

**ASSOCIATIONISM AND ELECTORAL PARTICIPATION:
A MULTILEVEL STUDY OF 2000 SPANISH GENERAL ELECTION**

Clara Riba and Anna Cuxart

Departament d'Economia i Empresa
Universitat Pompeu Fabra

clara.riba@upf.edu, anna.cuxart@upf.edu

ABSTRACT

This work presents an application of the multilevel analysis techniques to the study of the abstention in the 2000 Spanish general election. The interest of the study is both, substantive and methodological. From the substantive point of view the article intends to explain the causes of abstention and analyze the impact of associationism on it. From the methodological point of view it is intended to analyze the interaction between individual and context with a modelisation that takes into account the hierarchical structure of data. The multilevel study of this paper validates the one level results obtained in previous analysis of the abstention and shows that only a fraction of the differences in abstention are explained by the individual characteristics of the electors. Another important fraction of these differences is due to the political and social characteristics of the context. Relating to associationism, the data suggest that individual participation in associations decrease the probability of abstention. However, better indicators are needed in order to catch more properly the effect of associationism in electoral behaviour.

Keywords: electoral abstention in Spain, multilevel modelling,

Journal of Economic Literature Classification: C89, D72

*ASOCIACIONISMO Y PARTICIPACIÓN ELECTORAL. UN ESTUDIO MULTINIVEL DE LAS ELECCIONES GENERALES ESPAÑOLAS DEL AÑO 2000**

Clara Riba y Anna Cuxart

Departament d'Economia i Empresa
Universitat Pompeu Fabra

E-mail: clara.riba@upf.edu, anna.cuxart@upf.edu

RESUMEN

Este trabajo presenta una aplicación de las técnicas de análisis multinivel al estudio de la abstención en las elecciones generales españolas del año 2000. El interés del mismo es tanto sustantivo como metodológico. Desde el punto de vista sustantivo, se pretende explicar las causas de la abstención y analizar el impacto en la misma del asociacionismo, tanto de la pertenencia individual a asociaciones como de la densidad asociativa del contexto en el que se desarrolla la competición política. Desde el punto de vista metodológico, se pretende analizar la interacción entre individuo y contexto con una modelización que tenga en cuenta la estructura jerárquica de los datos. La mayor parte de los estudios realizados hasta el momento sobre la abstención en España o bien usan datos de un solo nivel o bien combinan datos individuales y contextuales ignorando la estructura jerárquica de los mismos. Dicho enfoque puede dar lugar a interpretaciones sesgadas que con la aplicación de modelos multinivel se hubieran podido evitar. El estudio multinivel de la abstención ha permitido validar los resultados obtenidos en análisis de un solo nivel y muestra que tan sólo una parte de las diferencias en la abstención se explica por las características individuales de los electores, mientras que otra parte importante es debida a diferencias en el contexto político y social de cada circunscripción. La densidad de afiliación partidista y el umbral electoral efectivo son variables contextuales que explican casi totalmente las diferencias de participación entre provincias una vez descontada la influencia debida a características individuales. Por lo que respecta al asociacionismo, parece que la participación individual en asociaciones disminuye la probabilidad de abstención. Queda, sin embargo pendiente la utilización de indicadores más finos que permitan distinguir los efectos de la pertenencia a distintos tipos de asociaciones.

* Comunicación presentada en el VI Congreso de la asociación Española de Ciencia Política y de la Administración, en el grupo de trabajo *Capital social, asociacionismo y participación política en España*. Barcelona, 18-20 de septiembre de 2003.

INTRODUCCIÓN

Participación electoral y democracia

La participación política y, en particular, la participación electoral, es considerada como uno de los pilares de la democracia. En democracia se accede al poder político mediante la lucha competitiva por el voto de los ciudadanos¹. La competición electoral es la que permite a los partidos difundir sus propuestas entre la ciudadanía y a los ciudadanos escoger entre ellos aquél que responda mejor a sus preferencias personales. En ese sentido, las elecciones son el instrumento mediante el cual la política conecta con la voluntad de los ciudadanos.

El estudio de la participación electoral es importante básicamente por dos razones. En primer lugar porque, desde el punto de vista normativo, se considera que una elevada participación es un elemento de legitimación del sistema político². Conocer los motivos por los que la gente vota o no vota puede facilitar la implementación de acciones encaminadas a incrementar la participación.

En segundo lugar, porque desde el punto de vista empírico, el nivel de abstención es uno de los factores que influyen en los resultados electorales y, por consiguiente, es uno de los aspectos que afectan la composición del gobierno y la definición de las políticas que éste implementa. Parece claro que los diferentes grupos sociales no participan en igual proporción en las elecciones y, en consecuencia, existe una desigualdad en la representación de determinados grupos en el parlamento. En palabras de Wolfinger y Rosenstone, "voters are not a microcosm of the entire body of citizens but a distorted sample that exaggerates the size of some groups and minimises that of the others" (Wolfinger & Rosenstone, 1980: 108).

La naturaleza social de la participación

La abundante literatura existente acerca de la participación política y electoral ha desarrollado distintas explicaciones y modelos de voto. Los modelos de carácter

¹ Véase en Schumpeter (1945) la concepción de democracia como competencia política por el poder.

² Aunque hoy en día parece existir un consenso generalizado sobre ese punto, los defensores de la teoría competitiva de la democracia consideran que una baja participación no es necesariamente mala ya que puede ser un reflejo de la confianza de los electores en el sistema político existente. Véase, por ejemplo, Dahl (1956) y Almond y Verba (1963). Sin embargo, tal como apunta Ranney (1983) la abstención puede ser también consecuencia de la apatía, la alineación o la insatisfacción y, por tanto, ser considerada como indicativa de un mal funcionamiento del sistema democrático.

sociológico vinculan la participación a las características socioestructurales de los electores. El modelo socio-psicológico atribuye la decisión de voto a variables de tipo psicológico, tales como la identificación partidista o el sentimiento de eficacia política, cuya conformación se habría realizado durante la época de socialización del individuo. El modelo de la elección racional considera que el elector, ante cada contienda electoral, evalúa los costes y beneficios que le reporta la victoria de cada una de las fuerzas políticas y decide votar o abstenerse en función del resultado. Si decide votar, lo hace por la opción que le reporta una mayor utilidad. Se abstiene cuando los costes superan a los beneficios esperados (*alienación*) o cuando considera varias opciones como equivalentes (*indiferencia*).

No obstante, ninguna de estas explicaciones resulta totalmente satisfactoria. Los modelos sociológicos son incapaces de dar cuenta de las variaciones de participación a lo largo del tiempo de un electorado que no ha cambiado sustancialmente sus características. Los modelos psicológicos, basados en identificaciones relativamente estables, no consiguen explicar las fluctuaciones de participación en elecciones de distinta índole. Finalmente, la paradoja del voto, según la cual la casi nula probabilidad de que un voto sea decisivo en el resultado llevaría a todos los electores a abstenerse racionalmente, pone en cuestión la validez explicativa de los modelos racionales.

¿Por qué sigue habiendo tantos ciudadanos que votan si son conscientes de la probabilidad de que su voto influya en el resultado es mínima? Rosenstone y Hansen (1993) argumentan que la participación política no puede ser explicada totalmente por las características y orientaciones de los ciudadanos, sino que debe tomarse en consideración la naturaleza social de la vida política. En ese sentido, los beneficios que los individuos obtienen de participar son dos tipos: Los derivados del resultado de la elección, que son *beneficios colectivos* asociados con bienes de naturaleza pública, y los llamados *beneficios selectivos* que se obtienen por el propio hecho de participar. Los beneficios selectivos son de naturaleza privada y entre ellos cabe citar la amistad, la camaradería, el sentirse aceptado y valorado o la satisfacción de cumplir con un deber cívico³.

En este trabajo se considera que, en principio, cualquier tipo de participación conlleva unos ciertos costes y genera la expectativa de unos determinados beneficios. En cierto sentido, el acto de votar es el resultado final de un proceso en el que el elector, en función de sus características, intereses y motivaciones, y en función también de los estímulos que recibe de su entorno, valora los pros y contras y decide votar por una

³ Para una discusión sobre las consecuencias que tiene para la actividad política la participación en asociaciones productoras de bienes públicos y en asociaciones productoras de bienes privados, véase Boix y Posner (2000).

opción política determinada o bien abstenerse cuando los costes son superiores a los beneficios esperados⁴. Desde esta perspectiva, cualquier hecho o circunstancia que incremente los beneficios o disminuya los costes incidirá, pues, en una mayor participación.

La naturaleza social de la participación política que destacan Rosenstone y Hansen (1993) pone de manifiesto la importancia de las redes sociales en la decisión de voto. Éstas actúan en una doble dirección. Por un lado, las redes sociales al facilitar la comunicación reducen los costes de información. Por otro, la pertenencia a redes sociales proporciona al individuo un aumento de los llamados beneficios selectivos. Ambos factores, disminución de costes y aumento de beneficios, aportan elementos para la superación de la paradoja del voto y facilitan la participación electoral.

Interés y objetivos del estudio

Tomar en consideración el entorno social del individuo requiere incorporar el contexto en los modelos de participación política. A pesar de que en los últimos tiempos los modelos de análisis que gozan de mayor consenso son aquellos que presuponen interacción entre factores de diferentes niveles (individual y agregado), los métodos estadísticos que tratan este tipo de información son todavía de uso poco generalizado en nuestro país. Véanse, por ejemplo, las diversas estrategias seguidas por Anduiza (1999), Boix y Riba (2000), Morales (2001) o Barreiro (2002) para combinar el efecto del contexto con las características individuales en sus respectivos trabajos sobre participación política. Así, Boix y Riba estudian la abstención en las elecciones generales en España entre 1982 y 1996 añadiendo al fichero de datos individuales una serie de variables de nivel provincial con valores constantes para todos los individuos de una misma provincia; Barreiro utiliza una estrategia similar en su análisis de la abstención en las elecciones generales españolas incorporando datos de las elecciones del año 2000. Anduiza analiza el fenómeno de la abstención en Europa occidental, planteando el análisis de la interacción entre efectos individuales y contextuales mediante la misma técnica que Boix y Riba, aunque proponiendo la utilización de modelos multinivel sin llegar a hacer uso de los mismos. El trabajo de Morales es una de las primeras contribuciones de autores españoles en las que se utiliza la modelización multinivel, pero no estudia la

⁴ Aquí se adopta una perspectiva un poco más amplia que la de la elección racional. La teoría de la elección racional considera el voto como el resultado de un cálculo coste-beneficio estrictamente individual y no toma en consideración los efectos contextuales. Véase Downs (1957) para las bases de los modelos racionales de voto y Riker y Ordeshook (1968) para el desarrollo de la teoría sobre el Cálculo de Votar.

abstención sino otro tipo de participación política (ser miembro de una organización política en la Unión Europea).

La técnica de la modelización multinivel está siendo reconocida como una buena manera de analizar la interacción entre variables de diferentes niveles y, en consecuencia, de tratar de un modo adecuado los datos que tienen una estructura jerárquica. Actualmente, es de amplia difusión en diversas áreas de investigación, como la educación o la epidemiología (Aitkin y Longford, 1986; Goldstein, 1991 y 1995; Byrk y Raudenbush, 1992), pero todavía escasamente aplicada en ciencia política. En esta área cabe destacar en el ámbito internacional los artículos de Jones, Johnston and Pattie (1992) y Steenbergen and Jones (2002).

El trabajo que se presenta a continuación aplica la metodología multinivel al estudio de la abstención en las elecciones generales españolas del año 2000. Se consideran dos fuentes de información, la individual y la contextual, y se analizan los datos con una técnica más idónea que la de la regresión a un solo nivel que se aplica en los estudios citados anteriormente⁵. Los objetivos específicos de este trabajo combinan el interés sustantivo por explicar la abstención con la preocupación por los aspectos metodológicos subyacentes en los análisis. Concretamente, son los siguientes:

1. Comprobar si los resultados sobre las causas de la abstención que se desprenden de la literatura sobre el comportamiento electoral se mantienen al aplicar la modelización multinivel (que incorpora una descomposición de la variación más completa que la de los modelos de regresión de un solo nivel).
2. Estudiar qué papel juega el asociacionismo en la participación electoral, tanto la pertenencia individual a asociaciones como la densidad asociativa del contexto en el que vive el individuo.
3. Comprobar si las diferencias de abstención que se observan entre circunscripciones tienen una componente atribuible a las características de sus respectivas comunidades autónomas.
4. Contrastar si los efectos de las variables individuales en la abstención difieren en función del contexto. Es decir, si los valores de los coeficientes de los modelos de regresión dependen del entorno político y social en el que el elector se desenvuelve.

⁵ Si bien es cierto que el efecto de contexto también se puede modelizar a través de variables dummies, dicho enfoque se basa en el supuesto poco realista de que las diferencias entre distintos contextos quedan totalmente explicadas por dichas variables. Los modelos multinivel son más adecuados ya que son más flexibles al admitir la existencia de heterogeneidad entre las unidades contextuales no explicada por las variables del modelo.

DATOS Y VARIABLES

Fuentes de datos

Los datos individuales utilizados en este trabajo provienen de las encuestas pre i post electoral de las elecciones generales de marzo del 2000 realizadas por el Centro de Investigaciones Sociológicas (CIS), estudio PANEL2382_2384. Aunque el número total de casos del panel es de 24.040, el estudio utiliza sólo los 5.283 que fueron incluidos en la post electoral.

Se trata de una muestra estratificada por comunidades autónomas y tamaño de hábitat. Ello da sentido tanto al estudio de la relación entre variables, principal objetivo de este trabajo, como también a la comparación de las variaciones de abstención entre comunidades y/o provincias.

El fichero de datos ha sido completado con información provincial y autonómica procedente de diferentes fuentes de datos agregados, entre los que cabe destacar el Instituto Nacional de Estadística, las Webs de partidos políticos y la Junta Electoral.

En las elecciones generales españolas del año 2000, se aprecia la existencia de diferencias provinciales entre la abstención real (medida como porcentaje de votantes sobre el censo electoral) y la abstención declarada (medida como el porcentaje de encuestados que declaran haber votado), en el sentido de que la abstención real supera a la declarada. La infraestimación de la abstención es un resultado conocido y frecuentemente analizado en la literatura sobre el tema⁶. Las razones que se esgrimen para justificarla son de dos tipos. Por un lado, existen explicaciones de carácter técnico, como deficiencias en la actualización del censo, errores de muestreo o errores derivados de un diseño deficiente de las encuestas, y, por otro, explicaciones relativas al comportamiento de los encuestados, como por ejemplo la tendencia a declarar que se ha votado por la opción ganadora, las resistencias a reconocer comportamientos socialmente no aceptados, las no respuestas o las deficiencias memorísticas. En el presente estudio esta subestimación es importante: En todas las provincias, a excepción de Ávila, la abstención real es superior a la declarada en el sondeo, siendo la diferencia superior a 15 puntos en 29 de ellas⁷.

⁶ Por ejemplo, Anduiza (1999), en un estudio sobre quince países europeos, estima una diferencia superior a 10 puntos entre el porcentaje de abstención real y el declarado en la encuesta.

⁷ La excepción de Ávila se podría explicar por la escasez de datos correspondientes a esta provincia en la muestra (24), hecho que conduce a una estimación muy poco precisa.

Imputación de valores faltantes

Puede decirse que, en promedio, casi la mitad de los encuestados en un sondeo dejan sin responder al menos una de las preguntas del mismo. Ante esta situación, una práctica habitual, conocida como *listwise deletion*, consiste en dejar fuera del análisis aquellos registros que contengan algún valor faltante. Otra práctica menos habitual consiste en la imputación simple de valores concretos a las no respuestas. Esta imputación puede hacerse aplicando criterios diversos. Por ejemplo, se imputa el valor de la media o de la mediana en el caso de variables numéricas o el de la moda en el caso de variables categóricas. En ocasiones se combina la información procedente de varias variables para formar tipologías de ciudadanos y predecir la respuesta más probable para cada tipología.

Ninguna de las dos alternativas, ni *listwise deletion* ni imputación simple, es del todo satisfactoria. Como ya se ha indicado, la aplicación de *listwise deletion* puede comportar la exclusión de un número considerable de casos. Ello es especialmente grave en los estudios electorales donde las no respuestas son abundantes debido a la privacidad de algunos de los temas consultados y donde el tipo de análisis aplicado, que involucra un gran número de variables, magnifica el problema. En el mejor de los escenarios, cuando los registros eliminados estuvieran repartidos aleatoriamente y tuvieran las mismas características que el resto de la muestra, *listwise deletion* sólo generaría una pérdida de información valiosa y, en consecuencia, estimaciones poco precisas. Pero esta situación ideal es poco frecuente y la aplicación de *listwise deletion* suele conducir a que la parte de la muestra con la que finalmente se realiza el análisis contenga un cierto sesgo⁸. En resumen, en muchas ocasiones, con esta práctica a la pérdida de la información hay que añadir la introducción de un sesgo de selección.

Por lo que respecta a la segunda alternativa, la imputación simple de datos faltantes no puede realizarse mecánicamente sino que requiere la toma en consideración aspectos de tipo sustantivo y de tipo metodológico. Entre los aspectos sustantivos, se debe considerar como un requisito el disponer de un buen conocimiento previo del tema de estudio con el objetivo de definir criterios de imputación solidamente sustentados.

⁸ La terminología estándar distingue entre tres tipos de procesos que conducen a la existencia de valores faltantes: Procesos *Missing Completely At Random* (MCAR), *Missing At Random* (MAR) y *NoIgnorable* (NI) según que los valores faltantes estén distribuidos completamente al azar, sean predecibles a partir de los valores observados o dependan también del valor de los datos no observados o *missings*, respectivamente (Sherman, R. P., 2000).

En cuanto a las consecuencias de carácter metodológico, se pueden distinguir las que hacen referencia a una sola variable y las que se refieren a la relación entre varias. En primer lugar, es un hecho conocido que la concentración de datos en los valores centrales de la distribución, reduce la variabilidad de los estadísticos. Por lo tanto, la imputación simple de datos en una variable reduce su varianza. Por lo que respecta a la relación entre variables en un modelo de regresión, los efectos son distintos si se imputan valores en variables explicativas o en la variable dependiente. Cuando el proceso de imputación se aplica solamente a las variables explicativas, se obtiene una mayor precisión en los intervalos de predicción. Sin embargo, la imputación realizada sobre la variable dependiente, da lugar a una reducción de los errores estándar de los coeficientes de la regresión, con las consecuencias que ello puede tener en la significación de los coeficientes de las variables (que pueden aparecer como significativas cuando en realidad no lo son). En resumen, la imputación simple de datos faltantes permite mantener al máximo la información de la encuesta pero puede introducir un sesgo si la imputación se realiza sin una buena base teórica. Incluso si los criterios de imputación para las no respuestas son en términos generales correctos, el procedimiento sobreestima el grado de certidumbre de las respuestas. En consecuencia, los errores estándar serán demasiado pequeños (King *et al.*, 2001). Por lo tanto, la imputación puede llevar a conclusiones erróneas sobre la significación de las relaciones entre variables.

La imputación de datos faltantes es un tema actual de investigación en estadística. En ese sentido es de destacar el desarrollo de las técnicas conocidas como de *imputación múltiple*. Referencias claves sobre este tema son: Rubin (1987) y Little y Rubin (1987). En el ámbito específico de la ciencia política debe mencionarse las aportaciones de Gary King *et al.* (2001). Sin embargo, la complejidad del proceso de implementación de la *imputación múltiple* y la dificultad de disponer de software comercial adecuado han impedido hasta el momento su uso generalizado.

En el fichero de datos de este estudio se detectaron inicialmente 509 de los 5283 registros con algún dato faltante, 52 de los cuales eran no respuestas a la pregunta sobre participación en las elecciones generales del 2000. Para evitar la pérdida de un número considerable de registros, se procedió a una imputación simple, diferenciando el proceso según el rol de la variable en el análisis. Para las variables independientes se imputó la mediana o la media según el tipo de variable, quedando 335 casos sin respuesta en alguna de las variables explicativas, correspondiendo la mayor parte de ellos a no respuestas en la pregunta sobre ocupación del encuestado.

Para la variable dependiente, en cambio, se procedió a una imputación en dos fases. En primer lugar, mediante la información de otras dos variables de la misma encuesta post electoral se imputó el valor de 10 casos para los que se disponía de un criterio

sólido. En una segunda fase se combinó el criterio anterior con la información procedente de la encuesta pre electoral. Ello permitió la imputación de 41 casos más, quedando finalmente solo dos casos sin respuesta en la variable dependiente⁹.

Variables

Aceptando, pues, la naturaleza social de la participación, las variables que interesan para estudiar la abstención electoral serán, además de las que han mostrado su capacidad explicativa en los modelos de voto contrastados en la literatura, aquellas que reflejen las relaciones e intereses sociales del individuo y la existencia de entramados sociales en el contexto donde se desenvuelve su actividad.

La variable dependiente de nuestro estudio es la abstención electoral, en sus dos vertientes, individual y agregada. En cuanto a las variables independientes, se podría clasificar al conjunto de factores que la teoría acepta que inciden en la participación en dos grandes grupos: los factores de tipo socioestructural y los factores de tipo político. En un grupo aparte se consideran los condicionantes institucionales que, si bien son de tipo político, no dependen de la actuación de los partidos y electores sino que constituyen las reglas del juego a las que todos los actores deben someterse.

Por otra parte, podemos hacer una segunda distinción de los factores que inciden en la participación según que éstos sean de tipo individual o contextual. En efecto, las particularidades individuales de los ciudadanos, ya sean sus características

⁹ Para aquellos encuestados que no respondieron cuando se les preguntó si fueron a votar en las elecciones del 12 de marzo del 2000, se ha utilizado la información contenida en la respuesta dada a otras dos preguntas que inquirían sobre si el individuo tenía ya decidida su participación o bien si dudaba y, en caso de duda, se pedía entre qué opciones dudaba. Nuestra opción fue considerar como abstencionistas a los individuos que respondieron que tenían decidido abstenerse; como votantes a los que respondieron que habían decidido el voto a un partido o coalición, y para el único que manifestó dudar entre voto y abstención, se tuvo en cuenta la respuesta a la segunda pregunta (donde manifestaba dudar entre voto en blanco y abstención), clasificándolo como abstencionista. Para la segunda fase de imputación de la variable dependiente se utilizó la información de las variables sobre intención de voto existentes en la encuesta preelectoral, donde se preguntaba sobre la intención de voto y, en caso de no respuesta, sobre la simpatía por una opción política. La estrategia consistió en combinar la información contenida en estas dos preguntas de la encuesta preelectoral con la de abstención declarada en la encuesta postelectoral. Para ello se estimaron los porcentajes de voto a fuerzas políticas y de abstención (calculados a partir de la encuesta *post*) para cada una de las categorías resultantes del cruce de las dos variables de intención de voto y simpatía de la encuesta *pre*). A cada uno de los restantes individuos sin valor en la variable dependiente se le imputó el valor más probable de su tipología. Los dos únicos casos sin respuesta corresponden a una tipología donde el porcentaje de abstencionistas era del 50%.

sociestructurales, sus predisposiciones, sus actividades o sus valoraciones políticas, hacen que unos tengan una mayor propensión a votar que otros. Sin embargo, un mismo individuo tendrá una mayor o menor probabilidad de votar en función de las características de su entorno y de la densidad del entramado social existente en el mismo. Distinguiremos, pues, entre factores individuales característicos de cada elector y factores contextuales propios del entorno donde se desarrolla la competición política.

	Socio estructurales	Políticas	Institucionales
Individuales	Tamaño de municipio Edad Nivel de estudios Situación laboral Ocupación	Extremismo político Interés por la política Frecuencia de conversaciones políticas Seguimiento de información política Grado de asociacionismo Satisfacción con el fun. de la democracia Valoración de la situación política Valoración de la situación económica Grado de confianza en los partidos Sentimiento de eficacia política	
Contextuales	Sectores de ocupación Nivel de renta	Densidad de afiliación a partidos Densidad asociativa Competitividad electoral	Número de escaños Número efectivo de partidos Umbral electoral efectivo

Tabla 1: Variables que afectan la participación electoral

Como resultado de esta doble clasificación, aparecen cinco grandes bloques de factores que inciden en la participación electoral: Las variables socioestructurales y las variables políticas, en sus dos niveles, individual y contextual, más las variables institucionales. En la Tabla 1 se indican las variables explicativas de la abstención que se han tomado en consideración en este estudio clasificadas en los cinco mencionados grupos. En algunos casos se han construido índices a partir de análisis factoriales. En el Anexo 1 se detallan los indicadores utilizados para operacionalizar cada una de estas variables.

Debe mencionarse aquí las dificultades halladas para definir variables que operacionalicen la riqueza de la vida social del individuo, en sus dos vertientes individual y agregada. En realidad, se trata de un conjunto de variables que conforman el llamado *capital social* (Putnam *et al.*, 1993). Pero el concepto de capital social no es simple ni tiene una única interpretación. Precisamente, la complejidad del concepto ha llevado a algunos investigadores a rechazar su utilización por

considerarlo un concepto “paraguas” bajo el que se cobijan varias dimensiones distintas. Las diferentes operacionalizaciones que se han propuesto para el mismo ponen de manifiesto esta multidimensionalidad del concepto (Torcal y Montero, 2000; Mota y Subirats, 2000).

En ese trabajo no se pretende entrar en la discusión sobre el concepto de capital social y sus posibles indicadores. La pretensión de las autoras es tratar de ver la influencia en la participación política (específicamente en la participación electoral) de algunas de las variables relacionadas con el concepto de capital social y, en particular, del asociacionismo. En concreto, se ha estudiado el efecto de las variables individuales *interés por la política*, *frecuencia de conversaciones políticas*, *seguimiento de información política* y *grado de asociacionismo* y las variables agregadas *densidad de afiliación a partidos* y *densidad asociativa*.

En el ámbito individual, para medir el asociacionismo, se ha construido una variable que indica el número total de tipos de asociaciones a las que pertenece el individuo. Se trata de un indicador poco fino puesto que trata del mismo modo todo tipo de pertenencia a asociaciones o entidades, cuando probablemente éstas no tengan efectos equivalentes. En ese sentido, puede ser importante la contribución de Méndez y Mota en este mismo grupo de trabajo que tiene por objetivo establecer una tipología de asociaciones en función de sus características, áreas de actividad y fines (Méndez y Mota, 2003). Posiblemente, las asociaciones pertenecientes a algunos de los grupos de esta tipología tengan una mayor capacidad explicativa que otras en el estudio de la abstención. En este trabajo las autoras han tratado de usar una variable individual que indique la militancia en partidos políticos o sindicatos, pero el intento no ha resultado positivo debido al reducido número de personas afiliadas a los mismos. En consecuencia se ha optado por un índice más amplio que incorpora la pertenencia a cualquier tipo de asociación¹⁰.

En el ámbito agregado, se ha podido utilizar el número de afiliados a los partidos políticos como medida de *densidad de asociacionismo político*. El indicador ideal para medir la densidad de afiliación a partidos habría sido una tasa de afiliación política (a cualquier partido). Ante la imposibilidad de obtener esta información, se ha optado por usar como indicador alternativo la tasa de afiliación a los dos principales partidos para los cuales sí disponemos de datos (publicados en su página Web por el PP y en la prensa por el PSOE). Esta variable, sin embargo, tiene el problema de tomar valores mucho más bajos que los que corresponderían a la afiliación política realmente existente en aquellas provincias donde existen partidos

¹⁰ Cabe citar que el indicador utilizado es de fácil interpretación y muestra una correlación altísima ($r = 0,97$) con el primer factor resultante de un análisis factorial de las correspondientes variables primarias.

de ámbito no estatal bien implantados, como ocurre en Cataluña y País Vasco. Al mismo tiempo, ante la dificultad para hallar datos de participación en asociaciones no políticas en el ámbito provincial, se ha optado por calcular un índice de asociacionismo provincial a partir de la agregación de los datos individuales de la encuesta. Finalmente, no ha sido posible hallar datos actualizados de afiliación sindical por provincias.

METODOLOGÍA

En este trabajo se ha estudiado la relación entre las variables explicativas y la abstención a través del análisis multinivel. El motivo para la aplicación de esta metodología se halla en la estructura de los datos y en el interés en estudiar la interacción del contexto con el comportamiento individual. Nuestros datos presentan una clara estructura jerárquica: individuos (unidades de primer nivel) agrupados en provincias (unidades de segundo nivel) que, a su vez, podrían agruparse en comunidades autónomas o unidades de tercer nivel. Tenemos, por lo tanto, una estructura de datos en tres niveles.

Steenberger y Jones (2002) destacan dos razones que aconsejan un análisis multinivel en situaciones similares. En primer lugar, el hecho de disponer de datos multinivel permite la combinación de diferentes niveles de análisis en un único modelo de variación para la variable respuesta y en el cual se especifican las variables predictoras de cada nivel. En segundo lugar, el análisis multinivel permite la investigación de la heterogeneidad causal, es decir, la posibilidad de interacción entre variables de diferentes niveles.

En el trabajo se utilizarán dos tipos de modelos: los llamados modelos de descomposición de la varianza y los modelos de regresión logística multinivel. El modelo más simple de descomposición de la varianza es el modelo de análisis de la varianza de efectos aleatorios y sin variables explicativas -llamado *null model* o *empty model* en la literatura anglosajona. La importancia de dicho modelo estriba en que ofrece una descomposición de la varianza de la variable respuesta en los dos o más niveles de variación posibles (Searle, Casella, y McCulloch, 1991; Longford, 1994). Por ejemplo, en el caso de individuos y provincias, cada observación individual (y_{ij}) se puede expresar como la suma de la media global (μ) más la desviación de la provincia respecto de esta media (efecto provincial u_j), más la desviación del individuo respecto de su provincia (efecto individual e_{ij}). Es decir, el residuo habitual de los modelos de un solo nivel se descompone en suma de dos

residuos independientes, el individual y el correspondiente a la provincia a la cual pertenece el individuo:

$$y_{ij} = \mu + u_j + e_{ij}$$

En este modelo μ representa la “verdadera” media poblacional, la que se obtendría a partir de todos los individuos de todos los grupos (en este caso, provincias), mientras que $\mu + u_j$ representa la “verdadera” media del grupo (provincia) j . Las distribuciones de probabilidad de las variables aleatorias u_j y e_{ij} se suponen independientes, de media cero y varianza σ_u^2 y σ_e^2 , respectivamente. En consecuencia, la varianza total de la variable respuesta es igual a la suma de las dos varianzas anteriores: $\sigma_u^2 + \sigma_e^2$. Este modelo permite estimar el llamado coeficiente de correlación *intra* (ρ) igual al cociente de la variación entre provincias y la variación total. Este coeficiente mide el porcentaje de variación total debido a las diferencias entre provincias, tomando valores entre 0 y 1. Cuanto más homogéneos sean los grupos (provincias) y más diferentes entre sí, mayor será el coeficiente de correlación *intra*.

Como ilustración, la Tabla 2 muestra la descomposición de la varianza para seis de las variables continuas utilizadas en este estudio, incluyendo la estimación del correspondiente coeficiente de correlación *intra*. En la tabla se observa que aproximadamente un 10% de la variación de la variable que mide la satisfacción con el funcionamiento de la democracia corresponde a diferencias de opinión entre provincias ($\rho = 0,096$)¹¹. En el otro extremo, la variable edad no presenta diferencias significativas entre provincias ($\rho = 0,002$). El resto de variables presentan una diferenciación provincial leve.

Si en el modelo de regresión lineal clásico (de un solo nivel) se admite la posibilidad de que los coeficientes (ordenada en el origen y pendiente) varíen entre unidades de segundo nivel según una distribución de probabilidad, se obtiene el modelo de regresión multinivel¹², también llamado de *coeficientes aleatorios*. En particular, la modelización de estos coeficientes mediante la dependencia de variables de segundo nivel, permite el estudio de la interacción entre variables de ambos niveles.

¹¹ La experiencia del cálculo del coeficiente de correlación *intra* en el ámbito de educación, por ejemplo el estudio de los resultados de exámenes de estudiantes de diferentes centros, proporciona valores de coeficiente de correlación *intra* por debajo de 0,2, siendo en la mayoría de casos cercanos a 0,1. En este sentido, un valor de ρ cercano a 0,1 debe considerarse no despreciable.

¹² Para más detalle sobre la formulación de estos modelos, véase Anexo 2.

Así mismo, si en el modelo de regresión logística clásico (de un solo nivel), se relaja la restricción de coeficientes fijos en la parte lineal del modelo, se obtiene el modelo logístico de regresión multinivel (Anexo 2).

	Satisfacción con funcionamiento de la democracia	Valoración situación política	Grado de confianza en los partidos	Sentimiento de eficacia política	Frecuencia de conversaciones políticas	Edad
<i>Varianza entre provincias</i> σ_u^2	.041	.034	.031	.031	.024	.548
<i>Varianza entre individuos</i> σ_e^2	.387	.502	.565	.673	.854	337.737
<i>Coficiente Correlación intra ρ</i>	.096	.063	.052	.046	.027	.002

Tabla 2: Descomposición de la varianza para las variables explicativas continuas

ANÁLISIS

La abstención, las provincias y las comunidades autónomas

Nuestro interés inicial era ver cómo la variación total de la abstención se descomponía en una parte propia de las características individuales y una parte debida al contexto, que aquí se explicita por la pertenencia a una determinada provincia y la pertenencia a una determinada comunidad autónoma.

En primer lugar se ha estimado la variación entre comunidades (Modelo I en la Tabla 3). Después, se ha estimado la variación entre comunidades y entre provincias dentro de las comunidades (Modelo II). De la lectura del modelo II parece desprenderse que la mayor parte de la variación de la abstención se debe realmente a la pertenencia a distintas comunidades autónomas, ya que la variación entre provincias estimada es no significativa. Ello tiene sentido teórico si se tiene en cuenta que existen partidos de ámbito no estatal que tienen especial incidencia en determinadas comunidades, especialmente en las llamadas comunidades históricas. Ahora bien, los resultados de

la estimación de ambos modelos no justifican la utilización del modelo más complejo, el de tres niveles, puesto que la mejora del estadístico que se utiliza para comparar modelos (*-2log likelihood*) es insignificante. Si se prescinde del tercer nivel, la variación provincial sí ha resultado ser significativa (Modelo III). Estos resultados llevaban a realizar el estudio en dos niveles, pudiendo escoger como segundo nivel las provincias o bien las Comunidades. Se ha optado por realizar el estudio a partir del modelo de provincias e individuos. Dos han sido las razones que justifican esta decisión. Desde el punto de vista teórico, interesa el análisis por provincias puesto que en España las provincias coinciden con las circunscripciones electorales. Desde el punto de vista técnico, tomando en consideración las provincias y no las comunidades como unidades de segundo nivel, es mayor el número de unidades disponibles (42 en lugar de 17) con la consiguiente mejora de la precisión en las estimaciones. En resumen, se tomó como modelo de partida el Modelo III, que considera dos niveles de variación: individuos y provincias.

	Modelo I (entre CCAA)	Modelo II (entre provincias y CCAA)	Modelo III (entre provincias)
<i>Efectos fijos</i>			
Ordenada en el origen	-1.657 (.105)	-1.658 (.108)	-1.687 (.077)
<i>Efectos aleatorios</i>			
Varianza de tercer nivel: entre comunidades, σ_v^2	.143 (.063)	.135 (.069)	
Varianza de segundo nivel: entre provincias, σ_u^2		.036 (.033)	.175 (.057)
<i>-2log likelihood</i>	4071.28	4068.58	4078.06

Tabla 3: Modelos logísticos de variación de la abstención sin variables explicativas

Factores explicativos de la abstención

Una vez tomada la opción de estimar un modelo explicativo de la abstención en dos niveles, provincias e individuos, se ha procedido a la estimación de un modelo logístico con las variables independientes especificadas anteriormente. En primer lugar, se ha introducido el bloque de variables socioestructurales. A continuación se

han añadido el resto de variables individuales, las de tipo político, y, finalmente, se ha completado el modelo con las variables contextuales.

	Modelo 1	Modelo 2	Modelo 3	Modelo 4
<i>Efectos fijos</i>				
Constante	-1,927 (,149)	-3,819 (,223)	-3,499 (,278)	-3,527 (,271)
Edad	-,036 (,003)	-,034 (,003)	-,035 (,003)	-,035 (,003)
Tamaño municipio	,117 (,027)	,121 (,027)	,121 (,027)	,117 (,027)
En paro	,337 (,129)			
Jubilado	,458 (,140)	,332 (,143)	,355 (,144)	,357 (,145)
Directivo	-,859 (,356)	-,738 (,365)	-,747 (,362)	-,742 (,363)
Nivel de estudios	-,124 (,032)			
Satisfacción con el funcionamiento de la democracia		,253 (,069)	,248 (,069)	,250 (,069)
Valoración de la situación política		,301 (,064)	,294 (,064)	,297 (,064)
Frecuencia de conversaciones políticas		-,457 (,050)	-,463 (,050)	-,465 (,050)
Grado de confianza en los partidos		-,243 (,050)	-,243 (,065)	-,241 (,065)
Sentimiento de eficacia política		-,172 (,061)	-,164 (,061)	-,165 (,061)
Umbral electoral efectivo			,028 (,011)	,030 (,010)
Densidad de afiliación a partidos			-,021 (,004)	-0,021 (,004)
<i>Efectos aleatorios</i>				
Varianza entre provincias	,216 (,067)	,086 (,038)	,009 (,015)	
<i>-2log likelihood</i>	3.665,06	2.795,24	2.758,34	2.754,01

Tabla 4. Modelos logísticos multinivel de predicción de la abstención

Después de una primera estimación exploratoria, se ha reestimado cada uno de los modelos incluyendo solamente aquellas variables que habían resultado ser significativas previamente. Los modelos finales son los que se reproducen en la Tabla 4.

El Modelo 1 confirma las relaciones ya establecidas en la teoría sobre comportamiento electoral entre las variables socioestructurales y la abstención. Así, al aumentar la edad disminuye la probabilidad de abstención; la abstención es mayor en zonas densamente pobladas; los parados y jubilados tienen mayor probabilidad de abstenerse que los que trabajan; los directivos tienen menor probabilidad de abstenerse que los trabajadores cualificados y semicualificados; y a mayor nivel de estudios, menor probabilidad de abstenerse. Los resultados, concuerdan, pues, con los que se obtienen en estudios de un solo nivel (individual).

El Modelo 2 resulta de añadir al modelo anterior las variables individuales de tipo político. Los resultados también coinciden con los que se esperarían desde el punto de vista teórico. Así, aquellos individuos que tienen un elevado sentimiento de eficacia política, así como los que más confían en los partidos, se abstendrán con menor probabilidad. También la frecuencia de conversaciones políticas favorece la participación. Por otra parte, cuanto peor es la valoración de la situación política o del funcionamiento de la democracia, más alta es la probabilidad de abstenerse. Cabe destacar que, con la introducción de las variables políticas en el modelo, han dejado de ser significativas la situación laboral y el nivel de estudios del encuestado.

Del Modelo 3, obtenido al añadir al modelo anterior las variables contextuales que han resultado ser significativas, se deduce que a mayor densidad de afiliación a partidos políticos, menor probabilidad de abstención, y que a mayor umbral electoral efectivo (es decir, a mayor porcentaje de votos necesarios para obtener un escaño), mayor es la probabilidad de abstenerse. Ambos resultados son coherentes con lo que sería de esperar desde el punto de vista teórico.

La comparación de los valores del estadístico $-2\log \textit{likelihood}$ para los distintos modelos muestra que el ajuste es cada vez mejor¹³. También puede observarse que la introducción de las variables de contexto en el modelo 3 hace disminuir la varianza entre provincias convirtiéndola en no significativa. Ello confirma que las dos variables contextuales del modelo explican prácticamente la totalidad de las

¹³ No obstante hay que tomar este estadístico con precaución puesto que en el caso de la regresión logística tiene sus limitaciones. En ese sentido, Snijders y Bosker (2000) recomiendan calcular también otros estadísticos que midan la bondad del ajuste.

diferencias provinciales en la abstención una vez que se ha descontado la influencia debida a las características individuales.

En los modelos anteriores se ha supuesto que los coeficientes de las variables eran constantes para todas las provincias. En análisis posteriores se relajó este supuesto, asumiendo la posible heterogeneidad en la relación causal que liga cada una de las variables con la abstención. De dichos análisis no se desprende evidencia empírica suficiente que sugiera rechazar los modelos de coeficientes constantes en favor de los de coeficientes aleatorios. Es decir, la relación causal parece ser constante a través de las provincias.

A efectos de comparación, se ha reproducido en la misma Tabla 4, bajo el título de Modelo 4, los resultados de la estimación del modelo de un solo nivel que incorpora las mismas variables explicativas que el modelo 3. Se puede observar que los resultados de las estimaciones de estos dos modelos son prácticamente iguales.

¿Qué ventaja tiene, pues, haber utilizado la modelización multinivel para este estudio? En primer lugar, aunque en este caso los resultados son equivalentes, no siempre es así. Si no se hubiera estimado el modelo en dos niveles no se tendría la misma confianza que se tiene ahora en la capacidad predictiva de las variables explicativas del modelo, puesto que es un hecho conocido que ignorar la estructura multinivel conlleva una subestimación de los errores estándar. En la tabla se puede observar que la subestimación en este caso es inapreciable. En segundo lugar, las sucesivas estimaciones de los modelos multinivel 1, 2 y 3 revelan que el bloque de variables individuales no explica totalmente las diferencias entre provincias en cuanto a la abstención y que el residuo correspondiente a cada provincia puede ser cuantificado y modelizado a través de las variables contextuales. Ahora sabemos que en el modelo 3 han desaparecido las diferencias entre provincias. La sola estimación del modelo 4 no permitiría llegar a este conocimiento, dejando abierta la posibilidad de la existencia de diferencias provinciales no explicadas. En definitiva, el proceso seguido permite validar los resultados del Modelo 4. En ese sentido, cabe destacar el papel exploratorio de la aproximación multinivel.

El rol del asociacionismo en la participación electoral

Según los resultados de la Tabla 4, para las elecciones generales del año 2000, el grado de asociacionismo individual y la densidad asociativa provincial no pueden considerarse variables explicativas de la abstención. Ya se había anticipado las limitaciones del indicador de asociacionismo individual y las dudas de las autoras sobre su capacidad explicativa. No es de extrañar que el indicador de densidad

asociativa provincial, obtenido a partir de las mismas variables de la encuesta, haya resultado también poco predictivo. La investigación sobre tipologías de asociaciones debería proporcionar en un futuro elementos para modelizar la relación entre participación en asociaciones y participación electoral de manera más idónea.

A pesar de que en el Modelo 2 la variable individual grado de asociacionismo resultara no significativa, ello no excluye la posibilidad de que sea explicativa. Únicamente indica que no aporta capacidad explicativa adicional. Con el objetivo de explorar la capacidad de esta variable, se estimó un modelo más simple en el que la probabilidad de abstención depende de únicamente dos variables: La edad y el grado de asociacionismo. Los resultados se muestran en la primera columna de la Tabla 5. El grado de asociacionismo resulta ser una variable significativa en la predicción de la abstención y su signo es negativo. Es decir, a mayor participación en asociaciones, menor probabilidad de abstenerse. En este mismo modelo se puede apreciar como existen diferencias significativas de abstención entre provincias, no explicadas por las variables individuales (varianza del residuo provincial = 0,179). Además, tales diferencias no pueden ser explicadas por diferencias de la densidad asociativa de cada provincia puesto que, como se puede apreciar en la segunda columna de la Tabla 5, el correspondiente coeficiente de la variable de segundo nivel es no significativo.

<i>Efectos fijos</i>		
Constante	-1,610 (,082)	-1,737 (,241)
Edad	-,027 (,002)	-,027 (,002)
Grado de asociacionismo	-,239 (,044)	-,241 (,044)
Densidad asociativa provincial		,187 (,335)
<i>Efectos aleatorios</i>		
Varianza entre provincias	,179 (,058)	,187 (,060)
<i>-2log likelihood</i>	3.659,22	3.655,71

Tabla 5. Modelo logístico multinivel: Asociacionismo y abstención

De este estudio se desprende, pues, que la participación en asociaciones favorece también la participación electoral, aunque el efecto pueda quedar oscurecido en los modelos generales de participación por la presencia de otras variables.

CONCLUSIONES

En este trabajo se ha aplicado la metodología multinivel al análisis de la abstención electoral en España. La estimación de los modelos con variables individuales socioestructurales y políticas de dos niveles (individuos y provincias) ha permitido confirmar los resultados presentes en la literatura, obtenidos mediante análisis de un solo nivel.

Además, el análisis realizado corrobora y complementa los resultados obtenidos en los estudios de un solo nivel para las elecciones generales del 2000. En efecto, la magnitud de los coeficientes, sus signos y sus correspondientes errores estándares son prácticamente coincidentes. No obstante, el estudio muestra que prácticamente no quedan variaciones provinciales en la abstención no explicadas por el conjunto de variables del modelo. La metodología multinivel ha permitido combinar el análisis individual con el contextual en un único modelo de dos niveles, de tal forma que la incorporación de variables contextuales ha comportado la reducción casi total de la variación residual provincial. No se ha detectado la existencia de heterogeneidad provincial en la relación entre las variables estudiadas y la abstención. Queda, sin embargo, pendiente un análisis más profundo de la interacción entre variables individuales y contextuales.

Por lo que respecta al asociacionismo, la participación en asociaciones parece tener un papel reductor en la abstención, aunque su efecto no se haya detectado en el modelo multivariante más completo. Probablemente ello sea debido a problemas de correlación entre variables explicativas. En cambio, han resultado ser significativas la variable individual frecuencia de conversaciones políticas y la variable contextual densidad de afiliación a partidos, ambas variables utilizadas en ocasiones como indicadoras del concepto de capital social. Estos resultados van en la dirección ya apuntada de la multidimensionalidad de dicho concepto.

Finalmente, cabe subrayar que aunque el análisis se ha realizado por provincias, la mayor parte de las diferencias provinciales en la abstención se deben a diferencias entre las comunidades autónomas a las cuales pertenecen. Ello se explica desde el punto de vista teórico por la diversidad de contextos en los que se desenvuelve la competición política, incluyendo sistemas de partidos diferenciados entre autonomías.

AGRADECIMIENTOS

Investigación financiada parcialmente por los proyectos del MCyT BFM2000-1064 y SEC2002-03364

ANEXO 1:

INDICADORES UTILIZADOS PARA LAS VARIABLES

Algunas de las variables se han obtenido directamente de respuestas a las correspondientes preguntas de la encuesta. En otras ocasiones, sin embargo, ha sido preciso construir un nuevo indicador a partir de las respuestas o se ha obtenido un índice mediante la aplicación de la técnica del análisis factorial a los datos de un conjunto de preguntas relacionadas entre sí.

Indicadores de variables individuales obtenidos directamente de preguntas de la encuesta:

- Abstención, con valor 0 si en individuo declara que vota y con valor 1 si declara que no vota por distintos motivos.
- Tamaño de municipio, igual al tamaño de la población en la que reside el encuestado agrupado en siete categorías ordinales.
- Edad, medida en años.
- Nivel de estudios, medido en 6 categorías ordinales.
- Situación laboral del encuestado. Esta variable se ha descompuesto en una serie de variables dicotómicas para lo cual se ha utilizado *trabaja* como categoría de referencia. Las restantes variables de situación laboral que se han creado son: *parado, jubilado, estudiante y sus labores*.
- Ocupación del encuestado. También se ha descompuesto la respuesta a la pregunta sobre la ocupación del individuo (o en su defecto del cabeza de familia) en una serie de variables dicotómicas tomando la categoría de los *trabajadores cualificados y semicualificados* como la de referencia. Además de ésta, las variables de ocupación obtenidas son otras 10: *Profesional o técnico; alto cargo o miembro del gobierno, propietario o gerente; cuadro medio; encargado o capataz; administrativo; comercial o vendedor; personal de servicios; trabajador no cualificado o agrícola; y militar*.
- Interés por la política, igual al interés por la política declarado por el encuestado medido en una escala de 4 puntos.

- Satisfacción con el funcionamiento de la democracia medida en una escala de 4 puntos.
- Valoración de la situación política general de España en una escala de 5 puntos
- Valoración de la situación económica general de España en una escala de 5 puntos
- Valoración de la situación económica familiar en una escala de 5 puntos.

Indicadores de variables contextuales obtenidas de fuentes oficiales

- Sectores de ocupación. Son tres variables que contienen los porcentajes de la población activa ocupados en los sectores de: *agricultura, ganadería y pesca; industria; y servicios*. El cuarto sector, porcentaje de ocupación en el sector de la *construcción*, se ha tomado como la categoría de referencia.
- Nivel de renta, igual a la renta por cápita provincial en 1998.
- Número de escaños, igual al número de escaños en el Congreso asignados a cada provincia.

Indicadores obtenidos mediante cálculos algebraicos

- Extremismo político. Para esta variable se han considerado dos indicadores alternativos: a) Valor absoluto de la distancia entre la autoubicación ideológica en el eje izquierda derecha del individuo y la media de las autoubicaciones ideológicas en la provincia; y b) Valor absoluto de la distancia entre la autoubicación ideológica en el eje izquierda derecha del individuo y la media de las autoubicaciones ideológicas en España.
- Grado de asociacionismo, igual a la suma de las respuestas a una serie de trece preguntas dicotómicas (1 = sí, 0 = no) sobre la pertenencia a distintos tipos de asociaciones. El enunciado de la pregunta es:

Señáleme, por favor, si es Ud. miembro o está afiliado a: Una asociación de vecinos; una asociación o grupo juvenil; un partido político; un sindicato; una asociación religiosa o parroquial; una organización empresarial; una asociación o colegio profesional; una asociación cultural o artística; una asociación o grupo ecologista; una asociación o club deportivo; una asociación de padres; una asociación de consumidores; una ONG o asociación de solidaridad.

Su valor indica a cuántos de estos tipos de asociaciones pertenece el individuo.

- Densidad de afiliación a partidos, igual a la suma del número de afiliados al PP y al PSOE en cada provincia por cada 1.000 habitantes.
- Densidad asociativa, índice de asociacionismo provincial obtenido como la media del grado de asociacionismo de los individuos de cada provincia.
- Competitividad electoral en la provincia medida mediante dos indicadores alternativos: Como la diferencia entre los porcentajes de voto obtenidos por los dos partidos más votados en las elecciones del 2000 o bien por el cálculo equivalente con los resultados de las elecciones generales anteriores.
- Número efectivo de partidos, indicador de fragmentación del sistema de partidos que se calcula como el inverso de la suma de los cuadrados de las proporciones de voto sobre votantes obtenidas por todos los partidos (Taagepera y Shugart, 1989).
- *Umbral electoral efectivo* medido mediante el siguiente procedimiento (Taagepera y Shugart, 1989):

Para cada circunscripción electoral en la que se eligen M escaños y compiten p partidos, se definen los umbrales máximo y mínimo

Umbral máximo = proporción de votos por encima de la cual el partido o candidato se asegura la obtención de un escaño. Se estima mediante la fórmula:

$$U_{\max} = 1 / (M+1)$$

Umbral mínimo = proporción mínima de votos que permite obtener un escaño cuando la elección se realiza en condiciones óptimas (con la oposición totalmente fragmentada). Se estima mediante la fórmula:

$$U_{\min} = 1 / (M*p)$$

El *umbral efectivo* es el promedio entre el umbral máximo y el umbral mínimo correspondiente cada provincia, si este promedio es superior al mínimo legal necesario para entrar en el reparto de escaños, o bien el mínimo legal en caso contrario.

Indices obtenidos mediante análisis factorial

Los índices que se presentan a continuación se han construido mediante un análisis factorial con extracción de los factores por el método de ejes principales, realizándose posteriormente un análisis de fiabilidad para comprobar la consistencia de las escalas que se han obtenido.

- Frecuencia de conversaciones políticas, índice obtenido a partir de las respuestas a las cuatro respuestas a la pregunta:
Durante la campaña electoral, ¿Hasta que punto las elecciones del 12 de marzo han sido tema de conversación en sus relaciones personales... En casa con su familia, con sus amigos, con sus vecinos, con sus compañeros de trabajo?
- Seguimiento de información política, índice obtenido a partir de las tres respuestas a la pregunta:
¿Podría decirme con qué frecuencia sigue Ud. la información política a través de los periódicos de información general? ¿Y a través de la televisión? ¿Y a través de la radio?
- Grado de confianza en los partidos, índice que sintetiza las respuestas a los tres ítems siguientes:
Gracias a los partidos la gente puede participar en la vida política
Los partidos sólo sirven para dividir a la gente
Los partidos se critican mucho entre sí, pero en realidad todos son iguales.
- Sentimiento de eficacia política, índice igual al primer factor obtenido de las respuestas a los tres ítems siguientes:
Por lo general, la política es tan complicada que la gente como yo no puede entender lo que pasa
Los políticos no se preocupan mucho de lo que piensa la gente como yo;
Esté quien esté en el poder, siempre busca sus intereses personales.

Preguntas utilizadas como información complementaria en la imputación de míssings.

Para la variable dependiente se utilizaron dos preguntas complementarias en el primer proceso de imputación. La primera es:

Personalmente, en las elecciones generales, ¿tenía usted decidido abstenerse, había decidido votar por un partido o coalición, dudó entre varios partidos o coaliciones, o dudó entre un partido o coalición y la abstención?

A los individuos que declararon que dudaban, se les hizo otra pregunta:

¿Podría decirme entre qué dos partidos u opciones dudó Ud.?

En el segundo proceso de imputación se utilizaron otras dos preguntas de la encuesta preelectoral. Las preguntas eran:

Si mañana se celebrasen elecciones generales, ¿a qué partido votaría Ud.?

Y a los que no contestaban a esta pregunta se les hacía también la siguiente:

¿Por cuál de los siguientes partidos o coaliciones siente Ud. más simpatía?

ANEXO 2: MODELIZACIÓN MULTINIVEL

Modelo de descomposición de la varianza de dos niveles

Si llamamos y_{ij} al valor de la variable numérica dependiente para el individuo i que pertenece a la provincia j , y en la formulación del modelo de regresión habitual (de un solo nivel, entre individuos) se admite que la ordenada en el origen β_{0j} pueda no ser fija sino variable de una provincia a otra, manteniendo sin embargo la pendiente β_1 fija, común a todos los individuos y provincias, se obtiene la siguiente expresión para la respuesta y_{ij} :

$$y_{ij} = \beta_{0j} + \beta_1 x_{ij} + e_{ij}$$
$$\beta_{0j} = \beta_0 + u_{0j}$$

Las hipótesis sobre las distribuciones de probabilidad de los efectos o residuos de nivel individual e_{ij} y nivel de provincia u_{0j} son:

$$e_{ij} \approx NI(0, \sigma^2) \quad u_{0j} \approx NI(0, \sigma_u^2)$$

En este modelo, al igual que en todos los que siguen, se supone que los residuos de cada nivel son incorrelacionados con los del resto de niveles.

Los parámetros de este modelo son cuatro: β_0 , β_1 , σ^2 y σ_u^2 , dos coeficientes y dos varianzas. El primer coeficiente corresponde a la ordenada promedio, el segundo a la pendiente común a todas las rectas.

Sustituyendo la expresión de la ordenada aleatoria en la primera ecuación se obtiene una expresión de la variación de la respuesta y_{ij} en que cabe distinguir la parte fija (promedio) de la parte aleatoria:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_{0j} + e_{ij}$$

La estimación de este modelo¹⁴ dará lugar a una recta promedio y a una serie de rectas paralelas, una para cada provincia (todas con la misma pendiente pero con diferente ordenada en el origen). El residuo provincial u_{0j} es el responsable de la separación de cada recta provincial respecto de la recta promedio. Así en la predicción de la respuesta para un cierto individuo perteneciente a la provincia j deberá añadirse a la predicción promedio la estimación del residuo o efecto provincial correspondiente.

El modelo anterior puede ampliarse a tres o más niveles si consideramos que la ordenada en el origen admite una componente de variación en dichos niveles. En nuestro caso, dada la estructura de los datos cabe plantearse la posibilidad de variación a nivel de comunidad autónoma, dando lugar a un modelo de tres niveles (individuo, provincia y comunidad autónoma).

Las hipótesis del modelo de regresión con ordenada aleatoria y pendiente fija ofrecen una descomposición básica de la variación de la variable respuesta en los dos niveles de variación posibles: entre grupos (provincias) y dentro de los grupos entre individuos.

$$\text{var}(Y_{ij}) = \text{var}(u_{0j}) + \text{var}(e_{ij}) = \sigma_u^2 + \sigma^2$$

La covarianza entre dos observaciones del mismo grupo (i y i' siendo $i \neq i'$) resulta ser igual a la varianza del efecto grupo (varianza del segundo nivel):

$$\text{cov}(Y_{ij}, Y_{i'j}) = \text{cov}(u_j + e_{ij}, u_j + e_{i'j}) = \text{var}(u_j) = \sigma_u^2$$

La correlación entre dos observaciones diferentes del mismo grupo será igual al cociente:

$$\rho = \rho(Y_{ij}, Y_{i'j}) = \frac{\sigma_u^2}{\sigma_u^2 + \sigma^2}$$

Este coeficiente conocido con el nombre de *coeficiente de correlación intragrupos* (intra-class correlation en inglés) admite las siguientes interpretaciones:

¹⁴ Si en la formulación de este Modelo de ordenada aleatoria se prescinde de la variable explicativa se obtiene el Modelo de Análisis de la varianza de efectos aleatorios (RANOVA). En tal caso el parámetro β_0 representará la media poblacional también llamada μ .

- La correlación entre dos observaciones escogidas al azar de un mismo grupo a su vez escogido al azar entre todos los grupos
- La proporción de variación total de Y que corresponde a variabilidad entre grupos
- Una medida del grado de homogeneidad de los grupos, de la dependencia dentro (intra) de los grupos

El coeficiente de correlación intragrupos ρ siempre es positivo y varía entre 0 y 1. Cuanto más difieran los grupos entre sí (medias diferentes) y más similares sean los individuos dentro de sus correspondientes grupos, mayor será ρ . La existencia de correlación intragrupos viola la hipótesis de independencia de los modelos clásicos (de un solo nivel) de regresión y conlleva una subestimación de los errores estándar. En consecuencia, cuanto mayor sea ρ menos fiables serán los resultados derivados de un análisis clásico. Valores de ρ alrededor de 0.10 aconsejan un análisis multinivel.

Modelo de regresión dos niveles con coeficientes aleatorios

Uno de los modelos que presenta mayor interés en cuanto a su aplicación es el modelo propiamente llamado de coeficientes aleatorios, puesto que admite la variación entre provincias tanto de la ordenada en el origen como del coeficiente de la variable explicativa o pendiente. La parte aleatoria de este modelo considera la variación de cada uno de estos coeficientes y la posible covarianza entre ambas variables aleatorias:

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + e_{ij}$$

$$\beta_{0j} = \beta_0 + u_{0j}$$

$$\beta_{1j} = \beta_1 + u_{1j}$$

Las hipótesis distribucionales sobre los residuos de ambos niveles son:

$$e_{ij} \approx NI(0, \sigma^2) \quad \begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \approx N(0, \Omega_u)$$

Puesto que los residuos de segundo nivel admiten correlación, en este caso se debe considerar la matriz de varianzas y covarianzas:

$$\Omega_u = \begin{bmatrix} \sigma_{u0}^2 & \sigma_{u10} \\ \sigma_{u10} & \sigma_{u1}^2 \end{bmatrix}$$

Sustituyendo en la primera ecuación las modelizaciones de los coeficientes, se obtiene una única ecuación en que cabe separar la parte fija de la parte aleatoria del modelo. La parte fija corresponde a la recta promedio que predice el valor de la respuesta para cada individuo en función del valor de la variable explicativa sin tener en cuenta el grupo al cual pertenece el individuo.

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_{0j} + u_{1j} x_{ij} + e_{ij}$$

Por último, la modelización de los coeficientes puede adquirir mayor complejidad si asumimos que no solamente tienen una variación entre grupos sino que dependen de variables de segundo nivel.

$$\beta_{0j} = \beta_{00} + \beta_{01} z_j + u_{0j}$$

$$\beta_{1j} = \beta_{10} + \beta_{11} z_j + u_{1j}$$

Este modelo de mayor complejidad contempla, pues, la interacción entre variables de dos niveles (*cross-level interactions*)

$$y_{ij} = \beta_{00} + \beta_{01} z_j + \beta_{10} x_{ij} + \beta_{11} z_j x_{ij} + u_{0j} + u_{1j} x_{ij} + e_{ij}$$

Modelo de regresión logística multinivel

Finalmente, se presentan los modelos de regresión logística multinivel que, por ser la variable dependiente dicotómica, son los adecuados para el estudio de la abstención. Nuestra variable dependiente y_{ij} es la abstención individual, que toma valor 1 si el individuo i de la provincia j se abstiene y 0 en caso contrario. Por π_{ij} se denota la probabilidad de abstención para el individuo i de la provincia j . La especificación del modelo logístico de dos niveles, con ordenada y pendiente aleatorios, sería:

$$\begin{aligned}
E(y_{ij}) &= \pi_{ij} \\
E\left(\log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right)\right) &= \beta_{0j} + \beta_{1j}x_{1ij} \\
\beta_{0j} &= \beta_0 + u_{0j} \\
\beta_{1j} &= \beta_1 + u_{1j}
\end{aligned}$$

Las hipótesis sobre las variables residuales son: el residuo de primer nivel e_{ij} es una variable binomial de parámetro π_{ij} y (u_{0j}, u_{1j}) , la variable residual del segundo nivel, tiene la matriz de varianzas y covarianzas Ω_u como se ha especificado en el modelo anterior.

REFERENCIAS

- Almond, G. y Verba, S. 1963. *The Civic Culture*. Princeton: Princeton University Press.
- Anduiza, E. 1999. *¿Individuos o sistemas? Las razones de la abstención en Europa Occidental*. Madrid: Centro de Investigaciones Sociológicas (CIS).
- Aitkin, M. and Longford, N. 1986. "Statistical modelling issues in school effectiveness studies", *Journal of the Royal Statistical Society A* 149, Part 1: 1-43.
- Barreiro, B. 2002. "La progresiva desmovilización de la izquierda en España: Un análisis de la abstención en las elecciones generales de 1996 a 2000", *Revista Española de Ciencia Política*, 6: 183-205.
- Boix C. y Posner, D. 2000. "Capital social y democracia", *Revista Española de Ciencia Política*, 2: 159-185.
- Boix, C. y Riba, C. 2000. "Las bases sociales y políticas de la abstención en las elecciones generales españolas: Recursos individuales, movilización estratégica e instituciones electorales", *Revista Española de Investigaciones Sociológicas*, 90: 95-128.
- Byrk, A. S. y Raudenbush, S. W. 1992. *Hierarchical Linear Models: Applications and data analysis methods*. Newbury Park: Sage Publications.
- Dahl, R. 1956. *A Preface to Democratic Theory*. Chicago: The University of Chicago Press.
- Downs, A. (1957) *An Economic Theory of Democracy*. New York, Harper & Row.
- Goldstein, H. 1991. "Nonlinear multilevel models with an application to discrete response data", *Biometrika*, 78: 45-51.
- Goldstein, H. 1995. *Multilevel Statistical Models*. 2nd ed. Kendall's Library of Statistics 3. London: Edward Arnold.
- King, G., Honaker, J., Joseph, A. y Scheve, K. 2001. "Analizing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation", *American Political Science Review* 95: 49-69.
- Jones, K., Johnston, R. J. y Pattie, C. J. 1992. "People, Places and Regions: Exploring the Use of Multi-Level Modeling in the Analysis of Electoral Data", *British Journal of Political Science*, 22: 343-380.
- Little, J. R. y Rubin, D. 1987. *Statistical Analysis with Missing Data*. New York: Willey
- Longford, N. T. 1994. "Random Coefficient Models", en Arminger y Sobel, eds., *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. New York: Plenum Press.
- Méndez, M. y Mota, F. 2003. "Las características organizativas de las asociaciones en España", comunicación presentada en el VI congreso de la AECPA, en Barcelona.

- Mota, F. y Subirats, J. 2000. "El quinto elemento: el capital social de la Comunidades Autónomas. Su impacto sobre el funcionamiento del sistema político autonómico", *Revista Española de Ciencia Política*, 2: 123-158.
- Morales, L. 2001, "Citizens In Politics: The Individual and Contextual Determinants of Political Membership in Western Countries", *WP 2001/164 Center for Advanced Study in the Social Sciences*, Juan March Institute. Madrid.
- Putnam, R., Leonardi, R. y Nanetti, R. 1993. *Making democracy work. Civic traditions in modern Italy*. Princeton: Princeton University Press.
- Ranney, A. 1983. "Nonvoting is not a Social Disease", *Public Opinion* October-November, 16-19.
- Riker, W. H. y Ordeshook, P. C. 1968. "A theory of the Calculus of Voting" *American Political Science Review*, 62: 25-42.
- Rosenstone, S. J. y Hansen, J. M. 1993. *Mobilization, Participation and Democracy in America*. New York: Macmillan.
- Rubin, D. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: Willey.
- Searle, S. R., Casella, G. y McCulloch, Ch. E. 1991. *Variance Components*. New York: Wiley Interscience.
- Sherman, R. P. 2000. "Tests of Certain Types of Ignorable Nonresponse in Surveys Subject to Item Nonresponse or Attrition", *American Journal of Political Science* 44: 356-368.
- Schumpeter, J. 1945. *Capitalism, Socialism and Democracy*. New York: Harper.
- Snijders, T. A. y Bosker, R. J. 2000. *Multivel Analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.
- Steenberger, M. R. y Jones, B. S. 2002. "Modeling Multilevel Data Structures", *American Journal of Political Science*, 46(1): 218-237.
- Taagepera, R. y Shugart, M. S. 1989. *Seats and Votes. The Effects and determinants of Electoral Systems*. New Haven: Yale University Press.
- Torcal, M. y Montero, J. R. 2000. "La formación y consecuencias del capital social en España", *Revista Española de Ciencia Política*, 2: 79-121.
- Wolfinger, R. y Rosenstone, S. J. 1980. *Who votes?* New Haven: Yale University Press.