

Predicciones ciudadanas de las elecciones presidenciales mexicanas, 2000-2024

Andreas Erwin Murr*

RESUMEN: La predicción científica de elecciones tiene una tradición larga y creciente en democracias bien establecidas; sin embargo, no ha progresado lo suficiente en América Latina. Para abonar a esto, analizamos las predicciones ciudadanas de las elecciones presidenciales mexicanas de 2000 a 2024. Desarrollamos nuevas herramientas estadísticas para estudiar si las predicciones ciudadanas revelan mejor los resultados electorales que el azar y que las encuestas de intención de voto. A partir de 55 encuestas, encontramos que los ciudadanos se vuelven mejores con el tiempo en predecir resultados electorales. Mientras que los ciudadanos, al inicio de la democracia en México, en el año 2000, predecían peor que el azar, en elecciones presidenciales recientes las han predicho de manera correcta. En la actualidad, dichas predicciones están a la par de las encuestas de intención de voto. A pesar de que los ciudadanos son ahora mejores que el azar al predecir el porcentaje de votos, siguen teniendo peores resultados que las encuestas de intención de voto. Para entender por qué, analizamos el sesgo y la varianza de las predicciones ciudadanas por primera vez en México. Encontramos que las principales razones por las que sus predicciones electorales han empeorado son la socialización bajo un régimen de partido único, la falta de experiencia con elecciones libres y justas, y las preocupaciones infundadas sobre si las elecciones son limpias o no.

Palabras clave: predicciones ciudadanas, transición democrática, predicción electoral, encuestas de opinión pública, socialización política.

Citizen forecasts of Mexican presidential elections, 2000-2024

ABSTRACT: Scientific election forecasting has a long and growing tradition in established democracies, but it has not made much progress in Latin America. To make progress, we examine citizen forecasts of Mexican presidential elections from 2000 to 2024. We develop novel statistical tools to study whether citizen forecasts are better at forecasting than chance and than vote intention polls.

* **Andreas Erwin Murr** es profesor investigador titular en la División de Estudios Políticos del Centro de Investigación y Docencia Económicas (CIDE). Carretera México-Toluca 3655, col. Lomas de Santa Fe, CDMX, 01210, México. Tels: +52 55 5727 y 9800-2234. Correo-e: andreas.murr@cide.edu. Profesor asociado, Department of Politics and International Studies, University of Warwick. Coventry, CV4 7AL, UK. Tel: +44 (0)24 765 72959. Correo-e: a.murr@warwick.ac.uk. ORCID: 0000-0002-9536-0118. Website: andreamurr.com. Agradezco a Ericka G. Rascón Ramírez y a dos revisores anónimos por sus valiosos comentarios y sugerencias, así como a Felipe de la O López por haberme indicado la encuesta utilizada en este trabajo. Traducción de Ana Inés Fernández y Andrea Ávila.

Artículo recibido el 3 de febrero de 2024 y aceptado para su publicación el 12 de noviembre de 2024.

Using 55 surveys, we find that citizens became better over time. While they forecasted who wins worse than chance at the dawn of democracy in 2000, they forecasted recent presidential elections correctly and are now on par with vote intention polls. While citizens are also better than chance at forecasting vote shares, they remain, however, worse than vote intention polls. To understand why, we analyze the bias and variance of citizen forecasts for the first time. We find that the main reasons for their worse performance are socialization under a one-party rule, inexperience with free-and-fair elections, and unwarranted concerns about the cleanliness of the elections.

Keywords: citizen forecasting, democratic transition, election forecasting, public opinion polls, political socialization.

La predicción científica de las elecciones tiene una tradición larga y creciente en las democracias bien establecidas. Actualmente, las predicciones electorales son comunes en Francia, Alemania, Gran Bretaña y Estados Unidos (Dassonneville y Tien, 2020; Fisher y Lewis-Beck, 2016; Jérôme y Graefe, 2021; Nadeau *et al.*, 2022). Se han usado, probado y comparado varios métodos de predicción; entre los más importantes están los mercados de apuestas, los modelos estructurales basados en variables políticas y económicas, y los sondeos de intención de voto y de predicciones ciudadanas. Al compararlos, la predicción ciudadana figura entre los más precisos (Graefe, 2014; Murr *et al.*, 2021).

Sin embargo, la predicción electoral no ha progresado lo suficiente en América Latina, Ratto *et al.* (2022: 7) señalan que “no hay mucho, salvo por artículos dispersos sobre campañas o datos sobre sondeos”, y añaden que “la bibliografía sobre predicciones en América Latina se ha enfocado en entender los sesgos de los sondeos en campaña (que llevan a errores de predicción en las intenciones de voto, a errores de muestreo y a errores en encuestas) y a corregirlos”. Si bien México lleva la delantera en la región con cuatro artículos sobre predicciones electorales, todos ellos se concentran en las encuestas de intención de voto (Ratto *et al.*, 2022: 9).

En respuesta a esta situación, Ratto *et al.* (2022) editaron un número especial sobre predicciones electorales en América Latina con investigaciones nuevas que utilizaron modelos estructurales, intenciones de voto o ambas cosas en Argentina, Brasil, Chile, Jamaica y Perú (Arce y Vera, 2022; Bertholini *et al.*, 2022; Bunker, 2022; Charles *et al.*, 2022; Ratto y Lewis-Beck, 2022). A pesar de estos esfuerzos, Ratto *et al.* (2022: 11) concluyen la introducción al número especial con el siguiente llamado:

Más allá de los enfoques presentados en esta colección, que involucran modelos estadísticos estructurales y encuestas de opinión pública más tradicionales, ¿sería posible aún explorar nuevos enfoques, como la predicción ciudadana, sobre la cual hay cada vez más bibliografía en otras partes del mundo democrático (Murr, 2016; Murr *et al.*, 2021)? La idea de la predicción ciudadana resulta bastante simple: preguntar a una muestra de votantes quién creen que va a ganar. ¿Qué revelarán los datos latinoamericanos al respecto? En este momento, no lo sabemos.

En este texto, responderemos al llamado examinando las predicciones ciudadanas de las elecciones presidenciales mexicanas de 2000 a 2024. La primera parte examina la bibliografía sobre predicciones electorales en general y predicciones ciudadanas en particular. Posteriormente, evaluamos la precisión de las predicciones ciudadanas en las elecciones presidenciales mexicanas más recientes de dos maneras: primero, analizando las predicciones ciudadanas de qué partido o candidato ganará la elección y, después, analizando las predicciones ciudadanas de la proporción de votos que obtendrá cada candidato. En el Estudio 1, examinamos si los ciudadanos predicen mejor que el azar y que las encuestas de intención de voto. En el Estudio 2, examinamos si los ciudadanos predicen mejor que el azar y que las encuestas de intención de voto los porcentajes de voto para cada candidato. Luego, analizamos qué ciudadanos son mejores que otros para predecir los porcentajes de voto. En el Estudio 3, examinamos qué ciudadanos tienen más sesgos y varianza en sus predicciones y cuáles menos. Por último, revisamos las implicaciones de esta investigación para predicciones electorales en México y otras democracias.

PREDICCIONES ELECTORALES EN DEMOCRACIAS BIEN ESTABLECIDAS

Las predicciones electorales pretenden predecir resultados electorales relevantes antes de la fecha de las elecciones (Lewis-Beck, 2005). Un resultado de interés típico es el porcentaje de votos que recibirá cada partido político, pero también son útiles otros resultados, como quién ganará la elección (Murr, 2011). La evaluación de las predicciones electorales por lo general se enfoca en la precisión y en el periodo de anticipación, aunque a veces también se toman en cuenta el ahorro de recursos y la reproducibilidad (Lewis-Beck, 1985). Para las predicciones sobre el ganador, la medida de precisión más común es si la predicción fue correcta o incorrecta. Para las predicciones de los porcentajes de voto, la medida de precisión más común es la raíz del error cuadrático medio (RMSE, por sus siglas en inglés) que, para cada partido o candidato, resta la predicción de la proporción de votos real obtenida, la eleva al cuadrado y luego le saca raíz cuadrada a la suma. Al comparar métodos de predicción, los investigadores revisan cuáles predijeron bien al ganador y cuáles no, o cuál tiene una RMSE más baja. Pero, como veremos más adelante, es mejor considerar también la incertidumbre de cada estimación al hacer una comparación de ese tipo.

Los enfoques más comunes a la predicción electoral incluyen mercados de apuestas, modelos estructurales,¹ y sondeos de intención de voto y de predicciones ciudadanas (Stegmaier y Norpoth, 2013). Como se indicó en la introducción, la predicción de elecciones en México se ha enfocado exclusivamente en encuestas de intenciones de voto (Beltrán y Valdivia, 1999; Moreno *et al.*, 2014; Cantú *et al.*, 2016; Bunker,

¹ El término “modelo estructural” en este caso se refiere a un modelo de regresión de un resultado electoral con variables económicas y políticas como predictores.

2020), pero las predicciones ciudadanas son otro enfoque de sondeo importante: se les pregunta a los ciudadanos quién creen que va a ganar o qué porcentaje de los votos creen que obtendrá un partido o candidato (Lewis-Beck y Skalaban, 1989; Boon, 2012). A diferencia de los sondeos de intención de voto, en los que se les pregunta a los ciudadanos sus propias intenciones de voto, las predicciones ciudadanas invitan a la gente a considerar las intenciones de voto propias y las de sus amigos, familiares, colegas, etc., al predecir el resultado de la elección (Leiter *et al.*, 2018); por lo tanto, las predicciones ciudadanas deberían ser más precisas que los sondeos de intención de voto, y muchas veces sí lo son (Graefe, 2014; Murr, 2017; Murr *et al.*, 2021).

LAGUNAS EN LA BIBLIOGRAFÍA SOBRE PREDICCIONES CIUDADANAS

La bibliografía sobre predicciones ciudadanas, desde nuestro punto de vista, tiene cuatro lagunas importantes.

1. ¿Cómo predicen los ciudadanos en las democracias emergentes?

La predicción ciudadana se ha estudiado bastante en las democracias bien establecidas, pero sabemos muy poco sobre las democracias emergentes. Las investigaciones han recabado gran cantidad de pruebas que demuestran que los ciudadanos predicen bien los resultados en varios sistemas políticos como Canadá, Francia, Alemania, Gran Bretaña y Estados Unidos (Dufresne *et al.*, 2022; Lewis-Beck y Tien, 1999; Murr, 2016; Murr y Lewis-Beck, 2022; Temporao *et al.*, 2019). Por lo anterior, desde hace varias décadas sabemos que las predicciones ciudadanas funcionan bien en sistemas presidenciales, semipresidenciales y parlamentarios. Así, entendemos cómo funcionan las predicciones ciudadanas en distintas circunstancias, como elecciones reñidas, victorias aplastantes, reelecciones o cambios en el gobierno (Lewis-Beck y Skalaban, 1989; Murr, 2015; 2016; Murr y Lewis-Beck, 2022). No obstante, sabemos poco sobre cómo predicen los ciudadanos en los albores o primeros años de una democracia.

A la fecha, sólo tenemos un estudio que analiza las predicciones ciudadanas en los inicios de la democracia en un país, aunque el escenario es inusual. Leiter *et al.* (2018) estudian las predicciones electorales de los alemanes del este y del oeste para la primera elección tras la unificación en 1990. Sorprende ver que los alemanes del este (que crecieron bajo un régimen autoritario) predijeron igual o mejor al ganador de la elección federal que los alemanes del oeste (que crecieron bajo un régimen democrático). En otras palabras, manteniendo estable el contexto institucional y electoral, los ciudadanos para quienes las elecciones democráticas eran nuevas predijeron igual de bien el resultado que los ciudadanos con una extensa experiencia democrática.

Sin embargo, resulta difícil generalizar este hallazgo para todas las democracias emergentes, pues el escenario de la Alemania unificada es distinto al de la mayoría de los países que transitan de un régimen autoritario a uno democrático. En Alemania,

la parte occidental tenía varias décadas de experiencia democrática y el nuevo país unificado adoptó su sistema político. Como resultado, los alemanes orientales tenían mucha historia electoral de la cual aprender y podían confiar en que el sistema político unificado funcionaría como lo hacía antes en la parte occidental. En contraste, en muchas democracias emergentes no hay elecciones de las cuales aprender o no se sabe hasta qué punto el nuevo sistema difiere del viejo, y todo eso dificulta predecir al ganador de la elección. Por lo tanto, necesitamos analizar más casos de democracias emergentes para entender si los ciudadanos pueden predecir elecciones democráticas inmediatamente y, si no, cómo los ciudadanos, de manera individual o colectiva, logran o no aprender a predecir elecciones. En la siguiente sección analizaremos el caso de México para responder estas cuestiones.

2. ¿Los ciudadanos predicen mejor los porcentajes de votos que el azar?

Los ciudadanos han logrado predecir al ganador mejor que el azar (Lewis-Beck y Skalaban, 1989; Lewis-Beck y Tien, 1999; Murr, 2016). Sin embargo, no tenemos pruebas ni herramientas estadísticas comparables para las predicciones ciudadanas de los porcentajes de votos. Lewis-Beck y Skalaban (1989) fueron los primeros en preguntarse si los ciudadanos predecirían mejor que el azar quién ganaría la elección presidencial de Estados Unidos. Gracias a la naturaleza bipartidista de dichas elecciones, revisaron si más de la mitad de los ciudadanos había acertado en su predicción del resultado y se dieron cuenta de que, en ocho elecciones, de 1956 a 1984, 69 por ciento de los ciudadanos en promedio había acertado. Después de eso, Lewis-Beck y Tien (1999) realizaron pruebas estadísticas basadas en una distribución binomial para descartar que el resultado fuera producto del azar. Mientras que ciertos investigadores han calculado las medidas de precisión para las predicciones ciudadanas de porcentajes de votos (Boon, 2012; Murr, 2017), no hicieron pruebas para averiguar si los ciudadanos acertaban más que el azar. No hay una prueba estadística para responder esta pregunta, pero en el siguiente punto la desarrollaremos.

3. ¿Los ciudadanos predicen mejor los porcentajes de votos que los sondeos de intención de voto?

Las predicciones ciudadanas de quién ganará han demostrado estar entre los modelos de predicción más precisos en Estados Unidos y Gran Bretaña (Graefe, 2014; Murr *et al.*, 2021). Por ejemplo, Murr *et al.* (2021) comparan predicciones ciudadanas con modelos basados en intención de voto para predecir las proporciones de escaños para cada partido en las elecciones generales británicas. Descubrieron que las predicciones ciudadanas dan mejores resultados que los sondeos de intención de voto. Mientras que las predicciones ciudadanas sobre quién ganará funcionan bien, sólo existen pruebas contrastantes sobre la calidad de las predicciones ciudadanas sobre porcentajes de voto u otros resultados (cuasi) continuos.

Por ejemplo, Boon (2012) les pidió a los ciudadanos que predijeran la proporción de votos para los principales partidos de las elecciones generales británicas de 2010. En esa ocasión, encontró que los ciudadanos predijeron mejor los resultados que los sondeos de intención de voto; sin embargo, para las elecciones generales británicas de 2015, Murr (2017) encontró lo contrario. Murr (2017) también analizó datos de Nueva Zelanda en 1999 y 2002 tras pedirles a los ciudadanos que predijeran el número de escaños que obtendría cada partido. En este caso, las predicciones directas para el número de escaños resultaron peores que los sondeos de intención de voto. De igual forma, en las elecciones federales alemanas de 2014, las predicciones ciudadanas de porcentajes de votos resultaron peores que las intenciones de voto y que los modelos estructurales, pero mejores que los mercados de apuestas (Ganser y Riordan, 2015; Graefe, 2016). En cambio, en las elecciones federales de 2017, las predicciones ciudadanas directas sobre los porcentajes de voto para el partido Alternativa para Alemania (AfD) predijeron mejor los resultados que otros métodos de preguntas indirectas, como la intención de voto (Lehrer *et al.*, 2019). Por lo tanto, los resultados en Alemania también son contrastantes entre sí.

En general, la mayoría de las pruebas que aquí se muestran se basan sólo en una o dos elecciones, y las comparaciones que se hacen comparan estimados de porcentajes e ignoran la incertidumbre de los estimados (con excepción de Lehrer *et al.*, 2019). En el siguiente punto, analizamos cuatro elecciones presidenciales al mismo tiempo y desarrollamos una herramienta estadística para probar si un enfoque funciona mejor que otro.

4. ¿Qué determina la capacidad para predecir porcentajes de voto?

Varios estudios realizados en democracias bien establecidas han observado diversas correlaciones entre las predicciones de los ciudadanos sobre quién ganará y el resultado real (Lewis-Beck y Tien, 1999; Murr, 2011; 2015), pero no sabemos mucho sobre lo que explica la capacidad de predicción de los porcentajes de voto en las democracias emergentes. Murr (2015), por ejemplo, analizó predictores de predicciones correctas sobre qué candidato presidencial ganaría cada estado en las elecciones presidenciales de 1952 a 2012. Entre otras cosas, los resultados revelaron que las personas de edad avanzada, con mayor nivel educativo y aquellas que no predicen que el ganador sea el partido con el que se identifican tienen mayor probabilidad de predecir correctamente al ganador. Asimismo, Murr (2015) observó que a los ciudadanos se les dificulta más predecir cambios en el ganador entre una elección y la siguiente, así como en elecciones con un margen de victoria estrecho. No obstante, no hay estudios comparables sobre predicciones ciudadanas de los porcentajes de voto. ¿Importan los mismos predictores para los porcentajes de voto? ¿Importan otros predictores distintos en las democracias emergentes? En la siguiente sección analizamos los predictores de errores en las predicciones de los porcentajes de voto

en cuatro elecciones presidenciales mexicanas. Hacemos dos aportaciones: proponemos un modelo para estimar el sesgo y la varianza en las predicciones de los porcentajes de voto, e incluimos predictores que deberían recoger características importantes en las democracias emergentes, como, por ejemplo, si las elecciones se perciben como limpias o no.

LAS ELECCIONES PRESIDENCIALES MEXICANAS COMO UN CASO DE IMPORTANCIA

¿Qué caso sirve para llenar esas cuatro lagunas? El caso mexicano puede ser de ayuda por dos razones principales: es una democracia emergente y existen datos de encuestas relevantes para analizar las predicciones ciudadanas en el país desde los inicios de la democracia hasta la actualidad.

En los años noventa, México vivió la transición de un régimen autoritario de partido único dominante hacia un régimen democrático pluripartidista competitivo. El Partido Revolucionario Institucional (PRI) ganó todas las elecciones presidenciales, para gobernador y las legislativas desde su fundación en 1929 hasta 1989. En ese año, el partido perdió por primera vez una gubernatura; en 1997, perdió la mayoría de escaños en la Cámara de Diputados, y en 2000, la presidencia. Por lo tanto, en términos de elecciones presidenciales, el año 2000 marca la transición hacia la democracia con sus primeras elecciones justas.

Identificamos 55 encuestas de predicciones ciudadanas que abarcan todas las elecciones presidenciales libres y justas de 2000 a 2024, lo que nos permite estudiar cómo predecían los ciudadanos en los inicios de la democracia y cómo sus predicciones han cambiado con el tiempo. Sobre todo, encontramos datos de encuestas sobre predicciones ciudadanas de quién ganará y qué porcentaje de votos recibirá cada candidato presidencial, así como las intenciones de voto de los ciudadanos. Así, podemos utilizar las herramientas estadísticas que desarrollaremos más adelante para comparar si las predicciones ciudadanas de quién ganará y de los porcentajes de voto predicen los resultados de esta elección mejor que una predicción aleatoria y que los sondeos de intención de voto. Las encuestas también incluyen un amplio conjunto de predictores, como variables sociodemográficas y políticas. Además, la herramienta estadística que propondremos a continuación nos permite analizar qué predicciones ciudadanas son mejores o peores que otras.

LA PRESENTE INVESTIGACIÓN

El presente estudio evalúa las predicciones ciudadanas de dos resultados: quién ganará la presidencia de México y qué porcentaje de votos recibirá cada candidato presidencial. En el Estudio 1 llevamos a cabo una prueba para saber si los ciudadanos predicen los resultados mejor que el azar, y luego otra para verificar si los ciudadanos predicen los resultados mejor que las encuestas de intención de voto. En el Estudio 2, llevamos a cabo una prueba para saber si los ciudadanos predicen los

porcentajes de voto para cada candidato mejor que el azar, y otra para saber si los ciudadanos predicen los porcentajes de voto para cada candidato mejor que las encuestas de intención de voto. Además, podemos comentar si la habilidad de los ciudadanos para predecir elecciones está mejorando con el tiempo, ya que contamos con datos de cinco elecciones presidenciales, en el caso de las predicciones de quién ganará, y con cuatro en el caso de las predicciones de los porcentajes de voto. Posteriormente, analizamos qué ciudadanos son mejores para predecir elecciones que otros. En el Estudio 3, analizamos el sesgo y la varianza en las predicciones ciudadanas de los porcentajes de voto para cada candidato. Si la capacidad de predicción de los ciudadanos varía, deberíamos ser capaces de explicar esa capacidad en términos de variables sociodemográficas y políticas.

ESTUDIO 1: PREDECIR QUIÉN GANARÁ LA PRESIDENCIA MEXICANA

El Estudio 1 examina si los ciudadanos predicen mejor que el azar al ganador o ganadora de las elecciones presidenciales en México. Se les preguntó a varias muestras representativas de ciudadanos antes de las elecciones quién creían que ganaría. Con base en sus respuestas, analizamos si predijeron mejor el resultado que una predicción aleatoria. Utilizando las respuestas para las cinco elecciones que sucedieron entre 2000 y 2024, también revisamos si los ciudadanos mejoran con el tiempo. Así, el Estudio 1 analiza si los ciudadanos predicen mejor que los sondeos de intención de voto quién ganará la presidencia de México. También se les preguntó a los encuestados por quién iban a votar y, con base en esas respuestas, analizamos si un sondeo de predicciones ciudadanas es mejor que uno de intenciones de voto para predecir al ganador. Para empezar, desarrollamos dicha prueba estadística y después la aplicamos a los datos. Gracias a que tenemos cinco elecciones, también estudiamos los cambios en el tiempo y cómo los ciudadanos van prediciendo la estabilidad o el cambio en el gobierno.

Pregunta 1: ¿Los ciudadanos predicen mejor que el azar?

La primera pregunta que analizamos es si los ciudadanos predicen mejor que el azar. Es importante porque queremos descartar golpes de suerte.

Datos

Encontramos un total de 55 encuestas en las que se les preguntó a los ciudadanos quién creían que ganaría la presidencia. Investigaciones Sociales Aplicadas² recopila y publica dichas encuestas. La cantidad de encuestados varía entre 900 y 4 100. Las encuestas abarcan las cinco elecciones presidenciales entre 2000 y 2024, cuyo periodo de anticipación oscila entre más de dos años y medio previos a la elección y nada de

² <https://invesoc.com/>

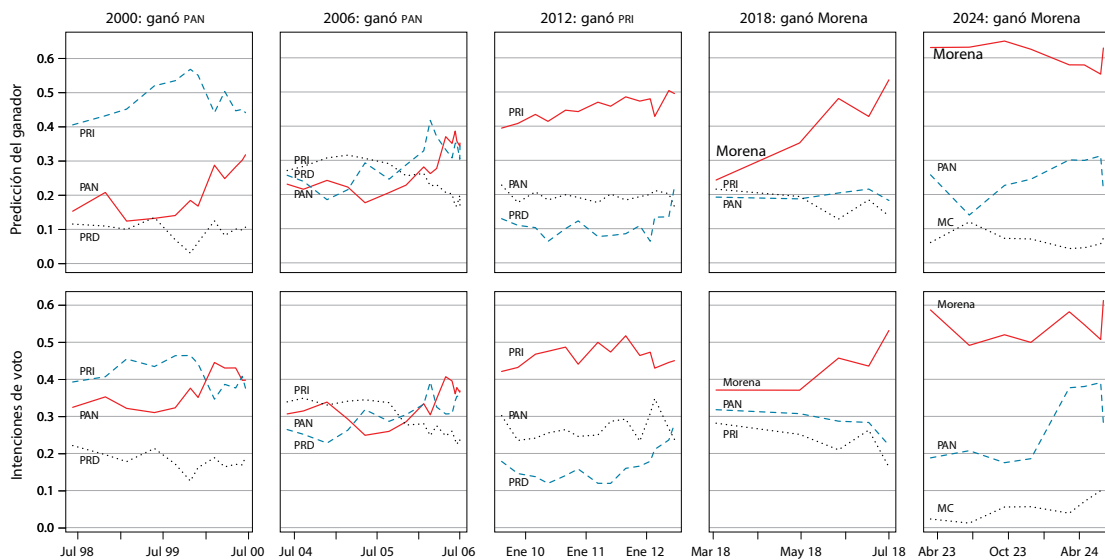
anticipación, pues en los casos de 2006, 2018 y 2024 se aplicaron encuestas de salida el día de las elecciones. Hasta donde sabemos, esta es la primera vez que se analizan las predicciones ciudadanas de encuestas de salida en una investigación de este tipo.

La formulación de las preguntas de la encuesta sobre predicciones ciudadanas varía ligeramente en función del periodo de aplicación. Alrededor de nueve meses antes de las elecciones, las preguntas se centran en los partidos; en adelante, una vez que los candidatos presidenciales son seleccionados, las preguntas se centran en los candidatos.

En general, se utilizaron las siguientes preguntas para medir las predicciones ciudadanas sobre quién va a ganar. La pregunta respecto al partido suele ser: “¿Qué partido cree que va a ganar las próximas elecciones para presidente de la república?”, mientras que la que gira en torno a los candidatos dice: “¿Y cuál de estos candidatos cree usted que ganará la próxima elección para presidente de república?”. Las preguntas de las encuestas de salida son similares a: “¿Y cuál de los candidatos cree usted que ganará esta elección para presidente de la república?”. La formulación de las preguntas es comparable con las aplicadas en otros países. En el Apéndice A presentamos la formulación común de preguntas para cada elección.

Para ilustrar los datos, los recuadros superiores de la gráfica 1 trazan la proporción a lo largo del tiempo de los ciudadanos que predijeron la victoria de un partido. En el año 2000, al inicio de la serie de tiempo, los ciudadanos fallaron de

GRÁFICA 1. Proporción de ciudadanos que predijeron la victoria de un partido y los que dijeron tener la intención de votar por un partido



Fuente: Elaboración propia con base en datos de Investigaciones Sociales Aplicadas (2024).

manera consistente al predecir qué partido ganaría: más de 40 por ciento de los ciudadanos respondió cada mes que el PRI iba a ganar, cuando, en realidad, el Partido Acción Nacional (PAN) resultó ganador. Después, a partir de agosto de 1999, cada vez más personas respondieron que el PAN sería el ganador. Sin embargo, dicho acercamiento a la respuesta correcta no fue suficiente para que la mayoría de los ciudadanos acertara.

La tendencia cambió en 2006: mientras que la mayoría de los ciudadanos predijo mal el ganador —primero el PRI y luego el Partido de la Revolución Democrática (PRD)—, finalmente acertaron en los últimos tres meses y predijeron que el PAN sería el vencedor. A partir de esas elecciones, los ciudadanos han predicho correctamente cada ganador de la presidencia: la victoria del PRI en 2012 y las de Morena en 2018 y 2024. Dichas predicciones son especialmente importantes porque son prueba de que los ciudadanos pueden predecir tanto la estabilidad como el cambio: predijeron correctamente la reelección de los partidos presidenciales en 2006 (PAN) y 2024 (Morena), así como cambios de partido presidencial en 2012 (PRI) y 2018 (Morena).

Asimismo, la gráfica 1 muestra que, a partir de las elecciones de 2012, la mayoría de los ciudadanos predicen correctamente la mayoría de las veces. Si bien ahondaremos sobre este punto más adelante, por ahora podemos señalarlo de manera cualitativa. Ambos resultados sugieren en conjunto que, a partir de las primeras elecciones de 2000 y 2006, las predicciones ciudadanas se han convertido en una herramienta de predicción confiable y válida. En la siguiente sección proponemos un modelo para comprobar estadísticamente si los ciudadanos predicen los resultados mejor, igual o peor que el azar en cada elección.

Un modelo de predicciones ciudadanas

Cada encuestado predice al ganador de la presidencia correcta o incorrectamente. Nos interesa la proporción poblacional desconocida θ de predicciones correctas. Por ejemplo, queremos saber si $\theta > .33$, es decir, si los ciudadanos son mejores que el azar para predecir cuál de los tres candidatos va a ganar.

Nuestros datos son el número de encuestados y que predijeron correctamente de un total de n encuestados. Modelamos estos datos con una distribución binomial

$$y \mid \theta \sim \text{Binomial}(y \mid n, \theta),$$

donde θ es el parámetro de interés.

Usamos la distribución a priori conjugada para θ , la distribución Beta. Establecemos que la distribución a priori sea uniforme entre 0 y 1. Gracias a que trabajamos con priors conjugadas, la inferencia a posteriori de los parámetros es sencilla (Gelman *et al.*, 2014: sección 2.1).

Un estimando de interés es $\theta - .33$, la diferencia poblacional en la posibilidad de acertar entre un ciudadano y una predicción aleatoria con tres partidos. Para calcular las distribuciones a posteriori de este y otros estimandos, extraemos 100000 puntos de la distribución a posteriori de θ y luego calculamos el estimando correspondiente para cada uno (Murr *et al.*, 2023).

CUADRO 1. Resúmenes de la distribución a posteriori de la proporción de predicciones ciudadanas acertadas y su diferencia respecto a una predicción aleatoria

Año	Predicciones correctas	Predicciones aleatorias	Diferencia
2000	22 [21, 23]	33	-11 [-12, -11]
2006	30 [29, 30]	33	-4 [-4, -3]
2012	46 [45, 46]	33	12 [11, 13]
2018	46 [44, 47]	33	12 [11, 13]
2024	62 [61, 63]	33	28 [27, 29]

Fuente: Elaboración propia con base en datos de Investigaciones Sociales Aplicadas (2024). *Nota:* Este cuadro reporta la media a posteriori y los intervalos a posteriori de 95% (entre corchetes).

Resultados

¿Los ciudadanos predicen mejor que el azar? Para cada elección, el cuadro 1 muestra la proporción poblacional θ de predicciones correctas, la proporción de la población que acertaría si todos predijeran aleatoriamente cuál de los tres candidatos ganaría, así como la diferencia entre ambos ($\theta - .33$). El cuadro 1 muestra una tendencia: con el pasar del tiempo, los ciudadanos han mejorado en sus predicciones. Tanto en las elecciones de 2000 como en las de 2006, las predicciones de los ciudadanos fueron peores que las predicciones aleatorias. En el año 2000, por ejemplo, las predicciones ciudadanas fueron 11 por ciento peores que las predicciones aleatorias (más o menos uno). No obstante, desde 2012 los ciudadanos han sido mejores que las predicciones aleatorias, y continúan mejorando. Si bien sólo 22 por ciento de los ciudadanos predijo correctamente en 2000 (más o menos uno), con el tiempo, dicha proporción ha incrementado monótonamente hasta 62 por ciento en 2024 (más o menos uno). Aunque algunos factores, como el margen de victoria, varían durante las elecciones, los resultados sugieren que después de un periodo inicial difícil, los ciudadanos saben cómo predecir al ganador de manera colectiva. Una vez demostrado que desde 2012 los ciudadanos predicen al ganador mejor

que el azar, a continuación, vamos a comparar dicho resultado con las encuestas de intención de voto.

Pregunta 2: ¿Las predicciones ciudadanas son mejores que las intenciones de voto?

La segunda pregunta que analizamos es si las encuestas de predicción ciudadana predicen mejor, igual o peor que las de intención de voto. Esto es importante porque, hasta ahora, los sondeos de intención de voto son la única herramienta de predicción que se ha empleado en las elecciones presidenciales mexicanas. Además, las encuestas de intención de voto son la comparación natural de las encuestas de predicciones ciudadanas, ya que sólo difieren en la pregunta planteada.

Datos

En las mismas 55 encuestas en las que se preguntó a los ciudadanos quién piensan que ganará las elecciones presidenciales, también se les preguntó por quién tenían intención de votar. Al igual que en el caso de las predicciones ciudadanas, la formulación de las preguntas de la encuesta de intención de voto varía ligeramente en función del periodo de aplicación. Alrededor de nueve meses antes de las elecciones, las preguntas se centran en los partidos; en adelante, una vez que los candidatos presidenciales son seleccionados, las preguntas se centran en los candidatos.

En general, las siguientes preguntas se utilizaron para medir las intenciones de voto. La pregunta respecto al partido suele ser: “Si tuviera usted que votar en este momento para elegir presidente de la república, ¿por cuál partido político votaría?”; mientras que la que gira en torno a los candidatos dice: “Si en este momento se celebraran las elecciones para presidente de la república, ¿por cuál candidatura votaría usted?”. Las preguntas de las encuestas de salida son similares a: “¿Por quién votó el día de hoy para presidente de la república?”. Una vez más, la formulación de las preguntas es comparable con las aplicadas en otros países. En el Apéndice A presentamos la formulación común de preguntas para cada elección.

Para ilustrar los datos, los recuadros inferiores de la gráfica 1 presentan el porcentaje de intención de voto para cada partido antes de cada una de las cinco elecciones presidenciales. Una comparación visual con los recuadros superiores (predicciones ciudadanas) demuestra que las intenciones de voto fueron más acertadas que las predicciones ciudadanas en 2000 y 2006. Aunque los ciudadanos predijeron una victoria para el PRI en el año 2000, cuando de hecho venció el PAN, sus intenciones de voto ya habían indicado seis meses antes de las elecciones que preferían al PAN. De manera similar, en 2006, las encuestas de intención de voto señalaron antes que las encuestas de predicción ciudadana que el PAN volvería a ser elegido. A partir de 2012, ambos tipos de encuestas predijeron correctamente

al ganador en cada ocasión. Para comprobar si las pruebas son más contundentes de lo que esperaríamos del azar, más adelante proponemos un modelo estadístico para comparar ambos tipos de encuestas.

Un modelo de comparación entre predicciones ciudadanas e intenciones de voto

Cada encuesta de predicción ciudadana o de intención de voto predice correcta o incorrectamente. Por un lado, una encuesta de predicción ciudadana acierta si la mayoría de los encuestados predice que va a ganar quien de hecho ganará; por otro lado, una encuesta de intención de voto acierta si la mayoría de los encuestados pretendía votar por el partido ganador. Nos interesa la proporción poblacional desconocida de las encuestas que predicen correctamente con predicciones ciudadanas θ_1 e intenciones de voto θ_2 . También nos interesa saber si la probabilidad de predecir correctamente con predicciones ciudadanas es mayor, igual o menor que con intenciones de voto. Esto es, estamos interesados en la diferencia $\theta_1 - \theta_2$.

Nuestros datos son el número de encuestas que predijeron correctamente y_1 (con predicciones ciudadanas) y y_2 (con intenciones de voto) de un total de n encuestas. Nuevamente modelamos estos datos con una distribución binomial y priors conjuguadas para θ_1 y θ_2 uniformes entre 0 y 1.

Un estimando de interés es $\theta_1 - \theta_2$, la diferencia poblacional en la probabilidad de acertar entre una encuesta de predicciones ciudadanas y una de intención de voto. Para calcular las distribuciones a posteriori de este estimando, muestreamos 100000 puntos de las distribuciones a posteriori de θ_1 y θ_2 , y luego calculamos la diferencia de cada uno.

Resultados

¿Las encuestas de predicciones ciudadanas son mejores que las de intención de voto? Para cada elección, el cuadro 2 muestra la proporción poblacional de encuestas que acertaron con predicciones ciudadanas θ_1 , con intenciones de voto θ_2 , y su diferencia $\theta_1 - \theta_2$. Asimismo, el cuadro 2 demuestra que, por lo menos en el año 2000, las encuestas de intención de voto predijeron mejor que las de predicciones ciudadanas. Si bien 7 por ciento de las encuestas de predicciones ciudadanas acertó; en el caso de las encuestas de intención de voto, la proporción asciende a 36 por ciento, una diferencia de 28 puntos porcentuales (con un intervalo de 95 por ciento que oscila entre 2 y 57 por ciento).

Dicho esto, tanto en 2006 como en 2024, ambos tipos de encuestas acertaron en igual proporción. En el resto de las elecciones, la diferencia a favor de las encuestas de intención de voto no es tan significativa como en el año 2000. En general, es más probable que las encuestas de intención de voto sean mejores que las de predicción ciudadana. No obstante, si nos basamos en las elecciones más recientes, es probable que ambos tipos de encuestas tengan la misma calidad.

CUADRO 2. Resúmenes de la distribución a posteriori de la proporción de encuestas acertadas de predicciones ciudadanas, de encuestas acertadas de intención de voto y su diferencia

Año	Predicciones ciudadanas	Intención de voto	Diferencia
2000	7 [0, 25]	36 [14, 61]	-28 [-57, -2]
2006	39 [18, 62]	56 [33, 77]	-17 [-47, 15]
2012	94 [78, 100]	94 [78, 100]	0 [-18, 18]
2018	71 [36, 96]	86 [54, 100]	-14 [-55, 26]
2024	90 [66, 100]	90 [66, 100]	0 [-27, 27]
En general	58 [45, 70]	72 [60, 83]	-14 [-31, 3]

Fuente: Elaboración propia con base en datos de Investigaciones Sociales Aplicadas (2024). *Nota:* Este cuadro muestra la media a posteriori y los intervalos a posteriori del 95% (entre corchetes).

ESTUDIO 2: PREDECIR LOS PORCENTAJES DE VOTO DE LOS CANDIDATOS PRESIDENCIALES MEXICANOS

El Estudio 1 indica que, con el tiempo, los ciudadanos mejoraron al predecir quién ganaría: aunque al principio predecían peor que las predicciones aleatorias, ahora están a la par con las encuestas de intención de voto. ¿Qué pasa con las predicciones de los porcentajes de voto? El Estudio 2 busca evaluar las predicciones ciudadanas sobre esta tarea aún más complicada. Se les preguntó a varias muestras representativas de ciudadanos antes de la elección qué porcentaje de voto creían que obtendría cada candidato presidencial. Con base en sus respuestas, el Estudio 2 analiza, en primer lugar, si predijeron los resultados mejor que una predicción aleatoria. Dado que entre 2000 y 2018 tenemos cuatro elecciones presidenciales, también estudiamos si las predicciones de los ciudadanos mejoran con el tiempo. En segundo lugar, el Estudio 2 analiza si los ciudadanos predicen mejor los porcentajes de voto que los sondeos de intención de voto. Para ambas pruebas, desarrollamos un procedimiento estadístico y luego lo aplicamos a los datos.

Pregunta 1: ¿Los ciudadanos predicen mejor que el azar?

La primera pregunta que analizamos es si los ciudadanos predicen mejor que el azar. Es importante porque queremos descartar golpes de suerte.

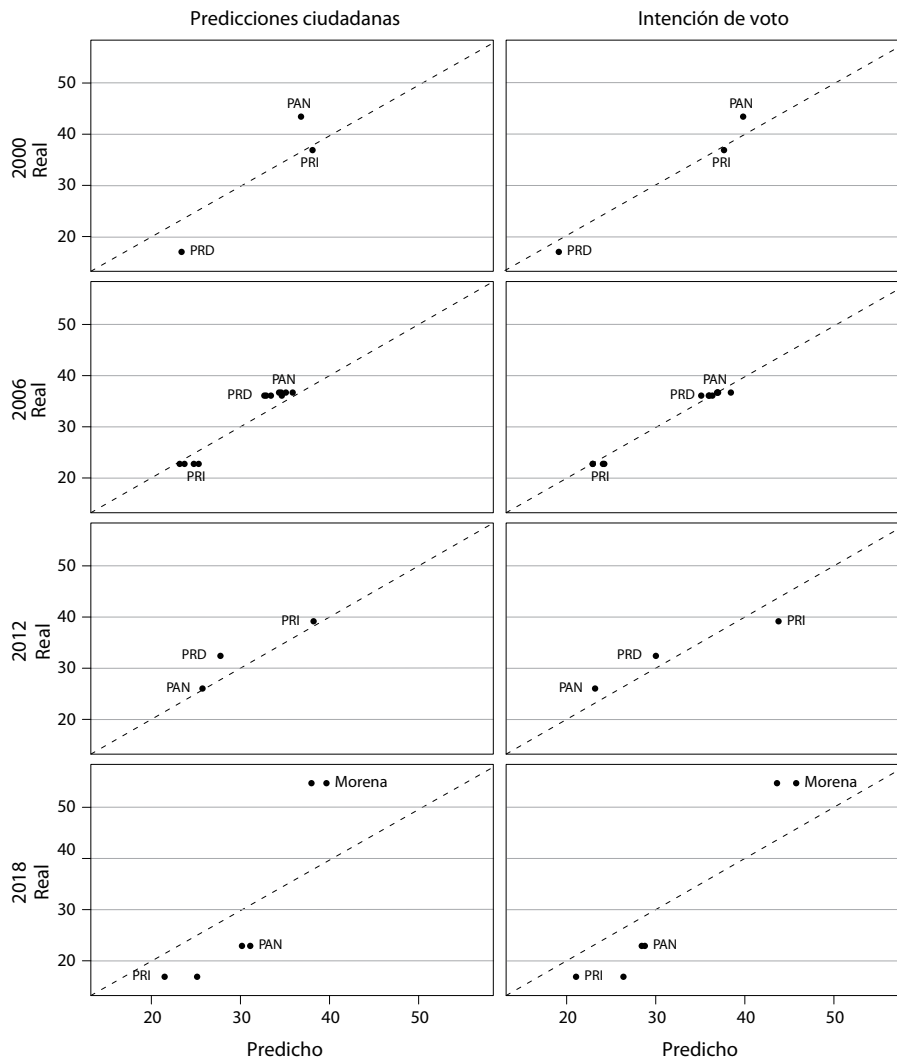
Datos

Encontramos un total de siete encuestas de Investigaciones Sociales Aplicadas en donde se les pidió a los ciudadanos que predijeran los porcentajes de voto que

obtendría cada candidato presidencial. La cantidad de encuestados varía entre 800 y 2300. Las encuestas abarcan cuatro elecciones presidenciales entre 2000 y 2018. Todas las encuestas se realizaron durante el mes previo a la elección, incluida una encuesta de salida. La pregunta de la encuesta por lo general se formuló así: “¿Qué porcentaje de votos cree que obtendrá [nombre del candidato]?”. Una vez más, esta formulación se asemeja a la que se usa en otros países. El Apéndice A presenta la formulación habitual de la pregunta para cada elección.

El recuadro izquierdo de la gráfica 2 muestra los porcentajes de voto reales y las

GRÁFICA 2. Porcentajes de votos reales y previstos de los partidos según el tipo de encuesta



predicciones para cada partido en las cuatro elecciones entre 2000 y 2018 a partir del promedio de las predicciones ciudadanas. En el año 2000, al inicio de la serie de tiempo, los ciudadanos predijeron que ganaría el partido equivocado: dijeron que el porcentaje de voto para el PRI sería ligeramente superior al del PAN. Desde entonces, todas las encuestas de predicción ciudadana han predicho correctamente quién ganará. El porcentaje de voto esperado del ganador final siempre es el mayor porcentaje de voto esperado. De hecho, la gráfica 2 muestra una fuerte correlación positiva entre las predicciones y los resultados reales. Entre mayor haya sido el porcentaje de voto predicho, mayor será el porcentaje real. Las predicciones de porcentajes de voto se acercan mucho a los resultados reales: los puntos se acercan mucho a la línea punteada, la cual indica una predicción perfecta. A continuación, proponemos un modelo estadístico para probar si los ciudadanos predicen los resultados mejor, igual o peor que el azar en cada elección.

Un modelo de predicción ciudadana

Cada encuestado predice el porcentaje de voto de los tres principales partidos de una elección. Las predicciones se agrupan en un vector y con tres componentes. Aquí modelamos esos datos con una distribución normal multivariada,

$$y \mid \mu, \Sigma \sim N(\mu, \Sigma)$$

donde μ es un vector columna de dimensión 3, y Σ es una matriz de covarianza de tres por tres. En cada elección, trabajamos con una muestra de n ciudadanos distribuidos de manera independiente e idéntica, y_1, \dots, y_n .

Usamos la distribución a priori conjugada para (μ, Σ) , la distribución normal inversa Wishart. La media a priori es 0 para cada partido, y la distribución inversa Wishart está parametrizada para que cada correlación tenga, marginalmente, una distribución uniforme. Como trabajamos con priors conjugadas, la inferencia a posteriori de los parámetros es sencilla (Gelman *et al.*, 2014: sección 3.6).

Un estimando de interés es la raíz del error cuadrático medio (RMSE) de las encuestas con predicciones ciudadanas de los porcentajes de voto. Si v_j denota el porcentaje de votos real de cada partido $j \in \{1, 2, 3\}$, el estimando de interés es, por lo tanto, $\text{RMSE}_{pre} = \sqrt{\sum_{j=1}^3 (v_j - \mu_j)^2}$. Para calcular las distribuciones posteriores de dicho estimando, muestreamos 100 000 puntos de la distribución posterior de μ y luego calculamos la RMSE_{pre} para cada uno.

Otro estimando de interés es $\frac{\text{RMSE}_{pre}}{\sqrt{\sum_{j=1}^3 (v_j - .33)^2}}$, la razón poblacional de la raíz del error cuadrático medio de una encuesta de predicciones ciudadanas y una predicción aleatoria. Si la razón es mayor a uno, entonces las predicciones ciudadanas aumentan la RMSE respecto a una predicción aleatoria. Si la razón es menor a uno, las predicciones ciudadanas reducen la RMSE respecto a una predicción aleatoria. Para calcular las

$$= \sqrt{\sum_{j=1}^3 (v_j - \mu_j)^2}$$

distribuciones posteriores de este estimando, usamos los 100000 puntos de la distribución posterior de $RMSE_{pre}$ y calculamos la razón para cada una.

Resultados

¿Los ciudadanos predicen mejor los porcentajes de votos que el azar? El cuadro 3 muestra, para cada elección, la raíz del error cuadrático medio (RMSE) de las predicciones ciudadanas, una predicción aleatoria y la razón entre ambas. El cuadro 3 muestra, por lo tanto, que los ciudadanos predicen mejor los porcentajes de votos que el azar en todas las elecciones. En general, los ciudadanos predicen con alrededor de la mitad del error de lo que se espera para el azar. Si analizamos la elección de 2006 y cada ciudadano hubiera predicho al azar el porcentaje de voto para cada candidato presidencial, su porcentaje de voto promedio predicho habría tenido una RMSE de 5.4 puntos porcentuales. Pero, en realidad, los ciudadanos tuvieron una RMSE de 2.3 puntos, con un intervalo posterior de 95 por ciento entre 2.0 y 2.6. Por lo tanto, la RMSE de los ciudadanos es alrededor de 40 por ciento del de una predicción aleatoria (el intervalo posterior de 95 por ciento oscila aproximadamente entre 40 y 50 por ciento). En general, en las cuatro elecciones, el cuadro 3 demuestra que una predicción aleatoria habría tenido una RMSE de 10.0 puntos, mientras que los ciudadanos lograron una de aproximadamente 5.3 puntos, más o menos 0.3, una reducción de alrededor de 50 por ciento. Por lo tanto, por primera vez, tenemos pruebas sustanciales de que los ciudadanos predicen los porcentajes de voto mejor que el azar.

CUADRO 3. Resúmenes de la distribución posterior de la raíz del error cuadrático medio de las predicciones de porcentajes de voto por predicción ciudadana, predicción aleatoria e intención de voto, y sus razones

Año	(1) Predicciones ciudadanas	(2) Predicciones aleatorias	Razón entre (1) y (2)	(3) Intención de voto	Razón entre (1) y (3)
2000	5.4 [4.9, 5.9]	11.3	0.5 [0.4, 0.5]	2.6 [1.2, 4.0]	2.3 [1.3, 4.4]
2006	2.3 [2.0, 2.6]	5.4	0.4 [0.4, 0.5]	1.8 [1.3, 2.2]	1.3 [1.0, 1.8]
2012	2.8 [2.4, 3.2]	6.6	0.4 [0.4, 0.5]	3.4 [1.3, 5.6]	0.9 [0.5, 2.1]
2018	10.9 [10.5, 11.3]	16.7	0.7 [0.6, 0.7]	7.4 [5.8, 9.0]	1.5 [1.2, 1.9]
En general	5.3 [5.1, 5.6]	10.0	0.5 [0.5, 0.5]	3.8 [3.0, 4.6]	1.5 [1.2, 2.2]

Fuente: Elaboración propia con base en datos de Investigaciones Sociales Aplicadas (2024). *Nota:* Este cuadro muestra las medias posteriores y los intervalos posteriores de 95% (entre corchetes).

Pregunta 2: ¿Los ciudadanos predicen mejor que los sondeos de intención de voto?

La segunda pregunta que nos hacemos es si las encuestas de predicción ciudadana predicen mejor, igual o peor que los sondeos de intención de voto. Es importante porque las encuestas de intención de voto son la única herramienta predictiva usada hasta ahora en las elecciones presidenciales mexicanas. Además, los sondeos de intención de voto son la comparación natural de las encuestas de predicción ciudadana, pues sólo difieren en la pregunta que se plantea.

Datos

Las mismas siete encuestas que les pidieron a los ciudadanos que predijeran los porcentajes de voto para cada candidato presidencial les preguntaron también su intención de voto. Ya se discutió en secciones previas la formulación de las preguntas de intención de voto.

El recuadro derecho de la gráfica 2 despliega los porcentajes de voto reales contra los predichos con base en las intenciones de voto. Al comparar ambos recuadros podemos ver que las intenciones de voto predicen mejor los resultados que las predicciones ciudadanas. En primer lugar, las intenciones de voto predicen bien el ganador de todas las elecciones, incluso en la elección del año 2000, en la que las expectativas de porcentajes de votos se equivocaron. Además, las intenciones de voto suelen acercarse más a la línea punteada que indica una predicción perfecta que las expectativas de porcentajes de votos. Pero, ¿acaso estos resultados se deben al azar y, si no es así, qué tanto mejor predijeron los resultados las intenciones de voto en comparación con las expectativas de porcentajes? Para responder esta pregunta, proponemos un modelo de intenciones de voto y, luego, una prueba estadística para saber si los ciudadanos son mejores, iguales o peores para predecir en cada elección que las intenciones de voto.

Un modelo de intención de voto

Cada encuestado puede decidir votar entre, por lo general, al menos tres candidatos presidenciales. Usamos la distribución multinomial para describir cada una de las observaciones de uno de los resultados k ,

$$x \mid n, \theta \sim \text{Multimodal}(n, \theta),$$

donde x es el vector del conteo de intención de voto para cada partido, y la suma de probabilidades $\sum_{j=1}^k \theta_j = 1$.

Usamos la distribución a priori conjugada para θ , la distribución Dirichlet (una generalización multivariada de la distribución Beta). Establecimos que la distribución a priori para θ sería uniforme. Como trabajamos con una prior conjugada, la inferencia posterior para los parámetros es sencilla (Gelman *et al.*, 2014: sección 3.4).

Un estimando de interés es $RMSE_{int} = \sqrt{\sum_{j=1}^3 (v_j - \theta_j)^2}$, la raíz del error cuadrático medio de la intención de voto para los tres principales partidos. Para calcular las distribuciones posteriores de este estimando, muestreamos 100 000 de la distribución posterior de θ y luego calculamos el estimando correspondiente para cada una.

Un estimando para comparar predicciones ciudadanas e intención de voto

Otro estimando de interés es $\frac{RMSE_{pre}}{RMSE_{int}}$, la razón poblacional de la raíz del error cuadrático medio de las predicciones ciudadanas y las intenciones de voto. Para calcular las distribuciones posteriores de este estimando, usamos los 100 000 puntos de la distribución posterior de $RMSE_{pre}$ y $RMSE_{int}$ calculamos la razón para cada uno.

Resultados

¿Las encuestas de predicción ciudadana son mejores que las de intención de voto? El cuadro 3 muestra para cada elección la raíz del error cuadrático medio (RMSE) de las predicciones ciudadanas, intención de voto y la razón entre ambas. También muestra que las predicciones ciudadanas predicen peor que las intenciones de voto en tres de cuatro elecciones. Por ejemplo, en las elecciones de 2018, las predicciones ciudadanas tuvieron una RMSE de 10.9 (más o menos 0.4), mientras que las intenciones de voto tuvieron una de 7.4 (más o menos 1.6). Como resultado, la RMSE de las predicciones ciudadanas fue aproximadamente 50 por ciento mayor que la de las intenciones de voto. La única excepción es la elección de 2012, con predicciones ciudadanas de alrededor de 90 por ciento del error de las intenciones de voto; sin embargo, el intervalo posterior de 95 por ciento oscila entre 50 y 210 por ciento. En general, para las cuatro elecciones, el cuadro 3 muestra que las intenciones de voto lograron una RMSE de alrededor de 3.8 puntos (más o menos 0.8), y las predicciones ciudadanas tuvieron una RMSE de 5.3 puntos (más o menos 0.3): un aumento de alrededor de 50 por ciento. El intervalo posterior de 95 por ciento para este aumento oscila entre 20 y 120 por ciento. Por lo tanto, en términos de RMSE, una encuesta de predicciones ciudadanas está aproximadamente a medio camino entre una encuesta de predicciones aleatorias y una de intención de voto.

ESTUDIO 3: PREDECIR A QUIEN PREDICE

La sección anterior sugiere que los ciudadanos en su conjunto predicen mejor los porcentajes de voto que el azar, aunque son peores que las encuestas de intención de voto. Sin embargo, este hallazgo agregado oculta diferencias individuales importantes. ¿Algunos ciudadanos predicen mejor que otros?

Gracias a investigaciones previas, sabemos, por ejemplo, que los ciudadanos con un mayor nivel educativo predicen mejor quién ganará, en comparación con personas con un nivel educativo inferior, manteniendo todo lo demás constante. Sin embargo, la medida del resultado de quién ganará es discreta (correcta/incorrecta) y por

lo tanto omite información. Por consiguiente, nos centramos en las predicciones ciudadanas continuas de los porcentajes de voto, puesto que resultan más informativas que las predicciones discretas de quién ganará. Hasta donde sabemos, esta es la primera vez que se analizan los errores de predicción continuos de los ciudadanos en una investigación de este tipo; en investigaciones anteriores, sólo se ha examinado si las predicciones aciertan o no. No obstante, consideramos que este estudio constituye un punto de partida para desarrollar un modelo de la capacidad de los ciudadanos para predecir los porcentajes de voto.

Predecir la capacidad ciudadana de predicción sobre quién ganará en las democracias bien establecidas

Al recurrir primero a las variables sociodemográficas, los estudios generalmente encuentran que, en las democracias bien establecidas, la edad y el nivel educativo determinan quién será mejor para predecir (Leiter *et al.*, 2018; Lewis-Beck y Tien, 1999; Murr, 2015). Nosotros sostenemos que esto es así porque, en las democracias bien establecidas, las personas de edad avanzada tienen más experiencia que la gente joven con la política en general, en especial con las elecciones. En relación con esto, el argumento de la educación dice que un mayor nivel educativo significa que la gente adquirió mejores habilidades para procesar la información y posee información importante para la predicción electoral. Por consiguiente, las personas de edad avanzada y con mayor nivel educativo deberían predecir mejor que la gente joven y con nivel educativo inferior.

En cuanto a las variables políticas, suele concluirse que las preferencias de los ciudadanos (identificación partidista e intención de voto) tienen una correlación positiva con sus predicciones (Granberg y Brent, 1983; Meffert *et al.*, 2011). Es decir, las personas que se identifican con un partido son más propensas a decir que ese partido va a ganar; así como las personas que pretenden votar por un partido también predicen que ese será el ganador. Aun así, en términos de causalidad, la conexión es teóricamente más clara en el caso de la identificación partidista: no esperaríamos que la identificación partidista se viera afectada por las expectativas sobre quién va a ganar las elecciones. Por el contrario, las intenciones de voto pueden estar influenciadas por consideraciones estratégicas en forma de voto útil (Fisher, 2004). Como consecuencia, las intenciones de voto —sin mayor análisis— no pueden ser percibidas como una causa de la precisión de las predicciones. A pesar de que las intenciones de voto pueden ser de utilidad para determinar predicciones, lo que nos interesa aquí son las explicaciones sobre la precisión, por lo que a continuación nos centramos en la identificación partidista.

En términos de características situacionales, hay dos hallazgos bien establecidos que atañen al margen de victoria de unas elecciones (Murr, 2011) y el periodo de anticipación de la predicción (Lewis-Beck y Tien, 1999). Como era de esperarse, es

más complicado predecir unas elecciones reñidas que aquellas con un gran margen de victoria; así como las elecciones más lejanas en el tiempo son más difíciles de predecir que las más cercanas. Los estudios que tomaron en cuenta el margen de victoria se han centrado en el resultado binario entre si un ciudadano acertó o no. Debido a que nos dedicaremos a los errores de predicción continuos de los ciudadanos en la siguiente sección, no vamos a seguir tomando en cuenta el margen de victoria, pero sí el periodo de anticipación. Asimismo, para reducir la heterogeneidad entre las encuestas, vamos a analizar encuestas con periodos de anticipación similares.

Predecir la capacidad de predicción de los ciudadanos para porcentajes de voto en las democracias emergentes

Mucho de lo anterior sugiere que las causas de predicción ciudadana deberían aplicar también para las democracias emergentes. Por ejemplo, seguimos esperando que la gente con identificación partidista sea más propensa a cometer un error —o errores más importantes— que la gente sin identificación partidista; y que los ciudadanos con mayor nivel educativo predigan mejor que aquellos con menor nivel educativo. No obstante, esperaríamos que el efecto de la edad en las democracias emergentes fuera contrario a su efecto en las democracias bien establecidas, así como también esperaríamos que la percepción de elecciones limpias influiría en la precisión de las predicciones ciudadanas.

El argumento mencionado anteriormente sobre la edad y la precisión de las predicciones se basa en que los ciudadanos crecieron en democracias. Nuestra hipótesis es que, en las democracias emergentes, la relación entre la edad y la precisión de las predicciones es la opuesta; como los ciudadanos de edad avanzada han pasado la mayor parte de su vida, si no es que toda, bajo un régimen autocrático, su experiencia en materia de política y elecciones proviene en su mayoría, si no exclusivamente, de ese tipo de régimen. Dicha experiencia debería ser una guía deficiente para la predicción de elecciones bajo un régimen democrático nuevo. Como consecuencia, la experiencia en política y elecciones de los ciudadanos de edad avanzada en las democracias emergentes es inútil para la predicción de elecciones. En cambio, los ciudadanos más jóvenes acuden a estas nuevas elecciones con una mentalidad más abierta y menos prejuicios. Por lo tanto, los ciudadanos más jóvenes en democracias emergentes deberían predecir mejor las elecciones que los de edad avanzada.

En relación con esto, un revisor anónimo sostuvo que, debido a que la transición democrática de México comenzó antes en unos estados que en otros, algunos ciudadanos tienen más experiencia que otros con las elecciones democráticas, así que los ciudadanos con más experiencia deberían predecir mejor que los menos experimentados. Dado que en las encuestas se indica el estado de residencia de los encuestados, podemos incluir un predictor que señale la cantidad de años que cada estado estuvo sin un gobernador del PRI, y realizar esta prueba suplementaria.

Además, tenemos la hipótesis de que la percepción de elecciones limpias debería influir en la precisión de las predicciones. Entonces, los ciudadanos que creen que las elecciones son limpias deberían predecirlas mejor que aquellos que sospechan lo contrario. Esto porque los expertos en materia electoral han determinado que las elecciones presidenciales mexicanas han sido limpias a partir del año 2000. De hecho, la institución encargada de garantizar elecciones libres, justas y democráticas, el Instituto Nacional Electoral (INE), ya estaba establecida en 1988 (con el nombre de Instituto Federal Electoral [IFE]), aunque se volvió completamente independiente después de la reforma electoral de 1996 (Garrido de Sierra, 2019), y en 2014 cambió su nombre a INE. El INE sigue siendo una de las instituciones con mayor prestigio tanto a nivel nacional como internacional. De este modo, los ciudadanos que sospechan de elecciones fraudulentas, cuando de hecho son limpias, deberían predecir peor que aquellos que perciben elecciones limpias. En la siguiente sección describimos las encuestas, las variables predictoras y de resultados y el modelo estadístico para poner a prueba esta y otras hipótesis.

Encuestas

Como señalamos antes, identificamos siete encuestas entre 2000 y 2018 que pedían a los ciudadanos predecir los porcentajes de voto de los candidatos presidenciales. Para lograr que los análisis de las predicciones fueran más comparables en todas las elecciones, seleccionamos una encuesta con el mismo periodo de anticipación por cada elección. De modo que el último día de cada encuesta fuera exactamente catorce días antes de la fecha de las elecciones. La cantidad de encuestados oscila entre 900 y 2 200.

Por desgracia, ninguna encuesta recopiló las predicciones de los porcentajes de voto para las elecciones de 2024, lo que nos deja con cuatro elecciones que analizar. Dichas cuatro elecciones abarcan distintos escenarios, tales como la reelección de candidatos (como en 2006), cambios en los gobiernos (como en 2012 y 2018), elecciones reñidas (como en 2006) y victorias aplastantes (como en 2018). En consecuencia, estas cuatro elecciones cubren características significativas que esperamos influyan en las predicciones ciudadanas. Es más sencillo predecir estabilidad que cambios, y es más complicado predecir elecciones con márgenes de victoria estrechos que aquellas con un amplio margen (Murr, 2011; 2015; 2016).

Predictores

En sintonía con las investigaciones previas y con nuestra propia teorización presentada, nuestros predictores incluyen variables sociodemográficas y variables políticas. Incluimos variables relacionadas con los siguientes aspectos de los encuestados:

- género (“mujer”: 0 = masculino; 1 = femenino),
- edad (“edad”: en años),
- educación (“educación”: 1 = sin educación formal / primaria; 2 = secundaria o vocacional o equivalente; 3 = bachillerato o equivalente; 4 = licenciatura o más),
- identificación partidista (“IP Morena”, “IP PAN”, “IP PRD”, “IP PRI”: 0 = no; 1 = sí).
- interés en política (“interés”: 0 = nada; 1 = poco; 2 = mucho),
- experiencia con gobernadores no priístas (“experiencia”: en años, con base en el estado en donde viva el encuestado), y
- percepción de qué tan limpia será la elección (“limpia”: suma de si las elecciones se consideran limpias o fraudulentas y si se cree que el IFE/INE garantizará elecciones imparciales, donde 0 = fraudulentas / no; 0.5 = no sabe; 1 = limpias / sí).

El Apéndice B ofrece la formulación típica de las preguntas para cada variable.

El cuadro 4 despliega estadísticas descriptivas de las variables sociodemográficas y políticas clave de cada encuesta. Las estadísticas descriptivas incluyen la media, la desviación estándar, el mínimo y el máximo. En general, la muestra parece representativa de la población. Por ejemplo, alrededor de la mitad de los encuestados son mujeres. Mientras que la mayoría de las estadísticas descriptivas son estables en el tiempo, a través de olas y elecciones, también observamos cambios importantes que pueden explicar la variación en la precisión de las predicciones ciudadanas. Por ejemplo, en 2018, tanto el interés en política como la percepción de que habría elecciones limpias estaban en el nivel más bajo. Dependiendo de la importancia relativa de esas variables para la precisión de las predicciones ciudadanas, esto puede explicar por qué las predicciones ciudadanas de 2018 tuvieron un error cuadrático medio tan grande en relación con las elecciones anteriores.

El cuadro 5 presenta los elementos únicos de la matriz de correlación de los predictores, junto con los errores estándar correspondientes (Lee, 2012: sección 6.1). No se encuentran fuertes correlaciones en la matriz de correlación: la correlación de -0.38 entre educación y edad es la mayor. Muchas correlaciones son indistinguibles de 0. Resulta interesante que la percepción de elecciones limpias tiene las correlaciones más débiles con los demás predictores. Por ejemplo, la correlación entre la percepción de elecciones limpias y la edad es de 0.03, entre experiencia y edad es de -0.06, y entre percepción de elecciones limpias y experiencia es de -0.00. Estas correlaciones nulas apoyan la afirmación de que esas tres variables miden dimensiones distintas. No existen pruebas de que dichas variables estén correlacionadas y, por lo tanto, que sean consecuencia una de la otra.

CUADRO 4. Las estadísticas descriptivas de la muestra son similares a la población en términos demográficos

Variable	Año	Media	Desv. est.	Mín.	Máx.	Variable	Año	Media	Desv. est.	Mín.	Máx.
Mujer	2000	0.49	0.50	0	1	IP PAN	2000	0.29	0.45	0	1
	2006	0.53	0.50	0	1		2006	0.34	0.47	0	1
	2012	0.50	0.50	0	1		2012	0.19	0.39	0	1
	2018	0.52	0.50	0	1		2018	0.19	0.39	0	1
Edad	2000	35.29	12.58	17	89	IP PRI	2000	0.31	0.46	0	1
	2006	39.20	14.22	17	86		2006	0.21	0.41	0	1
	2012	40.19	13.93	18	81		2012	0.35	0.48	0	1
	2018	39.12	13.20	17	75		2018	0.17	0.38	0	1
Educación	2000	2.29	1.08	1	4	Experiencia	2000	0.27	0.45	0	1.78
	2006	2.24	1.04	1	4		2006	0.72	0.81	0	2.78
	2012	2.26	0.91	1	4		2012	1.19	1.21	0	3.78
	2018	2.42	0.78	1	4		2018	1.55	1.39	0	4.78
Interés	2000	1.11	0.61	0	2	Limpias	2000	1.19	0.74	0	2
	2006	1.18	0.62	0	2		2006	1.53	0.63	0	2
	2012	1.13	0.58	0	2		2012	1.27	0.78	0	2
	2018	0.97	0.66	0	2		2018	1.06	0.71	0	2
IP PRD	2000	0.15	0.35	0	1						
	2006	0.31	0.46	0	1						
	2012	0.21	0.40	0	1						
IP Morena	2018	0.30	0.46	0	1						

Fuente: Elaboración propia con base en datos de Investigaciones Sociales Aplicadas (2024).

CUADRO 5. Triángulo inferior de la matriz de correlación

	Mujer	Edad	Educación	Interés	IP				Experiencia
					Morena	PAN	PRD	PRI	
Edad	-0.04 (0.01)								
Educación	-0.08 (0.01)	-0.38 (0.01)							
Interés	-0.04 (0.01)	0.01 (0.01)	0.16 (0.01)						
IP Morena	0.01 (0.03)	-0.14 (0.03)	0.04 (0.03)	0.07 (0.03)					
IP PAN	0.02 (0.01)	-0.07 (0.01)	0.06 (0.01)	0.05 (0.01)	-0.32 (0.03)				
IP PRD	-0.02 (0.01)	0.01 (0.01)	-0.04 (0.01)	0.05 (0.01)	-	-0.33 (0.01)			
IP PRI	0.02 (0.01)	0.08 (0.01)	-0.06 (0.01)	0.08 (0.01)	-0.30 (0.03)	-0.37 (0.01)	-0.33 (0.01)		
Experiencia	0.00 (0.01)	0.06 (0.01)	0.02 (0.01)	-0.06 (0.01)	0.12 (0.03)	-0.01 (0.01)	0.03 (0.01)	-0.08 (0.01)	
Limpias	0.02 (0.01)	0.03 (0.01)	-0.00 (0.01)	0.22 (0.01)	-0.08 (0.03)	0.10 (0.01)	-0.06 (0.01)	0.21 (0.01)	-0.00 (0.01)

Fuente: Elaboración propia con base en datos de Investigaciones Sociales Aplicadas (2004). *Nota:* Errores estándar entre paréntesis.

Resultado

Planteamos una nueva variable de resultado que no se había usado nunca en ningún estudio sobre la capacidad de los ciudadanos de predecir. Trabajamos con las predicciones ciudadanas continuas debido a que son más informativas que las predicciones discretas de quién gana. Hasta donde sabemos, esta es la primera vez que se estudian los errores de predicción continuos de los ciudadanos. En investigaciones anteriores sólo se había estudiado si los ciudadanos acertaban o no en sus predicciones.

Nuestro tratamiento de los errores de predicción es similar a la forma en la que se evalúan los modelos de predicción de porcentajes de voto: utilizamos el error cuadrático medio como punto de partida. En lugar de simplemente emplear el error cuadrático medio, utilizamos la descomposición del error cuadrático medio en la suma del sesgo cuadrático y la varianza, que resulta más informativa (Wasserman, 2004: 91).

El error cuadrático medio (MSE) puede escribirse de la siguiente manera:

$$\text{MSE} = \text{sesgo}^2 + \text{varianza}.$$

Al estimar tanto el sesgo como la varianza de quien predice, es posible entender de dónde proviene el MSE. El sesgo es la diferencia esperada entre el resultado real y el predicho. La varianza es la varianza de la predicción. Puesto que queremos examinar cómo difieren los ciudadanos tanto en sesgo como en varianza, a continuación proponemos un modelo de regresión de ambos parámetros.

Un modelo de sesgo y varianza

Cada encuestado tiene su propio sesgo y varianza al momento de predecir porcentajes de voto. Lo que queremos saber es cómo varían dichos sesgos y varianzas con características individuales como la edad, nivel educativo, identificación partidista, etcétera.

Nuestra variable de resultado es la diferencia entre el porcentaje de voto real v y el porcentaje predicho y . Como predictores x_i , usamos las siete variables mencionadas antes. Modelamos el resultado con una distribución normal y relacionamos cada parámetro con una variable independiente lineal,

$$\begin{aligned}v - y_i &\sim N(\mu_i, \sigma_i^2) \\ \mu_i &= x_i^T \beta \\ \sigma_i^2 &= \exp(x_i^T \gamma),\end{aligned}$$

donde μ_i es el sesgo (la diferencia esperada entre el porcentaje de voto real v y el porcentaje predicho y_i), y σ_i^2 es la varianza (del porcentaje de voto predicho y_i) para encuestado $i = 1, \dots, n$.

Elegimos la prior vaga $(\beta, \gamma) \sim N(0, 100 \cdot I_{16})$. Para estimar el modelo, codificamos el algoritmo métodos de Montecarlo basados en cadenas de Markov (MCMC) de Cepeda y Gamerman (2000) en el lenguaje de cómputo estadístico R. Ejecutamos el algoritmo MCMC para 1000 iteraciones y descartamos las primeras 100 como periodo de calentamiento. Las iteraciones no mostraron signos de no convergencia.

Hay tres estimandos de interés que constituyen las diferencias predictivas promedio en sesgo cuadrático, la varianza y el error cuadrático medio. Una diferencia predictiva compara el sesgo cuadrático, la varianza o el error cuadrático medio predicho entre dos encuestados —en este caso, un predictor ajustado al máximo más bajo de las cuatro elecciones, y otra con ese predictor ajustado al mínimo más alto de las cuatro elecciones—, pero idéntico para todos los demás predictores. Después, promediamos las diferencias predictivas de todos los n encuestados en los datos establecidos para calcular la diferencia predictiva promedio (Gelman *et al.*, 2020: sección 14.4). Para calcular la distribución a posteriori de este estimando, muestreamos 900 puntos de la distribución a posteriori de (β, γ) , luego calculamos la comparación predictiva promedio para cada uno.

Resultados

Los cuadros 7 a 9 del Apéndice C muestran cómo, para cada elección presidencial desde 2000 a 2018, la media y varianza (log) están asociadas con cada uno de los predictores para cada partido. Los cuadros contienen mucha información: con siete predictores para dos parámetros para tres partidos por cuatro años, el número de coeficientes equivale a $7 \times 2 \times 3 \times 4 = 168$. Con el fin de presentar esta información de manera más compacta, calculamos las comparaciones predictivas promedio como se explicó anteriormente. Luego, promediamos los resultados de los tres partidos y las cuatro elecciones, por separado para los tres estimandos de interés (sesgo cuadrático, varianza y error cuadrático medio).

Principales hallazgos

El cuadro 6 muestra las diferencias predictivas promedio en sesgo cuadrático, varianza y error cuadrático medio (MSE). Además, muestra dos resultados importantes. Primero, que las diferencias predictivas promedio para sesgo cuadrático y varianza por lo general difieren. Por ejemplo, para la identificación partidista, la diferencia predictiva promedio del sesgo cuadrático es aproximadamente el doble que la de la varianza. Esta y otras diferencias justifican de manera empírica nuestro enfoque propuesto para modelar sesgo y varianza al mismo tiempo.

Segundo, y más sustancialmente, el cuadro 6 presenta resultados significativos

CUADRO 6. Resúmenes de distribución a posteriori de las diferencias predictivas promedio para el sesgo cuadrático, varianza y error cuadrático medio (MSE)

	Sesgo ²	+	Varianza	=	MSE
Mujer (0 → 1)	+12.7 (0.1)		+5.3 (0.5)		+18.0 (0.5)
Edad (18 → 75)	+7.8 (0.3)		+5.5 (0.7)		+13.3 (0.8)
Educación (1 → 4)	-43.1 (0.2)		-38.7 (0.3)		-81.8 (0.4)
Interés (0 → 2)	+18.2 (0.2)		+6.0 (1.2)		+24.3 (1.2)
Partidista (0 → 1)	+285.5 (0.3)		+145.3 (6.3)		+430.8 (6.3)
Limpias (0 → 2)	+2.3 (0.2)		-35.5 (1.3)		-33.2 (1.3)
Experiencia (0 → 1.78)	-29.6 (0.1)		-17.6 (0.9)		-47.3 (0.9)

Fuente: Elaboración propia con base en datos de Investigaciones Sociales Aplicadas (2004). *Nota:* Medias a posteriori y errores estándar a priori (entre paréntesis).

sobre edad, experiencia con elecciones democráticas y percepción de elecciones limpias. Estos resultados son importantes porque ofrecen una explicación sobre la precisión de las predicciones en las democracias emergentes en específico. Explican por qué los ciudadanos, aunque en un principio en el agregado predijeron peor que los sondeos de intención de voto en México, han mejorado desde entonces.

Primero, examinaremos los resultados por edad. Como era de esperarse, los ciudadanos de edad avanzada predijeron peor que los jóvenes. En el promedio de los datos, en comparación con los ciudadanos de 18 años, los ciudadanos de 75 años tienen un sesgo cuadrático 7.8 puntos más elevado, una varianza 5.5 puntos más elevada y, por ello, un error cuadrático medio 13.3 puntos más elevado en los mismos niveles de los otros predictores.

Este hallazgo es importante porque Murr (2015) obtuvo el resultado opuesto en Estados Unidos: las personas de edad avanzada predicen mejor que la gente más joven. Nosotros sostuvimos antes que la razón de esto es que en Estados Unidos las personas de edad avanzada han experimentado más elecciones democráticas que los jóvenes. En contraste, para estas elecciones en México encontramos que la gente de edad avanzada predijo peor que la gente joven. La explicación sigue la misma lógica subyacente: la socialización temprana de las personas en los procesos democráticos influye en sus perspectivas futuras en la vida. Esto respalda el argumento de un efecto de cohorte en la predicción de elecciones.

En relación con esto, el cuadro 6 indica que, como era de esperarse, los ciudadanos que han vivido más tiempo bajo gobernadores no priístas predicen mejor que aquellos con menor experiencia. En la media de los datos, los ciudadanos con aproximadamente año y medio de experiencia con gobernadores no priístas tienen un sesgo cuadrático de 29.6 puntos más bajo, varianza 17.6 puntos más baja y, por lo tanto, error cuadrático medio 47.3 puntos más bajo que los ciudadanos sin dicha experiencia en los mismos niveles que los demás predictores. Este hallazgo es importante porque respalda el argumento de que la experiencia democrática es importante para la precisión de las predicciones.

En conjunto, ambos hallazgos sobre la edad y la experiencia democrática explican la baja precisión de los mexicanos para predecir en las primeras elecciones presidenciales, así como su creciente precisión a partir de entonces.

Por último, analizamos otro predictor que no había sido tomado en cuenta anteriormente en materia de predicciones ciudadanas: la percepción de que la elección será limpia. La idea es que si los ciudadanos perciben que las elecciones serán limpias, predicen mejor los porcentajes de voto de los candidatos. Como era de esperarse, el cuadro 6 muestra que aquellos que perciben elecciones limpias tienden a predecir mejor. En el promedio de los datos, en comparación con los ciudadanos que perciben elecciones fraudulentas, los ciudadanos que consideran que las elecciones serán limpias tienen un sesgo cuadrático 2.3 puntos más elevado, una

varianza 35.5 puntos más baja y, por lo tanto, un error cuadrático medio 33.2 puntos más bajo en los mismos niveles que los otros predictores. Una implicación de esto es que, mientras más ciudadanos perciban que las elecciones serán limpias, mejores serán las predicciones, siempre y cuando las instituciones pueden garantizar elecciones limpias en el futuro.

Hallazgos secundarios

Para tener un estudio completo, también discutimos las variables restantes. Los resultados para género, educación e identificación partidista encajan con hallazgos previos de democracias bien establecidas sobre las predicciones ciudadanas de quién ganará. El resultado para interés político, aunque no desconocido, es poco común.

El cuadro 6 evidencia que las mujeres predicen peor que los hombres, que los ciudadanos con mayor nivel educativo predicen mejor que aquellos con un nivel menor, y que la gente con identificación partidista predice peor que aquellos que no la tienen. Se han reportado hallazgos similares de predicciones sobre quién ganará en Austria, Alemania (Meffert *et al.*, 2011) y Estados Unidos (Lewis-Beck y Tien, 1999; Murr, 2015). Si bien las razones que justifican estos resultados están claras en el caso de la escolaridad e identificación partidista, siguen sin estarlo en el caso del género. Este hallazgo amerita nuevas investigaciones.

Asimismo, el cuadro 6 demuestra que los ciudadanos con mayor interés en la política predicen peor que los que no se interesan tanto en ella. Dicho resultado, aunque no desconocido, es inesperado, y es consistente con los resultados de predicción ciudadana sobre quién ganará en Estados Unidos. Hay una relación negativa a nivel nacional en la elección presidencial de 1956 (Lewis-Beck y Tien, 1999), y otra a nivel estatal en la elección presidencial en 1984 (Murr, 2015). Una posible explicación para esto es que, en este caso, el interés político capta el interés material o simbólico, lo que conduce a los ciudadanos a “hacerse ilusiones”. Sin embargo, se necesitan más investigaciones al respecto para analizar también más a fondo esta cuestión.

CONCLUSIÓN Y DISCUSIÓN


Hasta la fecha, la investigación sobre predicciones ciudadanas se ha enfocado en democracias bien establecidas y ha dejado de lado a las democracias emergentes. Si bien ese enfoque comparaba predicciones ciudadanas con otros métodos, carecía de las herramientas estadísticas necesarias para descartar que cualquier diferencia fuera resultado del azar. Y al explicar la capacidad de los ciudadanos para predecir, las investigaciones previas se centraban en un resultado discreto (si los ciudadanos predicen correcta o incorrectamente quién gana), pero ignoraban un resultado continuo más informativo (qué tan bien predicen los ciudadanos los porcentajes de votos para partidos o candidatos).

Nuestros estudios analizaron la predicción ciudadana en México desde los inicios de la democracia, en 2000, hasta 2024. Desarrollamos herramientas estadísticas para comparar las predicciones ciudadanas con predicciones aleatorias y con sondeos de intención de voto. Además, estudiamos qué variables sociodemográficas y políticas predicen el sesgo, la varianza y el error cuadrático medio de las predicciones ciudadanas para los porcentajes de voto. El Estudio 1 encontró que, de inicio, los ciudadanos predecían peor que el azar al ganador de la presidencia, pero que mejoraron con el tiempo y lograron predecir correctamente las últimas tres elecciones. Dicho estudio también concluye que, de inicio, los ciudadanos eran peores para predecir al ganador que las encuestas de intención de voto, pero que desde 2012 se emparejaron con estas últimas. El Estudio 2 encontró que, mientras que los ciudadanos predicen mejor que el azar los porcentajes de voto para cada candidato, han tenido peores resultados en este rubro que las encuestas de intención de voto.

En sintonía con la investigación sobre predictores de predicciones correctas en democracias establecidas, el Estudio 3 encontró que el género, la educación y la identificación partidista predicen los sesgos, la varianza y el error cuadrático medio para las predicciones de porcentajes de voto en México, una democracia emergente. Dicho estudio también encontró pruebas de la influencia del pasado hegemónico de partido único del PRI en los mexicanos, y que les tomó un tiempo, de manera individual y colectiva, ajustar sus convicciones. Los ciudadanos de edad avanzada, que crecieron bajo el régimen priísta, predicen peor que los más jóvenes. Este hallazgo apoya la idea del reemplazo generacional: los ciudadanos que tienen más experiencia con gobiernos no priístas en su estado de residencia predicen mejor que aquellos con menos experiencia en ese ámbito. Este hallazgo apoya la idea del ajuste de convicciones con el tiempo conforme los ciudadanos participan cada vez más en elecciones libres y justas. Así, los ciudadanos mexicanos tienen cada vez mayores preocupaciones sobre la celebración de elecciones limpias, a pesar de evidencias sobre lo contrario. Esto es importante porque descubrimos que aquellos que creen que las elecciones serán fraudulentas, predicen peor que los que no.

Desarrollamos algunas herramientas estadísticas nuevas y útiles para estudiar las predicciones ciudadanas. Asimismo, desarrollamos herramientas estadísticas para probar si las predicciones ciudadanas sobre quién ganará o sobre los porcentajes de voto son mejores que los sondeos de intención de voto. Desarrollamos también una herramienta estadística para estudiar el sesgo, la varianza y el error cuadrático medio de los ciudadanos al predecir porcentajes de votos. Vistas en conjunto, tales herramientas sientan bases estadísticas más sólidas para la predicción ciudadana y ofrecen oportunidades interesantes para aplicarlas en otros contextos. Los datos y el código R completo para replicar el análisis están disponibles

públicamente (Murr, 2025). Esperamos que otros investigadores utilicen estas herramientas estadísticas para estudiar las predicciones ciudadanas tanto en democracias bien establecidas como en democracias emergentes.

Al explotar los datos existentes y recopilar nuevos en distintos países a lo largo del tiempo, podremos estudiar las predicciones ciudadanas desde una perspectiva comparada. Esto nos permitirá probar varias hipótesis interesantes, tales como si las elecciones en democracias emergentes son más difíciles de predecir que las elecciones en democracias bien establecidas. Sospechamos que en América Latina existe una colección de datos inexplorada sobre predicciones ciudadanas. Para dar sólo un ejemplo, ya en 1957, Eduardo Hamuy, de la Universidad de Santiago de Chile, les hizo la siguiente pregunta a los residentes del área metropolitana de Santiago: “¿Qué candidato tiene más posibilidades de ganar la elección presidencial?” (Hamuy, 1958). De igual forma, en México algunas encuestas de Investigaciones Sociales Aplicadas les piden a los ciudadanos predecir otras elecciones federales (qué partido ganará la mayoría en las elecciones legislativas) y en elecciones estatales (qué candidato ganará la gubernatura). Por lo tanto, hay grandes oportunidades para estudiar más a fondo la predicción ciudadana en distintos tipos de elecciones y en diferentes contextos políticos. 

REFERENCIAS

- Arce, M. y S. Vera (2022), “Choosing the Lesser Evil: Forecasting Presidential Elections in Peru”, *Revista Latinoamericana de Opinión Pública*, 11(1), pp. 55-80.
- Beltrán, U. y M. Valdivia (1999), “Accuracy and Error in Electoral Forecasts: The case of Mexico”, *International Journal of Public Opinion Research*, 11(2), pp. 115-134.
- Bertholini, F., L. Rennó y M. Turgeon (2022), “Against All Odds: Forecasting Brazilian Presidential Elections in Times of Political Disruption”, *Revista Latinoamericana de Opinión Pública*, 11(1), pp. 129-147.
- Boon, M. (2012), “Predicting Elections: A ‘Wisdom of Crowds’ Approach”, *International Journal of Market Research*, 54(4), pp. 465-483.
- Bunker, K. (2020), “A Two-stage Model to Forecast Elections in New Democracies”, *International Journal of Forecasting*, 36(4), pp. 1407-1419.
- Bunker, K. (2022), “Forecasting Two-horse Races in New Democracies: Accuracy, Precision and Error”, *Revista Latinoamericana de Opinión Pública*, 11(1), pp. 81-108.
- Cantú, F., V. Hoyo y M.A. Morales (2016), “The Utility of Unpacking Survey Bias in Multiparty Elections: Mexican Polling Firms in the 2006 and 2012 Presidential Elections”, *International Journal of Public Opinion Research*, 28(1), pp. 96-116.
- Cepeda, E. y D. Gamerman (2000), “Bayesian Modeling of Variance Heterogeneity in Normal Regression Models”, *Brazilian Journal of Probability and Statistics*, 14(2), pp. 207-221.
- Charles, C.A.D., D. Dempster y T. Welcome (2022), “The Role of the Economy, Security and Party Leader Acceptance in Forecasting the 2020 General Election in Jamaica”, *Revista Latinoamericana de Opinión Pública*, 11(1), pp. 109-128.
- Dassonneville, R. y C. Tien (2020), “Introduction to Forecasting the 2020 US Election”, *ps: Political Science & Politics*, 54(1), pp. 47-51.

- Dufresne, Y., B. Jérôme, M.S. Lewis-Beck, A.E. Murr y J. Savoie (2022), “Citizen Forecasting: The 2022 French Presidential Election”, *ps: Political Science & Politics*, 55(4), pp. 730-734.
- Fisher, S.D. (2004), “Definition and Measurement of Tactical Voting: The Role of Rational Choice”, *British Journal of Political Science*, 34(01), pp. 152-166.
- Fisher, S.D. y M.S. Lewis-Beck (2016), “Special Symposium on Forecasting the 2015 British General Election”, *Electoral Studies*, 41, pp. 225-229.
- Ganser, C. y P. Riordan (2015), “Vote Expectations at the Next Level: Trying to Predict Vote Shares in the 2013 German Federal Election by Polling Expectations”, *Electoral Studies*, 40, pp. 115-126.
- Garrido de Sierra, S. (2019), *La reforma definitiva*, Ciudad de México, CIDE.
- Gelman, A., J. Hill y A. Vehtari (2020), *Regression and Other Stories*, Cambridge, Cambridge University Press.
- Gelman, A., J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari y D.B. Rubin (2014), *Bayesian Data Analysis*, 3a ed., Boca Raton, CRC Press.
- Graefe, A. (2014), “Accuracy of Vote Expectation Surveys in Forecasting Elections”, *Public Opinion Quarterly*, 78(S1), pp. 204-232.
- Graefe, A. (2016), “Forecasting Proportional Representation Elections from Nonrepresentative Expectation Surveys”, *Electoral Studies*, 42, pp. 222-228.
- Granberg, D. y E. Brent (1983), “When Prophecy Bends: The Preference-Expectation Link in U.S. Presidential Elections, 1952-1980”, *Journal of Personality and Social Psychology*, 45(3), pp. 477-491.
- Hamuy, E. (1958), “Political behavior in Chile, 1958”, Inter-university Consortium for Political and Social Research (distribuidor), DOI: <https://doi.org/10.3886/ICPSR07050.v1>.
- Investigaciones Sociales Aplicadas (2024), en: <https://Invesoc.com/> [fecha de consulta: 31 de julio de 2024].
- Jérôme, B. y A. Graefe (2021), “Forecasting the 2021 German Federal Election: An Introduction”, *ps: Political Science & Politics*, 55(1), pp. 61-63.
- Lee, P.M. (2012), *Bayesian Statistics: An Introduction*, Chichester, Wiley.
- Lehrer, R., S. Juhl y T. Gschwend (2019), “The Wisdom of Crowds Design for Sensitive Survey Questions”, *Electoral Studies*, 57, pp. 99-109.
- Leiter, D., A. Murr, E. Rascón Ramírez y M. Stegmaier (2018), “Social Networks and Citizen Election Forecasting: The More Friends the Better”, *International Journal of Forecasting*, 34(2), pp. 235-248.
- Lewis-Beck, M.S. (1985), “Election Forecasts in 1984: How Accurate Were They?”, *ps: Political Science and Politics*, 18(1), pp. 53-62.
- Lewis-Beck, M.S. (2005), “Election Forecasting: Principles and Practice”, *British Journal of Politics & International Relations*, 7(2), pp. 145-164.
- Lewis-Beck, M.S. y A. Skalaban (1989), “Citizen Forecasting: Can Voters See into the Future?”, *British Journal of Political Science*, 19(1), pp. 146-153.
- Lewis-Beck, M.S. y Tien, C. (1999), “Voters as Forecasters: A Micromodel of Election Prediction”, *International Journal of Forecasting*, 15(2), pp. 175-184.
- Meffert, M.F., S. Huber, T. Gschwend y F.U. Pappi (2011), “More than Wishful Thinking: Causes and Consequences of Voters Electoral Expectations about Parties and Coalitions”, *Electoral Studies*, 30(4), pp. 804-815.
- Moreno, A., R. Aguilar y V. Romero (2014), “Pre-election Poll Estimations in Mexico: In

- Search for the Main Sources of Error”, *Revista Latinoamericana de Opinión Pública*, 4, pp. 49-93.
- Murr, A.E. (2011), “Wisdom of Crowds? A Decentralised Election Forecasting Model that Uses Citizens’ Local Expectations”, *Electoral Studies*, 30(4), pp. 771-783.
- Murr, A.E. (2015), “The Wisdom of Crowds: Applying Condorcet’s Jury Theorem to Forecasting US Presidential Elections”, *International Journal of Forecasting*, 31(3), pp. 916-929.
- Murr, A.E. (2016), “The Wisdom of Crowds: What do Citizens Forecast for the 2015 British General Election?”, *Electoral Studies*, 41(1), pp. 283-288.
- Murr, A.E. (2017), “Wisdom of Crowds”, en K. Arzheimer, J.A.J. Evans y M.S. Lewis-Beck (eds.), *The SAGE Handbook of Electoral Behaviour*, Londres, SAGE, pp. 835-860.
- Murr, A. (2025), “Replication Data for: Citizen Forecasts of Mexican Presidential Elections, 2000-2024”, Harvard Dataverse, V1, DOI: <https://doi.org/10.7910/DVN/X5MXHA>
- Murr, A.E. y M.S. Lewis-Beck (2022), “Citizen Forecasts of the 2021 German Election”, *ps: Political Science & Politics*, 55(1), pp. 97-101.
- Murr, A.E., M. Stegmaier y M.S. Lewis-Beck (2021), “Vote Expectations Versus Vote Intentions: Rival Forecasting Strategies”, *British Journal of Political Science*, 51(1), pp. 60-67.
- Murr, A., R. Traunmüller y J. Gill (2023), “Computing Quantities of Interest and Their Uncertainty Using Bayesian Simulation”, *Political Science Research and Methods*, 11(3), pp. 623-632.
- Nadeau, R., B. Jérôme y M.S. Lewis-Beck (2022), “Introduction to Forecasting the 2022 French Presidential Election”, *ps: Political Science & Politics*, 55(4), pp. 683-686.
- Ratto, M.C. y M.S. Lewis-Beck (2022), “Argentinian Elections: Forecasting Outcomes”, *Revista Latinoamericana de Opinión Pública*, 11(1), pp. 17-37.
- Ratto, M.C., M.S. Lewis-Beck y E. Bélanger (2022), “Forecasting Elections in Latin America: An Overview”, *Revista Latinoamericana de Opinión Pública*, 11(1), pp. 5-13.
- Stegmaier, M. y H. Norpoth (2013), “Election Forecasting”, en R. Valelly (ed.), *Oxford Bibliographies in Political Science*, Oxford, Oxford University Press, en: <http://www.oxford-bibliographies.com/view/document/obo-9780199756223/obo-9780199756223-0023.xml> [fecha de consulta: 24 de julio de 2024].
- Temporao, M., Y. Dufresne, J. Savoie y C. van der Linden (2019), “Crowdsourcing the Vote: New Horizons in Citizen Forecasting”, *International Journal of Forecasting*, 35(1), pp. 1-10.
- Wasserman, L. (2004), *All of Statistics: A Concise Course in Statistical Inference*, Nueva York, Springer.

Apéndices

A. FORMULACIÓN DE LAS PREGUNTAS DE ENCUESTAS DE PREDICCIÓN CIUDADANA E INTENCIÓN DE VOTO

A continuación, presentamos las preguntas habituales que se hacen en cada elección para medir las predicciones ciudadanas de quién ganará la presidencia, sus intenciones de voto y las predicciones ciudadanas de los porcentajes de voto de cada candidato. Las preguntas que se usan dependen de si los candidatos ya fueron seleccionados (por lo general ya es así nueve meses antes de la elección). Antes de seleccionar a los candidatos, las preguntas se centran en los partidos; posteriormente, se usan las preguntas centradas en los candidatos. Las encuestas de salida se realizaron el día de la elección, cuando los votantes salen de las casillas electorales.

Predicciones ciudadanas de quién ganará

Versión centrada en el partido

- 2000: ¿Qué partido cree que va a ganar las próximas elecciones para presidente de la república?
- 2006: ¿Y qué partido político cree usted que ganará la próxima elección para presidente de la república?
- 2012: ¿Y qué partido político cree usted que ganará la próxima elección para presidente de la república?
- 2018: No se hizo la pregunta.
- 2024: ¿Qué partido cree usted que va a ganar la próxima elección para presidente de la república?

Versión centrada en el candidato

- 2000: Se hizo la pregunta, pero la formulación no está disponible.
- 2006: ¿Y cuál de estos candidatos cree usted que ganará la próxima elección para presidente de república?
- 2012: ¿Y quién cree usted que va a ganar la elección para la presidencia de la república?
- 2018: ¿Cuál de los candidatos a la presidencia de la república cree usted que ganará la elección?
- 2024: Independientemente de su preferencia partidista y su participación, ¿qué persona cree usted que va a ganar la elección para la presidencia de la república?

Versión de encuesta de salida

- 2000: No hay encuesta de salida disponible.
- 2006: ¿Y cuál de los candidatos cree usted que ganará esta elección para presidente de la república?
- 2012: No hay encuesta de salida disponible.

- 2018: ¿Cuál de los candidatos a la presidencia de la república cree usted que ganará la elección?
- 2024: Independientemente de por quién votó, ¿qué persona cree usted que va a ganar la elección para la presidencia de la república?

Intenciones de voto

Versión centrada en el partido

- 2000: Si tuviera usted que votar en este momento para elegir presidente de la república, ¿por cuál partido político votaría?
- 2006: Si tuviera que votar en este momento para elegir presidente de la república, ¿por cuál partido político votaría usted?
- 2012: Si tuviera que votar en este momento para elegir presidente de la república, ¿por cuál partido político votaría usted?
- 2018: No disponible
- 2024: Si tuviera que votar en este momento para elegir presidente de la república, ¿por qué partido votaría usted?

Versión centrada en el candidato

- 2000: Se hizo la pregunta, pero la formulación no está disponible.
- 2006: Si en este momento se celebraran las elecciones para presidente de la república, ¿por cuál candidato votaría usted si tuviera que elegir solamente entre...?
- 2012: Si en este momento se celebraran las elecciones para presidente de la república, ¿por quién votaría usted?
- 2018: Si en este momento se celebraran elecciones para presidente de la república, ¿por cuál de estos candidatos votaría usted?
- 2024: Si en este momento se celebrara la elección para la presidencia de la república, ¿por cuál candidatura votaría usted?

Versión de encuesta de salida

- 2000: No hay encuesta de salida disponible.
- 2006: ¿Por quién votó el día de hoy para presidente de la república?
- 2012: No hay encuesta de salida disponible.
- 2018: ¿Por quién votó usted hoy para la presidencia de la república?
- 2024: ¿Por cuál candidatura a la presidencia de la república votó usted hoy?

Predicciones ciudadanas de los porcentajes de voto

Versión centrada en el candidato

- 2000: Se hizo la pregunta, pero la formulación no está disponible.
- 2006: ¿Qué porcentaje de votos cree que obtendrá [F. Calderón / R. Madrazo / AMLO / Otros candidatos]?
- 2012: ¿Qué porcentaje de votos cree que obtendrá [Vázquez Mota / Peña Nieto / López Obrador / Quadri]?

- 2018: ¿Qué porcentaje de la votación nacional cree usted que finalmente obtendrá [Ricardo Anaya / José Antonio Meade / Andrés Manuel López Obrador / Jaime Rodríguez Calderón?]

Versión de encuesta de salida

- 2000: No hay encuesta de salida disponible.
- 2006: ¿Qué porcentaje de votos cree que obtendrá [F. Calderón / R. Madrazo / AMLO / Otros candidatos]?
- 2012: No hay encuesta de salida disponible.
- 2018: No hay encuesta de salida disponible.

B. FORMULACIÓN DE LAS PREGUNTAS PARA LAS VARIABLES SOCIODEMOGRÁFICAS Y POLÍTICAS

A continuación, presentamos la formulación de las preguntas para los datos sobre variables sociodemográficas y políticas, cuando estén disponibles. Los datos establecidos para la elección del año 2000 no incluyen formulaciones de preguntas, sólo etiquetas de variable.

Edad:

- 2006, 2018: ¿Cuántos años cumplidos tiene usted?
- 2012: Para el próximo 1º de julio, ¿cuántos años cumplidos va a tener usted?

Educación:

- 2006: ¿Hasta qué año de estudios llegó usted?
- 2012, 2018: ¿Hasta qué nivel de estudios llegó usted?

Interés en política:

- 2006, 2018: ¿Qué tanto interés tiene por informarse respecto a asuntos de política nacional?
- 2012: ¿Qué tanto interés tiene por informarse respecto al desarrollo de las campañas electorales?

Identificación partidista:

- 2006-2018: Independientemente de por quién ha votado, ¿con qué partido se identifica usted más?

Limpieza de las elecciones:

- 2006-2018: ¿Cree usted que las próximas elecciones federales serán limpias o que habrá fraude?
- 2006, 2012: ¿Cree que el IFE garantiza o no garantiza la imparcialidad en las próximas elecciones federales?
- 2018: ¿Cree que el INE garantiza o no garantiza la imparcialidad en las próximas elecciones federales?

C. MODELOS LINEALES HETEROCEDÁSTICOS DE ERRORES DE PREDICCIÓN

CUADRO 7. Modelo lineal heterocedástico del error de predicción de los porcentajes de voto del pan por elección presidencial de 2000-2018

	Error de predicción de los porcentajes de voto del PAN							
	2000		2006		2012		2018	
	Media	Varianza	Media	Varianza	Media	Varianza	Media	Varianza
Intercepto	17.29 (1.09) [15.22; 19.40]	4.91 (0.11) [4.72; 5.10]	15.38 (1.60) [12.28; 18.42]	5.03 (0.15) [4.77; 5.32]	7.79 (1.20) [5.42; 10.11]	3.72 (0.20) [3.34; 4.10]	9.69 (1.31) [7.09; 12.17]	4.02 (0.21) [3.59; 4.43]
Mujer	2.52 (0.47) [1.60; 3.53]	0.30 (0.04) [0.21; 0.38]	0.97 (0.57) [-0.13; 2.07]	0.08 (0.06) [-0.03; 0.19]	-0.76 (0.44) [-1.66; 0.11]	-0.08 (0.07) [-0.20; 0.07]	-0.03 (0.48) [-0.96; 0.89]	-0.14 (0.07) [-0.30; -0.00]
Edad	3.04 (1.94) [-0.75; 6.78]	0.07 (0.18) [-0.42; 0.28]	-3.39 (2.01) [-7.08; 0.62]	-0.23 (0.21) [-0.61; 0.17]	-0.33 (1.66) [-3.47; 2.89]	-0.35 (0.29) [-0.85; 0.18]	2.06 (1.98) [-1.92; 6.11]	2.46 (0.30) [1.91; 3.05]
Educación	0.92 (0.23) [-1.37; -0.46]	0.08 (0.02) [-0.12; -0.03]	-0.81 (0.30) [-1.38; -0.23]	-0.04 (0.03) [-0.10; 0.02]	-0.56 (0.26) [-1.10; -0.02]	-0.09 (0.04) [-0.18; -0.01]	-1.12 (0.31) [-1.74; -0.52]	-0.36 (0.05) [-0.46; -0.25]
Interés	0.14 (0.38) [-0.55; 0.90]	0.05 (0.03) [-0.02; 0.12]	0.09 (0.47) [-0.78; 1.00]	-0.09 (0.04) [-0.17; -0.01]	0.99 (0.38) [0.24; 1.73]	0.19 (0.06) [0.08; 0.30]	-0.77 (0.38) [-1.47; -0.02]	0.08 (0.07) [-0.05; 0.20]
Partidista	2.40 (0.54) [-3.47; -1.41]	0.29 (0.05) [0.20; 0.38]	5.15 (0.80) [3.59; 6.78]	1.08 (0.06) [0.97; 1.19]	11.01 (0.99) [8.97; 12.95]	1.60 (0.09) [1.43; 1.77]	13.83 (0.71) [12.41; 15.19]	0.55 (0.09) [0.39; 0.75]
Limpias	0.46 (0.32) [-0.12; 1.13]	0.11 (0.03) [-0.16; -0.05]	-0.02 (0.46) [-0.89; 0.86]	-0.10 (0.04) [-0.18; -0.01]	0.06 (0.29) [-0.50; 0.64]	-0.04 (0.04) [-0.13; 0.06]	0.11 (0.37) [-0.62; 0.85]	-0.12 (0.05) [-0.22; -0.02]
Experiencia	3.38 (0.48) [-4.32; -2.47]	-0.23 (0.04) [-0.31; -0.14]	-0.37 (0.32) [-1.01; 0.26]	-0.21 (0.04) [-0.28; -0.14]	0.53 (0.19) [0.17; 0.91]	0.12 (0.03) [0.07; 0.17]	0.12 (0.15) [-0.20; 0.41]	-0.08 (0.03) [-0.13; -0.02]
n	2206		1483		972		889	
Log. lik.	-8423.10		-5705.79		-3291.95		-3019.30	

Fuente: Elaboración propia con base en datos de Investigaciones Sociales Aplicadas (2004). Notas: Desviaciones estándar entre paréntesis. Intervalos de 95% entre corchetes. Log. lik. es el valor de verosimilitud logarítmica evaluado en las medias a posteriori.

CUADRO 8. Modelo lineal heterocedástico del error de predicción de los porcentajes de voto del PRD y Morena por elección presidencial de 2000–2018

	2000			2006			2012			2018		
	Media	Varianza		Media	Varianza		Media	Varianza		Media	Varianza	
Intercepto	9.95 (0.88) [8.22; 11.65]	4.75 (0.11) [4.54; 4.96]		15.29 (1.60) [12.11; 18.38]	5.13 (0.15) [4.83; 5.44]		10.67 (1.57) [7.56; 13.64]	4.24 (0.21) [3.77; 4.63]		17.75 (1.69) [14.46; 20.97]	5.30 (0.23) [4.82; 5.70]	
Mujer	0.26 (0.33) [-0.38; 0.86]	0.10 (0.05) [0.00; 0.18]		1.17 (0.60) [0.03; 2.33]	0.02 (0.06) [-0.08; 0.13]		-0.42 (0.52) [-1.45; 0.59]	-0.08 (0.07) [-0.21; 0.07]		0.32 (0.55) [-0.76; 1.43]	0.06 (0.07) [-0.07; 0.20]	
Edad	2.41 (1.43) [-0.57; 5.19]	0.34 (0.18) [-0.04; 0.71]		-2.70 (2.14) [-6.87; 1.66]	-0.03 (0.19) [-0.43; 0.30]		-2.42 (2.13) [-6.47; 1.93]	-0.15 (0.29) [-0.78; 0.37]		5.50 (2.36) [0.59; 9.75]	0.52 (0.31) [-0.08; 1.18]	
Educación	-0.27 (0.17) [-0.62; 0.07]	-0.11 (0.02) [-0.16; -0.08]		-0.61 (0.30) [-1.20; -0.07]	-0.06 (0.03) [-0.13; 0.00]		-0.48 (0.33) [-1.09; 0.19]	-0.08 (0.04) [-0.16; 0.01]		-0.16 (0.37) [-0.89; 0.53]	-0.33 (0.05) [-0.43; -0.23]	
Interés	0.16 (0.29) [-0.48; 0.70]	-0.00 (0.04) [-0.07; 0.08]		0.48 (0.47) [-0.40; 1.47]	0.05 (0.04) [-0.04; 0.13]		0.47 (0.50) [-0.45; 1.48]	0.07 (0.06) [-0.05; 0.19]		1.05 (0.46) [0.19; 1.97]	0.07 (0.07) [-0.05; 0.20]	
Partidista	23.30 (1.11) [21.20; 25.40]	1.73 (0.06) [1.62; 1.85]		3.43 (0.79) [1.93; 5.00]	0.93 (0.06) [0.83; 1.04]		2.49 (0.87) [0.92; 4.18]	0.66 (0.08) [0.50; 0.83]		-11.22 (0.56) [-12.41; -10.10]	-0.61 (0.08) [-0.76; -0.44]	
Limpias	-0.93 (0.25) [-1.41; -0.42]	-0.38 (0.03) [-0.44; -0.32]		-0.05 (0.49) [-0.97; 0.89]	-0.25 (0.04) [-0.34; 0.17]		0.85 (0.37) [0.15; 1.58]	-0.05 (0.04) [-0.13; 0.03]		1.15 (0.44) [0.34; 1.99]	-0.29 (0.06) [-0.41; -0.17]	
Experiencia	-1.46 (0.32) [-2.08; -0.84]	-0.43 (0.05) [-0.53; -0.34]		-1.16 (0.34) [-1.82; 0.51]	-0.14 (0.04) [-0.21; -0.07]		0.79 (0.24) [0.32; 1.24]	0.14 (0.03) [0.07; 0.19]		-0.57 (0.19) [-0.93; -0.21]	-0.02 (0.03) [-0.07; 0.04]	
n	2 206			1 483			972			889		
Log. lik.	-7908.43			-5692.34			-3451.19			-3158.47		

Fuente: Elaboración propia con base en datos de Investigaciones Sociales Aplicadas (2004). *Notas:* Desviaciones estándar entre paréntesis. Intervalos de 95% entre corchetes. Log. lik. es el valor de verosimilitud logarítmica evaluado en las medias a posteriori.

CUADRO 9. Modelo lineal heterocedástico del error de predicción de los porcentajes de voto del PRI por elección presidencial de 2000-2018

	Error de predicción de los porcentajes de voto del PRI							
	2000		2006		2012		2018	
	Media	Varianza	Media	Varianza	Media	Varianza	Media	Varianza
Intercepto	15.82 (1.14) [13.59; 18.07]	4.78 (0.11) [4.56; 5.01]	11.46 (1.17) [9.09; 13.78]	4.42 (0.15) [4.13; 4.70]	13.50 (1.60) [10.34; 16.47]	4.26 (0.19) [3.90; 4.63]	8.32 (1.22) [5.84; 10.69]	4.81 (0.20) [4.41; 5.24]
Mujer	1.91 (0.49) [0.99; 2.82]	0.25 (0.04) [0.17; 0.34]	0.84 (0.44) [0.00; 1.76]	0.18 (0.06) [0.07; 0.29]	-0.02 (0.55) [-1.11; 1.06]	0.09 (0.07) [-0.05; 0.22]	-0.52 (0.42) [-1.40; 0.27]	-0.39 (0.07) [-0.53; -0.23]
Edad	1.57 (1.93) [-2.27; 5.39]	0.19 (0.18) [-0.14; 0.56]	-1.44 (1.60) [-4.58; 1.80]	-0.40 (0.20) [-0.76; 0.00]	-2.93 (2.23) [-7.17; 1.49]	-0.91 (0.27) [-1.42; -0.44]	1.22 (1.66) [-2.08; 4.42]	0.02 (0.30) [-0.60; 0.54]
Educación	-0.95 (0.24) [-1.42; -0.46]	-0.09 (0.02) [-0.14; -0.05]	-0.48 (0.22) [-0.89; -0.05]	-0.17 (0.03) [-0.23; -0.11]	-1.10 (0.35) [-1.74; -0.42]	-0.10 (0.04) [-0.19; -0.03]	-0.15 (0.27) [-0.67; 0.36]	-0.44 (0.05) [-0.54; -0.35]
Interés	-0.46 (0.40) [0 1.22; 0.32]	-0.08 (0.04) [0 0.16; 0.02]	-0.18 (0.34) [0 0.85; 0.48]	0.001 (0.04) [0 0.09; 0.07]	1.19 (0.51) [0.16; 2.15]	0.13 (0.06) [0.01; 0.25]	0.28 (0.34) [0 0.34; 0.94]	0.05 (0.06) [0 0.06; 0.19]
Partidista	6.71 (0.62) [5.43; 7.86]	0.61 (0.05) [0.50; 0.69]	12.93 (1.02) [10.90; 14.86]	1.71 (0.06) [1.58; 1.84]	0.65 (0.68) [0 0.69; 2.02]	0.52 (0.07) [0.38; 0.66]	15.64 (0.79) [14.02; 17.13]	1.08 (0.10) [0.91; 1.24]
Limpias	0.022 (0.34) [-0.87; 0.44]	0.01 (0.03) [-0.05; 0.07]	0.30 (0.34) [-0.98; 0.33]	0.00 (0.04) [-0.08; 0.09]	0.10 (0.36) [-0.59; 0.80]	0.09 (0.04) [0.00; 0.17]	0.042 (0.32) [-1.02; 0.22]	0.020 (0.05) [-0.31; -0.09]
Experiencia	-2.55 (0.47) [-3.51; -1.61]	-0.22 (0.05) [-0.32; -0.12]	-0.12 (0.28) [-0.67; 0.41]	0.08 (0.03) [0.00; 0.14]	0.21 (0.26) [-0.32; 0.72]	0.19 (0.03) [0.13; 0.25]	-0.18 (0.15) [-0.48; 0.10]	0.09 (0.02) [0.05; 0.14]
n	2206		1483		972		889	
Log. lik.	-8431.92		-5343.60		-3507.07		-2905.2	

Fuente: Elaboración propia con base en datos de Investigaciones Sociales Aplicadas (2004). *Notas:* Desviaciones estándar entre paréntesis. Intervalos de 95% entre corchetes. Log. lik. es el valor de verosimilitud logarítmica evaluado en las medias a posteriori.