

¿Quién contesta qué y cuándo?

**Los efectos del cuestionario sobre la no respuesta en las preguntas de los
barómetros del CIS**

Lucía Medina (ICPS-UAB)

Robert Liñeira (Universidad de Edimburgo)

Jordi Muñoz (UB)

Guillem Rico (UAB)

TRABAJO EN CURSO

NO CITAR SI EN EL PERMISO DE LOS AUTORES

Investigación acogida a la convocatoria de ayudas para equipos de investigación para la explotación del banco de datos del CIS en su edición de 2013

1. Introducción

El estudio de los efectos de diseño es un clásico en la investigación de la metodología de encuestas. En realidad, los primeros estudios de opinión ya pusieron de manifiesto que la manera de presentar las preguntas a los entrevistados, las opciones de respuestas que se les ofrecían, e incluso, la propia posición de la pregunta en el cuestionario influían en la calidad y la cantidad de las respuestas (Rugg, 1941). La acumulación de investigaciones ha servido para agregar y refinar la evidencia en este sentido, y también para subrayar que los recursos y habilidades cognitivas de los individuos interfieren en su capacidad para responder a los cuestionarios (Schuman y Presser, 1979 y 1996; Schuman y Scott, 1987; Sudman y Bradburn, 1974; Tourangeau *et al.*, 2000).

Sin embargo, en este trabajo no nos vamos a interesar tanto por las características de los que no responden como por aquellas características de las preguntas que facilitan o dificultan la respuesta. Y, más en concreto, no sólo nos interesaremos por cuáles son las peculiaridades en el diseño de las preguntas y de los cuestionarios que influyen sobre la frecuencia de la no respuesta, sino también por la manera en que estos factores interactúan con los atributos de los individuos. En definitiva, nuestro interés último radica en saber qué características concretas de las preguntas y de los cuestionarios hacen que determinados individuos tengan una probabilidad distinta de responder.

Así pues, el objetivo de este estudio es doble. Por un lado, analizaremos el efecto que el diseño de las preguntas y de los cuestionarios tiene sobre la no respuesta a las preguntas acerca de actitudes y opiniones políticas de los barómetros del CIS. Por otro lado, estudiaremos si el diseño de las preguntas y los cuestionarios modula el efecto de los atributos sociodemográficos de los individuos sobre la no respuesta.

Para hacerlo, hemos recopilado información sobre las características de las preguntas, de los cuestionarios y de los individuos. Así, en una primera parte de esta investigación nos centramos en el efecto que tienen las características de las preguntas y de los cuestionarios sobre la cantidad de no respuesta. En la segunda parte, añadimos las características de los individuos para estimar en qué medida interactúan con las características de las preguntas.

2. ¿Qué sabemos sobre la no respuesta?

2.1. Los antecedentes

Los estudios sobre metodología de encuestas se han interesado por los problemas de no respuesta desde hace más de 60 años (Payne, 1949-50), sobre todo por sus sesgos potenciales en la estimación de parámetros (DeMaio, 1980; Rapoport, 1982). Estos estudios han distinguido entre la no respuesta como rechazo a la participación en una encuesta –*unit nonresponse*– y la no respuesta a preguntas concretas de un cuestionario –*item nonresponse*– (que constituye el objeto de estudio de esta investigación).

La mayoría de estudios sobre la no respuesta a preguntas en cuestionarios se han centrado en el efecto de las características del entrevistado (educación y estatus socioeconómico, principalmente) sobre la probabilidad de respuesta (Ferber, 1966; Francis y Busch, 1975; Faulkenberry y Mason, 1978; De Maio, 1980; Aneshensel *et al.*, 1982; Rapoport, 1982; Smith, 1982; Singer *et al.*, 1983; Bell, 1984; Bishop *et al.*, 1986; Billiet y Loosveldt, 1988; Feick, 1989; Gilljam y Granberg, 1983; Schuman y Presser 1996). Otros estudios se han centrado en los efectos del entrevistador (Singer y Kohnke-Aguirre, 1979), en los efectos de la redacción de las preguntas (Payne, 1949-50; Coombs y Boombs, 1976-77; Tourangeau *et al.*, 2000), en los efectos de los filtros (Messmer y Seymour, 1982), o en los efectos de la posición de la pregunta en el cuestionario (Dickinson y Kirzner, 1985; Sudman *et al.*, 1996).

Una práctica común en los estudios sobre la no respuesta ha consistido en tratar las categorías “no sabe” y “no contesta” como categorías equivalentes. Si no hubiera diferencias, estas dos respuestas indicarían simplemente diferentes grados en la voluntad de responder del entrevistado, lo que convertiría al “no sabe” en una especie de “no

contesta” agudo (Shoemaker *et al.*, 2001, 194). Sin embargo, diferentes estudios indican que existen diferencias entre estas dos categorías, y que las preguntas delicadas o sensibles, que interpelen sobre comportamientos socialmente no deseables o sobre temas personales, reciben mayor número de “no sabe”. En estas situaciones, los entrevistados pueden optar por evitar sus respuestas reales, para evitar sentirse violentados, y optar simplemente por no responder a este tipo de preguntas, lo que se registraría como “no contesta” (Sudman *et al.*, 1996; Tourangeau *et al.*, 2000; Shoemaker *et al.*, 2001).

En cambio, las preguntas que requieren un mayor esfuerzo cognitivo reciben un mayor número de “no sabe”, aunque el esfuerzo cognitivo está también correlacionado con la categoría “no contesta”. A menudo, las encuestas plantean preguntas que implican un grado elevado de esfuerzo cognitivo y que pueden ser difíciles de contestar. Una respuesta óptima por parte de los entrevistados requiere que éstos pasen por un proceso cognitivo constituido de cuatro fases (Tourangeau, 1984): 1) los entrevistados deben interpretar cuidadosamente el significado de cada pregunta; 2) deben buscar de manera exhaustiva en sus memorias para recabar toda la información relevante; 3) deben integrar de manera cuidadosa esta información en juicios que sintetizen sus opiniones, y 4) deben expresarlas de manera que transmitan de la manera más clara y precisa posible su significado. Dada esta dificultad, en muchas ocasiones, el entrevistado quizás no pueda afrontar los costos de generar una contestación válida y prefiera optar por no responder a lo que se le pregunta diciendo que no sabe. Así, el “no sabe” que se consigna puede ser entendido como una falta de capacidad para ofrecer una respuesta.

2.2. ¿Qué aporta esta investigación?

A pesar del número importante de investigaciones sobre la cuestión este estudio realiza una serie de aportaciones al estudio de la no respuesta.

Por un lado, realizamos una estimación sistemática de los efectos de los parámetros de las preguntas y de los cuestionarios. Para hacer una estimación robusta de estos efectos se necesita un elevado número de preguntas con características diferentes. Nuestra selección de preguntas a partir de los barómetros del CIS nos permite obtener un alto número de preguntas diferentes, y una alta variación en las principales características de preguntas y cuestionarios que pueden tener efectos sobre la cantidad de no respuesta.

En segundo lugar, el hecho de anidar las características de los individuos y de las preguntas nos permite avanzar en las estimaciones al uso. En efecto, es de esperar que no todas las características de las preguntas y de los cuestionarios tengan el mismo efecto en la probabilidad de no respuesta de los entrevistados. Anteriormente, ya hemos comentado que las categorías “no sabe” y “no contesta” podrían estar acogiendo a distintos tipos de entrevistados. Así, no sería sorprendente que las diferentes características de las preguntas (ofrecer categoría intermedia o no, longitud de la pregunta, etc.) afecte de manera diferenciada en función de las características sociodemográficas de los entrevistados, de su grado de implicación política, etc.

Sin embargo, estos efectos interactivos sólo se pueden estimar creando una base de datos jerárquica de pertenencia múltiple, que nos permita medir la incidencia que tienen las diferentes características de preguntas y cuestionarios sobre la probabilidad de respuesta de los individuos en función de sus características individuales.

3. La propuesta de un enfoque metodológico distinto

En el análisis de la no respuesta encontramos factores explicativos a diferentes niveles. Un primer gran grupo de factores se encuentra en el nivel de las preguntas. Se trata de aspectos del diseño y contenido de las preguntas, tales como la longitud del enunciado, el número y formato de las opciones de respuesta, la posición en el cuestionario o la temática abordada. El impacto de este tipo de variables se ha estudiado tradicionalmente mediante métodos experimentales, aunque a menudo insertados en encuestas de opinión sobre muestras representativas.

El segundo gran grupo de factores explicativos se sitúa al nivel de los individuos. Se trata en este caso de las características de los entrevistados que pueden afectar su capacidad y motivación para dar respuesta a las preguntas que se les formulan, fundamentalmente aspectos asociados a sus recursos cognitivos e identidades sociales, tales como el nivel de estudios, la clase social, la situación laboral o la autodefinición ideológica. El impacto de estas variables se ha estudiado fundamentalmente a través del análisis de encuestas de opinión, pero también mediante experimentos (Schuman y Presser, 1979 y 1996).

Este proyecto propone examinar la cuestión a partir de un enfoque integral y una metodología novedosa. En nuestro enfoque el efecto sobre la no respuesta de las características de las preguntas se examina teniendo en cuenta el efecto simultáneo de las características personales y en interacción con éstas. En última instancia lo que nos interesa es descubrir qué efectos tienen las características de las preguntas según las características de los individuos, cómo el impacto de los atributos sociodemográficos y actitudinales individuales se ve modulado por las peculiaridades en el formato y el contenido de las preguntas. Con este fin hemos querido aprovechar el abundante y

variado catálogo de encuestas del CIS para plantear una metodología basada en el análisis de un gran número de estudios de opinión. De esta manera tenemos acceso a preguntas de formatos y contenidos muy diversos, formuladas a muestras de individuos representativas de la población.

Así como los principales factores explicativos de la no respuesta se encuentran a dos niveles (preguntas e individuos), su análisis puede hacerse también a esos dos niveles. Pero si queremos sacar todo el provecho de la información disponible y, sobre todo, si nuestro interés ha de centrarse en la interacción entre características de las preguntas y características de los individuos, debemos combinar ambos. La originalidad de nuestra propuesta radica precisamente en la construcción de una base de datos que combina la información sobre el diseño de las preguntas con la información sobre los entrevistados, y que permite, en un último estadio, analizar la propensión a la no respuesta según el efecto combinado del diseño del cuestionario y las características individuales.

El análisis empírico procede en dos etapas. En un primer estadio calibramos los efectos del *wording* y del diseño del cuestionario sobre la no respuesta a nivel agregado. En este punto la unidad de análisis es la pregunta individual, y la variable dependiente el porcentaje de no respuesta obtenido sobre el total de entrevistados en esa pregunta.

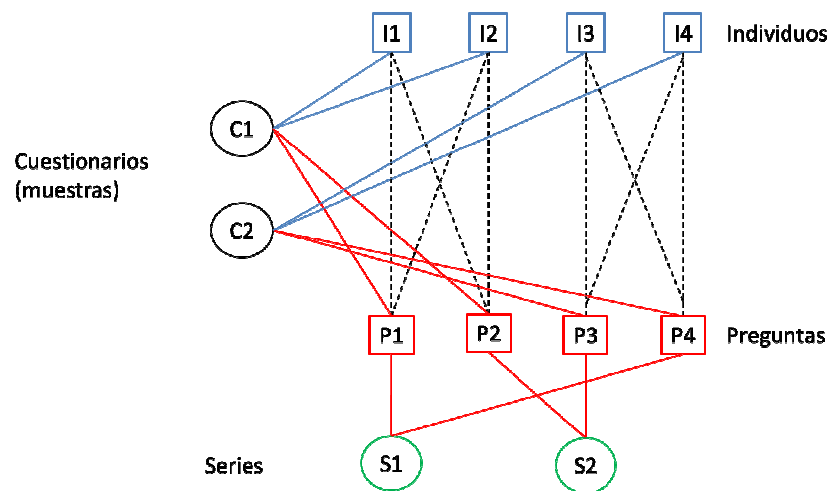
El objetivo consiste en conocer qué tipo de preguntas generan un nivel mayor o menor de “no sabe” y de “no contesta”.

En una segunda etapa analizamos el efecto conjunto de las características de los individuos y el diseño del cuestionario. Para ello construimos una base de datos que contiene información de los individuos de todas las encuestas seleccionadas junto con las variables de diseño de pregunta y cuestionario. En la matriz generada cada individuo aparece tantas veces como preguntas de las seleccionadas ha tenido que responder. O lo que es lo mismo, cada una de las preguntas seleccionadas aparece tantas veces como individuos había en la muestra a la que se administró. Estrictamente, las unidades de

análisis no son las preguntas ni los individuos sino el cruce de ambos; son lo que llamamos “individuos-pregunta”. En este punto nuestra variable dependiente es la no-respuesta del individuo a una pregunta concreta (1=no responde; 0=responde). Y las variables dependientes son las características individuales, en un nivel, y las características de la pregunta, en otro.

Se trata por tanto de una base de datos compleja, con varios niveles formando una estructura de clasificación cruzada en la que tanto preguntas como individuos aparecen anidados en cuestionarios (o muestras). Las preguntas están agrupadas también en series, ya que una misma pregunta puede aparecer en más de un cuestionario. Así pues es posible identificar hasta cinco niveles distintos: individuos-pregunta, individuos, preguntas, cuestionarios y series. El gráfico siguiente ilustra la enredada estructura de los datos analizados.

Cuadro 3.1. Estructura de los datos en el análisis de la no respuesta



En ambas etapas del análisis (agregado y multinivel) la estimación de los errores típicos de los coeficientes toma en cuenta que las observaciones no son independientes entre sí sino que están agrupadas en series. Las preguntas que no se repiten forman su propia serie.

3. 1. ¿A partir de qué trabajar? La construcción de una base de datos novedosa

Para la construcción de la base de datos seleccionamos las preguntas sobre actitudes y opiniones de carácter político incluidas en todos y cada uno de los barómetros mensuales del CIS de los últimos 10 años, hasta reunir un total de 110 barómetros realizados entre enero de 2003 y diciembre de 2012. Este marco temporal nos permite contar con un número suficiente de preguntas diferentes, que facilita la estimación de los efectos de los distintos factores considerados.

La razón fundamental para centrarnos en cuestiones referidas únicamente a actitudes y opiniones de carácter político es doble. Por un lado, las preguntas sobre elementos subjetivos suelen mostrar una amplia variación en la frecuencia de no respuesta lo que implica que, a diferencia de las preguntas sobre hechos objetivos, las preguntas sobre opiniones y actitudes se hallen ampliamente sujetas a los efectos de *wording* y de cuestionario (Converse y Presser 1983, Sudman y Bradburn 1982, Fowler 2008). Por otro lado, las preguntas de carácter político suelen verse claramente afectadas, no sólo por las características de la pregunta y del cuestionario, sino también por las características del contexto político y social en el momento de la recogida de datos (Berinsky 2004; Krosnick y Milburn 1990; Zaller 1992).

Se han excluido de nuestro análisis las preguntas que hacen referencia a comportamientos de los entrevistados. No están, por tanto, las preguntas sobre recuerdo o intención de voto en cualquier tipo de elecciones (que por otro lado ya han sido analizadas en profundidad de forma específica). Tampoco se incluyen en nuestra base de datos las preguntas de conocimiento, esto es, cuestiones en las que el conocimiento constituye el propio objeto de la pregunta. En estas el “no sabe” adquiere sentido de respuesta y por tanto no cabe considerarlo como no respuesta. Por esta razón quedan fuera las preguntas

sobre las valoraciones de los líderes políticos cuya formulación contempla su desconocimiento.

Siguiendo este criterio se han codificado un total de 1.746 preguntas, que constituyen las observaciones que empleamos en la primera etapa del análisis. La mayoría de las preguntas se ha formulado en más de uno de los cuestionarios analizados y por tanto aparece más de una ocasión en nuestra base de datos (referida a diferentes encuestas). En concreto, el 30 por ciento de las preguntas aparece solo una vez, y el 70 por ciento restante se repite entre dos y hasta en 110 ocasiones (aquellas que se repiten en todos los barómetros). Hay un promedio de 16 preguntas por barómetro, con un mínimo de 3 y un máximo de 57.

Para cada una de estas preguntas-cuestionario se ha registrado el porcentaje total de no respuesta y distinguiendo a su vez, de acuerdo con lo expuesto anteriormente, aquellos que no saben qué responder de los que rehúsan hacerlo (Sudman et *al.*, 1996; Tourangeau et *al.*, 2000; Shoemaker et *al.*, 2001). En el análisis tanto a nivel agregado como a nivel individual se examinan los “no sabe” y los “no contesta” por separado, así como la suma de ambos.

En cuanto al diseño del cuestionario, se ha codificado un gran número de variables referidas a aspectos formales y sustantivos o de contenido. A continuación se enumeran tan sólo las que han sido utilizadas en los análisis empíricos que se presentan en este primer trabajo, para las cuales se presentan también los descriptivos.

- Posición de la pregunta en el cuestionario: orden en el que aparece la pregunta en el cuestionario.
- Longitud de la pregunta: medida en número de palabras. El análisis de esta característica revela que sigue una relación no lineal con los niveles agregados de no-respuesta, por lo que se ha tomado su logaritmo.

- Pregunta básica vs. de “seguimiento” o “clarificación”: variable dicotómica que distingue las preguntas que inician un tema, o básicas (valor 1), de aquellas en las que, siguiendo a una pregunta básica, se continúa preguntando por diferentes aspectos de la misma cuestión con el fin de recabar más información (valor 0).
- La pregunta forma parte (1) o no (0) de una batería o filtro.
- Tipo de opciones de respuesta: distingue si las respuestas se dan en una escala y qué tipo de escala se emplea. Toma tres categorías: escala bipolar, escala unipolar y otros tipos (no es una escala).
- Grado de autonomía semántica: variable dicotómica que distingue las preguntas cuyas opciones de respuesta cuentan con autonomía semántica, ya sea total o parcial (1), de las que no (0). Una opción de respuesta tiene autonomía semántica cuando para su comprensión no es necesario conocer las demás opciones de respuesta. Las escalas carecen d autonomía semántica.
- Se ofrece una categoría central (1) o no (0) en las opciones de respuesta.
- Aliento a la respuesta: variable dicotómica que indica si el enunciado de la pregunta contiene (1) o no (0) un aliento a la respuesta.
- Se leen las opciones de respuesta: variable dicotómica que indica si el entrevistador lee en las opciones de respuesta en el enunciado (1) o no (0).
- Tipo de objeto por el que se pregunta: se han agrupado en tres categorías: las que hacen referencia al proceso político, o *politics* (preguntas sobre gobiernos, instituciones, partidos, elecciones, grupos de interés, movimientos sociales, etc.); a las políticas públicas, o *policy* (preguntas sobre economía, políticas sociales, familia, mujer, inmigración, seguridad, medio ambiente, cultura, asuntos exteriores, etc.); u otro temas no asimilables a las dos anteriores.
- Naturaleza del objeto: variable dicotómica que indica si la pregunta se refiere a un objeto concreto (1) o abstracto (0).
- Referencia temporal: en el enunciado de la pregunta se hace una referencia temporal explícita (1) o no se hace (0).

Tabla 3.1. Descriptivos de variables sobre las preguntas y los cuestionarios

Variable	Media	Desv. típ.	Mínimo	Máximo	N
Posición	31,234	24,006	2	117	1.638
Longitud (log)	3,247	0,453	0,693	4,890	1.635
Básica vs. seguimiento	0,706	0,456	0	1	1.746
Batería	0,518	0,500	0	1	1.746
Opciones de respuesta (categoría ref.: bipolar)					
Unipolar	0,096	0,294	0	1	1.746
No procede	0,255	0,436	0	1	1.746
Autonomía semántica	0,868	0,338	0	1	1.746
Categoría central	0,469	0,499	0	1	1.746
Tarjeta	0,198	0,399	0	1	1.746
Aliento a la respuesta	0,400	0,490	0	1	1.746
Se leen opciones respuesta	0,763	0,425	0	1	1.746
Tipo de objeto (categoría de ref.: proceso político)					
Políticas públicas	0,388	0,488	0	1	1.746
Otros	0,339	0,474	0	1	1.746
Objeto concreto vs. abstracto	0,398	0,490	0	1	1.641
Referencia temporal	0,170	0,375	0	1	1.746
Requiere conocimiento específico	0,398	0,490	0	1	1.641
Ámbito internacional vs. resto	0,081	0,273	0	1	1.746

- Requiere o no conocimiento específico: distingue aquellas preguntas que requieren para su respuesta de algún conocimiento específico (1) de las que no lo necesitan (0).
- Ámbito: la pregunta se refiere al ámbito internacional (1) o a otro ámbito (nacional, autonómico, local o a ninguno en particular) (0).

Con respecto a los individuos, nos centramos en aquellas características que la literatura ha destacado como factores que afectan a la probabilidad de respuesta (Schuman y Presser, 1996) y que están presentes en todos los barómetros de opinión del CIS. Estas variables pueden hacer referencia a los recursos cognitivos del individuo, su socialización, y su implicación y posicionamiento político.

- Edad: recodificada a un rango 0-1.
- Sexo
- Tamaño de municipio: recodificada a un rango 0-1.
- Nivel de estudios: sin estudios, estudios primarios, educación secundaria, formación profesional, estudios superiores.
- Situación laboral: trabaja, jubilado o pensionista, parado, estudiante, labores del hogar
- Clase social: trabajadores no manuales, empresarios y autónomos, agricultores, trabajadores manuales cualificados, trabajadores manuales no cualificados, profesionales y directivos.
- Autoubicación ideológica en el eje izquierda-derecha: escala de 0 a 10 recodificada en cuatro categorías: izquierda (posiciones 1-4), centro (5), derecha (6-10), no se ubica en el eje.
- Recuerdo de voto en las últimas elecciones generales: PSOE, PP, IU, otros partidos y coaliciones, abstención, no recuerda o no contesta.

A pesar de la complejidad de la tarea, la base de datos multinivel se ha construido a partir de la fusión de las variables relativas a las 1.746 preguntas agregadas y las respuestas y características individuales de todos los entrevistados a los que fueron administradas. La base de datos resultante contiene un total de 4.067.092 casos (individuos-pregunta).

4. Lo que nos dice el efecto del diseño del cuestionario sobre la no respuesta

¿Cómo afectan el diseño de las preguntas y del cuestionario al hecho de que las personas a las que se entrevista en un encuesta contesten o no a una pregunta concreta de tipo político? La respuesta a esta cuestión constituye el objetivo principal de este apartado y buscamos responderla a partir del análisis de la base de datos que hemos elaborado y que recoge información sobre las preguntas de los barómetros del CIS correspondientes al periodo 2003-2012.

Sin embargo, antes de comenzar se hace necesario conocer el alcance de la no respuesta, distinguiendo a su vez entre las personas que dicen no saber qué responder, de las que no quieren hacerlo y siguiendo, por tanto, las consideraciones de las aportaciones que diferencian entre la no respuesta debida al carácter delicado o sensible de lo que se pregunta, y la no respuesta debida a la falta de opiniones o conocimientos sobre lo que se interroga (Sudman *et al.*, 1996; Tourangeau *et al.*, 2000; Shoemaker *et al.*, 2001).

Como ya se ha señalado con anterioridad, la unidad de análisis de la base de datos y de este análisis inicial, la forman las preguntas acerca de las actitudes y opiniones políticas de las personas entrevistadas, descartándose en la construcción de la base aquellas otras preguntas sobre comportamientos o conocimientos acerca de la política, puesto que en éstas la no respuesta presenta un sentido distinto.

A partir de este planteamiento metodológico, la base arroja un total de 1.746 preguntas u observaciones de las cuales 1.635 contienen valores que indican el nivel de no respuesta por parte de los entrevistados. Es decir, para estas 1.635 preguntas existía la posibilidad,

aunque no ofrecida abiertamente, de que las personas entrevistadas no contestasen, bien porque que no supiesen qué decir o bien porque no quisiesen.

La tabla y los gráficos que siguen a continuación muestran precisamente el alcance de la no respuesta en su versión total y desglosada para estas preguntas.

El valor medio de la no respuesta se sitúa en el 12,4 por ciento, y su valor máximo llega incluso al 71,5, siendo éste el caso el de la pregunta en la que se pide a la persona entrevistada que ubique a Coalición Canaria en una escala de 1 a 10, donde el 1 representa la posición más extrema de izquierda, y el 10, la más extrema de derecha. Por su parte, el valor medio del no sabe gira en torno al 10,5 por ciento, con un valor máximo del 63,9. Se trata de nuevo de la pregunta sobre la posición ideológica de Coalición Canaria, aunque de cerca, y con un 61,1 por ciento de no respuesta debida también al no sabe, se sitúa la pregunta respecto del posicionamiento ideológico de UPyD. Por último, el valor medio del no contesta es mucho más modesto, un 1,9 por ciento, y su valor máximo tan sólo llega al 16,2 por ciento, correspondiendo a la pregunta sobre la ubicación del propio entrevistado en la escala izquierda-derecha.

Tabla 4.1. Estadísticos básicos de la no respuesta

	Media	Desviación típica	Coefficiente de variación	Valor mínimo	Valor máximo	N
% total no respuesta	12,4	9,3	0,8	0	71,5	1.635
% no sabe	10,5	8,2	0,8	0	63,9	1.626
% no contesta	1,9	2,8	1,5	0	16,2	1.633

Nota: el número de casos no coincide porque no todas las preguntas ofrecen a un tiempo la opción de responder que no se contesta o no se sabe.

Asimismo, la no respuesta total y la que concierne al no sabe presentan una desviación típica similar. En ambos casos los porcentajes concretos de no respuesta se separan del valor promedio en 9,3 y 8,2 puntos porcentuales, respectivamente, mientras que la no respuesta perteneciente al no contesta, lo hace en un 2,8. Sin embargo, la diferencia entre

estas cifras no debería llevarnos a engaño, ya que si se tiene en cuenta el coeficiente de variación para cada categoría de no respuesta, se aprecia que, independientemente de lo mucho que los datos diverjan de su valor medio, la mayor variabilidad se produce en el no contesta.

Pero quizás lo más importante es que el grueso de la no respuesta proviene del no sabe, mientras que el no contesta presenta un peso menor. Es decir, en términos agregados la mayoría de las personas entrevistadas que no responden a una pregunta de carácter político, alega que no sabe qué decir, mientras que una minoría se niega simplemente a contestar. En este sentido, los gráficos 4.1 a 4.3 reafirman este distinto comportamiento del no sabe y del no contesta. Los dos primeros gráficos prácticamente se superponen indicando la estrecha relación entre ambos fenómenos: la mayor parte de la no respuesta es producto de personas que no saben qué responder.

Gráfico 4.1. Distribución de los porcentajes totales de no respuesta

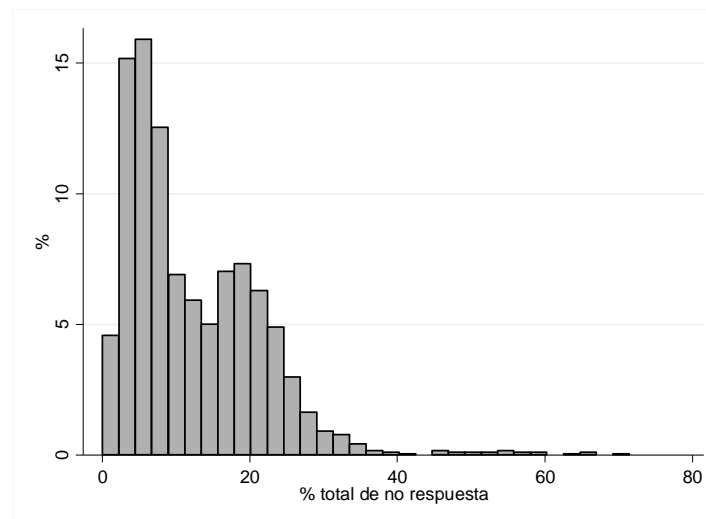


Gráfico 4.2. Distribución de los porcentajes relativos al no sabe

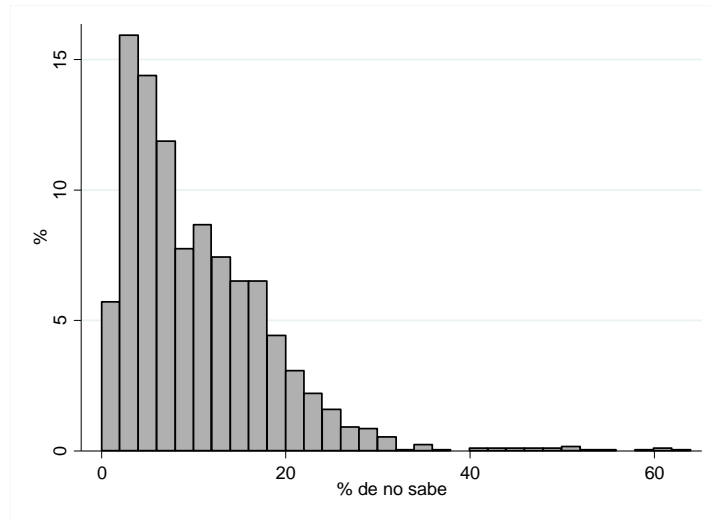
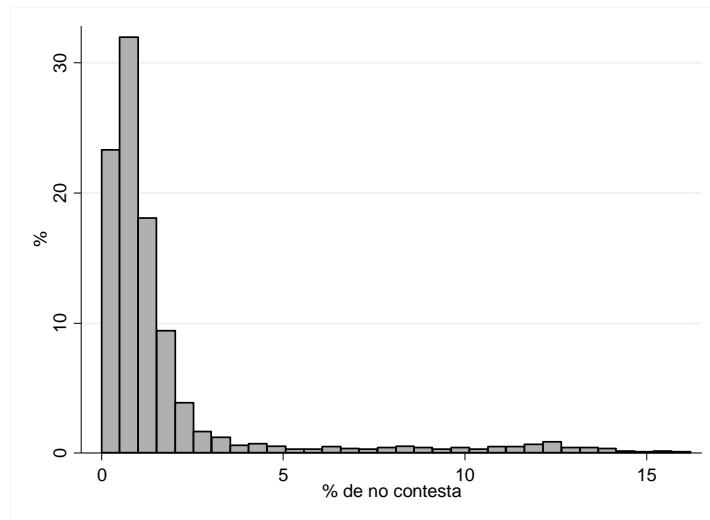


Gráfico 4.3. Distribución de los porcentajes relativos al no contesta



4.1. Algunas hipótesis acerca del efecto del diseño de las preguntas y del cuestionario sobre la no respuesta

A pesar de que parezca lógico e incluso probable que las personas no tengan siempre que ofrecer una respuesta a una pregunta concreta en una encuesta, esta investigación parte del supuesto de que la forma en que se diseña un cuestionario no es totalmente ajena a la circunstancia de que finalmente se responda o no. La forma en que se plantean las preguntas y cómo se organiza el propio cuestionario pueden desanimar o alentar la respuesta.

Concretamente, aspectos como la longitud de la pregunta, el lugar que ocupa en el cuestionario, el tipo de opciones de respuesta que se ofrecen y sus características, y otras particularidades del diseño de todos estos elementos, pueden influir en cómo se comporte el entrevistado. El objetivo de este apartado, consiste precisamente en apuntar y contrastar algunas hipótesis al respecto, hipótesis que se relacionan a continuación.

- **HP1a.** Cuanto más larga sea la pregunta, más difícil será para el entrevistado comprender su significado, y mayor será la no respuesta.

El fundamento lógico que subyace bajo este supuesto sugiere que una pregunta larga es mucho más difícil de entender que una pregunta más corta. Enunciados largos, complicados de seguir, pueden traer aparejada una mayor dificultad para enterarse y pensar sobre lo que se pregunta. Asimismo, se espera que el efecto de la longitud de la pregunta (medido en palabras) no sea lineal, sino que se suavice, ya que añadir más longitud no debe suponer el mismo incremento sobre la no respuesta pasado cierto umbral.

Sin embargo, podría darse el efecto contrario. También se podría plantear una hipótesis en la que la longitud de la pregunta supusiese un estímulo a la respuesta. En este caso, preguntas más largas pueden significar enunciados más clarificadores, la existencia de más elementos que aporten la información necesaria para responder, incluso el aliento a la propia respuesta. De ahí que hayamos formalizado la siguiente hipótesis alternativa:

- **HP1b.** Cuanto más larga sea la pregunta, más fácil será para el entrevistado comprender su significado, y menor será la no respuesta.

Una tercera hipótesis se relaciona con el lugar que ocupa la pregunta en el cuestionario, y que hemos formalizado como sigue:

- **HP2.** Cuanto más atrás se situó la pregunta en el cuestionario, mayor será el nivel de no respuesta.

El fundamento micro sobre el que se sustenta dicha afirmación tiene que ver con el llamado “efecto cansancio o fatiga” del entrevistado, efecto que habitualmente se advierte en los cuestionarios muy extensos. En estas situaciones, cuanto más tarde aparece la pregunta, más cansada se siente la persona entrevistada y más posibilidades existen de que no responda para evitar el esfuerzo que le supone generar una respuesta válida.

La lógica del cansancio también puede aparecer con las preguntas en batería donde una formulación inicial sirve para introducir la pregunta, que se aplica luego a una serie de ítems.

- **HP3a.** Las preguntas que forman parte de una batería producen niveles mayores de no respuesta.

Sin embargo, bien podría detectarse un comportamiento opuesto. Las preguntas que forman parte de una batería, donde el mismo mecanismo de pregunta-respuesta es repetido más de una vez, podrían allanar la comprensión y la comparación entre ítems, facilitando así la respuesta. De ahí que pueda contemplarse además, la siguiente hipótesis alternativa:

- **HP3b.** Las preguntas que forman parte de una batería producen niveles menores de no respuesta.

La cuarta hipótesis distingue entre el efecto de las preguntas que inician un tema, y que llamamos aquí “básicas”, frente a aquellas en las que se continúa preguntando por diferentes aspectos de la misma cuestión con el fin de recabar más información. A este último tipo las hemos denominado preguntas de “seguimiento” o “clarificación”. Un ejemplo de pregunta básica sería el siguiente: *Vamos a hablar ahora del hundimiento del petrolero Prestige frente a la costa gallega, y de la contaminación que ha provocado. ¿Podría decirme con qué interés ha seguido las informaciones y noticias sobre este suceso?* Y ejemplos de preguntas de seguimiento y/o clarificación son: *¿Le ha preocupado a Ud. mucho, bastante, poco o nada lo ocurrido?, ¿Cómo calificaría Ud. la situación creada por el hundimiento del petrolero y la contaminación que ha provocado?*

Partiendo de la premisa de que las preguntas básicas pueden resultar menos exigentes a la hora de ser contestadas que las preguntas de seguimiento o clarificación, para las que quizás se requiere tener mayor información o una opinión más formada, esperamos que las segundas desincentiven la respuesta frente a las primeras, de ahí que hayamos formulado la siguiente hipótesis:

- **HP4.** Las preguntas básicas producen niveles menores de no respuesta, en comparación con las preguntas de seguimiento o clarificación.

Por último, otras hipótesis hablan del aliento a la respuesta en el enunciado de la pregunta, las características de las opciones de respuesta o aspectos de la administración de la encuesta. Así,

- **HP5.** Las preguntas que contienen expresiones que animan a contestar, producen niveles menores de no respuesta.
- **HP6.** Las opciones de respuestas que cuentan con autonomía semántica total o parcial producen niveles menores de no respuesta, en comparación con las opciones sin esta autonomía.
- **HP7.** Las preguntas que contienen opciones de respuesta en forma de escala unipolar producen niveles menores de no respuesta, en comparación con las escalas bipolares.
- **HP8.** Las opciones de respuesta que contienen una categoría central producen niveles menores de no respuesta.
- **HP9.** En las preguntas en las que se ofrecen tarjetas a las personas entrevistadas, donde figuran escritas las opciones de respuesta para que las lean, los niveles de no respuesta son menores.
- **HP10.** Cuando el entrevistador ofrece leyéndolas las opciones de respuesta, los niveles de no respuesta son menores en comparación a cuando no las ofrece.

En casi todas estas hipótesis subyace la misma idea: ciertas particularidades del enunciado de la pregunta, de las opciones de respuesta o de cómo se gestiona la entrevista, facilitan el acto de contestar, principalmente porque reducen la complejidad potencial de aquello

por lo que se pregunta. En este sentido, la autonomía semántica y las escalas unipolares son un claro ejemplo de ello.

La autonomía semántica se refiere al grado de dependencia de significados entre opciones de respuesta. Así, resulta mucho más fácil responder si uno se considera católico, musulmán, judío o de otra confesión religiosa, que situarse en una escala de 0 a 10 para indicar, por ejemplo, cuánto se es de religioso. En el primer caso se trata de categorías de adscripción religiosa que no dependen las unas de las otras para su definición propia, mientras que en el segundo, el grado de religiosidad se establece después de comparar posiciones que se necesitan las unas a las otras para reflejar un continuo.

De forma similar, una escala bipolar obliga a la persona entrevistada a pensar en términos de “doble direccionalidad”, de dos polos con significados opuestos, aumentando con ello el esfuerzo cognitivo a la hora de generar una respuesta. Un caso típico consiste en la escala en la que se pide al encuestado que se sitúe en un continuo entre los siguientes polos “Cada uno se debe responsabilizar de su propio bienestar” y “El Estado es responsable del bienestar de todos los ciudadanos”.

Sin embargo, la hipótesis sobre la categoría central sigue un razonamiento distinto. No es que la existencia de este tipo de categoría haga menos difícil contestar *per se*. El esfuerzo cognitivo sigue siendo el mismo, pero de esta manera la persona entrevistada tiene a su alcance una categoría que actúa de refugio, que le permite dar una respuesta fácil y/o no comprometida si se siente presionada a responder para quedar bien.

Con el objeto de contrastar las hipótesis enumeradas hemos especificado un primer modelo de regresión de mínimos cuadrados ordinarios (MCO), donde la variable dependiente es el porcentaje de no respuesta total para cada una de las preguntas consideradas y las hipótesis esbozadas se han medido introduciendo las variables independientes correspondientes (véase el modelo 1 de la tabla 4.2). Asimismo, dado que

muchas de las preguntas se repiten y entonces las observaciones no son independientes, tal y como supone el modelo de regresión lineal, los errores típicos se han calculado teniendo en cuenta el hecho de que las observaciones están correlacionadas dentro de conglomerados (serie de preguntas), que a su vez son independientes entre sí.

Los resultados del modelo indican claramente que las características de las preguntas y el diseño de cuestionario influyen sobre la no respuesta de una manera notable, llegando a explicar un 32 por ciento de su varianza. Además, el signo de los coeficientes de regresión y su significatividad estadística permiten verificar la mayoría de las hipótesis apuntadas e incluso ponen de manifiesto la importancia, en cuanto a la magnitud de su efecto, de alguna de las variables consideradas.

Una pregunta larga, y que por tanto se entiende mejor, que se sitúe al principio del cuestionario, que sea básica y no de seguimiento o clarificación, en la que se aliente a responder y con opciones de respuesta semánticamente autónomas, en una escala unipolar y con categoría central, presenta niveles de no respuesta menores que otra en la que no se den todas estas circunstancias. En cambio, formar parte de una batería y las hipótesis relacionadas con las facilidades que se aportan durante la administración del cuestionario (ofrecimiento de tarjetas, lectura de opciones de respuesta) parecen no ejercer ningún tipo de influencia.

Asimismo, entre todos estos parámetros el efecto más importante se corresponde con la autonomía semántica. En una pregunta con esta característica la no respuesta disminuye en 10 puntos porcentuales, mientras que en segundo lugar destaca el estímulo a la respuesta que suponen las escalas unipolares frente a las bipolares. Concretamente, formular una pregunta a través de una escala unipolar disminuye la no respuesta en 5 puntos porcentuales, respecto de hacerlo a través de una escala bipolar. Por último, el aliento a la respuesta y la existencia de opciones de respuesta que contengan una

categoría central, también reducen la no respuesta de forma notable, excediendo su efecto los 3 puntos porcentuales en ambos casos.

Tabla 4.2. Los parámetros de la no respuesta

	Total no respuesta	
	(1)	(2)
Posición de pregunta en el cuestionario	0,065** (0,025)	0,072*** (0,020)
Longitud de pregunta (logaritmo)	-1,593* (0,753)	-1,246† (0,671)
La pregunta forma parte de una batería	1,829 (1,159)	0,665 (0,916)
Pregunta básica (<i>vs. de seguimiento o clarificación</i>)	-2,762† (1,644)	-1,020 (1,262)
Aliento a la respuesta	-3,499*** (0,839)	-2,955** (0,897)
Las opciones de respuesta tienen autonomía semántica	-10,157*** (1,733)	-9,075*** (1,581)
Tipo opción de respuesta (Categoría de ref. escala bipolar)		
Escala unipolar	-5,037* (2,478)	-3,944* (1,869)
Otros tipos	2,353* (0,941)	2,882** (0,877)
Se ofrece categoría central	-3,237** (0,979)	-2,916*** (0,710)
Se ofrece tarjeta	-0,998 (1,443)	1,792 (1,244)
Se leen las opciones de respuesta	-1,144 (0,752)	-0,912 (0,728)
Tipo de objeto (Categoría de ref. "Politics")		
Políticas públicas		-3,795*** (1,058)
Otros objetos		-0,051 (1,573)
Objeto concreto (<i>vs. abstracto</i>)		-1,059 (1,008)
Aparece referencia temporal en el enunciado de la pregunta		4,627* (1,860)
Se necesita conocimiento específico		5,362*** (0,866)
La pregunta se refiere al ámbito internacional		5,611*** (1,007)
Constante	29,312*** (3,507)	23,576*** (3,756)
R ²	0,316	0,405
Cambio en R ²		0,089
Error típico de la estimación (RMSE)	7,712	7,198
F	15,16	27,44
N	1.618	1.618
N de conglomerados (series de preguntas)	573	573

Entre paréntesis errores típicos agrupados por conglomerados de series de preguntas

*** p<0,001, ** p<0,01, * p<0,05, † p<0,1

Sin embargo, hasta aquí sólo hemos tenido en cuenta “aspectos formales”. Nuestras variables no nos dicen nada sobre el contenido de lo que se pregunta. Algunos elementos del diseño inciden en la no respuesta, aumentándola o disminuyéndola, porque estorban o facilitan la cognición y la formación posterior de una opinión. Pero puede que todo o parte de este efecto se deba a lo que se pregunta en sí, a la sustancia y no a la forma. De ahí que hayamos añadido en un segundo bloque, otro conjunto de variables que miden precisamente esta vertiente más sustantiva de las preguntas. Se trata de controlar el efecto de las características más formales de las preguntas y del diseño de cuestionario, a través de otras variables que buscan capturar la dificultad de contestar sobre según qué objetos o temas.

En la tercera columna de tabla 4.2 se muestran los coeficientes de regresión correspondientes a este segundo modelo. Además de las variables introducidas en el primer bloque, y que son las que se incluían en el primer modelo, se han añadido medidas que recogen:

- el tipo de objeto por el que se pregunta: política, instituciones y actores, cuestiones todas recogidas bajo el epígrafe de “politics”, frente a otros objetos como políticas públicas u otros temas no asimilables a estas dos categorías;
- la naturaleza del objeto: si se trata de un objeto concreto o abstracto;
- la existencia o no de referencias temporales en el enunciado de la pregunta;
- la necesidad o no de conocimiento específico para responder;
- y si la pregunta se refiere o no al ámbito internacional.

De este segundo modelo cabría destacar los siguientes aspectos. En primer lugar, la adición de este nuevo grupo de variables supone un aumento en el impacto agregado de la no respuesta muy modesto respecto del primer modelo: 0,9 puntos porcentuales.

En segundo lugar, la gran mayoría de las variables sobre el contenido ejercen el efecto esperado: cuando se necesita un conocimiento específico o se pregunta acerca de temas internacionales aumenta la no respuesta. La inclusión de referencias temporales lo mismo, pero aquí hay que tener en cuenta que el momento temporal que más deprime la respuesta es el futuro: elaborar predicciones sobre lo que ocurrirá sobre una determinada cuestión resulta mucho más difícil que evaluar, por ejemplo, hechos pasados. En cambio, las preguntas sobre políticas públicas influyen disminuyendo la no respuesta, quizás por lo cercano o tangible del objeto que representan, en comparación con la política y sus procesos.

Y en tercer y último lugar, y quizás también lo más interesante, es que este segundo bloque de variables de control permite separar hasta cierto punto entre el efecto de elementos vinculados al contenido, de los efectos más “formales” estrictamente atribuibles al diseño de las preguntas y el cuestionario. Los coeficientes de este último grupo de variables se “corrigen” y disminuye la magnitud de la mayoría. Pero lo más importante es que su efecto sigue siendo considerable y estadísticamente significativo. Así, las preguntas con opciones de respuesta semánticamente autónomas continúan reduciendo la no respuesta en cerca de 10 puntos porcentuales, las escalas unipolares lo hacen en 4 puntos, y el aliento a la respuesta y las categorías centrales contribuyen a emitir una respuesta válida en 3 puntos.

4.2. Distinguiendo entre el no sabe y el no contesta ¿Se comportan igual?

Hasta aquí hemos analizado los efectos del diseño de las preguntas y del cuestionario sobre la no respuesta, incluyéndose en ella indistintamente tanto la debida al no sabe como al no contesta. Pero ¿se comportan igual? ¿reaccionan de la misma forma y con la misma intensidad a las características de las preguntas y a cómo se ha organizado el cuestionario?

Si el no sabe se debe a la falta de conocimientos u opiniones y el no contesta responde al carácter delicado o sensible de lo que se pregunta, prevemos que su comportamiento tendría que ser un tanto diferente ante determinados estímulos. En concreto, aquellas preguntas que supusiesen un menor esfuerzo cognitivo tendrían que ejercer un mayor efecto sobre la reducción de la no respuesta en el caso del no sabe, en comparación con el no contesta.

Con el fin de verificar nuestras expectativas hemos vuelto a replicar los modelos de regresión lineal especificados en el apartado anterior, pero distinguiendo en esta ocasión entre ambos tipos de no respuesta. Así los dos primeros modelos de la tabla 4.3 tienen como variable dependiente el porcentaje de no sabe que recogen las preguntas incluidas en nuestra base de datos, y en los dos siguientes modelos de la misma tabla, la variable dependiente se refiere al porcentaje de no contesta.

Los resultados del modelo 1 señalan que el diseño de las preguntas y el cuestionario explican el 21 por ciento de la varianza del no sabe. Los coeficientes de regresión de este primer modelo indican que preguntas “básicas” con enunciados largos, en los que además se alienta a que se conteste, junto con opciones de respuesta autónomas y una categoría central provocan una reducción notable del no sabe, mientras que las preguntas que forman parte de una batería lo aumentan.

Por su parte, los resultados del modelo 3 revelan que el diseño de preguntas y cuestionarios explica el 64 por ciento de la varianza del no contesta, porcentaje mucho mayor en comparación con el que explicaba el no sabe. Y también muestran la relevancia de la magnitud y la significación estadística de algunos de los coeficientes de regresión, tal es el caso del aliento a la respuesta, la autonomía semántica, el ofrecimiento de una categoría central o el de una escala unipolar. En cambio la longitud del enunciado no despliega ningún efecto, y las baterías de preguntas, lejos de provocar la fatiga del entrevistado, favorecen que genere una respuesta válida.

Tabla 4.3. Los parámetros del no sabe y del no contesta

	No sabe		No contesta	
	(1)	(2)	(3)	(4)
Posición de pregunta en el cuestionario	0,046† (0,026)	0,045* (0,020)	0,018** (0,006)	0,026*** (0,004)
Longitud de pregunta (logaritmo)	-2,277* (0,987)	-2,026* (0,871)	0,607 (0,481)	0,719* (0,365)
La pregunta forma parte de una batería	2,968* (1,445)	1,662 (1,148)	-1,285* (0,564)	-1,136** (0,397)
Pregunta básica (vs. de seguimiento o clarificación)	-3,034† (1,630)	-1,250 (1,317)	0,351† (0,203)	0,267 (0,222)
Aliento a la respuesta	-2,265* (0,937)	-2,142* (0,978)	-1,143** (0,359)	-0,817** (0,299)
Tipo opción de respuesta (Categoría de ref. escala bipolar)				
Escala unipolar	-3,586 (2,283)	-3,177† (1,821)	-1,467** (0,509)	-0,897** (0,340)
Otros tipos	2,231* (1,000)	2,800** (0,914)	0,074 (0,225)	-0,013 (0,209)
Las opciones de respuesta tienen autonomía semántica	-6,138** (2,240)	-5,468** (1,946)	-4,086** (1,369)	-3,588*** (1,043)
Se ofrece categoría central	-2,242* (1,025)	-1,888* (0,798)	-1,111** (0,373)	-1,105*** (0,327)
Se ofrece tarjeta	-1,536 (1,219)	1,068 (1,067)	0,455 (0,411)	0,664† (0,386)
Se leen las opciones de respuesta	-0,869 (0,732)	-0,785 (0,699)	-0,324 (0,199)	-0,215 (0,162)
Tipo de objeto (Categoría de ref. "Politics")				
Políticas públicas		-4,000*** (1,205)		0,087 (0,317)
Otros objetos		-1,997 (2,145)		1,873* (0,727)
Objeto concreto (vs. abstracto)		-1,103 (1,137)		0,088 (0,281)
Aparece referencia temporal en el enunciado de la pregunta		4,634* (1,910)		-0,273 (0,322)
Se necesita conocimiento específico		4,544*** (0,949)		0,738** (0,239)
La pregunta se refiere al ámbito internacional		5,983*** (0,989)		0,189 (0,182)
Constante	25,273*** (4,054)	21,912*** (4,830)	4,525** (1,494)	2,206 (1,498)
R ²	0,205	0,308	0,635	0,689
Cambio en R ²		0,103		0,054
Error típico de la estimación (RMSE)	7,287	6,799	1,706	1,576
F	9,015	18,85	8,074	12,92
N	1.609	1.609	1.616	1.616
N de conglomerados	566	566	571	571
R ² cuando se introduce primero el segundo bloque de variables	0,112	0,308	0,186	0,689
Cambio en R ²		0,196		0,503

Entre paréntesis errores típicos agrupados por conglomerados de series de preguntas

*** p<0,001, ** p<0,01, * p<0,05, † p<0,1

Hasta aquí parece que nuestras expectativas no se cumplen. Los modelos explican mejor la variación del no contesta frente al no sabe, pero lo más relevante del caso es que prácticamente las mismas variables desempeñan un efecto similar sobre ambos fenómenos. El diseño de las preguntas y la forma de articular el cuestionario facilitan la respuesta tanto cuando creemos que no se responde porque no se sabe, como cuando no se hace porque no se quiere. E incluso el refugio que supondría la existencia de una categoría central parece funcionar mejor a la hora de inhibir la no respuesta debida al no sabe. En otras palabras, si las variables que recogen los aspectos más formales del diseño, en tanto que en su mayoría facilitadoras o no de la cognición, explican en mayor medida el comportamiento del no contesta, entonces la distinción inicial entre un no sabe debido a la falta de conocimientos u opiniones y un no contesta debido al carácter delicado o sensible de lo que se pregunta, parece no tener sentido.

Con todo, a los modelos comentados se ha vuelto a añadir en un segundo bloque o paso el conjunto de variables que aportan información más sustantiva sobre lo que se pregunta. Se vuelve a observar y se ratifica que el diseño de preguntas y cuestionarios facilitan sobre todo la respuesta en el caso del no contesta. Pero lo más importante es que el comportamiento del no sabe se halla mucho más sujeto a las dificultades de generar una opinión que entrañan preguntas sobre objetos complejos que requieren de una mayor competencia cognitiva.

Así, mientras que el modelo 3 respecto del 4 supone una mejora de 5 puntos porcentuales en la varianza explicada del no contesta, el modelo 2 añade una mejora de 10 puntos en la del no sabe, frente al modelo 1. Es decir, las variables que miden el tipo de objeto, el ámbito temporal y territorial en el que se sitúa, y los conocimientos de los que precisa, ayudan a explicar mejor porque no se responde con un no sé a determinadas preguntas. Por el contrario, y con la excepción de la necesidad de un conocimiento específico, las variables de carácter más sustantivo introducidas en este segundo paso, no ejercen

ningún efecto significativamente estadístico sobre la no respuesta de los que no quieren contestar.

Este resultado redonda en la idea de que detrás del no sabe existe realmente una dificultad para formarse y manifestar una opinión, pero globalmente parece claro que el diseño de las preguntas y de los cuestionarios tiene un impacto más importante sobre la variación en el no contesta que en el no sabe.

En un estudio similar sobre los factores explicativos del no sabe y del no contesta Shoemaker *et al.* (2002) también partían del supuesto de la distinción entre ambos fenómenos. Pensaban que se debían a razones diferentes (esfuerzo cognitivos vs. cuestión sensible). Sin embargo, aunque verificaron parte de sus premisas iniciales, se encontraron con un resultado inesperado: el esfuerzo cognitivo tal como lo medían también presentaba una relación positiva con el rechazo a contestar.

De ahí que argumentasen que este hallazgo quizás se debiese a un planteamiento erróneo. Las personas entrevistadas bien podrían comportarse de dos maneras diferentes: 1) se enfrentan desde un principio a la pregunta procesándola cognitivamente y se dan cuenta después de que no saben qué responder manifestándolo así, o 2) en el mismo instante en que se les pregunta perciben la dificultad de la cuestión y ni siquiera se molestan en pensar en ella, respondiendo que no quieren contestar.

Sea o no cierta esta distinción, en este apartado se ha comprobado que el diseño de las preguntas y cuestionarios incide de manera especial en la inhibición de la respuesta que se consigna con un no contesta, mientras que su impacto es más modesto reduciendo el no sabe. Este último parece mucho más ligado al contenido de la pregunta, a la dificultad sustantiva que entraña, mientras que el no contesta puede reducirse si se facilita la cognición a través de aspectos más formales.

5. Lo que nos dice el efecto conjunto de las características de los individuos y el diseño del cuestionario sobre la no respuesta: validando algunas hipótesis

5.1. Introducción

Para complementar el análisis realizado a nivel de pregunta, cuyos resultados hemos presentado en el apartado anterior, introducimos también la variación individual en la propensión a responder en nuestro análisis. Para ello, como se ha explicado, hemos construido una base de datos que agrupa los individuos de los 110 barómetros de interés: recogemos como variable dependiente la respuesta o no a las preguntas seleccionadas, y como variables explicativas combinamos factores individuales (características sociodemográficas y políticas de los encuestados) con las características de la pregunta que hemos tratado en el apartado anterior.

De este modo podremos ver hasta qué punto son importantes uno y otro tipo de factores para entender la no respuesta a las encuestas. Adicionalmente, este diseño nos permitirá analizar la interacción entre las características de las preguntas y de los individuos. Así, podremos entender qué tipo de encuestados se ven más afectados por los diversos efectos de la pregunta: se trata de una cuestión crucial, ya que si determinadas características de pregunta hacen crecer la no respuesta entre determinados perfiles de encuestados, pueden producirse sesgos importantes en los resultados de las encuestas. En otras palabras, si el efecto depresor de la respuesta no se distribuye de modo aleatorio sino sistemático entre los encuestados, esto puede tener consecuencias graves para las estimaciones.

5.2. Algunas hipótesis acerca del efecto de las características de los individuos sobre la no respuesta

Parece razonable pensar que no todos los individuos tienen la misma propensión a responder a las preguntas que les plantean los encuestadores, en especial cuando se trata de asuntos políticos. Básicamente, podemos pensar en factores relacionados con los recursos cognitivos de que disponen los ciudadanos, así como de su socialización e identificación política.

Los recursos cognitivos son importantes puesto que responder determinadas preguntas de encuesta requiere de ciertos conocimientos previos y un esfuerzo de abstracción por parte del encuestado. No todos los ciudadanos están en las mismas condiciones de realizarlo, y para unos el coste será mucho menor que para otros.

El factor más relevante son los recursos cognitivos, capturados fundamentalmente a través del nivel educativo de los ciudadanos. Así, esperamos que aquellos ciudadanos con mayor nivel de estudios tengan una mayor tendencia a contestar a las preguntas de encuesta sobre actitudes políticas. Adicionalmente, consideraremos la clase social como otra medida más de los recursos disponibles por parte de los entrevistados. Es de esperar que las clases sociales más altas respondan con mayor frecuencia a las preguntas políticas de encuesta. Finalmente, la situación laboral de los encuestados (activos o no) nos permitirá capturar también otra dimensión de los recursos: esperamos que los ciudadanos con una vida laboral activa tengan, en promedio, mayores niveles de contacto social, que les puedan proporcionar instrumentos relevantes a la hora de poder contestar con facilidad preguntas de encuesta.

Por otra parte, más allá de los recursos, hay factores de socialización y/o de contexto que pueden influir en la propensión de los ciudadanos a responder preguntas de encuesta de carácter político. En determinados entornos, o para determinados tipos de persona, las reservas a expresar a un desconocido las propias posiciones políticas pueden resultar mayores que en otros. Así, esperamos que las personas de mayor edad, y los residentes en municipios pequeños, tengan menos probabilidad de responder este tipo de preguntas.

En el caso del género también esperamos encontrar un efecto. Las mujeres tienen un nivel promedio de conocimiento político menor que los hombres, y su rol tradicional les lleva a no expresarlo del mismo modo que los hombres, por lo que esperamos también menores niveles de respuesta entre ellas.

Finalmente, ya que analizamos específicamente preguntas con contenido político, nos fijaremos también en las identidades políticas de los encuestados. Dinámicas de la opinión pública o del contexto pueden favorecer o dificultar la expresión de actitudes políticas por parte de determinados grupos de ciudadanos. Para capturar este efecto, incluimos la ideología y el recuerdo de voto. De entrada, podemos esperar que los ciudadanos que tienen una identificación ideológica tendrán más probabilidad de responder que los que no la tienen, mientras que los que recuerdan haber votado algún partido tendrán más probabilidades de responder que los abstencionistas o los que no expresan o no recuerdan su voto.

5.2.1. Los factores individuales de la no respuesta

Para contrastar las hipótesis sobre el efecto de los factores individuales realizamos análisis de regresión logística con la base de datos integrada, de individuos y preguntas.

Como variables dependientes usaremos, en primer lugar, una variable dicotómica que identifica los individuos que no respondieron a cada pregunta (1) frente a los que sí lo hicieron (0). En segundo lugar, separaremos la no respuesta en dos: por un lado los que respondieron (0) frente a los que registran la opción “no sabe” (1), y por otro frente a la opción “no contesta” (1). La distribución de frecuencias está representada en la tabla 5.1. Como se puede observar, de los 4.067.092 individuos-pregunta que contiene la base de datos, tenemos un 88 por ciento que dieron alguna respuesta a las preguntas, un 10 que eligieron la opción “no sabe” y cerca de un 2 por ciento que no contesta.

Tabla 5.1. Distribución de frecuencias de la no respuesta a nivel de individuos-pregunta

	Núm. observaciones	Porcentaje
Contesta	3.566.411	87,7
No sabe	425.262	10,5
No contesta	75.419	1,9
Total	4.067.092	100

La tabla 5.2 nos muestra los resultados de los modelos referentes a las variables individuales para contrastar las hipótesis mencionadas anteriormente. El primer modelo tiene como variable dependiente la no respuesta en general, el segundo el no sabe y el tercero el no contesta. ¿Qué nos dicen los resultados respecto a las hipótesis del nivel individual que hemos formulado?

Por lo que respecta a los recursos, vemos cómo el nivel de estudios tiene una gran influencia en la probabilidad de responder. A mayor nivel educativo de los individuos, mayor es su probabilidad de responder a las preguntas analizadas. Adicionalmente, aunque con un efecto menor, también observamos que los profesionales y directivos son el grupo ocupacional con mayor propensión a responder, y los agricultores el que menos, seguidos de los trabajadores manuales –especialmente los no cualificados. Los

encuestados que trabajan también tienen mayores probabilidades de responder que los jubilados, estudiantes y los dedicados al hogar.

Tabla 5.2. Los efectos de las características de los individuos sobre la no respuesta

	Total no respuesta b/sig.	No sabe b/sig.	No contesta b/sig.
Edad (0-1)	0,55***	0,48***	0,89***
Sexo (ref. mujer)	0,33***	0,39***	-0,02
Tamaño de municipio (0-1)	0,14***	-0,26***	0,54***
Ubicación ideológica eje izquierda-derecha (categoría de ref. centro)			
Izquierda (1-4)	-0,24***	-0,26***	-0,08*
Derecha (6-10)	-0,12***	-0,14***	0,05
No se ubica	1,08***	0,93***	2,01***
Recuerdo de voto en elecciones generales (categoría de ref. PSOE)			
PP	-0,08***	-0,08***	-0,03
IU	-0,09***	-0,12***	0,22***
Otros	-0,09***	-0,13***	0,17***
Abstención	0,23***	0,23***	0,29***
No recuerda / No contesta	0,22***	0,11***	0,76***
Nivel de estudios (categoría de ref. sin estudios)			
Estudios primarios	-0,34***	-0,38***	-0,05
Educación secundaria	-0,56***	-0,65***	-0,00
FP	-0,58***	-0,67***	0,03
Estudios superiores	-0,76***	-0,91***	0,04
Situación laboral (categoría de ref. trabaja)			
Jubilado/pensionista	0,17***	0,21***	-0,11***
Parado	0,07***	0,10***	-0,06*
Estudiante	0,18***	0,25***	-0,18***
Labores del hogar	0,10***	0,14***	-0,13***
Clase social (categoría de ref. profesionales y directivos)			
Trabajadores no manuales	0,05***	0,06***	0,04
Empresarios/autónomos	0,15***	0,17***	0,07*
Agricultores	0,41***	0,46***	0,10*
Trabajadores manuales cualificados	0,18***	0,21***	0,05
Trabajadores manuales no cualificados	0,22***	0,26***	-0,01
Constante	-2,72***	-2,80***	-5,46***
Observaciones	3.909.912	3.839.081	3.500.444
Pseudo-r2	0,099	0,092	0,134

* p < 0,05, ** p < 0,01, *** p < 0,001

Pero no es sólo en el ámbito de los recursos donde encontramos efectos relevantes: la edad tiene un efecto negativo sobre la probabilidad de responder (a más edad, menos respuesta), igual que el sexo (las mujeres tienden a responder menos) y el hábitat (se responde más en ciudades más grandes y menos en pueblos pequeños).

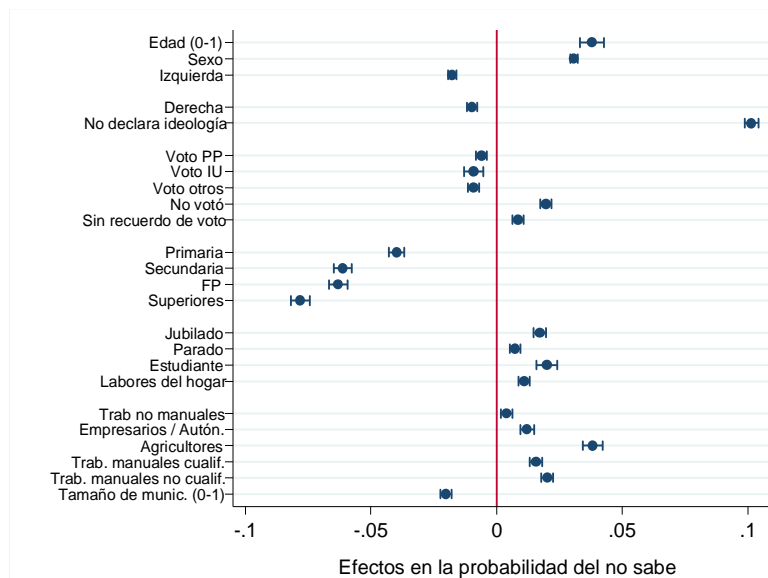
Finalmente, por lo que respecta a las identidades políticas, observamos que los tienen una posición ideológica clara, tanto a la izquierda como a la derecha, responden más que los encuestados que se ubican en el centro, mientras que los encuestados que no se ubican en este eje tienen una probabilidad mucho menor de responder. Algo similar ocurre con el recuerdo de voto: los abstencionistas y los que no expresan recuerdo de voto tienen menor probabilidad de responder, mientras que entre los votantes de los diversos partidos, las diferencias son menores siendo los electores del PSOE (categoría de referencia) los que menos responden de entre los que declaran voto a un partido.

Sin embargo, cuando separamos el análisis para distinguir entre el “no sabe” y el “no contesta”, observamos diferencias relevantes: las pautas generales no se repiten del mismo modo para ambos casos. Todas las variables pensadas para medir los recursos (estudios, clase social, situación laboral) tienen un efecto mucho más claro y fuerte en la probabilidad de elegir no sabe que en el caso del no contesta. De hecho, el nivel de estudios y la clase social no resultan significativos para explicar la probabilidad de contestar, mientras que la situación laboral presenta un efecto más atenuado que en el caso del no sabe. Esto nos indica que la opción no sabe es la que recoge principalmente las dificultades cognitivas, como cabría esperar.

Por otra parte, las mujeres tienen mayor probabilidad de decir que no saben que los hombres, mientras que en el caso del no contesta no detectamos diferencias significativas de género. Igualmente, el tamaño de municipio tiene efectos diferenciados: el no sabe es más probable de obtener en encuestados residentes en municipios pequeños, mientras

que el no contesta actúa en la dirección opuesta, y es más probable entre los que residen en grandes municipios. Por lo tanto, vemos cómo los efectos derivados del contexto y la socialización actúan en direcciones opuestas o al menos divergentes para una y otra categoría de la no respuesta. Finalmente, observamos también que los efectos de las identidades políticas (ideología y recuerdo de voto) son más marcados en la probabilidad de no contestar. Los gráficos 5.1 y 5.2 nos muestran los efectos marginales de las variables individuales sobre la probabilidad del no sabe y del no contesta respectivamente. En concreto, estos efectos expresan el cambio en la probabilidad del no sabe y el no contesta para un cambio unitario en cada una de las variables, dejando constantes el resto de variables del modelo en su valor medio. Se trata de una representación gráfica que nos ayuda a interpretar la magnitud de los efectos de modo mucho más claro que en el caso de los coeficientes logit.

Gráfico 5.1. Efectos marginales sobre el no sabe

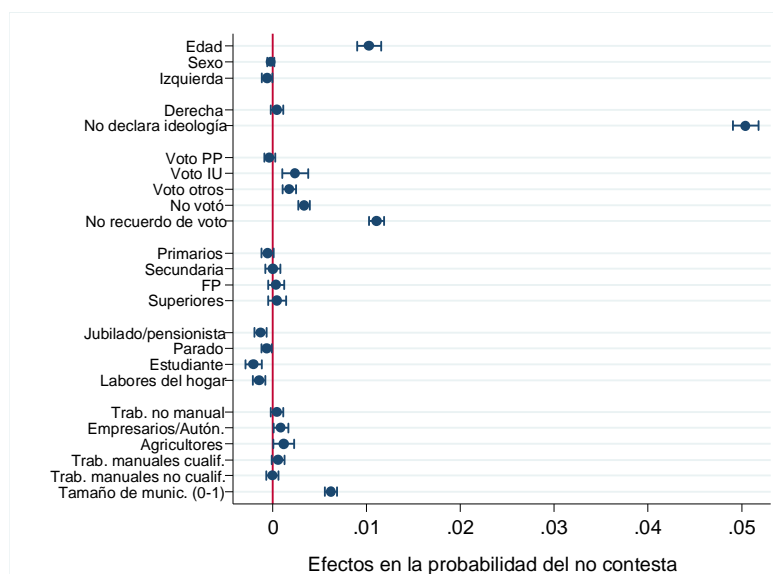


Así, vemos cómo el determinante más fuerte del no sabe es el nivel de estudios, a parte del hecho de no declarar ideología. La probabilidad de responder no sabe se reduce significativamente a medida que el encuestado tiene un mayor nivel de estudios, siendo la reducción en el caso de los estudios superiores casi de un -0,1 frente a los que no tienen

estudios, *ceteris paribus*. Las otras variables tienen efectos menos marcados, aunque los de la edad y el sexo no resultan tampoco desdeñables en su magnitud, que se acerca a 0,04: ser mujer, y ser de edad más avanzada aumentan la probabilidad del no sabe.

Por lo que respecta al no contesta, observamos cómo no hay ningún factor explicativo que tenga un impacto de magnitud similar al del no sabe, lo que está fundamentalmente relacionado con la distribución de la variable: recordemos que sólo un 1,85 por ciento de los individuos-pregunta tienen el valor no contesta. En este sentido, aparte de no declarar ideología (que es conceptualmente muy similar a la variable dependiente) encontramos efectos relevantes de la edad, el abstencionismo y el tamaño de municipio, en la línea de lo comentado anteriormente.

Gráfico 5.2. Efectos marginales sobre el no contesta



5.2.2. Los factores individuales y de pregunta sobre la no respuesta

Si ampliamos los modelos individuales que hemos analizado para incluir también en ellos las características de las preguntas, obtenemos unos resultados esencialmente

equivalentes a los que hemos visto hasta ahora. El ajuste de los modelos mejora sustancialmente, especialmente en el caso del no contesta, lo que va en la línea de los análisis agregados presentados en el apartado 4: los factores relacionados con la formulación y contenido de la pregunta tienen más efecto sobre el no contesta que sobre el no sabe.

En estos modelos encontramos también un efecto de cansancio: cuanto más adelante en el cuestionario esté una pregunta, mayores son las tasas de no respuesta. Por lo que respecta a la longitud de la pregunta, observamos un efecto negativo en el no sabe y positivo en el no contesta: las preguntas más largas reducen la probabilidad del no sabe, probablemente porque están mejor explicadas, pero aumentan la de no contestar, posiblemente por un efecto de cansancio. El hecho de estar o no en una batería también tiene este efecto ambivalente, pero en dirección opuesta: aumenta la probabilidad del no sabe, pero disminuye la del no contesta.

Las preguntas básicas, las escalas unipolares, las que tienen respuestas con autonomía semántica y las que incluyen una categoría central cosechan, *ceteris paribus*, menores niveles de no respuesta, tanto en global como para el no sabe y el no contesta. En cambio, cuando se muestra la tarjeta observamos más no respuesta (probablemente como efecto de la complejidad de la pregunta), mientras que el aliento a la respuesta parece funcionar bastante para reducir la no respuesta. Leer las opciones también tiene un efecto, pero mucho menor.

Finalmente, observamos también en estos modelos cómo las preguntas referidas al proceso político (enfrente de las de políticas públicas y otras) se contestan menos, al igual que ocurre con aquellas referidas a objetos abstractos, que requieren de conocimiento específico o que pertenecen al ámbito internacional, especialmente para el no sabe.

Tabla 5.3. Los efectos de las características de los individuos y de las preguntas sobre la no respuesta

	Total no respuesta	No sabe	No contesta
	b/se	b/se	b/se
Edad (0-1)	0,53***	0,47***	0,81***
Sexo (Mujer)	0,34***	0,40***	0,00
Tamaño de municipio (0-1)	-0,15***	-0,26***	0,52***
Ubicación ideológica eje izquierda-derecha (categoría de ref. centro)			
Izquierda (1-4)	-0,25***	-0,27***	-0,09*
Derecha (6-10)	-0,13***	-0,15***	0,04
No se ubica	1,17***	1,02***	2,54***
Recuerdo de voto en elecciones generales (categoría de ref. PSOE)			
PP	-0,07***	-0,07***	-0,03
IU	-0,06*	-0,10***	0,22***
Otros	-0,09***	-0,12***	0,13***
Abstención	0,25***	0,25***	0,30***
No recuerda / No contesta	0,26***	0,16***	0,73***
Nivel de estudios (categoría de ref. sin estudios)			
Estudios primarios	-0,39***	-0,42***	-0,14***
Educación secundaria	-0,60***	-0,68***	-0,08
FP	-0,64***	-0,72***	-0,09*
Estudios superiores	-0,84***	-0,97***	-0,09*
Situación laboral (categoría de ref. trabaja)			
Jubilado/pensionista	0,19***	0,23***	-0,08*
Parado	0,05***	0,07***	-0,09***
Estudiante	0,19***	0,25***	-0,17***
Labores del hogar	0,13***	0,17***	-0,08*
Clase social (categoría de ref. profesionales y directivos)			
Trabajadores no manuales	0,06***	0,07***	0,05
Empresarios/autónomos	0,14***	0,16***	0,06
Agricultores	0,44***	0,48***	0,14**
Trabajadores manuales cualificados	0,18***	0,21***	0,07*
Trabajadores manuales no cualificados	0,22***	0,27***	0,00
Posición de pregunta en el cuestionario	0,72***	0,66***	1,03***
Longitud de pregunta (log, 0-1)	-0,62***	-0,86***	0,40***
Pregunta básica (vs. de seguimiento o clarificación)	-0,06***	-0,10***	0,20***
La pregunta forma parte de una batería	0,11***	0,18***	-0,34***
Escala unipolar (ref. bipolar)	-0,34***	-0,31***	-0,73***
Otros tipos	0,41***	0,40***	0,45***

Las opciones de respuesta tienen autonomía semántica	-0,83***	-0,75***	-1,23***
Se ofrece categoría central	-0,28***	-0,23***	-0,67***
Se ofrece tarjeta	0,21***	0,16***	0,47***
Aliento a la respuesta	-0,31***	-0,27***	-0,54***
Se leen las opciones de respuesta	-0,06***	-0,05***	-0,01
Políticas públicas (ref. política)	-0,49***	-0,49***	-0,65***
Otros objetos	-0,07***	-0,15***	0,36***
Objeto concreto (vs. abstracto)	-0,14***	-0,16***	-0,05
Aparece referencia temporal	0,58***	0,63***	0,07*
Requiere conocimiento específico	0,60***	0,62***	0,35***
Ámbito internacional	0,61***	0,63***	0,10*
Constante	-1,85***	-1,84***	-5,21***
Observaciones	3.860.353	3.790.610	3.453.950
Pseudo-r2	0,148	0,135	0,286

* p < 0.05, ** p < 0.01, *** p < 0.001

5.3. Algunas hipótesis acerca de la interacción entre características individuales y del diseño de las preguntas y del cuestionario sobre la no respuesta

Una extensión lógica de la discusión anterior es la de la interacción entre variables individuales y de pregunta. Si determinadas características de las preguntas, como hemos visto ya en el apartado anterior, tienen efectos depresores de la respuesta, y al mismo tiempo determinadas características individuales también determinan la propensión a responder o no a las preguntas de encuesta, parece razonable esperar que ambos tipos de factores interactúen entre ellos.

Recordemos que hemos analizado, en el apartado anterior, el efecto de dos tipos de variables de pregunta: las relativas a las características formales de las preguntas, y las que hacen referencia al contenido de las mismas. En el caso de la interacción, esperamos que las características formales de las preguntas interactúen con los recursos, especialmente cognitivos: todas aquellas características que incrementan la complejidad de la pregunta, y que hacen la tarea de responderla más exigente desde el punto de vista cognitivo, afectarán especialmente a la probabilidad de responder de aquellos

encuestados con menores recursos cognitivos. Esperamos, pues, que los ciudadanos con menores niveles de estudios sean los más afectados por el cansancio y las dificultades que el formato de la pregunta puede entrañar.

Asimismo, hemos visto como determinadas características substantivas de las preguntas pueden deprimir los niveles de no respuesta: el ámbito al que se refieren, si requieren conocimiento específico, etc. Estas características, previsiblemente interactuarán con el nivel educativo, puesto que es de esperar que también en este caso encontremos mayores efectos entre los encuestados con menos recursos cognitivos.

Pero a diferencia de las variables formales, esperamos que las variables de contenido interactúen también con factores relacionados con la socialización de los encuestados. Nos fijaremos en especial en los efectos heterogéneos por género. Así, preguntas sobre el proceso político, o que tengan que ver con objetos abstractos y/o con determinados ámbitos pueden provocar proporcionalmente mayores tasas de no respuesta entre las mujeres.

En el siguiente apartado realizaremos el análisis empírico que nos permitirá contrastar estas hipótesis y así obtener una perspectiva completa sobre los determinantes de la no respuesta, considerando simultáneamente los efectos de la formulación y contenido de las preguntas, y los de los individuos, así como sus interacciones.

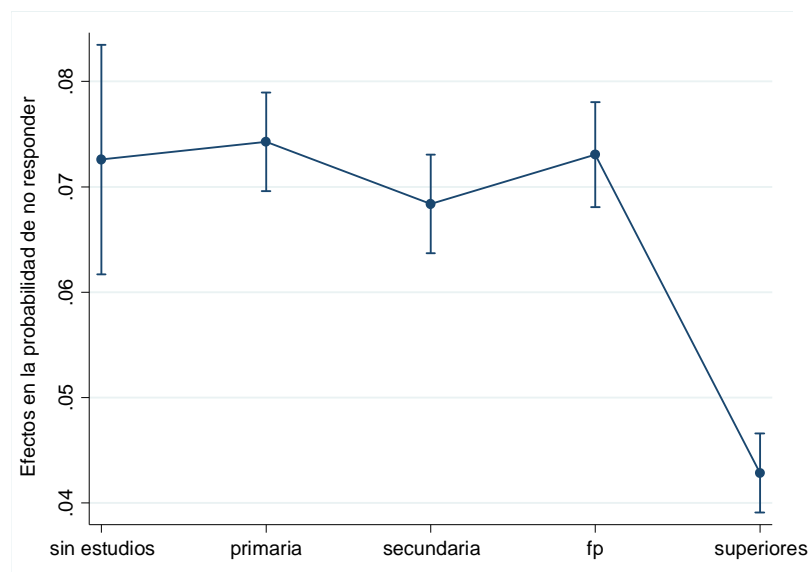
5.3.1 El efecto interactivo entre los factores individuales y los de pregunta sobre la no respuesta

Para contrastar las hipótesis de interacción, repetiremos los modelos de la tabla 5.3, incluyendo las interacciones relevantes en cada caso. Por simplicidad de análisis, consideraremos el efecto de las interacciones sobre la probabilidad de no ofrecer

respuesta a las preguntas consideradas, sin entrar en este apartado en la distinción entre el no sabe y el no contesta. Representaremos aquí gráficamente los efectos para una mejor comprensión de los resultados, dada la complejidad de interpretar directamente coeficientes de interacción en modelos no lineales como los que estimamos.

Nos fijaremos primero en los efectos condicionales de las características formales de la pregunta por nivel de estudios, para determinar hasta qué punto este tipo de factores afecta desproporcionalmente a determinados perfiles de encuestados. En primer lugar, vemos cómo la posición de la pregunta en el cuestionario, que recoge el efecto de cansancio, aumenta la probabilidad de no respuesta en todos los grupos, pero la intensidad del efecto es mucho menor para los encuestados con estudios superiores, que parecen mucho menos afectados por la fatiga.

Gráfico 5.3. Efecto de posición en el cuestionario por nivel de estudios



Hemos visto también cómo la longitud de la pregunta parecía ayudar a la respuesta: preguntas más largas cosechaban, *ceteris paribus*, mayores tasas de respuesta. Sin embargo, alargar la formulación de la pregunta para, presumiblemente, explicarla mejor, parece tener un efecto especialmente fuerte entre los encuestados con menor nivel de

estudios, sobre todo los que no tienen estudios terminados. Los encuestados con estudios superiores también ven reducida su probabilidad de no responder cuanto más larga es la pregunta pero a un ritmo menor.

Gráfico 5.4. Efecto de longitud de pregunta por nivel de estudios

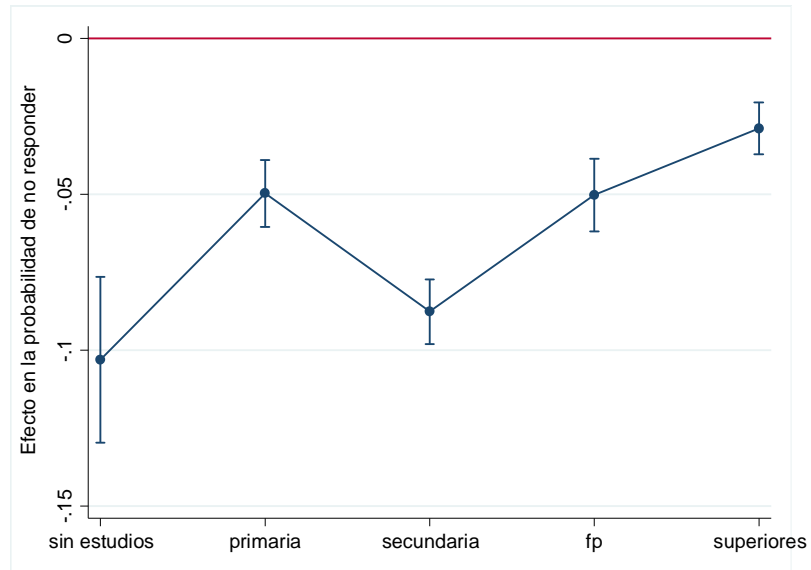
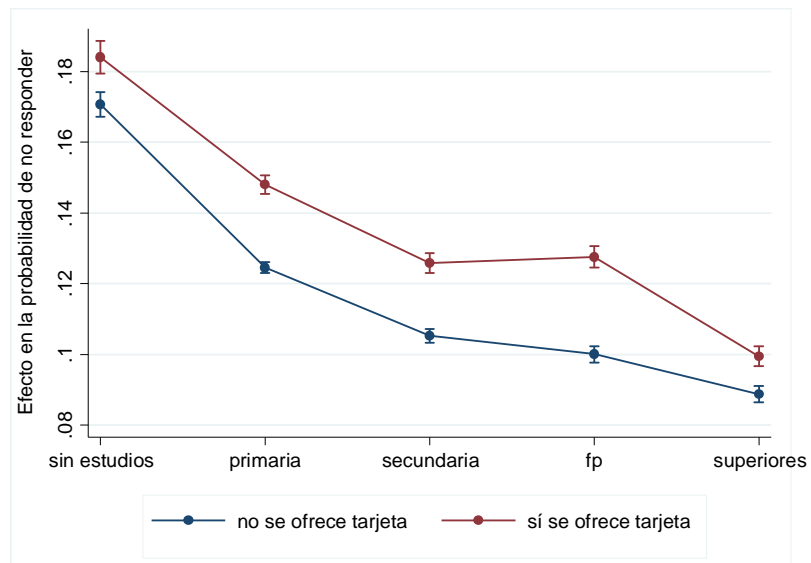
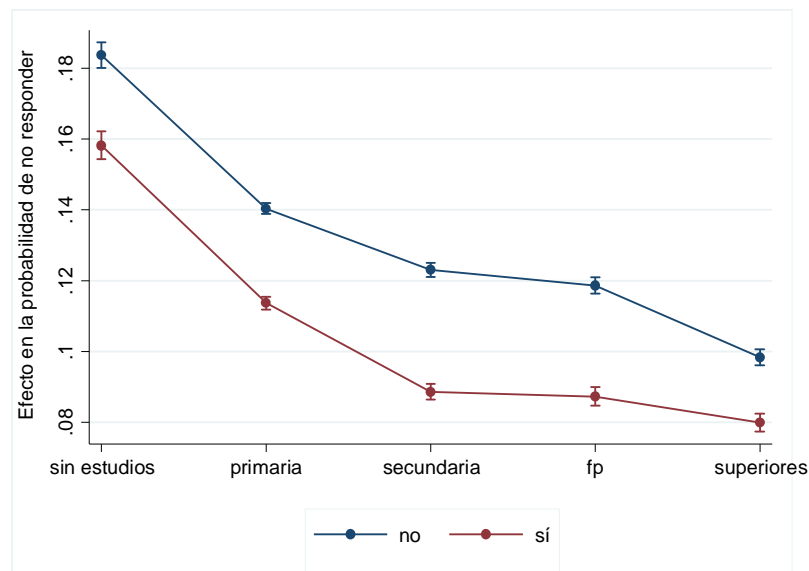


Gráfico 5.5. Efecto de tarjeta por nivel de estudios



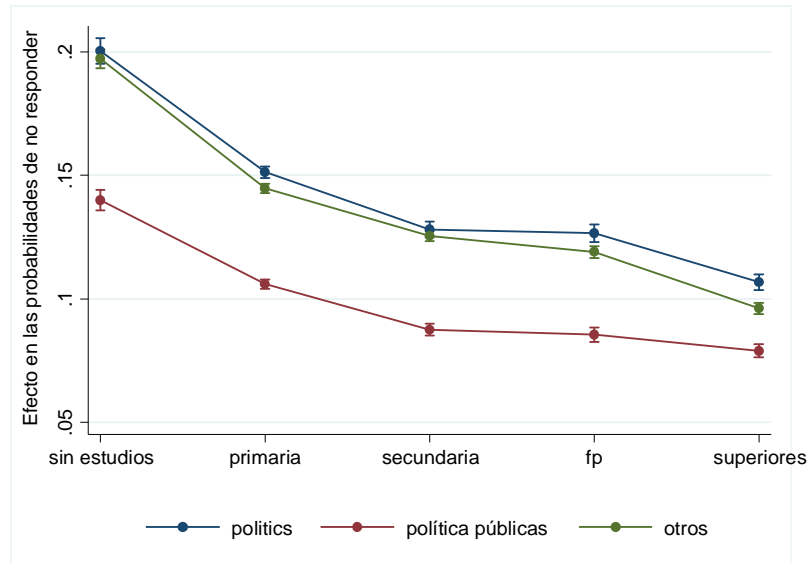
Estrategias de ayuda a la respuesta, como enseñar tarjetas o incluir alientos a la respuesta parecen tener un efecto similar entre todos los grupos educativos, algo menor entre los que tienen estudios superiores –que parten de tasas de no respuesta más bajas– pero en todo caso no muy diferentes. Se trata, pues, de estrategias que contribuyen a reducir la magnitud del problema de la no respuesta pero no parece que tengan un efecto diferencial que pudiese reducir los sesgos derivados de la mayor propensión a no responder de los encuestados con menores niveles educativos.

Gráfico 5.6. Efecto de aliento a la respuesta por nivel de estudios



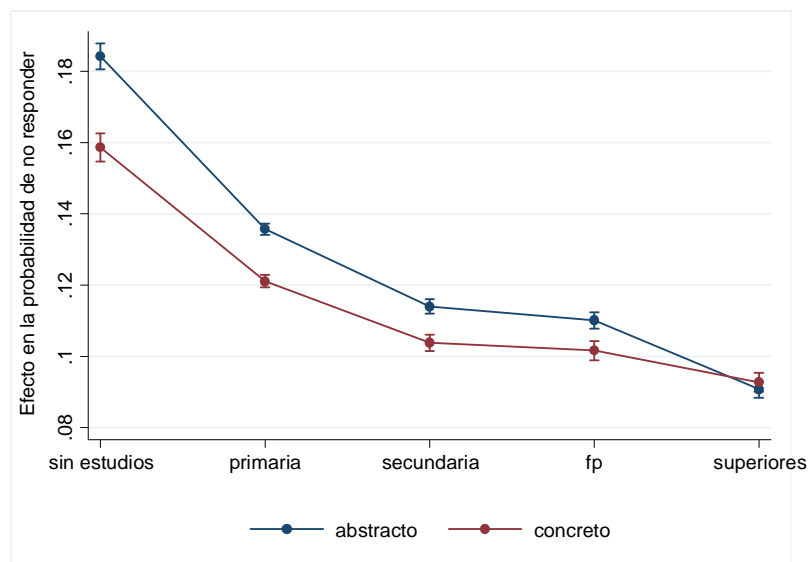
En cambio, los elementos relacionados con el contenido de las preguntas sí tienen efectos diferenciados en la no respuesta en función del nivel educativo. Así, como muestran los gráficos de las interacciones, observamos que el ámbito al que se refiere la pregunta (proceso político vs. políticas públicas) tiene un impacto algo mayor en la probabilidad de no responder de los encuestados con bajos niveles educativos, mientras que los que tienen estudios superiores se ven menos afectados por ello.

Gráfico 5.7. Efecto de ámbito por nivel de estudios



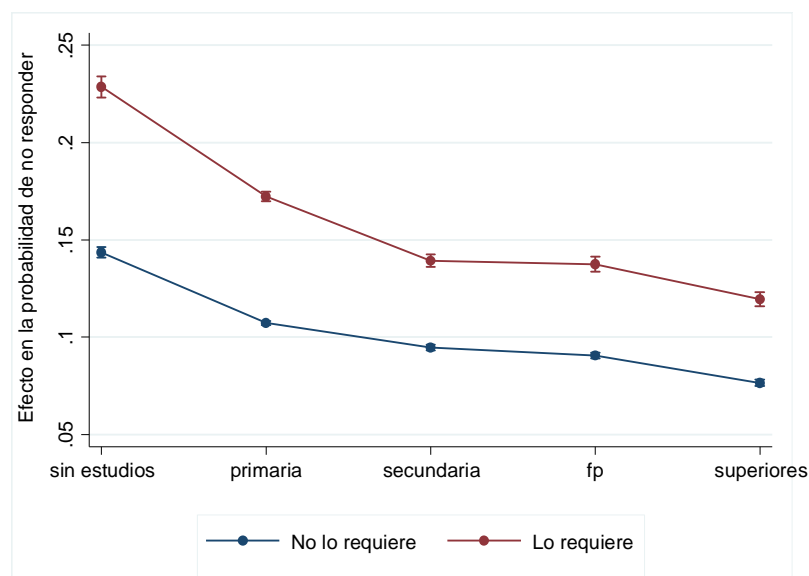
De igual modo, el hecho de que la pregunta se refiera a un objeto abstracto o concreto parece ser relevante para los encuestados con bajos niveles educativos, que sí tienen una probabilidad significativamente mayor de no responder cuando se trata de objetos abstractos. Por el contrario, la naturaleza del objeto no tiene efectos significativos sobre la probabilidad de responder o no de los titulados superiores.

Gráfico 5.8. Efecto de la naturaleza del objeto por nivel de estudios



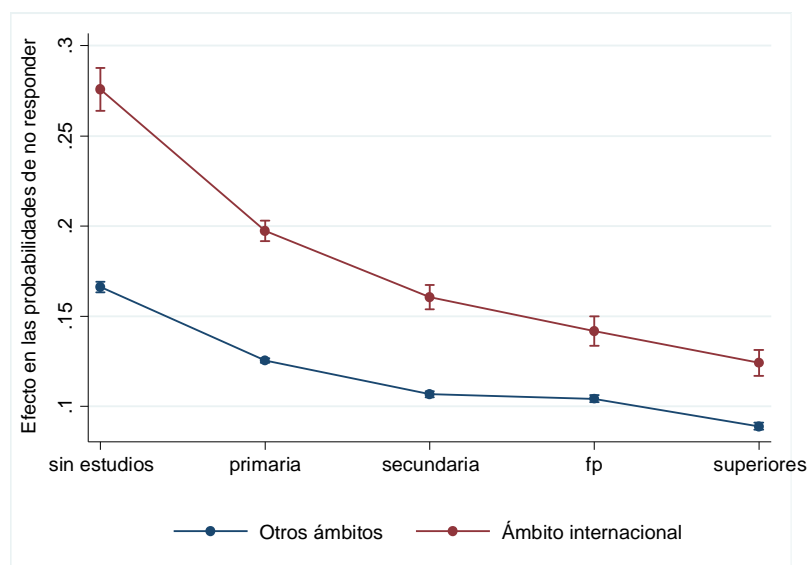
El hecho que se requiera un cierto conocimiento específico para poder contestar adecuadamente la pregunta es otro de los elementos que, según nuestro análisis, refuerza la no respuesta. Y lo hace para todos los grupos educativos, aunque con algo más de intensidad para aquellos entrevistados que no tienen estudios o que tienen solamente la educación primaria completada, y menos para los que tienen educación secundaria o más.

Gráfico 5.9. Necesidad de un conocimiento específico por nivel de estudios



Finalmente, hemos visto que las preguntas sobre cuestiones políticas del ámbito internacional recogían porcentajes significativamente más elevados de no respuesta. Este efecto sí está claramente condicionado por los recursos cognitivos de los encuestados: las tasas de no respuesta a las preguntas sobre cuestiones internacionales son mucho más elevadas, con respecto al resto, entre los ciudadanos con menores niveles de estudios. La diferencia es muy marcada para los que no tienen estudios y para los que solamente tienen estudios primarios, y mucho menos pronunciada para los que tienen estudios superiores.

Gráfico 5.10. Ámbito internacional por nivel de estudios



El hecho de que se requiera un cierto conocimiento específico para poder contestar adecuadamente la pregunta es otro de los elementos que, según nuestro análisis, refuerza la no respuesta. Y lo hace para todos los grupos educativos, aunque con algo más de intensidad para aquellos entrevistados que no tienen estudios o que tienen solamente la educación primaria completada y menos para los que tienen educación secundaria o más.

Si nos fijamos en los efectos condicionales por sexo del contenido de las preguntas, observamos que estos se concentran fundamentalmente en la cuestión del ámbito internacional vs. otros: si bien las mujeres tienen siempre mayor propensión que los hombres a la no respuesta, las diferencias son mucho más marcadas cuando se trata de cuestiones de política internacional. Por el contrario, el hecho que la pregunta se refiera a las políticas públicas vs. el proceso político no parece, en contra de nuestras expectativas, mitigar el 'gender gap' en la probabilidad de responder: las diferencias son relativamente constantes.

Gráfico 5.11. Ámbito político por sexo

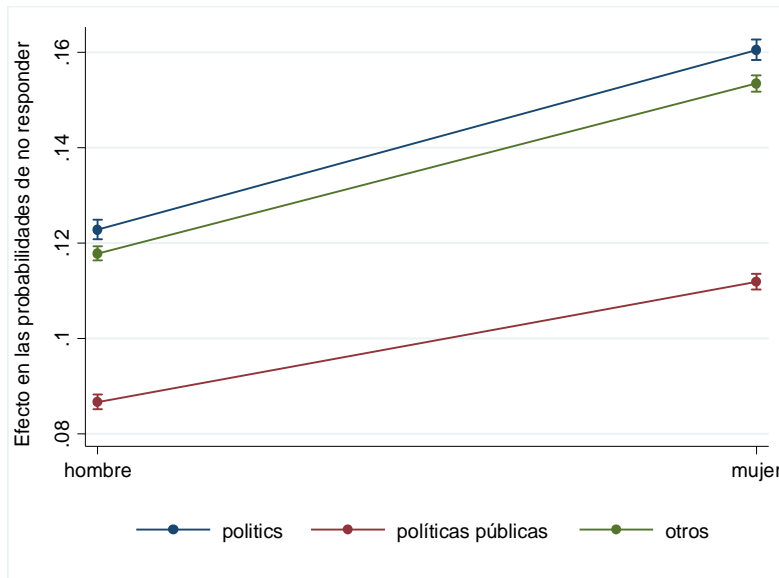
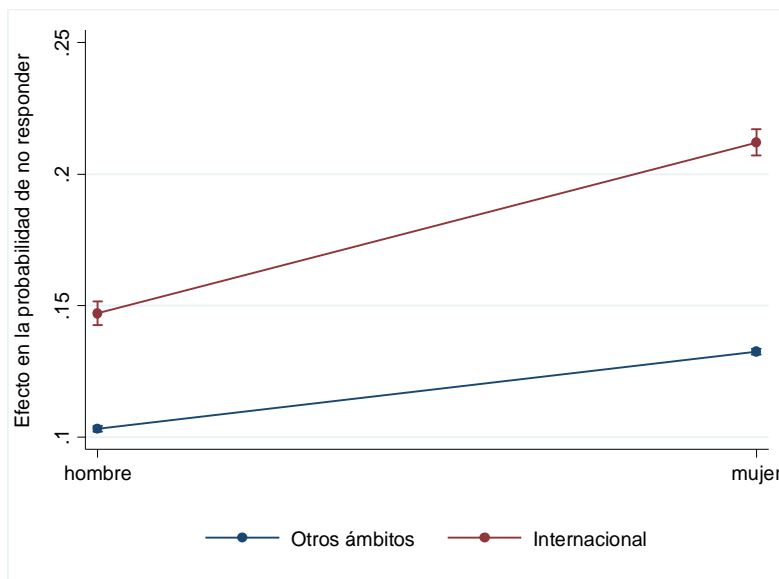


Gráfico 5.12. Ámbito internacional por sexo



En definitiva, hemos visto cómo determinadas características individuales condicionan la probabilidad de no respuesta a las preguntas con contenido político de las encuestas. Los factores, que son diferentes para explicar el no sabe y el no contesta, se refieren fundamentalmente a características individuales relacionadas con la disponibilidad de

recursos cognitivos, o con la socialización política de los ciudadanos. También hemos visto que estos efectos de las características individuales son complementarios a los efectos del diseño del cuestionario y, aún más específicamente, que en algunos casos ambas interactúan entre ellas. Esto es especialmente relevante, puesto que en la medida en que determinados elementos formales del planteamiento de las preguntas tienen efectos especialmente fuertes para los grupos que, *per se*, tienen tasas mayores de no respuesta, esto puede contribuir a reducir o minimizar los sesgos que la no respuesta introduce en las encuestas.

6. Conclusiones

Este proyecto ha tratado de explicar la variación de la no respuesta en las preguntas recogidas en los barómetros del CIS realizados en los diez últimos años. Para ello se han construido dos bases de datos. La primera recoge cuáles son las características de las preguntas y del cuestionario que afectan al nivel agregado de no respuesta. La segunda añade las características individuales de los entrevistados que participaron en dichos barómetros. La intención de esta segunda base de datos es identificar cómo interactúan los factores individuales y de pregunta/cuestionario para predecir la probabilidad de emitir una no respuesta.

Por lo que se refiere a las características de las preguntas y del cuestionario se han identificado diversos parámetros que afectan a la cantidad de individuos que se acogen a las categorías “no sabe” y “no contesta”. Entre los factores explicativos hemos distinguido entre aquellos que tienen que ver con la estructura de la pregunta y aquellos que tienen que ver con el contenido de la misma. Entre los primeros, hemos encontrado que tienen un efecto sobre la cantidad de no respuestas la posición de la pregunta en el cuestionario, el hecho de formar parte de una batería, el hecho de que las opciones de respuesta tengan autonomía semántica, el hecho de que la escala sea unipolar o bipolar, o el hecho de ofrecer una categoría central de refugio a los entrevistados. Todos estos factores afectan a la cantidad agregada de no respuestas.

Por lo que se refiere a las características relacionadas con el contenido de las preguntas hemos encontrado que las preguntas que se refieren a políticas públicas (más específicas) tienen una menor cantidad de no respuesta que las que se refieren a la política en abstracto. También aumenta la no respuesta el hecho de que la pregunta tenga una referencia temporal específica en su enunciado, el que se necesite un cierto conocimiento

específico para contestarla, o el hecho de que la pregunta se refiera al ámbito internacional. En cualquier caso, cabe destacar que aunque se producen algunas diferencias en el efecto de estas variables cuando se distingue entre las opciones “no sabe” y “no contesta”, la más importante de las diferencias tiene que ver con la capacidad explicativa de nuestros modelos. Es decir, la estructura y el contenido de las preguntas dan especial cuenta de la opción “no contesta”. Es esta categoría la que presenta mayor cantidad de variación entre preguntas, lo que permite una mayor capacidad explicativa del modelo.

En la segunda parte de esta investigación hemos introducido las características de los individuos en los modelos para analizar en qué medida los recursos individuales (cognitivos y otros), su contexto y socialización (sexo y edad), o su perfil político afectan a la probabilidad de no respuesta, y en qué medida estos factores interactúan con los factores de las preguntas. Confirmando mucha de la literatura anterior las variables funcionan en la dirección esperada (con efectos especialmente fuertes para el nivel educativo), pero los efectos de estos factores funcionan principalmente para explicar la probabilidad de que un individuo escoja el “no sabe” como opción de respuesta. En cambio, estas mismas características no suelen tener efecto sobre el “no contesta”, confirmando una de las principales conclusiones de esta investigación: mientras que la categoría “no contesta” tiene que ver fundamentalmente con la estructura y el contenido de las preguntas y el cuestionario, el “no sabe” está más relacionado con las características de los individuos.

Finalmente, la última cuestión que hemos dirimido, y una de las principales aportaciones de esta investigación tiene que ver con la interacción entre las características de las preguntas y las características de los individuos. Es decir, cuáles son los elementos de las preguntas y de los cuestionarios que aumentan la no respuesta en función de las características de los individuos. Una de las características que muestra efectos condicionales más importantes es el nivel educativo. Así, enseñar una tarjeta al

entrevistado, alentar una respuesta por parte del entrevistador, necesitar de conocimientos específicos para responder a una pregunta, o interrogar sobre cuestiones de política internacional tiene efectos decrecientes en función del nivel de estudios. Es decir, estos factores tienen un efecto mayor (en el sentido de disminuir la no respuesta) entre los individuos que tienen menos estudios, efecto que se va haciendo más pequeño a medida que aumentan los años de formación educativa.

Las diferencias son más exageradas en virtud de la posición de la pregunta en el cuestionario: el efecto cansancio afecta mucho menos al grupo de los que tienen estudios superiores que al resto de niveles educativos. Paradójicamente, la longitud de la pregunta no muestra la misma tendencia y no aparecen signos de grandes diferencias entre niveles de educación formal. Finalmente, destacar que sólo encontramos efectos interactivos entre las características de las preguntas y el sexo para las preguntas referidas a las políticas de ámbito internacional. Es decir, mientras que las mujeres son mucho más susceptibles que los hombres a no responder a las preguntas de política internacional, no encontramos efectos interactivos similares para la distinción entre las preguntas sobre el proceso político y las preguntas sobre políticas públicas.

Bibliografía

Aneshensel, C. S., R. R. Frerichs, V. A. Clark, y P. A. Yokopenic. 1982. "Measuring Depression in the Community: a Comparison of Telephone and Personal Interviews." *Public Opinion Quarterly* 46 (1): 110–121.

Bell, Ralph. 1984. "Item Nonresponse in Telephone Surveys: An Analysis of Who Fails to Report Income." *Social Science Quarterly* 65 (1): 207–15.

Berinsky, Adam J. 2004. *Silent Voices: Public Opinion and Political Participation in America*. Princeton University Press: Princeton.

Billiet, Jacques, y Geert Loosveldt. 1988. "Improvement of the Quality of Responses to Factual Survey Questions by Interviewer Training." *Public Opinion Quarterly* 52 (2): 190–211.

Bishop, George F., Alfred J. Tuchfarber, y Robert W. Oldendick. 1986. "Opinions on Fictitious Issues: The Pressure to Answer Survey Questions." *Public Opinion Quarterly* 50 (2): 240–250.

Coombs, C. H. y L. C. Coombs. 1976-1977. "'Don't' Know': Item Ambiguity or Respondent Uncertainty?" *Public Opinion Quarterly* 40, 497-514.

Converse, Jean M. y Stanley Presser. 1986. *Survey Questions: Handcrafting the Standardized Questionnaire*. Sage: Thousand-Oaks.

DeMaio, T. J. 1980. "Refusals: Who, Where and Why?" *Public Opinion Quarterly* 44: 223–233.

Dickinson, John R., y Eric Kirzner. 1985. "Questionnaire Item Omission as a Function of Within-group Question Position." *Journal of Business Research* 13 (1): 71–75.

Easton, David. 1965. *A Systems Analysis of Political Life*. New York: Wiley.

Feick, L. F. 1989. "Latent Class Analysis of Survey Question that Include don't Know Responses." *Public Opinion Quarterly* 53: 525-547.

Ferber, R. 1966. "Item nonresponse in a Consumer Survey." *Public Opinion Quarterly* 30: 399–415.

Fowler Jr., Floyd J. 2008. *Improving Survey Questions Design and Evaluation*. 4th edition. Sage: Thousand-Oaks.

Francis, J. D. y L. Busch. 1975. "What We Now Know about 'I don't Knows'." *Public Opinion Quarterly* 39: 207-218.

Gilljam, M. y D. Granberg. 1993. "Should We Take Don't Know for an Answer?" *Public Opinion Quarterly* 57: 348-57.

Krosnick, Jon A., y Michael A. Milburn. 1990. "Psychological determinants of political opinionation." *Social Cognition* 8(1): 49-72.

Messmer, D. J. y D. T. Seymour. 1982. "The Effects of Branching on Item Nonresponse." *Public Opinion Quarterly* 46: 270-277.

Payne, S. L. 1949-1950. "Case Study in Question Complexity." *Public Opinion Quarterly* 13: 653-658.

Rapoport, R. B. 1982. "Sex Differences in Attitude Expression: A Generational Explanation." *Public Opinion Quarterly* 46: 86-96.

Rugg, Donald. 1941. "Experiments in Wording Questions: II." *Public Opinion Quarterly* 5: 91-92.

Schuman, Howard y Jacqueline Scott. 1987. "Problems in the Use of Survey Questions to Measure Public Opinion." *Science* 236(4804): 957-959.

Schuman, Howard y Stanley Presser. 1996. *Questions and Answers in Attitude Surveys*. Nueva York: Academic Press.

Shoemaker, Pamela J., Martin Eichholz y Elizabeth A. Skewes. 2001. "Item Nonresponse: Distinguishing between don't Know and Refuse." *International Journal of Public Opinion Research* 14(2): 193-201.

Singer, E. y L. Kohnke-Aguirre. 1979. "Interviewer Expectation Effects." *Public Opinion Quarterly* 43: 245-260.

Singer, E., M. Frankel R. y M. B. Glassman. 1983. "The Effect of Interviewer Characteristics and Expectations on Response." *Public Opinion Quarterly* 47: 68-83.

Smith, T. W. 1982. "House Effects and the Reproducibility of Survey Measurements: A Comparison of the 1980 GSS and the 1980 American National Election Study." *Public Opinion Quarterly* 46: 54-68.

Sudman, Seymour y Norman M. Bradburn. 1974. *Response Effects in Surveys: A Review and Synthesis*. Chicago: Aldine.

Sudman, Seymour y Norman R. Bradburn. 1982. *Asking Questions*. San Francisco: Jossey-Bass.

Sudman, Seymour, Norman M. Bradburn y Norbert Schwarz. 1996. *Thinking about Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass.

Tourangeau, Roger, Lance J. Rips y Kenneth Rasinski. 2000. *The Psychology of Survey Response*. Cambridge: Cambridge University Press.

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press.