doi:10.5477/cis/reis.137.121

Sesgo de no-respuesta y modelos de superpoblación en encuestas electorales*

Nonresponse Bias and Superpopulation Models in Electoral Polls

José M. Pavía y Beatriz Larraz

Palabras clave

Predicciones electorales • Comportamiento electoral • Recuerdo de voto • Muestreo por conglomerados • Encuestas preelectorales • Simulación de Monte Carlo • Elecciones españolas

Resumen

El sesgo de no-respuesta (y, en menor medida, el error de respuesta) se ha convertido en la principal fuente de error de las predicciones electorales en España. Las técnicas de post-estratificación y los estimadores ratio utilizados actualmente por la industria demoscópica no muestran una capacidad suficiente para corregir los sesgos introducidos durante la recogida de datos. Este trabajo revela cómo un uso más eficiente de la información electoral extramuestral disponible permitiría mejorar sensiblemente la precisión de las estimaciones y muestra, utilizando técnicas de simulación, que ello podría venir acompañado de diseños muestrales más baratos. El estudio, no obstante, concluye que la especificación utilizada en esta investigación no constituye una panacea y señala que existe todavía margen para la corrección del sesgo de no-respuesta, apuntando diversas posibilidades de investigación futura.

Key words

Electoral forecasts • Voting behaviour • Recall vote • Cluster sampling • Pre-election polls • Monte Carlo simulation • Spanish elections

Abstract

Nonresponse bias (and, to a lesser extent, measurement error) has become the main source of error for electoral forecasts in Spain. Although the post-stratification techniques and ratio estimators currently used in the polling industry reduce deviations, they do not show enough capacity to mend the biases introduced when collecting data. This research reveals how a more efficient use of the electoral information available outside the sample could help to significantly improve the accuracy of predictions, and uses simulation techniques to show that this may be accompanied by less expensive sampling designs. The analysis, nevertheless, also concludes that the proposed specification is not a panacea and affirms that there is still scope for reducing nonresponse bias, pointing to several issues for future research.

Introducción[1]

La preocupación por las posibles consecuencias de la no-respuesta[2] en las investigaciones por encuesta es tan antigua como la propia disciplina. De acuerdo con Smith (1999): «Early research extends back to the

* La versión en inglés de este artículo puede consultarse en http://reis.cis.es y http://reis.metapress.com

[1] Agradecimientos: La investigación reportada en este artículo ha sido financiada por el Centro de Investigaciones Sociológicas a través de la concesión en 2009 de una Ayuda para la Investigación Sociológica por el trabajo «Un análisis para la mejora de la calidad predictiva en las estimaciones electorales a partir de datos de encuesta». Los autores desean agradecer a Alberto Penadés y Valentín Martínez la ayuda prestada, a dos evaluadores anónimos y al Consejo Editorial los comen-

José M. Pavía: Universitat de València | Jose.M.Pavia@uv.es
Beatriz Larraz: Universidad de Castilla-La Mancha | Beatriz.Larraz@uclm.es

emergence of polling in the 1930s and has been a regular feature in statistical and social science journals since the 1940s». A pesar de ello, sin embargo, no ha sido hasta décadas más recientes que la comunidad investigadora ha venido dedicando más atención a la cuestión, como consecuencia del deterioro que provoca en la calidad de las estimaciones la creciente falta de voluntad que manifiestan muchos entrevistados a participar en las encuestas (Singer, 2006).

A partir de la segunda mitad de la década de los ochenta del siglo XX se pudo ir constatando el incremento progresivo que las encuestas venían registrando en sus tasas de no-respuesta y fue establecida la falta de cooperación de los encuestados como principal causa de esta tendencia (De Leeuw y De Heer, 2002). Los esfuerzos de los investigadores se concentraron en tratar de entender las causas del fenómeno y en introducir nuevas ideas, como el concepto de aleatoriedad de la no-respuesta, para tratar de reducir la no-respuesta (en sus dos vertientes: no-contactados y rechazos) y/o de corregir sus consecuencias (e.g., Groves y Couper, 1998). Mientras, los científicos sociales y los responsables de las encuestas dirigían su atención a comprender y reducir la no-respuesta (e.g., Groves *et al.*, 1999), desde la perspectiva estadística, la investigación se centraba en tratar de minimizar sus consecuencias a través de técnicas como la imputación múltiple (e.g., Schafer, 1997; King *et al.*, 2001) y los métodos de ajuste y ponderación (e.g., Isaki *et al.*, 2004).

No obstante, una vez fue reconocida la ineficiencia de los costosos métodos utilizados para tratar de incrementar la participación (e.g., Curtin *et al.*, 2005), ser aceptada en la literatura anglosajona (e.g., Keeter *et al.*, 2000; Merkle y Edelman, 2002; Groves, 2006) la falta de una relación estable entre tasa de no-respuesta y sesgo de estimación y comprobar que los métodos estadísticos actualmente en uso no son capaces de corregir suficientemente el sesgo de no-respuesta[3], los principales retos a los que, de acuerdo con Groves *et al.* (2002), se enfrenta la metodología de encuestas en los inicios de este siglo consisten en: i) tratar de determinar bajo qué circunstancias la no-respuesta puede dañar las inferencias poblaciones y en ii) identificar los métodos que, en presencia de no-respuesta, pueden mejorar la calidad de las estimaciones muestrales. El documento que aquí se presenta trata de aportar algunas respuestas, dentro del contexto de las predicciones electorales, al segundo de los retos identificados en Groves *et al.* (2002).

Dentro de las encuestas, los sondeos electorales, y particularmente las predicciones electorales, ocupan un papel primordial, pues como señalan Martin *et al.* (2005), las previsiones electorales, en su dimensión sociológica y de opinión pública, ayudan a formar la imagen del conjunto del sector al poder ser cotejadas con un estándar externo de comparación: los propios resultados electorales. A pesar de ello, un análisis de los métodos actualmente empleados para generar predicciones electorales revela un uso poco eficiente de la información electo-

tarios realizados, y agradecen el soporte del Ministerio de Ciencia e Innovación (MICINN) a través del proyecto CSO2009-11246.

[2] En general en una encuesta se pueden encontrar dos tipos de no-respuesta: total o parcial. La no-respuesta total (o no-participación) se da cuando un individuo no puede ser localizado para ser entrevistado (cuya distribución entre las diferentes opciones se suele suponer que es básicamente al azar) o cuando este, después de ser contactado, declina ser entrevistado. Este segundo tipo de no-respuesta suele presentar distribución asimétrica entre las distintas opciones y es la principal fuente del llamado «sesgo de no-respuesta». La no-respuesta parcial (o al ítem) aparece cuando un sujeto accede a ser entrevistado pero no proporciona respuesta para determinadas preguntas.

[3] Como se acreditó dolorosamente con las predicciones realizadas, a partir de encuestas a salida de urna, en las presidenciales de Estados Unidos de 2000, cuyos errores, que fueron achacados mayoritariamente al sesgo de no-respuesta (e.g., Konner, 2003; Biemer *et al.*, 2003; Randall, 2008), generaron un gran revuelo.

ral extramuestral disponible. En efecto, los datos registrados en convocatorias electorales previas podrían ser explotados más intensamente mediante el uso de modelos de superpoblación (e.g., Valliant *et al*., 2000)[4].

En esa línea, cuando se trabaja con voto escrutado, el uso de modelos de predicción electoral basados en la congruencia que, a nivel de mesa o sección, presentan los resultados electorales entre elecciones sucesivas ha permitido mejorar significativamente la calidad de las predicciones con muestras sesgadas (e.g., Bernardo y Girón, 1992; Bernardo, 1997; Pavía-Miralles, 2005; Pavía *et al*., 2008), incluso aunque solo se disponga de resultados parciales de las mesas escrutadas (Pavía-Miralles y Larraz-Iribas, 2008), y ha mostrado muy recientemente (Pavía, 2010) su potencialidad con datos de encuestas a salida de urna. El objetivo del presente trabajo es estudiar y analizar la capacidad predictiva de esta estrategia cuando se trabaja con datos de encuestas electorales, y compararla con la metodología utilizada actualmente por la industria demoscópica, en concreto, con los métodos de post-estratificación y los estimadores ratio (e.g., Mitofsky y Murray, 2002; Mitofsky, 2003).

En particular, tomando como ejemplo las encuestas post-electorales[5] 2716 y 2720 del CIS (correspondientes, respectivamente, a las elecciones a la Asamblea de Madrid de 2007 y a las elecciones locales al Ayuntamiento de Barcelona de 2007)[6], se ha simulado, a partir de las poblaciones reales asociadas, una enorme cantidad de muestras bajo tres escenarios distintos de comporta-

[4] En un modelo de superpoblación se considera que la población objeto de estudio constituye una realización de una población subyacente mayor (superpoblación), donde las realizaciones individuales de cada miembro de la población presentan unos patrones de regularidad susceptibles de ser modelados estadísticamente y, por tanto, de ser utilizados para mejorar la calidad de las predicciones. Desde esta perspectiva, por ejemplo, si unas mismas elecciones pudieran ser repetidas indefinidamente en unas condiciones similares (emulando, en términos electorales, la película «El día de la marmota»), los resultados que se obtendrían en cada nueva votación serían distintos, aunque presentarían ciertas regularidades que permitirían reconocerlos como pertenecientes a una misma población subyacente. La idea puede ser extendida dinámicamente para buscar relaciones entre los resultados registrados en elecciones celebradas en distintos momentos del tiempo.

[5] El hecho de considerar como base para el análisis encuestas post-electorales, en lugar de pre-electorales, responde a la necesidad de valorar las distintas alternativas de estimación y de diseño muestral con el menor «ruido» posible y, sobre todo, utilizando un valor de comparación externo objetivo. Las respuestas recogidas en las encuestas post-electorales se corresponden con un hecho ya consumado, por lo que, teóricamente, las estimaciones que de ellas se deriven pueden ser comparadas directamente con los resultados realmente registrados en las elecciones. Las respuestas que recogen las encuestas pre-electorales, por el contrario, están sujetas a manifestaciones de intención, las cuales pueden variar entre la fecha de realización de la encuesta y el día de la votación, por lo que una parte de las posibles desviaciones que se observasen entre los verdaderos valores obtenidos en las elecciones y las estimaciones podría ser debida a cambios de estados de opinión entre el momento de la encuesta y el día de la votación y no a cuestiones técnicas, lo que dificultaría enormemente su valoración.

[6] Aunque las elecciones nacionales despiertan, en general, un interés superior entre la opinión pública, en este estudio se trabajó con una elección autonómica y una local, y no con unas generales, por razones logísticas y técnicas. Primero, el trabajar con unas elecciones generales comporta operar con más de 35.000 secciones censales, lo que hubiera implicado unos enormes costes en manejo de información. Segundo, debido al bajo número de diputados que se suelen repartir en la mayoría de las circunscripciones en unas elecciones generales las predicciones son menos sensibles a los sesgos de estimación y presentan, por tanto, un interés técnico inicial menor. Tercero, desde el punto de vista teórico, las dificultades para producir estimaciones precisas crecen inversamente con el tamaño de la población y directamente con el número de partidos; es decir, a mayor población menos costoso es (relativamente) generar estimaciones precisas, y a mayor número de partidos más difícil es acertar. Las elecciones a la Asamblea de Madrid fueron elegidas como ejemplo de una situación con un número efectivo de partidos bajo, pero cuyas estimaciones presentan una sensibilidad muy elevada. [En 2007 se repartieron 111 representantes a la Asamblea de Madrid.] Las elecciones al Ayuntamiento de Barcelona, por su parte, fueron elegidas para valorar el modelo en una situación teóricamente más desfavorable y en un entorno sociopolítico diferente. Desde el punto de vista técnico, el tamaño del electorado es sensiblemente menor y el número efectivo de partidos claramente mayor; mientas que políticamente, además del mayor fraccionamiento que presenta el electorado catalán, en Barcelona el partido mayoritario es el PSC frente al PP en Madrid.

miento de los encuestados al ser entrevistados y dos estrategias diferentes de diseño muestral. Cada una de las muestras ha sido analizada utilizando cuatro estimadores alternativos y la precisión de todas las estimaciones valorada mediante la comparación con los valores reales registrados. A la luz de los resultados, la introducción en el proceso de estimación de los resultados previos registrados en todas las secciones censales permitiría mejorar sensiblemente la precisión de las predicciones.

El resto del documento está estructurado como sigue. En la sección segunda se describen las características de las poblaciones utilizadas y de los criterios seguidos para generar las muestras en el ejercicio de simulación implementado. La sección tercera describe los estimadores utilizados. Las estimaciones obtenidas son analizadas y comparadas en la sección cuarta. En la sección quinta se aplican los estimadores a los datos recogidos en las encuestas 2716 y 2720. Finalmente, el último apartado ofrece un resumen de las conclusiones extraídas junto a una discusión crítica de las mismas. Dos Apéndices completan el documento. El primero de ellos describe cómo han sido generados los recuerdos de voto; mientras que el segundo ofrece todos los detalles de cómo fueron generadas las muestras y las respuestas de los encuestados.

Estrategias de muestreo y generación de muestras

De acuerdo con las notas metodológicas que acompañan a las encuestas electorales realizadas por el Centro de Investigaciones Sociológicas (CIS), el mecanismo seguido en la institución para la selección de unidades muestrales se define como un procedimiento polietápico estratificado por conglomerados «con selección de las unidades primarias de muestreo (municipios) y de las unidades secundarias (secciones) de forma aleatoria proporcional, y de las unidades últimas (individuos) por rutas aleatorias y cuotas de sexo y edad».

En esta investigación se ha tomado como base el procedimiento del CIS —en la forma descrita en Rodríguez Osuna (2005)— y han sido ensayadas dos estrategias diferentes de diseño muestral. Por una parte, se ha seguido un enfoque semejante al empleado en el CIS (que denotaremos CIS; véase Rodríguez Osuna, 1991 y 2005); mientras, por otra, se han simulado encuestas con un diseño muestral sensiblemente más sencillo y barato (que notaremos AL). En esta segunda estrategia se ha seleccionado aleatoriamente[7] un número más reducido de secciones que en las encuestas del CIS y se ha encuestado en cada sección seleccionada a una cantidad relativamente más grande de ciudadanos. Con ello, se ha pretendido estudiar el impacto que en la calidad de las estimaciones tendría utilizar un diseño muestral teóricamente más barato, por una necesidad de menores desplazamientos y encuestadores.

El muestreo por conglomerados (con las secciones censales como conglomerados) constituye, por tanto, la base de la técnica muestral utilizada en las encuestas CIS y AL analizadas en este trabajo[8]; por lo que disponer de las distribuciones de voto a nivel de sección censal (es decir, los votos obtenidos por cada partido en cada sección) es imprescindible para poder llevar a cabo el estudio de simulación planteado. Para poder obtener

[7] Mediante muestreo aleatorio sin reemplazamiento, con probabilidad de selección proporcional al tamaño de la sección y sin estratificación previa de la población.

[8] Este comentario podría llevar a la errónea conclusión de que la metodología propuesta en este documento no podría ser utilizada en encuestas telefónicas. Nada más lejos de la realidad. La metodología de muestreo propuesta con los diseños CIS y AL podría ser perfectamente adaptada para ser implementada mediante una encuesta telefónica, aceptando, no obstante, los problemas de cobertura que introducen las encuestas telefónicas. Los sesgos de cobertura, sin embargo, no plantearían ningún problema adicional pues podrían ser contemplados como no-respuesta.

las predicciones, sin embargo, no basta únicamente con los datos correspondientes a las elecciones cuyos resultados se pretenden estimar, sino que es necesario además: i) disponer de los resultados registrados en las elecciones previas en cada sección y ii) asociar a cada votante un recuerdo de voto[9].

Ambos conjuntos de resultados (los votos registrados en cada sección censal en las elecciones previas y actuales) son imprescindibles para poder generar las muestras: el recuerdo de voto es fundamental en los estimadores ratio y de post-estratificación empleados actualmente, y también en la metodología basada en modelos de superpoblación que se propone como alternativa en este trabajo.

Esta necesidad, sin embargo, plantea dos cuestiones previas que deben ser resueltas. Por una parte, es preciso establecer el mapa de relaciones entre las secciones censales de las elecciones actuales y previas; mientras, por otra, es necesario asignar un recuerdo de voto a cada encuestado potencial. Para resolver la primera de las cuestiones se ha utilizado la solución clásica, que consiste en comparar los códigos geográfico-administrativos y el número de electores asignados a las secciones como mecanismo que permite, mediante la aplicación de un conjunto de reglas (Pavía-Miralles, 2005: 1117-1118)[10], establecer la relación entre unidades de votación[11].

Tras haber sido establecidas las relaciones entre secciones, disponemos de las distribuciones marginales de voto en cada sección: los votos registrados en cada sección en las elecciones previas y actuales. Sin embargo, sería preciso disponer además del recuerdo de voto de cada votante, es decir, de las distribuciones cruzadas. El voto, sin embargo, es secreto, por lo que no es posible conocer con exactitud el comportamiento electoral presente o pasado a nivel individual. Para asignar un recuerdo de voto a cada sujeto, no obstante, se podría tratar de explotar el comportamiento electoral agregado registrado por los electores entre las elecciones sucesivas para, de ahí, inferir en cada sección los comportamientos electorales de cada grupo de electores. Esta cuestión, conocida como «the ecological inference problem» (King, 1997), es resuelta en este trabajo utilizando, en un proceso bietápico, técnicas de balanceado de matrices[12] (Pavía *et al.*, 2009), explotando la valiosa información que sobre transferencia de voto ofrecen las encuestas electorales 2716 y 2720. La idea básica consiste en calcular una estima-

[9] En el presente estudio nos hemos autorrestringido, a fin de simplificar los análisis, a utilizar únicamente los resultados previos registrados en la elección del mismo tipo celebrada en la circunscripción.

[10] En particular, la lista básica de reglas utilizada para rastrear y establecer las relaciones entre las secciones censales de las elecciones sucesivas ha sido: i) asumir una relación directa entre las secciones que aparentemente no han cambiado (implícitamente se asume la hipótesis de que el relativamente bajo número de entradas y salidas que registra la sección es aleatorio); ii) cuando dos (o más) secciones se combinan para crear una nueva sección, los resultados agregados de las unidades originales son asignados como valores históricos a las nuevas secciones; iii) a todas aquellas secciones que surgen de la división de una sección previamente existente se les asignan las proporciones de votos de la sección de origen como valores previos; y, iv) a las secciones de nueva creación (las cuales suelen estar localizadas en las áreas de expansión de las ciudades) se les asigna, como valores históricos, las proporciones correspondientes al distrito, la ciudad o, incluso, la circunscripción en la cual se encuentran.

[11] Actualmente, no obstante, con la creciente disponibilidad de información geográfica, la tarea del establecimiento de las correspondencias podría ser fácilmente automatizada, y mejorada, explotando la correlación espacial que presentan los resultados electorales, y extendida a situaciones donde la asignación manual es impracticable (Pavía y López-Quilez, 2012).

[12] En concreto se ha utilizado el método RAS, un procedimiento de aplicación mecánica muy respetuoso con la matriz inicial que recibe cobertura teórica a través de la teoría de la información y la entropía, ampliamente utilizadas en ciencia política para inferir comportamientos individuales a partir de valores agregados (e.g., Johnston y Pattie, 2000; Gschwend *et al.*, 2003).

ción de transferencia de votos entre opciones electorales a nivel de sección censal, de forma que dependiendo del grupo al que pertenezca el sujeto (opción política) se conoce su voto actual y su distribución de voto previo condicionada a su voto presente.

Tras aplicar el proceso anterior —cuyos detalles técnicos se ofrecen, mediante un ejemplo, en el Apéndice 1— se dispone para cada elector de cada sección censal de información sobre su comportamiento electoral actual y previo. Estas son las poblaciones que han sido utilizadas para simular las muestras[13].

De cada una de las poblaciones anteriores fueron extraídas 6.000 muestras de tamaño 1.000 (los mismos tamaños proyectados en las encuestas 2716 y 2720 tomadas como referentes). En concreto, mil muestras por cada uno de los seis escenarios que se obtienen al combinar las dos estrategias de diseño muestral propuestas (CIS y AL) con las tres hipótesis consideradas de comportamiento de los electores al ser entrevistados (sin error, SE, con sesgo de no-respuesta, SNR, y con sesgo de no-respuesta y error de respuesta, ER; véase el Apéndice 2).

El proceso de obtención de las muestras constó de tres etapas. En la primera etapa se seleccionaron las secciones censales. En la segunda etapa se seleccionó en cada sección censal el número de encuestados que correspondía de acuerdo al diseño muestral utilizado[14]. Y, en la tercera etapa, se recogió la respuesta de los individuos seleccionados en la segunda etapa. Los detalles de cómo se generaron las muestras y las respuestas de los individuos se ofrecen en el Apéndice 2.

Empero la información detallada ofrecida en el Apéndice 2, a fin de ilustrar cómo se generaron las muestras a continuación se explica, a modo de ejemplo, la mecánica que se siguió para obtener las respuestas en una muestra AL con sesgo de no-respuesta y error de respuesta de las elecciones locales de Barcelona. En primer lugar, entre las 1.482 secciones censales en que estaba dividido el electorado, se seleccionaron al azar 25 secciones (con probabilidad de selección proporcional al tamaño de la sección); para, a continuación, en cada sección seleccionada proceder como sigue. Entre todos los electores de la sección se elegía al azar un elector y se observaba su voto: supongamos que había votado por CiU. Entonces, para simular el sesgo de no-respuesta, se «lanzaba al aire una moneda» con probabilidad de sacar cara de alrededor del 20% (al ser votante de CiU)[15]. Si salía «cara» el elector era descartado y se procedía a elegir otro elector entre los ciudadanos no seleccionados todavía. Si salía «cruz» el elector era seleccionado para formar parte de la muestra. Este proceso se repetía hasta conseguir elegir 40 electores de la sección. En este punto, para simular el error de respuesta, a los 40 electores seleccionados se les sometía a una nueva prueba. En concreto, para cada elector se «lanzaba al aire una nueva moneda», con probabilidad de cara del 5%, de forma que si salía cruz sus datos poblacionales, sobre voto actual y recuerdo de voto, pasaban a formar parte directamente de la muestra; mientras que, si

[13] Nótese que en cada sección censal las distribuciones de votos registradas de estas poblaciones coinciden con las reales y aunque las distribuciones cruzadas son desconocidas, las estimadas logradas deberían estar próximas a las reales. En todo caso, todos los métodos de estimación analizados compiten, para ver cuál genera predicciones más próximas a los resultados reales, sobre las mismas poblaciones.

[14] En todas las muestras simuladas para la selección de los individuos han sido considerados únicamente los electores residentes. De hecho, para aproximarse lo máximo posible a la realidad, en la que los no residentes obviamente no pueden ser entrevistados, no se ha tenido en cuenta (ni para seleccionar, ni en ajustes, ni en comparaciones) los votos de los electores no residentes.

[15] Las probabilidades de no-respuesta de cada opción (que en el caso de los votantes de CiU se establecieron entre un 15 y un 25%) se determinaron a partir de la información contenida en las encuestas que sirvieron de referencia; en este caso la encuesta 2720.

salía cara, se le generaba al azar un recuerdo de voto y un voto actual.

Estimadores

Para cada una de las muestras simuladas se obtuvieron cuatro estimaciones alternativas. En primer lugar, y como referente de comparación, se obtuvo una estimación directa (DIR) resultado de trasladar a porcentajes las respuestas directas de voto de los encuestados. Este estimador permite valorar las «mejoras» que se derivan de incorporar el recuerdo de voto en el proceso de predicción. Además de este sencillo estimador, fueron utilizados otros tres estimadores. Todos ellos emplean el recuerdo de voto como información auxiliar para reducir los sesgos (muestrales o no muestrales) en la predicción y tratan de hacer un uso más eficiente de la información disponible. En concreto, un estimador de post-estratificación (PS), con «corrección a nivel individual» (Mitofsky y Murray, 2002), en la variante utilizada habitualmente en España; un estimador ratio-ponderado (RAT), con corrección a nivel circunscripción[16]; y, el estimador (HD) propuesto como alternativa en este trabajo, con correcciones a nivel de sección, en la variante expuesta en Pavía (2010). En todos los casos, los estimadores fueron construidos con el objetivo de obtener estimaciones de los porcentajes de voto válido que alcanzaría cada una de las formaciones que concurren a las elecciones. Los detalles para el cálculo de cada uno de estos estimadores se ofrecen a continuación.

El estimador directo de p_j, la proporción de votos válidos obtenidos por el partido j-ésimo, se obtiene de forma sencilla como cociente entre el número de votos que recibía en la encuesta la formación j-ésima y el número de encuestados que manifiestan su intención de votar. Es decir, la predicción directa (DIR) para p_j vendría dada por la ecuación (1).

$$\hat{p}_{\mathrm{DIR}_j} = \frac{v_j}{v} \qquad (j = 1, 2, \ldots r), \qquad (1)$$

donde v_j representa el número de electores que manifiestan en la encuesta que votarán por la opción j, v ($= \sum_j v_j$) aglutina al número total de sujetos que han manifestado que votarán en las elecciones presentes y r es el número de opciones (incluido blancos) que concurren en las elecciones.

A fin de corregir errores de no-respuesta, cobertura, de respuesta, etc., muchos institutos de opinión (incluido el CIS) suelen aplicar a sus predicciones correcciones por recuerdo de voto utilizando los resultados registrados en las elecciones previas. Entre ellas destacan las técnicas de post-estratificación[17], donde tras agrupar las respuestas en función del recuerdo de voto, cada observación es reponderada a fin de garantizar que aplicados los pesos resultantes a las respuestas sobre voto previo las estimaciones que se obtendrían coincidirían con los verdaderos valores registrados[18].

En concreto, dado que las muestras simuladas son autoponderadas, las nuevas ponderaciones vendrían dadas, para los votantes que eligieron la j-ésima opción electo-

[16] Los estimadores de tipo ratio tienen una elevada reputación y son recomendados en muchos libros de muestreo (véase Särndal *et al*., 2003). Un estimador de esta clase fue el utilizado para generar las predicciones en las encuestas a salida de urna de las elecciones presidenciales estadounidenses de 2000 (Mitofsky, 2003).

[17] Según Kalton y Kasprzyk (1986), en presencia de no-respuesta el estimador comúnmente denominado de post-estratificación debería denotarse «*the population weighting adjustment estimator*».

[18] Esta técnica, por tanto, también podría ser contemplada como un caso particular de un procedimiento de calibración donde los resultados previos son utilizados como variable auxiliar (e.g., Särndal, 2007).

ral en la elección previa[19], por la ecuación (2) y, para los electores que no votaron en la convocatoria anterior, por la ecuación (3).

$$\omega_{0j} = \frac{\pi_{j,0}}{\frac{v_{j,0}}{n}}, \qquad (j = 1, 2,\ldots, r) \qquad (2)$$

$$\omega_{0,r+1} = \frac{1 - \sum_{h=1}^{r} \pi_{h,0}}{\frac{n - v_0}{n}}, \qquad (3)$$

donde $v_{j,0}$ representa el número de votantes que manifestaron haber votado por la opción j en las elecciones previas, v_0 (= $\sum_j v_{j,0}$) denota el número total de sujetos que votaron en las elecciones previas, $\pi_{j,0}$ representa la proporción de votos (sobre censo) registrados por la opción j en la elección previa, y n es el tamaño efectivo de la muestra.

Los pesos obtenidos en (2) y (3) son utilizados para elevar los resultados de cada individuo de acuerdo con su recuerdo de voto. En concreto, si $v_{i,j}$ representa el número de electores que elegirán la opción j en las elecciones actuales (donde $v_j = \sum_i v_{i,j}$), habiendo manifestado haber elegido la opción i (con $i = 1, 2,\ldots, r + 1$) en las elecciones previas, se tiene que el número reponderado de votantes que obtendrá la opción j en las actuales elecciones vendrá dado por la ecuación (4) y, de ahí, a través de (5), las predicciones mediante el estimador PS para las proporciones de voto válido[20].

$$\hat{v}_j = \sum_{h=1}^{r+1} \omega_{0i}\, v_{i,j} \qquad (j = 1, 2,\ldots, r) \qquad (4)$$

$$\hat{p}_{PSj} = \frac{\hat{v}_j}{\sum_{h=1}^{r} \hat{v}_h}, \qquad (j = 1, 2,\ldots, r) \qquad (5)$$

Como alternativa al estimador PS (que realiza una corrección teniendo en cuenta los resultados individuales de todos los sujetos que participan en la encuesta), se podría obtener la estimación realizando la corrección a nivel agregado en el espíritu de los estimadores de razón (véase Särndal *et al.,* 2003: 180), donde los estimadores son balanceados para que cumplan la restricción impuesta por la covariable recuerdo de voto. En este caso la predicción es alcanzada a través del uso de sendos estimadores de ratio derivados de los datos de la encuesta y después de aplicar una corrección final para reducir la inconsistencia en las predicciones que introduce la conocida falta de insesgadez del estimador ratio.

En concreto, si $p_{j,0}$ representa la proporción de votos (sobre voto válido) registrados por la opción j en la elección previa y $\hat{p}_{DIR,j,0}$ la proporción directa estimada de votantes en la encuesta que declaran haber votado en la elección previa por la opción j (calculada sobre el total de encuestados que manifestaron haber votado en las elecciones previas), se trabajaría en un proceso en dos etapas. En primer lugar, se obtendría mediante (6) una predicción inicial para la proporción de la opción j, $\tilde{p}_j$, a través del estimador ratio. Para en una segunda etapa obtener el estimador RAT aplicando (7), que repondera todas las estimaciones individuales dadas por los estimadores $\tilde{p}_j$ para que sumen uno. Esta segunda etapa permite resolver el principal inconveniente del estadístico $\tilde{p}_j$, consistente en que conduce, casi con certeza, a estimaciones de suma no unitaria cuando se agregan las predicciones de todas las opciones[21].

[19] Para hacer menos densa la notación se suponen el mismo número de opciones en ambas elecciones.

[20] Al mismo estimador (aunque con diferente error) también se podría haber llegado adoptando un esquema de modelo de superpoblación sobre las probabilidades de cambio de opción política, con estas siendo función únicamente del voto previo (Aybar, 1998).

[21] De acuerdo con Särndal *et al.* (2003: 180), si las estimaciones de los valores de recuerdo de voto son muy sesgadas el sesgo puede ser incluso muy significativo.

$$\tilde{p}_j = \hat{p}_{\mathrm{DIR},j} \frac{p_{j,0}}{\hat{p}_{\mathrm{DIR},j,0}}, \qquad (j = 1, 2,\ldots, r) \quad (6)$$

$$\hat{p}_{\mathrm{RAT},j} = \frac{\tilde{p}_j}{\sum_{h=1}^{r} \tilde{p}_h}, \qquad (j = 1, 2,\ldots, r) \quad (7)$$

Por último, como alternativa a los estimadores anteriores, se ha testado el estimador de superpoblación HD[22], el cual es el auténtico motivador de esta investigación. Estudiar la capacidad del estimador HD para corregir el sesgo de no-respuesta y comparar su *performance* con los estimadores PS y RAT constituye, de hecho, el germen del presente trabajo. El estimador HD hace un uso teóricamente más intensivo de la información disponible, pues además de emplear las respuestas individuales de todos los encuestados, tiene en cuenta en qué sección censal está inscrito cada entrevistado y utilizada, a través de un modelo de superpoblación, la información histórica disponible de todas las secciones censales, y no únicamente de su agregación, como hacen los estimadores PS y RAT.

En esta investigación se ha utilizado el estimador HD en la versión introducida por Pavía (2010), quien propone seguir un procedimiento multietápico con estrategias de estimación diferentes para partidos grandes y pequeños[23]. Primero se estiman, corrigiendo por recuerdo de voto, las proporciones de voto que alcanzaría cada uno de los grandes partidos en cada una de las secciones muestreadas. Segundo, aprovechando la congruencia que a pequeña escala presentan los resultados electorales entre elecciones sucesivas, se predicen los porcentajes de las secciones no muestreadas, regresando sobre los valores previamente registrados en cada sección a partir de las estimaciones de las secciones muestreadas (Pavía-Miralles, 2005). Tercero, se agregan todos los resultados para generar una estimación a nivel de circunscripción para los grandes partidos (se consideraron PP y PSOE como las formaciones grandes en el caso de las elecciones a la Asamblea de Madrid y PSC, CiU y PP como las formaciones grandes en las elecciones al Ayuntamiento de Barcelona). Finalmente, se obtienen las predicciones HD de los partidos pequeños combinando las estimaciones obtenidas para los grandes partidos con las predicciones directas, $\hat{p}_{\mathrm{DIR},j}$,de los partidos pequeños.

En concreto, y adaptado a la situación actual, la mecánica del proceso propuesto en Pavía (2010) sería como sigue:

i) A partir de los datos de la encuesta, se obtiene en cada sección censal *s* muestreada y para cada partido político grande (j = 1, 2,.., G) estimaciones de las proporciones de votos actuales y pasadas.

$$\tilde{p}_{js} = \frac{v_{js}}{v_s}, \quad \tilde{p}_{js,0} = \frac{v_{js,0}}{v_{s,0}} \qquad (8)$$

$$(s = 1, 2,\ldots,N_s \text{ y } j = 1, 2,\ldots, G),$$

donde v_{js} ($v_{js,0}$) denota al número de encuestados que declaran votar por el partido j en la sección s en la elección actual (previa), $v_s = \sum_{j=1}^{r} v_{js}$ ($v_{s,0}$) representa el total de encuestados en la sección que declaran un voto en la sección s en la elección actual (previa) y N_s es el número de secciones muestreadas.

ii) Utilizando los valores reales registrados en la sección en las elecciones previas (y asumiendo que las entradas y salidas en el censo electoral de la sección son aleatorias)

[22] Las siglas HD (acuñadas en Pavía *et al.*, 2008) son el acrónimo de la expresión inglesa *historical data*. El estimador debe su nombre al uso intensivo a que somete a los datos históricos disponibles de pequeñas áreas.

[23] En Pavía (2010) se ofrece una justificación, basada en la ley de los grandes números, de por qué es conveniente seguir un enfoque diferente para partidos grandes y pequeños.

se obtiene una estimación, $\ddot{p}_{js}$, corregida por recuerdo de voto mediante[24]:

$$\ddot{p}_{js} = \tilde{p}_{js} + (p_{js0} - \tilde{p}_{js0}),$$
$$(s = 1, 2,\ldots, N_s \text{ y } j = 1, 2,\ldots, G), \quad (9)$$

donde $p_{js,0}$ denota la proporción de votos registrada en la sección s por el partido j en la elección previa.

iii) Se asume que existe a nivel de sección una relación lineal entre los porcentajes de votos actuales (p_{js}) y previos ($p_{js,0}$) para cada partido, con (para simplificar) perturbaciones normales de media cero y correlaciones constantes entre partidos y nulas entre secciones. Es decir:

$$p_{js} = \alpha_j + \beta_j p_{js,0} + \varepsilon_{js},$$
$$(s = 1, 2,\ldots, \text{N y } j = 1, 2,\ldots, r), \quad 10)$$

siendo ε_{js} perturbaciones normales de media cero verificando $E(\varepsilon_{js},\varepsilon_{j^*s^*}) = \delta_{ss^*}\sigma_{jj^*}$ (siendo δ la delta de Kronecher), α_j y β_j parámetros desconocidos y N el número de secciones de la circunscripción.

iv) Con las predicciones para los grandes partidos a nivel de sección obtenidas en ii) y definiendo mediante $\ddot{p}_{G+1,s} = 1 - \sum_{j=1}^{G} \ddot{p}_{js}$ a la proporción estimada para el resto de partidos en la sección s, se estiman —mediante el algoritmo iterativo propuesto en Pavía-Miralles (2005: 1121)[25]— los parámetros α_j y β_j (para $j = 1, 2,\ldots, G + 1$) de (10), de donde se obtiene, condicionando en los parámetros estimados, predicciones para las proporciones de votos de los principales partidos de todas las secciones, mediante:

$$\hat{p}_{js} = \hat{\alpha}_j + \hat{\beta}_j p_{js,0},$$
$$(s = 1, 2,\ldots, \text{N y } j = 1, 2,\ldots, \text{G}), \quad (11)$$

v) Una vez se dispone de estimaciones en cada sección los resultados son elevados a nivel de circunscripción para obtener mediante (12) las estimaciones HD para los grandes partidos y a través de (13) las estimaciones HD para las formaciones pequeñas:

$$\hat{p}_{\text{HD},j} = \sum_{s=1}^{\text{N}} \omega_s \hat{p}_{js}, \quad (j = 1, 2,\ldots, \text{G}), \quad (12)$$

$$\hat{p}_{\text{HD},j} = \frac{\hat{p}_{\text{DIR},j}\left(1 - \sum_{j=1}^{\text{G}} \hat{p}_{\text{HD},j}\right)}{\sum_{j=1+1}^{r} \hat{p}_{\text{DIR},j}},$$
$$(j = \text{G} + 1,\ldots, r) \quad (13)$$

donde ω_s es el peso de la s-ésima sección en la circunscripción, definido a través del producto de la tasa de participación registrada en la sección en las elecciones previas y el número de electores de la sección:

$$\omega_s = \frac{e_s t_{s,0}}{\sum_{h=1}^{\text{N}} e_h t_{h,0}}, \quad (s = 1, 2,\ldots, \text{N}) \quad (14)$$

con e_s el número de electores de la sección s-ésima y $t_{s,0}$ la tasa de participación registrada en la s-ésima sección en las elecciones previas.

Análisis de las simulaciones

De acuerdo con lo expuesto en las secciones anteriores, utilizando técnicas de

[24] La corrección de los sesgos de no-respuesta a nivel de sección hace el mecanismo más flexible, pues como señala Pavía (2010): «It allows for a different bias mechanism for each polling place and for the magnitude and even the direction of the bias to vary across locations».

[25] Para simplificar el proceso de estimación se admite que los errores en la medición de la variable dependiente son absorbidos por el término de error (véase Greene, 2003: 84). En Pavía-Miralles y Larraz-Iribas (2008) se puede encontrar un algoritmo alternativo que permite negociar el problema cuando los errores de medición en la variable dependiente son considerados de forma explícita.

TABLA 1. *Medidas de error utilizadas para valorar la bondad de ajuste de las estimaciones*

Descripción	Acrónimo	Expresión[a]
Error cuadrático medio *(Mean Square Error)*	MSE	$\frac{1}{r}\sum_{k}(p_k - \hat{p}_{x,k})^2$
Raíz del error cuadrático medio *(Root Mean Square Error)*	RMSE	$\sqrt{\frac{1}{r}\sum_{k}(p_k - \hat{p}_{x,k})^2}$
Error absoluto medio *(Absolute Mean Error)*	AME	$\frac{1}{r}\sum_{k}\lvert p_k - \hat{p}_{x,k}\rvert$
Error absoluto medio relativo *(Relative Mean Error)* o porcentaje medio de error	RME	$\frac{100}{r}\sum_{k}\frac{\lvert p_k - \hat{p}_{x,k}\rvert}{p_k}$
Entropía *(Entropy)*	ENT	$-\sum_{k} p_k \log\left(1 - \left\lvert\frac{p_k - \hat{p}_{x,k}}{100}\right\rvert\right)$

[a] r es el número de opciones políticas para los que se calcula el ajuste conjunto, p_k es el porcentaje real de votos observados para la k-ésima opción, y $\hat{p}_{x,k}$ es la estimación del porcentaje de votos obtenida aplicando el estimador correspondiente, con X=DIR, RAT, PS y HD.

Fuente: Elaboración propia.

Monte Carlo[26] una gran cantidad de muestras han sido simuladas bajo diferentes escenarios con el objetivo de valorar, del modo más neutral posible, la capacidad que presenta cada uno de los estimadores propuestos para corregir la incidencia del sesgo de no-respuesta en las predicciones. Se han generado 1.000 muestras en cada escenario y han sido obtenidos 4 estimadores de cada muestra. Ello supone disponer de 4.000 estimaciones para cada formación política por escenario, o equivalentemente, 24.000 conjuntos de estimaciones para cada elección. Dado el alto número de estimaciones disponibles es preciso recurrir a un resumen estadístico de los resultados para poder extraer conclusiones generales. Para ello se ha seguido un enfoque multilateral, estudiando tanto la precisión de las estimaciones correspondientes a cada partido como el grado de ajuste global entre estimaciones y valores reales. Así, además de calcular los estadísticos resumen habituales (media, mediana, desviación típica, primer y tercer cuartil, mínimo, máximo y coeficientes de variación, asimetría y curtosis) para valorar y comparar las distribuciones de las predicciones de cada partido y de sus correspondientes errores de estimación, también se ha evaluado el grado de ajuste conjunto que presentan estimaciones y valores reales. Para esto último se ha empleado toda la batería de medidas de error recogidas en la tabla 1, donde se detalla las expresiones de los estadísticos de ajuste utilizados. Las comparaciones de las distribuciones en cada escenario, calculadas sobre todas las muestras, de estos estadísticos han servido de base para valorar los distintos estimadores.

A fin de aligerar la notación, en la presentación de los resultados mostrados en tablas y figuras se ha utilizado un sistema de códigos empleando los acrónimos introducidos en las secciones precedentes: DIR, PS, RAT

[26] Los métodos de Monte Carlo abarcan una colección de técnicas ampliamente utilizadas que tratan de dar respuesta a problemas que no admiten una solución analítica mediante el uso de experimentos aleatorios repetidos.

y HD, para señalar el estimador; CIS y AL, para indicar la estrategia de diseño muestral; y, SE, SNR y ER, para aludir a las hipótesis sobre comportamiento de los electores al ser entrevistados.

Asamblea de Madrid

En este subapartado se analizan y presentan los principales resultados de las simulaciones obtenidas para las elecciones a la Asamblea de Madrid 2007. De todos los resultados, los correspondientes a las muestras generadas en los escenarios con sesgo de no-respuesta y error de respuesta deberían ser observados más detalladamente, ya que son los más próximos a las situaciones reales. En las tablas 2 y 3 y el gráfico 1 se ofrece un resumen de diversos aspectos relevantes para analizar las simulaciones. En la tabla 2 se presentan, clasificados por escenario y para cada estimador, los valores medios de las predicciones y sus sesgos asociados. En la tabla 3 se encuentran los valores medios de los indicadores de bondad de ajuste y una comparación de las distribuciones del grado de ajuste conjunto de los estimadores, realizada esta última midiendo el porcentaje de veces que cada estimador genera la mejor solución en términos de entropía. Finalmente, el gráfi-

TABLA 2. *Valores medios[a] de predicciones y errores de las estimaciones de porcentajes de voto. Elecciones a la Asamblea de Madrid 2007*

Escenario	Estimador	Porcentajes				Errores[c]			
		PP	PSOE	IU	RE	PP	PSOE	IU	RE
CIS_SE	DIR	53,43	33,13	8,80	4,65	0,29	–0,21	–0,07	–0,02
	PS	53,12	33,32	8,95	4,61	–0,02	–0,01	0,09	–0,06
	RAT	52,78	33,35	9,03	4,84	–0,36	0,01	0,17	0,17
	HD	53,16	33,15	8,94	4,75	0,02	–0,18	0,08	0,08
AL_SE	DIR	52,86	33,67	8,84	4,62	–0,28	0,34	–0,02	–0,04
	PS	52,89	33,65	8,85	4,61	–0,25	0,32	–0,01	–0,06
	RAT	52,71	33,59	8,88	4,82	–0,43	0,26	0,01	0,16
	HD	53,03	33,46	8,85	4,66	–0,11	0,13	–0,01	–0,01
CIS_SNR	DIR	51,13	34,72	9,30	4,84	–2,01	1,39	0,44	0,18
	PS	51,56	34,51	8,95	4,97	–1,58	1,18	0,09	0,31
	RAT	51,65	34,33	8,81	5,21	–1,49	0,99	–0,05	0,54
	HD	52,02	34,09	9,12	4,77	–1,12	0,76	0,26	0,10
AL_SNR	DIR	50,95	34,86	9,36	4,83	–2,19	1,52	0,49	0,17
	PS	51,54	34,61	8,88	4,97	–1,60	1,28	0,02	0,31
	RAT	51,75	34,43	8,67	5,16	–1,39	1,09	–0,20	0,49
	HD	52,11	34,22	8,99	4,68	–1,02	0,88	0,13	0,01
CIS_ER	DIR	49,99	34,08	10,11	5,83	–3,15	0,74	1,24	1,16
	PS	50,78	34,18	9,33	5,70	–2,35	0,84	0,47	1,04
	RAT	51,99	34,54	8,73	4,74	–1,15	1,20	–0,13	0,08
	HD	52,12	34,34	8,58	4,97	–1,02	1,00	–0,28	0,30
AL_ER	DIR	49,70	34,42	10,13	5,76	–3,44	1,08	1,27	1,09
	PS	50,69	34,41	9,27	5,64	–2,45	1,07	0,41	0,97
	RAT	51,98	34,70	8,61	4,71	–1,16	1,37	–0,25	0,04
	HD	52,07	34,54	8,52	4,87	–1,06	1,21	–0,35	0,20
Elecciones	Resultados[b]	**53,14**	**33,33**	**8,86**	**4,67**	—	—	—	—

[a] Valores medios a partir de 1.000 muestras simuladas.
[b] Porcentajes registrados en voto válido sobre población residente.
[c] Calculado como diferencia entre el porcentaje medio estimado y el real.

Fuente: Elaboración propia.

co 1 muestra, en forma de diagrama de cajas y bigotes y a modo de ejemplo, las distribuciones de las predicciones obtenidas siguiendo el diseño muestral CIS para el conjunto de muestras.

De los dos elementos que definen cada escenario, la disposición de los electores a participar en la encuesta es la cuestión que mayor incidencia tiene en la precisión de las estimaciones. De hecho, el impacto del diseño muestral podría ser clasificado incluso como menor. Ello no quiere decir, sin embargo, que sea nulo. Por término medio las estimaciones logradas con el diseño CIS son más ajustadas que las que se obtienen con la estrategia AL, si bien ambos conjuntos de predicciones presentan niveles bastante comparables. La reducción del número de puntos de muestreo y de la distribución espacial que conlleva la adopción del diseño AL (frente al CIS) se traduce, por una parte, en un ligero aumento en el sesgo de estimación de los porcentajes para los grandes partidos (véase la tabla 2) y, por otra (como era de esperar), en un ligero incremento en la variabilidad de las predicciones. Los cambios, sin embargo, no afectan por igual a todos los estimadores. El estimador HD es el que menos sufre el cambio de estrategia muestral tanto en términos de insesgadez como de variabilidad[27], mejorando su posicionamiento relativo con sus competidores, los estimadores PS y RAT (véase la tabla 3).

El impacto sobre las estimaciones de las hipótesis sobre actitud de los electores ante las encuestas es mucho más evidente. En condiciones ideales (es decir, cuando las muestras son generadas sin error) todos los estimadores producen, como era de esperar, predicciones altamente precisas, con un sesgo medio despreciable. Con carácter general, sin embargo, se puede afirmar que el estimador PS, apoyado en buena medida en sus mejores predicciones para las formaciones pequeñas, es el que registra superiores niveles de ajuste, seguido (con figuras muy similares) por los estimadores RAT y HD y por el estimador DIR; el cual presenta niveles de error claramente mayores. Empero, con diseño CIS, el estimador DIR genera la mejor solución un porcentaje de veces algo superior a lo que lo hacen los estimadores RAT y HD (véase la tabla 3), lo que combinado con lo anterior da idea de su menor robustez (véase el gráfico 1). Es decir, cuando erra, erra de más. Resumiendo, se puede afirmar que incluso en situaciones donde la estimación directa no presenta ninguna desventaja teórica la corrección por recuerdo de voto ayuda a mejorar la calidad de las estimaciones.

El comportamiento de los electores simulado en los escenarios ideales, sin embargo, dista mucho de ser realista. En presencia de no-respuesta asimétrica, las estimaciones DIR presentan un sesgo significativo y aunque el uso de la información auxiliar ofrecida por el recuerdo de voto reduce sensiblemente la magnitud del mismo, este todavía persiste de manera apreciable. De todos los métodos, no obstante, la estrategia HD es la que más reduce el sesgo de las estimaciones y la que presenta mejores figuras en la calidad de las estimaciones. Es más, se podría afirmar que cuanto mayor es el sesgo que presenta la muestra original mejor funciona, en términos relativos, el estimador HD. Así, por ejemplo, si en general el estimador HD es el mejor en algo más de un tercio de las muestras (véase la tabla 3), esta cifra crece hasta casi el 60% cuando se consideran las cien muestras con más error en el estimador directo. El estimador HD sale igualmente reforzado en las comparaciones dos a dos, invirtiendo la tendencia observada con las muestras ideales: mejora su situación respecto al estimador RAT y especialmente avanza en su comparación con el estimador PS. De hecho, la relativa baja variabilidad que presentan las estimaciones PS, unido a sus

[27] De hecho, en presencia de sesgo de no-respuesta el estimador HD ve incluso reducir sus niveles de variabilidad.

TABLA 3. *Resumen de las medidas de bondad de ajuste entre distribuciones conjuntas reales y estimadas. Elecciones a la Asamblea de Madrid 2007*

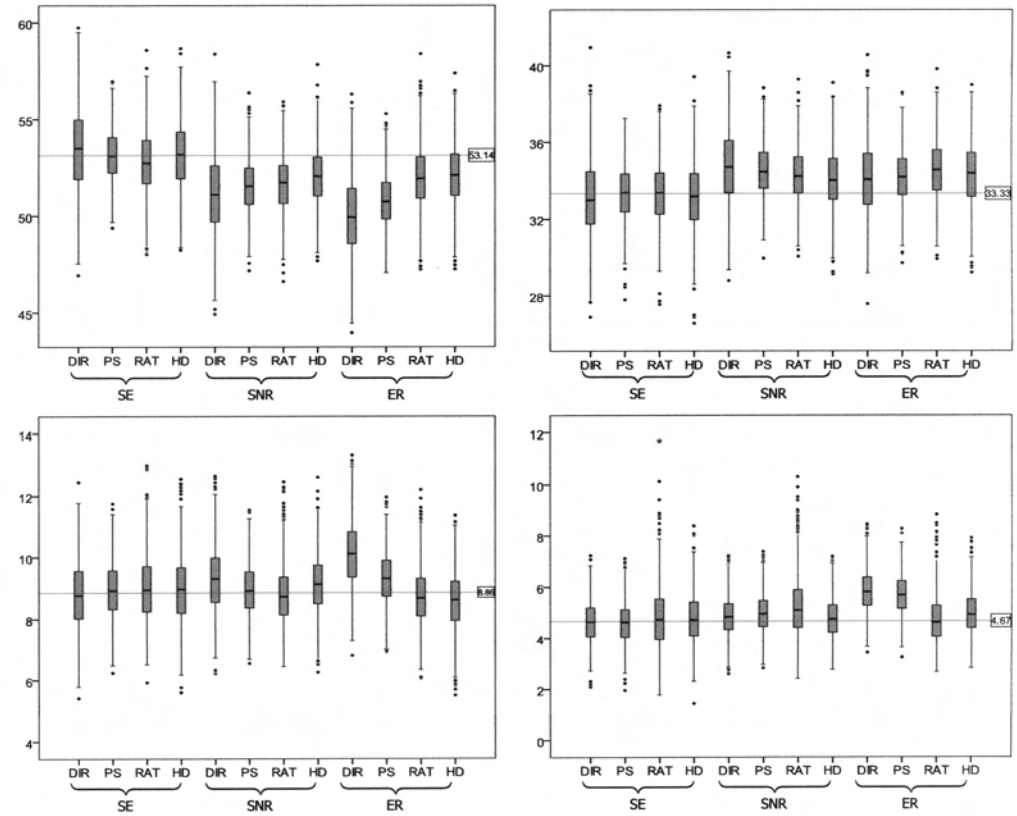
Escenario	Estimador	% éxito sobre 1.000 simulaciones[a]				Medidas de bondad de ajuste[b,c]				
		% éxito	PS *v* HD	PS *v* RAT	RAT *v* HD	ENT	MSE	RMSE	AME	RME
CIS_SE	DIR	24,0	—	—	—	1,62	2,70	1,47	1,24	8,12
	PS	33,0	65,3	64,7	—	1,07	1,35	1,05	0,90	6,81
	RAT	22,0	—	35,3	53,7	1,30	2,08	1,31	1,12	9,20
	HD	21,0	34,7	—	46,3	1,36	2,13	1,32	1,12	8,33
AL_SE	DIR	18,4	—	—	—	2,38	5,62	2,03	1,68	9,31
	PS	35,0	60,7	58,7	—	1,23	1,69	1,16	0,99	6,99
	RAT	23,0	—	41,3	49,2	1,41	2,29	1,37	1,17	9,08
	HD	23,6	39,3	—	50,8	1,41	2,22	1,35	1,15	8,20
CIS_SNR	DIR	25,1	—	—	—	2,05	4,01	1,76	1,46	8,37
	PS	18,8	43,8	51,7	—	1,50	2,20	1,33	1,12	7,20
	RAT	21,2	—	48,3	39,2	1,50	2,48	1,41	1,20	9,01
	HD	34,9	56,2	—	60,8	1,38	2,06	1,28	1,08	7,41
AL_SNR	DIR	22,5	—	—	—	2,68	6,73	2,22	1,83	9,72
	PS	22,1	42,6	50,4	—	1,62	2,50	1,42	1,20	7,56
	RAT	24,8	—	49,6	39,3	1,59	2,74	1,49	1,27	9,35
	HD	30,6	57,4	—	60,7	1,44	2,20	1,33	1,13	7,79
CIS_ER	DIR	16,4	—	—	—	2,51	5,74	2,21	1,87	13,12
	PS	12,1	30,2	29,5	—	1,87	3,19	1,66	1,40	10,25
	RAT	34,0	—	70,5	49,1	1,48	2,35	1,38	1,17	8,16
	HD	37,5	69,8	—	50,9	1,46	2,31	1,36	1,15	7,91
AL_ER	DIR	15,1	—	—	—	3,08	8,68	2,62	2,20	13,93
	PS	13,0	27,4	31,3	—	2,01	3,66	1,75	1,49	10,32
	RAT	35,4	—	68,7	45,0	1,56	2,62	1,45	1,23	8,36
	HD	36,5	72,6	—	55,0	1,51	2,47	1,41	1,20	8,11

[a] Porcentaje de muestras para las que el correspondiente estimador obtiene un mejor ajuste en términos de entropía.
[b] Valores medios a partir de 1.000 muestras simuladas.
[c] ENT: entropía; MSE: error cuadrático medio; RMSE: raíz de MSE; AME: error absoluto medio; RME: error relativo medio.
Fuente: Elaboración propia.

persistentes sesgos, provoca que el estimador PS sea el que presente los menores porcentajes de éxito en estos escenarios.

Las conclusiones con muestras ER son similares a las que se obtienen en los escenarios SNR, aunque en general se observa un cierto deterioro (de escasa cuantía) en los niveles de insesgadez en todos los estimadores. Destaca, no obstante, la convergencia que se produce entre los estimadores HD y RAT, aunque el orden de preferencia (HD, RAT, PS y DIR) obtenido en los casos con sesgo de no-respuesta continúa manteniéndose. En términos de ajuste conjunto el que más sufre de nuevo es el estimador PS (véase la tabla 3) que en la carrera de los cuatro estimadores solo funciona mejor en poco más de un 10% de las ocasiones, mientras que en las comparaciones dos a dos con ambos retrocede hasta situarse en una desventaja de 7 a 3. Los mayores problemas para la estrategia de post-estratificación se sitúan en la predicción de los dos grandes partidos, con los mayores errores concentrados en las estimaciones para PP. Se puede afirmar, por tanto, que el estimador PS es, de los tres estimadores, el que menos corrige los valores brutos y el que presenta predicciones más cercanas al estimador DIR.

GRÁFICO 1. *Representaciones Box-Whiskers del conjunto de las estimaciones obtenidas para PP (figura superior izquierda), PSOE (figura superior derecha), IU (figura inferior izquierda) y Resto (figura inferior derecha) de las simulaciones con diseño muestral CIS*



En cualquier caso, a pesar de que el estimador HD es el que mejor funciona tanto en los escenarios SNR como ER y de que en un gran porcentaje de casos los intervalos de confianza de las predicciones HD hubiesen contenido al verdadero valor[28] y que, por tanto, no existiría desde un punto de vista estadístico error (sino incertidumbre), los resultados medios del estimador HD están todavía a una distancia importante de los verdaderos valores. Por lo que existe todavía margen para la búsqueda de estimadores más precisos.

[28] Tomando como estimación de la varianza muestral a la varianza de las 1.000 simulaciones.

Locales de Barcelona

Las simulaciones obtenidas para las elecciones locales de Barcelona 2007 se analizan en esta subsección. Comparadas con las elecciones a la Asamblea de Madrid cabe destacar que el panorama electoral en Barcelona presenta una estructura multipartidista más evidente y que los tamaños de las poblaciones electorales son sensiblemente inferiores, aspectos ambos que hacen previsiblemente más complejo obtener estimaciones precisas en el caso actual. Las tablas 4 y 5 son respectivamente las homónimas para el caso actual de las tablas 2 y 3 ofrecidas en la subsección previa; mientras que el gráfico 2 ofrece las dis-

TABLA 4. *Valores medios[a] de predicciones y errores de las estimaciones de porcentajes de voto. Elecciones locales de Barcelona 2007*

Escenario	Estimador	Porcentajes						Errores[c]					
		PSC	CiU	PP	ICV	ERC	RE	PSC	CiU	PP	ICV	ERC	RE
CIS_SE	DIR	30,10	25,33	15,58	9,36	8,84	10,79	0,14	–0,13	–0,03	0,01	–0,01	0,02
	PS	30,19	25,40	15,78	9,17	8,87	10,59	0,23	–0,05	0,17	–0,17	0,02	–0,20
	RAT	30,05	25,29	15,68	9,34	8,74	10,90	0,09	–0,16	0,07	–0,01	–0,11	0,12
	HD	30,18	25,04	15,64	9,41	8,88	10,85	0,22	–0,42	0,03	0,06	0,03	0,08
AL_SE	DIR	29,97	25,30	15,64	9,49	8,81	10,79	0,01	–0,16	0,04	0,15	–0,04	0,00
	PS	30,06	25,47	15,82	9,24	8,85	10,56	0,10	0,01	0,22	–0,11	–0,01	–0,21
	RAT	29,93	25,44	15,72	9,36	8,74	10,81	–0,03	–0,02	0,12	0,01	–0,11	0,03
	HD	30,17	25,19	15,68	9,45	8,76	10,75	0,21	–0,26	0,08	0,10	–0,10	–0,03
CIS_SNR	DIR	33,99	26,46	9,20	11,46	9,11	9,78	4,03	1,01	–6,41	2,12	0,26	–1,01
	PS	31,27	25,11	13,01	11,87	8,63	10,11	1,31	–0,35	–2,59	2,52	–0,22	–0,67
	RAT	30,07	24,62	14,14	11,31	8,73	11,13	0,11	–0,83	–1,47	1,96	–0,13	0,36
	HD	30,02	24,84	14,85	11,44	9,08	9,77	0,06	–0,61	–0,75	2,09	0,23	–1,02
AL_SNR	DIR	34,14	26,36	9,16	11,56	9,00	9,78	4,19	0,91	–6,45	2,22	0,15	–1,02
	PS	31,38	25,10	12,93	11,95	8,55	10,09	1,42	–0,36	–2,68	2,61	–0,30	–0,69
	RAT	30,17	24,66	14,05	11,38	8,66	11,08	0,21	–0,79	–1,55	2,03	–0,19	0,29
	HD	30,11	24,83	14,76	11,55	8,98	9,77	0,15	–0,63	–0,85	2,20	0,13	–1,00
CIS_ER	DIR	32,80	25,49	9,45	11,72	9,45	11,09	2,84	0,04	–6,15	2,37	0,60	0,30
	PS	30,99	24,69	12,71	12,00	8,92	10,69	1,03	–0,77	–2,90	2,66	0,07	–0,09
	RAT	30,80	24,99	14,58	11,68	9,19	8,76	0,84	–0,47	–1,02	2,33	0,34	–2,02
	HD	30,18	24,60	14,92	11,01	8,87	10,42	0,22	–0,86	–0,69	1,66	0,02	–0,35
AL_ER	DIR	32,96	25,55	9,44	11,62	9,43	11,00	3,00	0,10	–6,17	2,27	0,58	0,22
	PS	31,05	24,75	12,67	11,96	8,92	10,65	1,09	–0,71	–2,93	2,61	0,07	–0,13
	RAT	30,78	25,03	14,56	11,66	9,21	8,76	0,82	–0,43	–1,04	2,31	0,36	–2,02
	HD	30,22	24,58	14,91	10,98	8,90	10,41	0,26	–0,87	–0,70	1,63	0,05	–0,37
Elecciones	Resultados[b]	**29,96**	**25,46**	**15,61**	**9,35**	**8,85**	**10,77**	—	—	—	—	—	—

[a] Valores medios a partir de 1.000 muestras simuladas.
[b] Porcentajes registrados en voto válido sobre población residente.
[c] Calculado como diferencia entre el porcentaje medio estimado y el real.

Fuente: Elaboración propia.

tribuciones de las estimaciones obtenidas siguiendo, en este caso, el diseño muestral AL.

Las conclusiones obtenidas con las simulaciones para las elecciones de Barcelona vienen a reforzar los resultados obtenidos con los datos de las elecciones de Madrid. A pesar de cambiar a una situación de partida más compleja con una población menor y un panorama político más fraccionado, las tendencias generales y los resultados globales obtenidos previamente se han visto confirmados. En concreto, respecto a las comparaciones entre muestreo CIS y muestreo AL, los resultados para las elecciones de Barcelona vienen a confirmar las conclusiones alcanzadas en las elecciones de Madrid. En general, la selección de un menor número de puntos de muestreo (unido a la realización de una mayor cantidad de entrevistas en cada punto) afecta ligeramente a la precisión de las estimaciones, siendo el estimador HD el que menos (si algo) sufre el cambio de estrategia muestral. En general, por tanto, se podría afirmar que deberían ser criterios económicos (y de precisión en la estimación de otras variables) los que habrían de ser ponderados para decidir, a través de un análisis costes-beneficios, la oportunidad de seguir

TABLA 5. *Resumen de las medidas de bondad de ajuste entre distribuciones conjuntas reales y estimadas. Elecciones locales de Barcelona 2007*

Escenario	Estimador	% éxito sobre 1.000 simulaciones[a]				Medidas de bondad de ajuste[b,c]				
		% éxito	PS *v* HD	PS *v* RAT	RAT *v* HD	ENT	MSE	RMSE	AME	RME
CIS_SE	DIR	20,3	—	—	—	1,53	3,07	1,65	1,35	9,14
	PS	35,3	72,0	66,0	—	1,10	1,69	1,23	1,02	7,19
	RAT	25,0	—	34,0	58,0	1,26	2,39	1,45	1,19	8,56
	HD	19,4	28,0	—	42,0	1,34	2,43	1,47	1,22	8,49
AL_SE	DIR	16,8	—	—	—	1,80	4,14	1,89	1,55	10,20
	PS	35,8	63,9	63,8	—	1,15	1,81	1,27	1,06	7,51
	RAT	24,0	—	36,2	48,1	1,30	2,51	1,49	1,23	8,78
	HD	23,4	36,1	—	51,9	1,29	2,29	1,43	1,19	8,43
CIS_SNR	DIR	0,2	—	—	—	3,18	13,03	3,56	2,77	17,86
	PS	17,4	28,0	28,0	—	1,56	3,82	1,90	1,57	11,73
	RAT	38,1	—	72,0	46,0	1,36	3,15	1,68	1,39	10,56
	HD	44,3	72,0	—	54,0	1,30	2,69	1,57	1,31	10,04
AL_SNR	DIR	0,3	—	—	—	3,43	14,73	3,76	2,95	18,81
	PS	15,2	22,9	27,9	—	1,69	4,29	2,01	1,68	12,37
	RAT	33,8	—	72,1	39,0	1,46	3,49	1,77	1,48	11,13
	HD	50,7	77,1	—	61,0	1,34	2,91	1,63	1,35	10,42
CIS_ER	DIR	1,4	—	—	—	2,71	10,91	3,25	2,50	16,93
	PS	10,9	19,0	33,0	—	1,62	4,23	2,00	1,64	12,18
	RAT	24,2	—	67,0	30,0	1,47	3,70	1,85	1,54	12,07
	HD	63,5	81,0	—	70,0	1,25	2,31	1,45	1,21	8,91
AL_ER	DIR	0,8	—	—	—	3,00	12,29	3,43	2,67	17,54
	PS	13,6	18,3	35,0	—	1,67	4,39	2,03	1,67	12,32
	RAT	19,9	—	65,0	24,2	1,51	3,80	1,87	1,57	12,25
	HD	65,7	81,7	—	75,8	1,25	2,31	1,44	1,20	8,86

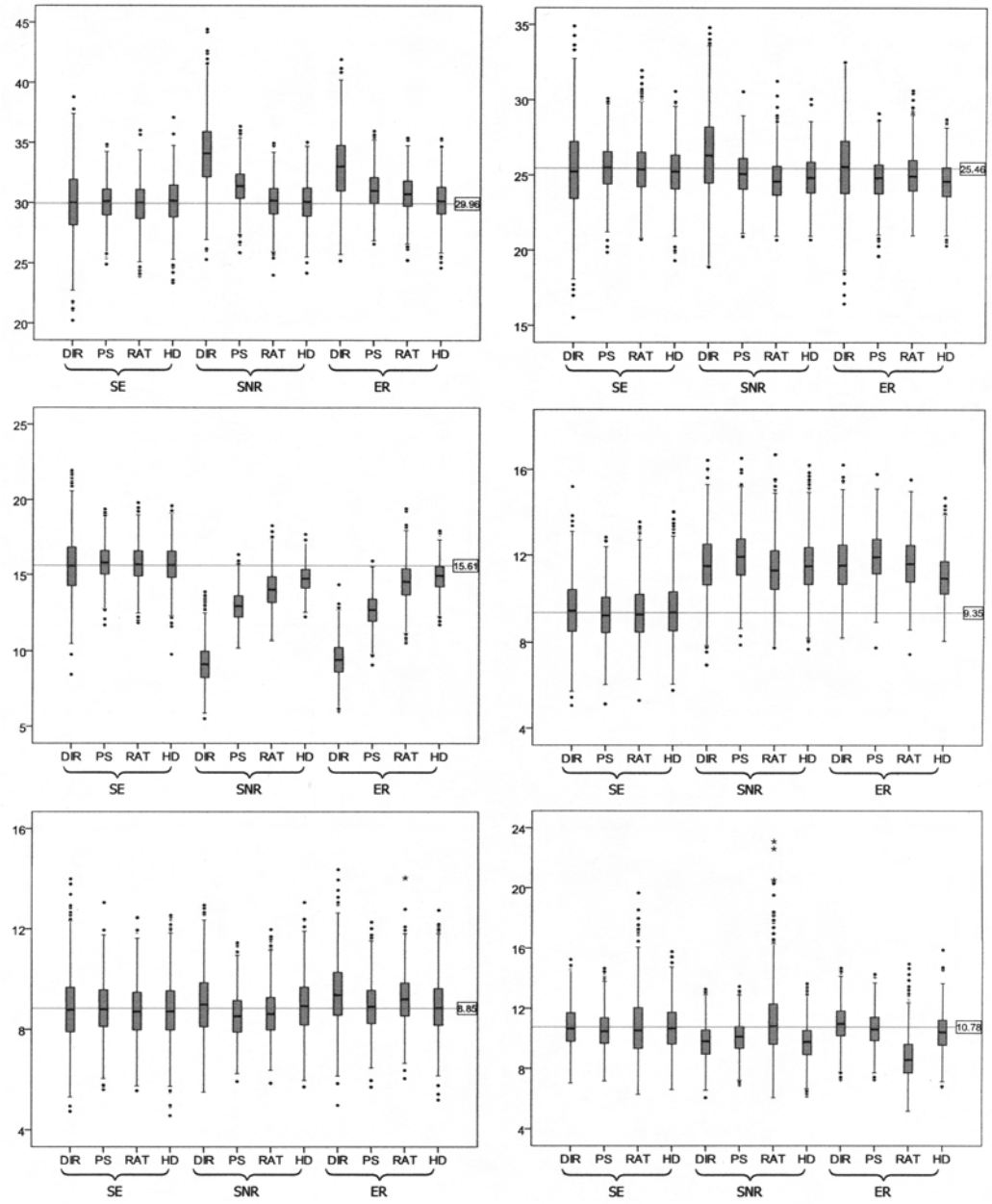
[a] Porcentaje de muestras para las que el correspondiente estimador obtiene un mejor ajuste en términos de entropía.
[b] Valores medios a partir de 1.000 muestras simuladas.
[c] ENT: entropía; MSE: error cuadrático medio; RMSE: raíz de MSE; AME: error absoluto medio; RME: error relativo medio.
Fuente: Elaboración propia.

el plan de muestreo CIS o un plan de muestreo más sencillo como el AL ensayado en este estudio; especialmente si el estimador adoptado es el HD.

Asimismo, centrándonos en el impacto que sobre las predicciones induce el comportamiento de los electores, encontramos que los resultados van en la línea a los logrados en la sección previa. El estimador PS es, de nuevo, el que genera mejores estimaciones en condiciones ideales. Ello, sin embargo, no inhabilita a sus competidores, pues, al igual que ocurría en el caso de Madrid, las predicciones obtenidas con estos son igualmente muy precisas (véanse las tablas 4 y 5)[29]. Por otro lado, el estimador HD vuelve a emerger como el mejor cuando las muestras son obtenidas con sesgo de no-respuesta y error de respuesta. En este caso, además, su ventaja comparativa con el estimador RAT se ve incrementada sustancialmente respecto a las elecciones de Madrid, mientras que su superioridad con respecto al estimador PS es ya enorme. Muy probablemente, el crecimiento en cuanto a su ventaja relativa del estimador HD sea debido al hecho

[29] De hecho, el estimador RAT funciona en esta ocasión incluso mejor que el PS en términos de sesgo.

GRÁFICO 2. *Representaciones Box-Whiskers del conjunto de las estimaciones obtenidas para PSC (figura superior izquierda), CiU (figura superior derecha), PP (figura central izquierda), ICV (figura central derecha), ERC (figura inferior izquierda) y Resto (figura inferior derecha) de las simulaciones con diseño muestral AL*



de que en este caso se ha trabajado con muestras significativamente más sesgadas, cuestión que, como fue comentada con anterioridad, parece favorecer al estimador HD. En efecto, como se puede apreciar en la tabla 5 observando las predicciones medias logradas con el estimador DIR, en los escenarios SNR y ER las muestras simuladas presentan enormes niveles de sesgo, que en el caso del PP superan incluso los seis puntos.

Por partidos, merece especial mención las estimaciones logradas para ICV y CiU, en la parte negativa, y para PSC y PP, en la positiva. Los mayores errores de predicción se concentran en ICV: ninguno de los estimadores empleados parece ser capaz de reducir significativamente los sesgos que presentan sus datos brutos (véase el gráfico 2). El caso de CiU es algo distinto. CiU tiende a presentar sesgo positivo en las muestras generadas con sesgo de no-respuesta y, aunque este hecho parece ser detectado por las distintas estrategias, lo cierto es que es corregido en exceso, obteniéndose, en general, predicciones con un sesgo negativo significativo. En ambos casos, sería interesante estudiar la posible distribución espacial de los apoyos o determinar si el tema se debe a una cuestión puntual consecuencia de las transferencias de votos concretos que se produjeron entre las elecciones de 2003 y 2007. En el otro lado de la balanza valdría la pena situar a las predicciones logradas para PSC y PP. En este caso, a pesar de presentar los datos brutos sesgos que podrían ser calificados de espectaculares, se observa que todos los estimadores realizan importantes correcciones alcanzándose en el caso de las obtenidas con el estimador HD predicciones bastante precisas.

Para finalizar, no obstante, es preciso notar que empero la evidente superioridad que presenta el estimador HD en las situaciones más realistas y de que sus valores medios están ahora, en general, más cerca de los verdaderos valores, los comentarios del párrafo anterior muestran que existe todavía margen de mejora en la búsqueda de estrategias más precisas. En cualquier caso, parece evidente que en presencia de no-respuesta asimétrica la estrategia HD supera claramente a las estrategias PS y RAT utilizadas actualmente por la industria demoscópica.

Estimando con datos reales: encuestas CIS 2716 y 2720

Los resultados de las simulaciones presentados en la sección anterior muestran que, en presencia de sesgo de no-respuesta, el estimador HD es la opción que genera con mayor frecuencia la mejor estimación de la distribución de votos. En este apartado se ofrecen las predicciones que se obtendrían si los cuatro estimadores analizados en esta investigación fuesen aplicados a los datos recogidos en las encuestas que han servido de referente en este estudio.

Hasta ahora, en este trabajo se ha considerado exclusivamente un solo tipo de no-respuesta: total. En las encuestas reales, sin embargo, es muy habitual que también se produzca otro tipo de no-respuesta: parcial. Es decir, encontrar individuos para los que solo se recogen respuestas a algunas de las preguntas del cuestionario. En estas situaciones el analista puede optar entre dos alternativas. Excluir del análisis a aquellos individuos para los que no se dispone de respuesta en todas las variables relevantes o adoptar un esquema teóricamente más eficiente y utilizar la información muestral disponible para predecir (imputar) los valores no disponibles de las variables relevantes.

En nuestro caso, las respuestas relevantes para realizar las estimaciones son las contestaciones a los ítems sobre voto actual y recuerdo de voto. Las encuestas 2716 y 2720, aunque planeadas de tamaño 1.000, tuvieron tamaños efectivos de 969 y 974, respectivamente. De los encuestados, sin embargo, solo 807 y 728 ofrecieron una respuesta válida a los dos ítems requeridos para

generar las estimaciones; mientras que algunos más, 881 y 826 respectivamente, fueron los que facilitaron respuesta sobre su voto actual. A la luz de estos datos, se han considerado tres escenarios distintos para, utilizando los datos recogidos en las encuestas, generar predicciones. En el primer escenario son utilizados exclusivamente los datos brutos de aquellos individuos que informan sobre su comportamiento electoral actual y previo. Sin embargo, al no ser este enfoque teóricamente eficiente (pues supone prescindir de una gran cantidad de información todavía disponible en la encuesta), se han considerado dos escenarios adicionales en los que las respuestas faltantes relevantes (No sabe/No contesta) son estimadas mediante el uso de técnicas de imputación. En concreto, un segundo escenario, de imputación simple, donde se ha imputado un recuerdo de voto a los individuos que dan respuesta sobre su voto actual pero no reportan recuerdo; y un tercer escenario, de imputación doble, donde se ha procurado imputar voto actual y recuerdo de voto de aquellos encuestados para los que no se dispone de cualquiera de tales respuestas.

El problema, por tanto, se traslada a elegir un método de imputación entre los que han sido propuestos en la literatura: imputación por regresión, aleatoria, de la media, múltiple, por criterio experto, *hot-deck*, *cold-deck* (e.g., Schafer, 1997; Särndal y Lundström, 2005; Galván y Medina, 2007). En este

TABLA 6. *Predicciones y errores de las estimaciones de porcentajes de votos a las elecciones a la Asamblea de Madrid 2007 a partir de datos de la encuesta CIS 2716*

Escenario[a]	Estimador	Porcentajes				Errores[c]			
		PP	PSOE	IU	RE	PP	PSOE	IU	RE
Elecciones Oct-2003	Resultados[b]	48,47	39,04	8,51	3,98	—	—	—	—
Datos Brutos n = 807	DIR	46,55	37,67	11,18	4,59	–6,58	4,34	2,32	–0,07
	PS	48,17	35,47	11,12	5,24	–4,97	2,14	2,26	0,57
	RAT	48,92	33,72	11,23	6,13	–4,22	0,39	2,37	1,46
	HD	50,23	32,93	11,94	4,91	–2,91	–0,41	3,08	0,24
Imputación Simple n = 881	DIR	47,65	36,42	10,67	5,26	–5,49	3,08	1,81	0,60
	PS	48,93	35,10	10,51	5,47	–4,21	1,76	1,65	0,80
	RAT	49,80	33,87	10,85	5,48	–3,34	0,53	1,99	0,82
	HD	50,68	32,91	10,99	5,42	–2,45	–0,42	2,12	0,75
Imputación Doble n = 929	DIR	48,51	35,41	10,81	5,27	–4,62	2,07	1,95	0,60
	PS	49,57	34,33	10,46	5,64	–3,57	1,00	1,60	0,97
	RAT	50,19	33,38	10,64	5,78	–2,94	0,05	1,78	1,11
	HD	51,12	32,39	11,08	5,40	–2,01	–0,94	2,22	0,74
Elecciones 2007	Resultados[b]	**53,14**	**33,33**	**8,86**	**4,67**	—	—	—	—

[a] Datos brutos: sin imputación, solo se utilizan los individuos para los que se dispone de observación de ambas variables: voto actual y recuerdo de voto; Imputación simple: variable recuerdo de voto imputada para los entrevistados con voto actual observado; Imputación doble: cualquiera de las variables no observadas de voto actual o recuerdo de voto es imputada; n, tamaño efectivo de la muestra utilizada.

[b] Porcentajes de voto válido para población residente.

[c] Calculado como diferencia entre el porcentaje estimado y el real.

Fuente: Elaboración propia a partir de datos CIS.

TABLA 7. *Medidas de bondad de ajuste a los resultados reales de las predicciones a la Asamblea de Madrid derivadas de encuesta CIS 2716*[a,b]

	Datos brutos				Imputación simple				Imputación doble			
Estimador	ENT	RMSE	AME	RME	ENT	RMSE	AME	RME	ENT	RMSE	AME	RME
DIR	5,31	4,11	3,33	13,27	4,23	3,29	2,74	13,19	3,42	2,73	2,31	12,47
PS	3,66	2,94	2,48	13,38	3,06	2,46	2,11	12,25	2,45	2,07	1,78	12,14
RAT	2,70	2,53	2,11	16,78	2,20	2,00	1,67	11,96	1,81	1,81	1,47	12,41
HD	1,99	2,13	1,66	11,65	1,69	1,68	1,44	11,51	1,63	1,61	1,48	11,87

[a] Datos brutos: sin imputación, solo se utilizan los individuos para los que se dispone de observación de ambas variables: voto actual y recuerdo de voto; Imputación simple: variable recuerdo de voto imputada para los entrevistados con voto actual observado; Imputación doble: cualquiera de las variables no observadas de voto actual o recuerdo de voto es imputada; n, tamaño efectivo de la muestra utilizada.

[b] ENT: entropía; RMSE: raíz de MSE; AME: error absoluto medio; RME: error relativo medio.

Fuente: Elaboración propia.

caso, y en aras a combinar operatividad y significatividad, se ha optado por emplear imputación por criterio experto (Särndal y Lundström, 2005) utilizando como variables informativas, entre otras, la valoración de líderes, la autoubicación ideológica del encuestado, su cercanía a los partidos, la ubicación ideológica de los partidos por el encuestado, el voto del encuestado en otras elecciones y su valoración de las políticas o de los resultados registrados.

Las tablas 6 y 8 presentan, respectivamente, los resultados de las predicciones obtenidas para la Asamblea de Madrid y locales de Barcelona de 2007 utilizando cada uno de los tres conjuntos de datos disponibles (sin imputación, con imputación simple y con imputación doble); mientras que las tablas 7 y 9 muestran los estadísticos de ajuste de las distribuciones estimadas.

Como puede deducirse fácilmente observando la tabla 6, los resultados brutos de la encuesta 2716 son enormemente sesgados y aunque las imputaciones reducen significativamente los sesgos, estos permanecen todavía sensiblemente altos. En el caso de Madrid las predicciones obtenidas con el estimador HD son las que más se ajustan a los resultados reales con una notable ventaja sobre las estimaciones RAT, que se sitúan en segundo lugar, y las predicciones PS, que aunque mejoran los resultados brutos están muy lejos de los verdaderos valores (véase la tabla 7). En los resultados con imputaciones, se observa que éstas reducen el sesgo y mejoran la precisión global de los estimadores. La mejora en la calidad de las predicciones, sin embargo, no incide por igual en todas las estrategias. El estimador HD es el que menos se beneficia de este esfuerzo extra, mientras el resto de estimadores obtienen tras la imputación avances sensibles en sus niveles globales de ajuste.

En el caso de Barcelona los resultados son algo diferentes. En primer término destaca que, en general, los datos brutos son algo menos sesgados (aunque todavía presentan importantes niveles de sesgo) y resalta, especialmente, la enorme sobrecorrección que introduce incluir el recuerdo de voto en las estimaciones de CiU y PP y, también, PSC. Ello probablemente explique por qué, sin imputación, todos los estimadores que corrigen con recuerdo de voto funcionan, en esta ocasión, peor que el estimador directo (véase la tabla 8).

Entre los tres estimadores que utilizan recuerdo de voto, el estimador RAT es el que funciona claramente peor, con el estimador PS siendo el que mejores resultados obtiene

TABLA 8. *Predicciones y errores de las estimaciones de porcentajes de votos a las elecciones locales de Barcelona 2007 a partir de datos brutos de la encuesta CIS 2720*

Escenario[a]	Estimador	Porcentajes						Errores[c]					
		PSC	CiU	PP	ICV	ERC	RE	PSC	CiU	PP	ICV	ERC	RE
Elecciones 2003	Resultados[b]	33,57	21,45	16,12	12,07	12,81	3,97	—	—	—	—	—	—
Datos brutos n = 728	DIR	31,41	26,99	11,21	11,38	8,66	10,36	1,45	1,54	–4,40	2,03	–0,19	–0,43
	PS	27,65	22,47	16,17	12,59	8,82	12,30	–2,31	–2,99	0,57	3,24	–0,03	1,51
	RAT	26,15	21,16	17,43	10,51	9,02	15,73	–3,81	–4,30	1,82	1,16	0,17	4,95
	HD	27,85	22,17	17,50	11,16	9,26	11,07	–2,11	–3,28	1,89	2,81	0,40	0,29
Imputación simple n = 826	DIR	32,31	26,58	9,37	11,66	8,99	32,31	2,36	1,12	–6,24	2,32	0,14	0,31
	PS	28,01	22,24	15,50	12,61	8,90	12,73	–1,95	–3,21	–0,10	3,27	0,05	1,95
	RAT	26,03	20,77	17,81	10,39	8,95	16,04	–3,93	–4,68	2,21	1,04	0,10	5,26
	HD	27,39	22,78	17,41	11,91	9,18	11,33	–2,57	–2,68	1,80	2,57	0,33	0,55
Imputación doble n = 888	DIR	33,26	26,90	8,83	11,50	9,24	10,27	3,31	1,44	–6,78	2,15	0,39	–0,52
	PS	27,36	22,06	17,00	12,65	8,96	11,97	–2,59	–3,40	1,39	3,30	0,11	1,19
	RAT	24,88	20,10	19,97	9,87	8,91	16,28	–5,08	–5,36	4,37	0,52	0,05	5,49
	HD	27,76	23,18	17,91	11,55	9,28	10,31	–2,19	–2,28	2,30	2,20	0,43	–0,47
Elecciones 2007	Resultados[b]	29,96	25,46	15,61	9,35	8,85	10,77	—	—	—	—	—	—

[a] Datos brutos: sin imputación, solo se utilizan los individuos para los que se dispone de observación de ambas variables: voto actual y recuerdo de voto; Imputación simple: variable recuerdo de voto imputada para los entrevistados con voto actual observado; Imputación doble: cualquiera de las variables no observadas de voto actual o recuerdo de voto es imputada; n, tamaño efectivo de la muestra utilizada.

[b] Porcentajes de voto válido para población residente.

[c] Calculado como diferencia entre el porcentaje estimado y el real.

Fuente: Elaboración propia a partir de datos CIS.

TABLA 9. *Medidas de bondad de ajuste a los resultados reales de las predicciones al Ayuntamiento de Barcelona derivadas de la encuesta CIS 2720[a,b]*

	Datos brutos				Imputación simple				Imputación doble			
Estimador	ENT	RMSE	AME	RME	ENT	RMSE	AME	RME	ENT	RMSE	AME	RME
DIR	1,79	2,17	1,67	11,15	2,27	2,92	2,08	13,56	2,77	3,27	2,43	15,39
PS	2,04	2,14	1,77	12,02	1,96	2,18	1,76	12,23	2,34	2,33	2,00	13,08
RAT	3,24	3,22	2,70	16,92	3,46	3,44	2,87	17,79	4,32	4,16	3,48	20,52
HD	2,12	2,12	1,80	11,57	2,08	2,00	1,75	11,15	1,91	1,85	1,65	10,64

[a] Datos brutos: sin imputación, solo se utilizan los individuos para los que se dispone de observación de ambas variables: voto actual y recuerdo de voto; Imputación simple: variable recuerdo de voto imputada para los entrevistados con voto actual observado; Imputación doble: cualquiera de las variables no observadas de voto actual o recuerdo de voto es imputada; n, tamaño efectivo de la muestra utilizada.

[b] ENT: entropía; RMSE: raíz de MSE; AME: error absoluto medio; RME: error relativo medio.

Fuente: Elaboración propia.

en términos de entropía (véase la tabla 9). Su ventaja respecto al estimador HD es, sin embargo, más bien exigua, pudiendo ser incluso cuestionada. En efecto, por una parte, se observa que, con imputación doble, el estimador HD mejora al estimador PS en términos de entropía; mientras, por otra, las diferencias entre ambos conjuntos de estimadores (PS y HD) son más bien mínimas. La elección del mejor, por tanto, depende del índice que se utilice para medir el grado de ajuste. De hecho, en las estimaciones sin imputación, cada estimador registra los mejores índices de ajuste para dos de los cuatro indicadores; mientras en el resto de escenarios, el estimador PS solo produce el mejor índice en una ocasión. En todo caso, la posible ventaja relativa del estimador PS, en este caso, se sustentaría en las mejores estimaciones que genera este estimador para el porcentaje de voto para el PP.

Conclusiones

Cuando se trabaja con voto escrutado, los modelos de superpoblación basados en la congruencia que presentan los resultados agregados a pequeña escala (mesas, secciones, colegios electorales) de elecciones sucesivas han ayudado a mejorar sensiblemente las predicciones que se obtienen con muestras sesgadas. Paralelamente, se ha observado que el sesgo de no-respuesta constituye un inconveniente creciente en las encuestas electorales, donde las muestras sesgadas son la norma. El objetivo de este trabajo ha consistido en estudiar la capacidad predictiva de estos métodos en un entorno de investigación por encuestas y en evaluar su solvencia frente a los estimadores utilizados actualmente por la industria. Como elemento añadido, además, se ha tratado de dar una respuesta a la cuestión de si el uso de un procedimiento de selección muestral (previsiblemente menos costoso económica y temporalmente) donde se seleccionasen un mayor número de electores en un menor número de secciones censales (muestreo AL) podría ser implementado sin menoscabo para la calidad de las estimaciones.

Para responder a estas cuestiones se ha realizado un complejo ejercicio de simulación donde han sido generadas una ingente cantidad de muestras de dos elecciones de diferentes ámbitos (Asamblea de Madrid 2007 y Ayuntamiento de Barcelona 2007) bajo distintos escenarios de comportamiento de los electores al ser entrevistados. Así mismo, se han analizado los datos reales (sin y con imputación para reducir la no-respuesta parcial) de dos encuestas post-electorales (2716 y 2720) del CIS. Cuatro estimadores distintos han sido utilizados para generar las estimaciones. Un estimador directo (DIR) que traslada a porcentajes los resultados brutos de la encuesta y tres estimadores que hacen uso de las respuestas de recuerdo de voto para tratar de mejorar las estimaciones: el estimador ratio ponderado (RAT), el estimador de post-estratificación (PS) y el estimador de superpoblación basado en la congruencia que a nivel de sección dibujan los resultados electorales de elecciones sucesivas (HD).

A la luz de los resultados, se ha logrado un valioso conjunto de conclusiones con un amplio alcance práctico. A pesar de los diferentes ámbitos geográficos y políticos considerados en los dos ejercicios de simulación implementados y de las diferentes hipótesis consideradas sobre respuesta asimétrica por parte de los electores, las conclusiones obtenidas en las simulaciones para ambas elecciones son muy similares. En concreto, se puede decir que:

i) Todos los estimadores generan estimaciones altamente precisas en condiciones ideales (sin sesgos ni error de respuesta).

ii) El uso del recuerdo de voto mejora la calidad de las estimaciones, con el estimador PS produciendo las soluciones más próximas a los datos brutos.

iii) El estimador PS es, de los tres estimadores elaborados usando recuerdo de voto, el que produce las mejores predicciones en condiciones ideales.

iv) El estimador PS, sin embargo, sufre mucho en presencia de sesgo de no-respuesta y sus predicciones empeoran a medida que el sesgo crece.

v) El estimador HD se revela claramente como el mejor estimador tan pronto se consideran situaciones más realistas (aparece sesgo de no-respuesta y error de respuesta en las muestras).

vi) El estimador RAT genera soluciones bastante similares (por término medio) al estimador HD, aunque en general de peor calidad.

vii) Las distancias de precisión relativas entre el estimador HD y los estimadores RAT y PS se incrementan a medida que crece el sesgo de no-respuesta.

viii) A pesar de la evidente superioridad que presenta el estimador HD en las situaciones más realistas parece existir todavía margen de mejora para la búsqueda de estrategias más precisas.

ix) En general, la selección de un menor número de secciones (unido a la realización de una mayor cantidad de entrevistas por sección) empeora ligeramente la precisión de todas las estimaciones, aunque de manera desigual por estimador. El estimador HD es el que menos sufre el cambio de estrategia muestral.

x) En general, el uso combinado del estimador HD y la estrategia muestral AL permitiría incrementar la calidad de las estimaciones reduciendo los costes económicos.

Los resultados con los datos reales derivados de las encuestas 2716 y 2720 apuntan en la misma dirección a los obtenidos en las simulaciones. Por una parte, en la encuesta 2716 el estimador HD domina claramente a sus competidores y es capaz de reducir sensiblemente el enorme sesgo de los datos originales. Por otra parte, a partir de los datos de la encuesta 2720, se obtiene que ambos estimadores PS y HD generan predicciones comparables, mejorando las prestaciones del estimador HD a medida que disponemos de datos originales más sesgados tras la imputación.

A la vista de los resultados obtenidos en este estudio, la recomendación evidente es que el estimador HD debería ser potenciado frente a las estrategias PS y RAT. Casi con certeza tal decisión conllevaría una mejora media en la precisión de las predicciones. De este trabajo también se deduce que la calidad de las estimaciones no se vería alterada en modo especialmente significativo si se adoptase un diseño muestral en el que se muestreasen menos secciones y más electores por sección, especialmente si ello viniese acompañado del uso del estimador HD. Por lo que, dado que previsiblemente la adopción de un plan de este tipo sería económicamente menos costoso, el único impedimento que, *a priori*, podría desaconsejar su adopción sería que tal cambio afectase a la validez de la estimación de otras cuestiones que también se suelen preguntar en las encuestas electorales (como valoración de líderes). Esta potencial limitación, sin embargo, podría ser adecuadamente superada mediante selecciones intencionadas de secciones; cuestión que sería perfectamente asumible dentro del enfoque HD, pues este no precisa de una selección aleatoria para su aplicación. La adopción del enfoque HD posibilitaría, además, recoger las ventajas de utilizar muestras intencionadas, que permitiesen garantizar, *a priori*, una adecuada representación de todo el espectro sociopolítico en la encuesta.

A pesar de lo manifestado en el párrafo anterior, el estimador HD no es una panacea. Por una parte, el hecho de que el estimador HD sea el que mejor funciona por término medio con muestras realistas no garantiza que siempre vaya a generar mejores predicciones que sus competidores RAT y PS.

Mientras, por otro lado, como muestran las cifras de sesgo que aún presenta este estimador, todavía existe margen de mejora en este contexto. En tal sentido, sería interesante explorar si la consideración explícita de la componente geoespacial de los datos y/o un uso más eficiente de la información muestral y poblacional podrían conducir a la obtención de predicciones más precisas.

Aparte de las cuestiones reseñadas en el párrafo anterior, que conllevarían el seguir otras sugerentes vías de exploración empleando enfoques alternativos, dentro del camino inaugurado con este trabajo hay, al menos, un par de cuestiones que todavía están por resolver y que deberían ser abordadas en el futuro, como son: el decidir el diseño muestral que mejor se adaptaría a las características del estimador HD; y el determinar la estrategia de cálculo para la varianza muestral más adecuada teniendo en cuenta las características del estimador HD.

Referencias

Aybar, C. (1998): *Modelos de Superpoblación y Poblaciones Discretas Finitas*, Tesis Doctoral, Valencia, Universitat de València.

Bernardo, José M. (1997): «Probing Public Opinion: The State of Valencia Experience», en C. Gatsonis, J. S. Hodges, R.E. Kass, R. McCulloch, P. Rossi y N. D. Singpurwalla (eds.), *Bayesian Case Studies 3*, Nueva York: Springer-Verlag.

— y F. Javier Girón (1992): «Robust Sequential Prediction from Non-random Samples: the Election Night Forecasting Case», en J. M. Bernardo, J. O. Berge, A. P. Dawid y A. F. M. Smith (eds.), *Bayesian Statistics 4*, Oxford: Oxford University Press.

Biemer, Paul, Ralph Folsom, Richard Kulka, Judith Lessler, Babu Shah y Michael Weeks (2003): «An Evaluation of Procedures and Operations Used by the Voter News Service for the 2000 Presidential Election», *Public Opinion Quarterly,* 67: 32-44.

Curtin, Richard, Stanley Presser y Eleanor Singer (2005): «Changes in Telephone Survey Nonresponse over the Past Quarter Century», *Public Opinion Quarterly,* 69: 87-98.

De Leeuw, Edith y Wim de Heer (2002): «Trends in Household Survey Nonresponse: A Longitudinal and International Comparison», en R. M. Groves, D. A. Dillman, J. L. Etinge y R. J. A. Little (eds.), *Survey Nonresponse*, Nueva York: Wiley.

Galván, Marco y Fernando Medina (2007): *Imputación de Datos: Teoría y Práctica*, Santiago de Chile: United Nations Publications.

Greene, William H. (2003): *Econometric Analysis*, New Jersey: Prentice Hall.

Groves, Robert M. (1989): *Survey Errors and Survey Costs*, Nueva York: John Wiley & Sons.

— (2006): «Nonresponse Rates and Nonresponse Bias in Household Surveys», *Public Opinion Quarterly*, 70: 646-675.

—, Don A. Dillman, John L. Eltinge y Roderick J. A. Little (2002): *Survey Nonresponse*, Nueva York: Wiley.

—, Eleanor Singer, Amy D. Corning y Ashley Bowers, (1999): «A Laboratory Approach to Measuring the Joint Effects of Interview Length, Incentives, Differential Incentives, and Refusal Conversion on Survey Participation», *Journal of Official Statistics*, 15: 251-268.

— y Mick P. Couper (1998): *Nonresponse in Household Interview Surveys*, Nueva York: Wiley.

Gschwend, Thomas, Ron Johnston y Charles Pattie (2003): «Split-Ticket Patterns in Mixed-Member Proportional Election Systems: Estimates and Analyses of their Spatial Variations at the German Federal Election, 1998», *British Journal of Political Science*, 33: 109-128.

Isaki, Cary T., Julie H. Tsay y Wayne A. Fuller (2004): «Weighting Sample Data Subject to Independent Controls», *Survey Methodology*, 30: 35-44.

Johnston, Ron y Charles Pattie (2000): «Ecological Inference and Entropy-Maximizing: An Alternative Estimation Procedure for Split-Ticket Voting», *Political Analysis*, 8: 333-345.

Kalton, Graham y Daniel Kasprzyk (1986): «The Treatment of Missing Data», *Survey Methodology*, 12: 1-6.

Keeter, Scoot, Carolyn Miller, Andrew Kohut, Robert M. Groves y Stanley Presser (2000): «Consequences of Reducing Nonresponse in a National Telephone Survey», *Public Opinion Quarterly,* 64: 125-48.

King, Gary (1997): *A Solution to the Ecological Inference Problem: Reconstructing Individual Beha-*

vior from Aggregate Data, Princeton, NJ: Princeton University Press.

—, James Honaker, Anne Joseph y Kenneth Scheve (2001): «Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation», *American Political Science Review*, 95: 49-69.

Konner, Joan (2003): «The Case for Caution. This System Is Dangerously Flawed», *Public Opinion Quarterly*, 67: 5-18.

Martin, Elizabeth A., Michael W. Traugott y Courtney Kennedy (2005): «A Review and Proposal for a New Measure of Poll Accuracy», *Public Opinion Quarterly*, 69: 342-369.

Merkle, Daniel M. y Murray Edelman (2002): «Nonresponse in Exit Polls: A Comprehensive Analysis», en R. M. Groves, D. A. Dillman, J. L. Etinge y R. J. A. Little (eds.), *Survey Nonresponse*, Nueva York: Wiley.

Mitofsky, Warren J. (2003): «Voter News Service after the Fall», *Public Opinion Quarterly*, 67: 45-58.

— y Edelman Murray (2002): «Election Night Estimation», *Journal of Official Statistics*, 18: 165-179.

Pavía, José M. (2010): «Improving Predictive Accuracy of Exit Polls», *International Journal of Forecasting*, 26: 68-81.

—, Bernardi Cabrer y Ramón Sala (2009): «Updating Input-Output Matrices: Assessing Alternatives through Simulation», *Journal of Statistical Computation and Simulation*, 79: 1467-1498.

—, Beatriz Larraz y José María Montero (2008): «Election Forecasts Using Spatio-Temporal Models», *Journal of the American Statistical Association*, 103: 1050-1059.

— y Antonio López-Quilez (2012): «Spatial Vote Redistribution in Redrawn Polling Units», en evaluación.

Pavía-Miralles, José Manuel (2005): «Forecasts from Non-Random Samples: The Election Night Case», *Journal of the American Statistical Association*, 100: 1113-1122.

— y Beatriz Larraz-Iribas (2008): «Quick Counts from Non-Selected Polling Stations», *Journal of Applied Statistics*, 35: 383-405.

Randall, J. Jones, Jr. (2008): «The State of Presidential Election Forecasting: The 2004 Experience», *International Journal of Forecasting*, 24: 310-321.

Rodríguez Osuna, Jacinto (1991): *Métodos de Muestreo,* Cuadernos Metodológicos del CIS, 1. Madrid: Centro de Investigaciones Sociológicas.

— (2005): *Métodos de Muestreo. Casos Prácticos,* Cuadernos Metodológicos del CIS, 6. Madrid: Centro de Investigaciones Sociológicas.

Rosén, Bengt (1997a): «Asympotic Theory for Order Sampling», *Journal of Statistical Planning and Inference*, 62: 135-158.

— (1997b): «On Sampling with Probability Proportional to Size», *Journal of Statistical Planning and Inference*, 62: 159-191.

Särndal, Carl-Erik (2007): «The Calibration Approach in Survey Theory and Practice», *Survey Methodology*, 33: 99-119.

— y Sixten Lundström (2005): *Estimation in Surveys with Nonresponse*, Chichester: John Wiley & Sons.

—, Bengt Swenson y Jan Wretman (2003): *Model Assisted Survey Sampling*, Nueva York: Springer.

Schafer, J. L. (1997): *Analysis of Incomplete Multivariate Data*, Boca Ratón, Florida: Chapman & Hall/CRC.

Singer, Eleanor (2006): «Introduction: Nonresponse Bias in Household Surveys», *Public Opinion Quarterly*, 70: 637-645.

Smith, Tom W. (1999): «Developing Nonresponse Standards», *International Conference on Survey Nonresponse*, Portland, October, 1999. http://cloud9.norc.uchicago.edu/dlib/nonre.htm.

Valliant, Richard, Alan H. Dorfman y Richard M. Royall (2000): *Finite Population Sampling and Inference. A Prediction Approach*, Nueva York: John Wiley & Sons.

RECEPCIÓN: 27/04/2010
REVISIÓN: 22/10/2010
APROBACIÓN: 03/06/2011

Apéndices

Apéndice 1: Estimación de las distribuciones de voto cruzadas (recuerdo de voto)

La estrategia de cálculo de las distribuciones de voto en cada sección censal se ha implementado mediante un proceso en dos etapas. En una primera etapa, se ajustaron las matrices de transferencia que se derivan de las encuestas 2716 y 2720 para que fueran coherentes con los resultados reales registrados a nivel de circunscripción. En una segunda etapa, se estimó una matriz de transferencia en cada sección ajustando en cada sección censal la matriz obtenida para el conjunto de la circunscripción[30].

A fin de ilustrar en qué consiste el proceso, considérense dos elecciones sucesivas con solo 3 opciones electorales (A, B y C) en las que participaron un millón de votantes. Supóngase que en las primeras elecciones los resultados que cosechó cada opción, A, B y C, fueron, respectivamente, 500.000, 400.000 y 100.000; mientras que en las siguientes elecciones fueron 450.000, 440.000 y 110.000. Admítase que se realizó una encuesta a 1.000 electores en la que cada encuestado fue interrogado por su voto en las elecciones actuales y previas y que los resultados se clasificaron en una tabla de doble entrada como la que aparece a continuación, donde por filas aparecen los recuerdos de voto y por columnas los votos en las elecciones actuales (por ejemplo, 55 electores manifestaron haber elegido la opción B en las elecciones previas y la opción A en las presentes elecciones).

	A	B	C	
A	400	80	10	490
B	55	350	15	420
C	3	10	77	90
	460	440	100	1.000

Es obvio que esta matriz de transferencia no es completamente coherente con los resultados agregados registrados en ambas elecciones (por ejemplo, el recuerdo de voto que se derivaría de la encuesta para la opción C sería de 90.000 votantes). No obstante, tomando como punto de partida la distribución cruzada que se deriva de la encuesta, se podrían aproximar las transferencias de voto que hubo entre elecciones entre las distintas opciones imponiendo que la suma por filas y columnas coincidan con los valores realmente registrados. En concreto, aplicando a nivel de circunscripción el método RAS (véase, e.g., Pavía *et al.*, 2009), se obtendría la siguiente matriz de transferencia:

	A	B	C	
A	399.517	89.735	10.748	500.000
B	47.393	338.698	13.909	400.000
C	3.090	11.567	85.343	100.000
	450.000	440.000	110.000	1.000.000

Una vez calculada la estimación de la distribución de voto cruzada a nivel de circunscripción se aplica la misma metodología a nivel de sección (tomando como referente la distribución cruzada estimada para la circunscripción). Así, supóngase que en una circunscripción de 1.000 votantes, 700, 250 y 50 eligieron, respectivamente, las opciones A, B y C en las elecciones previas y que en las elecciones actuales fueron 630, 309 y 61. Aplicando a estos datos el balanceado de

[30] Desde una perspectiva práctica-heurística es ineficiente trabajar con todos los posibles comportamientos electorales. Por lo que, a fin de hacer más manejable el problema, en el caso de las elecciones a la Asamblea de Madrid, se consideraron como posibles categorías votar por PP, PSOE, IU, Otros o en Blanco, Abstenerse, o No votar por edad (esta última opción solo para recuerdo de voto); y, en el caso de las elecciones locales de Barcelona, votar por PSC, CiU, PP, ERC, ICV, Otros o en Blanco, Abstenerse, o No votar por edad (esta última opción solo para recuerdo de voto).

matrices con la distribución estimada en el paso anterior como distribución inicial se obtendría como estimación de la distribución cruzada de esta sección:

	A	B	C	
A	590	99	11	700
B	38	204	8	250
C	2	6	42	50
	630	309	61	1.000

La estimación de estas distribuciones para cada sección censal permite que una vez seleccionado un votante (eligiendo una sección al azar y dentro de la sección, también al azar, un elector) se le pueda asignar un voto actual y un recuerdo de voto.

Apéndice 2: Selección de secciones y simulación de respuestas

Los detalles relativos a la selección de secciones y número y respuestas de los encuestados se ofrecen a continuación. Para el muestreo CIS, se seleccionaron en la Comunidad de Madrid 101 secciones, eligiéndose en cada sección seleccionada alrededor de 10 (entre 8-12) electores, y en Barcelona 100 secciones eligiendo 10 electores por sección. La selección de secciones se realizó siguiendo el diseño polietápico estratificado por conglomerados característico de las encuestas del CIS (véase Rodríguez Osuna, 2005). Por otro lado, en el diseño AL y para ambas elecciones, se seleccionaron 25 secciones (con probabilidades de selección proporcionales al tamaño de la sección; Rosén, 1997a, 1997b) y se *entrevistaron* 40 electores en cada sección.

Una vez elegidas las secciones que componen las muestras se aborda la última etapa: la obtención de las respuestas de los entrevistados. Para ello, se ha seguido un criterio jerárquico de complejidad en el comportamiento de los electores al ser entrevistados. Partiendo del caso ideal en el que todos los electores contactados acceden a responder con veracidad, se ha ido profundizando en el realismo de las muestras simuladas hasta generar situaciones donde el comportamiento de los electores reproduce los dos tipos de errores no muestrales con mayor incidencia en la precisión de las estimaciones (Groves, 1989; Särndal y Lundström, 2005; Pavía, 2010) y más habituales en los estudios por encuesta: el sesgo de no-respuesta y el error de respuesta. En concreto, se han simulado muestras de tres tipos: i) muestras sin error (que notaremos SE), donde cada encuestado reporta fielmente su voto emitido en las elecciones actuales y en las previas[31]; ii) muestras con sesgo de no-respuesta (SNR), donde existen distribuciones asimétricas de no-respuesta para los votantes de cada partido, es decir, situaciones en las que la probabilidad de que un elector contactado esté dispuesto a responder a las preguntas depende de su voto y es distinta según su opción electoral; y, iii) muestras con sesgo de respuesta y error de respuesta (que denotaremos ER), donde además de producirse el comportamiento señalado en el punto ii) una parte de los encuestados reporta respuestas incorrectas.

La selección de los individuos se realizó, en todos los casos, por muestreo aleatorio simple (sin reposición). En las muestras SNR y ER, no obstante, cada sujeto inicialmente *contactado* era sometido a una prueba dicotómica o Bernouilli[32] para determinar si pasa-

[31] Aunque hay electores que no votan y, por tanto, su opción electoral es no votar y no pueden reportar ningún «voto emitido», a fin de simplificar el lenguaje durante todo el documento cuando se hable de voto de un elector se entenderá que la opción no votar está incluida como una posibilidad.

[32] Una prueba de Bernouilli sirve para resolver un problema de decisión dicotómico en el cual mediante el «lanzamiento de una moneda» (con probabilidades de cara y cruz no necesariamente iguales) se decide (a partir del resultado del lanzamiento) si el sujeto seleccionado realiza o no determinada acción.

ba a formar parte de la muestra final. En el caso iii), además, los sujetos finalmente seleccionados eran sometidos a una segunda prueba de Bernouilli para decidir si reportaban o no respuestas verídicas.

En el caso i), en cada sección seleccionada, se eligieron al azar m_s electores (con m_s en torno a 10 para muestreo CIS o igual a 40 con diseño AL) y sus votos en las elecciones presentes y pasadas recopilados. En los casos ii) y iii), por su parte, en primer lugar se ordenaron aleatoriamente los electores de la sección y de acuerdo con la ordenación obtenida a cada sujeto se le sometió a una prueba Bernouilli con la que se iba decidiendo si el votante pasaría o no a formar parte de la muestra. Este proceso se aplicó de forma secuencial hasta obtener una muestra de tamaño m_s en la correspondiente sección o hasta que todos los electores de la sección hubiesen pasado por la prueba de Bernouilli. La probabilidad de que cada sujeto superase esta primera prueba de Bernouilli (y, por tanto, pasase a formar parte de la muestra) no era la misma para todos los votantes y dependía de su comportamiento electoral. En concreto, las probabilidades de pertenecer a la muestra asignadas a cada sujeto fueron derivadas comparando los resultados registrados en las correspondientes elecciones y las estimaciones que se obtenían directamente en las encuestas de referencia. Las probabilidades fueron elegidas para que las muestras generadas replicaran un sesgo de no-respuesta similar al realmente observado en las encuestas que sirvieron de modelo. En el caso de las elecciones a la Asamblea de Madrid, se supuso que en cada sección entre un 50 y un 60% de los electores que no votaron declinarían participar en la encuesta, de donde se estimó que entre los que votaron estos porcentajes se situarían entre un 15 y un 25%, de los cuales alrededor de un 70% (65-75%) serían votantes del PP y en torno a un 22% (20-25%) votantes socialistas, con aproximadamente un 5% (3-7%) votantes de IU y un 3% (0-5%) votantes de otras opciones. En el caso de Barcelona, por su parte, se supuso que en cada sección entre un 65 y un 55% de los electores que no votaron declinarían participar en la encuesta, estimándose que este porcentaje se situaría entre un 10 y un 20% para los votantes del PSC, entre un 25-15% entre los de CiU, del orden del 60-50% en los votantes del PP, 23-17% en ERC, 8-2% para ICV, 40-50% en Otros y sobre el 3% de los votantes en blanco. La aplicación de estos porcentajes sobre los resultados reales iba a permitir alcanzar muestras simuladas que arrojaban sesgos similares a los realmente registrados en las encuestas de referencia, con los rangos porcentuales anteriores permitiendo flexibilizar la estructura de no-respuesta en cada sección dentro de un patrón general para el conjunto de la circunscripción.

Finalmente, en las muestras tipo iii) los votantes seleccionados fueron adicionalmente sometidos a una segunda prueba de Bernouilli, para decidir si el elector reportaría su voto real[33]. A los votantes que superaban esta segunda prueba Bernouilli, se les generaba al azar una respuesta para su voto actual y para su recuerdo de voto.

[33] La probabilidad de que un encuestado no reportara sus verdaderos valores se fijó *ad hoc* en un 5%.