

# 02

## ¿Se pueden mejorar las predicciones electorales?

Nuevas metodologías de análisis de los sondeos electorales y de los históricos de resultados

*Can election forecast be improved?  
New methodologies for the analysis of electoral surveys and historical results*

**Davide Fabrizio**

Coneto

### **Resumen / Abstract**

En este artículo presentamos dos ejemplos de nuevas técnicas de análisis, que pueden ofrecer respuestas nuevas al investigador y que están orientadas a la mejora de la capacidad predictiva de los resultados electorales. En primer lugar, la reconstrucción de la variable "intención de voto"; y en segundo lugar, cómo el análisis de una serie temporal permite, una vez reconocidos los factores que condicionan el voto, mejorar sensiblemente el acierto en las predicciones de intención de voto.

*The paper presents two examples of new analysis techniques that can provide researchers with new solutions and that are oriented to improve the predictions of election results. First, we will focus on reconstructing the "voting intention" variable; and, secondly, on how the analysis of a temporary series allows, once the factors conditioning the vote have been spotted, to improve accuracy in the predictions of voting intention.*

### **Palabras clave / Keywords**

Análisis electoral. Comunicación Política. Intención de voto.  
*Election analysis. Political Communication. Voting intention.*

## 1. Introducción

Estimar un resultado electoral no es nunca una tarea sencilla. Las razones principales son esencialmente dos, en parte relacionadas: por un lado la falta de series robustas de datos de partida, por el otro el retraso, en comparación con otros sectores como el marketing o la economía, en los avances de desarrollo de modelos explicativos y predictivos del comportamiento de voto de los electores.

Si analizamos datos de encuesta, por ejemplo, tenemos que tener en cuenta algunas de las principales distorsiones asociadas a los sondeos: el error de muestreo<sup>1</sup> o de cobertura, por ejemplo, no resultan eliminables, con la posibilidad de añadir márgenes de error importantes debidos a una equivocada formulación de las preguntas o ponderación de los datos, o a una escasa formación del entrevistador que puede llevar a una cuota mayor de respuestas no sinceras por parte del entrevistado (Ceri, 1997; Díaz de Rada, 2010). Por otro lado, los resultados de las elecciones anteriores constituyen una ayuda indispensable para el investigador en su objetivo de predecir los resultados electorales: pero en este caso las series históricas disponibles, votando cada 4 años, son muy limitadas, con claros reflejos en las posibilidades de desarrollo de nuevos modelos de análisis, sobre todo longitudinales (o sea el análisis de las series temporales).

Pero, sobre todo, hay que conocer bien los *drivers* que pueden mover el voto de los electores y medir su importancia en determinar la decisión de voto final (Lau y Redlawsk 2006): el peso que puede tener la comunicación política y las campañas electorales (Martínez i Coma, 2008), el peso de factores más estructurales como las características socio-demográficas y el comportamiento de voto pasado, el peso de variables como la valoración del gobierno y la satisfacción relacionada con la situación económica. Un conocimiento completo de estos factores es la clave para hacer predicciones acertadas. Pero, ¿cómo podemos reducir el margen de error asociado a las predicciones electorales? El desarrollo de nuevos modelos matemáticos y estadísticos es una de las respuestas.

En este artículo presentamos dos ejemplos de nuevas técnicas de análisis, que pueden ofrecer nuevas respuestas al investigador y que están orientadas a la mejora de la capacidad predictiva de los resultados electorales.

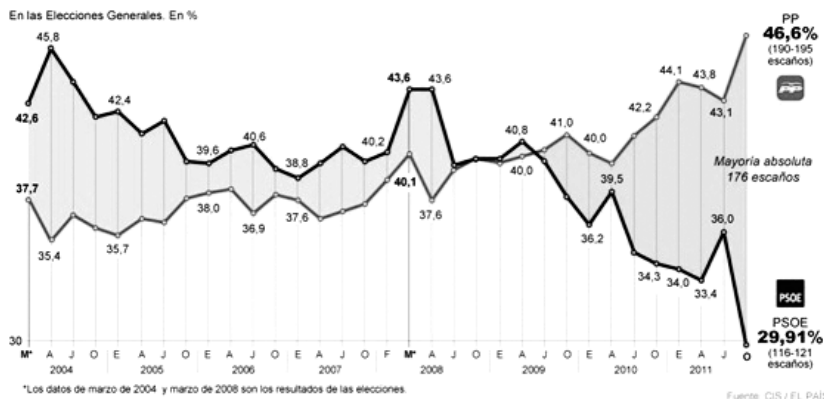
En el primer caso introducimos un modelo transversal que permite, en el análisis de las encuestas electorales, la reconstrucción de la variable "intención de voto". Esta técnica es muy útil para corregir algunas distorsiones de los datos de las encuestas, que se notan sobre todo en los meses antecedentes a la fecha de las elecciones. En el segundo caso presentamos un modelo longitudinal, o sea un

análisis de una serie temporal que permite, una vez reconocidos los factores que condicionan el voto, mejorar sensiblemente el acierto en las predicciones.

## 2. Modelo transversal: reconstrucción de la intención de voto de las encuestas

La gran pregunta que todo el mundo se ha hecho en campo electoral en el curso de este último año ha sido la siguiente: ¿cuál será el resultado de las elecciones generales? Todavía no conocemos la respuesta exacta (en el momento en el cual se escribe este artículo faltan todavía dos semanas a la cita electoral del 20-N), porque en el escenario persisten elementos de incertidumbre que todavía no se han aclarado. Pero los sondeos del último año y el resultado de las elecciones Autonómicas y Municipales del 22-M nos han dado pistas muy claras: no hay ninguna duda sobre el ganador de las próximas elecciones, la única duda está relacionada con la consecución o menos de la mayoría absoluta de los escaños.

En el curso del año 2011 hemos visto una cierta variación en la estimación del voto: en verano, por ejemplo, una recuperación del PSOE parecía difícil pero posible, gracias al nombramiento a candidato a la Presidencia del Gobierno de Alfredo Pérez Rubalcaba. Como vemos en la Figura 1, que ofrece las estimaciones de la intención de voto del CIS en sus encuestas desde abril 2004 hasta octubre 2011, en la encuesta de Julio 2011 se registra una fuerte recuperación del PSOE, que reduce su distancia con el PP a 7,1 puntos (desde los 10,4 puntos de abril). Pero, ¿se trata de una recuperación real o se debe a otros factores de distorsión? En otras palabras, si se hubiese votado en Julio 2011, ¿el PP habría ganado con una ventaja de 7 puntos porcentuales? La respuesta es sencilla: probablemente no. Vamos a analizar el porqué.



La literatura politológica habla de diferentes tipos de voto, entre los cuales destacan el voto de pertenencia y el voto de opinión (existe también el voto de intercambio, que no analizamos porque tiene un peso más limitado en España con respecto a otros países) (Mannheimer & Sani 1987). En el primer caso estamos hablando de un voto que tiende a la continuidad: el comportamiento electoral se basa principalmente en la identificación partidista, una vinculación más bien estable a largo plazo. Esto significa que, si consideramos un sondeo electoral, la decisión de voto del entrevistado depende en gran medida de algunas componentes estructurales como las variables políticas (ubicación en el eje izquierda/derecha, voto en las elecciones anteriores) y socio-demográficas (la edad, el nivel de estudio, la profesión, el área de residencia). En el segundo caso, el voto de opinión, estamos hablando de un consenso que supone una apertura mayor a la movilidad partidaria y al cambio político: el comportamiento electoral depende de las evaluaciones de los electores sobre propuestas programáticas de los partidos y sobre las percepciones relacionadas con temas de actualidad social, política y económica. Retomando el ejemplo del sondeo, en este caso la decisión de voto depende de las respuestas relacionadas con las variables de opinión (valoración de la situación económica, de la actuación del gobierno) o contingentes (opiniones sobre otros temas de actualidad en el debate público o en la campaña electoral). En la Figura 2 podemos ver una representación del esquema analizado.

En los meses anteriores a las elecciones, las variables contingentes son las que más influyen sobre la declaración de voto de los entrevistados en las encuestas: una nue-

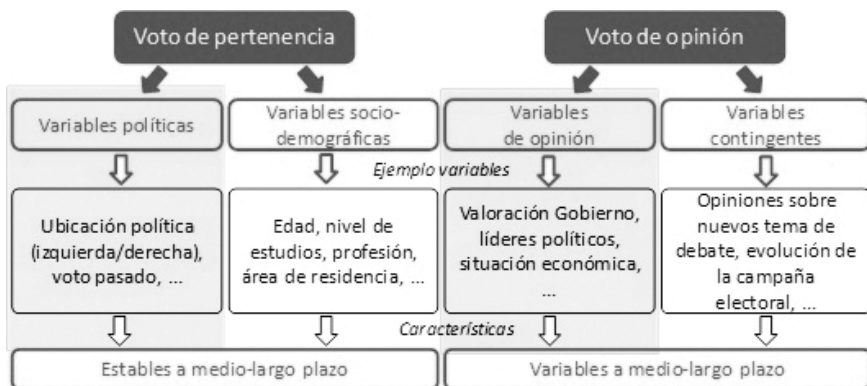


Figura 2. Tipos de voto y variables relacionadas con la decisión de voto.

va situación política (un nuevo candidato, un cambio en la composición del gobierno) o un evento exógeno inesperado (un escándalo político, un desastre medioambiental, etc.) pueden 'alterar' temporalmente las declaraciones de voto de los entrevistados, produciendo cambios significativos en las intenciones de voto. Se trata en cualquier caso de cambios temporáneos, que van a ser reabsorbidos en el corto-medio plazo.

Hagamos un ejemplo relacionado con los datos de la Figura 1: analizando los datos de la encuesta del CIS de julio 2011 se nota como muchos electores con perfil abstencionista o con colocación de izquierda radical (y algunos también con perfil de centro-derecha), declaran de repente, una vez hecha pública la renuncia de Zapatero y la nominación de Rubalcaba, una intención de voto 2012 a favor del PSOE. El efecto 'novedad' les lleva a una verdadera y provisional 'conversión'. Pero esta condición no es constante el tiempo: en las semanas siguientes la situación evoluciona y vuelven a ganar más peso los componentes estructurales, como por ejemplo las variables políticas, socio-demográficas, y las variables de opinión sobre la política y la economía. Si ganan peso las componentes estructurales puede cambiar la declaración y el comportamiento de voto final del entrevistado: al final un elector que nunca va a votar; por cuanto pueda estar satisfecho con la candidatura de Rubalcaba, muy difícilmente acudirá a las urnas el 20-N. Lo mismo pasa con un elector del PP insatisfecho del candidato Rajoy y con una buena valoración de Rubalcaba: en la cabina electoral no se atreve a dar un cambio tan radical hacia el PSOE, optando por una solución de continuidad. Es la clásica elección del mal menor; porque la distancia ideológica con los socialistas es demasiado grande en el caso concreto para ser colmada en esta circunstancia.

Seguramente la crisis financiera y económica a partir del año 2008 ha cambiado de manera radical las percepciones de los votantes en todos los países democráticos: la evaluación de la situación económica se ha convertido en el factor más explicativo del comportamiento electoral de los ciudadanos, casi una componente 'estructural' de la decisión de voto. De hecho, hace unos años eran muy complicados los movimientos entre campos políticos distintos (izquierda-derecha): se trataba de una decisión muy difícil de tomar para cualquier elector. Esta conversión radical (del todo asimilable a una conversión religiosa) existía, pero era muy limitada en la práctica: normalmente estos procesos suponían recorridos lentos y graduales, conflictos internos que ponían en discusión las creencias y las elecciones pasadas de las personas. Pues, por las razones explicadas, el votante privilegiaba soluciones de continuidad, que no pusieran en peligro sus seguridades. Prácticamente todos los análisis de flujos de voto en España, históricamente, señalaban esta tendencia, con elevados porcentajes de fidelidad.

Con la crisis todo ha cambiado, y este proceso de ‘conversión’ se ha acelerado: la insatisfacción hacia la gestión de la situación económica por parte del gobierno lleva una parte del electorado a apostar por un cambio radical<sup>3</sup>, independientemente de sus valores políticos. Esto explica por qué, el 20-N, los populares obtendrán su mejor resultado en la historia y los socialistas el peor<sup>4</sup>.

Estos ejemplos ponen en evidencia la necesidad de elaborar nuevos modelos de análisis de los datos de sondeos electorales, que permitan tener en cuenta la variación temporal del peso de las variables explicativas del voto: como hemos visto, en los meses anteriores a las elecciones, los electores se apoyan mucho en evaluaciones coyunturales (como por ejemplo el efecto ‘novedad’ de la candidatura Rubalcaba), pero más adelante sus creencias, sus valores, sus actitudes (las fuerzas disposicionales) y sus evaluaciones sobre la situación económica volverán a ganar peso y espacio en el mecanismo de elección del voto.

Con el uso de algoritmos matemáticos y estadísticos, se puede recrear en cualquier momento, analizando los datos de los sondeos electorales, las condiciones del elector en el día del voto para las elecciones generales del 20-N. Esto significa reconstruir la intención de voto declarada por los entrevistados, gracias a un modelo capaz de contestar a la siguiente pregunta: ¿cambiaría la intención de voto actual declarada por los entrevistados (cuando todavía faltan meses o semanas a la fecha del voto) si le aplicáramos el peso de los factores explicativos del día de la votación? Esto significa que el modelo puede detectar en cualquier momento la probabilidad de que un elector convierta efectivamente su intención en un comportamiento de voto real. Este resultado se puede obtener con modelos de regresión logística, que (1) reducen la aportación de las variables contingentes a la declaración de voto de los entrevistados, (2) atribuyen un peso mayor a los componentes estructurales y de opinión y (3) ‘reconstruyen’ una nueva estimación de la intención de voto.

El modelo es el fruto de un análisis en profundidad sobre centenares de encuestas personales y telefónicas de España y de los principales países en los últimos diez años. La aplicación de la técnica de corrección de los datos a diferentes encuestas pone en evidencia algunas regularidades, que se pueden notar comparando las respuestas originales de los entrevistados a la pregunta “¿Qué partido votaría?” y la “reconstrucción” de la respuesta conseguida con el modelo. En general los aprendizajes más significativos son los siguientes:

**I. Baja la MOVILIDAD: los flujos de voto de un partido al otro disminuyen. El nivel de volatilidad es más limitado, mientras la fidelidad gana relevancia. En muchos ca-**

sos, cuando un entrevistado declara un voto pasado para un partido x y una intención de voto para un partido y, el modelo detecta como más probable un nuevo voto a favor de x;

2. Sube la **DESACTIVACIÓN**: los electores que no aprueban la actuación de su partido de referencia se refugian más fácilmente en la abstención que en el cambio de partido. Se trata de una defección parcial, un primer pase hacia un cambio más radical (que podría convertirse en efectivo en la cita electoral siguiente). Es la situación de un entrevistado que declara un voto pasado para un partido x y una intención de voto para un partido y: el modelo detecta como más probable una abstención;

3. Sube la **ACTIVACIÓN** de los indecisos: en realidad muchos electores que se declaran indecisos tienen las ideas bastantes claras, pero todavía no quieren descubrir sus cartas. De hecho, en muchos casos los indecisos presentan características estructurales definidas (comportamiento de voto pasado, ubicación ideológica, valoración de los líderes políticos) que permiten al modelo predecir una elevada probabilidad de un voto a un partido determinado.

La aplicación de este modelo matemático ofrece una respuesta importante para mejorar las predicciones electorales, gracias a técnicas capaces de “incorporar” y tener bajo control algunos de los efectos más analizados y estudiados en la literatura politológica internacional.

Uno de estos efectos es el llamado *bandwagon* o de ‘subirse al carro del ganador’ (Mehrabian, 1998; Goidel & Todd, 1994), efecto que en este caso se aplica a las declaraciones de voto y no tanto al voto en concreto. Algunas personas, de hecho, declaran su voto por aquellos candidatos o partidos que los sondeos y la opinión pública en general ven como ganadores probables (o, como en el caso presentado anteriormente, por aquellos candidatos que se benefician de una fuerte visibilidad temporánea). Se trata de una modificación de las opiniones de acuerdo al punto de vista de la mayoría, a la base de la cual, según las ciencias psicológicas, encontramos la ‘deseabilidad social’ (Roccatto, 2003): los entrevistados ajustan sus opiniones a lo que (creen) el entrevistador y más en general la opinión pública espera de ellos. No cabe duda de que la mayoría de los españoles califique de mala o muy mala la gestión del gobierno del PSOE. Pero, si un elector pensara de manera diferente, no sería sencillo ‘nadar contracorriente’ en una entrevista telefónica (y en muchos casos tampoco merecería la pena). El modelo consigue reducir el peso del efecto *bandwagon*. Otro efecto contabilizado con el uso del modelo es el llamado voto útil, cuyo peso se ve ampliado sensiblemente. De hecho los entrevistados, en un momento lejano al día de la elección, tienden a ex-

presar un voto libre, que en muchos casos se dirige hacía partidos menores, sin darse cuenta que al final serán casi 'obligados' a votar de forma estratégica. Para evitar que su voto sea desperdiciado, direccionarán su consenso hacia preferencias *second-best*, o sea hacia partidos que pueden ganar las elecciones o que pueden superar la barrera legal del 3% a nivel de circunscripción (y participar a la repartición de los escaños).

¿Cuáles pueden ser las ventajas tangibles obtenidas gracias a la aplicación de este modelo matemático-estadístico? Las ventajas se pueden medir bajo tres perspectivas diferentes:

1. El análisis de la intención de voto nos ofrecerá resultados y estimaciones más adherentes a la realidad, que no sufren de las distorsiones proporcionadas por el peso, variable en el tiempo, de los factores explicativos de la decisión de voto. Conseguiremos de esta forma contestar eficazmente a la pregunta ¿cuál sería el resultado de las elecciones generales si se votara hoy?

2. La aplicación del modelo permite obtener segmentaciones del electorado más eficientes. El tema de los indecisos es emblemático: el modelo detecta los verdaderos indecisos, que son una porción más reducida de los que se declaran indecisos en las encuestas. No es solamente un tema numérico: la composición de los indecisos en la variable original y en la variable corregida con el modelo es significativamente diferente, ofreciendo enfoques alternativos a los partidos políticos que organizan acciones de campaña ad hoc finalizadas a la recuperación del segmento.

3. Las tendencias temporales de la intención de voto en encuestas seguidas enseñarán una variabilidad menor. Se evitará de esta forma la andadura esquizofrénica de algunas tendencias de las encuestas, ofreciendo un panorama con volatilidad menor, más coherente, que ayudará a detectar y a medir los verdaderos efectos de los eventos de la campaña mediática (ofreciendo una información útil a los expertos en comunicación y marketing político).

La tabla 1 muestra un ejemplo del resultado de la aplicación de esta técnica a las bases de datos de las últimas encuestas publicadas por el CIS antes de las elecciones Autonómicas del 22-M. Se trata de las encuestas de intención de voto en las 13 autonomías, publicadas en la web del CIS unas semanas antes del voto. En este caso hemos reelaborado las bases de datos a nivel individual y hemos reestimado la intención de voto con el modelo matemático. Los resultados son claros: la aplicación de esta técnica reduce -en algunos casos de forma notable- el margen de error de los sondeos (error medio por partido<sup>5</sup>).



Comunidad	Error medio por partido		Dif. B-A
	CIS (A)	Reconstrucción voto (B)	
Extremadura	1,1	1,1	0,0
Castilla-La Mancha	1,1	0,9	-0,2
Baleares	3,5	2,1	-1,4
Aragón	3,4	2,1	-1,3
Cantabria	2,6	2,6	0,0
Asturias	7,2	7,3	0,1
Comunidad de Madrid	1,3	1,0	-0,3
Comunidad Valenciana	3,6	2,2	-1,4
Murcia	2,6	2,2	-0,4
Castilla y León	1,6	1,2	-0,4
La Rioja	0,8	0,9	0,1
Ceuta	2,5	1,7	-0,8
Melilla	8,2	4,5	-3,7

Tabla 1. Error medio por partido en las encuestas publicadas por el CIS de ámbito autonómico antes de las elecciones del 22-M (CIS y nuevo modelo matemático).

### 3. Modelo longitudinal: análisis de series temporales para estimar los resultados electorales

Los modelos econométricos (o análisis de series temporales) consisten en una simulación matemática de la realidad, que nos ayuda a conocer por qué han sucedido las cosas de una forma determinada, así como a utilizar este conocimiento para predecir un comportamiento futuro, como por ejemplo pronosticar un resultado electoral con un buen nivel de precisión. Se utiliza en este caso el modelo de regresión múltiple, que estudia la relación existente entre una variable dependiente (el resultado electoral) y variables independientes (variables económicas, valoraciones de del gobierno, factores exógenos, estacionalidades, radicación de los partidos en el territorio, etc.). La fórmula sería:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \epsilon$$

Donde Y es la variable dependiente o explicada, X<sub>1</sub>...X<sub>n</sub> son las variables independientes o explicativas y ε se denomina perturbación aleatoria (es el error asociado en todos los modelos: no podemos esperar captar toda la influencia de unas variables en un modelo, por muy elaborado que éste sea). Estos modelos se

han utilizado con mucho éxito en las elecciones presidenciales estadounidenses (Lewis-Beck & Rice, 1992), gracias a la presencia de dos elementos de ventaja con respecto a otros países: un formato bipartidista estable en el tiempo (Republicanos y Demócratas, sin otros competidores relevante con la única excepción de Ross Perot en la carrera presidencial de 1992) y una serie temporal bastante larga en comparación con otros países (tenemos por ejemplo 15 puntos de observación si analizamos las elecciones entre 1952 y 2008). En los países europeos estas técnicas todavía no se utilizan mucho (el formato partidista más variable es uno de los problemas principales), pero no faltan intentos y experimentos exitosos en los últimos años en el Reino Unido, Francia, Alemania y España (Lewis-Beck & Jérôme, 2010).

Tomamos como ejemplo de referencia el caso de EE.UU, siguiendo el trabajo desarrollado por Michael Lewis-Beck en EE.UU. en ocasión de las elecciones presidenciales 2008 (Lewis-Beck & Tien, 2008). En este caso vamos a explicar el resultado de las elecciones en el voto popular del partido del presidente incumbente<sup>6</sup> (V) en el período 1952-2008 a través de 5 variables independientes:

1. P: la popularidad del Presidente en el mes de julio del año de las elecciones (t), medida a través de las encuestas Gallup.

2. G: la variación del producto interior bruto (PIB) en porcentaje desde el cuarto trimestre del año anterior a las elecciones (t-1) al segundo trimestre del año de las elecciones (t).

3. E: el presidente incumbente está presente (1) o no (0,5) en las elecciones.

4. J: la tasa de crecimiento del empleo en los primeros 3 años y medio de mandato del presidente.

5. I: la ventaja del incumbente (1 si el candidato del partido incumbente es el presidente, 0 si el candidato del partido incumbente tiene una relación clara con el presidente anterior, -1 si no la tiene).

Analizando los valores de las variables independientes y de la dependiente desde el año 1952, observamos como podemos explicar el resultado del partido del presidente incumbente (V) con la siguiente ecuación<sup>7</sup>:

$$V = 33,01 + 0,25*P + 1,51*(G \times E) + 0,50*J + 2,14*I + ?$$

Esto significa que el resultado del partido del presidente incumbente está determinado por un nivel base de 33 puntos (o sea una cuota mínima de consenso que puede obtener independientemente de otros factores políticos y económicos, y que depende de la visibilidad adquirida en los cuatro años de mandato), al cual sumamos 0,25 puntos por cada subida de popularidad de un 1%, 1,51 puntos por cada subida del PIB de un 1% (0,75 puntos si el presidente incumbente no está presente en las elecciones), 0,50 puntos por cada subida de la tasa de crecimiento del empleo de un 1% y 2,14 puntos si el candidato del partido incumbente es el presidente. La misma ecuación nos permite estimar el resultado de las elecciones presidenciales si se votara hoy. A día de hoy (noviembre 2011), los valores de las variables independientes son los siguientes:

- P: 42% (dato Gallup).
- G: +1,50%.
- E = 1 (Obama estará presente en las elecciones del 2012).
- J: -0,24%.
- I = 1 (Obama es el candidato del partido incumbente).

Si ponemos los valores en la fórmula, el resultado estimado para Obama sería el siguiente:

$$V = 33,01 + 0,25 \cdot 42 + 1,51 \cdot (1,5 \times 1) + 0,50 \cdot (-0,24) + 2,14 \cdot 1 + ? = 47,8\%$$

Como vemos, actualmente Obama no ganaría la reelección<sup>8</sup>. Es todavía pronto para hacer un pronóstico, pero sabemos que el nivel de recuperación económica (en términos de crecimiento del PIB, de la tasa de empleo y de la correlacionada popularidad del Presidente) serán la clave que decidirá las próximas elecciones, en gran parte de forma independiente respecto a la conducción de la campaña electoral por parte de los Demócratas y de los Republicanos. Una recaída en la recesión (el tanto temido *double-dip*), podría resultar fatal para el candidato demócrata, mientras buenas noticias económicas (crecimiento de la economía y recuperación de empleo) serían la palanca para mejorar su valoración y, consecuentemente, aumentar las probabilidades de reelección.

En el gráfico siguiente (figura 3) resumimos la información principal aportada por el modelo econométrico. La línea azul representa el dato real, mientras la línea roja corresponde al ajuste del modelo, o sea a la variación del resultado real explicada por el modelo. Las áreas coloreadas representan el peso de las diferentes variables explicativas en las elecciones presidenciales estadounidense 1952-2008: como vemos hay un nivel base del 33% (como hemos visto anteriormente: por cuanto la situación pueda ser mala, el partido del presidente incumbente tiene una

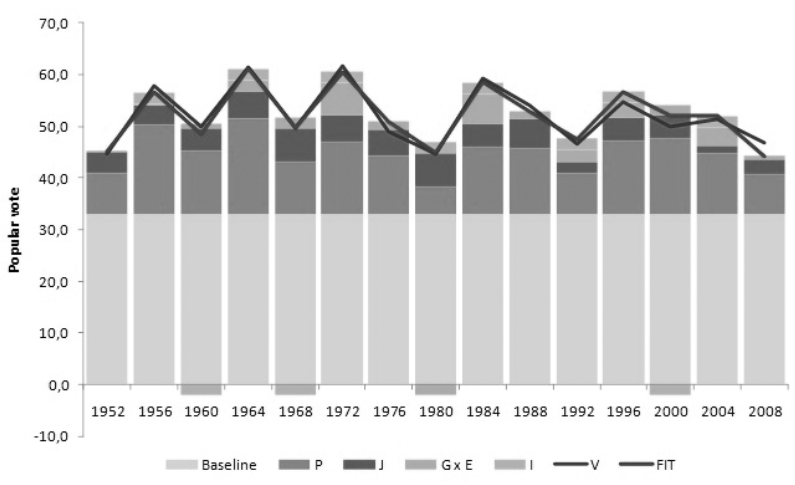


Figura 3. Modelo econométrico explicativo del resultado de las elecciones presidenciales de EE.UU. 1985-2008

base de consenso estable cuantificable en un 33%), a la cual se suman las áreas debidas a la popularidad del Presidente (azul), al crecimiento del empleo (roja), a la variación del PIB (naranja) y la ventaja del incumbente (verde).

La aplicación de estos modelos en el contexto electoral de otros países puede ofrecer resultados muy interesantes, con predicciones muy fiables a pesar de la limitada numerosidad de los puntos de observación. La gran utilidad es la posibilidad de detectar los factores estructurales capaces de mover el voto de los ciudadanos: esto significa que podemos predecir un resultado electoral -con un buen nivel de precisión- con mucha antelación respecto al día de la votación. Esto no significa que las campañas electorales y las componentes de corto plazo no muevan voto, pero son principalmente las componentes de largo plazo a orientar las preferencias políticas.

#### 4. Conclusiones

Como hemos visto el análisis de los sondeos electorales lleva mucho retraso en comparación con los avances de los modelos predictivos en otros sectores como el marketing o la economía. Un poco depende de los límites del método de recogida de la información (las encuestas), un poco de la escasez de materia prima a nivel de fuente (los datos electorales, sobre todo si queremos realizar aná-

lisis longitudinales). En cualquier caso, pese a los elementos en contra, se puede seguir en el camino de desarrollo de nuevas técnicas para mejorar las predicciones electorales. La innovación en este campo sigue siendo fundamental. En esto nos ayuda la aplicación de una lógica matemática y de las técnicas estadísticas: descubrir nuevos elementos de evaluación, que nos pueden proporcionar una información fundamental para circunscribir los factores explicativos del voto y predecir con un mayor nivel de precisión los resultados electorales. Esto es lo que hemos intentado ofrecer en este artículo: un análisis de dos diferentes técnicas para mejorar la estimación de los resultados electorales (cuando trabajamos con encuestas y con históricos de resultados).

Pero el punto más importante es que los dos modelos tienen puntos de convergencia: el resultado electoral, en un 80%, depende de factores estructurales, como la tendencia electoral histórica de una determinada área o el ciclo económico o de gobierno. No es que las campañas no muevan votos: pero mueven menos consensos de lo que habitualmente nos hacen creer. Porque el consenso se construye durante los años de una legislatura, es una sedimentación lenta y progresiva que una lluvia de unas cuantas semanas, o sea la campaña electoral, no puede erosionar fácilmente.

Y en estos últimos meses estamos notando como sube el peso de los factores estructurales de medio-largo plazo en la decisión de voto de los ciudadanos: la crisis económica ha acelerado el desgaste de los partidos al gobierno en todos los países (pensemos en España, EE.UU., Francia, Italia, Reino Unido, Alemania), y al mismo tiempo convierte en muy complicado cualquier intento de recuperación en el horizonte temporal limitado de una campaña electoral. El consenso se pierde de manera muy sencilla y es muy complicado recuperarlo. Lo estamos viendo en la campaña para las elecciones de 20-N. La comunicación de Rubalcaba, que ha empezado oficiosamente su campaña en agosto, no ha servido prácticamente para nada: frente a un desgaste tan amplio y consolidado del ejecutivo nuevas promesas electorales no son suficientes, como no funcionan campañas de alerta contra los recortes que practicarían los adversarios una vez al gobierno. Una vez perdida la credibilidad, y al final de un ciclo político, cualquier campaña electoral pierde gran parte de su potencial efecto: pero hay que reconocer los factores que mueven el consenso, es importante medir la aportación de cada uno al resultado final. Solamente de esta forma podremos conseguir nuestro objetivo, o sea mejorar la capacidad predictiva en el ámbito electoral. Los modelos matemáticos-estadísticos nos pueden proporcionar una ayuda fundamental en este sentido, porque pueden agujerear en profundidad, encontrando nuevas capas de información que un análisis superficial de los datos no dejaría entrever.

## Referencias

Ceri, P. (1997), *Politica e sondaggi*, Turín, Rosenberg&Sellier.

Díaz de Rada, V. (2010), *Comparación entre los resultados proporcionados por encuestas telefónicas y personales. El caso de un estudio electoral*, Madrid, (CIS) Centro de Investigaciones Sociológicas.

Gelman, A. & King, G. (1993), Why Are American Presidential Election Campaign Polls So Variable When Votes are So Predictable?, en *British Journal of Political Science*, 23, n. 1, (409-451).

Goidel, R. K. & Todd, G. S. (1994), The Vanishing Marginals, the Bandwagon, and the Mass Media, en *The Journal of Politics*, 56, (802-810).

Lau, R.R. & Redlawsk, D.P. (2006), *How Voters Decide: Information Processing in Election Campaigns*, Cambridge, Cambridge University Press.

Lewis-Beck, M.S & Jêrôme, B. (2010) (eds.), European Election Forecasting (Special Section), en *International Journal of Forecasting*, 26, n. 1.

Lewis-Beck, M.S. & Rice, T.W. (1992), *Forecasting Elections*, CQ Press.

Lewis-Beck, M.S & Tien, C. (2008) The Job of President and the Jobs Model Forecast: Obama for '08?, en *PS: Political Science & Politics*, (687-690).

Mannheimer, R. (eds.) (1989), *I sondaggi elettorali e le scienze politiche. Problemi metodologici*, Milán, Franco Angeli.

Mannheimer, R. & Sani, G. (1987), *Il mercato elettorale. Identikit dell'elettore italiano*, Bologna, Il Mulino.

Martínez i Coma, F. (2008), *¿Por qué importan las campañas electorales?*, Madrid, CIS (Centro de Investigaciones Sociológicas).

Mehrabian, A. (1998), Effects of Poll Reports on Voter Preferences, en *Journal of Applied Social Psychology*, 28, (2119-2130).

Natale, P. (2004), *Il sondaggio*, Roma-Bari, Laterza.

Roccatò, M. (2003), *Desiderabilità Sociale e Acquiescenza. Alcune Trappole delle Inchieste e dei Sondaggi*, Turín, LED Edizioni Universitarie.

## Notas

(1) Merece la pena recordar que en una muestra de 1.000 electores (casi todos los sondeos rondan alrededor de estos números), el margen de error estadístico, o sea la diferencia posible entre el resultado del sondeo y el valor real en la población, se coloca en un  $\pm 3\%$  (un intervalo del 6%). Esto significa que, si a un partido se atribuye el 30% de las preferencias, en el 95% de los casos el resultado real podría estar entre el 27% y el 33% de los votos. Y, en el 5% de los casos que queda, el dato podría estar fuera de este intervalo (Natale, 2004).

(2) Es el caso de las convenciones de los partidos americanos: cuando las convenciones de los Demócratas o Republicanos terminan, se disparan sus estimaciones en las encuestas de los días siguientes. Este efecto se agota en las semanas siguientes (Gelman y King 1993).

(3) Cambio facilitado, hay que subrayarlo, después de dos legislaturas de gobierno del mismo partido (la alternancia es normal después de ocho años de gobierno de los socialistas).

(4) En cualquier caso los movimientos que deciden el resultado de la mayoría de las elecciones siguen siendo los de los indecisos y de los electores de partidos menores.

(5) Se trata de la suma, para cada partido, de los valores absolutos de las diferencia entre dato estimado y dato real, dividida por el número de partidos en competición (Mannheimer, 1989).

(6) O sea el partido del presidente que ha gobernado en la última legislatura: por ejemplo en el año 2000 en EE.UU. el incumbente era Al Gore, en 2004 George W. Bush, en 2008 John McCain, en será 2012 Barack Obama.

(7) El ajuste del modelo, medido con el  $R^2$ , es muy elevado y llega al 94%.

(8) Suponiendo la ausencia de candidatos fuertes de terceros partidos o independientes.