

Las encuestas en la elección presidencial de México, 2018

The Polls in the Mexican Presidential Election, 2018

Ricardo de la Peña Mena¹

Recibido el 27 de agosto de 2018.

Aceptado el 29 de julio de 2019.

Resumen

En este ensayo se pretende hacer un balance del rendimiento de las encuestas en ocasión de la elección presidencial de 2018. Para ello, en un primer apartado, se definen las fuentes de información pertinentes que dan cuenta del fenómeno. En un segundo apartado se repasan los conceptos teóricos sobre cómo debe medirse la exactitud de las encuestas electorales. En el tercero, se descubren los defectos de mediciones que no se apegan a los procedimientos ortodoxos del campo, con miras a profundizar, en el apartado siguiente, en el rendimiento logrado por las encuestas en ocasión de los comicios presidenciales de 2018. En un apartado posterior, se analizan algunas cuestiones que revelaron aspectos todavía problemáticos de explicar. Finalmente, se evalúa el rendimiento logrado este año en comparación con el que se observó en otras ocasiones.

Palabras clave

México, presidente, encuestas, elecciones, exactitud, pronóstico.

Abstract

This essay intends to take stock of the performance of the presidential elections polls of 2018. In a first section, the relevant sources of information that account for the phenomenon are defined. In a second section the theoretical concepts are reviewed on how the accuracy of the electoral surveys should be measured. In a third section the defects of measurements that do not conform to the orthodox procedures of the field are discovered, with a view to deepen, in the following point, in the performance obtained by the polls on the occasion of the presidential elections of 2018, although leaving in a latter section some aspects that revealed still problematic aspects to explain. Finally, the performance achieved in this year is evaluated in comparison with that observed in other occasions.

Keywords

Mexico, president, polls, elections, accuracy, forecast.

¹ Presidente ejecutivo de Investigaciones Sociales Aplicadas (ISA), secretario de organización de la Sociedad Mexicana de Estudios Electorales, y coordinador de la Comisión de Relaciones Académicas del Colegio de Especialistas en Demoscopia y Encuestas (CEDE). ricartur@gmail.com

“No debe haber barreras a la libertad de investigación... No hay lugar para el dogma en la ciencia. El científico es libre, y debe ser libre de hacer cualquier pregunta, dudar de cualquier afirmación, buscar cualquier evidencia, corregir cualquier error.”

J. Robert Oppenheimer²

Introducción

La noche del 1 de julio de 2018, los contendientes a la presidencia de la República que no resultaron favorecidos por el voto ciudadano reconocieron rápidamente este hecho, lo que se confirmó horas después con el anuncio del consejero presidente del Instituto Nacional Electoral (INE) sobre los resultados del conteo rápido institucional, que daba una clara y contundente victoria al candidato que resultó oficialmente ganador.

El resultado que se había pronosticado para el triunfador no fue preciso ni respecto a la amplitud de su victoria ni mucho menos a la altura final lograda por cada contendiente. El margen de victoria fue mayor al estimado por la mayoría del electorado, especialistas y actores políticos. Incluso el promedio de las encuestas y las proporciones de los dos candidatos derrotados más importantes resultaron claramente inferiores a lo esperado.

A pesar de lo anterior, el balance público con respecto a las encuestas resultó afortunado esta vez, pues remarcó que las mediciones hechas públicas habían observado correctamente quién ganaría y que lo haría por un amplísimo margen. Eso propició una complacencia que llevó a afirmar que “2018 fue, en general, un buen año para las encuestas en nuestro país”.³ Pero, ¿lo fue? Hay evidencia para afirmar lo contrario.

En principio y antes de pasar a los datos, se debe recordar que todo balance serio tiene luces y sombras que dibujan una imagen de claroscuro; el de las encuestas en la elección presidencial de México en 2018 no es, y no tiene por qué ser, la excepción. Así, el objetivo de este ensayo fue la realización de un análisis empírico para determinar cuál fue el rendimiento de los ejercicios de medición por encuesta en ocasión de la elección para presidente de la República de 2018, atendiendo tanto lo bueno como lo malo de las encuestas científicas, no con el ánimo de sumarse al aplauso fácil, ni con la intención de defenestrar al campo o sus operadores. El propósito fue, por tanto, otro: apuntar consensos, pero advertir riesgos; alimentar la construcción de un canon pero, al mismo tiempo, abrir espacios al espíritu innovador dentro de esta actividad.

Antes que nada y con esta finalidad en mente, fue necesario determinar cuál era el conjunto de estudios sobre el que se podía y era confiable realizar un análisis, así como cuál era la fuente pertinente de información por utilizar sobre sus estimaciones. En seguida, hubo que discutir cuáles eran los indicadores usuales disponibles para la evaluación de la exactitud de las encuestas electorales *versus* los resultados, lo cual constituyó el objetivo central del estudio. El cuerpo

2 “There must be no barriers to freedom of inquiry... There is no place for dogma in science. The scientist is free, and must be free to ask any question, to doubt any assertion, to seek for any evidence, to correct any errors” (Oppenheimer citado en Lincoln Barnett, “J. Robert Oppenheimer”, *Life*, vol. 7, núm. 9, edición internacional, Time Incorporation, Nueva York, EE. UU., 24 de octubre, 1949, p. 58.

3 Lorena Becerra y Rodrigo León, “Nitidez en la contienda”, *Reforma Revista R*, Reforma, Ciudad de México, México, 8 de julio, 2018, p. 14.

central del ensayo se destinó a la aplicación de estos indicadores sobre la colección de estudios seleccionados para el análisis.

Una vez finalizado el ejercicio descriptivo, se propuso un mecanismo para colapsar diversos estimadores disponibles en un único indicador que buscara servir como una suerte de “medidor universal” de la exactitud de las encuestas. Pero, ¿era posible dar cuenta en un único dato de la riqueza de una distribución? Ello no siempre parecía factible, pues todo indicador que se postulara enfrentaría limitaciones, al reducir varias unidades de información a una sola.

¿A qué se llegó en este ensayo? Por una parte, se pudo ver que, luego de un caos de datos iniciales producto de mediciones disímbricas, al final se consolidó y reivindicó la ortodoxia en los métodos de investigación. En este sentido, si lo que se revisó fue si las encuestas habían alcanzado el objetivo de dar cuenta del formato de la contienda y pronosticar correctamente el orden de llegada de los competidores, se concluyó que los ejercicios en vivienda, en general, habían cumplido satisfactoriamente, como lo mostró la visión predominante en el público atento.

Sin embargo, si se atendía a la exactitud de las mediciones con respecto al resultado, la capacidad anticipatoria fue, en esa ocasión, relativamente reducida e inferior a la que se había logrado en otras ocasiones durante este siglo,⁴ lo que no dejó de referir la existencia de problemas y límites a la herramienta y su empleo, lo que obligó a concluir que existían aspectos que deberían revisarse y mejorar.

El hecho de que no haya una crítica pública hacia las encuestas en 2018 no quiere decir que sea pertinente adoptar una actitud de autocomplacencia. En lugar de ello, debiera avanzarse en una crítica interna en el campo demoscópico, que advierta y analice errores, posibles orígenes de éstos, así como alcances y riesgos que representan para el futuro inmediato.

Igualmente hay que considerar que los problemas de estimación se presentan en situaciones específicas, sin que existan reglas generales que permitan saber, *ex ante* y con cualquiera que sea el criterio de acierto o exactitud que se observe, cuándo y dónde ocurrirán fallos, y cuándo y dónde las encuestas estarán en lo correcto.

Fuentes de información

Hoy día existe un repositorio institucional de las encuestas que son entregadas a la autoridad administrativa electoral nacional, en el que se indica si cumplen con los requisitos legalmente demandados para su publicación.⁵ En este almacén informativo permanente, de acceso universal y gratuito, se pueden consultar los resultados, las bases de datos de los estudios, así como conocer los datos técnicos básicos, los responsables de su realización, patrocinio y difusión, y su costo. Esta es la principal fuente informativa que se empleó para recuperar los datos sobre las encuestas dadas a conocer durante 2018 sobre la elección por la presidencia de la República en México, y la que se utilizó como fuente definitiva para corroborar los resultados de los estudios analizados en este ensayo.

4 Cfr. Ricardo de la Peña, “Las encuestas presidenciales en México durante el siglo XXI”, *Revista Mexicana de Estudios Electorales*, núm. 15, Sociedad Mexicana de Estudios Electorales, D. F., México, enero-junio, 2015, pp. 78-114.

5 Instituto Nacional Electoral (INE), “Encuestas electorales: elecciones federales ordinarias 2017-2018. Estudios entregados”, INE. Disponible en <<http://www.ine.mx/voto-y-elecciones/encuestas-electorales/elecciones-federales-ordinarias-2017-2018-estudios-entregados/>>, [fecha de consulta: julio, 2018].

Existen otras fuentes que pueden ser consultadas, las cuales suelen ser más expeditas para la disposición pública de resultados y bases de datos de los estudios durante las campañas. Una de ellas, sin duda la mejor y más completa, es el *Observatorio Electoral*, que se halla a cargo del Colegio de Especialistas en Demoscopia y Encuestas, A. C. (CEDE).⁶ Aunada a ella, para revisar las estimaciones hechas públicas, se pudo recurrir también a la principal agregadora de encuestas: *Oraculus*,⁷ surgida precisamente para el proceso electoral 2017-2018.

El listado de encuestas elegidas para este ensayo no correspondió al de ninguna de esas fuentes. El almacén del INE no es selectivo: incorporó toda medición que le fue reportada sin distinción alguna, por lo que entre los múltiples estudios que conjuntó hubo que discriminar aquellos que resultaría pertinente considerar para llevar a cabo un análisis con sentido.

Por lo anterior, se tomó como guía para el análisis la selección de estudios incluida en el *Observatorio Electoral* del CEDE, la cual incorporó las encuestas nacionales relativas a las preferencias para la elección por la presidencia de la República tomando en cuenta dos criterios. El primero, que las encuestas fueran realizadas por empresas conocidas que formaran parte de las agrupaciones gremiales (o que hubieran sido reportadas a la autoridad en su momento), y se excluyó estudios que, por los objetivos o vínculos de su realizador, hubieran podido servir no para informar sino para intentar conducir las creencias o decisiones de los electores (un fenómeno conocido como efecto *herding*). El segundo, que fueran el resultado de ejercicios por muestreo, realizados mediante entrevistas personales en vivienda, a partir de los cánones de la investigación demoscópica. Por ende, esta compilación resultó ser una base informativa más homogénea y, gracias a ello, científicamente más útil que otras fuentes informativas sobre la materia.

El CEDE terminó incluyendo en su listado no sólo las encuestas dadas a conocer en fecha próxima a su realización, sino mediciones que fueron reportadas de forma tardía, pero bajo la responsabilidad de empresas con integrantes que formaban parte de este Colegio. En un caso, sin embargo, los resultados de una empresa no vinculada a agrupaciones gremiales, *Conteo*, fueron incorporados en los primeros meses, pero no al cierre del proceso.

En el listado de estudios usado para este análisis, se subsanó esa inconsistencia incluyendo la serie completa de la empresa en cuestión. De esta manera, se contabilizaron 67 encuestas, frente a 64 incluidas en el acervo definitivo del CEDE (véase la Gráfica 1).⁸

Asimismo, se asumió la corrección de las estimaciones producto de la encuesta de mayor tamaño, siguiendo el criterio que la propia casa responsable, Berumen y Asociados, privilegiara en su reporte oficial, al INE, de los resultados de su estudio.⁹

Adicionalmente, se conjuntaron los datos relativos a seis ejercicios de encuesta de salida, publicados luego del cierre oficial de las casillas en todo el territorio nacional y durante la hora

6 Colegio de Especialistas en Demoscopia y Encuestas (CEDE), "Observatorio Electoral 2018", CEDE. Disponible en <<http://cede.org.mx/web2016/observatorio-electoral/>>, [fecha de consulta: 16 de julio, 2018].

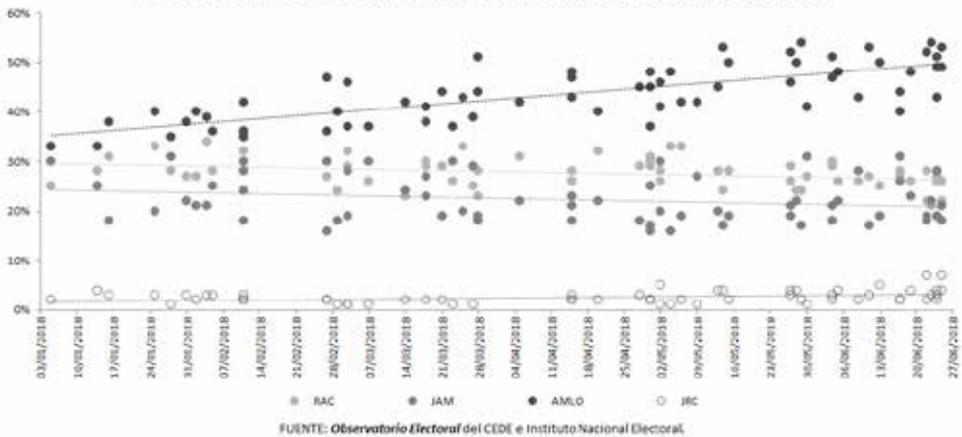
7 Oraculus, "Poll of Polls", *Oraculus*. Disponible en <<https://oraculus.com.mx/>>, [fecha de consulta: 16 de julio, 2018].

8 En archivo electrónico se pone a disposición la lista y datos de los estudios considerados en este análisis. Cfr. Ricardo de la Peña, "Encuestas en vivienda para la elección por la presidencia de la República en México 2018" [base de datos], *Ricardo de la Peña*. Disponible en <<http://www.ricartur.com/databases/EncuestasPresidenciales2018.txt>>, [fecha de consulta: 14 de agosto, 2018].

9 Disponible en Instituto Nacional Electoral (INE), "Estudios entregados a la Secretaría Ejecutiva", INE. Disponible en <<https://www.ine.mx/voto-y-elecciones/encuestas-electorales/elecciones-federales-ordinarias-2017-2018-estudios-entregados/berumen-2/>>, [fecha de consulta: 16 de julio, 2018].

inmediata posterior para dar cuenta de las tendencias de votación en la jornada electoral, que fueron reportados a la autoridad electoral competente.

Gráfica 1. Estimaciones por encuesta en vivienda para la elección presidencial 2018



Junto con las estimaciones puntuales por contendiente, se incluyeron dos datos adicionales: uno, la diferencia o margen de ventaja del primer lugar respecto del segundo en la elección, mas no en la encuesta (porque puede ser distinto), y dos, un indicador de la concentración esperada del voto, con base en las estimaciones reportadas: el número efectivo de contendientes (N o N^* si se agrupan proporciones para componentes menores), que correspondía al inverso de la sumatoria de los cuadrados de las proporciones estimadas para cada uno de los candidatos.¹⁰

Hay que mencionar que en estos cálculos, como en las estimaciones de exactitud que posteriormente se explicaron, invariablemente se consideraron solamente los casos definidos en las encuestas, y los votos válidos por contendiente registrados en las elecciones; de esta manera se satisfizo el objetivo central, que era generar una base de comparación homogénea.

Cómo medir la exactitud de encuestas

Como lo que se pretendía era llevar a cabo un análisis empírico original, no fue necesario realizar ni siquiera una somera revisión de la bibliografía previa en la que se efectuaran ejercicios similares, pues ello no resultaba pertinente, considerando que no se estaban tomando dichos análisis previos como soporte o elemento por cotejar, ni se estaba buscando validar el análisis a partir de una comparación con dichos estudios, sino que se pretendía realizar un ejercicio empírico que partiera de fuentes primarias y comparar entre ellos diversos estimadores (definidos en el texto), ocurridos en diversos eventos por cotejar.

10 Markku Laakso y Rein Taagepera, "Effective Number of Parties: A Measure with Application to West Europe", *Comparative Political Studies*, núm. 12, SAGE, Oslo, Noruega, 1979, pp. 3-27.

Al menos en ese acercamiento, el marco para la teorización partió de elementos conceptuales, así como de procedimientos y métodos analíticos propios de la metrología y la estadística, más que de consideraciones teóricas convencionales de la ciencia política. Por supuesto que era viable pasar en algún análisis posterior a un ejercicio de reflexión más propio del campo politológico o sociológico, para ir hacia la búsqueda de los factores que pudieran estar detrás de la exactitud misma de las encuestas, o que explicaran los patrones de votación de los ciudadanos. Pero esa era una tarea diferente, que no se planteó tratar en este ensayo.

En lo referente a los elementos conceptuales y metodológicos peculiares que la investigación demoscópica ha desarrollado en el nivel mundial, a lo largo de décadas, hay que recordar, por una parte, que es de sobra conocido el juicio de Warren Mitofsky (como son los criterios para la medición de la exactitud de los ejercicios por encuesta): “después de más de cincuenta años de encuestas electorales [...] no ha sido adoptada por la comunidad de encuestadores una métrica estándar para medir la exactitud de las encuestas”.¹¹ Esa afirmación fue hecha hace ya más de un cuarto de siglo y sigue manteniéndose vigente.

Por otra parte, si se consideran los diversos criterios disponibles para la medición de la exactitud de las encuestas, se encontrará que existen divergencias entre la lectura que hace el público interesado y la que realizan los especialistas.

Los legos —analistas políticos, actores en la arena electoral, comunicadores— han estado más atentos a lo “atinado” del “pronóstico” y por ende han puesto énfasis en dos criterios: primero, si se “acertó” al ganador de la contienda, y segundo, si se reportó un ordenamiento de los adversarios que resultara el correcto. Los investigadores que han apuntado a la pertinencia de una valoración de las mediciones electorales por encuesta como pronósticos hubieran debido, también, atenderse primordialmente a estos criterios.

Los especialistas que desean sujetarse a reglas más científicas que el mero orden de llegada en la “carrera de caballos” suelen buscar métodos estadísticos más científicos, que les permiten la apreciación de la exactitud de las mediciones. Desde hace ya siete décadas se cuenta con varios medidores sobre la adecuación entre las encuestas y los resultados, contenidos los más usuales de ellos en un conocido estudio de Fredrick Mosteller,¹² que fue elaborado para evaluar los resultados de las encuestas en Estados Unidos en la elección presidencial de 1948.

De los diversos medidores ahí presentados, el campo profesional estadounidense ha tendido a privilegiar el empleo de dos: el que mide la diferencia absoluta entre lo previsto y lo real para los dos mayores contendientes respecto del total de votos —que suele dividirse entre dos para compensar el hecho de que se trata de dos estimadores altamente relacionados, sobre todo en sistemas bipartidistas, en donde lo que uno gana otro lo suele perder ($M5/2$), y el más completo y propio para sistemas pluripartidistas, correspondiente a la desviación media en puntos porcentuales entre lo previsto y lo real para todos los contendientes, sin tener en cuenta el signo ($M3$; o $M3^*$, si para su cálculo se agrupan componentes menores, tomando por tales aquellos para los que el error relativo esperado sería mayor que la proporción medida, lo que ocurre cuando dicha proporción es menor a 10%).

11 Warren J. Mitofsky, “Review: Was 1996 a Worse Year for Polls than 1948?”, *Public Opinion Quarterly*, vol. 62, núm. 2, Oxford University Press, Oxford, Reino Unido, 1998, pp. 230-249.

12 Fredrick Mosteller, “Measuring the Error”, en Fredrick Mosteller, Herbert Hyman, Philip J. McCarthy, Eli S. Marks and David B. Truman, *The Pre-election Polls of 1948. Report to the Committee on Analysis of Pre-Election Polls and Forecasts*, Social Science Research Council, Nueva York, EE. UU., 1949, pp. 54-80.

Dado que estos estimadores de (in)exactitud no están exentos de problemas,¹³ en el presente siglo han surgido diversas propuestas para su mejoramiento, algunas orientadas a modelos bipartidistas. Hay también otra búsqueda más reciente, directamente encaminada a atender el reclamo de un mejor estimador para sistemas multipartidarios:¹⁴ la media del error normalizado de las estimaciones por cada contendiente respecto del dato real (E), definido como el cociente del error absoluto observado entre el error esperado, que desde luego disminuye —de manera no lineal, sino parabólica— a medida que aumenta el tamaño de la muestra disponible y que la proporción de ocurrencia del evento considerado se reduce. Este estimador otorga así un peso equilibrado a la (in)exactitud por los diversos contendientes sin importar su tamaño y toma en cuenta la magnitud de la muestra observada para ponderar la exactitud lograda, por lo que refleja la relevancia del impacto de la ampliación del tamaño de muestra y la relación costo-beneficio lograda con ello.

En este análisis no se eliminó ninguno de los criterios fundamentales disponibles de evaluación de las encuestas, y se evitó jerarquizarlos. Por ende, se renunció de antemano al tradicional ordenamiento en tablas donde se presentaban las estimaciones, producto de encuestas finales, según mayor o menor grado de exactitud conforme a un criterio (privilegiado por el autor del análisis) que muchas veces se modificaba de una a otra elección, dependiendo de lo que conviniere al interés de quien escribía. Es claro que el orden que se derivaba de la aplicación de uno u otro criterio no era idéntico y que, en muchos casos, dicho orden incluso resultaba contradictorio.

Los criterios incorporados al estudio fueron cinco, dependiendo de si el ejercicio demoscópico detectaba acertadamente al ganador como líder en la encuesta; de si el ordenamiento que arrojaba coincidía con el que se presentaba en la realidad; del nivel de proximidad en la determinación de la ventaja que lograba el ganador, medido por $M5/2$, y del nivel de exactitud logrado para todos los componentes relevantes, medido por $M3^*$. Adicionalmente se estimó E , para amortiguar los efectos que los diferenciales de tamaño de muestra pudieran tener en los niveles esperables de exactitud de los estudios.

Se ha dejado a los lectores las tareas de elegir y valorar a qué criterios debiera darse un mayor o menor peso y de, en todo caso, cotejarlos con otros análisis con similares pretensiones, algunos publicados en fechas recientes.¹⁵ En este texto simplemente se anotó lo que resultó de la aplicación de estos principios. Al final, se caminó al encuentro de algún método de síntesis de los diversos criterios existentes, cuyas características se precisaron en el apartado final de este mismo artículo, y que partió de la agregación de cuatro de los cinco criterios de determinación de la exactitud de las encuestas anotados. Éstos fueron: si se detectó acertadamente al ganador como líder en la encuesta, si el ordenamiento coincidió con el que se presentó en la realidad, el nivel de proximidad en la determinación de la ventaja que logró el ganador (medido por $M5/2$), y el nivel de exactitud logrado para todos los componentes relevantes (medido por $M3^*$). Por

13 Al respecto, véase Ricardo de la Peña, "Cómo se mide la exactitud de las encuestas electorales", *Política y Cultura*, núm. 44, Universidad Autónoma Metropolitana (Xochimilco)-División de Ciencias Sociales y Humanidades-Departamento de Política y Cultura, d. f., México, 2015, pp. 217-247.

14 Ricardo de la Peña, "Una alternativa para estimar la exactitud de las encuestas", *Política y Cultura*, núm. 49, Universidad Autónoma Metropolitana (Xochimilco)-División de Ciencias Sociales y Humanidades-Departamento de Política y Cultura, d. f., México, 2018, pp. 123-156.

15 Aun quien suscribe este ensayo ha analizado en repetidas ocasiones la bibliografía sobre el tema en escritos que se citan en este mismo texto. Se puede agregar el reciente capítulo final del libro de Alejandro Moreno, *El cambio electoral. Votantes, encuestas y democracia en México*, Fondo de Cultura Económica, Ciudad de México, México, 2018.

razones de universalización del indicador, se dejó fuera del proceso de agregación el medidor ponderado conforme el tamaño de muestra (E).

El mérito del método

Al inicio del año electoral aparecieron, sobre todo en redes sociales, miles de ejercicios que tenían la pretensión de ser considerados como “encuestas” y que supuestamente reflejaban las preferencias del electorado, pero que en realidad generaban un ambiente de confusión y descrédito. Sin embargo, a medida que la campaña fue avanzando, estas mediciones arbitrarias fueron perdiendo presencia y atención. Aparecieron entonces ejercicios —al menos en el así llamado “círculo rojo”—¹⁶ que de verdad cumplían con los cánones científicos y, sobre todo hacia el final, basados en las encuestas realizadas en vivienda y en métodos tradicionales, propios de la ortodoxia en el campo demoscópico.

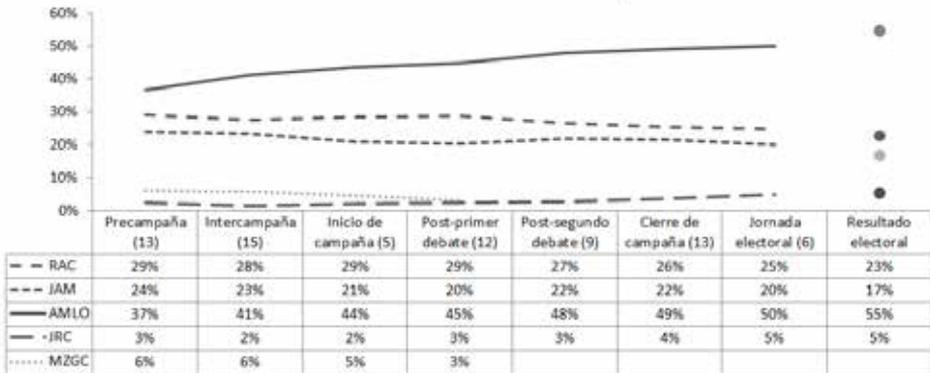
Esto fue consecuencia de múltiples sucesos: 1) lo increíble y contradictorio de los datos que arrojaban los procedimientos de suma de respuestas de usuarios de redes sociales, sin control ni método alguno, ya fuera que hubieran sido realizados bajo la responsabilidad de un periodista conocido o simplemente por usuarios deseosos de hacer su propia medición; 2) la insistencia de los profesionales de la actividad, los encargados de espacios de agregación de datos de encuestas y algunos comunicadores responsables e informados, sobre la necesidad de diferenciar y llevar a cabo un filtro de la enorme cantidad de números que circulaban para destacar aquellos que realmente brindaban información de calidad, y finalmente 3) la coherencia de la narrativa que el conjunto de encuestas tradicionales iba generando, y que otorgaba creciente credibilidad a sus resultados.

Al respecto, al hacer un recuento de las encuestas serias en vivienda publicadas durante 2018, relativas a las preferencias para la elección por la presidencia de la República, agrupándolas en períodos pertinentes (Gráfica 2), se pudo descubrir que la estimación de intenciones de voto por Andrés Manuel López Obrador fue creciendo sistemáticamente a lo largo del proceso, y que su ventaja respecto del contendiente más cercano se fue ampliando: de 37% y ocho puntos en precampaña, a 44% y quince puntos al arranque de la campaña, hasta cerrar en 49% y 24 puntos antes de la elección, y un poco más en los ejercicios a la salida de casillas.

Eso significó no sólo que paulatinamente las encuestas en conjunto fueron dando un dibujo próximo a lo que fue el resultado, sino que detectaron una creciente concentración del voto que terminó dando lugar a una contienda mucho menos disputada que en ocasiones anteriores.

16 Concepto coloquial por el que se conoce a los segmentos de la población con capacidad de intervención o influencia en la toma de decisiones.

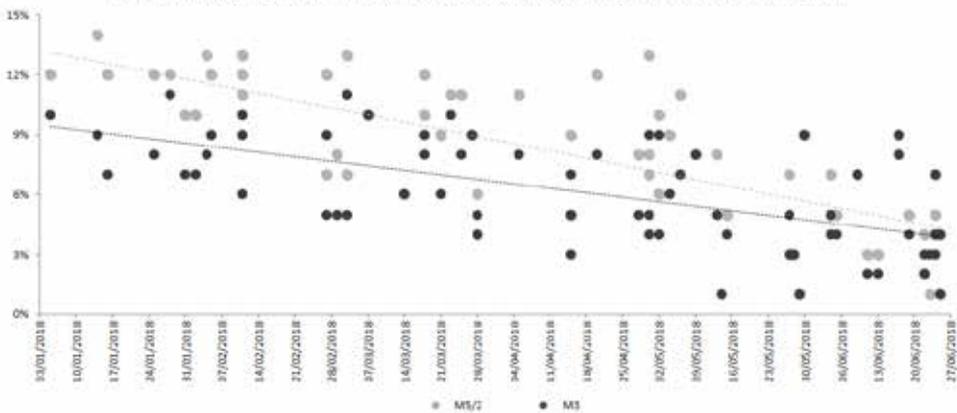
Gráfica 2. Estimaciones por encuesta en vivienda para la elección presidencial de México, 2018



FUENTE: Observatorio Electoral de CEDE e Instituto Nacional Electoral

Lo que se vio no fueron las estimaciones puntuales de cada encuesta dada a conocer durante el proceso electoral, sino la inexactitud de éstas respecto del resultado final (Gráfica 3), por lo que pudo advertirse que hubo una tendencia a la reducción del error y que, aunque no se llegó a cerrar la diferencia con el resultado, asintóticamente las encuestas tendieron a reflejar la distribución final.

Gráfica 3. Inexactitud en las estimaciones por encuesta en vivienda para la elección presidencial 2018



FUENTE: Cálculos propios con base en Observatorio Electoral del CEDE e Instituto Nacional Electoral.

Resulta pertinente regresar a otros estudios y dejar de lado, por inútiles, una posible recuperación y valoración de los miles de ejercicios sin sentido, método, calidad, precisión o exactitud

esperable que circularon. Simplemente hay que decir que muchas veces su difusión no estuvo exenta de un interés de los particulares involucrados en su divulgación por hacer propaganda respecto de uno u otro de los contendientes. En muchos casos, tan sólo fueron producto de la curiosidad de personas carentes de la debida calificación para llevar adelante una medición científica. Sin embargo, sí vale la pena mencionar un par de casos que mantuvo presencia y conservó cierta credibilidad, aun cerca del término del proceso, a pesar incluso de sus resultados. Nos referimos a ejercicios de medición mediante encuestas por redes sociales, particularmente Facebook, y a las encuestas automatizadas por vía telefónica.

Las encuestas realizadas por México Elige con el objetivo expreso de “medir las preferencias rumbo a la elección a Presidente de México” tomaron como marco muestral a las “personas de 18 años en adelante que radican en México y sean usuarios de la red social Facebook”. La selección ciertamente no representaba al conjunto de los electores y era un factor de sesgo potencial que causaba alejamiento entre estimaciones y resultados. A esto se sumó que no existía claridad para el público respecto del procedimiento específico de selección de entrevistados y no se garantizaba que los datos no fueran producto de una auto-selección de respondientes, ni que existieran mecanismos que permitieran controlar razones de aceptación o rechazo a los varios miles de entrevistas levantadas en cada toma de información reportada por esta empresa.¹⁷

Al final, el claro distanciamiento que hubo entre las estimaciones obtenidas por esa serie de encuestas y el resultado de la elección presidencial llevó a los responsables de esos estudios a reconocer que “no fueron suficientemente acertadas”, por lo que su “método innovador necesita ajustes”.¹⁸ Con ello se confirmó que lo deseado era realmente disponer de datos que dieran en la diana, sin reparar en la pertinencia científica de los métodos adoptados.

Fue claro que los resultados de estos ejercicios confirmaban lo inadecuado de realizar mediciones mediante la compilación de respuestas de usuarios de redes sociales, por más amplia cobertura que estas hubieran tenido, lo que tan sólo reflejó la incompatibilidad entre la población observada por esos estudios y el universo de los electores.

Massive Caller realizó investigaciones con el objetivo expreso de “generar información estadísticamente válida sobre las tendencias electorales” mediante entrevistas robotizadas que fueron contestadas directamente por informantes seleccionados a partir de un listado de números telefónicos fijos y móviles. Con este procedimiento se intentó alcanzar al grueso de los electores. Conforme sus propios reportes a la autoridad administrativa electoral nacional,¹⁹ la tasa de respuesta observada en esos estudios rondó el 10%, con un rechazo en seis de cada siete contactos efectivos, sin que la empresa hubiera reportado el haber realizado los ejercicios estadísticamente pertinentes para ponderar los casos observados respecto a parámetros demográficos externos, ni compensado las tasas de respuesta diferenciadas,²⁰ lo que pudo haber propiciado sesgos en las estimaciones, dado el método de entrevista adoptado. En todo caso, sus datos nacionales, que no fueron cabalmente de dominio público, durante la mayor parte de la campaña (hasta los últimos

17 Instituto Nacional Electoral (INE), *op. cit.*

18 Federico Arreola, Sergio Zaragoza y Aldo Campuzano, “La encuesta en Facebook de México Elige no fue la más acertada; la mejor fue la de *Reforma*; después, la de *El Financiero*”, *SDP Noticias*, Ciudad de México, México, 1 de julio, 2018.

19 Instituto Nacional Electoral (INE), *op. cit.*

20 *Idem.*

días) distaron mucho del resultado y —aun con un inusitado cambio final— no alcanzaron la exactitud de muchas de las mediciones en vivienda.

Esta propuesta enfrentó, además, cuestionamientos relacionados con la neutralidad de sus estudios, al delatarse que había un contrato millonario con uno de los contendientes para la elección presidencial.²¹ En el mismo sentido, y cerca del cierre de campaña, se hizo una denuncia, avalada por el INE, referente a que Massive Caller llevaba a cabo encuestas de empuje (*push polls*) en contra de otros de los contendientes,²² lo que contravenía las normas internacionales relativas a la operación de las empresas demoscópicas, violaba regulaciones electorales del país, y constituía una acción contraria a la debida ética de investigación.

Resulta difícil saber con claridad si los fallos de esas encuestas por vía telefónica, hechas mediante procedimientos automatizados, se debieron a los efectos probados en múltiples ocasiones del método mismo de aproximación, a inadecuaciones entre la población observada y el universo de electores, a desequilibrios sistemáticos en la estructura socio-demográfica de los respondientes respecto del colectivo que se pretendía observar, a patrones sesgados de rechazo, a falseamientos en las respuestas, o a muchos otros factores posibles.

Pero ello no necesariamente descarta el empleo de esos métodos para la obtención futura de información fidedigna sobre preferencias del electorado. Simplemente reflejó que las estimaciones realizadas en esa ocasión por la empresa mencionada no tuvieron los controles requeridos ni atendieron a la pulcritud debida que garantizara la producción de estimaciones con la exactitud reclamada.

Los errores de medición

Respecto de las llamadas “encuestas finales”, que por tradición son aquellas publicadas durante las tres semanas previas a la elección. De las encuestas serias, realizadas en vivienda, hubo trece ejercicios que cumplieron con las condiciones para ser incluidos en el recuento. Doce de ellos fueron encuestas convencionales, con muestras de entre mil y menos de mil quinientos casos, y una poseía una muestra mucho mayor (de decenas de miles de casos), con expansiones para la estimación simultánea de elecciones de ejecutivos estatales.

21 Verificado 2018, “Anaya sí contrató a la encuestadora Massive Caller”, *Milenio Diario*, Ciudad de México, México, 2 de mayo, 2018. Disponible en <<http://www.milenio.com/elecciones-mexico-2018/anaya-si-contrato-a-la-encuestadora-massive-caller>>, [fecha de consulta: 16 de julio, 2018].

22 Notimex, “INE ordena a empresas dejar de hacer llamadas telefónicas contra López Obrador”, *El Economista*, Ciudad de México, México, 26 de junio de 2018. Disponible en <<https://www.eleconomista.com.mx/politica/INE-ordena-a-empresas-dejar-de-hacer-llamadas-telefonicas-contra-Lopez-Obrador-20180626-0071.html>>, [fecha de consulta: 16 de julio, 2018].

Tabla 1. Estimaciones finales por encuesta en vivienda, publicadas antes de la elección presidencial de México, 2018

| Encuestadora | Fecha de cierre | Casos en muestra | Estimación reportada | | | | MV* | N |
|------------------|-----------------|------------------|----------------------|------|------|-----|------|-----|
| | | | RAC | JAM | AMLO | JRC | | |
| ARCOP | 22/06/2018 | 1000 | 28.0 | 18.0 | 52.0 | 2.0 | 24.0 | 2.6 |
| Berumen | 11/06/2018 | 13317 | 25.8 | 16.7 | 54.7 | 2.8 | 28.9 | 2.5 |
| BGC | 25/06/2018 | 1000 | 26.0 | 21.0 | 49.0 | 4.0 | 23.0 | 2.8 |
| Consulta | 19/06/2018 | 1000 | 25.5 | 22.5 | 48.1 | 3.9 | 22.6 | 2.9 |
| Conteo | 17/06/2018 | 1100 | 26.7 | 31.1 | 40.0 | 2.2 | 13.3 | 3.0 |
| Demotecnia | 13/06/2018 | 1200 | 25.0 | 19.0 | 50.0 | 5.0 | 25.0 | 2.8 |
| El Financiero | 23/06/2018 | 1420 | 21.0 | 22.0 | 54.0 | 3.0 | 33.0 | 2.6 |
| Eukoll | 22/06/2018 | 1600 | 21.9 | 19.2 | 51.7 | 7.2 | 29.8 | 2.8 |
| GEA-ISA | 17/06/2018 | 1070 | 28.4 | 26.4 | 43.6 | 1.6 | 15.2 | 2.9 |
| Parametría | 25/06/2018 | 1000 | 22.0 | 18.0 | 53.0 | 7.0 | 31.0 | 2.7 |
| Reforma | 24/06/2018 | 1200 | 27.0 | 19.0 | 51.0 | 3.0 | 24.0 | 2.7 |
| Suasor | 24/06/2018 | 1000 | 26.1 | 28.4 | 43.2 | 2.3 | 17.1 | 3.0 |
| Varela | 24/06/2018 | 1080 | 28.1 | 18.8 | 49.2 | 3.9 | 21.1 | 2.8 |
| Promedio | | | 25.5 | 21.5 | 49.2 | 3.7 | 23.7 | 2.8 |
| Resultado | | | 22.9 | 16.9 | 54.8 | 5.4 | 31.9 | 2.6 |

* El margen de victoria es respecto al resultado electoral, no al ordenamiento estimado en la encuesta.
Fuente: cálculos propios con base en *Observatorio Electoral* del CEDE e Instituto Nacional Electoral.

En promedio, esas trece encuestas mostraban una contienda con un voto elevadamente concentrado; ubicaban, a lo sumo, un número efectivo de contendientes de tres, con un López Obrador que llegaba a una intención de voto apenas por debajo de la mitad de los sufragios potenciales, y casi 24 puntos encima de su más cercano perseguidor. Diez de las encuestas coincidían en detectar que el segundo lugar le pertenecía a Ricardo Anaya, mientras que tres mostraban a José Antonio Meade en este segundo lugar. Esa fue la primera fuente de controversia previa y de error posterior (Tablas 1 y 2).

Si el orden de los contendientes fue la primera fuente de equivocaciones, una segunda causa de error fue la distancia entre los candidatos. En promedio, las encuestas mostraban una brecha casi ocho puntos inferior a la real, por lo que el estimador de inexactitud $M5/2$ fue de 4.3. Sistemáticamente, aunque en diferente magnitud, todas las encuestas subestimaron el voto por López Obrador, y sobreestimaron el voto por Meade, lo que provocó un inusitado error medio en la medición de estos dos candidatos respecto de lo observado a lo largo de este siglo, como se verá más adelante en este ensayo.

Tabla 2. Inexactitud de las estimaciones finales por encuesta en vivienda para la elección presidencial de México, 2018

| Encuestadora | Inexactitud en la estimación * | | | | Orden | M5/2 ^o | M3 | E |
|---------------|--------------------------------|------|-------|------|-------|-------------------|-----|-----|
| | RAC | JAM | AMLO | JRC | | | | |
| ARCOP | 5.1 | 1.1 | -2.8 | -3.4 | 1 | 4.0 | 3.1 | 3.2 |
| Berumen | 2.9 | -0.2 | -0.1 | -2.6 | 1 | 1.5 | 1.5 | 5.9 |
| BGC | 3.1 | 4.1 | -5.8 | -1.4 | 1 | 4.5 | 3.6 | 2.7 |
| Consulta | 2.6 | 5.6 | -6.7 | -1.5 | 1 | 4.7 | 4.1 | 2.8 |
| Conteo | 3.8 | 14.2 | -14.8 | -3.2 | 0 | 9.3 | 9.0 | 7.2 |
| Demotecnia | 2.1 | 2.1 | -4.8 | -0.4 | 1 | 3.5 | 2.4 | 1.4 |
| El Financiero | -1.9 | 5.1 | -0.8 | -2.4 | 0 | 1.4 | 2.6 | 2.6 |
| Enkoll | -1.0 | 2.3 | -3.1 | 1.8 | 1 | 2.1 | 2.1 | 1.9 |
| GEA-ISA | 5.5 | 9.5 | -11.2 | -3.8 | 1 | 8.4 | 7.5 | 6.3 |
| Parametría | -0.9 | 1.1 | -1.8 | 1.6 | 1 | 1.4 | 1.4 | 1.3 |
| Reforma | 4.1 | 2.1 | -3.8 | -2.4 | 1 | 4.0 | 3.1 | 2.9 |
| Suasor | 3.2 | 11.5 | -11.6 | -3.1 | 0 | 7.4 | 7.4 | 5.8 |
| Varela | 5.2 | 1.9 | -5.6 | -1.5 | 1 | 5.4 | 3.6 | 2.7 |
| Promedio | 2.7 | 4.7 | -5.7 | -1.7 | 1 | 4.4 | 3.9 | 3.6 |
| Resultado | 22.9 | 16.9 | 54.8 | 5.4 | | | | |

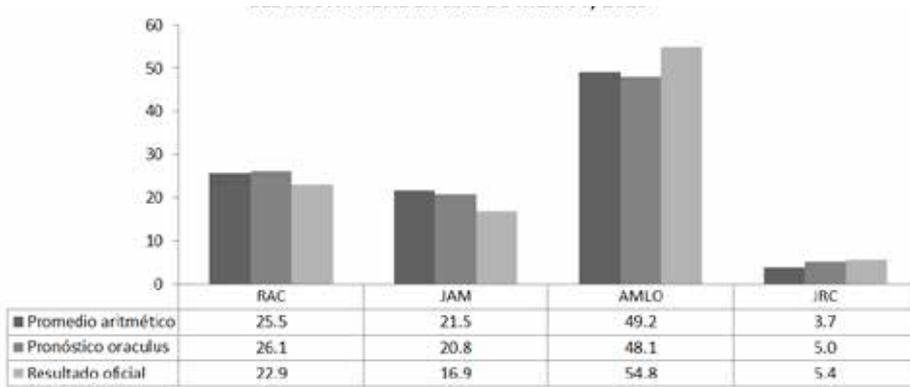
* La inexactitud en la estimación equivale a la diferencia en puntos porcentuales entre estimación y resultado. ^o La inexactitud en la estimación de la ventaja es respecto al resultado electoral, no el de la encuesta.

Fuente: cálculos propios con base en *Observatorio Electoral* del CEDE e Instituto Nacional Electoral.

Cuando se consideró el error medio para los cuatro contendientes, la situación no mejoró mucho: *M3* se ubicó en 3.9, y aun las mejores estimaciones mostraron un error medio cercano a punto y medio, aunque las hubo con hasta siete o más puntos de error promedio.

Cuando lo que se observó fue el error estandarizado (*E*), que tomó en cuenta el tamaño muestral y la proporción relativa por cada contendiente, las mediciones más desviadas tendieron a compensar en parte su error, disminuyendo así la distancia con otras mediciones. Empero, en lo referente a la encuesta con mayor número de casos —cuyo costo superó al de series completas de encuestas difundidas en el periodo—, la disposición de una muestra significativamente mayor la afectó de manera importante, pues su mayor tamaño no le otorgó una exactitud suficiente para evitar que se convirtiera en una de las menos exactas de las encuestas finales, según ese cálculo de error estandarizado.

Gráfica 4. Estimaciones finales por encuesta en vivienda elección presidencial de México, 2018



FUENTE: Observatorio Electoral del CEDE, oraculus.mx e Instituto Nacional Electoral.

Se pudo haber pensado que tal vez en esa ocasión la revisión particular de las encuestas no fue tan afortunada, pero que los ejercicios de las agregadoras cumplieron mejor con los objetivos de exactitud pretendidos, sobre todo en aquellos casos en que se partió de modelos complejos de estimación que tomaron en cuenta comportamientos históricos y no simples medias.

Sin embargo, cuando se revisó la exactitud que esos nuevos actores lograron, se descubre que sus estimaciones puntuales finales estuvieron apenas un poco más cercanas al resultado que el promedio aritmético de las encuestas, y más alejadas que la mayoría de las encuestas en particular. Esto fue lo que ocurrió, por ejemplo, con el pronóstico de Oraculus, una de las agregadoras más serias por su metodología, que observó un *M5/2* de 5.0, y un *M3* de 3.6 (véase Gráfica 4), cuando se esperaba que se presentara una competitividad superior a la real e incluso a la estimada conforme al promedio de las encuestas.

Se debe conceder, sin embargo, que Oraculus (así como otras agregadoras) no estuvo apostando por una estimación puntual, sino por un intervalo. Pero incluso así, cuando se cotejó el intervalo pronosticado con el resultado, se detectó que en dos de cuatro contendientes el resultado se hallaba claramente fuera del rango esperado: López Obrador por debajo, Meade por encima, y Anaya apenas fuera del límite inferior (Gráfica 5). Fue obvio que ello no había sido sólo responsabilidad de la agregadora, sino secuela de la inexactitud de las encuestas que tomó como base. También se consideró que este modelo pudo haber pecado de un exagerado optimismo al cerrar excesivamente sus horquillas al final, a pesar de la evidencia que mostraba que el mejor estimador del resultado de una elección presidencial no había sido regularmente la media de las estimaciones por encuesta.

Gráfica 5. Pronóstico de la agregadora oraculus.mx con base en los resultados de la encuesta en vivienda para la elección presidencial 2018



FUENTE: oraculus.mx e Instituto Nacional Electoral.

A continuación, se presenta una breve revisión de las encuestas publicadas luego de la elección, durante la misma jornada electoral y poco después del cierre de casillas. Seis encuestas de salida reportaron sus estimaciones de forma casi inmediata y entregaron tanto reportes como bases de datos a la autoridad electoral administrativa del país.

Tabla 3. Estimaciones por encuesta de salida publicadas al cierre de casillas sobre la elección presidencial de México, 2018

| Encuestadora | Estimación reportada | | | | MV * | N |
|------------------|----------------------|-----|------|-----|------|-----|
| | RAC | JAM | AMLO | JRC | | |
| BGC | 25 | 22 | 49 | 4 | 24 | 2.8 |
| Buendía & Laredo | 24 | 20 | 51 | 5 | 27 | 2.8 |
| Consulta | 25 | 24 | 47 | 4 | 22 | 2.9 |
| El Financiero | 27 | 18 | 49 | 6 | 22 | 2.9 |
| GEA-ISA | 27 | 21 | 49 | 3 | 22 | 2.8 |
| Parametría | 22 | 17 | 56 | 5 | 34 | 2.5 |
| Promedio | 25 | 20 | 50 | 5 | 25 | 2.8 |
| Resultado | 23 | 17 | 55 | 5 | 32 | 2.6 |

* El margen de victoria es respecto al resultado electoral, no al ordenamiento estimado en la encuesta.
Fuente: *Observatorio Electoral* del CDE e Instituto Nacional Electoral.

Aunque todas las encuestas acertaron en el orden en que terminaron los contendientes, en sus reportes originales siguieron mostrando los mismos síntomas de inexactitud detectados en las encuestas preelectorales: una mayor competitividad de la que realmente se registró, y un margen de victoria menor al que se produjo, como resultado de subestimación del ganador y de una tendencia a la sobreestimación del tercer lugar. También los errores medios de estos estudios, con un *M5/2* de 3.8 y un *M3* de 2.9, resultan particularmente elevados (Tablas 3 y 4).

Esta persistencia de los sesgos pareció indicarnos que factores exógenos pudieron haberlos propiciado, con lo que se descartó problemas de detección de votantes entre los electores o giros de última hora en preferencias, dado que ambos elementos se encontraban excluidos en mediciones simultáneas a la votación. A la luz de la distancia en el resultado, la relevancia de este hecho hubiera podido parecer menor, lo que se percibió de manera diferente al considerarse que cuatro de los seis ejercicios publicados reportaron un ganador que no alcanzaba la mayoría absoluta de los votos —cosa que sí ocurrió y que supuso un triunfo que le otorgó al vencedor una enorme legitimidad.

Tabla 4. Inexactitud de estimaciones por encuesta de salida publicadas al Cierre de casillas sobre la elección presidencial de México, 2018

| Encuestadora | Inexactitud en la estimación * | | | | M5/2 ^o | M3 |
|------------------|--------------------------------|-----|------|------|-------------------|-----|
| | RAC | JAM | AMLO | JRC | | |
| BGC | 2 | 5 | -6 | -1 | 4.0 | 3.5 |
| Buendía & Laredo | 1 | 3 | -4 | 0 | 2.5 | 2.0 |
| Consulta | 2 | 7 | -8 | -1 | 5.0 | 4.5 |
| El Financiero | 4 | 1 | -6 | 1 | 5.0 | 3.0 |
| GEA-ISA | 4 | 4 | -6 | -2 | 5.0 | 4.0 |
| Parametría | -1 | 0 | 1 | 0 | 1.0 | 0.5 |
| Promedio | 2.0 | 3.3 | -4.8 | -0.5 | 3.8 | 2.9 |
| Resultado | 23 | 17 | 55 | 5 | | |

La inexactitud en la estimación equivale a la diferencia en puntos porcentuales entre estimación y resultado. La inexactitud en la estimación de la ventaja es respecto al resultado electoral, no el de la encuesta.
Fuente: *Observatorio Electoral* del CEDE e Instituto Nacional Electoral.

Las dudas cartesianas

Una revisión más detallada que la mera comparación de las estimaciones agregadas, producto de las encuestas finales contra los resultados, nos llevó a advertir la potencial existencia de peculiaridades que debían tomarse en cuenta para una evaluación completa del fenómeno de las encuestas en la temporada. Una de ellas fue la posibilidad de que hubieran existido, en el curso de las últimas semanas, giros en sentido encontrado en las preferencias.

Al ver la Tabla 5, que comparaba series de estudios de nueve distintas encuestadoras, realizados después de cada uno de los debates entre candidatos a la presidencia de la República —pero sin tomar en cuenta la magnitud y sentido de los sesgos sistemáticos que pudieron haber afectado cada serie—, se descubrió que, mientras era clara la aproximación de estimaciones al resultado en las mediciones posteriores al segundo debate, en diversas mediciones, luego del tercer debate, se había registrado un giro a favor del candidato que terminó en tercer lugar y en contra del líder, lo que provocó que las mediciones finales se hubieran alejado más del resultado que las estimaciones inmediatas previas, posteriores al segundo debate.

Tabla 5. Inexactitud en las estimaciones por encuesta en vivienda para la elección presidencial 2018

| Encuestadora | Post-primer debate | | | Post-segundo debate | | | Post-tercer debate (final) | | |
|---------------|--------------------|-------|-----|---------------------|-------|-----|----------------------------|-------|-----|
| | Orden | M5/2* | M3 | Orden | M5/2* | M3 | Orden | M5/2* | M3 |
| BGC | 1 | 11.5 | 7.1 | 1 | 7.0 | 5.1 | 1 | 4.5 | 3.6 |
| Consulta | 1 | 7.7 | 4.9 | 1 | 5.1 | 4.2 | 1 | 4.7 | 4.1 |
| Conteo | 0 | 9.8 | 9.1 | 0 | 8.9 | 8.9 | 0 | 9.3 | 9.0 |
| El Financiero | 1 | 6.0 | 3.9 | 1 | 3.0 | 3.1 | 0 | 0.6 | 2.6 |
| GEA-ISA | 1 | 12.8 | 9.3 | 1 | 7.5 | 5.1 | 1 | 8.4 | 7.5 |
| Parametría | 1 | 7.9 | 4.9 | 1 | 0.9 | 0.6 | 1 | 0.4 | 1.4 |
| Reforma | 1 | 7.0 | 4.4 | 1 | 3.0 | 2.6 | 1 | 4.0 | 3.1 |
| Suasor | 0 | 8.3 | 7.8 | 0 | 7.5 | 7.4 | 0 | 7.4 | 7.4 |
| Varela | 1 | 8.8 | 5.6 | 1 | 5.4 | 3.7 | 1 | 5.4 | 3.6 |
| Promedio | 1 | 8.9 | 6.3 | 1 | 5.3 | 4.5 | p | 4.9 | 4.7 |

* La inexactitud en la estimación de la ventaja es respecto al resultado electoral, no el de la encuesta. Fuente: cálculos propios con base en *Observatorio Electoral* del CEDE e Instituto Nacional Electoral.

Eso mismo ocurrió claramente en tres de las nueve mediciones, dos de las cuales fueron de las que finalmente resultaron más exactas. Incluso pudo haber sido la causa de provocar en otra medición una equivocación en el orden de llegada de los contendientes, pues ello solamente resultó errado en esa medición final —a diferencia de otras series cuya inversión del orden fue sistemática. De hecho, y contrariamente a lo que podría decir la intuición y a lo esperado, en esta elección el paquete de las encuestas cuyas estimaciones resultaron más cercanas al resultado incluyó a algunas de las realizadas más de tres semanas antes de la elección, y no solamente a aquellas realizadas en los días previos al cierre de campañas.

Otro punto que debiera atenderse es que la exactitud de la medición en el agregado no necesariamente significó que se reprodujera fehacientemente la realidad. Si una misma casa encuestadora, Parametría, aportó las estimaciones más exactas, tanto en su medición preelectoral como en la de salida, cuando se revisaron las proporciones de respaldo al ganador, según

variables demográficas de ambos estudios, se descubrió que no existía coincidencia en sus distribuciones (Tabla 6). Por eso resultó difícilmente sostenible que ambas encuestas reflejaran correctamente, al mismo tiempo, el reparto de apoyos hacia el candidato ganador según género. De haber asumido como ciertos estos datos, se habría requerido una explicación (que se antoja compleja) de por qué existieron notables diferencias entre electores y votantes en sus patrones de preferencia por género, que, sin embargo, se compensaron mutuamente en el agregado. Esta duda se reforzó cuando se consideró que otras encuestas de salida no habían confirmado una brecha de género de la magnitud detectada por el estudio de marras, aun cuando sus datos hubieran sido ya ajustados al resultado.²³

Tabla 6. Votación por Lopez Obrador según la encuesta final y de salida de *Parametría* para la elección presidencial de México, 2018, conforme variables demográficas

| Demográficos | Encuesta final | Encuesta de salida |
|---------------|----------------|--------------------|
| Fecha | 25/06/2018 | 01/07/2018 |
| Población | Electores | Votantes |
| Casos | 1000 | 1321 |
| Total | 53% | 56% |
| Hombres | 55% | 65% |
| Mujeres | 51% | 49% |
| 18 a 25 años | 51% | 55% |
| 26 a 35 años | 55% | 63% |
| 36 a 45 años | 59% | 56% |
| 46 a 55 años | 59% | 56% |
| 56 y más años | 45% | 55% |

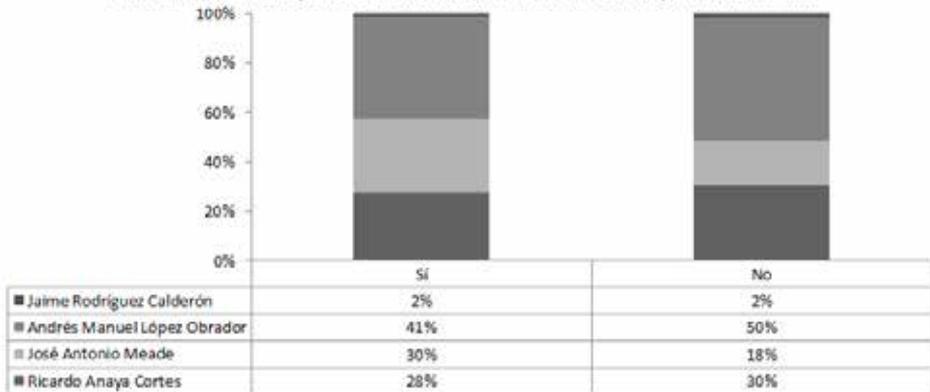
Fuente: base de datos en el *Observatorio Electoral* del CEDE y encuesta de salida de *Parametría*.

Al explorar potenciales causas de los sesgos en la mayoría de las estimaciones por encuesta, pudiera aportarse un elemento como posible generador de sesgos en las mediciones por encuesta para la elección presidencial de 2018: aunque la confianza en el proceso electoral y en los resultados que se darían fueron mayores hacia el cierre del proceso, todavía al final de las campañas los seguidores del candidato ganador resultaron menos propensos a manifestar abiertamente por quién votarían, que aquellos que respaldaban a otros candidatos. Ese hecho pudo propiciar —al menos en algunas mediciones y para algunas casas encuestadoras— un ocultamiento de las intenciones de los entrevistados que disminuyera tendencialmente la proporción favorable

23 Véase Francisco Abundis, "¿Quiénes eligieron a AMLO como presidente?", *Milenio Diario*, Ciudad de México, México, 10 de julio, 2018, y Alejandro Moreno, "Lucha contra la corrupción, el principal reto de López Obrador", *El Financiero*, Ciudad de México, México, 3 de julio, 2018.

a quien resultó triunfador (Gráfica 6). Eso finalmente explicó, además, por qué este sesgo se mantuvo todavía en las encuestas de salida.

Gráfica 6. Intención declarada de voto para la presidencia de la República según si el entrevistado siente que puede decir libremente en público por quién piensa votar



FUENTE: GEA-ISA, Encuesta Nacional (15-17 de junio de 2018, 1070 casos)

La exactitud comparada

¿Qué tan buenos o malos fueron los saldos de las encuestas en la elección presidencial de 2018, en términos de la exactitud lograda? Ello pudo medirse al cotejarlos con los indicadores de exactitud correspondientes a las otras elecciones presidenciales ocurridas este siglo en el país, y al tomarse en consideración, exclusivamente, encuestas hechas públicas durante las tres semanas inmediatamente anteriores a cada elección (Tabla 7).²⁴

24 En archivo electrónico se pone a disposición la lista y datos de los estudios considerados en este análisis. Cfr. Ricardo de la Peña, "Encuestas en vivienda para las elecciones por la presidencia de la República en México 2000-2018" [base de datos], *Ricardo de la Peña*. Disponible en <<http://www.ricartur.com/databases/ EncuestasPresidencialesXXXI.txt>>, [fecha de consulta: 14 de agosto, 2018].

Tabla 7. Inexactitud de las estimaciones finales por encuesta en vivienda para las elecciones presidenciales en México (2000-2018)

| | | | | | | | | | |
|------|---------------------------|-----|------|-------|------------------|----------------|------|-----|-----|
| 2000 | Inexactitud por candidato | | | | Ganador correcto | Orden correcto | M5/2 | M3 | E |
| | VFQ | FLO | ccs | Otros | | | | | |
| | -3.6 | 2.3 | 0.9 | 0.4 | 46% | 46% | 3.2 | 2.6 | 2.4 |
| 2006 | Inexactitud por candidato | | | | Ganador correcto | Orden correcto | M5/2 | M3 | E |
| | FCH | RMP | AMLO | Otros | | | | | |
| | -2.5 | 3.6 | -1.6 | 0.5 | 46% | 46% | 1.4 | 2.2 | 1.9 |
| 2012 | Inexactitud por candidato | | | | Ganador correcto | Orden correcto | M5/2 | M3 | E |
| | JVM | EPN | AMLO | GQT | | | | | |
| | -1.0 | 3.4 | -3.2 | 0.9 | 100% | 92% | 3.6 | 2.6 | 2.1 |
| 2018 | Inexactitud por candidato | | | | Ganador correcto | Orden correcto | M5/2 | M3 | E |
| | RAC | JAM | AMLO | JRC | | | | | |
| | 2.7 | 4.7 | -5.7 | -1.7 | 100% | 77% | 4.2 | 4.0 | 3.6 |

Fuente: *El papel de las encuestas en las elecciones federales. Memorias del Taller Sumiya 2000*, Instituto Federal Electoral - Asociación Mexicana de Agencias de Investigación de Mercado y Opinión Pública - Colegio Nacional de Actuarios. México, 2001. *Memorias del seminario: encuestas y elecciones 2006*, Instituto Federal Electoral - Asociación Mexicana de Agencias de Investigación de Mercado y Opinión Pública - Asociación Mundial de Investigadores de la Opinión Pública - Consejo de Investigadores de la Opinión Pública. México, 2010. *Séptimo Informe que presenta la Secretaría Ejecutiva al Consejo General del Instituto Federal Electoral respecto del cumplimiento del Acuerdo C G411/2011, por el que se establecen los lineamientos así como los criterios generales de carácter científico que deberán observar las personas físicas y morales que pretendan ordenar, realizar y/o publicar encuestas por muestreo, encuestas de salida y/o conteos rápidos durante el Proceso Electoral 2011-2012*, Instituto Federal Electoral. México, 26 de julio de 2012, Observatorio Electoral del CEDE e Instituto Nacional Electoral.

Conforme a esos datos, en 2000 y 2006 las encuestas presentaron más de la mitad de las veces equivocaciones en la detección del ganador y, por consiguiente, en el orden de los contendientes. Mientras, en 2012 y 2018 no hubo encuestas que se equivocaran en el ganador, pero en 2018 erró el orden de llegada 23% de las encuestas, el triple que lo observado seis años antes —lo que, en estos términos, hizo que en 2012 hubiera ocurrido la mejor previsión en el presente siglo.

De acuerdo con los datos que se presentaron en este ensayo, 2018 fue la ocasión en que las encuestas finales para una elección presidencial tuvieron mayores niveles de inexactitud en este siglo: un M5/2 de 4.3, contra 3.6 observado seis años antes, e incluso frente a menores niveles de ocasiones previas; un M3* de 3.9 (recalculado para componentes mayores de diez puntos), contra valores de 2.6 en dos ocasiones y de 2.2 en 2006, y un E de 3.6, contra una media de 2.1 en pasadas elecciones.

Debe mencionarse que, por una parte, la inexactitud observada para la estimación particular del ganador, desviada en 5.7 puntos promedio, fue superior a cualquiera observada antes. Por

otra parte, la valoración de quien ocupó el tercer lugar, de 4.7, fue en general mayor a cualquiera previa, y en particular, la valoración del partido que encabezaba la coalición que lo postuló, el cual siempre ha sido sobreestimado.

Hacia un índice de rendimiento

Cuando se desagregaron los resultados por encuestadora para las cuatro elecciones presidenciales ocurridas en el presente siglo (Tabla 8), se vio que existía una elevada variabilidad de quienes concurren a aportar estimaciones finales. Más de la mitad de las responsables participó en una única ocasión, bien porque las casas no tuvieron patrocinio o interés en hacerlo, bien porque fueron proyectos coyunturales, surgidos al calor de una contienda. Sólo cuatro instancias han participado en todas las ocasiones, dos de ellas con los mismos responsables (aunque hay un responsable que ha participado en todas desde medios diferentes). Otras tres marcas que concurren recientemente lo hicieron dos veces antes, mientras que seis casas tuvieron apenas su primera estimación en 2018. Únicamente una encuestadora, GEA-ISA, ha reportado correctamente al ganador y el orden de contendientes en todas las elecciones.

Ante esa situación, fue difícil realizar un análisis comparativo, toda vez que no se dispuso de un procedimiento que permitiera agregar en un único medidor, y comparar sin ambigüedad, el rendimiento de las encuestadoras. Este último debiera generarse, de ser posible, tomando en cuenta diversos indicadores y estimadores de exactitud. Aunque si se pretendiera alcanzar al respecto un consenso básico entre los especialistas sería necesario un arduo esfuerzo de concertación, sí fue posible proponer un índice para medir el rendimiento de las estimaciones producto de encuestas, que respondiera a lo intuitivo y a la vez permitiera reducir los criterios disponibles sobre (in)exactitud a un único indicador que hiciera viable tanto la comparación entre estudios como un ordenamiento unívoco.

Tabla 8. Inexactitud promedio por encuestadora en estimaciones finales en vivienda para las elecciones presidenciales en México (2000-2018)

| Encuestadora | Encuestas finales | Ganador correcto | Orden correcto | M5/2 | M3* | E |
|-------------------|-------------------|------------------|----------------|------|-----|-----|
| AC Nielsen | 1 | 0% | 0% | 4.8 | 2.8 | 2.8 |
| Alduncin | 2 | 100% | 100% | 0.8 | 2.9 | 2.9 |
| ARC OP | 4 | 100% | 75% | 2.2 | 1.9 | 1.7 |
| Berumen | 1 | 100% | 100% | 1.5 | 1.5 | 5.9 |
| BGC | 3 | 100% | 100% | 3.1 | 2.9 | 2.2 |
| Buendia & Laredo | 1 | 100% | 100% | 5.2 | 3.1 | 2.9 |
| CEO | 2 | 0% | 0% | 3.3 | 2.6 | 2.8 |
| Con Estadística | 1 | 100% | 100% | 5.5 | 3.6 | 2.8 |
| Consulta Mitofsky | 3 | 67% | 67% | 3.5 | 3.0 | 2.2 |

| | | | | | | |
|-----------------------|-----------|------------|------------|------------|------------|------------|
| Consultores y MP | 2 | 50% | 50% | 3.0 | 2.2 | 1.9 |
| Conteo | 1 | 100% | 0% | 9.3 | 9.0 | 7.2 |
| Covarrubias | 1 | 100% | 100% | 2.1 | 1.3 | 1.2 |
| Democracy Watch | 1 | 100% | 100% | 0.8 | 1.7 | 1.6 |
| Demotecnia | 4 | 75% | 75% | 2.0 | 2.4 | 2.0 |
| El Financiero | 1 | 100% | 0% | 0.6 | 2.6 | 2.6 |
| El Universal | 1 | 0% | 0% | 1.3 | 1.6 | 1.3 |
| Eukoll | 1 | 100% | 100% | 1.1 | 2.1 | 1.9 |
| GAUSSC | 1 | 100% | 100% | 2.3 | 1.9 | 1.5 |
| GEA ISA | 4 | 100% | 100% | 4.3 | 3.4 | 3.0 |
| IXDEMERC | 2 | 50% | 50% | 3.8 | 4.0 | 3.5 |
| IPSOS-BIMSA | 1 | 100% | 100% | 3.6 | 2.2 | 1.2 |
| MERCAEI | 1 | 100% | 100% | 0.3 | 1.1 | 0.9 |
| MUXD | 1 | 0% | 0% | 3.9 | 4.8 | 1.9 |
| Parametría | 3 | 67% | 67% | 2.3 | 2.3 | 1.7 |
| Pearson | 1 | 0% | 0% | 5.6 | 3.6 | 3.1 |
| Reforma (actual) | 1 | 100% | 100% | 4.0 | 3.1 | 2.9 |
| Reforma (previo) | 3 | 33% | 33% | 2.6 | 1.9 | 2.3 |
| Suasor | 1 | 100% | 0% | 7.4 | 7.4 | 5.8 |
| Varela | 1 | 100% | 100% | 5.4 | 3.6 | 2.7 |
| Zogby | 2 | 50% | 50% | 3.2 | 3.5 | 2.7 |
| Reforma (total) | 4 | 50% | 50% | 2.9 | 2.2 | 2.4 |
| Reforma/El Financiero | 4 | 50% | 25% | 2.1 | 2.1 | 2.4 |
| TOTAL | 52 | 73% | 65% | 3.1 | 2.9 | 2.5 |

Fuente: *El papel de las encuestas en las elecciones federales. Memorias del Taller Sumiya 2000*, Instituto Federal Electoral - Asociación Mexicana de Agencias de Investigación de Mercado y Opinión Pública - Colegio Nacional de Actuarios, México, 2001. *Memorias del seminario: encuestas y elecciones 2006*, Instituto Federal Electoral - Asociación Mexicana de Agencias de Investigación de Mercado y Opinión Pública - Asociación Mundial de Investigadores de la Opinión Pública - Consejo de Investigadores de la Opinión Pública, México, 2010. *Séptimo Informe que presenta la Secretaría Ejecutiva al Consejo General del Instituto Federal Electoral respecto del cumplimiento del Acuerdo C G411/2011 por el que se establecen los lineamientos así como los criterios generales de carácter científico que deberán observar las personas físicas y morales que pretendan ordenar, realizar y/o publicar encuestas por muestreo, encuestas de salida y/o conteos rápidos durante el Proceso Electoral 2011-2012*, Instituto Federal Electoral, México, 26 de julio de 2012. *Observatorio Electoral del CEDE e Instituto Nacional Electoral*.

Dada la diversidad de dimensiones que simultáneamente se ha demandado a una encuesta para calificar su exactitud, sería deseable construir un indicador sintético que permitiera estudiar las variaciones en el tiempo y el espacio de la eficacia que diversas encuestas presentaron para alcanzar su objetivo de aproximarse al resultado del evento electoral futuro. El indicador debía capturar por medio de un único número el desempeño relativo de la encuesta. Igualmente, de-

bía poder medir la contribución de ésta a la eliminación de la incertidumbre en torno al futuro resultado de la elección; de este modo, podía dar cuenta de la información anticipatoria que el estudio proporcionó.

El Índice de Rendimiento (IR) que se propuso, por encuestadora o para un conjunto de encuestas, consideró cuatro criterios y otorgó un máximo de diez puntos en cada uno de ellos. Aunque era relevante, se excluyó del cálculo de este índice el error estándar observado, debido a que su empleo no fue generalizado y a que el IR se aproximaba más a una medición de la eficiencia, al uso racional de medios para el logro de los fines predictivos de los estudios. Tampoco se consideró el número de veces que las encuestas fueron publicadas. Dos de los criterios atendieron al orden en que se estimó a los contendientes: uno se puntúa con diez o cero, dependiendo si se reportó como líder al ganador de una contienda; otro, diez o cero, si se reportó el orden correcto de los contendientes. Otros dos criterios corresponderían a los medidores tradicionales de exactitud: uno, la diferencia de diez menos el *M5/2* observado; el segundo, la diferencia de diez menos el *M3** obtenido. Ambos pueden detentar valores negativos simultáneamente.

Cuando se observó este índice en el caso de cada elección, se descubrió que la dualidad en la lectura de lo ocurrido en 2018 era real: altos valores en algunos criterios y reducidos en otros. Empero, el IR se ubicó por encima del promedio observado en el presente siglo (Tabla 9), aunque por debajo del rendimiento logrado en 2012, sin duda el mejor registrado hasta el momento del análisis entonces (aunque, paradójicamente fuera muy cuestionado). A pesar de que algunas de las peores mediciones se reportaron en 2018, también se registraron varias de las mejores, lo que ubicó indudablemente la más reciente experiencia indudablemente por arriba de la media histórica, debido a que no se equivocó al señalar al ganador, como sí ocurrió en muchas de las encuestas publicadas en 2000 y en 2006.

Tabla 9. Índice de rendimiento en las estimaciones finales en vivienda en las elecciones presidenciales en México (2000-2018)

| Elección | Número de encuestas | Ganador correcto | Orden correcto | M5/2 | M3* | IR |
|----------|---------------------|------------------|----------------|------|-----|-----|
| 2000 | 13 | 4.6 | 4.6 | 6.8 | 7.4 | 5.9 |
| 2006 | 13 | 4.6 | 4.6 | 8.6 | 7.8 | 6.4 |
| 2012 | 13 | 10.0 | 9.2 | 6.4 | 7.4 | 8.3 |
| 2018 | 13 | 10.0 | 7.7 | 5.8 | 6.1 | 7.4 |
| TOTAL | 52 | 7.3 | 6.5 | 6.9 | 7.1 | 7.0 |

Fuente: cálculos propios con base en la Tabla 7 de este documento.

El promedio de los valores de los diferentes criterios dio el índice del rendimiento, que se elevó hasta diez cuando la estimación fue perfecta; es decir, cuando correspondió puntualmente con el resultado. Esta fue una forma pertinente y adecuada de aproximarse al rendimiento de las encuestadoras, mejor que la observación de un único evento, cuya exactitud dependía de lo fortuito. Aunado a lo anterior, se debió considerar que, al juzgar el desempeño logrado en varios

ejercicios, se posibilitaba que sesgos de origen estocástico tendieran a compensarse y pudieran quedar como remanentes las diferencias con los resultados que, si eran regulares, permitieran estimar un rendimiento secular de la encuestadora (Tabla 10).

Tabla 10. Índice de rendimiento por encuestadora en las estimaciones finales en vivienda para las elecciones presidenciales en México (2000-2018)

| Encuestadora | Encuestas finales | Elección presidencial | | | | Indicadores particulares | | | | IR |
|-------------------|-------------------|-----------------------|------|------|------|--------------------------|----------------|------|-----|-----|
| | | 2000 | 2006 | 2012 | 2018 | Ganador correcto | Orden correcto | M5/2 | M3* | |
| AC Nielsen | 1 | 3.1 | | | | 0.0 | 0.0 | 5.2 | 7.3 | 3.1 |
| Alduncin | 2 | 9.3 | 8.9 | | | 10.0 | 10.0 | 9.2 | 7.1 | 9.1 |
| ARCOP | 4 | 9.6 | 9.4 | 6.2 | 8.2 | 10.0 | 7.5 | 7.8 | 8.1 | 8.3 |
| Berumen | 1 | | | | 9.3 | 10.0 | 10.0 | 8.5 | 8.5 | 9.3 |
| BGC | 3 | | 9.4 | 8.2 | 8.0 | 10.0 | 10.0 | 6.9 | 7.1 | 8.5 |
| Buendía & Laredo | 1 | | | 7.9 | | 10.0 | 10.0 | 4.9 | 6.9 | 7.9 |
| CEO | 2 | 2.9 | 4.2 | | | 0.0 | 0.0 | 6.7 | 7.5 | 3.5 |
| Con Estadística | 1 | | | 7.8 | | 10.0 | 10.0 | 4.6 | 6.5 | 7.8 |
| Consulta Mitofsky | 3 | | 4.0 | 8.3 | 7.8 | 6.7 | 6.7 | 6.5 | 7.1 | 6.7 |
| Consultores y MP | 2 | 3.0 | 9.4 | | | 5.0 | 5.0 | 7.0 | 7.8 | 6.2 |
| Conteo | 1 | | | | 2.9 | 10.0 | 0.0 | 0.7 | 1.0 | 2.9 |
| Covarrubias | 1 | | | 9.2 | | 10.0 | 10.0 | 7.9 | 8.8 | 9.2 |
| Democracy Watch | 1 | 9.4 | | | | 10.0 | 10.0 | 9.2 | 8.3 | 9.4 |
| Demotecnia | 4 | 9.1 | 3.4 | 9.6 | 8.6 | 7.5 | 7.5 | 8.0 | 7.6 | 7.7 |
| El Financiero | 1 | | | | 6.7 | 10.0 | 0.0 | 9.5 | 7.5 | 6.7 |
| El Universal | 1 | | 4.3 | | | 0.0 | 0.0 | 8.7 | 8.4 | 4.3 |
| Enkoll | 1 | | | | 9.2 | 10.0 | 10.0 | 9.0 | 8.0 | 9.2 |
| GAUSSC | 1 | 9.0 | | | | 10.0 | 10.0 | 7.8 | 8.2 | 9.0 |
| GEA-ISA | 4 | 9.1 | 9.7 | 7.6 | 6.0 | 10.0 | 10.0 | 5.8 | 6.7 | 8.1 |
| INDEMERC | 2 | | 3.9 | 7.3 | | 5.0 | 5.0 | 6.2 | 6.0 | 5.6 |
| IPSOS-BIMSA | 1 | | | 8.6 | | 10.0 | 10.0 | 6.4 | 7.9 | 8.6 |

| | | | | | | | | | | |
|-----------------------|----|-----|-----|-----|-----|------|------|-----|-----|-----|
| MERCAEI | 1 | | | 9.7 | | 10.0 | 10.0 | 9.7 | 9.0 | 9.7 |
| MUND | 1 | 2.8 | | | | 0.0 | 0.0 | 6.2 | 5.2 | 2.8 |
| Parametría | 3 | | 3.8 | 8.2 | 9.6 | 6.7 | 6.7 | 7.7 | 7.7 | 7.2 |
| Pearson | 1 | 2.7 | | | | 0.0 | 0.0 | 4.4 | 6.5 | 2.7 |
| Reforma (actual) | 1 | | | | 8.2 | 10.0 | 10.0 | 6.1 | 6.9 | 8.2 |
| Reforma (previo) | 3 | 3.1 | 4.3 | 9.2 | | 3.3 | 3.3 | 7.4 | 8.1 | 5.5 |
| Suasor | 1 | | | | 3.8 | 10.0 | 0.0 | 2.6 | 2.7 | 3.8 |
| Varela | 1 | | | | 7.8 | 10.0 | 10.0 | 4.6 | 6.5 | 7.8 |
| Zogby | 2 | 3.0 | 8.7 | | | 5.0 | 5.0 | 6.8 | 6.5 | 5.8 |
| Reforma (total) | 4 | 3.1 | 4.3 | 9.2 | 8.2 | 5.6 | 5.6 | 6.7 | 6.4 | 6.1 |
| Reforma/El Financiero | 4 | 3.1 | 4.3 | 9.2 | 6.7 | 6.3 | 6.3 | 6.3 | 6.5 | 6.3 |
| TOTAL | 52 | 5.9 | 6.4 | 8.3 | 7.4 | 7.3 | 6.5 | 6.9 | 7.1 | 7.0 |

Fuente: cálculos propios con base en la Tabla 9 de este documento.

Conclusiones y perspectivas

Atendiendo a los datos anteriores, se puede decir, a contrapelo de otras opiniones externadas por especialistas,²⁵ que 2018 fue, en general, un mal año para las encuestas nacionales en nuestro país. Pocos analistas mediáticos se dieron cuenta de ello porque, como de costumbre, el balance suele apoyarse en subjetividades y no en evidencias, pero el análisis de los datos concretos permitió afirmar que las mediciones demoscópicas de 2018 fueron menos exactas que las publicadas en elecciones presidenciales anteriores, en este siglo.

Ello se debió, en gran medida, a que la lectura del público en general, los legos, suele atender como criterio primordial la capacidad de las encuestas para detectar correctamente al líder de la contienda y, en todo caso, a la distancia que logra respecto de su competidor más cercano. En tanto, los especialistas (como se dijo anteriormente en este mismo texto), privilegiaron la exactitud en la estimación de todos los componentes en una contienda teniendo como segundo dato la distancia entre los dos primeros lugares, y atendieron sólo en forma marginal lo correcto del ordenamiento indicado por los números. Para los especialistas, una equivocación ocurre realmente cuando las estimaciones por encuesta se ubican fuera del rango de error estadísticamente tolerado; mientras que, para la gente común, se yerra cuando se “canta” mal al ganador.²⁶

25 Como la referida en la nota 2 de este mismo ensayo.

26 Claire Durand, “¿Todavía podemos tener confianza en las encuestas?”, conferencia presentada en el *Segundo Seminario de Estudios Sobre Opinión Pública*, Benemérita Universidad Autónoma de Puebla-CEDE Colegio de Especialistas en Demoscopia y Encuestas-Asociación Mundial de Investigadores de la Opinión Pública-Sociedad Mexicana de Estudios Electorales-Instituto Nacional Electoral-ISA Investigaciones Sociales Aplicadas-Universidad Nacional Autónoma de México-Universidad Veracruzana-Universidad Autónoma de Nuevo León-Asociación Mexicana de Agencias de Investigación de Mercados-Asociación Mexicana de Ciencias Políticas-Asociación Mexicana de Investigadores de la Comunicación, Puebla, Puebla, 25 de octubre, 2018.

La revisión de los datos sobre la exactitud de las encuestas publicadas para la elección presidencial de 2018 mostró que todas ellas estuvieron acertadas al detectar como líder a quien resultó ganador, aunque algunas erraron en el orden de llegada del segundo y tercer lugares. Esa fue la parte brillante que se reflejó en los juicios más ligeros. Sin embargo, al cotejar los niveles de exactitud logrados en términos estadísticos, se descubrió que el grado de desviación al estimar el margen de ventaja del ganador y la exactitud en el cálculo de las proporciones para los distintos componentes tuvieron una distancia media mayor que la observada en ocasiones anteriores en este mismo siglo.

El hecho de que por lo atinado en la detección del ganador no se haya manifestado una crítica hacia las encuestas en la temporada electoral 2018 no quiso decir que sea pertinente adoptar una actitud de autocomplacencia. En lugar de ello, debiera avanzarse en una crítica interna al campo demoscópico, que advierta y analice errores (y los posibles orígenes de estos), así los alcances y riesgos que representen para el futuro inmediato. Y eso vale tanto para las encuestas como para las agregadoras de encuestas un actor que llegó para quedarse; aunque, a pesar de su propia lectura de lo ocurrido,²⁷ no tuviera el éxito deseado en sus pronósticos.

¿Por qué fallaron las encuestas? El deseo cuasi-mágico de muchos analistas y críticos externos, e incluso de profesionales de la investigación, es encontrar explicaciones puntuales y fehacientes sobre las razones de este hecho, cuando así ocurre. Sin embargo, pareciera que no fuera posible encontrarlas, al menos no siguiendo las conclusiones del análisis reciente de Durand, sobre las encuestas en distintas naciones, que se ha citado antes.²⁸ La autora señala acertadamente que “cuando hay grandes errores, todas las encuestas cometen el mismo error”, lo que la ha llevado a afirmar que existen situaciones en las que fallan las encuestas donde no han fallado antes y no van a fallar después. Esto significa que, en situaciones específicas, en casos particulares, puede haber un problema de estimación sin que exista una norma ni reglas generales que permitan saber *ex ante* cuándo y dónde ocurrirán fallos, o cuándo y dónde las encuestas estarán en lo correcto, cualquiera que sea el criterio de acierto o exactitud que se observe.

De esta manera, ¿es factible disponer de pronósticos confiables y exactos de los resultados electorales como producto de las encuestas? Sabemos que una encuesta en particular, que se publique en un momento determinado, no constituye un pronóstico en sí. La evidencia mostró que, como fue el caso en esta ocasión, mediciones en un momento más anticipado a una elección pudieron resultar más próximas al resultado que aquellas que terminaron siendo las estimaciones finales, aunque las ya famosas tablas comparativas y de ordenamiento que suelen presentarse luego de una elección soslayaran esta realidad.

Por lo tanto, la tarea encaminada a disponer de un pronóstico más exacto debiera recaer en las agregadoras, a pesar de que estas también hayan dejado fuera el dato real en varios de los intervalos calculados para esta elección. Ante ello, tal vez lo correcto sería retomar la modestia y el realismo extremo para recuperar la necesidad de reportar intervalos mucho más amplios que correspondan efectivamente a lo conocido, y así dejar un amplio margen de incertidumbre.

Hicimos un ejercicio: tomamos el promedio aritmético de las encuestas finales publicadas y estimamos la desviación estándar de sus estimaciones, las cuales resultaron ser mayores a las

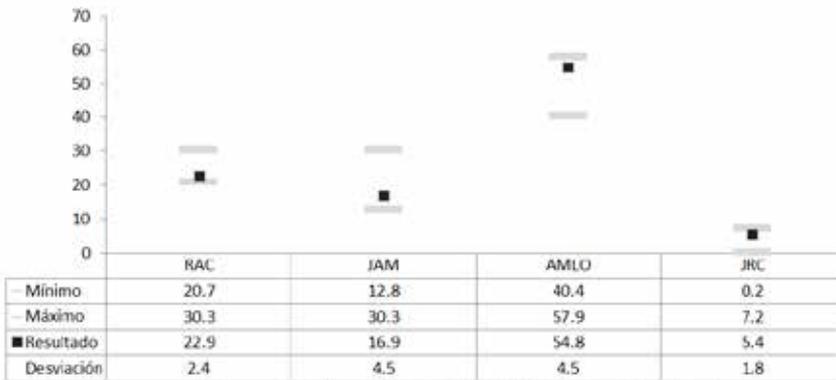
27 Nación 321, ¿Para qué sirven las agregadoras de encuestas y cuál fue su papel en las elecciones de 2018? [video en YouTube], 17 de julio, 2018. Disponible en <<https://www.youtube.com/watch?v=WSqg95eC35M>>, [fecha de consulta: 18 de julio, 2018].

28 Claire Durand, *op. cit.*

observadas en elecciones anteriores. Después estimamos los intervalos con una confianza de 95%, restando y sumando 1.96 desviaciones estándar al dato puntual. Los intervalos estimados fueron muy amplios: 17 puntos para López Obrador y Meade, y diez puntos para Anaya. Si se hacía lo propio en cuanto al margen de ventaja, lo esperable sería una distancia de entre 12 y 35 puntos.

Estos rangos resultaron amplísimos y no clarificaban el ordenamiento entre el segundo y el tercer contendientes. Empero, claramente incluían las proporciones que realmente se observaron (Gráfica 7). En consecuencia, se eliminó el problema de estimación equivocada propia de la comparación de cada encuesta contra el resultado, e incluso el de las horquillas estrechas, reportadas al final del camino por las propias agregadoras. Probablemente lo que cabe es reconocer con humildad que lo único que se conoce y se puede decir antes de una elección son datos todavía afectados por una elevada incertidumbre, y atenerse a ello.

Gráfica 7. Intervalos estadísticos al 95% de confianza con base en los resultados de las encuestas en vivienda para la elección presidencial 2018



FUENTE: Cálculos propios con base en Observatorio Electoral de CEDE e Instituto Nacional Electoral.

Se debe agregar que, en lo mediato, el éxito coyuntural de las encuestas tradicionales no significa una victoria definitiva. El desacierto de las mediciones por encuesta mediante nuevas tecnologías no descarta que estos métodos resulten pertinentes y adecuados para la obtención futura de información fidedigna sobre preferencias del electorado. Pero ello demandará cumplir con los principios éticos de la investigación científica, así como seguir rigurosos criterios y procedimientos que garanticen el control de calidad y fiabilidad de las estimaciones. Ese es el reto futuro, pues pensar que nada va a cambiar sería jugar al avestruz y perder.

Referencias

ABUNDIS, Francisco, "¿Quiénes eligieron a AMLO como presidente?", *Milenio Diario*, Ciudad de México, México, 10 de julio, 2018.
 ARREOLA, Federico, ZARAGOZA, Sergio y CAMPUZANO, Aldo, "La encuesta en Facebook de México Elige no fue la más acertada; la mejor fue la de *Reforma*; después, la de *El Financiero*", *SDP Noticias*, 1 de julio, 2018.

- BARNETT, Lincoln, "J. Robert Oppenheimer", *Life*, vol. 7, núm. 9, edición internacional, Time Incorporation, Nueva York, EE. UU., 24 de octubre, 1949, p. 58.
- BECERRA, Lorena y LEÓN, Rodrigo, "Nitidez en la contienda", *Reforma Revista R*, Ciudad de México, México, 8 de julio, 2018, p. 14.
- Colegio de Especialistas en Demoscopia y Encuestas (CEDE), "Observatorio Electoral 2018", *CEDE*. Actualmente no disponible, pero originalmente en <<http://cede.org.mx/web2016/observatorio-electoral/>>, [fecha de consulta: 16 de julio, 2018].
- DURAND, Claire, "¿Todavía podemos tener confianza en las encuestas?", conferencia presentada en el *Segundo Seminario de Estudios Sobre Opinión Pública*, Benemérita Universidad Autónoma de Puebla-CEDE Colegio de Especialistas en Demoscopia y Encuestas (CEDE)-Asociación Mundial de Investigadores de la Opinión Pública-Sociedad Mexicana de Estudios Electorales-Instituto Nacional Electoral-ISA Investigaciones Sociales Aplicadas (ISA)-Universidad Nacional Autónoma de México-Universidad Veracruzana-Universidad Autónoma de Nuevo León-Asociación Mexicana de Agencias de Investigación de Mercados-Asociación Mexicana de Ciencias Políticas-Asociación Mexicana de Investigadores de la Comunicación, Puebla, México, 25 de octubre, 2018.
- Instituto Nacional Electoral (INE), "Encuestas electorales: elecciones federales ordinarias 2017-2018 estudios entregados", *INE*. Disponible en <<http://www.ine.mx/voto-y-elecciones/encuestas-electorales/elecciones-federales-ordinarias-2017-2018-estudios-entregados/>>, [fecha de consulta: 16 de julio, 2018].
- Instituto Nacional Electoral (INE), "Estudios entregados a la Secretaría Ejecutiva", *INE*. Disponible en <<https://www.ine.mx/voto-y-elecciones/encuestas-electorales/elecciones-federales-ordinarias-2017-2018-estudios-entregados/berumen-2/>>, [fecha de consulta: 16 de julio, 2018].
- LAAKSO, Markku y TAAGEPERA, Rein, "Effective Number of Parties: A Measure with Application to West Europe", *Comparative Political Studies*, núm. 12, SAGE, Oslo, Noruega, 1979, pp. 3-27.
- MITOFSKY, Warren J., "Review: Was 1996 a Worse Year for Polls than 1948?", *The Public Opinion Quarterly*, vol. 62, núm. 2, Oxford University Press, Oxford, Reino Unido, 1998, pp. 230-249.
- MORENO, Alejandro, *El cambio electoral. Votantes, encuestas y democracia en México*, Fondo de Cultura Económica, Ciudad de México, México, 2018.
- MORENO, Alejandro, "Lucha contra la corrupción, el principal reto de López Obrador", *El Financiero*, Ciudad de México, México, 3 de julio, 2018.
- MOSTELLER, Fredrick, "Measuring the Error", en MOSTELLER, Fredrick, HYMAN, Herbert, MCCARTHY, Philip J., MARKS, Eli S. y TRUMAN, David B. (eds.), *The Pre-election Polls of 1948. Report of the Committee on Analysis of Pre-election Polls and Forecasts*, Bulletin 60, Nueva York, EE. UU., Social Science Research Council, 1949, pp. 54-80.
- Nación 321, ¿Para qué sirven las agregadoras de encuestas y cuál fue su papel en las elecciones de 2018? [video de YouTube], mesa redonda transmitida el 17 de julio, 2018. Disponible en <<https://www.youtube.com/watch?v=Wsqg95eC35M>>, [fecha de consulta: 18 de julio, 2018].
- Notimex, "INE ordena a empresas dejar de hacer llamadas telefónicas contra López Obrador", *El Economista*, Ciudad de México, México, 26 de junio, 2018. Disponible en <<https://www.economista.com.mx/politica/INE-ordena-a-empresas-dejar-de-hacer-llamadas-telefonicas-contra-Lopez-Obrador-20180626-0071.html>>, [fecha de consulta: 16 de julio, 2018].
- Oraculus, "Poll of Polls", *Oraculus*. Disponible en <<https://oraculus.com.mx/>>, [fecha de consulta: 16 de julio, 2018].
- PEÑA, Ricardo de la, "Cómo se mide la exactitud de las encuestas electorales", *Política y Cultura*, núm. 44, Universidad Autónoma Metropolitana (Xochimilco)-División de Ciencias Sociales y Humanidades-Departamento de Política y Cultura, D. F., México, otoño de 2015, pp. 217-247.
- PEÑA, Ricardo de la, "Encuestas en vivienda para la elección por la presidencia de la República en México 2018" [base de datos], *Ricardo de la Peña*. Disponible en <<http://www.ricartur.com/databases/EncuestasPresidenciales2018.txt>>, [fecha de consulta: 14 de agosto, 2018].
- PEÑA, Ricardo de la, "Encuestas en vivienda para las elecciones por la presidencia de la República en México 2000-2018"

- [base de datos], *Ricardo de la Peña*. Disponible en <[http://www.ricartur.com/databases/ EncuestasPresidencialesSXXI.txt](http://www.ricartur.com/databases/EncuestasPresidencialesSXXI.txt)>, [fecha de consulta: 14 de agosto, 2018].
- PEÑA, Ricardo de la, "Las encuestas presidenciales en México durante el siglo XXI", *Revista Mexicana de Estudios Electorales*, núm. 15, Sociedad Mexicana de Estudios Electorales, d. f., México, enero-junio, 2015, pp. 78-114.
- PEÑA, Ricardo de la, "Una alternativa para estimar la exactitud de las encuestas", *Política y Cultura*, núm. 49, Universidad Autónoma Metropolitana (Xochimilco)-División de Ciencias Sociales y Humanidades-Departamento de Política y Cultura, d. f., México, primavera de 2018, pp. 123-156.
- Verificado 2018, "Anaya sí contrató a la encuestadora Massive Caller", *Milenio Diario*, Ciudad de México, México, 2 de mayo, 2018. Disponible en <<http://www.milenio.com/elecciones-mexico-2018/anaya-si-contrato-a-la-encuestadora-massive-caller>>, [fecha de consulta: 16 de julio, 2018].