

Puntos de Referencia

Edición online
N° 466, octubre 2017

Un modelo de votante probable para la encuesta CEP

Ricardo González T. / Bernardo Mackenna C.

Resumen

La introducción de la inscripción automática y el voto voluntario el año 2012 generó un reto importante para las organizaciones encuestadoras chilenas, ya que introdujo una fuente de incertidumbre adicional e inédita. Uno de los principales problemas que están enfrentando las encuestas a la hora de producir pronósticos electorales es el de distinguir entre los encuestados quienes participarán de la elección de quienes optarán por abstenerse.

En este texto desarrollamos un procedimiento para identificar a los votantes probables. Este mecanismo busca identificar a quienes con alta probabilidad acudirán a las urnas, a través de un conjunto de respuestas a preguntas sobre su comportamiento tanto en la próxima elección como en las pasadas.

Este documento, además describe algunos modelos de este tipo utilizados internacionalmente y sus problemas, además de los avances recientes realizados en Chile.

Nuestro modelo utiliza sólo cuatro preguntas: intención de participación, edad, interés en elección, y voto en elecciones pasadas. Con ellas, logramos aproximarnos razonablemente bien a los niveles de participación observados, tanto a nivel nacional como por sexo y edad, en las elecciones municipales de 2012 y 2016 y la primera vuelta presidencial de 2013, utilizando la información de las encuestas pre-electorales del CEP solamente.

A pesar de ello, observamos estos resultados con cautela, ya que nada garantiza que las relaciones estimadas con datos de elecciones pasadas se observen también en el futuro, además de que los modelos de participación requieren de una serie de elecciones para su validación. Además, sigue pendiente la tarea de evaluar si el modelo descrito puede también producir estimaciones razonables de los resultados de las elecciones.

Esperamos que la publicación de nuestra metodología, la primera que se publica en el contexto actual (con voto voluntario) y que puede ser replicada por cualquier interesado, invite a las demás organizaciones encuestadoras a dejar de lado la opacidad con la que manejan sus procedimientos y participen de un diálogo abierto que permita promover la investigación sobre los métodos de votantes probables y mejorar su eficacia hacia adelante.

Ricardo González T. Economista, Pontificia Universidad Católica de Chile. Coordinador Programa de Opinión Pública del CEP.

Bernardo Mackenna C. Sociólogo, Pontificia Universidad Católica de Chile, profesor e investigador en el Instituto de Sociología de la Pontificia Universidad Católica de Chile.

Introducción

La introducción de la inscripción automática y el voto voluntario, además de haber traído consigo una gran transformación al sistema electoral chileno, ha significado un gran desafío para las organizaciones encuestadoras en nuestro país, particularmente, en cuanto a su capacidad de aproximarse a los resultados de las elecciones. En efecto, uno de los principales problemas que enfrentan las encuestas a la hora de producir pronósticos electorales es el de distinguir entre los encuestados quienes participarán de la elección de quienes optarán por abstenerse. Al igual que en la experiencia internacional, los datos sugieren que los chilenos tienden a “sobre-reportar” su intención de asistir a las urnas. Por ejemplo, tras la elección municipal de 2016, en que según los registros oficiales, participó el 35 por ciento de la población en edad de votar, el 55 por ciento de los encuestados por el CEP en noviembre y diciembre de ese año declaró haber sufragado. Esta tendencia al sobre-reporte obliga a los encuestadores a buscar métodos para ajustar sus proyecciones, identificando a los llamados “votantes probables”. En este documento discutiremos una técnica para distinguir a los encuestados que eventualmente votarían de aquellos que no, en base a los datos de la encuesta CEP.

Este texto se divide en cuatro partes. La primera discute qué es un “votante probable” y describe algunos modelos de este tipo utilizados internacionalmente y sus problemas, además de los avances recientes realizados en Chile. La segunda propone un modelo “votante probable”, basado en las preguntas de las encuestas pre-electorales del CEP de 2012, 2013 y 2016. La tercera compara las estimaciones de este modelo con los “registros oficiales” a nivel nacional para analizar su desempeño en la práctica. La cuarta, en fin, concluye.

1. ¿Qué es el votante probable?

El procedimiento del “votante probable” busca identificar a los votantes a través de un conjunto de respuestas a preguntas sobre su comportamiento tanto en la próxima elección como en las pasadas.

Una de las dificultades asociada a este método es que los encuestados tienden a sobre-reportar su decisión de acudir a las urnas por la deseabilidad social de la respuesta, es decir, hay una tendencia de las personas de afirmar que votarán porque ellas creen que los “buenos ciudadanos” lo hacen, aun cuando no sea su interés hacerlo. Por esa razón, quizás la dificultad más importante que enfrenta el método del votante probable para las organizaciones encuestadoras es encontrar un conjunto de preguntas que permitan discriminar aquellas personas que sí asistirán a las urnas de aquellas que dijeron que lo harían, pero no les interesa hacerlo.

En la experiencia internacional, hay varios métodos de votante probable. Quizás el más popular es el desarrollado por Paul Perry de la organización Gallup (Perry 1960, 1973, 1979). Este procedimiento utiliza las respuestas a las siguientes siete preguntas:

- ¿Cuánto ha pensado en la elección? (Mucho, algo)
- ¿Conoce dónde vota la gente de su barrio? (Sí)
- ¿Votó en la elección local pasada? (Sí)
- ¿Con qué frecuencia vota? (Siempre, casi siempre)
- ¿Planea votar la elección del 2017? (Sí)
- En una escala de 10 puntos, ¿con qué probabilidad votará? (7-10)
- ¿Votó en la elección presidencial pasada? (Sí)

Si los encuestados dan la respuesta en paréntesis, el método le asigna un uno y cero en caso contrario. Al final, se obtiene el puntaje total del encuestado sumando el puntaje asignado por este procedimiento, de modo tal que los que alcanzan un puntaje más alto son más propensos a votar que aquellos que alcanzan un puntaje más bajo. Luego, se vislumbran dos escenarios: uno de participación baja, el cual se estima a partir de los encuestados que obtuvieron siete puntos en la escala descrita, y otro de participación alta, que considera a los encuestados que obtuvieron seis y siete puntos. En cada caso, las preferencias electorales que se consideran para los cálculos corresponden a la de los encuestados con puntajes de seis o siete, dependiendo del escenario, mientras que se descartan las respuestas de quienes obtienen puntajes de cinco o inferiores. El puntaje de corte, seis o siete, se obtuvo del análisis de las elecciones previas, pero también de lo que esperan los investigadores antes de la elección, lo que podría suponer un problema si es que la participación esperada cae dentro del rango del seis. Por ejemplo, si por algún motivo, se cree que la participación electoral estará cerca de 46 por ciento y la categoría seis contiene 9 puntos porcentuales de votantes probables, entre 42 y 51 por ciento, ¿cuál cuatro por ciento considerar para el cálculo de las preferencias electorales? El método no contiene una regla que permita discernir qué porcentaje considerar, lo que requiere, en el mejor de los casos, implementar un procedimiento adicional o, en el peor de los casos, aplicar ajustes arbitrarios, que no puedan ser reproducibles por otros investigadores.

La investigación en esta área ha girado en torno al modelo descrito, en parte por su simpleza, su validación posterior a través de sucesivas elecciones en Estados Unidos y porque el procedimiento es público. De hecho, Gallup sigue utilizando esta metodología (Gallup 2014) y el Pew Research Center

emplea un procedimiento parecido (Dimock et al. 2001). Sin embargo, la reticencia de las compañías encuestadoras en el mundo a publicar sus métodos ha limitado los avances de la investigación en esta área (Traugott 2015).

En Chile, la implementación de los modelos de votante probable es reciente. Antes de la entrada en vigencia de la ley de inscripción automática y voto voluntario, el votante probable de las encuestas correspondía a quienes estaban inscritos en los registros electorales, puesto que el voto era obligatorio. A pesar de que no todos los inscritos votaban (Cox y González 2016), el porcentaje de votantes superaba el 90 por ciento de los inscritos, suficiente

“Antes de la entrada en vigencia de la ley de inscripción automática y voto voluntario, el votante probable de las encuestas correspondía a quienes estaban inscritos en los registros electorales, puesto que el voto era obligatorio.”

para tener una aproximación razonable al electorado que acudiría a las urnas el día de la elección. En las primeras elecciones después de implementada la reforma del 2012, buena parte de las organizaciones encuestadoras no implementaron estos modelos, posiblemente porque no había elecciones previas bajo el nuevo sistema que permitieran ajustar las respuestas a las preguntas de las encuestas con los datos oficiales de participación electoral. Michael Traugott, profesor de la Universidad de Michigan en Ann Harbor y experto en uso y validez de las encuestas políticas y predicción de elecciones, decía acerca del caso chileno que “un cambio de un sistema de voto obligatorio a uno voluntario implica que la

participación se reducirá, haciendo la estimación del votante probable un proceso más complicado”, ya que durante la transición de un sistema a otro “puede que no existan modelos adecuados de comportamiento de voto por subgrupos de población que guíen la definición o especificación del votante probable” (Traugott 2015).

A diferencia de las elecciones con voto voluntario pasadas, este año las organizaciones encuestadoras están estimando modelos de votante probable, pero tal como en la experiencia internacional, no han revelado los detalles de sus procedimientos a la fecha. Algunas compañías muestran las preguntas que utilizan, pero no la forma en que las combinan para generar el votante probable. En Chile, las compañías que revelan parte de la información están mezclando, de forma misteriosa todavía, las respuestas a las siguientes preguntas:

- Declaración de intención de voto en las elecciones presidenciales de 2017.
- Declaración de participación en las últimas elecciones con voto voluntario (algunos consideran las municipales de 2016 solamente y otras esa junto a la elección presidencial de 2013).
- Declaración de interés en la próxima elección presidencial (algunos usan la respuesta directa a una pregunta sobre tal interés, mientras otros consideran identificación con la política, que no sabemos si es identificación ideológica o partidaria).

Por supuesto, también hay compañías que no revelan ningún detalle de los procedimientos utilizados. En ambos casos, tampoco sabemos si los métodos de votante probable de las organizaciones encuestadoras en Chile pueden producir estimaciones razonables si se comparan con los pocos datos oficiales disponibles de las elecciones pasadas con voto voluntario.

2. El modelo

Para elaborar el modelo de votante probable tomaremos las preguntas del módulo de participación electoral de las encuestas pre-electorales CEP, esto es, los estudios de opinión N° 67, 70 y 77, los que corresponden a julio-agosto 2012, septiembre-octubre 2013, y julio-agosto 2016, respectivamente.

A continuación, describiremos los pasos y preguntas utilizadas para construir el votante probable para la encuesta CEP. El diagrama de flujo de la página siguiente, resume el procedimiento.

PASO 1: *Se clasificará como votante a la persona que haya respondido “Sí, con toda seguridad irá a votar” a la pregunta “Y en su caso, ¿irá Ud. a votar en las próximas elecciones presidenciales de 2017?”*

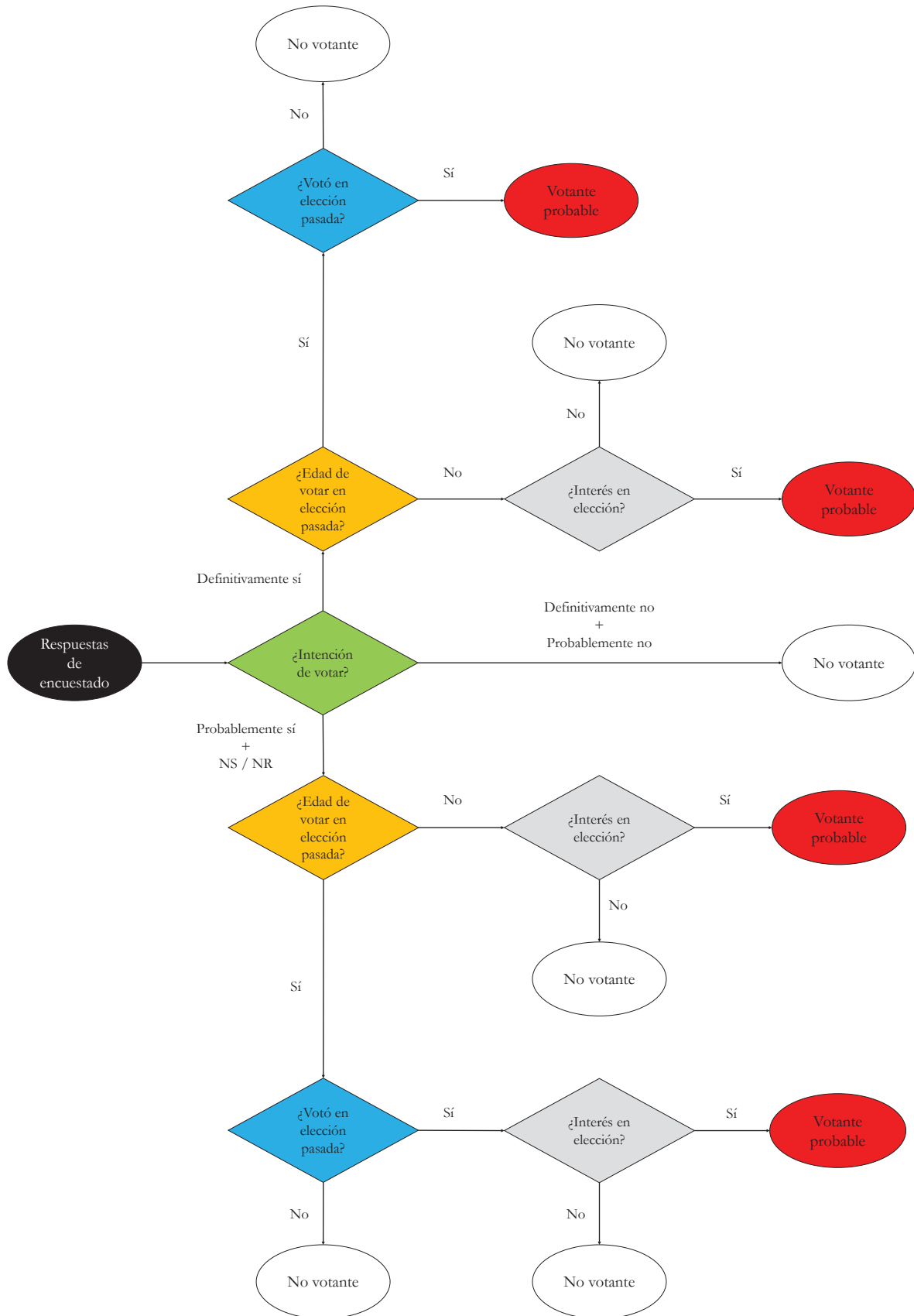
El problema de utilizar las respuestas a esta pregunta sin ajustes, como decíamos antes, es que los individuos tienden a sobre-reportar la decisión de votar por deseabilidad social. Esto significa que varias personas que afirman que votarán con seguridad podrían no hacerlo.¹ Para reconocer ese hecho, introducimos el siguiente ajuste.

PASO 2: *Se clasificará como no-votante a la persona que haya respondido “Sí, con toda seguridad irá a votar” a la pregunta “Y en su caso, ¿irá Ud. a votar en las próximas elecciones presidenciales de 2017?”, pero que no haya votado en la elección presidencial pasada.*

La literatura muestra que las personas con mayores probabilidades de votar son aquellas que han

¹ En la encuesta CEP, anterior a las elecciones municipales del 2016, un 27 por ciento de los encuestados declaró que votaría con toda seguridad, inferior al 35 por ciento que acudió a votar en esas elecciones, lo que podría sugerir evidencia de sub-reportaje. Sin embargo, en esa misma encuesta, un 24 por ciento afirmó que probablemente si votaría en esas elecciones, por lo tanto, un 51 por ciento tenía intenciones de participar en esa votación, lo que es evidencia de sobre-reportaje.

DIAGRAMA DE FLUJO



votado en el pasado, principalmente, porque votar parece ser un hábito que se adquiere tras haber participado en elecciones previas (Green y Shachar 2000; Kanazawa 2000; Denny & Doyle, 2009). Así, es posible que aquellas personas que declaran que votarán con seguridad, pero no han votado en la elección previa, no acudan a las urnas, por eso los clasificamos como no-votantes. Sin embargo, este ajuste descarta a un grupo de jóvenes que no eran elegibles para sufragar cuatro años atrás porque no tenían la edad para hacerlo. Para reconocer ese hecho, incorporamos la siguiente regla.

Paso 3: *Los jóvenes que no tenían edad para votar en la elección presidencial pasada se clasificarán como votantes si respondieron “Sí, con toda seguridad irá a votar” a la pregunta “Y en su caso, ¿irá Ud. a votar en las próximas elecciones presidenciales de 2017?” y han declarado estar “muy” o “bastante” interesados en la elección presidencial de 2017.*

Naturalmente, a los jóvenes que no tenían edad para votar en la elección presidencial pasada no les podemos exigir haber votado en el pasado, ya que no eran elegibles. También sabemos que la intención de votar tiene sobre-declaración, por lo tanto, es necesario que los jóvenes cumplan un requisito adicional para atenuar este problema. Ese requisito es el interés en la elección futura. Si declaran estar muy o bastante interesados en la elección presidencial de 2017, además de que declararon votar con seguridad en la elección futura, entonces se clasificarán como votantes.

Paso 4: *Se clasificará como votante a la persona que haya respondido “Probablemente sí” a la pregunta “Y en su caso, ¿irá Ud. a votar en las próximas elecciones presidenciales de 2017?”, haya votado en la elección presidencial pasada y haya declarado estar “muy” o “bastante” interesado en la elección presidencial de 2017.*

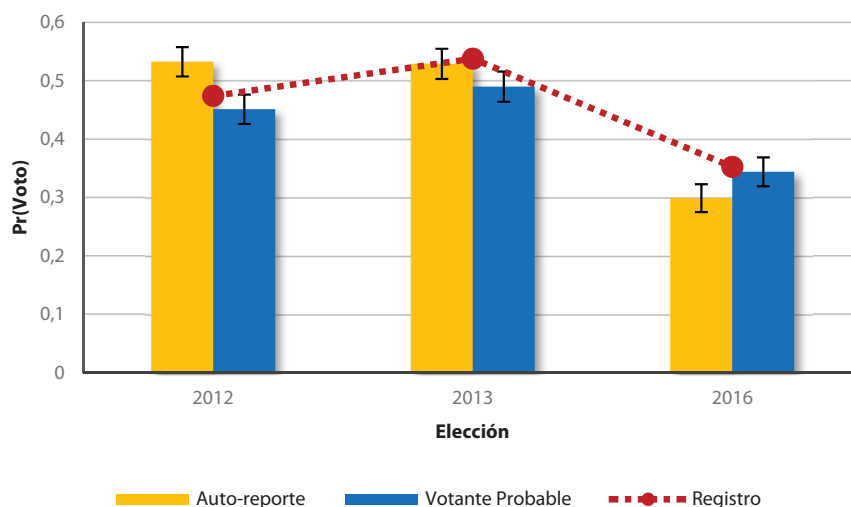
Pero también hay encuestados que no están seguros de votar en las elecciones próximas. Naturalmente, habrán algunos que acudirán a las urnas, mientras otros no lo harán. El grupo de electores inseguros de sufragar corresponde a quienes responden “Probablemente sí” a la pregunta “Y en su caso, ¿irá Ud. a votar en las próximas elecciones presidenciales de 2017?”. Por los motivos expuestos anteriormente, nos parece probable que alguien inseguro, pero que haya votado en la elección presidencial pasada y que esté muy o bastante interesado en la elección futura, eventualmente se decidirá por votar el día de la elección.

Paso 5: *Se clasificará como votante a la persona que no sepa o no haya respondido la pregunta “Y en su caso, ¿irá Ud. a votar en las próximas elecciones presidenciales de 2017?”, haya votado en la elección presidencial pasada y haya declarado estar “muy” o “bastante” interesado en la elección presidencial de 2017.*

Tal como en el caso anterior, podríamos decir que aquellos individuos que responden “no sé” o no contestan la pregunta “Y en su caso, ¿irá Ud. a votar en las próximas elecciones presidenciales de 2017?”, están inseguros de participar en las elecciones. Por eso, creemos que alguien inseguro, pero que haya votado en la elección presidencial pasada y que esté muy o bastante interesado en la elección futura, probablemente decidirá votar el día de la elección.

Por último, aquellos que respondan “Probablemente no” o “No, con toda seguridad” a la pregunta “Y en su caso, ¿irá Ud. a votar en las próximas elecciones presidenciales de 2017?” se clasificarán como no-votantes, sin importar su comportamiento pasado o su interés en las elecciones futuras.

Ahora analizaremos cuán bien (o mal) se relacionan las cifras del votante probable recién descrito con los registros oficiales.

GRÁFICO 1: Estimaciones de participación electoral usando el votante probable

Nota: Porcentaje de encuestados que declaran que votarán con toda seguridad, indicador de votante probable, con intervalos de confianza del 95%, y datos de participación de registros oficiales, según elección. Las estimaciones no consideran el uso de ponderados, por lo tanto, estas cifras difieren ligeramente de lo reportado en las conferencias de prensa.

Fuente: elaboración propia a partir de datos de CEP, SERVEL e INE.

3. ¿Cuán precisas son las estimaciones del votante probable?

Cabe destacar, en primer lugar, que no podemos comparar directamente las estimaciones del votante probable con los datos de registro a nivel individual porque el SERVEL no publica esa información. Las cifras agregadas de participación electoral, a nivel nacional, por sexo y edad (año de nacimiento) son públicas y por esa razón, sólo podemos comparar nuestro modelo con ellas. También contrastaremos las cifras del votante probable con la intención de votar declarada por los propios individuos, reportadas tradicionalmente por las encuestas, sin ajustes de ningún tipo.

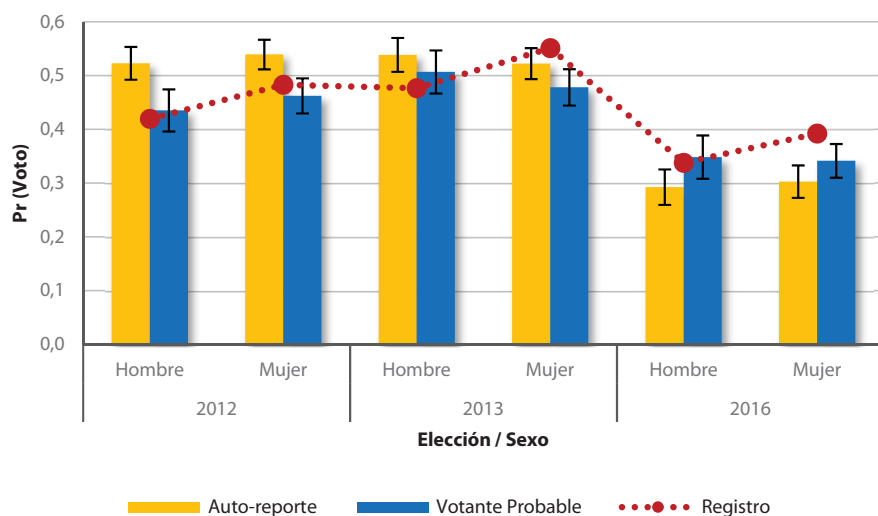
En el Gráfico 1 presentamos los promedios de participación electoral predicha siguiendo dos estrategias de estimación. Para cada elección, la primera barra (“Auto-reporte”) exhibe el promedio de participación calculado a partir de lo declarado

por los encuestados mismos en las encuestas pre-electorales, sin ningún ajuste, y la segunda barra (“Votante Probable”) corresponde a los resultados de la estrategia de estimación descrita en la sección anterior. Luego, las cifras son contrastadas con la tasa de participación electoral, calculada a partir de la cantidad de votos emitidos de acuerdo con el SERVEL y el tamaño de población en edad de votar según proyecciones del INE².

En general, el desempeño de nuestra estimación de votante probable muestra un ajuste razonable: subestima ligeramente la participación en 2012 y 2013, y da un pronóstico acertado para el 2016. En promedio, logra mejorar el ajuste a los datos de registro si lo comparamos con el obtenido a partir del uso exclusivo de las cifras de auto-reporte.

El Gráfico 2 exhibe la participación electoral por sexo que resulta de la intención de acudir a las urnas que reportan los encuestados exclusivamente y la del modelo de votante probable (sin ponderación). A simple vista, al mirar la distribución de participación por sexo de la figura, el votante probable se aproxima razonablemente a los datos de registro. No obstante, en 2013 y 2016 el modelo subestima ligeramente la participación femenina. Las razones detrás de estas diferencias requieren estudios más profundos, pues los indicadores que utilizamos para conformar la variable de votante probable no muestran diferencias significativas por sexo.

² Debido a la persistencia de problemas en los registros de inscritos de SERVEL, preferimos basarnos en la población en edad de votar, que es, además, una práctica común en estudios de participación electoral (Geys, 2006).

GRÁFICO 2: Estimaciones de participación electoral usando el votante probable, por sexo

Nota: Porcentaje de encuestados que declaran que votarán con toda seguridad, indicador de votante probable, con intervalos de confianza del 95%, y datos de participación de registros oficiales, según sexo y elección. Las estimaciones no consideran el uso de ponderados, por lo tanto, estas cifras difieren ligeramente de lo reportado en las conferencias de prensa.

Fuente: elaboración propia a partir de datos de CEP, SERVEL e INE.

sugiere que el modelo de votante probable descrito en este texto provee un reflejo razonablemente fidedigno de la conducta electoral de los chilenos. De hecho, el error absoluto promedio de predicción (diferencia entre la estimación con la encuesta versus datos de registro) se reduce de 4 puntos porcentuales (usando la pregunta de intención de participación declarada por los mismos encuestados sin ajustes) a 2,6 por ciento con nuestro indicador de votante probable. Aún queda pendiente la tarea de evaluar si esta estrategia es capaz de mejorar la precisión de los estimadores de los resultados de las elecciones.

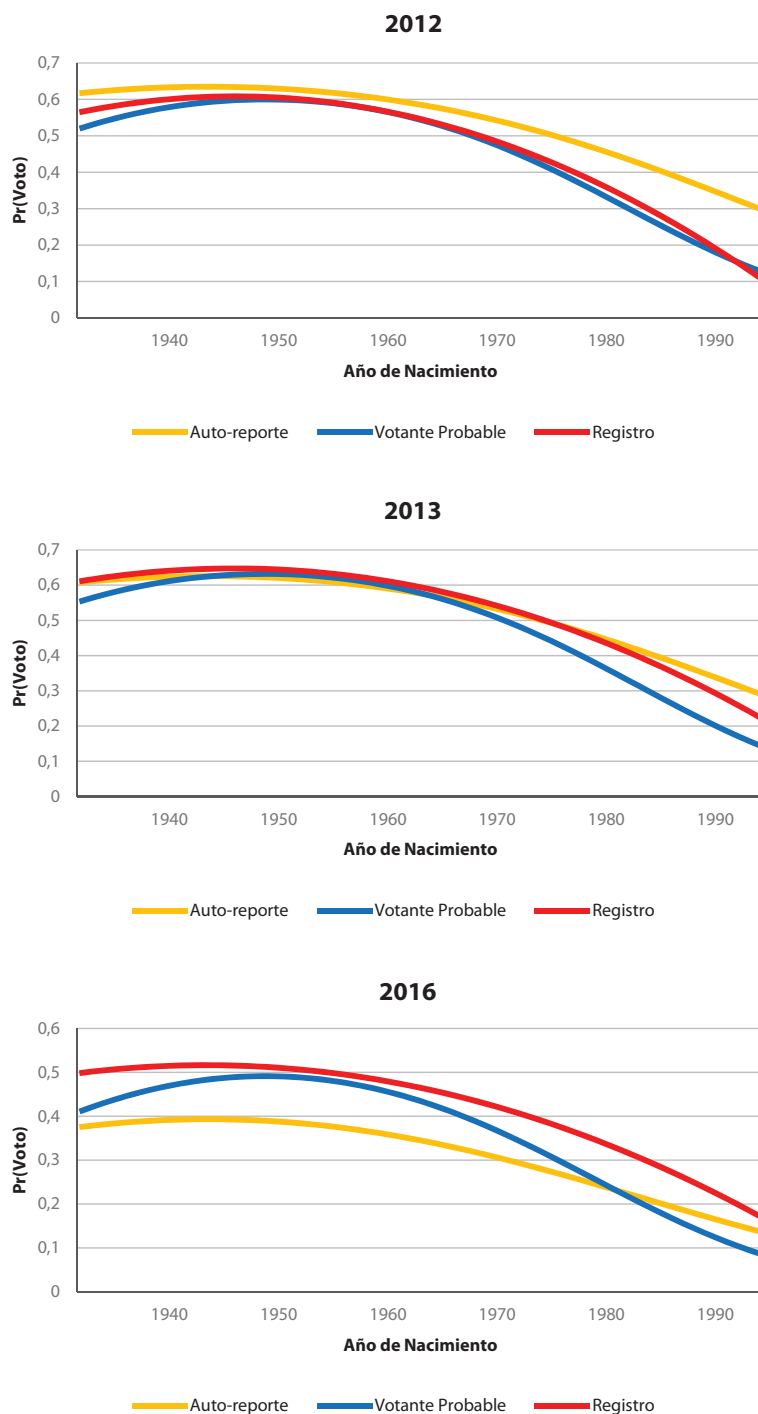
Finalmente, el Gráfico 3 ilustra la participación electoral por año de nacimiento, estimada a partir de la intención de ir a votar que reportan los encuestados, sin ajustes, y el votante probable. La distribución por año de nacimiento nos lleva a una conclusión similar: el indicador de votante probable se ajusta razonablemente bien a los datos de registro. De hecho, en promedio, este modelo produce pronósticos algo más precisos de los datos de registro, aunque estas diferencias resultan estadísticamente significativas a los niveles convencionales de confianza para la gran mayoría de los grupos de edad.

En síntesis, nuestra evaluación de esta estrategia de estimación de votantes probables entrega resultados prometedores: con sólo cuatro preguntas (intención de participación, edad, interés en elección, y voto en elecciones pasadas) podemos predecir la participación electoral con relativa precisión, lo que

4. Comentarios finales

En este texto hemos elaborado un modelo de votante probable con el objeto de definir al grupo de encuestados que con probablemente acudirá a votar y al que se abstendrá. Hemos mostrado que el modelo produce estimaciones cercanas a los niveles de participación observados en las elecciones municipales de 2012 y 2016 y la primera vuelta presidencial de 2013, utilizando la información de las encuestas pre-electorales del CEP, y se aproxima con razonable precisión a la participación electoral por sexo y edad. A pesar de ello, observamos estos resultados con cautela, ya que nada garantiza que las relaciones estimadas con datos de elecciones pasadas se observen también en el futuro. En palabras de Traugott (2015), "los nuevos modelos de participación pueden requerir una serie de elecciones para su validación y establecer confianza en ellos." Además, sigue pendiente la tarea de evaluar

GRÁFICO 3: Estimaciones de participación electoral usando el votante probable, por año de nacimiento



Nota: Porcentaje de encuestados que declaran que votarán con toda seguridad, indicador de votante probable y datos de participación de registros oficiales, según año de nacimiento y elección. Las estimaciones no consideran el uso de ponderados, por lo tanto, estas cifras difieren ligeramente de lo reportado en las conferencias de prensa.

Fuente: elaboración propia a partir de datos de CEP, SERVEL e INE.

si los modelos aquí descritos pueden también producir estimaciones razonables de los resultados de las elecciones.

Por otra parte, esperamos que la publicación de nuestra metodología, la primera que se publica en el contexto actual (con voto voluntario) y que puede ser replicada por cualquier interesado, invite a las demás organizaciones encuestadoras a dejar de lado la opacidad con la que manejan sus procedimientos y participen de un diálogo abierto que permita promover la investigación sobre los métodos de votantes probables y mejorar su eficacia hacia adelante.

Por último, dado que las estimaciones del modelo se aproximan razonablemente bien a los niveles de participación efectivos y que el CEP, a la fecha de publicación de este documento, ha realizado dos encuestas pre-electorales, más de algún lector se estará preguntando cuántos chilenos votarán en la elección de noviembre de 2017 y sus características, al menos por sexo y edad. Sin embargo, todavía no es posible responder estas preguntas con precisión. Es bien sabido que los problemas de medición en las encuestas electorales se minimizan a medida que se acerca el día de la elección (Traugott 2015), por lo tanto, habrá que esperar a la encuesta más cercana a ese día para hacer un análisis similar al realizado en este texto para las elecciones de este año.

5. Referencias

Artículos y documentos

- Cox, L. y González, R. (2016). "Cambios en la participación electoral tras la inscripción automática y el voto voluntario". CEP, Debates de Política Pública N° 14. <http://bit.ly/2dwgrQ2>
- Denny, K., & Doyle, O. (2009). Does voting history matter? Analysing persistence in turnout. *American Journal of Political Science*, 53(1), 17-35.
- Dimock, M., Keeter, S., Schulman, M. y Miller, C. (2001). A Voter Validation Experiment: Screening for Likely Voters in Pre-Election Surveys. *Report of the Pew Research Center*. <http://www.people-press.org/files/2001/05/12.pdf>.
- Geys, B. (2006). Explaining voter turnout: A review of aggregate-level research. *Electoral Studies*, 25(4), 637-663.
- Green, D. P., & Shachar, R. (2000). Habit formation and political behaviour: Evidence of consuetude in vo-

ter turnout. *British Journal of Political Science*, 30(4), 561-573.

Kanazawa, S. (2000). A new solution to the collective action problem: The paradox of voter turnout. *American Sociological Review*, 433-442.

Perry, P. (1960). The Election Procedures of the Gallup Poll. *Public Opinion Quarterly*, 24: pp. 531-42.

Perry, P. (1973). A Comparison of the Preferences of Likely Voters and Likely Nonvoters. *Public Opinion Quarterly*, 37: pp. 99-109.

Perry, P. (1979). "Certain Problems in Election Survey Methodology." *Public Opinion Quarterly*, 43: pp. 312-25.

Traugott, M. W. (2015). Métodos alternativos para la estimación de resultados electorales. *Estudios Públicos*, (137), 7-42.

Prensa y sitios web

- Gallup. 2014. "Understanding Gallup's Likely Voter Procedures for Presidential Elections." <http://www.gallup.com/poll/111268/how-gallups-likely-votermodels-work.aspx>.