segundo hace mención a la colaboración de la unidad muestral elegible en la encuesta (responde, no contacto, negativa), y el tercer bloque corresponde a anotaciones útiles para la gestión del trabajo (citas, observaciones etc.). El cuestionario informático de Hoja de Ruta sigue un grafo programado, de forma que se garantiza la completitud y la consistencia entre los diversos ítems. Los agentes de campo para las entrevistas personales iban provistos de una PDA con una versión idéntica de hoja de ruta a la que tenían los entrevistadores telefónicos.

Antes de pasar al análisis de los datos, habría que comenzar, quizás, concediendo que las incidencias contempladas en la hoja de ruta de ECV04 para categorizar las diversas situaciones que pueden presentarse en relación tanto a la no-respuesta como a la respuesta han resultado insuficientes. Frente a la profusión con que se ha tratado la relación entre marco y muestra en el caso de las viviendas, materializada en más de 20 situaciones diferentes factibles, tanto para las familias como para los individuos solo hay una categoría provisional, 'pendiente', y tres finales, 'encuestada', 'negativa' y 'ausencia prolongada'.

En ese mismo sentido, una de las indeterminaciones que más afectó al cálculo de las tasas de respuesta en la ECV2004 proviene de las entrevistas telefónicas. Con probabilidad la no existencia de una rúbrica específica para el establecimiento de contacto y/o la determinación de vivienda ocupada enmascara bajo un mismo resultado ('ausencia prolongada') dos categorías muy diferentes: elegible no contactado y elegibilidad desconocida.

Tabla 3. Tasas de respuesta de la ECV04 de la C.A. de Euskadi

	Viviendas	Individuos
Tasa de respuesta	70,8	68,5
Tasa de cooperación	79,6	79,7
Tasa de contacto	88,9	86,0
Tasa de rechazo	18,1	17,5

La tasa de respuesta se calcula como el número de entrevistas completas entre el número de unidades de la muestra elegibles, ésto es, las que definen la tasa de elegibilidad. La tasa de respuesta está en 71% para las familias y 69% para los individuos (tabla 3). La escasa diferencia entre ambos colectivos es debida a que una negativa o una ausencia familiar llevaba pareja una ausencia o negativa individual la mayoría de las ocasiones. En cambio, la colaboración de un informante familiar no era garantía de disponibilidad a completar la encuesta por parte del miembro familiar finalmente seleccionado, de ahí que la tasa de respuesta individual se vea mermada en más de dos puntos con respecto a la familiar.

La tasa de cooperación es la proporción de todos los casos entrevistados de las unidades elegibles contactadas alguna vez (son unidades elegibles los entrevistados, las negativas y los no contactados). Los 9 puntos de diferencia con respecto a la tasa de respuesta familiar y los más de 11 para la tasa individual se deben a que en esta tasa no se consideran las ausencias prolongadas, es decir, los no contactados en ninguna ocasión.

La tasa de rechazo es la proporción de todos los casos en los que una unidad rechaza ser entrevistado o interrumpe la entrevista una vez iniciada sobre los potencialmente elegibles. En este sentido, no es complementaria de la tasa de cooperación, ya que, a diferencia, en la tasa de rechazo se incluyen los casos de elegibilidad desconocida. Casi dos de cada diez hogares se negaron a colaborar con la encuesta.

La tasa de contacto sería la proporción de los contactados, hayan completado o no el cuestionario, sobre el total de unidades muestreadas. Lógicamente, su complementario es la tasa de no contacto, y, en este caso, representa una parte considerable de los elegibles: 11% para las familias y 14% en el caso de los individuos.

Resulta difícil valorar la calidad de los resultados representados por estas tasas. Está generalmente aceptado que la no respuesta ha crecido a lo largo del tiempo. Por ejemplo, De Leeuw y De Heer (2001) demuestran como para una larga lista de encuestas de oficinas estadísticas la no respuesta es creciente a lo largo del tiempo, tanto por el incremento de los no contactados como de las negativas a participar.

Las oficinas estadísticas no acostumbran, o solo recientemente, a difundir información sobre las incidencias habidas en la fase de recogida de información de sus diferentes operaciones. Tampoco está suficientemente estandarizada la metodología de cálculo de tasas, cuyos problemas comienzan con harta frecuencia en la propia selección y definición de las incidencias que las sustentan. Reforzando estos impedimentos, es igualmente cierto que ligeras variaciones, cuando no sustanciales, en el diseño y contexto en el que se mueven la encuesta, dificultan en extremo la comparabilidad de los resultados, tanto más cuando venga referida a un contexto internacional. Entre las variaciones más importantes cabe citar:

1. el diseño muestral: unidad de muestreo (familia o individuo), sustitución en campo, o aceptación de informante proxy;
2. el sponsor: su aceptación social, la obligatoriedad legal o no de responder;
3. la encuesta: el tema de la encuesta, la carga de respuesta de los cuestionarios o el método de encuestación;
4. el trabajo de campo: estrategia de contacto, uso de incentivos, duración del trabajo de campo, carga de trabajo por agente, monitorización del trabajo y uso de controles;
5. circunstancias de la encuesta, el clima social en el que se desarrolla la encuesta, período del año en que tiene lugar, eventos sociales particulares.

Un largo repertorio que obliga a tomar con cautela cualquier confrontación de indicadores de calidad entre dos encuestas.

La propia ECV es un buen ejemplo de la meticulosidad con la que se deben llevar a cabo las comparaciones. La anterior encuesta es de 1999 y se realizó exclusivamente con entrevistas personales. El informe final de incidencias de aquel año recoge cuatro categorías: negativas, viviendas vacías, ausencia prolongada y otras causas, no habiendo desglose entre titulares y sustitutas. Los resultados para los hogares se muestran en la tabla siguiente.

Tabla 4. ECV, tasas por tipo según año y método de encuestación (%)

	ECV1999		ECV2004	
	Entr.	Total	Entr.	Entr.
	Personal		Personal	Telefóni
Tasa de respuesta	75,3	70,8	54,5	78,4
Tasa de cooperación	86,8	79,6	83,7	78,4
Tasa de contacto	86,7	88,9	65,1	100,0
Tasa de rechazo	11,4	18,1	10,6	21,6

Las conclusiones derivables de esta tabla encajan en la tendencia general experimentada en otros países (tabla 4). La evolución de los cinco años es negativa. La tasa de respuesta desciende 4,5 puntos mientras que el rechazo incrementa su presencia relativa más de un 50%, resintiéndose, por extensión, la colaboración de los encuestados, 7 puntos menos en la tasa de colaboración.

Las importantes discrepancias entre las dos ediciones tienen su origen, en parte, en consideraciones previas. Al no existir no contacto en las encuestas telefónicas, tal y como se apuntó más arriba, y quedar cubierta toda la muestra elegible, la tasa de respuesta por el método telefónico es significativamente más alta que la conseguida a través de entrevistas personales. Si se recalculan las tasas obviando la probabilidad de contacto de uno y otro método, se cumple con la experiencia general de que, para encuestas comparables, las entrevistas cara a cara tienden a obtener más altas tasas de respuesta que las entrevistas telefónicas (De Leeuw, 2005)

El problema conceptual es solo uno de los elementos que viene a justificar los resultados de ambas encuestas.

El cambio de diseño muestral podría explicar en cierta medida la caída en la tasa de respuesta. En 1999 el muestreo partía de una selección de secciones en base a tipologías y dentro de cada sección se muestreaba un número preestablecido de viviendas. La facilidad del desplazamiento por la proximidad geográfica de las viviendas en la sección propiciaba la contactabilidad del hogar recurriendo a sucesivas visitas, incluso por encima del número aconsejado.

El muestreo aleatorio de 2004, obliga a planificar detalladamente las rutas del día, dificultando, por falta de oportunidad, el menudeo de visitas. En resumen, el costo en tiempo de una encuesta mediante visita personal ha sido superior en 2004. Es evidente si se compara la tasa de contacto y de respuesta de 1999 con las obtenidas por entrevista personal en 2004.

Igualmente la tasa de rechazo se ha visto influida por el método de encuestación en los resultados de ambos años. Si se comparan los resultados para el mismo método, ligeramente, pero en el año 2004 la tasa de rechazo ha descendido en relación a la anterior edición. Es la entrevista telefónica en la que abundan las negativas, situando la tasa general en niveles superiores a los de 1999.

El Instituto Nacional de Estadística (INE) ha publicado un informe titulado Evaluación de la Falta de Respuesta en la Encuesta de Condiciones de Vida 2004. En él se lleva a cabo un análisis pormenorizado de las incidencias en campo, centrándose especialmente en aquellas relacionadas con la falta de respuesta. Pese al nombre la temática de la Encuesta de Condiciones de Vida del INE no es la misma que la de Eustat, pero puede constituir un buen ejemplo de comparación.

De la información aportada en ese informe y siguiendo la misma metodología del ISER para el cálculo de tasas, se pueden deducir las tasas que aparecen a continuación a nivel español y que corroboran, básicamente los indicadores constatados en la encuesta de la C.A. de Euskadi (tabla 5). El mejor comportamiento de la tasa de respuesta en la ECV de Euskadi es compensado con una menor tasa de rechazo en el conjunto español. En el mismo informe se señala que la C.A. de Euskadi es la comunidad en la que habitualmente la proporción de muestra realmente conseguida es más baja. En su caso, de las 960 encuestas previstas se recogió información de 486 titulares, el 50,6%, frente al 58,3% de titulares obtenidas en ECV04 de Eustat.

Tabla 5. Tasas por tipo de la Encuesta de condiciones de Vida 2004 del INE

	%
Tasa de respuesta	69,7
Tasa de cooperación	80,0
Tasa de contacto	87,2
Tasa de rechazo	16,6

Fuente: Elaboración propia a partir de INE, Evaluación de la falta de respuesta en la Encuesta de Condiciones de Vida 2004.

Otro ejemplo. La Encuesta Social Europea, esponsorizada por la Comisión Europea, es una encuesta dirigida a exponer y explicar la interacción entre el cambio de las instituciones europeas y las actitudes, creencia y patrones de comportamiento de sus poblaciones; actualmente cubre 23 países y emplea las más rigurosa metodología. Los diferentes países participantes fueron recomendados sobre la necesidad de adoptar una serie de medidas correctoras durante el período de trabajo de campo para maximizar la tasa de respuesta.

Tabla 6. Tasas de respuesta y rechazo de la European Social Survey (2002-03).

	Tasa de respuesta	Tasa de rechazo		Tasa de respuesta	Tasa de rechazo
Finlandia	73,3	20,9	Polonia	72,2	19,6
Grecia	79,6	16,9	Portugal	68,8	26,9
Hungría	70,3	15,1	Eslovenia	71,8	15,3
Irlanda	64,4	22,9	España	53,6	35,3
Israel	70,9	21,3	Suiza	33,0	55,1
Países Bajos	67,8	26,2	Gran Bretaña	55,0	30,6

Fuente: Work package 7. Data-Based quality assessment in the ESS. Part 2. Non-response and fieldwork efforts in the ESS: results from the analysis of call record data.

Finalizada la primera ola, la mitad de los países superaban el 70% de tasa de respuesta que se había planteado como objetivo, aunque en promedio, 65,1% todavía se encontraban lejos. La ECV04 se encuentra, pues, dentro de unos parámetros razonables, particularmente en lo que se refiere a las negativas, el 18% de ECV04 está, en la gran mayoría de los casos, por debajo de lo registrado en otros países (tabla 6).

Características de la no respuesta

Un problema serio producido por las incidencias en la fase de recogida de información reside en que las estimaciones de las características de la población pueden resultar sesgadas. Esta situación ocurre cuando algunos grupos de la población están sub o sobre representados y, además, se comportan de manera diferente con respecto a las características investigadas en la encuesta. La calidad de los datos y la comparabilidad quedan comprometidos en la medida en que los entrevistados difieran de los no entrevistados. En estas situaciones hay que prever algún procedimiento de corrección del sesgo de la no-respuesta.

De entre los varios procedimientos que existen para analizar la no respuesta, uno de los más habituales consiste en utilizar la información disponible en la propia muestra proveniente del marco muestral. Se trata de comparar la tasa de respuesta entre distintas variables de la muestra evidenciando los estratos mejor y peor representados. Este método asume la correlación entre las variables del marco y las estadísticas claves de la encuesta. (Groves y Brick, 2006).

El Directorio de Viviendas, que ha servido como marco para la selección de la vivienda, ha sido posteriormente unido con el registro de población, para seleccionar la persona que debía de contestar al cuestionario individual. El marco muestral finalmente obtenido contiene información sobre el Territorio y municipio de las unidades seleccionadas, año de construcción de la vivienda, tamaño de la familia, sexo y edad de la persona de referencia y sexo y edad de la persona seleccionada para contestar el cuestionario individual.

Dado que las diferencias entre las incidencias de hogares e individuos son escasas, ya que las segundas se acumulan sobre las primeras, el análisis se efectúa preferentemente con los hogares.

Por lo que respecta a las variables de estratificación muestral se observa el distinto comportamiento de los tres Territorios (tabla 7). Las diferencias afectan sobre todo a la parte de la no respuesta correspondiente al contacto. Gipuzkoa aparece como el Territorio más propenso a colaborar, gracias sobre todo a una alta contactabilidad. De hecho, su tasa de respuesta hubiera sido bastante más amplia de no haberse registrado tanta incidencia en vivienda. A Bizkaia le sucede prácticamente lo contrario. Presenta la proporción más alta de no respuesta con una frecuencia de no contactados 10,3 puntos (48,3%) sobre el 38% que cabría esperar de acuerdo a un reparto equitativo según su peso en la muestra. Araba, por su parte, es el Territorio con mayor presencia comparativa de negativas (34%), aunque el contacto es relativamente asequible.

Algo similar sucede con el tamaño familiar. En el conjunto de la no respuesta no se observan grandes diferencias entre la distribución muestral teórica y la respuesta o la falta de respuesta. Sin embargo, en cada estrato del tamaño familiar, es el comportamiento divergente de las dos categorías de no respuesta, el contacto y el rechazo, el que equilibra la no respuesta total.

La diferencia más acusada se produce en las familias unipersonales donde parece muy difícil establecer el contacto, 10 puntos porcentuales separan la población esperada con respecto a la observada. Sin embargo, una vez producido la posibilidad de obtener un rechazo es menor, comparativamente, que en otros estratos.

Tabla 7. Viviendas titulares por Territorio Histórico, tamaño familiar y año de construcción de la vivienda según distribución muestral, respuesta, no respuesta y tipo de no respuesta (%).

Territorio	Muestra	Respuesta	No respuesta		
	100	100	Total	No Contacto	Negativas
	100	100	100	100	100
Bizkaia	38,0	37,1	41,3	48,3	36,9
Gipuzkoa	32,0	32,6	26,6	22,6	29,0
Araba	30,0	30,3	32,1	29,1	34,0
Tamaño familiar	100	100	100	100	100
1 persona	18,1	15,6	20,0	28,3	15,0
2 personas	24,0	24,3	23,8	20,9	25,6
3-4 personas	45,6	48,1	44,2	40,4	46,6
5 y más personas	12,2	12,1	11,9	10,4	12,9
Año de construcción de la vivienda	100	100	100	100	100
Hasta 1940	8,9	8,3	7,8	10,9	5,9
De 1941 a 1970	43,1	41,9	45,2	43,9	45,9
1971 y posterior	47,9	49,9	47,1	45,2	48,2

A medida que el tamaño familiar aumenta hay un descenso relativo del no contacto, mientras las negativas se mantienen ligeramente por encima de lo esperado en los estratos superiores a una persona.

Existe igualmente cierta relación entre la antigüedad de la vivienda y la probabilidad de llevar a cabo la encuesta. El volumen de negativas es considerablemente más bajo entre familias que viven en las viviendas más antiguas, más difíciles de contactar, que cuando se trata de viviendas de construcción más reciente. Pero la relación entre antigüedad de la vivienda y participación en la encuesta no es directa. Los valores para los otros dos estratos no reflejan diferencias considerables y son, en todo caso, menos positivos en las viviendas construidas entre 1941 y 1970, donde se constata la tasa de respuesta más débil y la de rechazo comparativamente más alta a pesar de que la contactabilidad no difería en exceso de la mostrada por las viviendas más recientes.

Además de las tres variables que han dado origen a la estratificación de la encuesta, el fichero muestral disponía de variables demográficas relativas al sexo y la edad de la persona designada para responder. Como la unidad muestral primera es la vivienda, caso de que la familia prevista en la muestra no coincidiera con la que realmente residía en el momento de encuestación, se procedía a un sorteo entre los miembros de la nueva familia para designar el respondiente individual. Tanto en un caso como en el otro se designaba una persona de reserva ante la eventualidad de que el inicialmente designado no pudiera responder (por ausencia prolongada o incapacidad).

Las variables socio-demográficas, per se, no determinan la no respuesta, más bien es el hecho de que personas con determinadas características sean más propensas a llevar ciertos estilos de vida o hábitos que determinan la probabilidad de encontrarlos en casa y, además, convencerle para tomar parte en la encuesta (Groves y Couper 1998). Analizándolas se puede obtener más información sobre una posible sobre representación o sub representación de ciertos grupos en la encuesta. Si la no respuesta es aleatoria, habría que esperar que la distribución por edad y sexo de quienes responden no sea significativamente diferente de aquellos que no lo hacen.

En la tabla siguiente figuran los resultados (tabla 8). Se tabulan, en este caso, las características de los individuos que han respondido y de los que no han respondido solo de aquéllos en que se constató que la familia residente coincidía con la fijada en la muestra (se recuerda que la entrevista era al hogar no a la familia concreta que residía en la vivienda).

Se han excluido de la tabulación, por lo tanto, los no contactos y negativas de familias diferentes a las esperadas por no disponer de información acerca del sexo y edad de la persona encargada de contestar el cuestionario individual. Por este procedimiento se ha filtrado el 9% de la no respuesta (115 casos de 1.255) bajo la hipótesis de similar distribución en las características de sexo y edad.

Tabla 8. Individuos titulares por sexo y sexo y edad según distribución muestral, respuesta, no respuesta y tipo de no respuesta (%).

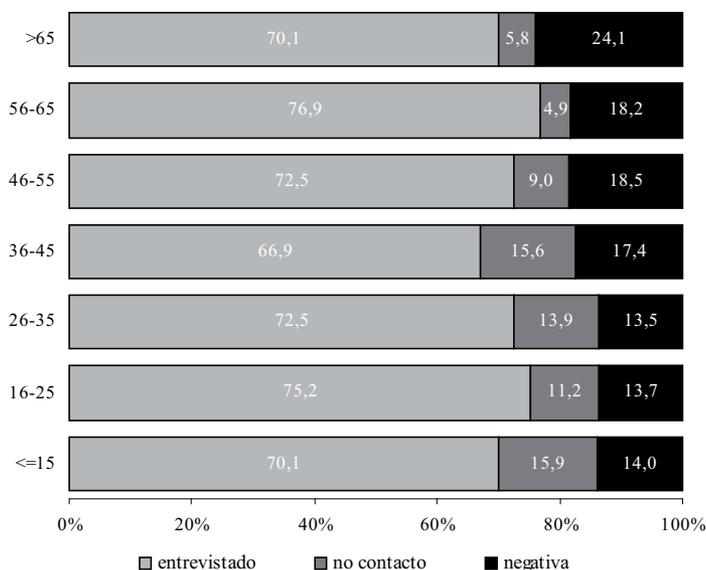
	Muestra	Respuesta		No respuesta	
		Total	No Contacto	Negativas	
Sexo	100	100	100	100	100
Mujeres	54,8	56,5	52,9	50,1	54,4
Varones	45,2	43,5	47,1	49,9	45,6
Mujeres	100	100	100	100	100
≤15	8,5	8,6	8,0	10,3	6,8
16-25	6,8	6,9	6,5	9,8	4,8
26-35	13,7	13,0	12,6	18,1	9,8
36-45	16,0	14,8	17,8	23,5	14,8
46-55	13,1	13,5	12,3	11,8	12,6
56-65	11,4	13,3	10,0	5,9	12,1
>65	30,6	29,8	32,9	20,6	39,2
Varones	100	100	100	100	100
≤15	10,1	9,6	11,8	18,7	7,5
16-25	8,8	10,4	7,6	7,9	7,5
26-35	15,7	15,7	14,6	20,2	11,1
36-45	16,7	14,1	18,1	24,1	14,7
46-55	14,5	14,9	14,7	12,8	15,9
56-65	12,5	14,0	10,6	6,4	13,2
>65	21,7	21,2	22,6	9,9	30,2

No se observan comportamientos claramente diferenciados por sexo en la respuesta ni en la no respuesta total. Tanto la distribución de la primera como de la segunda ofrece porcentajes favorables a las mujeres, pero en cantidad poco significativa. Las mujeres señalan una no respuesta 1,9 puntos inferior (52,9% frente a 54,8%) y una respuesta superior en 1,7 puntos (56,5% frente a 54,8%) a la que cabría esperar si no existieran diferencias entre ambos sexos. La discrepancia entre ambas cifras se puede atribuir al desigual reflejo de las incidencias de viviendas (no localizadas, desaparecidas...) de titulares de uno y otro sexo.

Las distribuciones por sexo y edad de la muestra teórica y de los respondientes son prácticamente simétricas. La no respuesta se concentra en el grupo de edad de 36 a 45 y más de 65 años en el caso de que la encuestada sea mujer, y en los mismos grupos de edad más el de los menores de 15 años, cuando el encuestado es un hombre.

Pero la composición de la no respuesta opera de forma diferente dependiendo de qué grupo de edad se trate. En el de 36 a 45 años, la no respuesta es básicamente debida al no contacto, mientras en el grupo más mayor, el de 66 y más años, la no respuesta es fruto del rechazo a colaborar.

Extendiendo esta situación al resto del grupo de edades se puede afirmar que hasta los 45 años la no respuesta tiene un carácter básico de no contacto, mientras a partir de esa edad comienza a funcionar la negativa. Excepto para los dos grupos de edad señalados, en que el predominio de uno de las dos componentes es claro, para el resto el resultado tiende al equilibrio.

Gráfico 3. Individuos titulares elegibles por grupo de edad según respuesta y tipo de no respuesta

Respecto al grupo de varones menores de 15 años, el abultado no contacto obedecerá a variaciones aleatorias propias de tamaños muestrales reducidos, aún siendo una característica propia del grupo de edad.

Otra perspectiva, la del reparto de la respuesta y no respuesta dentro de cada grupo de edad, haga, quizás, más evidente la relación entre los diferentes componentes de la respuesta dentro de cada grupo de edad. Como se puede apreciar en el gráfico 3 la tasa de respuesta, que es en el fondo lo que aparece representado, es variable dentro de cada grupo de edad. La proporción de negativas, la tasa de rechazo, crece considerablemente a partir de los 36 años, a la vez que se constata un considerable descenso de la tasa de no contacto. En el grupo de 36 a 45 años la unión del no contacto y negativas provoca que la tasa de no respuesta total se haga máximo, 33,1%. Próximos al máximo se encuentran los grupos de edad extremos, menores de 16 y mayores de 65 años, ambos con un 29,9% de no respuesta.

Para el análisis de la no respuesta se ha utilizado una última variable deducible de la información contenida en el fichero muestral, el tamaño del municipio de residencia.

Tabla 9. Familias titulares por tamaño de municipio de residencia, respuesta, no respuesta y tipo de no respuesta (%).

	Muestra	Respuesta		No respuesta	
		Total	No Contacto	Negativas	
		100	100	100	100
10.000 habitantes	19,3	19,2	14,6	10,4	17,1
10.001-40.000 habitantes	23,9	24,9	22,3	20,9	23,2
40.001-100.000 habitantes	13,6	13,6	14,7	17,8	12,9
>100.000 habitantes	43,2	42,3	48,4	50,9	46,9

La no respuesta se acumula en los municipios de mayor tamaño, los mayores de 40.000 habitantes y, particularmente, en los de 100.000 o más (tabla 9). En ellos se registra la mayor diferencia entre tamaño muestral y porcentaje de no respuesta, 5,2 puntos porcentuales. Además, la acumulación es debida a un exceso de no contacto en ambos estratos, que se agrava para los municipios de más de 100.000 habitantes por una mayor presencia de negativas.

El comportamiento de los municipios más pequeños es el contrario. Constatan una menor frecuencia de no respuesta que la esperable, pero mantienen el no contacto como elemento clave de la no respuesta, en este caso por la escasez de no contactos en comparación con los otros dos estratos. En definitiva, una vez más, la pauta la está marcando los contactos.

Para finalizar este apartado la tabla siguiente resume la relación entre las variables analizadas y la no respuesta (tabla 10). Según la prueba χ^2 y pese a las diferencias de tasas, solo existe una relación significativa entre mujeres y edad

y varones y edad, en resumen entre la edad, con el tipo no contacto. Al límite de la significación también con el no contacto el Territorio, el municipio y el año de construcción de la vivienda.

En cambio no existe asociación entre las negativas y ninguna de las variables elegidas, lo que provoca que tampoco exista asociación con la no respuesta total.

Tabla 10. Prueba χ^2 para las distintas variables con respecto a la muestra teórica, según tipo de no respuesta

	No respuesta total	No contacto	Negativas
Territorio	0,51	0,06	0,65
Tamaño familiar	0,97	0,07	0,87
Año de construcción de la vivienda	0,88	0,73	0,55
Sexo	0,70	0,35	0,93
Mujeres y edad	0,99	0,05	0,60
Varones y edad	0,99	0,00	0,42
Tamaño del municipio	0,58	0,07	0,89

Conclusiones

La explotación de la información de la Hoja de Ruta de la Encuesta de Condiciones de Vida 2004 ha permitido poner de manifiesto algunas cuestiones de interés sobre el marco muestral y la no respuesta.

En primer lugar, ha demostrado la importancia de disponer de marcos actualizados y completos así como de intensificar esfuerzos en la fase de campo para dar solución a todas las incidencias que inevitablemente se presentan.

Respecto a los errores de cobertura cabe señalar:

1.- En el caso de que se utilicen métodos mixtos de encuestación es deseable que el apunte final sobre el estado de una unidad se haga en campo. A este respecto la encuestación telefónica genera incertidumbre sobre la elegibilidad o la contactabilidad de la unidad muestral, influyendo en los resultados del cálculo de las tasas de respuesta.

2.- Pese a lo anterior, la frecuencia de todos los tipos de incidencias en el marco es aparentemente independiente del tipo de método de encuestación que se emplee.

3.- La repercusión de las incidencias varía notablemente entre las distintas categorías de cada una de las variables de estratificación. Viviendas antiguas, familias unipersonales y municipios pequeños son más proclives a presentar incidencias de marco.

Frente a la no respuesta total es fundamental tener definidas estrategias para minimizar el impacto durante la fase de campo. En este proceso se considera imprescindible hacer acopio de información sobre la recogida de información, analizarla y planificar conforme a los resultados los mejores calendarios y horarios para potenciar la contactabilidad, así como establecer las características de la no respuesta para concentrar esfuerzos en aquellos subgrupos que lo precisen.

Una monitorización continua sobre las tasas de respuesta permitirá tomar las acciones correctoras oportunas en el debido tiempo.

En cuanto a las características de la no respuesta de la ECV04 la conclusión fundamental es que el elemento clave es la falta de contacto más que la negativa a colaborar.

Las diferencias a nivel territorial, Gipuzkoa la más perjudicada, de tamaño familiar, particularmente las familias unipersonales, y, en menor medida, el año de construcción de la vivienda, las más antiguas, evidencian problemas de falta de contacto.

A nivel individual, las variables citadas más el sexo y la edad conjuntamente, así como el tamaño del municipio de residencia están significativamente relacionados con el no contacto. Hasta los 45 años la no respuesta tiene un carácter básico de no contacto, mientras a partir de esa edad comienza a funcionar la negativa. En los municipios grandes resulta más difícil conseguir un contacto.

Referencias

- Biemer Paul B. y Liberg Lars E. (2003). *Introduction to Survey Quality: Wiley Series in Survey Methodology*. New York: John Wiley and Sons.
- De Leeuw, Edith D., De Heer, W. (2001). Trends in household survey nonresponse: A longitudinal and international comparison. Págs: 41-54. En: RM. Groves, D.A. Dillman, J.L. Eltinge, y R.J.A. Little. (eds). *Survey nonresponse*. New York: Wiley.
- De Leeuw, E.D. (2005). *Nuevas metodologías en la recogida de datos, diseño de cuestionarios y calidad de las encuestas*. Vitoria-Gasteiz: EUSTAT.
- Díaz de Rada V. (2000). *Problemas originados por la no respuesta en investigación social: definición, control y tratamiento*. Pamplona: Universidad Pública de Navarra.
- Dillman, D.A., Glenn, P., Tortora, R. Swift, K., Kohrell, J. y Berck, J.. Response rate and measurement differences in Mixed Mode Surveys. Using Mail, Telephone, Interactive Voice Response and Internet. URL: <http://www.sesrc.wsu.edu/dillman/papers/> (15/05/2006).
- Evaluación de la falta de respuesta en la Encuesta de Condiciones de Vida 2004. Madrid: INE. URL: <http://www.ine.es/daco/daco42/condivi> (15/05/2006).
- Groves, R.M. y Couper M.P. (1998). *Nonresponse in household interview surveys*. New York: John Wiley and Sons.
- Groves, R.M. (1989). *Survey errors and survey costs*. New York: Jon Wiley&Sons.
- Groves, R.M. y Brick, J.M.. Practical Tools for nonresponse bias studies. Cardiff: European Conference on Quality in Survey Statistics 2006. (fotocopia).
- Sánchez Carrión, J.J. (2000). *La bondad de la encuesta: el caso de la no respuesta*. Madrid: Alianza Editorial.
- Lynn, P, Beerten, R., Laiho, J. y Martin J. (2002). Towards standarisation of usrvey outcome categories and response rate calculation. *Research in Official Statistics*. 1/2002 61-84.
- Lynn, P., Beerten, R., Laiho, J. y Martin J. (2001). 'Recommended Standar Final Outcome Categories and Standar Definitions of Response Rate for Social Survey'. Working Papers of the Institute for Social and Economic Research, paper 2001-23. Colchester: Univeristy of Essex. URL: <http://www.iser.essex.ac.uk/pubs/workpaps/> (30/04/2006).
- On the definition on response rates (1982). Council of American Survey of response rates. URL: <http://www.casro.org/resprates.cfm> (15/05/2006)
- Standars definitions. Final Disposition of case codes and outcome rates for surveys (2006). The American Association for Public Opinion Research. URL: http://www.aapor.org/default.asp?page=survey_methods/standards_and_best_practices (15/05/2006)
- Vehovar, V. (1999). Field substitution and unit nonresponse. *Journal of Official Statistics*. 15 (2) 335-350.

