

# La imputación múltiple como alternativa al análisis de la no respuesta en la variable intención de voto<sup>1</sup>

*Cristina Rivas, María del Mar Martínez Rosón y Purificación Galindo*

En este artículo se aborda la falta de respuesta en la pregunta sobre intención de voto en la encuesta preelectoral del CIS para las elecciones generales de 2000. El objetivo de este trabajo es testar un modelo de imputación múltiple aplicado a los datos faltantes de la pregunta sobre intención de voto en tres comunidades autónomas (CC AA) con diferentes porcentajes de falta de respuesta: Asturias, Castilla y León y Galicia. La principal conclusión del trabajo es que el procedimiento de imputación múltiple resulta efectivo para la estimación de la no respuesta cuando las variables del cuestionario permiten diferenciar a todas las categorías de la intención de voto.

*Palabras clave: imputación múltiple, encuestas preelectorales, falta de respuesta, intención de voto, actitudes políticas.*

## INTRODUCCIÓN

Los análisis realizados a partir de datos de encuesta en distintas disciplinas se enfrentan habitualmente a la existencia de datos faltantes. Ante la falta de información en algunas preguntas de las encuestas, la práctica más común que llevan a cabo los investigadores es obviar los datos faltantes y trabajar únicamente con la información disponible<sup>2</sup>.

1. La realización de este trabajo ha sido posible gracias a los fondos recibidos del Centro de Investigaciones Sociológicas a través de las Ayudas a la Investigación 2006. Asimismo, agradecemos los comentarios de Aníbal Pérez-Liñán y de los evaluadores anónimos que sin duda han ayudado a mejorar el texto.
2. Método *listwise* o *pairwise deletion*. Según Schafer (1997), cuando el volumen de registros incompletos es un porcentaje reducido (5% o menos), la eliminación de estos registros es una solución razonable a los problemas de datos faltantes. Para una descripción más detallada de los problemas de la eliminación de casos véase Little y Rubin (1987).

El principal inconveniente es que se reduce el número de individuos de la muestra, corriendo el riesgo de que las estimaciones realizadas sean sesgadas si la distribución de la falta de respuesta no es completamente aleatoria, algo que ocurre con frecuencia<sup>3</sup>. Otra alternativa disponible para los investigadores es tratar de recuperar esos casos, sustituyendo el valor perdido por otro dato calculado por el propio investigador a partir del resto de los casos. Se trata de aplicar un determinado procedimiento estadístico que permita asignar un valor al dato faltante; los más habituales son los métodos de imputación simple donde a cada valor perdido se le asigna un único valor completo<sup>4</sup>. Frente a esta opción es posible utilizar métodos de imputación múltiple, poco habituales en el mundo de las ciencias sociales y políticas que, a diferencia de la imputación simple para cada valor faltante, generan más de un posible valor completo<sup>5</sup>. En este artículo se trata de comprobar el funcionamiento de la imputación múltiple como modo de recuperar datos faltantes en la variable de intención de voto a partir de la encuesta preelectoral del Centro de Investigaciones Sociológicas (CIS) del año 2000. El objetivo central de este trabajo es probar la efectividad del procedimiento de imputación múltiple desarrollado por Rubin en 1987 con datos de carácter electoral y determinar si este tipo de técnicas pueden contribuir de forma positiva a paliar el frecuente problema de la falta de respuesta en las encuestas de opinión.

## DATOS, VARIABLES Y ANÁLISIS PREVIOS A LA IMPUTACIÓN

En los últimos estudios preelectorales realizados por el CIS para las elecciones generales el porcentaje de los individuos incluidos en las categorías No sabe (Ns) y No contesta (Nc) asciende al 24% en 1996, 27% en 2000, 31% en 2004 y 27% en 2008.

Los datos utilizados en este trabajo proceden de la encuesta preelectoral realizada por el CIS para las elecciones generales del año 2000<sup>6</sup>. Para realizar el análisis se han seleccionado tres CC AA, Asturias, Castilla y León y Galicia, en las que probar el modelo. En cada una de estas comunidades han sido entrevistados 546, 2.731 y 1.799 ciudadanos, respectivamente, de modo que el análisis se realiza con 5.076 entrevistas. El criterio de selección de estas tres comunidades se basa en el porcentaje de falta de respuesta en la intención de voto,

3. Por ejemplo, en el caso de los estudios electorales la falta de respuesta no se distribuye de igual modo entre los partidos políticos. En España para el Partido Popular (PP) durante muchos años su voto real fue mayor que el pronosticado por los sondeos electorales, lo que reflejaba que los potenciales votantes de este partido eran más reticentes a declarar su voto en las encuestas que los electores de otras fuerzas políticas. Ante este tipo de situaciones los investigadores y analistas políticos tratan de corregir estos aspectos, realizando estimaciones de los datos a partir de otras variables relacionadas con la intención de voto, lo que se conoce como "cocina" de los datos (Anduiza y Bosch, 2004: 84).
4. Algunas de las técnicas de imputación simple más utilizadas son la imputación por medias no condicionadas, imputación con variables ficticias o mediante regresión. Sobre este tema véase Medina y Galván (2007: 25-29).
5. Sobre imputación múltiple de datos faltantes en ciencia política véase King, Honaker, Joseph y Scheve (2001).
6. Se trata de una encuesta de ámbito nacional realizada a la población española de ambos sexos mayor de 18 años, con una muestra total de 24.040 casos. Dado que el objetivo de este trabajo es probar la eficacia de un método novedoso en ciencias sociales, consideramos que el año de realización de la encuesta, 2000, no es significativo ni determinante en nuestros resultados.

corregida con la variable de simpatía partidista que cada una de ellas presenta en la encuesta<sup>7</sup>. Asturias como ejemplo de región con baja tasa de falta de respuesta a la pregunta de intención de voto más simpatía (4,9% de datos faltantes), Castilla y León como representante de comunidades con una falta de respuesta intermedia (10,9%) y, por último, Galicia con un porcentaje elevado (17%)<sup>8</sup>. De este modo, estaríamos tomando muestras representativas que recogen un intervalo amplio de falta de respuesta en la variable objeto de estudio.

La variable dependiente en este análisis es la intención de voto corregida con las respuestas sobre simpatía hacia algún partido político en las elecciones generales de 2000, clasificada en las siguientes categorías: Partido Popular (PP), Partido Socialista Obrero Español (PSOE), Izquierda Unida (IU), Partido Nacionalista Vasco (PNV), Convergencia i Unió (CiU), otros partidos<sup>9</sup>, en blanco, no votaría, Ns y Nc. La intención de voto más simpatía se construye asignando a todos aquellos sujetos que dicen no saber todavía a qué partido votar en las elecciones generales de 2000 el partido por el que sienten una mayor afinidad o simpatía y al que tendría mayor probabilidad de votar en el supuesto de acudir a las urnas el día de las elecciones<sup>10</sup>. De este modo, los individuos en los que se profundizará en este estudio son aquellos de los que es difícil conocer si participarán o no en los comicios electorales y, en caso de hacerlo, por qué partido votar.

Las técnicas de recogida de datos a partir de las encuestas de opinión permiten elaborar modelos explicativos del comportamiento electoral, basados no sólo en las características sociodemográficas de los electores, sino también en sus actitudes políticas. Por eso las variables independientes utilizadas para la caracterización e imputación de los datos faltantes de la variable dependiente son relativas a valores o actitudes políticas que pueden influir en el voto de los ciudadanos<sup>11</sup>. Se distinguen tres grupos de actitudes o valores políticos: el primero referido a las actitudes hacia partidos y actores políticos; el segundo relativo a la implicación política del individuo, y el tercero a la ideología e identificación partidista.

- 
7. En este artículo consideraremos dato faltante tanto a los sujetos que no contestan a la pregunta sobre intención de voto como a los que dicen no saber a qué partido votar ni por qué partido sienten más simpatía o cercanía, puesto que consideramos que estos últimos, en realidad, no quieren declarar su voto.
  8. La falta de respuesta es menor que la registrada a nivel nacional, dado que al considerar como variable dependiente la intención de voto más simpatía, algunos de los datos faltantes iniciales han podido ser recuperados.
  9. Se incluyen dentro de esta categoría a todos aquellos partidos con menos del 1% de respuestas. El hecho de colocar bajo el mismo paraguas ("otros partidos") a un gran número de fuerzas políticas implica que los resultados de los análisis posteriores sean pobres para esta categoría, dada la heterogeneidad interna de la misma, pero el reducido tamaño de la muestra para la gran mayoría de fuerzas políticas que componen esta categoría ha llevado a tomar esta decisión.
  10. Somos conscientes de que no todos los que declaran simpatía o cercanía por un partido votan a ese partido. En la decisión de voto entran en juego otros muchos factores como el "voto útil" o "voto estratégico", que en el caso de IU o los partidos minoritarios se verá más acentuado. A la vista de los resultados electorales y de las encuestas de opinión de los últimos años, se ha puesto de manifiesto que, por ejemplo, IU recibe un mayor porcentaje de votos en las encuestas que su porcentaje real de voto. Y es que son muchos los que, sintiéndose próximos a IU, deciden votar a otra fuerza política de izquierdas, PSOE, que consideran tiene mayores opciones de ganar unas elecciones que IU, de este modo sienten que su voto no está perdido. Sobre voto útil o estratégico véase Lago (2005, 2006) y Oñate (2002).
  11. Sobre actitudes políticas y comportamiento electoral, véase, entre otros: Font, Montero y Torcal (eds.) (2006); Anduiza y Bosch (2004); Del Castillo (ed.) (1994).

La ideología izquierda-derecha y la identificación partidista son las variables que, habitualmente, más influencia ejercen sobre el comportamiento electoral y más particularmente sobre el voto a un partido, mientras que variables relativas al interés por la política pueden ejercer una mayor influencia, no tanto en el voto, como en la participación electoral. En este trabajo se incluyen también este tipo de preguntas, puesto que se considera que pueden contribuir a caracterizar al grupo de los “No votaría”<sup>12</sup>. En la Tabla 1 se presentan las variables empleadas en los distintos análisis realizados en este trabajo.

TABLA 1.  
VARIABLES EN EL ANÁLISIS

#### **Actitudes hacia partidos y actores políticos**

- Confianza en líderes políticos: J. M. Aznar y J. Almunia. Respuestas: mucha, bastante, poca o ninguna confianza.
- Valoración de líderes políticos (escala 0-10, donde 0 significa que lo valora muy mal y 10 muy bien): J. Almunia, J. M. Aznar y F. Frutos.
- Valoración de la gestión del PP durante los últimos cuatro años. Respuestas: muy buena, buena, regular, mala o muy mala.
- Valoración de la gestión del PSOE durante los últimos cuatro años. Respuestas: muy buena, buena, regular, mala o muy mala.

#### **Implicación política**

- Interés por la política. Respuestas: mucho, bastante, poco o nada.
- Frecuencia con la que habla de política. Respuestas: habitualmente, de vez en cuando, rara vez, nunca o casi nunca.
- Frecuencia con la que sigue la información política a través de la prensa, televisión y radio. Respuestas: todos o casi todos los días, cuatro o cinco días por semana, dos o tres días por semana, sólo fines de semana, de vez en cuando, nunca o casi nunca.

#### **Ideología e identificación partidista**

- Ubicación ideológica del entrevistado. Escala de 1 a 10, donde 1 significa izquierda y el 10 derecha.
- Probabilidad de votar a: IU, PP, PSOE y BNG\*. Escala de 0 a 10, donde 0 significa “con toda seguridad, no le votaría nunca” y el 10 “con toda seguridad, le votaría siempre”.

\* Variable incluida en los análisis para Galicia.

Fuente: elaboración propia a partir de cuestionario del CIS (encuesta número 2.382).

Antes de proceder a probar el método de imputación, es necesario conocer si los individuos que no responden a la pregunta sobre intención de voto constituyen un grupo homogéneo con un comportamiento y características distintas a las del resto del electorado. Si esto fuese así, la imputación carecería de sentido, ya que estos sujetos no serían semejantes al electorado de ninguna opción política y, por tanto, tampoco podríamos probar la efectividad del método, ya que

12. Se ha prescindido de las variables sociodemográficas (sexo, edad y nivel de estudios, que son las únicas disponibles en el preelectoral de 2000), porque el modelo de imputación requiere que la variable a imputar y el vector de covariables que se emplee para la imputación de los datos faltantes estén altamente correlacionados. Las variables sociopolíticas tienen una correlación más alta que las sociodemográficas con la variable intención de voto, de ahí la decisión de no incluirlas en el modelo.

los datos imputados en la variable intención de voto no se ajustarían a la realidad de los datos faltantes, que tendrían características diferentes a los votantes de un partido político.

Esta comprobación se ha realizado mediante la aplicación de un análisis *cluster* (no jerárquico, k-medias), que es una técnica multivariante que permite agrupar a los individuos en función de las similitudes existentes entre ellos de forma que los sujetos pertenecientes al mismo grupo sean más parecidos entre sí que los pertenecientes a otros grupos<sup>13</sup>. Los resultados del análisis de conglomerados k-medias<sup>14</sup> muestran que los individuos que no saben o no contestan el partido por el que van a votar tienen un comportamiento similar al del resto del electorado español (figura 1 del Anexo). En el conjunto nacional sus actitudes políticas son parecidas y mantienen el mismo patrón de comportamiento político en las variables incluidas en el análisis y esto es así tanto para los Ns como para los Nc.

Dado que para probar la eficacia de la imputación múltiple nos centraremos únicamente en tres CC AA, es necesario comprobar previamente que las personas que no declaran su intención de voto en estas regiones tengan también un comportamiento similar al del resto del electorado que sí manifiesta su voto. Los resultados de los análisis revelan que los Ns y Nc tanto en Asturias, Castilla y León como en Galicia tienen actitudes políticas similares a las de aquellos que dicen el partido por el que van a votar (figura 2 del Anexo)<sup>15</sup>. Por tanto, el hecho de que el colectivo de Ns y Nc no formen un grupo homogéneo con unas características propias diferentes a las del resto del electorado en estas tres CC AA permite proceder al análisis de datos faltantes en la variable intención de voto mediante el procedimiento de imputación múltiple.

## APLICACIÓN DE LA IMPUTACIÓN MÚLTIPLE

Para estimar los datos faltantes en la variable intención de voto, se utiliza un procedimiento de imputación múltiple (Rubin, 1987) específico para variables de carácter nominal

- 
13. Para conocer si los Ns y Nc en la variable intención de voto constituyen un grupo homogéneo o no, con unas características propias, se ha utilizado un análisis de homogeneidad, HOMALS, que pertenece al grupo de técnicas exploratorias de análisis multivariante no lineal que trabaja con variables de carácter nominal (Gifi, 1990). Su objetivo principal es identificar un espacio de pocas dimensiones que permita resumir y representar la estructura de asociaciones entre dos o más variables nominales, así como las similitudes entre los sujetos pertenecientes a estas categorías. Esta técnica se ha empleado, principalmente, con el fin de transformar las variables independientes de carácter ordinal o nominal en variables numéricas a partir de las puntuaciones factoriales obtenidas por los sujetos en el HOMALS y emplearlas como variables de clasificación en el análisis *cluster*.
  14. Los cuatro grupos obtenidos del análisis de conglomerados, utilizando como variables de clasificación las puntuaciones factoriales obtenidas para cada uno de los sujetos en el análisis de homogeneidad, se presentan en la tabla 1 del Anexo. A modo de exploración se han realizado dos análisis de conglomerados k-medias, solicitando tres y cinco grupos. Se observó que el análisis con cuatro grupos ofrecía *clusters* más estables y coherentes con la literatura sobre comportamiento electoral, por lo que se optó por utilizar esta opción.
  15. En Castilla y León y Galicia, que cuenta con un mayor porcentaje de datos faltantes (10,9% y 17,0%, respectivamente), el patrón de comportamiento político se ajusta de forma más clara al del conjunto general de los encuestados a nivel nacional (Figura 1 del Anexo). En Asturias, pese a que el número de personas que no manifiestan su intención de voto es bajo (4,9% de datos faltantes en la variable intención de voto), los Ns y Nc también se dispersan, como en el resto de casos, por todo el gráfico lo que indica que estos sujetos están presentes en los diferentes grupos de votantes de la CCAA.

implementado por SAS<sup>16</sup>. El objetivo de los modelos de imputación múltiple es crear predicciones para cada uno de los datos faltantes, no la explicación causal o la interpretación de los parámetros obtenidos. La imputación múltiple estima para cada dato faltante  $m$  valores posibles, creando  $m$  bases de datos completas<sup>17</sup>. En cada una de esas bases de datos completas los valores observados son los mismos, pero los datos faltantes presentan valores imputados diferentes que reflejan distintos niveles de incertidumbre. Finalmente la imputación múltiple requiere aplicar un modelo estadístico que combine las  $m$  imputaciones realizadas para cada dato faltante de manera que se obtenga un único valor estimado para cada valor perdido.

Para realizar la imputación múltiple es necesario especificar un modelo en el que exista una correlación alta entre la variable a imputar y el vector de covariables que se empleará para la imputación de los datos faltantes. Atendiendo a las características de nuestra variable a imputar, la intención de voto, se aplica el modelo de función discriminante específicamente diseñado para el empleo de variables nominales<sup>18</sup> que asume que los datos faltantes son perdidos al azar, MAR (*missing at random*). Ésta es una condición necesaria para la imputación múltiple, en la medida en que los datos observados puedan predecir los datos faltantes. De este modo, aunque las probabilidades de responder a la pregunta sobre intención de voto puedan ser menores entre los votantes de un determinado partido, el hecho de que al menos algunos individuos pertenecientes a ese partido contesten a preguntas sobre identificación partidista e ideología permitirán predecir su intención de voto (King *et. al*, 2001: 51). No obstante, el incumplimiento de este supuesto de aleatoriedad de los datos faltantes introducido por Rubin (1976), según estudios más recientes (Schafer, 1999), no influiría en los resultados del modelo de imputación múltiple utilizado.

La imputación múltiple mantiene las ventajas de la imputación simple (como la posibilidad de trabajar con bases de datos completas) y, además, mejora o corrige algunas de sus desventajas. Por ejemplo, corrige la desventaja de la imputación simple respecto a la subestimación de la verdadera varianza cuando la proporción de datos faltantes es elevada (Rao y Shao, 1992). Como las  $m$  imputaciones son repeticiones bajo un modelo de predicción de valores faltantes, el análisis de los valores imputados puede combinarse de forma sencilla para realizar inferencias que tengan en cuenta la variabilidad del muestreo y por consiguiente la variabilidad de los datos faltantes (López Vázquez, 2005). Y, si además se tiene en cuenta que las  $m$  imputaciones proceden de más de un modelo, la incertidumbre de encontrar el modelo correcto queda diluida por la variabilidad en las inferencias a través de los modelos utilizados para las  $m$  imputaciones (Little y Rubin, 1987 y 1989). Sin embargo, el método de imputación múltiple también tiene algunos inconvenientes, el principal es el esfuerzo computacional que requiere su aplicación, a diferencia de las técnicas de imputación simple.

Para evaluar la efectividad de la imputación múltiple en nuestros datos es necesario trabajar con bases de datos completas, es decir, bases en las que todas las variables que intervienen

16. Para una mayor información véase Yuan (2000) y SAS Institute Inc. (2004, capítulo 44).

17. El número de imputaciones realizadas suele estar entre tres o diez. Por defecto, SAS realiza tres imputaciones.

18. El método de la función discriminante para variables nominales es implementado por SAS 9.1, aunque se encuentra en fase experimental. El método de imputación múltiple se ha realizado aplicando el comando PROCMI de SAS.

en la imputación tengan todos los datos. El único modo de conocer si un dato ha sido imputado correctamente es conociendo el valor real del dato; por este motivo, se ha trabajado con una base de este tipo. Sobre esta base de datos completa, en cada una de las CC AA objeto de análisis se ha provocado de forma aleatoria la eliminación del dato real de la variable intención de voto, convirtiéndolo en un dato faltante<sup>19</sup>. Siguiendo este patrón se han creado cuatro variables de intención de voto donde se elimina un 5%, 10%, 17% y 25% de los datos de dicha variable<sup>20</sup>. De este modo, se podría conocer como funcionaría la imputación múltiple en cada una de las CCAA si su porcentaje de falta de respuesta en la intención de voto fuese menor, igual o mayor al porcentaje real de datos faltantes que han tenido cada una de ellas en la encuesta. La distribución de los datos completos y faltantes ficticios que se imputan en cada una de las CCAA se presenta en la tabla 2.

TABLA 2.

NÚMERO DE CASOS Y DATOS FICTICIOS A IMPUTAR POR CC AA

Comunidad autónoma	Nº casos en la muestra	Nº casos completos	Datos faltantes ficticios			
			5%	10%	17%	25%
Asturias	546	325	16	41	53	81
Castilla y León	2.731	1.418	76	143	245	357
Galicia	1.799	847	41	96	140	214

*Fuente:* elaboración propia con datos del CIS (estudio 2.382).

Por otro lado, diseñamos dos modelos de imputación con variables diferentes en cada uno de ellos. En el modelo 1 se incluyen tres grupos de variables: valoración de líderes y partidos políticos, identificación partidista e ideología e implicación política. En el modelo 2 dejamos fuera las variables de implicación política, por considerar que este grupo de variables pueden ser más efectivas para la determinación de la participación política que para la intención de voto en sí misma (Anduiza y Bosch, 2004: 126).

## EVALUACIÓN DE LOS RESULTADOS DE LA IMPUTACIÓN MÚLTIPLE

Los resultados de la imputación múltiple de los datos faltantes ficticios en cada una de las CC AA son evaluados tratando de dar respuesta a cuatro cuestiones: a) la efectividad de la imputación en función del porcentaje de datos faltantes que tenga la base de datos; b) el modelo de variables creado para la imputación; c) funcionamiento del método en las diferentes CC AA; y d) efectividad según opción política o electoral.

19. Esto sólo es posible si las personas que Ns o Nc tienen unas características políticas similares a las del resto de personas de la encuesta que responden a la pregunta sobre intención de voto. Si esto es así, eliminar aleatoriamente datos completos no sesgaría los resultados de la imputación.

20. Los tres primeros porcentajes se corresponderían con la falta de respuesta real en la variable intención de voto más simpatía en las CC AA seleccionadas.

TABLA 3.  
PORCENTAJE DE ÉXITO DE LA IMPUTACIÓN

Comunidad autónoma	5%		10%		17%		25%		Efectividad media	
	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2
Asturias	62,5	81,3	78,0	73,2	83,0	81,1	74,1	80,2	74,0	79,0
Castilla y León	81,6	88,2	85,3	88,8	84,9	84,9	83,8	83,5	83,9	86,4
Galicia	87,8	90,2	91,7	91,7	83,6	85,0	100,0	98,6	90,8	91,4

M1: modelo 1. M2: modelo 2.

Fuente: elaboración propia con datos del CIS (estudio 2.382).

### *Efectividad de la imputación según el porcentaje de datos faltantes*

A priori, es fácil pensar que el método de imputación múltiple tendría un mayor éxito cuando el número de datos faltantes es menor. Pero a la vista de los resultados obtenidos (tabla 3), el éxito de la imputación múltiple es similar en cada una de las CC AA, independientemente del porcentaje de datos faltantes que tenga la variable intención de voto. Es más, se puede decir que no existe un patrón de comportamiento que explique la efectividad del modelo aplicado en función del número de datos faltantes. Al contrario de lo que se podía pensar, porcentajes de falta de respuesta bajos, en la práctica totalidad de los casos analizados, conducen a un peor funcionamiento de la imputación. Un ejemplo claro de ello es el modelo 1 de Asturias, donde para un porcentaje de datos faltantes del 5%, el éxito alcanzado es del 62,5%, muy lejos del obtenido para tasas de no respuesta superiores al 10%. Igualmente en Galicia la imputación resulta mucho más efectiva cuando el porcentaje de falta de respuesta es del 25%.

La técnica de imputación mantiene unos resultados homogéneos al interior de las CC AA, independientemente del porcentaje de datos faltantes. Las tasas de éxito conseguidas por la imputación múltiple en los dos modelos aplicados oscilan entre el 62,5% y el 100% y, por tanto, la evaluación del método de imputación múltiple es bastante positivo.

### *Efectividad en función del modelo de variables*

El éxito de la imputación múltiple mejora sustancialmente para la mayor parte de los casos en el modelo 2, cuando se dejan fuera las variables relativas a la implicación política (interés por la política, frecuencia con la que habla de política y seguimiento de la información política en los medios de comunicación). Esto resulta mucho más acentuado en el caso de Asturias, donde para un porcentaje de falta de respuesta del 5% el éxito de la imputación aumenta en 20 puntos porcentuales de un modelo a otro (tabla 3). En cambio, con el 10% y el 17% de datos faltantes, el método empeora ligeramente respecto al modelo 1. En Castilla y León, el modelo 2 resulta más efectivo que el 1 cuando los porcentajes de falta de respuesta son inferiores o iguales al 10%; en el resto de casos la efectividad del

procedimiento de imputación se mantiene constante. En Galicia, no incluir en la función discriminante las variables de implicación política aumenta, para la mayor parte de los casos, las probabilidades de éxito de la imputación, si bien no alcanza la máxima efectividad conseguida por el modelo 1 cuando la falta de respuesta es del 25%.

En definitiva, y salvo pequeñas excepciones, podemos decir que la eficacia del método de imputación múltiple empleado depende del modelo de variables creado para la imputación. En los casos analizados, la implicación política contribuye en mayor medida a una asignación errónea de valores a los datos faltantes.

#### *Funcionamiento del método en las diferentes CC AA*

A la vista de los resultados expuestos en la tabla 3, parece claro que la imputación de la intención de voto es más eficiente en unas CC AA que en otras. Independientemente del modelo considerado y los porcentajes de datos faltantes existentes en la variable a imputar, las probabilidades de que la imputación sea exitosa son muy altas en Galicia, con una probabilidad media de éxito alrededor del 90% en ambos modelos. Algo que también sucede en Castilla y León con una efectividad media entre el 83,9% y el 86,4%. En cambio, en Asturias, la imputación múltiple tiende a funcionar peor en ambos casos, con una tasa de acierto inferior al 80%. Una de las posibles explicaciones al mayor éxito de la imputación conseguido en Galicia está en la presencia de un partido nacionalista con peso en la vida política de la región. Es posible pensar que la incorporación de la variable probabilidad de votar al BNG en el modelo de imputación haya contribuido de forma favorable a la asignación correcta de valores perdidos a la categoría "otros partidos", categoría con una mayor dificultad de identificación en las otras dos CC AA.

#### *Efectividad de la imputación múltiple según opción política o electoral*

El hecho de que los resultados obtenidos en la imputación múltiple sean favorables no significa que el procedimiento testado sea igual de efectivo para todas las opciones políticas o electorales en cada una de las CC AA (tablas 4, 5 y 6). El procedimiento de imputación múltiple aplicado resulta mucho más exitoso en el caso de los partidos mayoritarios, PP y PSOE, y en buena medida IU, que cuando se trata de las categorías "otros partidos" (excepto en el caso de Galicia), voto en blanco o no votaría. En términos generales la efectividad más alta la encontramos en el PP (con porcentajes de acierto en la imputación de valores a este partido superiores al 95%, excepto en Asturias)<sup>21</sup>, si bien la imputación múltiple también funciona de manera satisfactoria para el PSOE con porcentajes de éxito medio entre el 80% y el

---

21. A diferencia de lo que sucede en Galicia y Castilla y León, feudos del PP durante muchos años, en Asturias, y para las elecciones generales de 2000, el margen de votos del PP respecto del PSOE fue relativamente bajo.

90%<sup>22</sup>. En el caso de IU, los resultados obtenidos varían significativamente de unas CC AA a otras. La presencia de dos partidos de izquierda en el panorama electoral español puede contribuir al aumento de la tasa de error en la imputación para ambos partidos, y principalmente para el más minoritario, IU. Aunque el perfil de votante de IU y del PSOE suele ser diferente, el uso del “voto útil” en el electorado de IU se traduce, con frecuencia, en un mayor incremento de votos al PSOE. El hecho de que un individuo manifieste su simpatía e identificación partidista por IU no significa que finalmente decida votar por él en los comicios electorales.

En lo que respecta a las opciones políticas restantes, el fracaso de la imputación es evidente. Cuando se trata de imputar valores para la categoría “otros partidos”, la efectividad del proceso fracasa de manera contundente en Castilla y León, con porcentajes inferiores al 25% y es exitosa en Galicia con un 87% de aciertos. Esto pone de manifiesto la importancia de las variables introducidas en el modelo, la presencia en la arena política gallega de un partido nacionalista fuerte, como es el caso del BNG con un 19% de potenciales votantes en la base de datos completa considerada, hace que la incorporación en el análisis de la variable probabilidad de votar al BNG haya favorecido su efectividad, algo de lo que carecen Castilla y León y Asturias. En esta última, ninguno de los datos faltantes eliminados de la variable intención de voto correspondían a votantes de otros partidos; esto es comprensible si tenemos en cuenta que en la base de datos completa de esta comunidad tan sólo un 1,5% de los entrevistados manifiesta su intención de votar por otro partido que no sea PP, PSOE o IU. Quizá, en Castilla y León, donde el porcentaje de votantes a otros partidos es algo mayor que el existente en Asturias, preguntar por la probabilidad de votar por Unión del Pueblo Leonés (UPL) y por Tierra Comunera-Partido Nacionalista Castellano (TC-PNC) contribuyese a mejorar la efectividad de la imputación de la categoría otros partidos.

TABLA 4.

PORCENTAJE DE ÉXITO DE LA IMPUTACIÓN EN ASTURIAS SEGÚN OPCIÓN POLÍTICA

Partido político	5%		10%		17%		25%		Efectividad media	
	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2
IU	33,3	66,7	75,0	62,5	80,0	80,0	66,7	88,9	63,8	72,5
PP	83,3	83,3	86,7	73,3	100,0	94,4	89,7	93,1	89,9	86,0
PSOE	66,7	100,0	75,0	81,3	90,5	90,5	80,6	87,1	82,9	89,7
En blanco	,0	,0	,0	,0	,0	,0	,0	,0	,0	,0
No votaría	**	**	100,0	100,0	50,0	50,0	50,0	50,0	66,7	66,7
(n)	10	13	32	30	44	43	60	65	146	151
% éxito total	62,5	81,3	78,0	73,2	83,0	81,1	74,1	80,2	74,0	79,0

\*\* En la selección aleatoria de los casos no fue seleccionado ningún individuo para esta categoría.

M1: modelo 1. M2: modelo 2.

Fuente: elaboración propia con datos del CIS (estudio 2.382).

22. Pese a estos resultados satisfactorios, una tasa de éxito menor en el PSOE lleva a pensar que el potencial electorado de este partido tiene actitudes menos definidas que las de los posibles electores del PP.

La imputación múltiple como alternativa al análisis de la no respuesta variable...

109

TABLA 5.

PORCENTAJE DE ÉXITO DE LA IMPUTACIÓN EN CASTILLA Y LEÓN SEGÚN OPCIÓN POLÍTICA

Partido político	5%		10%		17%		25%		Efectividad media	
	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2
IU	100,0	100,0	90,0	80,0	71,4	66,7	74,1	77,8	83,9	81,1
PP	93,3	97,8	93,0	94,2	97,7	95,3	96,3	94,7	95,1	95,5
PSOE	63,6	81,8	83,8	91,9	85,5	92,8	88,5	89,4	80,3	89,0
Otros partidos	,0	50,0	**	**	20,0	20,0	10,0	,0	10,0	23,3
En blanco	**	**	,0	25,0	,0	,0	10,0	10,0	3,3	11,7
No votaría	75,0	25,0	33,3	50,0	38,9	33,3	26,3	31,6	43,4	35,0
(n)	62	67	122	127	208	208	299	299	691	701
% éxito total	81,6	88,2	85,3	88,8	84,9	84,9	83,8	83,5	83,9	86,4

\*\* En la selección aleatoria de los casos no fue seleccionado ningún individuo para esta categoría.

M1: modelo 1. M2: modelo 2.

Fuente: elaboración propia con datos del CIS (estudio 2.382).

TABLA 6.

PORCENTAJE DE ÉXITO DE LA IMPUTACIÓN EN GALICIA SEGÚN OPCIÓN POLÍTICA

Partido político	5%		10%		17%		25%		Efectividad media	
	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2
IU	**	**	100,0	100,0	40,0	60,0	100,0	83,3	80,0	81,1
PP	100,0	100,0	95,1	100,0	96,1	98,7	100,0	100,0	97,8	99,7
PSOE	100,0	100,0	94,4	77,8	83,3	79,2	100,0	100,0	94,4	89,3
Otros partidos	76,9	84,6	85,7	78,6	83,3	87,5	100,0	100,0	86,5	87,8
En blanco	,0	,0	,0	,0	,0	,0	100,0	100,0	25,0	25,0
No votaría	**	**	,0	100,0	14,3	,0	100,0	77,8	38,1	59,3
(n)	36	37	88	88	117	119	214	211	455	455
% éxito total	87,8	90,2	91,7	88,8	83,6	85,0	100,0	98,6	90,8	91,4

\*\* En la selección aleatoria de los casos no fue seleccionado ningún individuo para esta categoría.

M1: modelo 1. M2: modelo 2.

Fuente: elaboración propia con datos del CIS (estudio 2.382).

A la vista de los resultados obtenidos para el voto en blanco, resulta claro que la imputación múltiple no es capaz de asignar correctamente su valor en la práctica totalidad de los casos. La falta de efectividad del método para esta opción política es general, independientemente de la CC AA de la que se trate. Tan sólo es exitoso en Galicia cuando el porcentaje de datos faltantes es del 25% (tabla 6). Esta falta de acierto puede ser atribuida a dos cosas. En primer lugar, al escaso número de votantes en blanco que existe en las tres CC AA en las bases a imputar, siempre considerando los datos completos, 3,7% en Asturias, 2,3% en Castilla y León y 1,9% en Galicia, lo que puede implicar una mayor dificultad del procedimiento para encontrar características comunes a estos sujetos. En segundo

lugar, a la falta de preguntas en el cuestionario que capten el comportamiento y las características sociopolíticas de este grupo de electores.

Tampoco resultan especialmente satisfactorios los resultados obtenidos para la abstención. Si bien cabría pensar que la efectividad de la imputación para los “no votaría” fuese mayor en el modelo 1 (en el que se incluyen variables de implicación política) que en el 2 (en el que se excluyen dichas variables), lo cierto es que no siempre es así en todas las CC AA. La imputación múltiple en esta categoría resulta mucho más efectiva en Asturias (66,7% de aciertos como media), sin existir diferencias entre un modelo y otro, que en Castilla y León y Galicia. Por su parte, Castilla y León, con los peores resultados en la abstención, mejora sustancialmente la efectividad con la aplicación de las variables incluidas en el modelo 1 que en el 2, reduciéndose el éxito de la imputación en un 8% (tabla 5). Sin embargo, la situación inversa se produce en Galicia, donde el procedimiento aplicado funciona mejor cuando se excluye de la imputación las variables sobre implicación política (tabla 6).

Por tanto, no podemos mantener, para los casos analizados, que las variables relativas a la implicación política (modelo 1) ayuden a identificar mejor a los abstencionistas; esto requiere de un análisis más detallado de los casos mal estimados en cada una de las CC AA. Por consiguiente, y a la vista de estos resultados, podemos concluir que la efectividad del procedimiento de imputación múltiple difiere significativamente según la opción política de la que se trate. Ahora bien, si realmente existe este efecto partido del que parece depender el éxito o fracaso del modelo de imputación, cabría pensar que el principal problema podría resolverse con el diseño del cuestionario.

## ANÁLISIS DE LOS ERRORES DEL PROCEDIMIENTO DE IMPUTACIÓN MÚLTIPLE

Hasta ahora hemos visto cuál es la efectividad de la imputación múltiple según cuatro aspectos: porcentaje de datos faltantes, variables que intervienen en la imputación, CC AA y opción política imputada. Pero nos queda saber en qué casos o situaciones el procedimiento de imputación múltiple ha fallado. Es decir, trataremos de conocer qué valores estimados son los que conducen a una mayor confusión en cada una de las CC AA, de forma que se pueda valorar si los errores cometidos por la imputación tienen una explicación razonable en lo que a comportamiento electoral se refiere o no.

En el caso de Asturias, los principales errores cometidos en la imputación se concentran en IU (tabla 7). El procedimiento imputa erróneamente a votantes del PSOE y votantes en blanco el voto de IU. Este tipo de error en el primer caso puede deberse al “voto útil” o “voto estratégico”, donde el individuo concede mayor importancia a ganar la elección que a sus propias ideas políticas. En cambio, el hecho de que a un buen número de potenciales votantes en blanco se les asigne como dato imputado IU revela que en Asturias parte de este electorado, en el año 2000, tenía un perfil parecido al de los votantes de IU. Otro de los errores más frecuentes en esta comunidad son los producidos en la

estimación de casos del PP al PSOE y viceversa. Un seguimiento de estos individuos demuestra que se trata de sujetos poco implicados en política, sin una identificación partidista clara y con una ideología de centro. Un error menos explicable, desde el punto de vista del comportamiento político, es el que se produce en la predicción de votantes del PP a IU y al contrario, que podrían ser considerados los casos más extremos. Por último, se puede decir que los abstencionistas en Asturias tienen un perfil más parecido al de los votantes de partidos políticos de izquierda que de derecha. El procedimiento de imputación tiende a asignarles a estos casos, principalmente, el voto de IU y, en menor medida, el del PSOE.

TABLA 7.

DISTRIBUCIÓN DE LOS ERRORES EN ASTURIAS

Voto imputado	Voto observado									
	IU		PP		PSOE		En blanco		No votaría	
	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2
IU			0	0	9	4	4	5	2	4
PP	1	2			2	3	1	1	0	0
PSOE	2	2	5	2			2	3	2	1
Otros partidos	2	1	0	0	0	0	1	0	1	0
En blanco	1	0	0	4	1	2			1	1
No votaría	2	1	1	2	2	0	3	2		
(n)	8	6	6	8	14	9	11	11	6	6

M1: modelo 1. M2: modelo 2.

Fuente: elaboración propia con datos del CIS (estudio 2.382).

En Castilla y León se mantienen buena parte de los errores que la imputación comete en Asturias. Las principales confusiones vienen dadas en los casos de votantes del PSOE a los que se les imputa el voto de IU, y viceversa. Algo parecido sucede con los votos del PSOE imputados como voto al PP y, al contrario, aunque teniendo en cuenta los resultados de la tabla 8, podemos decir que son más los errores cometidos con el PSOE que con el PP, lo que indicaría que el grupo de electores del PSOE para los que la imputación falla tiene un perfil sociopolítico próximo al del electorado del PP.

Otra de las principales fuentes de error de la imputación en esta comunidad se debe al voto a otros partidos que el programa tiende a considerar votantes del PP. La presencia en la arena electoral castellano-leonesa de dos partidos políticos regionales, UPL y TC, con un importante número de simpatizantes, contribuye a la falta de efectividad de la imputación para esta categoría. Si se tiene en cuenta que TC se caracteriza por ser un partido de izquierdas y UPL un partido de centro, más próximo al PSOE que al PP, sorprende la imputación a estos sujetos del voto del PP, pero un seguimiento de estos individuos revela que tienen un perfil ideológico de centro-derecha, manifiestan una mayor

probabilidad de votar al PP frente al PSOE o IU y valoran positivamente a J. M. Aznar. Este error podría subsanarse en buena parte si en el cuestionario preelectoral se incluyesen algunas preguntas relativas a la identificación partidista para UPL y TC, como sucede en otras CC AA con partidos nacionalistas fuertes o que obtienen en las elecciones un mayor número de votos.

El voto en blanco, al igual que sucede en Asturias, está asociado principalmente a IU, lo que pone de manifiesto que votantes en blanco y de IU tienen perfiles sociopolíticos parecidos. Los errores en la imputación de los “no votaría” en Castilla y León se reparten de forma similar por los tres principales partidos.

TABLA 8.  
DISTRIBUCIÓN DE LOS ERRORES EN CASTILLA Y LEÓN

Voto imputado	Voto observado											
	IU		PP		PSOE		Otros partidos		En blanco		No votaría	
	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2
IU			1	2	13	5	1	1	7	5	6	6
PP	4	1			15	10	8	9	3	3	10	9
PSOE	6	8	5	9			2	1	3	2	7	7
Otros partidos	1	2	7	4	2	0			1	2	4	5
En blanco	2	1	4	2	1	0	3	2			3	4
No votaría	1	3	2	5	5	8	1	2	2	3		
(n)	14	15	19	22	36	23	15	15	16	15	30	31

M1: modelo 1. M2: modelo 2.

Fuente: elaboración propia con datos del CIS (estudio 2.382).

En Galicia, a diferencia de lo sucedido en las otras dos CC AA analizadas, los errores debidos a la presencia de dos partidos políticos de izquierda, IU y PSOE, apenas existen. Se mantienen las confusiones en la asignación del voto entre PP y PSOE, siendo, en este caso, este reducido número de sujetos del PP mal imputados los que tienen un perfil más similar al del electorado del PSOE. Pese a que la inclusión en el modelo de imputación de la variable probabilidad de votar al BNG ha contribuido de forma notoria al éxito de la imputación en esta comunidad, todavía existe algún error a la hora de estimar el voto a otros partidos, asignándole el voto al PP o al PSOE. Una posible explicación a este problema podría encontrarse nuevamente, y como sucede en el caso de IU, en el “voto útil” o “voto estratégico”, o bien en la existencia de un voto dual.

En definitiva, el análisis de los errores cometidos por el método de imputación múltiple ponen de manifiesto que para que el procedimiento sea efectivo es imprescindible contar con un buen instrumento de medición de las actitudes políticas de todo el electorado y con una selección adecuada de las variables que intervienen en la imputación.

TABLA 9.  
DISTRIBUCIÓN DE LOS ERRORES EN GALICIA

Voto imputado	Voto observado											
	IU		PP		PSOE		Otros partidos		En blanco		No votaría	
	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2	M1	M2
IU			0	0	0	0	0	0	0	0	0	0
PP	0	0			3	2	4	2	2	2	3	4
PSOE	1	0	6	0			4	3	1	1	1	0
Otros partidos	1	1	0	0	1	5			3	2	3	5
En blanco	0	0	0	1	1	0	0	1			0	0
No votaría	1	2	0	0	0	2	1	2	0	1		
(n)	3	3	6	1	5	9	9	8	6	6	7	9

M1: modelo 1. M2: modelo 2.

Fuente: elaboración propia con datos del CIS (estudio 2.382).

## CONCLUSIONES

Esta investigación ha tratado de conocer mejor la naturaleza de la no respuesta en las investigaciones sociales, centrando su interés en el comportamiento electoral y, más concretamente, en la intención de voto de los individuos en los estudios preelectorales.

La primera conclusión que se extrae de este trabajo es que aquellas personas que se niegan a decir por qué partido van a votar no son un grupo homogéneo con unas características particulares, sino que presentan actitudes políticas similares a las del electorado de cualquier otra opción política. El objetivo de este trabajo era probar la efectividad de un modelo de imputación múltiple con el fin de conocer si este tipo de prácticas pueden paliar en lo posible la falta de respuesta en preguntas sobre intención de voto. Los resultados de la imputación demuestran principalmente tres cosas: a) que el porcentaje de datos faltantes en la variable intención de voto no es determinante en la efectividad de la imputación. Como se ha podido comprobar, un menor porcentaje de datos faltantes no necesariamente se traduce en una mayor efectividad de la imputación y viceversa, mayor falta de respuesta no se corresponde con peores resultados de la imputación; b) el papel que juegan las variables incluidas en el modelo, la inclusión o no de un determinado grupo de variables, demuestra su influencia en la efectividad de la imputación; y c) la falta de variables en el cuestionario preelectoral que permitan identificar quiénes son o qué características tienen algunos grupos de votantes contribuyen también a la falta de efectividad de la imputación. Este análisis ha puesto de manifiesto cómo, aun partiendo de información completa para las variables incluidas en la imputación, no en todos los partidos y categorías puede ser imputado el voto con una probabilidad de acierto similar. El principal problema se encuentra en los casos de

individuos que señalan que su voto será en blanco. No se han encontrado preguntas en el cuestionario que puedan caracterizar a estos individuos de forma más precisa. Lo mismo sucede con los abstencionistas y con los votantes a otros partidos. Este último es el caso más revelador y donde se demuestra que la existencia de variables que permitan identificar a este grupo de votantes conduce a una imputación exitosa, como ha sido el caso de Galicia con la incorporación de variables que identifican a los votantes del BNG.

Por esta razón se cree que una investigación posterior sobre los elementos y características que mejor definen a estos colectivos sería muy apropiada. Esto ayudaría a diseñar un cuestionario que incluya preguntas o ítems que permitan posteriormente mejorar la identificación de estos ciudadanos. Por tanto, una de las conclusiones que se puede derivar de esta circunstancia es que las variables que se incluyen en el procedimiento de imputación múltiple deben ser variables que permitan diferenciar a todas las categorías de la variable intención de voto.

En definitiva, los resultados de esta investigación ponen de manifiesto que el procedimiento de imputación múltiple en la variable intención de voto puede ser satisfactorio y una posible solución al problema de la falta de respuesta parcial en los estudios políticos. Aunque para obtener resultados satisfactorios con la imputación múltiple es necesario contar con buenas variables predictoras. Los análisis realizados muestran qué problemas generan ciertas preguntas y cómo éstas pueden resultar insuficientes o inapropiadas para realizar imputaciones. No obstante, hay que señalar que existen errores para los que la solución resulta más difícil; el caso más claro lo encontramos en los votantes de IU y el uso del “voto útil” o “voto estratégico”.

Una línea de investigación futura llevaría a profundizar sobre cuál sería la efectividad del método de imputación múltiple aplicado en este trabajo cuando hay datos faltantes en las variables que intervienen en la imputación, algo que con frecuencia sucede en las encuestas de opinión y particularmente en las encuestas electorales.

### Referencias

- Anduiza, E. y Bosch, A. 2004. *Comportamiento político y electoral*. Barcelona: Ariel.
- Del Castillo, P. (ed.) 1994. *Comportamiento político y electoral*. Madrid: CIS.
- Font, J., Montero, J. R. y Torcal, M. 2006. *Ciudadanos, asociaciones y participación en España*. Madrid: CIS.
- Gifi, A. 1990. *Nonlinear Multivariate Analysis*. Nueva York: John Wiley & Sons.
- King, G., Honaker, J., Joseph, A. y Scheve, K. 2001. “Analyzing Incomplete Political Science Data”, *American Political Science Review*, 95: 49-69.
- Lago Peña, I. 2005. “El voto estratégico en las elecciones autonómicas andaluzas de 1996”, *Revista Española de Investigaciones Sociológicas*, 109: 155-177.
- Lago Peña, I. 2006. “El voto estratégico en las elecciones generales en España (1977-2000)”, *Sistema*, 192: 125-127.
- Little, R. y Rubin, D. B. 1989. “The analysis of social science data with missing values”, *Sociological Methods and Research*, 18: 292-326.

- Little, R. y Rubin, D. B. 1987. *Statistical Analysis with Missing Data*. Nueva York: John Wiley & Sons.
- López Vázquez, V. 2005. *Comparación de los métodos de imputación con respecto al poder de separación del modelo de regresión logística*, tesis de grado de Maestro en Ciencias Matemáticas. Universidad de Puerto Rico.
- Medina, F. y Galván, M. 2007. "Imputación de datos: teoría y práctica", *Serie Estudios estadísticos y prospectivos*, 54. Santiago de Chile: CEPAL.
- Oñate Rubalcaba, P. 2002. "Sistema electoral y voto estratégico", en I. Crespo (coord.), *La campaña electoral de 2000: partidos, medios de comunicación y electores*, vol. 2. Madrid: Tirant Lo Blanc.
- Rao, J. N. K. y Shao, J. 2002. "Jackknife Variance Estimation with Survey Data under Hot-Deck Imputation", *Biometrika*, 79: 811-822.
- Rubin, D. B. 1976. "Inference and Missing Data", *Biometrika*, 63: 581-592.
- Rubin, D. B. 1987. *Multiple Imputation for Nonresponse in Surveys*. Nueva York: John Wiley & Sons.
- SAS Institute Inc. 2004. *SAS/STAT 9.1. User's Guide*. Cary, NC: SAS Institute Inc, EE UU.
- Schafer, J. L. 1997. *Analysis of Incomplete Multivariate Data*. Londres: Chapman & Hall.
- Schafer, J. L. 1999. "Multiple imputation: a primer", *Statistical Methods in Medical Research*, 8: 3-15.
- Yuan, Y. 2000. "Multiple Imputation for Missing Data: Concepts and New Development", *Proceedings of the Twenty-Fifth Annual SAS Users Group*. SAS Institute Inc.

## ANEXO

TABLA 1.

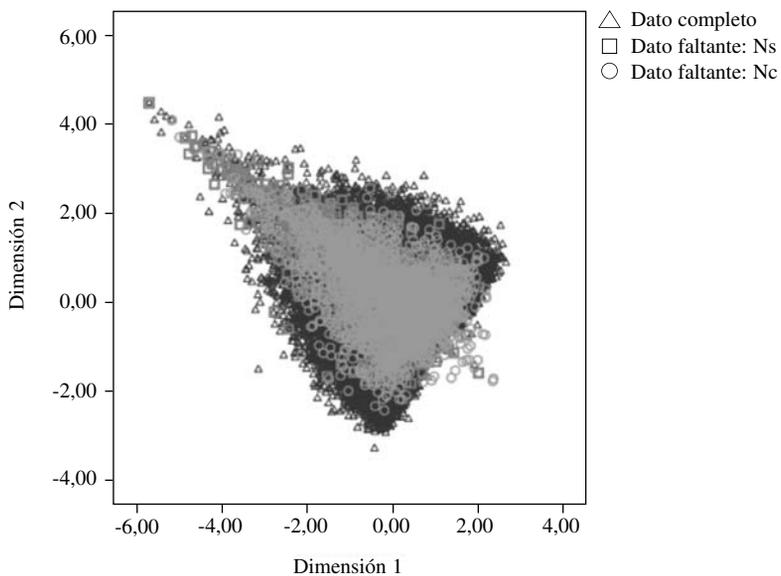
DISTRIBUCIÓN DE LOS INDIVIDUOS SEGÚN INTENCIÓN DE VOTO Y GRUPO DE PERTENENCIA

	<b>Grupo 1</b>	<b>Grupo 2</b>	<b>Grupo 3</b>	<b>Grupo 4</b>
IU	0,8	2,4	11,6	1,9
PP	78,4	15,8	2,7	39,8
PSOE	1,1	23,6	47,6	9,6
PNV	0,2	0,3	2,9	0,6
CiU	3,8	1,0	3,5	2,8
Otros partidos	2,5	3,1	10,0	6,4
En blanco	1,7	2,6	3,3	4,1
No votaría	3,7	26,2	9,3	20,1
Ns	1,6 (n=100)	11,9 (n=340)	2,5 (n=216)	5,8 (n=330)
Nc	6,2 (n=389)	13,0 (n=374)	6,9 (n=601)	8,9 (n=506)
(n)	6.243	2.869	8.703	5.688

Fuente: elaboración propia con datos del CIS (estudio 2.382).

FIGURA 1.

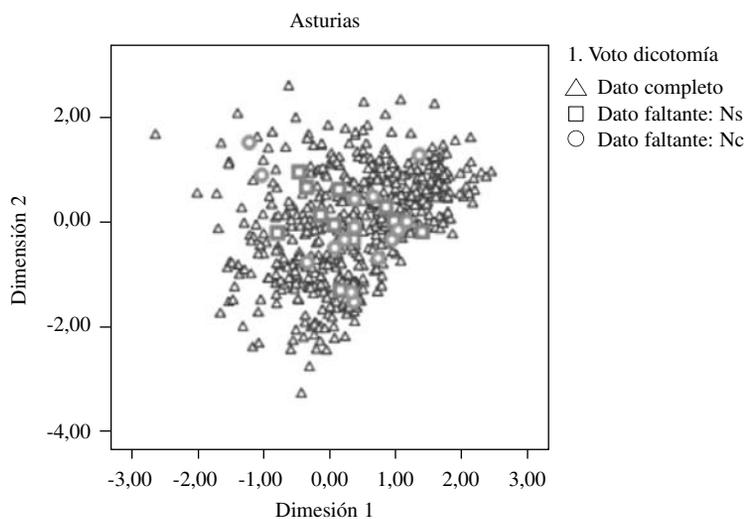
DISTRIBUCIÓN DE LOS Ns Y Nc EN EL PRIMER PLANO FACTORIAL DEL ANÁLISIS HOMALS A NIVEL NACIONAL

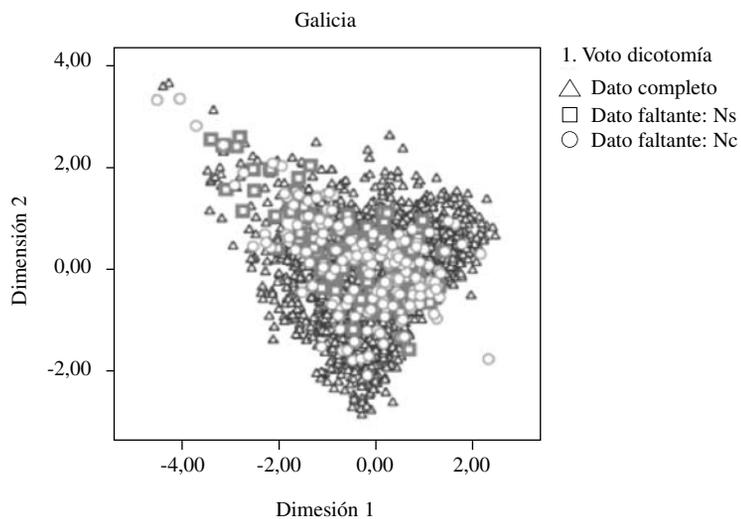
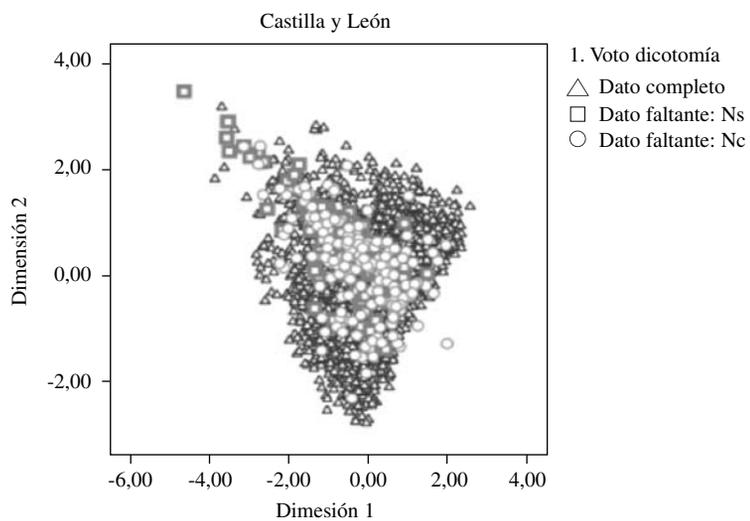


Fuente: resultados HOMALS con datos del CIS (estudio 2.382).

FIGURA 2.

DISTRIBUCIÓN DE LOS Ns Y Nc EN EL PRIMER PLANO FACTORIAL DEL ANÁLISIS HOMALS EN ASTURIAS, CASTILLA Y LEÓN Y GALICIA





Fuente: resultados HOMALS con datos del CIS (estudio 2.382).

Presentado para evaluación: 15 de enero de 2008

Aceptado para publicación: 29 de septiembre de 2009

CRISTINA RIVAS PÉREZ, Universidad de Salamanca

crisrivas@usal.es

Es máster en Estudios Latinoamericanos e investigadora del Instituto de Iberoamérica en la Universidad de Salamanca. Su principal línea de investigación es el estudio de la cultura política de la elite parlamentaria y de los ciudadanos en la región andina.

MARÍA DEL MAR MARTÍNEZ ROSÓN, Universidad de Salamanca

rososon@usal.es

Es doctora en Ciencia Política por la Universidad de Salamanca y actualmente es becaria postdoctoral del Center for Inter-American Policy and Research (CIPR) en la Universidad de Tulane (Nueva Orleans, EE UU). Su principal línea de investigación es el análisis de la carrera política de los legisladores en la región centroamericana, así como la influencia de la ambición política sobre el comportamiento y la representación.

M<sup>a</sup> PURIFICACIÓN GALINDO VILLARDÓN, Universidad de Salamanca

pgalindo@usal.es

Es profesora y directora del Departamento de Estadística en la Universidad de Salamanca. Ha sido profesora visitante en 14 universidades europeas y americanas. Su principal línea de investigación está relacionada con los métodos multivariantes en general y la visualización de datos, métodos Biplot en particular. Sus investigaciones más recientes tienen que ver con los métodos de representación visual y su aplicación a diferentes campos, incluyendo economía y ciencias sociales.