

Evaluación de la precisión de las encuestas electorales. Desafíos y retos metodológicos

Lauro Mercado Gasca*

Introducción

Desde hace 30 años, las encuestas electorales en México han sido una fuente invaluable de información sobre la opinión pública en general, y el comportamiento electoral en particular. Conforme se fue abriendo el sistema político mexicano, y la democracia electoral fue ganando terreno a la hegemonía del PRI, las encuestas se convirtieron en uno de los actores relevantes de este proceso. En un primer momento, las encuestas a la salida de las casillas (*exit polls*) y los conteos rápidos de resultados electorales, fueron una fuente alternativa para conocer los resultados de una elección. Independiente del órgano regulador de las elecciones, el cual carecía de independencia e imparcialidad. Varios “fraudes electorales” se evitaron de esta manera, comenzando con la forzada aceptación por parte del PRI, del triunfo del PAN en la elección para gobernador de Baja California de 1989. Ernesto Ruffo logró así ser el primer gobernador no priista en la historia de México.

En 1988 el presidente Carlos Salinas de Gortari (CSG) había ganado la elección presidencial con un fuerte reclamo por parte de los partidos de oposición de haber hecho un fraude electoral y ser, por tanto,

* Director General de Mercaei.

un presidente ilegítimo. En este contexto, una de las líneas de política pública seguida por CSG para legitimarse fue la de ir “abriendo el sistema político”, permitiendo el triunfo de los partidos de la oposición en las elecciones donde éstos ganaran. Esto, a contrapelo de los intereses de los líderes más conservadores dentro de su partido, el PRI. Fue así como las encuestas electorales y los *exit polls* se convirtieron en instrumentos para contra restar la resistencia del PRI a perder puestos de elección popular. Lo que ocurrió en Baja California fue la primera de varias batallas en este sentido.

A lo largo de los siguientes seis años las encuestas fueron ganando presencia ante la opinión pública, tanto que, para la elección presidencial de 1994 no fue sorpresa que se publicaran cerca del día de la elección ocho encuestas nacionales con las estimaciones de las preferencias electorales para cada candidato, además de realizarse varias encuestas a la salida de las casillas el día de la jornada electoral. Las preferencias favorecían claramente al candidato del PRI, lo cual también otorgaría un bono de legitimidad para el candidato del oficialismo.¹ Este desarrollo de las encuestas fue incluso legislado pues, como se verá más adelante, la regulación de 1993 obligaba a las empresas a publicar solo encuestas con criterios metodológicos científicos como el de una muestra representativa. Todas estimaron, aunque con diferentes márgenes de ventaja, que Ernesto Zedillo, candidato del PRI, ganaría por una cómoda mayoría la elección presidencial de ese año. Fue notable que las distintas mediciones registraron de manera bastante consistente la resistencia al cambio por parte del electorado, favoreciendo al otrora partido oficial con su voto de una manera muy apabullante. En una proporción cercana a dos a uno los resultados oficiales favorecieron al candidato del PRI, Ernesto Zedillo, sobre su más cercano contendiente, el candidato del PAN, Diego Fernández de Ceballos.

Cuando la Cámara Nacional de la Industria de la Radio y Televisión dio a conocer el resultado de tres ejercicios simultáneos de *exit polls*,

¹ “A nivel Federal, hasta 1997 las encuestas en México fueron un instrumento que servía para mostrar que el régimen político mexicano era cada vez más democrático, por cuanto las encuestas de opinión mostraban que el PRI se mantenía en la delantera de la preferencia electoral y que, en virtud de ciertos eventos, por ejemplo, el debate de 1994, la oposición subía en la preferencia del electorado, aunque no lo suficiente para ganar la elección” (Kuschik 2002, 124).

realizados para conocer el ganador de dicha elección, nadie se sorprendió de que el ganador fuera el candidato del PRI, resultado que fue corroborado más tarde mediante el conteo oficial de los votos por parte del IFE (Nares 2018).

En contraste con la tensión, la amenaza de violencia, los reclamos de fraude y la incertidumbre que caracterizaron la noche de la elección presidencial en 1988, la noche de los comicios presidenciales de 1994 transcurrió en calma y sin mayores sobresaltos. Parecía que México transitaba exitosamente hacia una democracia electoral confiable. Y quedó claro que las encuestas jugaron bien su papel.

En suma, el contexto político de la época fue el momento preciso en el cual se visibilizó la necesidad de contar con información confiable sobre los procesos electorales en el país. El cuestionado proceso electoral de 1988 y las sospechas de fraude, el surgimiento del IFE y su independencia del gobierno ya para las elecciones federales de 1997, una creciente competitividad electoral y pluralismo político, y la profesionalización de los medios de comunicación, fueron parte del contexto en el cual se ciernen los estudios demoscópicos en México.²

Y una vez liberadas las ataduras políticas autoritarias, la dinámica propia del juego electoral se desató y comenzó a crecer con mayor fuerza cada vez. A este escenario se le sumó el descontento social debido a la grave crisis económica de 1995, el fortalecimiento de liderazgos e instituciones políticas, notablemente Cuauhtémoc Cárdenas con el recién creado PRD, y la aparición de otros, en el espectro ideológico opuesto, como Vicente Fox, con un PAN más pragmático que ideológico, como había sido hasta entonces. Las consecuencias fueron que, en las elecciones federales intermedias de 1997, el PRI perdió por primera vez en su historia la mayoría en el Congreso Federal, y paralelamente, Cuauhtémoc Cárdenas ganó la elección de Jefe de Gobierno del Distrito Federal, igualmente un triunfo inédito de la oposición. Y tres años más tarde, Vicente Fox del PAN, arrebató al PRI, por primera vez en la historia, la Presidencia de la República.

² Valdés (1994) añade también el desarrollo de una cultura política en la ciudadanía mexicana como un elemento que posibilitó el surgimiento y consolidación de las encuestas electorales en México (a nivel federal) a partir del año 1994.

Con relación a las encuestas, a diferencia de lo ocurrido en 1994 donde hubo consenso en las mismas sobre quién sería el ganador, en la elección del año 2000 no hubo tal consenso. Algunas agencias de investigación de opinión o casas encuestadoras estimaron que el PRI ganaría la elección, mientras que otras mediciones estimaron que sería el PAN el ganador. Este hecho generó mucha controversia y avivó un tema recurrente conocido como “guerra de las encuestas”. Sin embargo, durante la jornada electoral todos los *exit polls* realizados, públicos o no, coincidieron desde alrededor de la una de la tarde, que Vicente Fox sería el siguiente presidente de México. Ese consenso en las encuestas de salida, el hecho de que el principal candidato opositor ganara y el reconocimiento por parte del presidente Zedillo primero y el PRI después, de su derrota, derivaron en un ánimo festivo, de cambio. Y la guerra de encuestas perdió relevancia para la opinión pública. Sin embargo, entre los especialistas, los resultados mixtos que mostraron las encuestas electorales en el año 2000 abrieron una nueva perspectiva sobre el papel de las encuestas en México, principalmente, su uso político y propagandístico. Con ello también se buscaron explicaciones a los fallos que tuvieron varias de las encuestas en acertar al ganador (Beltrán 2000). Aunque la explicación más básica se centró en el tema del ocultamiento del voto en el contexto de un régimen autoritario, otras lecturas lo atribuyeron a las características de las propias encuestas, entre ellas su financiamiento y sus metodologías, elementos que finalmente contribuían a su uso propagandístico; ello puede explicar también que hubo encuestas que sí pudieron estimar con certeza al ganador de la elección. Sin embargo, las críticas a las divergencias entre las encuestas no trascendieron demasiado a la opinión pública, quien, como se mencionó, se encontraba más interesada en el triunfo de Fox.

Esta calma e indiferencia hacia las encuestas luego de la jornada electoral no se volvería a vivir sino hasta la elección de Andrés Manuel López Obrador en el 2018. En las siguientes dos elecciones presidenciales, las encuestas jugaron un rol altamente polémico y fueron usadas por los políticos y los medios de comunicación para atizar los conflictos postelectorales en 2006 y en 2012. A finales de 2005 nadie dudaba que Andrés Manuel López Obrador ganaría la elección presidencial de 2006, gracias a su buen desempeño como Jefe de Gobierno del Distrito Federal que se tradujo en altos niveles de popularidad y aprobación

de su gobierno. Esto lo logró en parte gracias a una exitosa campaña de comunicación política, donde dejó plasmada su imagen como la de un político honesto, efectivo, cercano a la gente y preocupado por las clases más desfavorecidas y las poblaciones vulnerables como lo son las personas de la tercera edad y las madres solteras. Aun así, una efectiva campaña negativa por parte del PAN a partir de inicios del 2006, donde agregó información nueva al electorado y, en resumen, que logró posicionar la idea en muchos mexicanos de que AMLO era un peligro para México. A esto se le sumaron errores de la propia campaña de AMLO. El resultado fue que de la cómoda ventaja en las preferencias electorales de que gozaba en diciembre de 2005, la elección se empatara entre él y el candidato del PAN, Felipe Calderón, hacia marzo de 2006. Al final el candidato del PAN, Felipe Calderón, ganó la elección por poco más de medio punto porcentual. Como veremos, aproximadamente la mitad de las últimas encuestas publicadas daban como ganador a AMLO, mientras que la otra mitad daba como ganador a Felipe Calderón. Amparado en la mitad de las encuestas que lo favorecían AMLO no aceptó los resultados oficiales, acusó fraude, se declaró presidente legítimo, y tomó en protesta la Avenida Reforma por varios meses. La derrota del candidato opositor, y su declaración de fraude, hicieron que la crítica hacia las encuestas fuera intensa. En el año 2012 la polémica volvió al ruedo. En esta ocasión el candidato del PRI ganó por un margen menos estrecho que el de 2006, pero mucho más cerrado del que anunciaban las encuestas publicadas. Derrotado nuevamente el candidato opositor, AMLO, el ataque a las encuestas y a los encuestadores fue despiadado. Siendo el argumento que habían inflado artificialmente la ventaja de Enrique Peña Nieto, con lo que evitaron que se detonara el “voto estratégico” en favor de AMLO. El debate sobre el uso propagandístico de las encuestas, las encuestas falsas, y la influencia de las encuestas sobre la decisión del voto renació con nueva fuerza erosionando la credibilidad del gremio de encuestadores políticos.

Aunado a ese descrédito, en los siguientes años se sumaron distintas crisis de las encuestas en el extranjero, el triunfo del Brexit, el inesperado triunfo de Donald Trump, y el sorpresivo triunfo del NO en el plebiscito en Colombia para decidir el fin de la guerra contra las FARC, fueron los ejemplos más sonados. Y llega entonces la elección presidencial en México de 2018, donde AMLO logra una aplastante victoria

con más del 50% de los votos. A lo largo de la campaña todas las encuestas dieron por mayor o menor margen, la ventaja a AMLO, y los *exit polls* coincidieron en medir el esperado triunfo arrollador. En esta ocasión, como en 1994 o en 2012, todas las encuestas daban como ganador a quien resultó triunfante. Y, al haber ganado el candidato opositor al oficialismo, la percepción pública fue que las encuestas, ahora sí, habían hecho bien su trabajo. Sin embargo, ¿realmente lo hicieron? En realidad, hubo serias discrepancias entre las estimaciones de los resultados y el resultado oficial. Este hecho no debe ser desestimado por los especialistas, más allá de la percepción pública.

El presente capítulo tiene por objeto mostrar un panorama de las encuestas electorales en México a lo largo de los últimos 25 años y los desafíos y retos que han enfrentado y enfrentan aún hoy con relación a la precisión de sus estimaciones. El foco del estudio son las encuestas para elecciones presidenciales publicadas por las distintas casas encuestadoras o medios de comunicación, normalmente éstas se realizan alrededor de 15 días antes de la elección, pero en ocasiones la última encuesta de una firma se realizó con un mes de antelación. Abonaremos a la discusión teórica intentando distinguir entre precisión estadística y precisión popular, vía los ciudadanos, políticos, y medios de comunicación que consumen esta información.

El foco en encuestas presidenciales se debe, primeramente, por cuestiones de espacio y, en segundo lugar, porque es el área mejor documentada. El estudio comprenderá por tanto las cinco elecciones presidenciales comprendidas entre 1994 y 2018. El análisis revelará un problema recurrente, como industria, que no ha mejorado en estos 25 años. A saber, que la varianza en las estimaciones finales de la elección entre casas encuestadoras, se ha mantenido amplia a lo largo de estos años. El problema pasa desapercibido cuando el ganador lo hace por un amplio margen. Pero si las elecciones son cerradas, la diferencia de estimaciones se hace evidente, en el peor de los casos con algunas agencias dando a un candidato como ganador y otras a otro, casos 2000 o 2006. O el margen de ventaja de la primera fuerza era menor a la esperada, caso 2012. Como hemos visto hay un tercer ingrediente del que depende que estas discrepancias se traduzcan en polémica, conflictos políticos y descrédito para los encuestadores. Esto es, si gana o pierde el candidato opositor al régimen. En 2000 y 2018 este hecho

generó indiferencia hacia las discrepancias. En contraste, en 2006 y 2012, las discrepancias fueron leña para el fuego.

El capítulo presenta la siguiente estructura, en la primera parte presentamos elementos del debate teórico sobre cómo se debe medir la precisión de una encuesta. En la segunda, seleccionamos cinco métricas específicas y sometemos a examen los resultados de las encuestas previas más cercanas a la elección presidencial del periodo de interés. En tercer lugar, tras el panorama presentado, ponemos a debate los aspectos más importantes que han sido presentados en la palestra de la opinión pública por medios de comunicación, académicos, políticos y por el propio gremio encuestador. Subdividimos este apartado en los aspectos regulatorios, las causas de la discrepancia entre casas encuestadoras y entre estas (en su conjunto) y los resultados electorales, y finalmente, las principales posturas sobre la precisión de las encuestas electorales que se han dado entre los actores del gremio. En cuarto lugar, reseñamos brevemente el nuevo enemigo para la credibilidad de las encuestas, a saber, las encuestas por Facebook y las telefónicas por robot dadas sus grandes carencias metodológicas por ser encuestas no probabilísticas. Es el mismo debate que se presentó en los años noventa cuando algunos encuestadores insistían en hacer encuestas en calle, en puntos de afluencia, hasta el 2000 se dieron estos casos, pero reeditado y empaquetado tecnológicamente para ocultar la falsedad de sus estudios. Por último, se presentan las conclusiones del capítulo y algunas sugerencias para seguir con la mejora continua de la industria.

¿Cómo debe medirse la precisión de las encuestas electorales?

En todo el mundo, el debate sobre las encuestas lleva aparejada la discusión sobre la precisión de sus estimaciones, pues sin duda, estas dan sentido a su ejercicio y México no ha sido la excepción. Ahora bien, ¿cuál debe de ser el indicador o la métrica que debe definir la precisión de las encuestas electorales?, ¿atinarle a la candidatura ganadora?, o más bien, independientemente de acertar al ganador, la métrica clave es que la estimación electoral de la encuesta, a pocos días de la

elección, se encuentre dentro del intervalo de confianza especificado por el error muestral de dicha encuesta (esto siempre y cuando no haya habido ningún acontecimiento importante entre la medición y el día de la elección). ¿Importa sólo la diferencia entre la estimación de la encuesta y el resultado oficial de la primera fuerza?, ¿O de la diferencia de la primera y la segunda fuerza?, ¿O del promedio del error de las estimaciones de todos los partidos?, ¿Es relevante medir el desempeño de las encuestas en su generalidad y no cada una de manera independiente?

Nos guste o no, las respuestas a estas preguntas dependerán de quién las responda. A manera de simplificación, podemos pensar que hay al menos tres grupos de personas con diferentes tipos de intereses y conocimiento, que derivan en tres grupos de respuestas. El primer grupo sería el de los especialistas: directivos e investigadores en agencias de investigación de opinión pública, actuarios, especialistas en opinión pública, científicos políticos, sociólogos, psicólogos sociales, etc. El hilo conductor entre ellos es el haber sido entrenados en metodología de la investigación científica y particularmente en estadística. El segundo grupo estaría conformado por personas usuarias de la información generada por las encuestas electorales, principalmente los estrategas de campañas políticas, los políticos mismos y los medios de comunicación. En general los primeros tienden a saber interpretar mejor una encuesta y a distinguir metodologías correctas e incorrectas. En los medios de comunicación existe todo un rango de nivel de *expertise*, desde quienes no conocen nada de encuestas hasta quienes se pueden considerar casi expertos. En el tercer grupo se encuentra la opinión pública en general, normalmente poco enterada de temas metodológicos y, por ende, indefensa para poder distinguir encuestas serias de sondeos carentes de validez metodológica.

Según la opinión pública

A esta población lo que más le interesa conocer es “la carrera de caballos”, es decir quién va ganando, por cuánto, quién va en segundo lugar, etc. Y tener una estimación de qué tanto podrían cambiar las estimaciones de una encuesta de ese día hasta la jornada electoral. La norma en este segmento es tener poco o nulo conocimiento metodológico.

Por ello, desafortunadamente, le atribuye la misma validez a una encuesta levantada en el momento por una estación de radio o TV, o alguna de las que salen en los portales noticiosos donde te invitan a votar por uno de los candidatos o candidatas, que a una encuesta telefónica en vivienda o vía Facebook que sólo alcanzan a tocar a la mitad del electorado, o a una encuesta telefónica automatizada por robot, donde los respondientes se autoseleccionan, al igual que en las otras modalidades mencionadas en este párrafo. Ninguna de esas modalidades **representa** a la opinión pública ni al electorado en su conjunto. Tampoco provienen de una muestra probabilística por lo que no podemos decir que representen a alguna población en específico. Sin embargo, les dan la misma validez que una encuesta metodológicamente seria, que para el caso de México debe ser, además de basarse en una muestra probabilística, preferentemente ser una encuesta cara a cara en la vivienda del entrevistado (no en puntos de afluencia), donde cada hogar fue seleccionado siguiendo una serie de etapas previas sucesivas de selección aleatoria de manzanas, secciones electorales, municipios, etc. Y dentro del hogar se selecciona también aleatoriamente a un habitante de este. Este mismo desconocimiento por parte del público, genera que no le preste atención a la nota metodológica, la cual debe dar los elementos para poder saber qué tan confiable es esa encuesta, para qué lugar geográfico, para qué fecha, etc. Esto es, debe incluir el universo de estudio, el número de entrevistados, el diseño muestral, las fechas y el modo de levantamiento, así como el margen de error y el nivel de confianza estadística del estudio. Como resultado, es muy fácil que un ciudadano llegue al final de una campaña electoral confundido y desconfiado de “las encuestas”, porque para este ciudadano todas las encuestas entran en la misma canasta.

Según los estrategias de campaña, candidatos y medios de comunicación

Este segmento de actores suele tener un mayor nivel de conocimiento sobre la metodología de encuestas. Normalmente estos actores sí revisan la nota metodológica y pueden distinguir entre una encuesta seria que representa al electorado de su interés, de otra que fue elaborada violando las reglas necesarias. Adicionalmente, estos públicos se

interesan por información valiosa, adicional a la carrera de caballos, que trae la encuesta y que puede servir para el diseño de las campañas electorales de sus candidatos, o bien para explicarle a las audiencias porqué las preferencias electorales se están dando de una manera y no de otra. Estos temas pueden estar relacionados con el nivel de conocimiento y las opiniones que despierta cada candidato, la percepción que tienen las personas acerca de sus autoridades, identificar si existe un potencial voto de premio o castigo, conocer la extensión y profundidad de las simpatías o antipatías hacia los diferentes partidos políticos, o bien, saber la importancia que le atribuyen diferentes segmentos sociales a los diferentes problemas del país, estado, municipio o distrito. Quien sabe distinguir las encuestas buenas de las malas, y quien sabe explotar la información que suelen acompañar el dato de la preferencia electoral, puede tener una clara ventaja competitiva contra candidatos rivales, o medios noticiosos de la competencia. A pesar de lo mencionado, uno se encuentra de todo, hay candidatos o medios de comunicación que hacen caso omiso de los criterios mencionados y leen la información que quieren leer, o comunican la información como la quieren comunicar. Aún hay camino por recorrer en la “profesionalización” del uso de las encuestas electorales en este segmento.

Según los especialistas

En términos generales puede decirse que los especialistas califican las encuestas con base en qué tanto se apegan a los cánones metodológicos aceptados por la academia y la industria. A saber, que el marco muestral con el que se diseñará la muestra está limpio y completo, es decir, no deja fuera a ninguna parte de la población que se desea estudiar. Que el diseño muestral sea probabilístico. Esto es, que todos y cada uno de los miembros del universo bajo estudio tengan una probabilidad conocida y diferente a cero de haber sido seleccionados para participar en la encuesta. Y que la selección de las personas a entrevistar sea aleatoria, o se realice siguiendo mecanismos que garanticen la aleatoriedad de los seleccionados. Adicional al tipo de selección de la muestra, debe considerarse el tamaño de dicha muestra, esto es, la cantidad de entrevistas a realizar. A parte del tema muestral, deben cuidarse una serie de criterios metodológicos para lograr captar

fielmente la opinión pública del momento. A saber, el cuidado en el diseño de las preguntas y la estructura del cuestionario a aplicar. La calidad del trabajo en campo, cuidando que los encuestadores realicen su labor, que lo hagan siguiendo las instrucciones dadas, que la tasa de rechazo a la encuesta no sea muy alta, etc. Y finalmente el cuidado en el procesamiento de los datos recopilados, su adecuada ponderación y uso para análisis estadísticos.

Ahora bien, en teoría, si todos los elementos mencionados fueron realizados con el cuidado debido, los resultados que arroje una encuesta electoral, específicamente la estimación de los porcentajes de preferencia electoral hacia cada uno de los candidatos y partidos políticos, si la misma se realizó a pocos días de la elección, y en esos días no surgió nueva información relevante para los electores, dichos porcentajes estimados deben de corresponder a los resultados oficiales de la elección, dentro del intervalo de confianza dado por el margen de error esperado.

Métodos para estimar la precisión de las encuestas

Existen diversas formas para comparar el mayor o menor acercamiento entre las estimaciones de una encuesta y el resultado electoral (De la Peña 2016a, Moreno 2011 y 2018, Zuckermann 2012a y 2012b). Para efectos del presente capítulo utilizaremos 4 criterios.

Criterio 1: La encuesta acertó al ganador.

Aunque sabemos que técnicamente no es una condición indispensable para juzgar una encuesta, lo incluimos debido a que, al final del día, éste es el dato que más le importa a la opinión pública, a la clase política, a los medios de comunicación, y por supuesto a los contendientes de una campaña electoral.

Criterio 2: Si el resultado oficial de la primera fuerza se encuentra contenido dentro del intervalo de confianza, definido por el margen de error muestral, de la estimación de la encuesta previa, en el presente capítulo calificaremos a esa encuesta o esa estimación como precisa.

Este criterio puede o no coincidir con el anterior. Es decir, puede darse el caso en que una encuesta estime que el ganador va a ser x candidato, y acierte. Sin embargo, el porcentaje con el que ganó el candidato es menor o mayor que el intervalo de confianza establecido por la encuesta. Esta será una encuesta certera mas no precisa. Puede darse el caso al revés, es decir, que la elección se encuentre tan cerrada que una encuesta puede dar como candidato ganador al candidato A, pero gana el candidato B. Sin embargo, lo cierto es que ambos candidatos caían dentro del intervalo de confianza de la estimación. En este caso tendremos una encuesta no certera pero precisa. En realidad, en este segundo caso lo correcto es que el encuestador no reporte ganador alguno, porque estadísticamente hablando la encuesta solo alcanza a definir que existe un “empate técnico” entre los candidatos punteros. Esto es, la encuesta será precisa, más no tendrá la capacidad de ser certera con relación a quién ganará.

Criterio 3: La similitud de la **distancia entre la primera y la segunda fuerza** en los resultados oficiales, comparados con la misma distancia en la encuesta.

Otra métrica que permite diferenciar las encuestas más y las menos precisas es medir quienes estimaron más adecuadamente la diferencia entre la primera y la segunda fuerza electoral. Cuando una contienda electoral es muy competida, y la distancia en cuestión es muy corta, alcanzar la mayor precisión posible es vital. Como se mencionó arriba, sabiendo siempre que la metodología de las encuestas implica un margen de error, si la distancia de la preferencia entre dos candidatos o candidatas es menor de cuatro puntos porcentuales, frecuentemente la casa encuestadora debe anunciar que existe un “empate técnico”, es decir, que estadísticamente hablando es imposible decir cuál de los candidatos va ganando.

Criterio 4: El error total promedio de cada encuesta. Esto es, la raíz de la suma de cuadrados de la diferencia entre la estimación en la encuesta por cada partido versus el resultado oficial para cada partido. Dividido entre el número de partidos. Para el caso de este estudio recodificamos en “otros” al resto de partidos pequeños.

Más allá de las implicaciones políticas y mediáticas que tiene acertar al candidato ganador, acercarse lo más posible a replicar el porcentaje de votos del ganador o los puntos porcentuales que separarán al

primero del segundo lugar, desde un punto de vista más técnico, la encuesta más precisa es aquella que tiene el menor error en la estimación de todos los porcentajes de preferencia estimados, no solo uno o dos. En la medida que el error promedio de cada encuesta se acerca a cero, mayor es la precisión global de dicho estudio.

Calificación de las encuestas presidenciales en México

Con base en estos cuatro criterios vamos a evaluar la evolución de la precisión de las encuestas electorales en México. Como ya se mencionó, haremos la comparación entre las últimas encuestas publicadas antes de cada una de las cinco elecciones presidenciales ocurridas entre 1994 y 2018.

Criterio 1: Acierto al ganador

Como podemos observar en la gráfica 1, en las elecciones de Ernesto Zedillo por el PRI en 1994, de Enrique Peña Nieto por el PRI en 2012 y de Andrés Manuel López Obrador por Morena en 2018, hubo consenso entre las casas encuestadoras acerca de quién sería el ganador de la elección. En este sentido, podríamos decir que, en tres de las últimas cinco elecciones presidenciales, el gremio de las encuestas electorales públicas fue acertado en sus estimaciones. Contrastan las dos elecciones donde ganó el PAN la presidencia. En el año 2000 con Vicente Fox y en el 2006 con Felipe Calderón.

En el año 2000, la mayoría —diez de 16 encuestas públicas— estimaron que el partido en el poder ganaría la elección; sin embargo, Francisco Labastida Ochoa del PRI fue derrotado por más de seis puntos porcentuales, y con ello terminó la hegemonía que por más de 70 años había mantenido su partido.

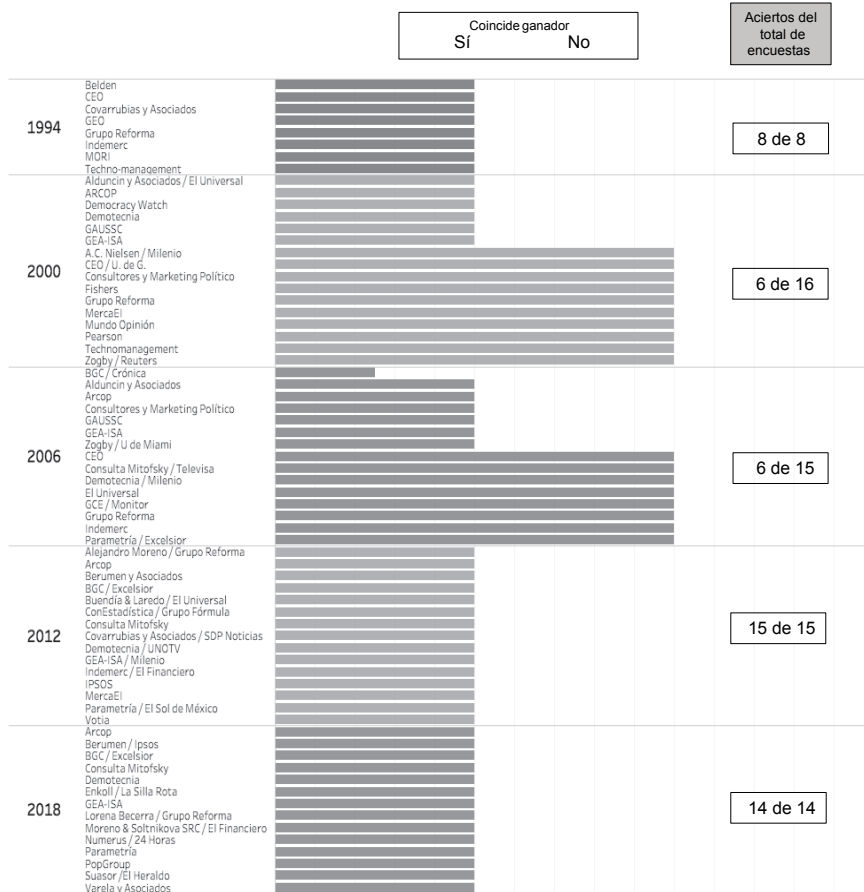
En el 2006, tampoco hubo consenso entre las casas encuestadoras. Similar al 2000, seis de 15 encuestas publicadas cerca de la elección le daban el triunfo a Felipe Calderón, mientras que 8 se lo daban a AMLO, y una encuesta dio empate. Si a esta publicación global de

resultados con mayoría de encuestas diciendo que ganaría AMLO, le sumamos la excitación política del momento, con los simpatizantes de cada candidato creyendo en las encuestas que favorecían a su candidato, y criticando las otras como propaganda. Cuando el resultado oficial le dio el triunfo a Felipe Calderón por siete décimas de punto, López Obrador salió de inmediato a protestar al Zócalo de la Ciudad de México y sus simpatizantes le creyeron que le habían hecho fraude. Es la opinión del autor que dado lo cerrado de la contienda, el mensaje individual y común debió haber sido que había un empate técnico y que cualquiera podía ganar. Pero, la falta de consenso por parte del gremio en su conjunto, en la métrica más relevante para la opinión pública, esto es, quién ganará la elección, mermó la confianza hacia las encuestas.

Cabe mencionar que es importante distinguir entre precisión y credibilidad hacia las encuestas. La primera es una medida objetiva mientras que la segunda es una medida subjetiva. Si comparamos la reacción pública hacia las encuestas entre la elección del 2000 y la del 2006, observamos que si bien, en ambos casos no hubo consenso sobre quién sería el ganador, e incluso en ambos casos la mayoría de las encuestas anunciaba que ganaría quien al final perdió, en el año 2000 no hubo polémica ni crítica hacia las encuestas, al menos no por parte de la opinión pública. Mientras que, en el 2006, las encuestas fueron fuertemente cuestionadas. Esto es, la reacción del líder inconforme y la de sus simpatizantes pesó más en la credibilidad de las encuestas que los resultados objetivos de las mismas, en ambos casos.

Como veremos al analizar los siguientes criterios, el criterio de quién ganará la elección es muy limitado para evaluar adecuadamente el desempeño de la precisión de las encuestas.

Gráfica 1. Años con coincidencia y sin coincidencia en acertar al ganador



Fuente: Elaboración propia con datos de resultados electorales oficiales del IFE y el INE, complementado de diversas otras fuentes: artículos académicos de revistas arbitradas, memorias de simposios post electorales entre especialistas, notas de periódico o revistas, Wikipedia. Se utilizaron varias fuentes para cruzar información y verificar su validez. Ver referencias específicas en Bibliografía.

Dicho de otra manera, observamos que la “calificación” y la “credibilidad” sobre la labor de las encuestas por parte de los medios de comunicación y la opinión pública se basa en el cumplimiento o no de la función de “pronóstico” que esperan de las mismas. Al simplificar de tal manera el desempeño de las encuestas, se desvía la atención de otros aspectos fundamentales, tales como las metodologías, los tamaños de muestra o los métodos de recolección de la información.

Criterio 2: Distancia con respecto al resultado oficial de la primera fuerza. Intervalo de confianza.

Las gráficas 2 a la 6 nos presenta un panorama menos halagüeño con relación a la precisión de las encuestas presidenciales preelectorales que el de la “predicción” del ganador. Estas gráficas parten de tres supuestos: la homologación del margen de error al $\pm 3\%$, el uso de la preferencia efectiva en las encuestas, y en los resultados oficiales. Primero, se partió de un criterio *pseudo*-arbitrario para definir el margen de error de las encuestas analizadas. A todas se les asignó un margen de error de ± 3.0 puntos porcentuales.³

Partiendo de esta base, las siguientes gráficas muestran la estimación de la preferencia efectiva⁴ por cada uno de los dos candidatos punteros, hecha por cada casa encuestadora que publicó su información. Una línea negra vertical de seis puntos porcentuales, tres por arriba y tres por debajo de la estimación puntual reportada por cada agencia. Esa línea negra es la representación gráfica del intervalo de confianza de la encuesta, es decir, si el resultado oficial queda incluido dentro del intervalo de confianza de una estimación, podemos decir, con solidez estadística, que la encuesta estimó con precisión el

³ Se adoptó el criterio de ± 3 puntos porcentuales por las siguientes razones: 1) El margen de error no fue siempre reportado. 2) La gran mayoría de las notas metodológicas asignan el margen de error “teórico muestral” como si el diseño de muestra utilizado hubiera sido una selección aleatoria simple. Lo cierto es que para encuestas cara a cara en vivienda es necesario realizar diseños muestrales complejos. Y cuando esto ocurre, es necesario multiplicar el margen de error que usualmente se utiliza por un número que es el coeficiente de diseño o def. el cual suele aumentar el tamaño real del margen de error. 3) Adicionalmente, la idea de un solo margen de error para todas las estimaciones de una encuesta es una simplificación ya que cada porcentaje estimado en un ejercicio de estos para conocer la preferencia electoral de una población tiene su propio margen de error asociado. Debe reconocerse que los reportes de encuestas enviados al INE para la elección de Presidente tuvieron, en general, una mayor calidad y detalle que lo observado en el pasado. Incluso, varias casas encuestadoras sí reportan el def. y el error asociado específico para cada porcentaje de preferencia. Sin embargo, esto aún no es la norma. 4) Por el alcance del estudio no se cuenta con la información necesaria para que el autor estimara correctamente todos los márgenes de error específicos. 5) Finalmente, debido que la mayoría de las encuestas nacionales reportadas presentan errores muestrales de 3% o alrededor. Se juzgó apropiado utilizar ese tamaño de error muestral y aplicárselo a las 68 encuestas reportadas.

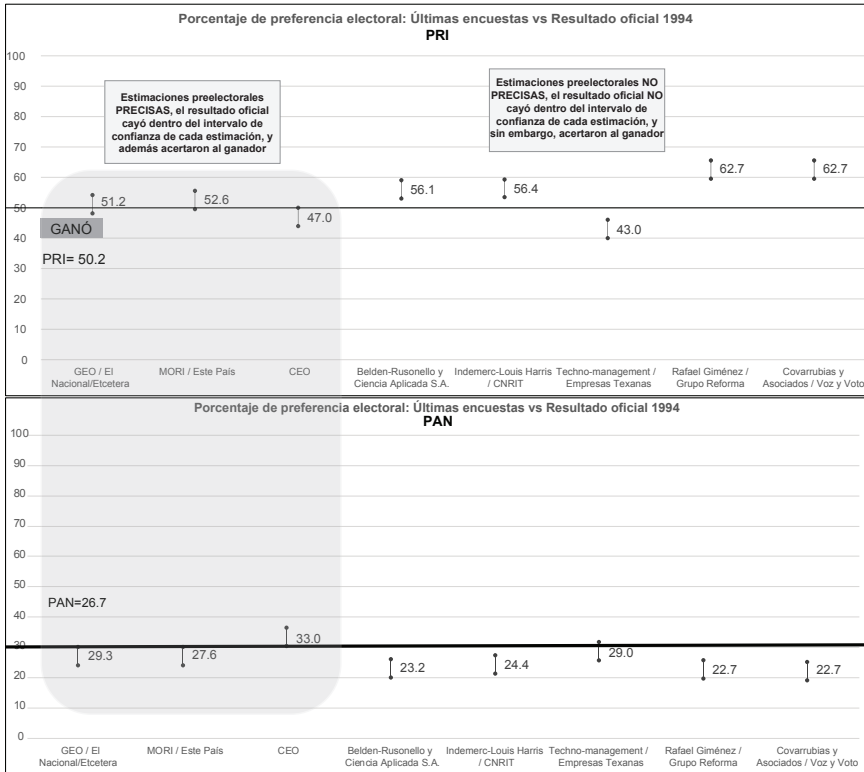
⁴ La preferencia efectiva se obtiene de descontar la no respuesta (ninguno, no sabe, no contestó) y recalcular a 100% los porcentajes de quienes sí manifestaron su preferencia electoral.

porcentaje de votos que obtendría ese candidato. Cada gráfica, para cada uno de los dos partidos punteros en cada año electoral presidencial, contiene una línea roja señalando el porcentaje efectivo oficial de votos que obtuvo ese candidato.⁵ Cuando las líneas verticales negras intersectan la línea roja, podemos decir que esa estimación es estadísticamente significativa, o en términos del presente capítulo, que es una estimación precisa. Como se ha señalado, esta aseveración se mantiene, desde el punto de vista técnico o estadístico, independientemente de que la encuesta haya o no acertado al candidato ganador. Por el contrario, aquellas estimaciones cuyo intervalo de confianza no intersecte el resultado oficial, no serán estimaciones precisas que podamos extrapolar al resto de la población. Y entre más alejadas se encuentren de dicha línea roja, mayor imprecisión tendrán.

Bajo esta perspectiva, **la elección de 1994**, a pesar de que las ocho encuestas públicas acertaron al ganador, únicamente tres, las de GEO, MORI y CEO, califican para decir que estimaron con precisión estadística el porcentaje de votos que obtuvo el candidato ganador. Los intervalos de confianza de las otras cinco encuestas quedaron fuera del resultado oficial.

⁵ Para todas las comparaciones, también utilizamos el porcentaje efectivo de la preferencia en la votación, es decir, eliminamos los votos por candidatos no registrados, los votos nulos y los votos en blanco, y recalculamos a 100% la preferencia efectiva obtenida por cada candidato o candidata.

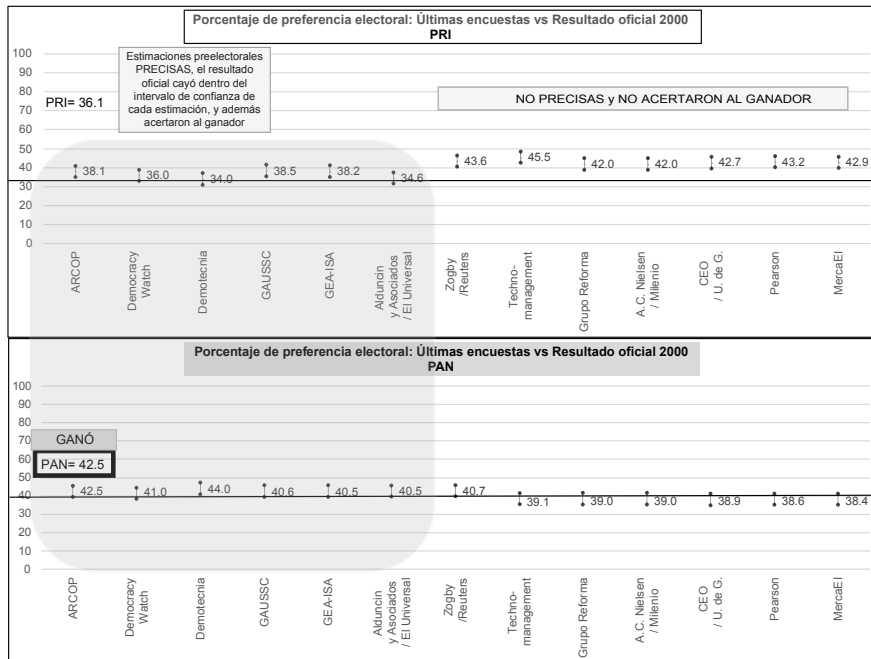
Gráfica 2. Error distancia primera y segunda fuerza: elección presidencial 1994



Fuente: Elaboración propia con las fuentes de datos de encuestas y resultados electorales ya mencionados.

En la **elección del 2000**, seis casas encuestadoras, de 16, acertaron a decir que Vicente Fox ganaría, y ese mismo número lo hizo con una precisión estadísticamente correcta. Dichas empresas fueron Arcop, Demotecnia, GEA-ISA, Democracy Watch, GAUSSC y Alduncin y Asociados quien publicó en el periódico El Universal. Las otras diez encuestas subestimaron el porcentaje de votos que obtendría Vicente Fox, sobre estimando la preferencia por Francisco Labastida del PRI, y todos los intervalos de confianza quedaron fuera del resultado oficial.

Gráfica 3. Error distancia primera y segunda fuerza: elección presidencial 2000

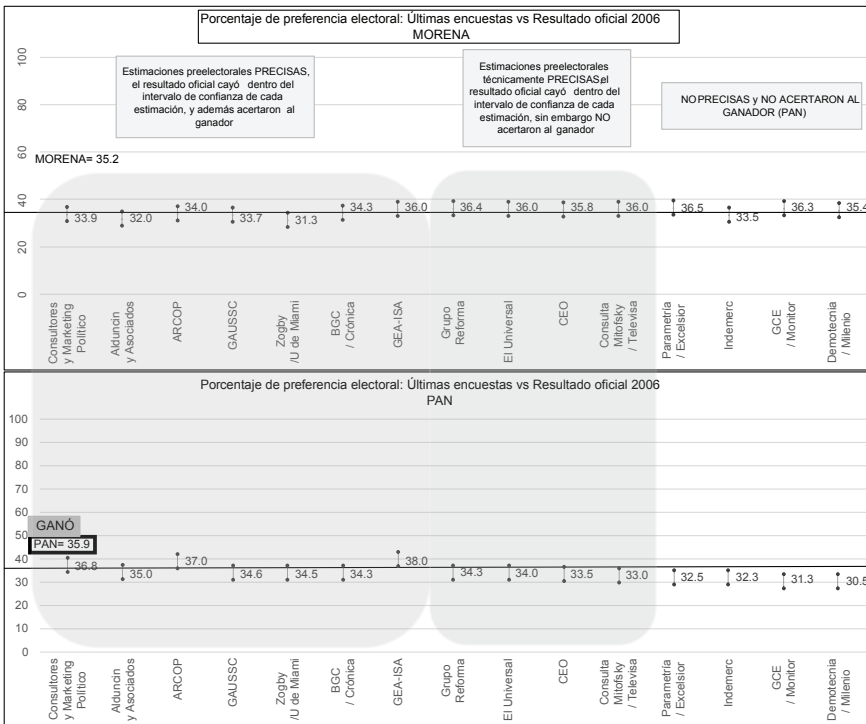


Fuente: Elaboración propia con las fuentes de datos de encuestas y resultados electorales ya mencionados.

La elección del 2006 fue un caso muy particular. Por un lado, como ya se mencionó, ha sido la elección más polémica de las cinco que estamos estudiando. Esta polémica fue en parte gestada por las encuestas ya que la mayoría de ellas estimó erróneamente que AMLO sería el ganador de la contienda. Sin embargo, paradójicamente, ha sido la elección con mayor precisión técnica que han tenido las encuestas. Siete de las quince encuestas publicadas se encontraban dentro del intervalo de confianza, calificándoseles de técnicamente precisas, y adicionalmente acertaron al ganador. Otras cuatro encuestas también fueron precisas, estadísticamente hablando, pero no acertaron al ganador. Por último, cuatro encuestas más no acertaron tampoco al ganador, y sólo las estimaciones para AMLO fueron precisas, las estimaciones para Calderón subestimaron su preferencia. Como se mencionó arriba, la polémica y la crisis política derivada de la elección tuvo más que ver con un problema de comunicación que de técnica de las encuestas. A toro pasado el mensaje que las encuestadoras y

los medios de comunicación debieron empeñarse en transmitir era uno afirmando que las mediciones daban un “empate técnico” y que “cualquiera de los dos candidatos punteros podía ganar” ya que las encuestas no tienen la potencia de mira como para poder determinar un ganador cuando la contienda está así de cerrada. Con ese escenario la historia post electoral pudo haber sido diferente. Si la gente hubiera sabido que la elección estaba empatada y que cualquiera podía haber ganado por muy pequeño margen, la credibilidad de que se había realizado un fraude se hubiera podido quizás desactivar. Sabemos, sin embargo, que eso no fue lo que pasó. Lo que ocurrió fue que la mayoría de las encuestas y por ende de las voces mediáticas anunciaron un virtual triunfo de AMLO. Al no ocurrir éste, la tierra estaba fértil para la protesta social.

Gráfica 4. Error distancia primera y segunda fuerza: elección presidencial 2006

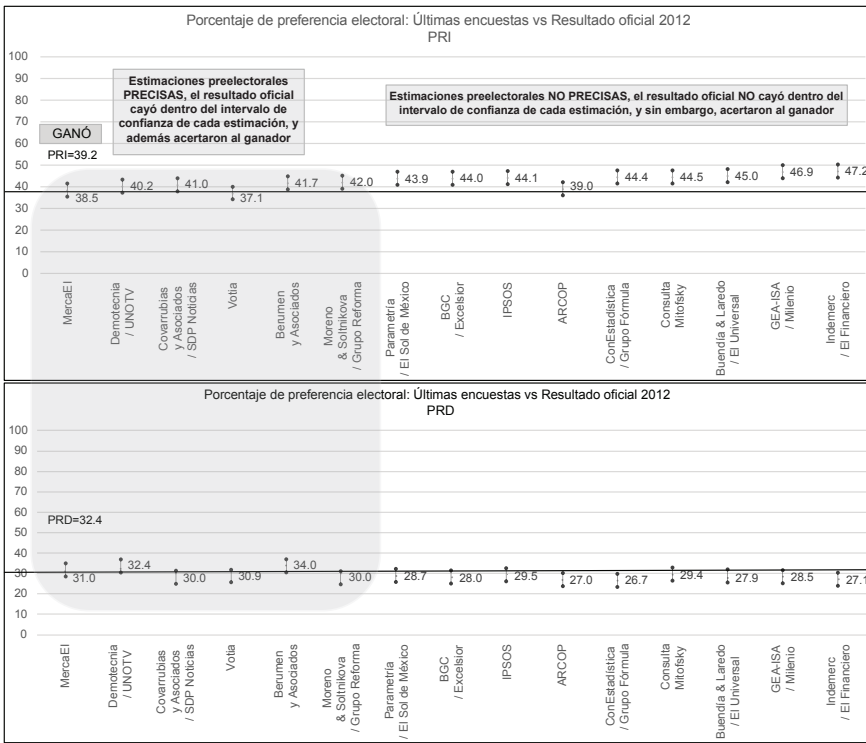


Fuente: Elaboración propia con las fuentes de datos de encuestas y resultados electorales ya mencionados.

En **las elecciones de 2012**, el mensaje principal enviado por las encuestas en su conjunto fue que el candidato del PRI ganaría por un amplio margen. Sin embargo, nueve de las quince encuestas públicas finales se encontraron fuera del intervalo de confianza. Sólo Demotecnia y Mercaei, tuvieron una precisión cuyo error fue menor a un punto porcentual respecto al porcentaje de votos que obtuvo EPN. Covarrubias y Asociados, Votia, Grupo Reforma y Berumen y Asociados también fueron muy certeros con la primera fuerza con errores de entre dos y poco menos de tres puntos porcentuales, pero claramente dentro del intervalo de confianza. Ese año, la sobre estimación de la preferencia por el PRI, por parte de ocho de las quince encuestas públicas, y con diverso grado de error, fue utilizada por AMLO para descalificar a las encuestas, y decir que se habían confabulado para desincentivar el voto estratégico en su favor. Esto es, en la medida que las personas percibieran que la ventaja del PRI no era insuperable, aquellos simpatizantes del PAN que sabían que su candidata perdería, pero que rechazaban más al PRI que a AMLO, podrían haber votado “estratégicamente” por AMLO y así darle el triunfo. El argumento cobró fuerza sobre todo porque en ese año las encuestadoras más visibles fueron las que aun acertando al ganador, sobre estimaron el tamaño de su triunfo. El debate sobre la influencia de las encuestas sobre el comportamiento electoral ha sido amplio y largo, y puede decirse que no se ha cerrado definitivamente hacia ninguno de los argumentos de quienes piensan que sí influye y quienes piensan que no.

Así, en 2012, las casas encuestadoras se enfrentaron al juicio popular y al descrédito por parte de la opinión pública. Aunque podrían examinarse con mayor detalle las causas de las imprecisiones en el error ventaja, la elección de 2012 visibilizó también el uso irresponsable de la información de las encuestas por parte de los medios de comunicación, su manipulación, e incluso, la falta de cumplimiento con los ordenamientos legales (Abundis *et al.*, 2014, 24-25). Se evidenció también la necesidad de actualizar el marco regulatorio y se sentó un precedente -previo a la explosión del uso de redes sociales- para la ciudadanía, sobre la necesidad de aprender a leer las encuestas y demandar la utilización de criterios científicos en su realización.

Gráfica 5. Error distancia primera y segunda fuerza: elección presidencial 2012

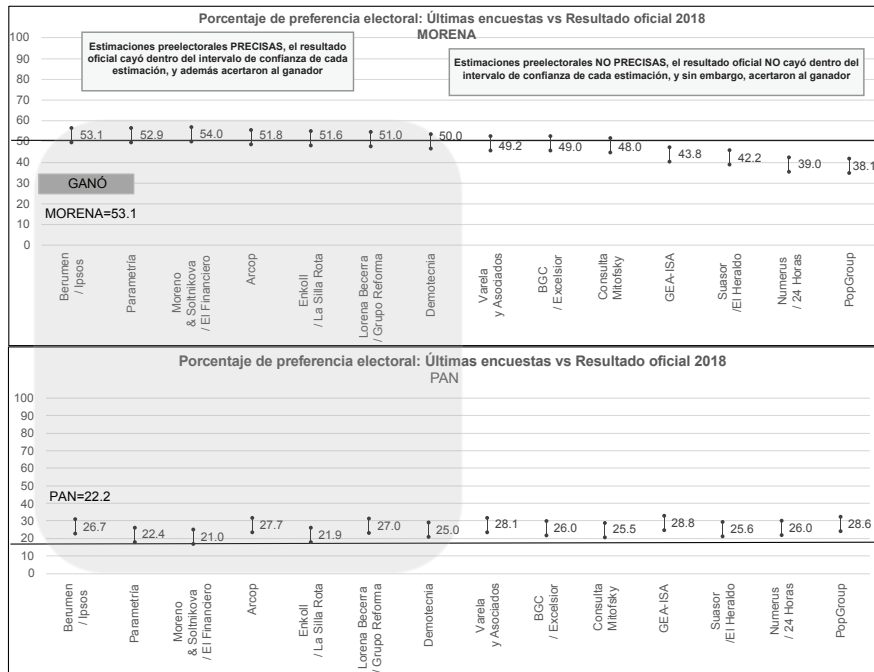


Fuente: Elaboración propia con las fuentes de datos de encuestas y resultados electorales ya mencionados.

Finalmente, en las elecciones de 2018, si bien todas las encuestas estimaron que AMLO ganaría la elección, nuevamente se observó un sesgo a subestimar la preferencia por López Obrador por parte de la mayoría de las casas encuestadoras. De las doce encuestas finales, tres obtuvieron un error menor de uno por ciento entre la encuesta y el porcentaje de votos obtenido por AMLO: Berumen/Ipsos, Parametría, y el diario El Financiero, dos más, Arcop y Enkoll, tuvieron un error alrededor de dos puntos porcentuales, y dos más, Grupo Reforma y Demotecnía estuvieron cerca de los tres puntos porcentuales de error. Siete mediciones el total quedaron dentro del intervalo de confianza. Sin embargo, debido a que el candidato opositor al partido en el poder, AMLO, ganó abrumadoramente la elección, pasó desapercibido para la opinión pública y los medios de comunicación el hecho de que

otras siete casas encuestadoras fueron imprecisas, con errores de entre el cuatro y hasta quince puntos porcentuales de subestimación de la preferencia por AMLO.

Gráfica 6. Error distancia primera y segunda fuerza: elección presidencial 2018



Fuente: Elaboración propia con las fuentes de datos de encuestas y resultados electorales ya mencionados.

En resumen, existe un contraste importante entre el mensaje global que mandaron como industria las encuestas en cada elección, y su desempeño técnico. En 1994, el mensaje fue de certidumbre en cuanto a que el candidato del PRI ganaría con una cómoda ventaja, sin embargo, la precisión de las estimaciones fue subóptima. En el 2000, no hubo un mensaje consensado, aunque predominó el mensaje de que ganaría el PRI ya que se subestimó la preferencia por el PAN. En el 2006, aunque los números arrojaban empate técnico, el mensaje predominante fue que ganaría AMLO, quien fue sobre estimado ligeramente por más encuestadoras. En el 2012, el mensaje debió ser similar al de 1994, consenso sobre el triunfo del PRI. Pero como las encuestas

más públicas sobre estimaron a EPN, su triunfo generó polémica. Finalmente, el 2018 fue un poco más parecido a 1994 en el sentido del consenso hacia el ganador. Y fue también parecido en la baja precisión puntual de las cifras, aunque este fenómeno ya no fue tan generalizado como en 1994.

Criterio 3: Distancia entre la primera y la segunda fuerza

Esta métrica es de las más populares ya que integra en la misma tanto la estimación del ganador, como de lo cerrado o abierta que se estima la competencia. Varias de las observaciones realizadas con el criterio 2 se refuerzan al analizar la gráfica 7. Y añadimos nuevos hallazgos. Para empezar, se observa un ligero empeoramiento en la precisión técnica de las estimaciones previas a cada elección, esto medido mediante el porcentaje de encuestas técnicamente precisas, en términos de la distancia entre los punteros, frente a las que se salen de los intervalos de confianza.

Para clarificar este criterio, usemos como ejemplo 1994. Ese año el candidato del PRI le ganó al candidato del PAN por 23 puntos porcentuales. Ernesto Zedillo obtuvo el 50.2% de los votos efectivos, contra 26.7% que obtuvo Diego Fernández de Ceballos. El periódico Reforma estimó que Zedillo ganaría con una ventaja de 40 puntos porcentuales, lo cual sobre estimaba la diferencia en favor de Zedillo en 17 puntos. En contraste, la encuesta de MORI estimó acertadamente una ventaja de 25 puntos. Una sobre estimación de dos puntos que entran perfectamente dentro del intervalo de confianza de la encuesta. En conclusión, en esta métrica se aprecia la sobre estimación reportada hacia el candidato del PRI.

Así, siguiendo esta lógica, tenemos que en 1994 sólo dos de ocho encuestadoras estimaron con precisión estadística la ventaja del primero sobre el segundo lugar de la contienda. Las otras seis sobre estimaron dicha diferencia, en todos los casos en favor del candidato del PRI.

En el año 2000, doce de dieciséis encuestas previas sobre estimaron a Francisco Labastida del PRI. Una sobre estimó la ventaja que lograría

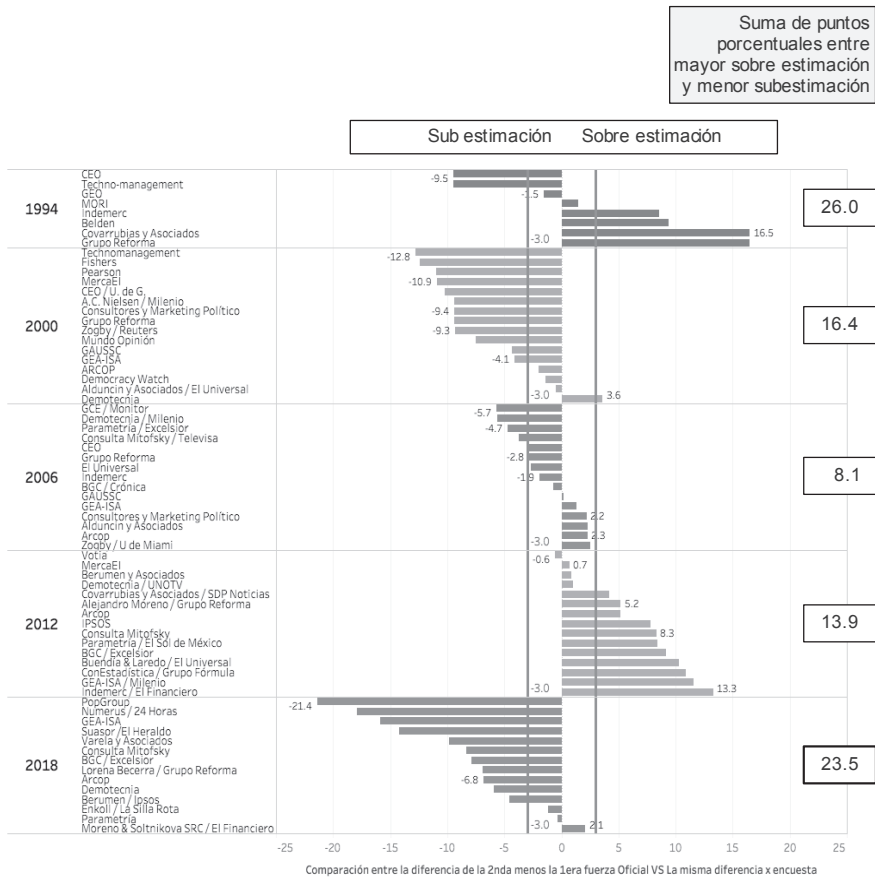
Vicente Fox sobre Labastida. Únicamente las encuestas del Arcop, Al-duncin y Asociados en El Universal y Democracy Watch lograron estimar la ventaja de Fox dentro del intervalo de confianza estadístico. La falta de precisión técnica de la mayoría de las encuestas en esta ocasión fue más un tema de discusión dentro del gremio que de generar una polémica en la opinión pública.

Esta métrica se vuelve más relevante entre más cerrada es una elección. Y en este sentido el desempeño técnico de la mayoría de las encuestas en la elección de 2006 fue de alto nivel. Once de las quince encuestas realizadas indicaban un empate técnico entre los candidatos punteros. Incluso las cuatro mediciones que se salieron del intervalo de confianza establecido en el presente capítulo lo rebasaron por porcentajes mucho menores a los observados en otros años. Como se mencionó en el apartado anterior, éste logro de las mediciones electorales en México se vio opacado por la forma en que se comunicaron y politizaron los resultados de las encuestas.

Las elecciones de 2012 y 2018 volvieron a contrastar con ese buen desempeño logrado en 2006. En 2012, a excepción de cuatro casas encuestadoras, Votia, Mercaei, Berumen y Asociados y Demotecnia, las otras once encuestas sobre estimaron la ventaja del candidato del PRI sobre el candidato del PRD. Y en 2018, salvo las encuestas de Parame-tría, Enkoll y El Financiero, las once restantes subestimaron la preferencia por el candidato de Morena sobre la del PAN.

En suma, con excepción de las mediciones realizadas en 2006, en las otras cuatro elecciones presidenciales bajo estudio, en promedio sólo una de cada cuatro encuestas publicadas tuvo la precisión estadística a la hora de estimar la distancia entre la primera y la segunda fuerza electoral. Estos datos no nos hablan de una evolución hacia mediciones más precisas, sino de un estancamiento en la profesionalización del gremio encuestador.

Gráfica 7. Error distancia entre primera y segunda fuerza electoral



Fuente: Elaboración propia con las fuentes de datos de encuestas y resultados electorales ya mencionados.

Un último punto para destacar de la gráfica 7 es que, si observamos la distancia entre los extremos de las estimaciones, podemos deducir con ese sólo número la falta de precisión en la medición de la ventaja en cada año. Si partimos de la base de que las diferencias en general no deberían de superar el intervalo de confianza, esto es, no deberían de ser mayores a los seis puntos porcentuales, el ya mencionado +/- 3%, vemos que el peor año fue 1994, con una distancia de 26 puntos porcentuales entre la mayor sobre estimación y la mayor subestimación en el conjunto de encuestas para ese año. El segundo peor año en este sentido fue el 2018, con 23.5 puntos porcentuales de apertura.

Los años de precisión intermedia serían el 2000 con 16.4 puntos porcentuales, y el 2012 con 13.9 puntos. Paradójicamente, el año electoral con mayor precisión técnica de las encuestas, el 2006 con una distancia de sólo 8.1 puntos porcentuales (no muy lejos de los 6 aceptables), fue también el más polémico y crítico políticamente, y junto con el 2012, los que más han golpeado la credibilidad de las encuestas electorales en México.

Otros análisis (Moreno 2018) coinciden en señalar que uno de los aspectos más interesantes para la demoscopia en México, visibilizados en la cerrada elección de 2006, fue la pregunta de si acaso el tamaño de la ventaja entre los contendientes es un factor que influirá también en el tamaño del error, sin que éste realmente muestre qué tan precisas fueron las encuestas.⁶ De esta manera, es posible observar que las elecciones presidenciales de mayor controversia por sus resultados (años 2006 y 2012) son las que presentaron errores promedio (del total de encuestas publicadas) menores. Tomando en cuenta la serie de datos para las elecciones de México desde 1994 y hasta 2018, y concentrándonos en las elecciones presidenciales, las elecciones menos cuestionadas (años 1994, 2000 y 2018) presentaron mayores errores promedio que las referidas anteriormente. Ergo, no se observa una correlación directa entre el nivel de precisión de las encuestas y la evaluación que de las mismas hace la opinión pública.

Algunas conclusiones preliminares con base en la evidencia mostrada hasta aquí, son:

- a) La mayor utilidad que han demostrado tener las encuestas es a nivel de acertar al ganador. De las cinco elecciones, podemos decir que en cuatro las encuestas midieron lo que iba a pasar. Porque incluso en la elección cerrada de 2006, el mensaje de las encuestas era un acertado empate técnico. La única excepción fue la elección del 2000, cuya mejor explicación pareciera ser la espiral del silencio, o el voto oculto, frente a la envergadura del cambio de régimen que implicaba la derrota presidencial del PRI, por primera vez en 70 años.

⁶ Adicionalmente, Moreno (2018) señala que el número de opciones políticas influirá también en el tamaño del error, de forma que mientras más contendientes existan el error será menor.

- b) En contraste con el criterio de acertar al ganador, la evaluación de la precisión estadística de las encuestas arroja un balance más bien negativo. Descontando la extraordinaria medición de las preferencias en el 2006, el porcentaje promedio de encuestas que miden con precisión estadística el porcentaje de votos que obtendrá la primera fuerza, es de 38%. Esto es, solo 1 de cada 3 encuestadoras por elección han medido con precisión al ganador. Y este dato es tan bajo en 2018 como en 1994.
- c) En consonancia con esta baja incidencia de encuestadores con mediciones precisas, se observa que, nuevamente excluyendo el 2006, la precisión para medir la diferencia de puntos porcentuales entre la primera y la segunda fuerza electoral es aún más baja que la métrica anterior. En este caso, solo 1 de cada 4 encuestadores han medido con precisión estadística esta distancia. Y al igual que la métrica anterior, no se observa ninguna mejoría entre 1994 y 2018.

Tabla 1. Porcentaje de aciertos en los tres primeros criterios

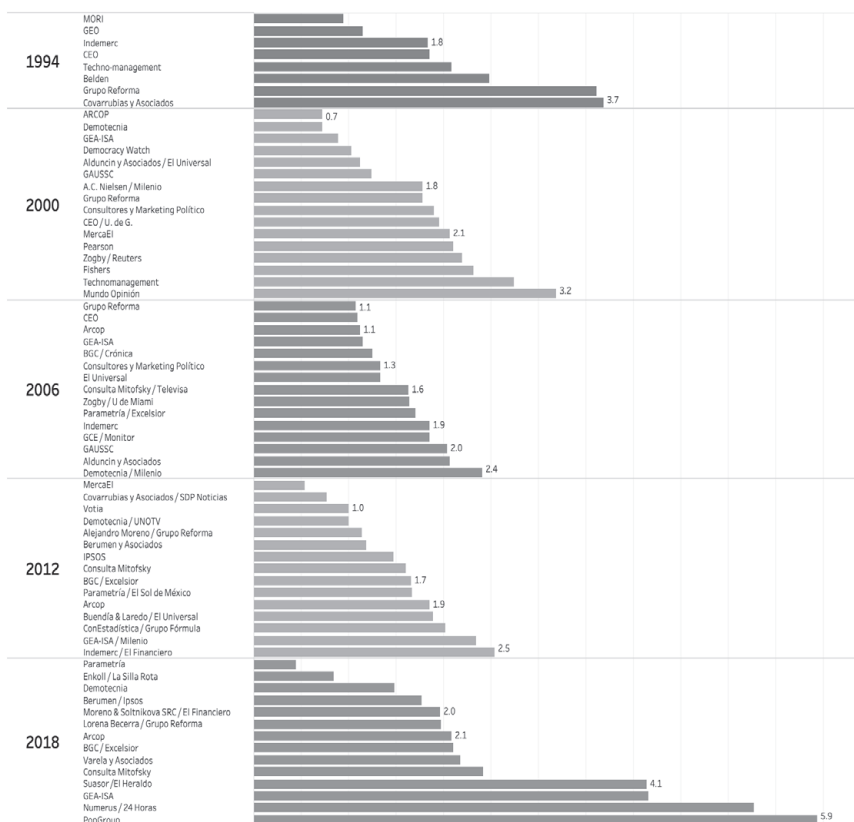
Hallazgos sobre la evaluación de la precisión de las encuestas			
Criterio	1	2	3
Definición	Aciertos al ganador	Precisión 1era fuerza	Precisión diferencia de saldos en 1era y 2da fuerzas
Cálculos del indicador	% de aciertos entre el total de encuestas finales	% de estimaciones con precisión de la 1era fuerza electoral, entre el total de encuestas finales	% de estimaciones con precisión de la diferencia de saldos entre 1era y 2da fuerzas, entre el total de encuestas finales
1994	100	38	25
2000	38	38	25
2006	47	73	73
2012	100	40	27
2018	100	36	21
Promedio global	77	45	34

Fuente: Elaboración propia con las fuentes de datos de encuestas y resultados electorales ya mencionados.

Criterio 4: El error total promedio de cada encuesta.

Este último criterio por presentar es el más técnico, el que mide con mayor rigor la precisión de una encuesta. No sólo viendo si coincidió con el ganador, o la distancia entre el porcentaje de preferencia obtenido por el ganador versus la encuesta, o bien, la distancia entre los dos punteros. En el caso del criterio número 4, entran al juego los porcentajes de todos los partidos políticos que hayan contendido en la elección. Como se mencionó más arriba, entre menor sea el error promedio total de cada encuesta, puede decirse que mayor fue su precisión con relación al resultado oficial de la elección con relación a cada fuerza electoral. Quizás esta métrica no es tan relevante como las otras en cuanto a su relación con los medios de comunicación, la opinión pública o la dinámica electoral *per se*. Sin embargo, puede constituir una forma de evaluar el desempeño de las diferentes casas encuestadoras ya que, si ha sido difícil tener precisión con uno de los porcentajes, y más con dos de ellos, tener precisión en cuatro (los partidos pequeños se sumaron en otros) lo es aún más. Esta puede servir más bien para que tanto el público en general, como otros tipos de usuarios y consumidores de encuestas, puedan tener una referencia sólida de quien es quien en las encuestas: poder identificar quienes tienen mayor trayectoria, mejor récord histórico de precisión, quienes han logrado las mejores mediciones en las diferentes elecciones bajo estudio.

Gráfica 8. Error total promedio por encuesta



Fuente: Elaboración propia con las fuentes de datos de encuestas y resultados electorales ya mencionados.

La siguiente tabla resume quienes lograron el mejor desempeño en cada una de las cinco últimas elecciones, usando como criterio de corte aquellos que hayan logrado un error promedio total de 1.2 puntos porcentuales o menos. Los resultados hablan por sí solos:

Tabla 2. Casas encuestadoras con mayor precisión en el error total promedio (1994-2018)

	1994	2000	2006	2012	2018
MORI	0.9				
GEO	1.2				
Demotecnia		0.7			
Arcop		0.7			

	1994	2000	2006	2012	2018
GEA-ISA		0.9			
Democracy Watch		1.0			
Alduncin y Asociados		1.1			
GAUSSC		1.2			
Reforma			1.0		
CEO			1.1		
Arcop			1.1		
GEA-ISA			1.2		
Mercaei				0.5	
Covarrubias				0.8	
Votía				1.0	
Demotecnia				1.0	
Reforma				1.1	
Berumen y Asociados				1.2	
Parametría					0.4
Enkoll					0.8

Fuente: Elaboración propia con las fuentes de datos de encuestas y resultados electorales ya mencionados.

La regulación de las encuestas electorales

Si bien es un debate abierto, podemos señalar que una de las principales motivaciones en la regulación de las encuestas parte de la suposición de que éstas tienen un efecto en el comportamiento político y pueden incidir en las preferencias de los electores. Partiendo de esa premisa y observando las discrepancias descritas arriba entre las estimaciones de una casa encuestadora y otra, el terreno es fértil para la sospecha de que lo que se publica como encuesta es en realidad propaganda inventada, disfrazada de ejercicio científico. La autoridad electoral buscó entonces mecanismos para tratar de controlar esa situación de lo que hoy se conocería como *fake news*, en la década de los noventa y la primera de este siglo serían encuestas falsas o “patito”.

Con la creación del IFE en 1990, se contempló la necesidad de regular las encuestas electorales y esto se hizo en 1993 incluyendo disposiciones específicas en una nueva ley electoral, el Cofipe. Esta primera regulación, podríamos decir que se limitó a tratar de sustentar la información de las encuestas mediante la obligación para las encuestadoras de entregar los estudios realizados durante la campaña al IFE, esto sin especificar mayores requerimientos. Más adelante, en el año de 1996, la legislación también incluyó que las encuestas sobre preferencias electorales publicadas y entregadas al IFE deberían adoptar y demostrar criterios científicos en cuanto a su metodología y muestreo.

La creciente competitividad en las elecciones, las transformaciones en los medios de comunicación y la opinión pública, entre otros factores, generaron incentivos para aumentar la regulación para las encuestas electorales. Debe destacarse que el IFE no se fue por la libre a la hora de definir los criterios científicos que deberían incluirse en los requisitos que demandaran, sino que, por ejemplo, en el año 2008, el gremio de los encuestadores estuvo involucrado en la definición de dichos criterios científicos. Cabe mencionar que el papel de la autoridad electoral en la regulación de las encuestas electorales ha sido, en palabras de Lorena Becerra, actual encuestador del Grupo Reforma, “reactivo y poco anticipatorio” (Becerra 2017). Actualmente esto se refleja en la ausencia de una regulación de cara a los nuevos tipos de encuestas que la tecnología permite realizar, pero que justamente carecen de los criterios científicos mencionados.

Los principales instrumentos en la regulación son la obligación de entregar al IFE el estudio en cuestión, con sus resultados, características técnicas (diseño muestral, cuestionario utilizado, tasa de rechazo, etc.), la base de datos, la especificación de quién solicitó, quién pagó, y quién realizó la encuesta, así como cuál fue el costo de esta.

Si bien los criterios mencionados han favorecido la desaparición de las encuestas patito, y de alguna manera la profesionalización de la industria, ya que, al ser públicos los estudios y la metodología, cada agencia queda expuesta a una auditoría, no sólo de la autoridad electoral, sino de sus colegas, de la academia, y de cualquier persona interesada en el tema. De esta forma, el uso de malas prácticas se ha ido desterrando. Un elemento positivo adicional fue el poder bajar el “periodo de veda” de siete a tres días naturales antes de la elección, relacionado con la autorización para publicar encuestas. Aun así, el establecimien-

to de periodos de veda, que para México sigue siendo de tres días, ha hecho que entre el periodo de levantamiento y la publicación de la última encuesta antes de la veda, los últimos resultados reportados por las casas encuestadoras reflejen las preferencias de aproximadamente diez días antes de la elección. Desde el punto de vista de las casas encuestadoras ésta sigue siendo una asignatura pendiente ya que entre más cercano sea el levantamiento de la encuesta al día de la elección, mayor probabilidad hay de que ningún evento modifique las preferencias estimadas de las que realmente ocurren el día de la elección. Cabe mencionar que aproximadamente la mitad de los países del mundo donde se llevan a cabo elecciones, no tienen ningún periodo de veda, ver el estudio de WAPOR “La libertad para publicar estudios de opinión” (<http://wapor.org/freedom>). La siguiente tabla resume los rasgos principales de la evolución de la regulación de encuestas en México.

Tabla 3. Principales hitos en la reglamentación de encuestas electorales en México (1993-2014)

Año	Suceso	Ley o reglamento	Características
1993	Inicio de la regulación de encuestas	Código Federal de Instituciones y Procedimientos Electorales	Obligación de entregar al IFE el estudio que respalde los resultados de las encuestas electorales publicadas.
1996	Cambios en la normatividad del Cofipe	Código Federal de Instituciones y Procedimientos Electorales	Inclusión de disposiciones para que toda persona física o moral que pretenda llevar a cabo y publicar encuestas por muestreo sobre preferencias electorales adopte criterios de carácter científico para su realización.
2008	Reformas al Cofipe	Código Federal de Instituciones y Procedimientos Electorales	Se agrega la obligación de consultar al gremio de la demoscopía sobre la definición de los criterios científicos aplicables a una encuesta.
2014	Reforma político electoral (Normatividad vigente)	Establecimiento en el capítulo VII del Reglamento de Elecciones de la Ley General de Instituciones y Procedimientos Electorales	Determinación de las principales reglas para las encuestas electorales, las cuales se especifican de manera muy clara como obligación en el reporte de los estudios y que incluye: Criterios científicos en las encuestas por muestreo o sondeos de opinión; Veda electoral de tres días anteriores a la elección; Entrega de copia del estudio con la información que la sustente; Registro de personas físicas o morales interesadas (patrocinadores, empresas y solicitantes de su difusión); Especificación de metodología, diseño de cuestionario y copia del cuestionario aplicado / Monitoreo de publicaciones impresas; y costo total del estudio, entre otras.

Fuente: Portal del INE (2017).

La discusión: consensos y disensos sobre las causas de la falta de precisión

En un mundo ideal, todas las encuestas electorales públicas realizarían sus estudios siguiendo rigurosamente los cánones metodológicos, cuidando cada uno de sus procesos, y todos obtendrían resultados dentro de cuyos intervalos de confianza se ubiquen las preferencias de los resultados electorales oficiales. Sin embargo, en el mundo real, como lo acabamos de constatar, las discrepancias estadísticas se dan en dos de cada tres encuestas con relación a la estimación de la primera fuerza electoral, y en tres de cada cuatro encuestas cuando lo que se mide es la distancia entre el primero y segundo lugar en la contienda electoral. Abajo intentaremos sintetizar los argumentos que se han esgrimido en las discusiones dentro y fuera del gremio de los encuestadores y académicos. Repasaremos primero aquellos puntos donde se ha llegado a un cierto consenso, y luego a puntos en donde dicho consenso no existe.

Causas de las discrepancias: consensos

a) Error en la muestra: se trata de un elemento esencial para poder garantizar que los resultados de una encuesta sean representativos de la población, esto es, hay consenso en que se debe utilizar una muestra probabilística que, por definición, le dé una probabilidad conocida y diferente de cero a cada ciudadano mexicano con credencial de elector. Y que la muestra sea seleccionada utilizando el principio de aleatoriedad en cada etapa. Más allá de eso, la propia legislación en la materia menciona que los estudios sobre preferencias electorales deben estar basados en criterios científicos, y entre ellos, se destaca el correcto diseño muestral de acuerdo con los objetivos de la investigación en cuestión.

Centrándonos como lo ha sido en nuestro caso de estudio (elecciones presidenciales en México), un diseño muestral adecuado debiera tener como marco muestral las secciones electorales del INE, un tamaño de muestra y una dispersión geográfica que abarque secciones urbanas, rurales y mixtas acordes a los objetivos del estudio, y múltiples etapas en la selección de unidades muestrales siguiendo mecanismos aleatorios en cada etapa. También debieran de quedar claras las

probabilidades de selección asignadas a las diferentes unidades (para asegurarse que se apliquen los ponderadores correctos) y que se defina el efecto de diseño esperado derivado del diseño muestral específico utilizado.

b) Error en el instrumento: El **fraseo** de las preguntas suele ser uno de los elementos que puede tener mayor probabilidad de sesgar las respuestas de los entrevistados. También el **orden** en el cual se realiza la pregunta electoral pudiera incidir. Esto es particularmente delicado en el caso de las encuestas preelectorales, pues el diseño del instrumento y la labor del encuestador debe garantizar siempre la secrecía cuando se utiliza el método de boletas. Es decir, que se le ofrezca un mecanismo de que el entrevistado vote de manera confidencial. La entrega al INE de los cuestionarios de los estudios permite analizar en qué medida el error total promedio por encuesta se vio afectado por el fraseo, el lugar de la preferencia electoral en el cuestionario y el grado de confidencialidad otorgado al respondiente a la hora de manifestar su preferencia electoral.

c) Error en el levantamiento de campo: Dos son las principales fuentes de error en esta área. Primero, deficiencias en la logística de campo o en la capacitación de los encuestadores. Segundo, una mala supervisión de campo. La consecuencia de lo primero es una mala selección de los entrevistados que cayeron en la muestra, una mala dinámica para lograr la aceptación de la entrevista, lo que afecta el tamaño del rechazo, o bien una aplicación inadecuada del cuestionario. Estas fallas pueden corregirse mediante la supervisión *in situ*. La segunda fuente, la mala supervisión, puede traer consecuencias mucho más graves. Desde la imposibilidad de corregir errores como los mencionados, hasta la imposibilidad de detectar encuestadores fraudulentos, aquellos que en vez de hacer su trabajo de ir a la localidad designada, a la sección especificada, a la manzana seleccionada, al hogar especificado y seleccionar a la persona designada, y aplicar el cuestionario adecuadamente al entrevistado, deciden hacer las cosas a su manera, y esto puede implicar desde que se pongan a aplicar el cuestionario en puntos de afluencia dentro de la sección electoral, hasta el falseamiento total de las entrevistas, llenándolas en el hotel o en su casa! Este último es el más grave y dañino para la calidad de una encuesta. Y, hasta hace pocos años era un problema ampliamente extendido en el medio de los

encuestadores de a pie, tanto para temas electorales como de mercado. De hecho, es la hipótesis del autor, que la principal fuente de error de la sobre estimación de las encuestas en las elecciones de 2012, fue ésta. Que el volumen de trabajo que tuvieron las encuestadoras más visibles fue tal, que perdieron el control en la supervisión de su campo. Y que esto generó un círculo vicioso: los encuestadores falseadores suelen tratar de replicar lo que ven en los medios de comunicación, y como lo que veían era que Peña Nieto tenía una amplia ventaja, el falseamiento hacía que se replicara esa ventaja. Creándose así un círculo vicioso de una verdad auto construida y auto alimentada. Evidentemente es difícil de probar, pero cuento con elementos para creer que esta hipótesis no es errada.

Afortunadamente este problema se encuentra en vías de extinción. Uno de los elementos más positivos que han traído las nuevas tecnologías al ramo de los estudios de mercado y opinión pública ha sido la utilización de tabletas y software CAPI (*Computer Assisted Personal Interviewing*) para el levantamiento de información. Ello ha permitido reducir a un nivel mínimo el problema de falseamiento de datos por parte del entrevistador. Esto gracias a que mediante este procedimiento es posible hacer una “supervisión virtual” mediante la cual es posible seguir remotamente y en tiempo real los movimientos de cada encuestador (vía GPS), es posible tener toda otra serie de métricas como el tiempo que tarda en aplicar un cuestionario, o entre un cuestionario y otro, y sobre todo, es posible recopilar evidencia adicional de su desempeño, ya sea mediante fotos del entrevistador al aplicar la entrevista, y la grabación parcial o total de la misma entrevista. Con toda esta riqueza de información, es posible no solo erradicar los cuestionarios falseados, sino supervisar el desempeño y corregir cualquier proceso que esté realizando mal el entrevistador en campo. Gracias a esta tecnología la calidad y precisión de las encuestas cara a cara en vivienda tendrá una indudable mejoría.

Causas de las discrepancias: disensos

a) Error en el análisis por aplicación de modelos: ¿Se deben o no utilizar modelos para reportar la preferencia electoral? Sin ser un consenso absoluto, la práctica mayoritariamente aceptada en el gremio

es el reportar las frecuencias simples (luego de aplicar el ponderador muestral correspondiente), incluyendo el porcentaje de aquellos que no contestaron la pregunta de preferencia electoral. Y esto junto con lo que se conoce como “preferencia efectiva”, esto es, se elimina la no respuesta y se recalculan los porcentajes a 100%. Sin embargo, hay encuestadores que prefieren reportar un modelo propio de “votantes probables”, donde mediante diferentes preguntas filtro, estiman la preferencia sólo de aquellos que creen que tienen una alta probabilidad de ir a votar. Otros encuestadores diseñan modelos aún más sofisticados utilizando diversas variables incluidas dentro del cuestionario, e incluso información electoral histórica. Lo cierto es que ningún modelo ha demostrado ser sistemáticamente infalible a la fecha. Y lo peor, muchas veces no se reporta de manera precisa cómo se realizó la estimación. En estos casos al menos deberían de reportarse las frecuencias simples, y el modelo de estimación propietario. De esta forma habría una base de comparabilidad más justa. Actualmente se compara lo que reporta cada casa encuestadora, ya sea la preferencia efectiva o la estimación de su modelo. Un ejemplo de este fenómeno se dio en las elecciones de 2012, donde cuatro de las encuestas públicas reportaron con modelos: Demotecnia, BGC, Con Estadística e Indermerc. De éstas, únicamente la encuesta de Demotecnia situó sus estimaciones dentro del intervalo de confianza esperado. El resto de las encuestas que utilizaron modelos sobreestimó las preferencias para el candidato Enrique Peña Nieto entre 5 y 9 puntos porcentuales. Aquí el peligro puede provenir de la falta de transparencia.

b) ¿Falseamiento de datos por parte del entrevistado, espiral del silencio o miedo institucional?: Uno de los debates más añejos acerca de la precisión de las estimaciones de las encuestas preelectorales, y que es anterior a los aspectos técnicos y metodológicos, es el de la veracidad de las respuestas por parte de los entrevistados. Aunque para algunos pueda ser una discusión fútil porque difícilmente podrá evidenciarse, las experiencias a nivel nacional e internacional parecen confirmar la importancia de su análisis. Las amplias diferencias registradas entre los sondeos preelectorales y los resultados de las votaciones podrían encontrar su explicación en la llamada espiral del silencio, un concepto sociológico que denota que la opinión pública tenderá a expresar en las encuestas las opiniones percibidas como mayoritarias,

aunque esté en contra de ellas y actúe diferente a la hora de votar. La motivación principal para este comportamiento es la fuerza de la conformidad al grupo, o el miedo a la exclusión del mismo grupo social. (Noëlle-Neumann 1995)

Uno de los casos más paradigmáticos referentes a esta cuestión lo ilustra el referéndum en Colombia sobre los acuerdos de paz en el año 2016, en donde las divergencias entre los sondeos y los resultados finales ¡llegaron a casi 20 puntos! (Palomino 2016) En un tema que genera grandes consensos como la paz, los votantes probablemente no expresaron en las encuestas preelectorales sus verdaderas intenciones de voto por miedo al rechazo social.

Aunque se trata de un argumento recurrente para justificar los sesgos en las estimaciones, parece ser relevante específicamente en lo tocante a temas que suscitan gran controversia, ¿A qué se deben estos sesgos?, es, como se mencionó, ¿Un problema de respuestas socialmente aceptables? ¿Se trata de un voto oculto?, ¿Hay falseamiento hacia el entrevistador por temor a perder un beneficio ofrecido por un partido o institución gubernamental? ¿Existe movilización del voto a último minuto?, ¿Hay segmentos de votantes que no son captados por las encuestas?, o simplemente, ¿El entrevistado proporciona información falsa? Todas estas cuestiones deben tomarse en cuenta como factores que puedan afectar la precisión de las estimaciones en encuestas preelectorales y que, a la fecha, no existe consenso entre los académicos y los practicantes de la opinión pública sobre cuando un fallo en las encuestas se debe a este fenómeno de la opinión pública, y cuándo dichas fallas tienen un origen distinto.

c) *¿Quién paga manda? O el uso propagandístico de las encuestas:* Uno de los temas que causan mayor suspicacia, en todo el mundo es la credibilidad de las encuestas. Bien podemos afirmar, que parte del descrédito de las encuestas electorales se debe a la difusión de encuestas “a modo”, cuyo objetivo es simplemente propagandístico. Aunque esto se encuentra dentro de la legalidad (cuando se trata de metodologías válidas), la publicación de estos resultados causa confusión en la opinión pública, la cual difícilmente puede diferenciar entre encuestas patrocinadas y estudios independientes. Aunque las regulaciones han ido cambiando con el objetivo de evitar tales confusiones, vale la pena resaltar la labor de profesionalización que realizan las asociaciones del

gremio, tales como la AMAI (Asociación Mexicana de Inteligencia de Mercados) o el CEDE (Colegio de Especialistas en Demoscopia y Encuestas). Otros aspectos de gran relevancia han sido la recopilación de estudios, con estándares metodológicos por parte del INE y de los institutos electorales locales, la cual ha puesto candados para la publicación de encuestas “patito”. Podemos decir que este problema ha sido erradicado, al menos de las últimas publicaciones previas a la elección presidencial.

Indudablemente, estos avances en la regulación de encuestas pre-electorales han podido detener la difusión de encuestas espurias cuyo objeto es la propaganda. Sin embargo, nuevos dilemas sobre el uso propagandístico de las encuestas se ciernen sobre la capacidad de las instituciones para regularlas. En primer lugar, el hecho de que, aunque no se reporten como auspiciadores, ciertas encuestas responden a intereses particulares y los patrocinadores permanecen ocultos. En segundo lugar, los nuevos métodos de recolección originan recovecos que las regulaciones actuales no pueden alcanzar. Así las nuevas tecnologías y las transformaciones de la opinión pública llevan un paso adelante. Es necesario, al respecto afianzar los mecanismos institucionales para sancionar las malas prácticas en la materia.

Independientemente de estos factores, no parece que sea un tema el evitar el uso propagandístico de las encuestas, puesto que será intrínseco al ritmo y el propósito de las campañas electorales; es labor de los medios de comunicación y de la opinión pública definir los criterios para su publicación y/o difusión, el conocimiento de los estándares de calidad y los alcances y las limitaciones de cada ejercicio y/o metodología.

d) Las encuestas preelectorales y los resultados de la elección no tienen por qué parecerse. La encuesta mide a todo el electorado, no a quienes van a votar. Este argumento parte de una lógica que en principio es correcta pero que empíricamente, al menos en la mayor parte de las ocasiones, no se cumple. Para que fuera válido este argumento tendría que observarse un caso en el que **todas** las últimas encuestas preelectorales estiman un ganador por cierto margen, y el resultado es diferente, gana otro candidato cuyo porcentaje de votos no se intersecta siquiera con el intervalo de confianza de dichas encuestas. En los cinco casos de estudio aquí analizados, en todos ellos hubo algunas casas

encuestadoras que estimaron correctamente, dentro del intervalo de confianza, la distancia hacia el primer lugar, la diferencia entre el primero y el segundo lugar, e incluso el error promedio total de la encuesta fue cercano o menor a un punto porcentual. Entonces, el hecho de que algunas casas encuestadoras sí se encuentren dentro del intervalo de confianza y otras no puede deberse a otras razones, mas no debido al argumento aquí esgrimido.

e) Las encuestas preelectorales no predicen el resultado de una elección porque son una fotografía del momento de un fenómeno dinámico. Este argumento es, como el anterior, igualmente lógico y en teoría correcto. Sin embargo, es necesario complementar esta teoría con una premisa adicional, esto es, a mayor distancia del día de la elección, mayor es la probabilidad de que nueva información, eventos inesperados, o la misma dinámica de las campañas electorales, provoquen cambios en las preferencias electorales de los ciudadanos. Y en ese sentido, entre más lejos de la elección se tome la fotografía, mayor probabilidad existe de que la misma sea diferente del resultado final. Pero, bajo la misma lógica, la probabilidad de nueva información o eventos que puedan modificar la preferencia será menor entre más cerca se mida de la elección. Y, si no existe nueva información entre el día del levantamiento de la encuesta, alrededor de ocho días antes del día de la elección, y la jornada electoral, entonces no debería de haber diferencias estadísticamente significativas entre una encuesta cercana a la elección y el resultado oficial. Es por esta razón que para el presente capítulo se utilizaron únicamente las últimas encuestas publicadas. Nuevamente, la evidencia descarta la validez de este argumento como causa de las discrepancias entre las estimaciones de diferentes casas encuestadoras. Y, al igual que el punto anterior, sostenemos que tratar el error como una condición apriorística parece un argumento peligroso que no permite abonar a la discusión sobre la mejora continua en la precisión de los ejercicios demoscópicos.

f) A las encuestas electorales se les debe evaluar por su metodología, no por sus resultados. Existen criterios técnicos indispensables para la correcta realización de la encuesta, que van desde su diseño, como un tamaño de muestra suficiente, fraseos adecuados de las preguntas, hasta la capacitación de los encuestadores y las garantías de confidencialidad para los entrevistados, entre otras que impactarán decididamente

en la precisión de las estimaciones. La evaluación en la precisión de las estimaciones sí ha sido probada por la metodología que utilizan. Es de esta manera, como en México las encuestas en vivienda han demostrado ser las que presentan las estimaciones más confiables.

El argumento de que a las encuestas electorales se les debe evaluar por su metodología no por sus resultados tiene al menos dos vertientes. Por un lado, una encuesta que cumple con los cánones metodológicos aceptados por la ciencia, es una encuesta correcta. Pero puede darse el caso, como se dio en la elección de 2006, la estimación puede dar de ganador a un candidato, pero por estar en margen de error, el que gane otro candidato será correcto desde un punto de vista técnico, aunque la opinión pública lo condene.

En ocasiones se ha argumentado también que siempre existe el riesgo de que la muestra seleccionada haya caído en las colas de la curva normal. Es decir que no cayó dentro del 95% de confianza estipulado en la metodología, esto es, cayó en el 5% y por ende la estimación que resulte estará fuera del intervalo de confianza aceptado. En realidad, esta es una falacia, normalmente la persona encargada de seleccionar la muestra, la valida contra diferentes parámetros para evitar que haya obtenido una muestra fuera del intervalo de confianza, y segundo, escoge la que menor error tiene frente a esos parámetros.

Por lo tanto, un diseño y selección muestral metodológicamente correcto y validado, con la garantía de estándares de calidad en el resto de los procesos que componen una encuesta, deben traducirse en encuestas que representen la distribución de las opiniones y actitudes en la sociedad, es decir, deben ser precisas. Una muestra representativa garantiza, dentro del margen de error y el intervalo de confianza, que los resultados de la encuesta serán representativos de la población.

Aun así, a veces se cuestiona este argumento diciendo que, si las diferentes encuestas se realizan con base en diseños y metodologías similares, por qué presentan en ocasiones resultados tan dispares. Aquí la respuesta más evidente es que no todas las encuestas están apeguándose a los mismos estándares de calidad, un mal cuestionario, un campo descuidado, etc. arruinan, como hemos visto, la precisión de las estimaciones. Ante lo cual, la solución es que exista la posibilidad de auditar los procesos de las encuestas. El INE ha avanzado mucho en este camino, pero quedan elementos sueltos como la calidad de la re-

colección de datos en campo. En estos casos entonces, la calidad de los resultados será la consecuencia del rigor con que se llevaron a cabo los postulados metodológicos a la hora de realizar su encuesta cada casa encuestadora.

La segunda vertiente del argumento de la pre-eminencia de la metodología sobre los resultados, es que es posible tener precisión en los resultados habiendo realizado una encuesta metodológicamente incorrecta. Tal es el caso de las encuestas por Internet vía Facebook o por teléfono vía robot, las cuales, como veremos más adelante, no son estudios representativos, y sin embargo es posible que en ocasiones acierten a tener estimaciones correctas, como hace 20 años lo eran las encuestas en calle.

Nuevas fuentes de error: la amenaza digital

Para la elección presidencial del año 2018, la explosión de las nuevas tecnologías, el uso de redes sociales, el avance en la cobertura de Internet y, sobre todo, la telefonía celular, dieron pie a la expansión en México en el uso de nuevos métodos de recolección de información para encuestas los que, por desconocimiento popular, se han hecho pasar como estudios demoscópicos serios. Esta tendencia ha venido creciendo también en el resto del mundo.

Es importante decir que las encuestas telefónicas con robot y las encuestas por Facebook incumplen los principios básicos del muestreo probabilístico y por ende no son representativas de ninguna población. Por la importancia como amenaza a la conformación de un gremio de encuestadores en México que generen estimaciones metodológicamente sólidas, precisas y confiables, vale la pena repasar aquí lo que debe ser una encuesta de esta naturaleza y evidenciar las diferencias que existen frente a estos nuevos métodos.

La estadística inferencial estableció desde hace muchas décadas los postulados básicos para poder utilizar la información de un subgrupo de individuos, y poder generalizar las características encontradas en ese grupo al resto de la población, con un margen de error conocido y

derivado del mismo diseño de la muestra. Este subgrupo de individuos se le llamó muestra y la población bajo estudio se le llamó universo.

El principal postulado para poder hacer el tipo de inferencias mencionado es tomar la información de una **muestra probabilística**. En una muestra probabilística todos los individuos del universo de personas objeto de estudio deben tener una probabilidad conocida y diferente a cero de haber sido seleccionados en la muestra.

De este postulado se deriva que para que los resultados obtenidos de la muestra a la que se le aplica la entrevista **representen** al conjunto de la población que se quiere estudiar se debe, primero, contar con un marco muestral que incluya al 100% de los individuos que busca representar. Si lo anterior no se cumple, se pierde la posibilidad de inferir, de los resultados de la encuesta a la realidad de la población. A esta deficiencia en el marco muestral se le conoce como “error de cobertura” y entre más grande sea este error, menor será la representatividad de los hallazgos hacia, en este caso, los ciudadanos mexicanos. Lo más crítico es cuando las personas quienes no tuvieron la posibilidad de ser seleccionadas en la muestra, poseen características diferentes a las personas sí seleccionadas, y que ello implique diferencias en las opiniones de ambos grupos.

Si bien existen diferentes mecanismos para seleccionar la muestra de un universo de interés, desde la selección aleatoria simple hasta una muestra multietápica, por conglomerados, etc., el factor común es que ya sea en uno o en cada uno de varios pasos, la selección de unidades se hace de manera aleatoria. Y una vez seleccionada la muestra, esa debe ser la muestra de personas a ser entrevistadas. Para ello es necesario regresar varias veces para localizar a un entrevistado o entrevistada seleccionada. Para garantizar el cumplimiento de la muestra.

Entonces, lo primero que se requiere para hacer una encuesta es definir el universo de personas que nos interesa estudiar. Para el caso de la opinión pública, el universo de interés son todos los ciudadanos, es decir, todas las personas mayores de 18 años y que por tanto están legalmente autorizados a participar en los asuntos públicos del país, esto es, todos los ciudadanos en general, y en particular para temas electorales, todos los ciudadanos con credencial de elector vigente.

Siguiendo la definición mencionada, para que una encuesta de opinión en México sea representativa de la opinión de la ciudadanía

mexicana, debe estar basada en una muestra probabilística de los ciudadanos, ergo, *todos los ciudadanos deben tener una probabilidad conocida y no nula de ser seleccionados en la muestra*. Si esto no se cumple, nos encontramos entonces en lo que se conoce como “error de cobertura”. Todas las encuestas por Facebook tienen un error de cobertura de al menos 40%, ya que solo el 93% del 65% de ciudadanos que tienen acceso a Internet en México, usa Facebook. El error de cobertura en encuestas telefónicas es de alrededor del 20%, siempre y cuando se combine telefonía residencial y telefonía móvil. La primera sola tiene un error de cobertura de 55% a nivel nacional. Los investigadores deben ser conscientes de que a nivel estatal las coberturas mencionadas tienen variaciones muy fuertes. En todo caso, las diferencias de cobertura de Internet y uso de Facebook, o de uso de telefonía residencial y/o móvil, genera diferencias en los tipos de poblaciones a las que nos dirigimos. Diferencias en educación, ingreso y edad, mismas que se ha demostrado afectan las estimaciones sobre mediciones como la aprobación presidencial. (Abundis 2019)

El otro elemento clave, y aún más grave, es que la muestra, una vez seleccionada, es la que debe ser entrevistada. Si esto no se cuida, caemos en el conocido “error de auto-selección”. Esto ocurre cuando se deja abierta a la voluntad de los individuos participar o no en la encuesta. Este error elimina por completo la selección aleatoria, la selección probabilística de la muestra, y por ende aniquila cualquier posibilidad de inferir estadísticamente hablando, los resultados de la encuesta para describir la población que deseamos estudiar. Puede sobre representar a personas con mayor interés en política, o a otros tipos de personas, pero en cualquier caso deja de ser una encuesta representativa de los ciudadanos en su conjunto. El hacer pasar una encuesta como representativa de la ciudadanía, cuando fue recolectada mediante mecanismos de “auto selección”, es una de las cinco prácticas que condena abiertamente la Asociación Americana de Investigación en Opinión Pública (AAPOR por sus siglas en inglés).⁷ Tanto las

⁷ Las otras cuatro prácticas condenadas son: solicitar contribuciones o dinero a los respondientes de una encuesta, usar los contactos de una encuesta como prospectos de venta, relevar la identidad de un respondiente sin su consentimiento, y hacer “push polls”, esto es, telemarketing negativo contra un candidato, disfrazando la llamada como una encuesta de investigación legítima (AAPOR 2019).

encuestas por Facebook como las encuestas telefónicas por robot utilizan el mencionado mecanismo de “auto selección”.

Este debate de encuestas probabilísticas versus encuestas no probabilísticas es de total vigencia y preocupación para la industria demoscópica internacional. Una prueba muy informativa de esto es un número especial que sacó AAPOR en su revista *Public Opinion Quarterly*, donde el número entero está dedicado a un *Symposium* sobre este preciso tema. En uno de los artículos principales, que es luego comentado por varios especialistas, los autores realizaron una serie de experimentos para probar qué tipo de encuesta tenía mayor sesgo con respecto a los parámetros reconocidos, en este caso los parámetros fueron los datos del Censo. Probaron encuestas telefónicas probabilísticas con una tasa de respuesta baja, contra encuestas vía Internet mediante el uso de paneles, bases de datos con personas dispuestas a responder un cuestionario a cambio de una remuneración. Y dentro de éste se utilizaron diferentes métodos para ajustar la muestra desde ponderarla para compensar desbalances sociales, hasta el uso de propensity score matching para obtener una “muestra” apropiada. La conclusión central del estudio fue que, a pesar de la baja tasa de respuesta de la encuesta telefónica, pero probabilística, este método tuvo un sesgo dos y media veces menor que los levantamientos no probabilísticos, independientemente de los ajustes que se les buscó hacer (Dutwin y Buskirk 2017).

En congruencia con estos postulados metodológicos, la forma correcta de hacer investigación de opinión pública en México hoy en día es mediante encuestas cara a cara a una persona en su vivienda que haya sido seleccionado/a mediante una muestra probabilística de hogares. En México normalmente se utilizan como marcos muestrales las AGEBS del INEGI o las secciones electorales del INE.

Y a pesar de lo anterior, al observar la tabla 3, podemos notar cómo los sondeos a través de Internet han ido ganado terreno, especialmente, en los países desarrollados. Esto, por un lado, nos habla de que estas nuevas metodologías son aceptadas cada día más como referentes en la opinión pública en contextos determinados. Por citar un ejemplo, el porcentaje de encuestas por Internet en la elección presidencial de los Estados Unidos aumentó del 15% registrado en 2012 al 53% registrado en 2016.

Tabla 4. Penetración de nuevas tecnologías y su uso en la medición de la opinión pública

Elección/ Referéndum	Fecha	Periodo de análisis	Número de encuestas	% Web	% IVR (robot)	Uso de Internet %	Móviles con In- ternet por cada 100 habitantes 2017
EUA	-	-	-	-	-	87	132.4
Presidenciales 2012	06-nov-12	ene-nov 2012	406	15	33	-	-
Presidenciales 2016	08-nov-16	ene-nov 2016	407	53	12	-	-
Gran Bretaña	-	-	-	-	-	95	88.2
Referéndum Escocia	18-sep-14	ene-sep 2014	67	75	-	-	-
Elecciones parlamentarias	07-may-15	abr-may 2015	95	76	-	-	-
Brexit	23-jun-16	ene-jun 2016	127	71	-	-	-
Canadá	-	-	-	-	-	91	72
Elecciones parlamentarias	19-oct-15	ago-oct 2015	78	36	36	-	-
México	-	-	-	-	-	65	80
Elección presidencial	01-jul-18	Ene-jun 2018	-	-	-	-	-

Fuente: Elaboración propia y actualización con base en datos de Claire Durand (2017) y Banco Mundial (2019).

Sin embargo, existe una brecha digital entre los países, lo cual dificulta para efectos empíricos su comparación. En otras palabras, es difícil equiparar la utilidad de un sondeo por Internet en un país donde el uso de Internet es cercano al 100% como en los países de Europa Occidental a uno como México, en donde se ubica alrededor del 65% de las personas mayores de 18 años.

Cabe mencionar que los avances tecnológicos que representa la difusión y acceso a encuestas por Internet no han repercutido en un progreso en la certidumbre o infalibilidad de los métodos tradicionales y novedosos de recolección de la información. Esto quedó de manifiesto en diversas partes del mundo. En ocasión de procesos electorales o

referendo sobre cuestiones políticas cruciales, como lo fueron el Brexit, el plebiscito sobre los acuerdos de paz en Colombia, o las elecciones presidenciales de Estados Unidos -los tres casos en 2016- las diferencias entre las estimaciones de encuestas, tradicionales y no tradicionales, y los resultados, fueron causa de escarnio mundial y preocupación para las casas encuestadoras.

Las elecciones de 2018 en México abrieron paso a la expansión de estas técnicas de encuestas por Internet y de robots telefónicos. La lógica fundacional de algunos de sus promotores era que la universalización del uso de tecnologías como el teléfono celular, el acceso a Internet o el uso de redes sociales como Facebook, habían convertido a las encuestas en vivienda en una metodología obsoleta y costosa. Tras los fallos en la estimación del error ventaja registrados en la elección de 2012, así como en las elecciones para gobernador de los años 2015 y 2016, que se dieron en paralelo a varios aciertos al ganador de encuestas por Facebook o telefónicas Robot, algunos sectores de la opinión pública les dieron el beneficio de la duda. Pero en 2018, tomando como parámetro el error ventaja, las diferencias entre los distintos métodos fueron más evidentes. Dos de las encuestas realizadas en vivienda tuvieron un error ventaja de solo 1 punto, mientras que las encuestas realizadas con otros métodos fueron de hasta 18 puntos. Sin duda, estos márgenes de error ponen en duda la presunción de estas nuevas casas de sondeos, vía Facebook o telefónicas por robot, de ser las más precisas del mercado.

Con la reglamentación de las encuestas electorales en años recientes, las publicaciones de las encuestas “patito” y su uso propagandístico, se ha venido limitando; pero ahora surgen como nueva amenaza a la credibilidad de las encuestas el surgimiento de estos nuevos métodos de recolección de datos. Los nuevos actores son ajenos al tradicional gremio de los encuestadores de opinión pública, en muchas ocasiones con desconocimiento de las metodologías científicas, que promueven las encuestas por Facebook y con robots telefónicos gracias a una considerable reducción en sus costos y tiempos de realización. Sin invocar posiciones ludistas, el reto para el gremio de encuestadores de la opinión pública es evidenciar la diferencia en precisión entre los diferentes tipos de metodologías.

Conclusión

a) La precisión de las encuestas no se califica igual entre especialistas que entre no especialistas. Para el primer grupo importa más el uso de una metodología correcta, la realización con rigurosa calidad de los diferentes procesos que implica una encuesta, y la expectativa de que, cerca de una elección y en ausencia de nueva y relevante información para el electorado, la estimación de la encuesta debe tener la suficiente precisión como para que el resultado oficial caiga dentro del intervalo de confianza de la encuesta. Para los no especialistas, la mayoría de los políticos, candidatos y partidos, columnistas, reporteros y medios de comunicación, y la opinión pública en general, precisión quiere decir fundamentalmente “atinarle” al ganador, y si acaso, “predecir” el número de puntos porcentuales que separan al primero del segundo lugar. Bajo una óptica de especialista, la evolución de la precisión de las encuestas electorales en México no la ha llevado a mejorar su precisión en los últimos 25 años. Desde la óptica de la opinión pública, los resultados han sido más positivos que negativos, y en la actualidad la imagen de las encuestas goza de una credibilidad renovada gracias a que todas las encuestas predijeron el triunfo de López Obrador.

b) Con mayor precisión, para los especialistas, el análisis de la precisión de las últimas encuestas públicas de las últimas cinco elecciones presidenciales en México arroja como conclusión que, si bien ha habido mejores y peores momentos, y mejores o peores encuestadores en cada elección, en su conjunto se puede decir que en general predomina la imprecisión sobre la precisión. Con respecto a la diferencia entre el resultado oficial del candidato ganador, y la estimación de las últimas encuestas públicas, solo 1 de cada 3 encuestadores en promedio logró precisión estadística en tanto que el dato oficial se encontrara dentro del intervalo de confianza de sus estimaciones. De manera similar, la diferencia entre la primera y la segunda fuerza electoral en los resultados oficiales solo alcanzó a ser medido con precisión en el promedio de 4 elecciones nacionales, en solo 1 de cada 4 encuestas públicas. Como se mencionó, la elección del 2006 y sus mediciones fueron un caso atípico por lo cerrado de la elección y por la precisión estadística lograda por 3 de cada 4 encuestadores. Sin embargo, para efectos del análisis de conjunto del periodo, consideramos que este caso es un

outlier, y al eliminarlo del promedio obtenemos una foto más nítida de la precisión promedio de las encuestas.

c) Para los no especialistas, la imagen que debería tener la opinión pública de las encuestas en México debería ser muy positiva. De hecho, en cuatro de las cinco elecciones estudiadas las encuestas predijeron lo que iba a ocurrir. En tres ocasiones todas se manifestaron por un candidato ganador que ganó. En una ocasión, las mediciones daban empate técnico, y la diferencia real de siete décimas de punto no hay encuesta de mil a 3,000 casos capaz de detectar esa diferencia. Por lo que cuenta como buena predicción. La excepción fue el año 2000, donde la mayoría de las encuestas predijo el triunfo del PRI, y el ganador fue el PAN. Cuatro de cinco no es un mal récord. Sin embargo, la percepción en la opinión pública se ve afectada no solo por los resultados objetivos de las encuestas. Sino que depende del resultado de la elección, y de cómo los líderes y candidatos califiquen el desempeño de las encuestas. Esto es, para la opinión pública las elecciones de 1994 y 2018 fueron un éxito para las encuestas ya que todas las encuestas coincidieron en “atinarle” al ganador y los candidatos opositores no cuestionaron el resultado. El hecho de que hayan sido las ocasiones con mayores niveles de imprecisión no fue relevante para la opinión pública. En contraste, la elección de 2010, donde el 80% de los encuestadores estimaron dentro del intervalo de confianza el resultado, fue la más polémica y donde más ayudaron las encuestas a las dificultades políticas post electorales ya que la mayoría estimó que ganaría Andrés Manuel López Obrador, en vez de decir que había un empate técnico y que cualquiera de los candidatos punteros podía ganar. El año de 2012 también fue altamente polémico a pesar de que todas las encuestas acertaron al ganador. Aquí el problema fue que la mayoría de las encuestas sobre estimaron al candidato del PRI, y al darse a conocer los resultados oficiales, AMLO protestó porque perdió por un porcentaje mucho menor de lo que habían “predicho” las encuestas. Y en contraste, en un escenario donde las encuestas no presentaron un mensaje unificado como fue el 2000, debido a que el retador del sistema ganó, no hubo polémica post electoral contra las encuestas. En suma, la precisión es objetiva y subjetiva a la vez, y las consecuencias de la interacción de ambos con el resultado real de la elección y el comportamiento de los candidatos, determina si se juzga positivo o negativo el desempeño

de las encuestas ese año, independientemente de la precisión real de las mismas. Bajo esta perspectiva, la percepción pública, luego de las elecciones de 2018, es que las encuestas en México ya mejoraron.

c) Tal como hemos venido repasando, son numerosas las aristas en el tema de la precisión de las encuestas. Es imposible, de forma determinística, suponer que, de forma general, los problemas de precisión obedezcan exclusivamente a una sola causa. De hecho, en el presente capítulo se revisaron ocho posibles causas o argumentos que se han dado en la discusión que podrían explicar las discrepancias observadas.

d) Los primeros cuatro criterios tienen que ver con los cuatro procesos centrales necesarios para realizar una encuesta, a saber, el adecuado diseño de la muestra probabilística, el correcto diseño de la estructura del cuestionario y el fraseo de las preguntas, el riguroso cuidado en cuanto a calidad y veracidad del levantamiento de campo, y finalmente el procesamiento de la base de datos y el análisis de estos. Si estos cuatro procesos son realizados a la perfección, y el resultado de la encuesta no entra dentro del intervalo de confianza de los resultados oficiales de la elección entonces, debemos juzgar como correcta la encuesta, pero adicionalmente debemos de buscar las causas de la discrepancia.

e) Que en la parte del análisis se utilice un **modelo de estimación** diferente al convencional de reportar la preferencia bruta y la efectiva. Hay muchos ejemplos de cómo esos modelos en ocasiones le funcionan muy bien al encuestador, pero en otras ocasiones el mismo modelo fracasa. Un ejemplo, en 2012 si IPSOS hubiera reportado su preferencia efectiva en vez de su modelo, hubiera entrado dentro del grupo de encuestadores más precisos de esa elección.

f) Fuera de la aplicación de una metodología correcta y un riguroso cuidado de los procesos de la encuesta, y si se ha reportado la preferencia bruta y la efectiva, y aun así hubo discrepancias, hay tres hipótesis que podrían explicarlo. Son tres fenómenos que en ocasiones pueden ocurrir, pero que es necesario demostrar con evidencia que dicho fenómeno fue el verdadero causante de las discrepancias de esa elección. Ninguna de estas hipótesis debe tomarse como una regla general que ocurre en todas las elecciones. Estas son:

- i. **Falseamiento de datos por parte del entrevistado**, ya sea por un fenómeno de espiral del silencio donde le es socialmente complicado revelar su auténtica preferencia, o bien existe coerción por parte de algún partido político o institución de gobierno para votar de determinada manera, y teme revelar su verdadera preferencia. Este fenómeno puede ocurrir, pero es necesario identificar cuando puede atribuirse como causa de la discrepancia y cuando no.
- ii. Que exista una **diferencia entre el perfil de los votantes y el de las personas entrevistadas** en la encuesta, que representan a la totalidad de la ciudadanía. Este fenómeno puede ocurrir a veces, pero es necesario demostrar que todas las demás encuestas tenían una estimación similar entre ellas y los resultados oficiales fueron diferentes. Dejando aparte a quienes hayan publicado su modelo de votantes probables, que a veces funciona y a veces no.
- iii. Una prueba similar es necesaria para verificar que las discrepancias se deben a que **“las preferencias en el electorado cambiaron”** y que la encuesta no predice porque las preferencias son dinámicas. Es decir, ninguna encuesta pública debe contradecir ese diagnóstico, y debe haber evidencia de que algún evento o información entre el levantamiento de campo de las encuestas y el día de la elección provocó el cambio en las preferencias.
- iv. Ahora bien, más allá de las explicaciones, en su conjunto, la industria de las encuestas no ha mejorado mucho en cuanto a precisión en los últimos 25 años. Hay señales positivas y negativas sobre lo que pudiera ocurrir en adelante. Por el lado positivo, el cambio generalizado de levantamiento en campo de papel y lápiz a CAPI, prácticamente está erradicando la práctica del falseamiento de entrevistas en campo y mejorando la calidad del desempeño de los encuestadores en general. Adicionalmente, la exposición pública de las metodologías y bases de datos de todas las encuestas públicas gracias a la regulación electoral ha hecho que aumente el uso de mejores prácticas en diseño de muestra y pregunta electoral, así como el uso o no de modelos. En conjunto, estos elementos empujaron a las empresas encuestadoras a tener los más altos estándares de calidad en los cuatro rubros metodológicos y de procesos mencionados.
- v. Por el lado negativo, la irrupción de las encuestas vía Facebook y telefónicas con robot, que están expandiéndose en el mercado

vertiginosamente por su velocidad de respuesta y su precio económico, empujarán a la credibilidad de las encuestas a una nueva crisis. Como se detalló más arriba, al no tener cobertura adecuada, y sobre todo no originarse en muestras probabilísticas para la selección de entrevistados, sino usar mecanismos de auto selección de los entrevistados, cancelan su posibilidad de ser encuestas probabilísticas y por ende de ser representativas de las opiniones de la ciudadanía. En suma, esta tendencia amenaza con impedir la posible recuperación de credibilidad de las encuestas vía mejores prácticas metodológicas, levantamientos en campo hiper cuidados, que seguramente se traducirían en una mayor precisión promedio de las encuestas en vivienda en México.

- vi. Un desarrollo adicional que actualmente podemos calificar de incierto (más que como algo positivo o negativo), y que sin duda impactará el tema de la investigación de la opinión pública por medio de encuestas, es la creciente presencia de la Inteligencia Artificial y el Big Data en todos los rubros del quehacer humano. Si nuevos modelos de Aprendizaje Automatizado y Aprendizaje Profundo logran generar modelos predictivos mucho más potentes que lo que actualmente conocemos, mediante la integración y análisis de múltiples fuentes de información, entonces, quizá, el tema de este artículo pase a ser un documento histórico pero irrelevante para la práctica futura de la mercadotecnia política.
- vii. Por lo pronto, la principal recomendación de mejora inicial es expandir la regulación actual del INE a todas las encuestas públicas de opinión pública, como por ejemplo las de aprobación de autoridades, aunque no sean electorales. Aumentando la vigilancia y regulación de la realización y publicación de encuestas incluyendo por supuesto las que utilizan las nuevas tecnologías. En paralelo es importante seguir dando a conocer las características de las encuestas metodológicamente serias al público en general y a los medios de comunicación. Por último, fomentar la profesionalización de la labor de los nuevos actores de la demoscopia en México, los jóvenes investigadores del futuro. Labor que está comenzando a realizar la AMAI. ¶