

# La Falla de las Encuestas en las Elecciones Argentinas de 2019

## Un Análisis en Perspectiva Comparada Internacional

### The Failure of the Polls in the 2019 Argentine Elections

### An Analysis in International Comparative Perspective

**José Eduardo Jorge**

E-mail: jjorge@perio.unlp.edu.ar



**Universidad Nacional de La Plata**

La Plata, Argentina

**Ernesto Marcelo Miró**

E-mail: marcelomiro@unnoba.edu.ar



**Universidad Nacional del Noroeste de  
la Provincia de Buenos Aires**

Junín, Argentina

**August 2021**

**Cita:** Jorge, J. E. y Miró, E. M. (2021). La Falla de las Encuestas en las Elecciones Argentinas de 2019. Un análisis en Perspectiva Comparada Internacional. A publicarse en Contreras, Claudio (Comp.). *Democracia y Verdad*. Buenos Aires: Prometeo.

**Resumen:** El desempeño de las encuestas electorales en la Argentina alcanzó en 2019 su punto más bajo desde la recuperación de la democracia en 1983. Los errores fueron mucho mayores que los observados en los últimos años en democracias maduras, donde se habla de una crisis de este tipo de sondeos. Luego de subestimar ampliamente la ventaja que obtendría el Frente de Todos sobre Juntos por el Cambio en las primarias de agosto –y de ajustar el diseño de sus trabajos sobre la base de ese resultado–, las encuestadoras locales pasaron a sobrestimar con el mismo exceso esa ventaja en la elección general de octubre. Aquí analizamos ambos fallos en perspectiva comparada internacional. Utilizamos los datos de intención de voto a presidente para calcular indicadores de error y compararlos con los promedios internacionales actuales y de décadas pasadas, y con los errores en varios ciclos electorales en EE.UU. y el Reino Unido. Examinamos los cambios en la industria de las encuestas, la caída de las tasas de respuesta y las causas de error identificadas por nuevos estudios académicos y trabajos encarados por asociaciones profesionales del campo de investigación de la opinión pública. Esta literatura muestra que los sondeos preelectorales no son hoy más inexactos que en el pasado y subraya la importancia de los errores "no de muestreo", en particular los que proceden de un sesgo de no-respuesta, cuando la probabilidad de participar en una encuesta es menor para ciertos grupos, como las personas con poco interés por la política o con baja confianza social o institucional. Este tipo de errores y sus causas se comprenden hoy mucho mejor y es posible diseñar procedimientos para aumentar la exactitud de las estimaciones. Discutimos además el impacto de las encuestas inexactas en las estrategias políticas y en la narrativa de los medios sobre la marcha de una elección, así como la dificultad de encuestadoras, periodistas y actores políticos para reconocer y comunicar los niveles de incertidumbre asociados a los pronósticos electorales. Nuestro análisis pone en evidencia limitaciones básicas de la industria de las encuestas y la investigación de la opinión pública en la Argentina, como la falta de una asociación profesional y el incipiente desarrollo de los estudios sobre comportamiento electoral. Las encuestadoras se vieron sorprendidas por el

“voto castigo” al gobierno debido a la crisis económica, cuando este era el escenario más probable según un conjunto no muy grande pero significativo de estudios empíricos sobre el “voto económico” en la Argentina y América Latina. Los grandes errores de 2019 tienen implicaciones para la confiabilidad de los sondeos sobre temas de la agenda pública, que podrían mostrar un cuadro distorsionado de las opiniones de la sociedad argentina.

**Palabras clave:** Elecciones en Argentina, Encuestas Electorales, Errores en las Encuestas

**Abstract (English):** The performance of polls in the 2019 Argentine elections reached its lowest point since democratic restoration in 1983. Their errors were much greater than those observed in recent years in mature democracies, where there is talk about a crisis of this type of surveys. In the primaries, Argentine pollsters widely underestimated the advantage of the Everyone's Front opposition alliance over the governing Together for Change. But after adjusting their methods, in the general election they overestimated that advantage by similar margins. We analyze both failures in an international comparative perspective. Vote intention data for president was used to calculate error indicators and compare them with current international averages and those of past decades, and with errors in several electoral cycles in the US and the UK. We examine ongoing changes in the survey industry, falling response rates, and causes of error identified by new academic studies and work commissioned by professional associations in the field of public opinion research. This literature shows that pre-election polls are no more inaccurate today than in the past and underlines the importance of "non-sampling" errors, particularly those produced by non-response bias, when the probability of participating in a poll is lower for certain groups, such as people with little interest in politics or low social or institutional trust. These types of errors and their causes are now better understood and procedures can be designed to improve the accuracy of estimations. We also discuss the impact of inaccurate polls on political strategies and the media narrative of elections, as well as the difficulty for pollsters, journalists and political actors to acknowledge and communicate the levels of uncertainty associated with electoral predictions. Our analysis highlights basic limitations of the polling industry and public opinion research in Argentina, such as the lack of a professional association and the incipient development of electoral behavior studies. Pollsters were surprised when voters punished the government because of the economic crisis, even though it was the most likely scenario according to a not very large but significant set of empirical studies on economic voting in Argentina and Latin America. The substantial errors in 2019 have implications for the reliability of issue polling, which may show a distorted picture of opinions in Argentine society.

**Keywords:** Elections in Argentina, Polling, Survey Error

El desempeño de las encuestas de intención de voto en la Argentina alcanzó en las elecciones de 2019 su punto más bajo desde la recuperación de la democracia en 1983. Los errores de pronóstico fueron mucho mayores que los observados en los últimos años en democracias de larga duración, donde hay una percepción extendida de una crisis de esta técnica de medición de la opinión.

El caso argentino presenta además el rasgo inusual –y agravante- de un yerro por partida doble, pues luego de subestimar ampliamente la ventaja que obtendría el Frente de Todos sobre Juntos por el Cambio en las primarias de agosto –y de ajustar el diseño de sus trabajos sobre la base de ese resultado-, las encuestadoras locales pasaron a sobrestimar con el mismo exceso esa ventaja en la elección general de octubre.

En este trabajo analizamos ambos fallos en perspectiva comparada internacional. En años recientes han surgido importantes investigaciones sobre el alcance y las causas de los nuevos errores de los sondeos en democracias maduras. Esos estudios revelan que las encuestas preelectorales no son hoy más inexactas que en el pasado, pese a que los cambios de la industria –con la caída de las tasas de respuesta, la rápida expansión de los relevamientos online y el impacto disruptivo del uso del celular sobre las entrevistas telefónicas- puedan sugerir lo contrario. La idea de una “era dorada” de este tipo de estudios, dominada por los trabajos de campo presenciales, es un mito que no resiste la prueba empírica.

Desde los años 40 hasta el presente, los márgenes de error han sido siempre más grandes que los estimados –y declarados- a partir de un uso simplista de la teoría del muestreo. La causa principal de esa discrepancia son los errores “no de muestreo”. En particular –si bien no únicamente-, los que proceden de un “sesgo de no-respuesta”, cuando la probabilidad de participar en un sondeo es menor para ciertos grupos de la población, como las personas no interesadas en la política.

Todo sugiere que estas y otras fuentes de error no muestral no son exclusivas de las metodologías actuales, sino que han afectado a las encuestas desde que hicieron su aparición hace casi una centuria. Estos tipos de error y sus causas se comprenden hoy mucho mejor, de modo que es posible concebir y ensayar procedimientos para aumentar la exactitud de las predicciones.

En lo que sigue recurrimos a los datos de los sondeos de intención de voto a presidente publicados en la prensa argentina, a fin de evaluar el desempeño de los pronósticos en las dos elecciones de 2019. El análisis está enfocado en la contienda entre el Frente de Todos y Juntos por el Cambio, para los que disponemos de estimaciones finales a horas de los comicios. Calculamos los indicadores típicos de error empleados en el campo de investigación y los comparamos con los promedios internacionales de elecciones recientes y de décadas pasadas, y con los errores observados en varios ciclos electorales de EE.UU y el Reino Unido, dos países intensamente estudiados.

Mostramos luego cómo los márgenes de error reportados por las encuestas preelectorales –en la Argentina y en todo el mundo- no solo se circunscriben a los que derivan del muestreo, sino que son calculados de un modo que tiende a subestimarlos, pasando por alto complejidades de

diseño y aún muestras no probabilísticas, para las que no pueden computarse. Los estudios empíricos sugieren que, debido sobre todo a los errores “no de muestreo”, los márgenes “reales” duplican a los reportados, un resultado consistente con lo observado en 2019 en nuestro país.

Después de examinar las transformaciones actuales de la industria y sus potenciales efectos problemáticos sobre la exactitud de las encuestas, dirigimos nuestra atención a las causas de error no muestral.

Empezamos por analizar los casos de las competencias presidenciales de 2016 y 2020 en EE.UU. y la elección general de 2015 en el Reino Unido, donde la actuación de los sondeos, tras las críticas recibidas, fue objeto de investigaciones independientes. Exponemos las hipótesis sobre las posibles fuentes de error, la evidencia y los métodos usados para su contraste empírico y las conclusiones obtenidas, según las investigaciones de los comités de expertos creados por la *American Association for Public Opinion Research* (AAPOR) y el *British Polling Council* (BPC).

Seguidamente ampliamos y sistematizamos el análisis de las fuentes de error no muestral, a partir de un conjunto de estudios académicos que abarcan varios ciclos electorales de estos dos países.

Los hallazgos del campo de investigación nos permiten plantear hipótesis sobre las posibles causas de los errores de los sondeos argentinos en 2019, así como evaluar la plausibilidad de las explicaciones brindadas por quienes los condujeron. Por comparación con los casos previamente examinados, este ejercicio trae a un primer plano limitaciones básicas de la industria de las encuestas y de los estudios académicos sobre las actitudes y conductas políticas en nuestro país. La falta de una asociación profesional de autorregulación de la actividad –similar a la AAPOR o el BPC- y el desarrollo incipiente de la investigación del comportamiento electoral de los argentinos –vinculado a su vez a la ausencia de una encuesta académica periódica para los ciclos electorales, como las que existen en otros países- restringen la posibilidad de indagar aspectos como los móviles del voto y el proceder de los sondeos en una elección dada.

La escasa atención general que reciben en la Argentina los trabajos académicos en esta área de conocimiento afectó a las propias encuestadoras, que se vieron sorprendidas por el llamado “voto castigo” al gobierno originado en la mala situación económica, cuando este era el escenario más probable según un conjunto no muy grande pero significativo de estudios empíricos sobre el llamado “voto económico” en la Argentina y América Latina.

En la sección final discutimos la forma en que las encuestas inexactas distorsionan la narrativa de la elección por parte de los medios y las decisiones de los actores políticos y económicos. El periodismo, a su vez, suele pedir más exactitud de la que es factible y seleccionar datos –o cambios insignificantes en las mediciones de intención de voto- que mejor se ajustan a su narrativa, influyendo con su propia acción en las estrategias políticas. Estos procesos subrayan la dificultad de los distintos actores para reconocer y comunicar los niveles de incertidumbre que son inherentes a los pronósticos electorales.

Los márgenes “reales” de error que los tests electorales ponen en evidencia tienen implicaciones para la confiabilidad de los sondeos sobre temas de la agenda pública, que al no pasar por ningún examen semejante podrían instalar un cuadro sesgado de las opiniones de la sociedad. Con los grandes errores observados en 2019 y la ausencia de un ente profesional que fije estándares de calidad y un código de ética, esta posibilidad constituye un riesgo cierto en nuestro país.

### **Los errores de las encuestas en las primarias y la elección presidencial**

Durante los dos meses de campaña previos a las elecciones Primarias, Abiertas, Simultáneas y Obligatorias (PASO) del 11 de agosto, las encuestadoras preveían para la carrera presidencial una modesta ventaja del candidato del Frente de Todos, Alberto Fernández, sobre el de Juntos por el Cambio, Mauricio Macri, que buscaba su reelección. En la víspera de las PASO, los sondeos anticipaban en promedio una diferencia de alrededor de 5 puntos a favor de Fernández (Tabla 1).

La narrativa dominante en los medios de comunicación era que Macri, apelando a una estrategia de polarización, podía remontar un resultado de ese tenor en las primarias y ganar la elección general del 27 de octubre –o una eventual segunda vuelta-, pese la crisis económica y las políticas de austeridad que caracterizaron a la segunda mitad de su gestión.

Los 16 puntos de ventaja que Fernández obtuvo en las PASO causaron gran sorpresa y superaron ampliamente las expectativas del Frente de Todos, para el que la encuesta publicada más favorable vaticinaba poco más de 9 puntos de distancia. Según la prensa, también los sondeos privados de ambas coaliciones habían errado por mucho el resultado (Clarín, 27/9/2019).

El consenso unánime fue que, a más de dos meses de la elección general, la competencia había quedado prácticamente definida. Los mercados, que venían apostando a la reelección de Macri, reaccionaron con nerviosismo. Las acciones, los bonos y el peso se desplomaron.

A fin de evaluar el desempeño de las encuestas, en la Tabla 1 hemos compilado las estimaciones disponibles más próximas a la fecha de los comicios. La mayor parte de estos datos fueron producidos por las encuestadoras durante el periodo de veda electoral –en el que se prohíbe su difusión- e informados por *Clarín* el 11 de agosto. Están circunscriptos a las dos principales fórmulas para presidente y vice: Alberto Fernández / Cristina Fernández de Kirchner y Mauricio Macri / Miguel Ángel Pichetto.

Solo se difundió información fragmentaria sobre el diseño de las encuestas: modo de administración –telefónico, online, presencial o mixto-, tamaño de muestra y margen de error declarado. Estos detalles no se suministraron para los datos generados dentro del lapso de veda, por lo que la tabla incluye como referencia las características de diseño reportadas por esas mismas firmas en su trabajo previo.

**Tabla 1**  
**Elecciones Primarias, Abiertas, Simultáneas y Obligatorias (PASO) del 11 de agosto de 2019**  
**Encuestas finales: estimaciones publicadas para las dos principales fórmulas presidenciales**

Encuestadora	Fernández Fernández	Macri Pichetto	Dife- rencia	Error Absoluto		Modo	Nº de casos	Margen de Error (informado)
				Medio	En el Margen			
<b>Resultado</b> (sobre votos válidos)	<b>47,79</b>	<b>31,80</b>	<b>15,99</b>					
Gustavo Córdoba & Asociados <sup>a</sup>	43,5	34,2	9,3	3,35	6,69	T <sup>c</sup>	--	--
Federico González & Asociados <sup>a</sup>	44,5	36,8	7,7	4,15	8,29	P-T <sup>b</sup>	1.600 <sup>b</sup>	+/- 2,5 <sup>b</sup>
Trespuntozero <sup>a</sup>	42,6	36,1	6,5	4,75	9,49	T <sup>c</sup>	1.200 <sup>c</sup>	--
Raúl Aragón & Asociados <sup>a</sup>	42,2	35,9	6,3	4,85	9,69	T <sup>c</sup>	2.500 <sup>c</sup>	--
Dicen <sup>b</sup>	43,7	38,1	5,6	5,20	10,39	T <sup>b</sup>	2.243 <sup>b</sup>	+/- 2,1 <sup>b</sup>
Marketing & Estadística <sup>a</sup>	41,0	36,0	5,0	5,50	10,99	O <sup>b</sup>	3.287 <sup>b</sup>	+/- 1,7 <sup>b</sup>
Synopsis <sup>a</sup>	42,3	38,4	3,9	6,05	12,09	--	--	--
Quality Research <sup>a</sup>	45,0	41,7	3,3	6,35	12,69	--	--	--
Opinaia <sup>a</sup>	39,2	36,1	3,1	6,45	12,89	O <sup>b</sup>	3.051 <sup>b</sup>	+/- 1,8 <sup>b</sup>
Opina Argentina <sup>a</sup>	39,0	36,0	3,0	6,50	12,99	T-P <sup>c</sup>	1.850 <sup>c</sup>	--
Management & Fit <sup>d</sup>	41,2	39,2	2,0	7,00	13,99	P-T-O <sup>d</sup>	2.000 <sup>d</sup>	+/- 2,2 <sup>d</sup>
<b>Promedio</b>	<b>42,20</b>	<b>37,14</b>	<b>5,06</b>	<b>5,46</b>	<b>10,93</b>			
<b>Promedio del Error con Signo</b>	<b>-5,59</b>	<b>+5,34</b>						

**Nota:** En las PASO los porcentajes se calculan sobre el total de votos "válidos", que incluyen los sufragios en blanco.

Las referencias contiguas a los nombres de cada encuestadora corresponden a la fuente periodística que informó sus porcentajes finales para las dos principales fórmulas presidenciales. **Modo de recolección de datos:** T, telefónico; P, presencial; O, online.

<sup>a</sup> Clarín, 11/8/2019. <sup>b</sup> Clarín, 2/8/2019. <sup>c</sup> Página 12, 2/8/2019. <sup>d</sup> Clarín, 1/8/2019. Ver el detalle de las fuentes de datos periodísticas al final de la Bibliografía

Nuestra tabla contiene exclusivamente las estimaciones que ya incorporaban el voto proyectado de los “indecisos”. Asumimos además que las encuestadoras computaron estos porcentajes finales sobre el total de votos “válidos” –que en las PASO, a diferencia de la elección del 27 de octubre, comprenden a los sufragios en blanco- y que, o bien excluyeron de la base del cálculo porcentual a los entrevistados que era improbable que votaran, o asignaron a todos los entrevistados la condición de probables votantes.

Con estos supuestos, podemos calcular dos indicadores ampliamente utilizados para evaluar la exactitud de este tipo de sondeos. Esto nos permitirá cotejar el desempeño de las encuestas en las PASO con el que alcanzaron después en la elección general, así como compararlo con los estándares mundiales –recientes e históricos- y los yerros observados en los últimos años en elecciones de otros países.

Una de estas medidas es el error absoluto medio (EAM) en la predicción del voto de las dos principales fórmulas presidenciales (Jennings y Wlezien, 2018, p. 278). Resulta de computar para cada una la diferencia en valor absoluto entre el porcentaje de los votos que les vaticinó una encuesta y el que lograron en la elección  $|\% \text{ Encuesta} - \% \text{ Elección}|$  y obtener luego el promedio de los resultados de ambas fórmulas. El valor absoluto –simbolizado por las barras verticales- implica en la práctica prescindir del signo negativo que pueda arrojar una diferencia. Así, para la primera encuestadora de la Tabla 1, el EAM surge de calcular  $(|43,5 - 47,79| + |34,2 - 31,80|)/2$ , es decir  $(4,29 + 2,40)/2$ , que arroja el valor 3,35. Este es el error absoluto medio más pequeño de las once firmas del cuadro, que hemos ordenado según un ranking de exactitud. El EAM máximo de la tabla es de 7 puntos porcentuales y el promedio de todas las encuestadoras es 5,46.

Nótese que, promediando entre las once firmas, el error “con signo” en la predicción del voto es -5,59 para la fórmula del Frente de Todos y + 5,34 para la de Juntos por el Cambio. Esto indica el grado en que las encuestas en conjunto subestimaron el voto de Fernández y sobrestimaron el de Macri. El promedio de estos dos errores, prescindiendo del signo negativo del primero, arroja también el EAM global de 5,46.

Una segunda medida es el “error absoluto en el margen”, es decir, el error en predecir la ventaja del partido o candidato que obtuvo más votos sobre su inmediato seguidor (Ibíd., pp. 279-80). Este indicador suele ser más relevante que el anterior en varios contextos, como es predecir si una elección presidencial arrojará un resultado decisivo o hará falta una segunda vuelta. Surge de calcular la diferencia en valores absolutos entre el margen de ventaja  $(\% \text{ Fernández} - \% \text{ Macri})$  anticipado por una encuesta menos el mismo margen  $(\% \text{ Fernández} - \% \text{ Macri})$  observado en la elección. En el caso de la primera encuestadora de la Tabla 1, el cómputo es  $|(43,5 - 34,2) - (47,79 - 31,80)|$ , es decir  $|9,3 - 15,99|$ , que arroja el valor 6,69. El error absoluto en el margen más grande de la tabla es 13,99 y el promedio de las once firmas es 10,93. Como veremos, se trata de un error muy grande –aunque no sin precedentes- para los estándares internacionales.

Tras este fracaso colectivo, la credibilidad de los encuestadores quedó seriamente afectada, un hecho acentuado por su habitual alto perfil como comentaristas políticos en los programas de

televisión. Los grandes diarios hablaron de “papelón” y “pifie generalizado” (Clarín, 27/9/2019; La Nación, 23/8/2019). El Financial Times (2/9/2019) apuntó al “desastroso” error de pronóstico como causa del pánico posterior de los mercados.

Las firmas de opinión hicieron una pausa en la difusión de trabajos, encararon su propio diagnóstico sobre las razones del yerro y ensayaron explicaciones tentativas ante los requerimientos de la prensa (La Nación, 12/8, 23/8 y 5/10/2019; Clarín, 28/8 y 27/9/2019). Uno de los puntos más mencionados fue el predominio de técnicas de relevamiento basadas en Internet y telefonía fija –omitiendo el creciente uso del celular-, que dificultaban la obtención de muestras representativas del conjunto del electorado, como los sectores de menores ingresos y educación, que votaban mayormente al Frente de Todos. Los bajos presupuestos disponibles, se indicó, volvían prohibitivo el tradicional muestreo de áreas y las entrevistas presenciales.

Un argumento frecuente fue la existencia de un voto “oculto” o “vergonzante” por la principal coalición opositora, que los sondeos no habrían podido detectar. Otras razones invocadas fueron las bajas tasas de respuesta –que, otra vez, habrían afectado en mayor medida a segmentos proclives al Frente de Todos, como los jóvenes-, una alta proporción de indecisos que habría optado finalmente por esta coalición, cambios de voto de último momento e incluso entrevistados que no quisieron revelar su voto o dieron respuestas deliberadamente falsas.

Hubo referencias a una conducta de “manada” por parte de firmas que, al obtener estimaciones que se desviaban del discurso dominante de paridad entre las dos principales coaliciones, juzgaron que sus datos podían ser erróneos y, ante el riesgo de quedar expuestas, ajustaron sus estudios para conformar sus resultados al consenso existente.

Veremos que todos estos argumentos coinciden con los esgrimidos habitualmente en otras democracias, aunque las investigaciones independientes muestran que los grandes yerros de los últimos años se han originado en distintas -y a menudo múltiples- causas y dan poco respaldo a explicaciones como el voto vergonzante, que tienden a excusar a los responsables de los sondeos.

Las encuestadoras argentinas no hicieron mención sin embargo a un punto problemático subrayado por esas investigaciones: los modelos utilizados para predecir la probabilidad de votar de los entrevistados, que en nuestro caso podrían haber tenido un impacto adicional por la distinta participación electoral observada en las PASO y la elección general.

Apuntando a la elección del 27 de octubre, según los reportes periodísticos, las firmas de nuestro país adecuaron sus muestras sobre la base de la información proporcionada por los resultados de las primarias. Para mejorar la representatividad, algunas optaron por modos mixtos de relevamiento, combinando en especial las entrevistas telefónicas u online con otras presenciales. Hubo quienes dejaron de difundir sus estimaciones para concentrarse en sus clientes privados, mientras el tropiezo general de las firmas establecidas estimuló el ingreso de otras nuevas (Reuters, 9/9/2019; Bloomberg, 16/10/2019).

Pero la fallida experiencia no condujo a mejores pronósticos. La gran mayoría de las encuestadoras que divulgaron sus datos predijo desde el principio que Alberto Fernández ampliaría su ventaja. Sobre el filo de la elección, la diferencia anticipada entre el candidato del



Frente de Todos y el de Juntos por el Cambio era, promediando entre todos los sondeos, de 17,7 puntos porcentuales a favor del primero (Tabla 2). Las urnas arrojaron empero una diferencia de algo menos de 8 puntos.

Nuestra tabla reúne las estimaciones de quince firmas, con datos más completos que en las PASO sobre el diseño de los trabajos. El error absoluto medio total en la predicción del voto de las dos fórmulas es 4,87 –apenas inferior al de las primarias-, mientras el error con signo es +3,86 para el Frente de Todos y -5,89 para Juntos por el Cambio. El error absoluto en el margen de ventaja asciende en promedio a 9,75, solo 1,18 puntos por debajo del registrado en las PASO.

Los yerros de las predicciones para el 27 de octubre fueron, en consecuencia, de magnitud similar a los del 11 de agosto, solo que esta vez en la dirección contraria. En lugar de subestimar los votos de Fernández y sobrestimar los de Macri, como ocurrió en las primarias, ahora las encuestas sobrestimaron los votos del primero y subestimaron los del segundo. En esta ocasión hubo empero dos firmas –la brasileña *Atlas Intel* y *Trespuntozero*- que lograron una exactitud razonable, con errores absolutos en el margen de 1,94 y 3,04 puntos.

La Tabla 3 permite apreciar el grado en que varió entre las PASO y la elección general el desempeño de seis encuestadoras locales que divulgaron sus datos en ambas oportunidades. Solo una de ellas –*Trespuntozero*- mejoró sus predicciones hasta reducir los indicadores de error a niveles aceptables. Otras dos vieron empeorar sus pronósticos en forma significativa.

### **¿Crisis de las encuestas? Percepción pública y realidad estadística**

Al comentar el traspíe, los medios no dejaron de aludir a las críticas recibidas por los sondeos preelectorales en otros países, tras una serie de fallas que para algunos denotarían una crisis de la industria (Sturgis, 2019; Cassino, 2016; Zukin, 2015). En 2016, dos episodios habían erosionado la confianza en las encuestas: los triunfos de la opción por abandonar la Unión Europea en el referéndum británico conocido como *Brexit* y de Donald Trump sobre Hillary Clinton en la elección presidencial de EE.UU. Ambos desenlaces sobrevinieron como grandes sorpresas, pues los sondeos anticipaban de modo consistente los resultados opuestos.

La elección norteamericana de 2020, en la que Joe Biden derrotó a Trump, volvió a mostrar que la industria de las encuestas –y los medios de comunicación y analistas que las presentan e interpretan para su consumo por el público- tienen por delante más problemas que resolver (Clinton et al., 2021; Merkle, 2021; Silver, 2020; Leonhardt, 2020). Los estudios predijeron correctamente quién ganaría la presidencia, pero no previeron la magnitud del apoyo que conservaba Trump a pesar de su mala gestión de la pandemia del Covid-19, que había costado por entonces la vida de más de 200 mil estadounidenses. Aunque Biden ganó con claridad, los sondeos sobrestimaron la ventaja de los demócratas en las contiendas para presidente, senadores y gobernadores. La falla en la estimación del voto popular en el nivel nacional no solo fue mayor que la de 2016, sino la más grande desde que Ronald Reagan se impuso a Jimmy Carter en 1980 por una diferencia tan amplia como imprevista.

**Tabla 2**  
**Elección presidencial del 27 de octubre de 2019**  
**Encuestas finales: estimaciones publicadas para las dos principales fórmulas**

Encuestadora	Fernández Fernández	Macri Pichetto	Dife- rencia	Error Absoluto		Modo	Nº de casos	Margen de Error (informado)
				Medio	En el Margen			
<b>Resultado</b> (sobre votos afirmativos)	<b>48,24</b>	<b>40,28</b>	<b>7,96</b>					
Atlas Intel <sup>a</sup>	49,5	39,6	9,9	0,97	1,94	--	--	--
Trespuntozero <sup>a</sup>	48,2	37,2	11,0	1,56	3,04	O	1.100 <sup>b</sup>	--
Giacobbe & Asociados <sup>a</sup>	51,3	36,3	15,0	3,52	7,04	O-P	--	--
Opinaia <sup>a</sup>	50,0	34,0	16,0	4,02	8,04	O	--	--
CIGP <sup>a,b</sup>	51,4	35,1	16,3	4,17	8,34	O-T	1.202	+/- 3,5
Ricardo Rouvier <sup>a,b</sup>	52,3	34,3	18,0	5,02	10,04	T	1.200	+/- 2,8
Dicen <sup>a,b</sup>	51,9	33,2	18,7	5,37	10,74	T	550	+/- 4,2
Oh Panel <sup>a,b</sup>	52,0	33,0	19,0	5,52	11,04	O	1.250	--
Gustavo Córdoba & Asociados <sup>a,c</sup>	52,2	32,7	19,5	5,77	11,54	P	1.200	+/- 2,8
Circuitos <sup>a</sup>	53,6	33,7	19,9	5,97	11,94	--	--	--
Opina Argentina <sup>a,d</sup>	52,0	32,0	20,0	6,02	12,04	O	1.700	+/- 3,0
CB Consultora <sup>a,e</sup>	55,3	34,9	20,4	6,22	12,44	O	7.931	+/- 1,1
Clivajes <sup>a,e</sup>	53,7	33,2	20,5	6,27	12,54	O	1.238	+/- 2,8
CEOP <sup>f</sup>	53,6	33,0	20,6	6,32	12,64	P	1.605	--
Federico González & Asociados <sup>a,b</sup>	54,4	33,6	20,8	6,42	12,84	P-T	1.800	+/- 2,4
<b>Promedio</b>	<b>52,09</b>	<b>34,39</b>	<b>17,71</b>	<b>4,87</b>	<b>9,75</b>			
<b>Promedio del Error con Signo</b>	<b>+3,86</b>	<b>-5,89</b>						

**Nota:** En la elección general los porcentajes se calculan sobre los votos “afirmativos”, que excluyen los sufragios en blanco.  
<sup>a</sup> Clarín, 27/10/2019. <sup>b</sup> Clarín, 18/10/2019. <sup>c</sup> Clarín, 13/10/2019. <sup>d</sup> Clarín, 16/10/2019. <sup>e</sup> Clarín, 9/10/2019. <sup>f</sup> Página 12, 19/10/2019.  
 Ver el detalle de las fuentes de datos periodísticas al final de la Bibliografía

**Tabla 3**  
**Encuestadoras que publicaron estimaciones en las PASO y la elección de octubre para las dos principales fórmulas presidenciales: variación de los indicadores de error**

Encuestadora	Error Absoluto Medio			Error Absoluto en el Margen		
	PASO	Octubre	Diferencia	PASO	Octubre	Diferencia
Trespuntozero	4,75	1,56	-3,19	9,49	3,04	-6,45
Opinaia	6,45	4,02	-2,43	12,89	8,04	-4,85
Opina Argentina	6,50	6,02	-0,48	12,99	12,04	-0,95
Dicen	5,20	5,37	0,18	10,39	10,74	0,35
Federico González & Asociados	4,15	6,42	2,28	8,29	12,84	4,55
Gustavo Córdoba & Asociados	3,35	5,77	2,43	6,69	11,54	4,85
<b>Promedio</b>	<b>5,06</b>	<b>4,86</b>	<b>-0,20</b>	<b>10,12</b>	<b>9,71</b>	<b>-0,42</b>

Nuevos estudios académicos revelan empero que la idea de que este tipo de yerros se han vuelto más prevalentes en los últimos años no se corresponde con los hechos. La investigación más exhaustiva es la de los politólogos Will Jennings y Christian Wlezien, que reunieron una base de datos de 30.000 encuestas electorales en democracias de Europa, EE.UU., Canadá, América Latina y Asia entre 1942 y 2017. Sus análisis muestran que el desempeño de los sondeos recientes no diverge de los estándares de décadas previas. Para 220 elecciones en 32 países, el trabajo cuenta con datos de relevamientos de la semana anterior a los comicios. El error absoluto medio, calculado para cada año sobre todas las elecciones y partidos que obtuvieron al menos 5% de los votos, ha oscilado desde 2000 en torno de 2%, cifra similar al 2,1% entre las décadas de los 40 y los 70 (Jennings y Wlezien, 2018, p. 280).

Estos indicadores no son comparables directamente con los presentados para las PASO y las presidenciales argentinas, que alcanzan solo a las dos principales coaliciones por falta de suficientes estimaciones de última hora para los partidos con menos caudal de votos. Como prevé la teoría del muestreo, las predicciones para los partidos pequeños tienen menores márgenes de error que las de los grandes y, de haberse incluido en el cálculo del EAM, hubieran tenido el efecto de reducirlo.

Según Jennings y Wlezien, el EAM sube a 2,6 puntos porcentuales para los partidos con 20% o más de los votos en 11 elecciones de democracias maduras celebradas entre 2015 y 2017 (Tabla 4). Esa cifra es similar al promedio histórico de 2,3 puntos correspondiente al periodo 1942-2014. El EAM de los dos comicios argentinos de 2019 -esta vez comparable, al incluir solo las fórmulas con 20% o más de los sufragios- duplica esos promedios internacionales.

La Tabla 4 indica también el error absoluto en el margen de ventaja entre el primer y el segundo puesto en varias elecciones de EE.UU., Reino Unido y Francia. Los datos proceden de estudios independientes, en tres casos comités creados *ad hoc* para investigar el proceder de las encuestas (Clinton et al., 2021; Kennedy et al., 2017, 2018; Sturgis et al., 2016).

Podemos ver que las predicciones de nivel nacional para la elección presidencial de 2016 en EE.UU. fueron exactas según los patrones históricos. Su error absoluto en el margen de 2,2 puntos porcentuales es inferior a los 2,9 puntos de 2012 -cuando subestimaron la distancia con la que Barack Obama se impuso a Mitt Romney- y al promedio de 2,7 puntos del periodo 1992-2016. Los sondeos nacionales indicaron correctamente que Hillary Clinton aventajaba a Donald

Trump en el voto popular por una diferencia cercana a los 2,1 puntos porcentuales que arrojaron las urnas.

**Tabla 4**  
**Errores de las estimaciones de las encuestas preelectorales**  
**Las elecciones argentinas de 2019 en la comparación internacional**

	<b>Error Absoluto en el Margen</b>	<b>Error Absoluto Medio</b>
<b>Argentina 2019. Elección presidencial: dos principales coaliciones</b>		
Elección del 27 de octubre	9,7	4,87
PASO del 11 de agosto	10,9	5,46
<b>Promedios mundiales: partidos con más de 20% de los votos <sup>a</sup></b>		
Elecciones presidenciales y legislativas		
Promedio 1942-2014	--	2,30
Democracias maduras: 2015-2017 <sup>b</sup>	--	2,60
<b>Elecciones recientes: partidos en el primer y segundo puesto</b>		
EE.UU. Elecciones presidenciales		
2020. Nivel nacional <sup>c</sup>	4,5	
2020. Nivel estadual <sup>c</sup>	5,1	
2016. Nivel nacional <sup>d</sup>	2,2	
2016. Nivel estadual <sup>d</sup>	5,1	
2012. Nivel nacional <sup>d</sup>	2,9	
2012. Nivel estadual <sup>d</sup>	3,5	
Promedio 1992-2016. Nivel nacional <sup>d</sup>	2,7	
1980. Nivel nacional <sup>d</sup>	6,0	
Reino Unido		
2017: Elección general <sup>e</sup>	5,4	
2016: Referéndum por el Brexit <sup>f</sup>	6,1	
2015: Elección general <sup>g</sup>	6,6	
1992: Elección general <sup>h</sup>	9,0	
1970: Elección general <sup>i</sup>	6,6	
Francia 2017: Elección presidencial (2ª ronda) <sup>j</sup>	10,2	

<sup>a</sup> Jennings y Wlezien, 2018, p. 281. Media de los promedios anuales de error calculados sobre todas las elecciones de cada año. <sup>b</sup> Ibídem. Error promedio de 11 elecciones: Gran Bretaña 2015 y 2017; Francia 2017 (1ª y 2ª ronda); EE.UU. 2016; Dinamarca 2015; Grecia 2015; Canadá 2015; Irlanda 2016; España 2016; Australia 2016; Islandia 2016. <sup>c</sup> Clinton et al., 2021. <sup>d</sup> Kennedy et al., 2017, 2018. <sup>e</sup> Cálculos propios a partir de datos de Cowley y Kavanagh, 2018, p. 270. <sup>f</sup> Cálculos propios a partir de datos de Curtice, 2016b. <sup>g</sup> Cálculos propios a partir de datos de Sturgis et al., 2016, p. 12. <sup>h</sup> Cálculos propios a partir de Jowell et al., 1993, p. 242. <sup>i</sup> Cálculos propios a partir de Abrams, 1970, pp. 318-19. <sup>j</sup> Enten, 2017b.

El error de 5,1 puntos de las encuestas estaduales fue sin embargo mucho más grande y señala una causa principal de la percepción del fracaso general de la industria por parte del público. Aunque estos relevamientos auguraban una carrera competitiva con una leve ventaja de Clinton, los realizados en Pensilvania, Michigan y Wisconsin –de tradición demócrata– subestimaron el voto por Trump, quien se adjudicó los tres por márgenes estrechos, volcando a su favor el colegio electoral (Kennedy et al., 2017).

En 2020, en cambio, las encuestas sobrestimaron ampliamente la ventaja de Biden sobre Trump, tanto en el nivel nacional como en el estadual (Clinton et al., 2021). El error de 4,5 puntos para todo el país duplica al de 2016 y es el mayor de los últimos 40 años. En 1980,

cuando la industria reunía un número mucho menor de encuestadoras, este error -calculado a partir de solo cuatro estudios- había sido de 6 puntos.

El Reino Unido es un caso instructivo, en el que tres elecciones sucesivas muy relevantes –dos generales (en 2015 y 2017) y el referéndum del Brexit en 2016- se consideraron notorios fracasos de los sondeos. En 2015, para sorpresa de todos, los conservadores liderados por el primer ministro David Cameron lograron un triunfo categórico y obtuvieron la mayoría en la cámara de los comunes, luego de que los sondeos pronosticaran un parlamento equilibrado, subestimando el voto conservador y sobrestimando el laborista. Una comisión investigadora impulsada por el *British Polling Council* (BPC) –que nuclea a las principales firmas- y la *Market Research Society* (MRS), examinó las causas del fallo e hizo recomendaciones para mejorar las metodologías de medición (Sturgis et al., 2016).

Pero la industria volvió a naufragar en 2017, cuando la primera ministra Theresa May, alentada por encuestas muy favorables, llamó a elecciones anticipadas, buscando ampliar la mayoría de los Tories en la cámara de los comunes y lograr un mandato más robusto para negociar la salida de la Unión Europea –aprobada por referéndum el año anterior- (Cowley y Kavanagh, 2018). La jugada resultó desastrosa. A la inversa de lo ocurrido en 2015, esta vez los sondeos subestimaron el voto de los laboristas y sobrestimaron el de los conservadores, que perdieron escaños al superar al partido liderado por Jeremy Corbyn por apenas 2,5 puntos porcentuales. May quedó así políticamente debilitada.

Como un evento único, diferente de una elección general, el referéndum del Brexit constituyó desde el principio un desafío para las firmas de medición de la opinión, carentes de experiencia previa que sugiriera posibles ajustes metodológicos para mejorar las estimaciones (Curtice, 2016a). El grueso de los sondeos predijo el triunfo del *Remain* –la opción por permanecer en la UE-, pero fue la alternativa opuesta –el *Leave*– la que recibió casi 52% de los sufragios.

Esta serie de fallos se ve reflejada en los grandes errores absolutos en el margen de ventaja de la Tabla 4, que son de 6,6 puntos porcentuales en 2015, 6,1 en 2016 y 5,4 en 2017. Sin embargo, estos niveles de inexactitud no eran desconocidos para la política británica. El máximo error –de 9,0 puntos- se produjo en la elección general de 1992, mientras que en 1970 había sido de 6,6 puntos.

Los ingentes errores de 2019 en Argentina solo son comparables con el observado en la segunda ronda de la elección francesa 2017, que Emmanuel Macron ganó por más de 30 puntos de diferencia sobre Marine Le Pen. Aunque las encuestas daban a Macron una gran ventaja, la diferencia final fue todavía mayor y el error en el margen ascendió a 10,2 puntos (Enten, 2017b).

De cara a la primera ronda, la contienda electoral francesa había sido considerada “demasiado cerrada” para pronosticar (Enten, 2017a). Los sondeos previos mostraban una diferencia de apenas cuatro o cinco puntos porcentuales entre el primer y el cuarto puesto. Pero los resultados concordaron casi exactamente con las previsiones, un hecho que fue recibido con alivio tras la inquietud sobre el futuro de la industria creada por los tropiezos en EE.UU. y el Reino Unido (Henley, 2017; Schultheis, 2017).

En este contexto, el grueso error en el margen de la segunda ronda pasó en gran medida desapercibido para el público y los círculos políticos. Prosser y Mellon (2018, pp. 765-67) –dos codirectores de la encuesta académica *British Election Study* (BES)- destacan que esta impresión distorsionada se debe a la idea de que las encuestas fallan cuando pronostican el ganador equivocado, en lugar de atender a la magnitud del error. Los sondeos franceses de la segunda ronda fueron mucho más inexactos que los de 2016 en EE.UU., pese a que éstos, a diferencia de los primeros, vaticinaron mal el triunfador.

### Márgenes de error reales y declarados

Una conclusión de este examen comparativo internacional es que los estudios preelectorales son hoy “tan (in)exactos como siempre” (Ibíd., p. 765). Sus niveles de error son y han sido en todas las épocas sustancialmente mayores que los declarados a partir de cálculos simplistas. Casi todos los márgenes de error reportados por las firmas comerciales se basan en la ecuación de cómputo para el muestreo aleatorio simple, que para poblaciones grandes -y con un nivel de confianza de 95%- es la siguiente (Groves et al., 2009, pp. 97-106; Cochran, 1977, pp. 18-30):

$$\text{Margen de error de muestreo (MOSE)} = \pm 1,96 \times \sqrt{\frac{p \times (1 - p)}{n}}$$

La abreviatura MOSE procede del inglés *Margin of Sampling Error*. Aquí  $p$  es la proporción –del voto para un partido, por ejemplo- estimada por un sondeo y  $n$  es el tamaño (número de entrevistados) de la muestra. La ecuación alcanza su valor máximo cuando  $p=0,50$  y este es típicamente el margen informado por las encuestadoras. Ilustrando con la última firma de la Tabla 1, *Management & Fit* (Clarín, 1/8/2019), el resultado es:

$$\begin{aligned} \text{MOSE} &= \pm 1,96 \times \sqrt{\frac{0,50 \times (1 - 0,50)}{2000}} \\ &= \pm 1,96 \times \frac{0,50}{44,71} = \pm 0,022 \text{ (ó } \pm 2,2 \text{ puntos porcentuales)} \end{aligned}$$

El nivel de confianza de 95% -implícito en el valor 1,96, derivado de la distribución estadística conocida como “campana de Gauss”- significa que, si la misma encuesta se realizara 100 veces, en 95 esperaríamos que la proporción “real” de votos de ese partido esté dentro del intervalo  $0,50 \pm 0,022$  –en porcentajes, entre 52,2% y 47,8%- surgido de la encuesta actual.

Una cuestión no siempre comprendida es que, como se desprende de la ecuación, el MOSE varía según la proporción de votos, disminuyendo para valores mayores y menores de 0,5. Usando el ejemplo de la misma firma, sería de  $\pm 2,1$ ,  $2,0$ ,  $1,8$  y  $1,3$  puntos para proporciones de votos de 0,4, 0,3, 0,2 y 0,1 respectivamente.

Una práctica incorrecta es el uso de los intervalos de confianza de los porcentajes de voto de dos partidos para inferir si la ventaja de uno sobre otro es o no significativa. Se suele asumir que, si los intervalos de cada partido se superponen, el orden “real” entre ambos podría ser inverso al que surge del estudio. Ejemplos de este razonamiento se encuentran en Clarín, 2/8 y 16/10/2019. Pero estimar el margen de error de la ventaja de un partido sobre otro es un problema específico, que requiere una aproximación estadística diferente (Prosser y Mellon, 2018, p. 782), como ilustramos en la nota (1) al final del trabajo.

A fin de tomar la muestra de individuos a entrevistar, las encuestas suelen emplear diseños complejos, cuyo MOSE es mayor que el informado usando como aproximación la ecuación del muestreo aleatorio simple (Cochran, 1977). Muchas, como la mayoría de las conducidas en Internet, no se basan en muestras probabilísticas, que dan a todos los miembros de la población la misma probabilidad de ser seleccionados. Para ellas, en rigor, no pueden calcularse márgenes de error de muestreo. En la mayor parte de los sondeos online y muchos de los telefónicos, el marco muestral –el listado exhaustivo de los miembros de la población, del que se toma la muestra- excluye a segmentos considerables de la sociedad que no tienen acceso a esas tecnologías y cuyos integrantes tienen así una probabilidad cero de ser elegidos.

En síntesis, los márgenes de error que informan las encuestadoras se circunscriben a los derivados del muestreo y tienden a ser subestimados al basarse en cálculos simplistas que ignoran complejidades de diseño, marcos muestrales deficientes y métodos de selección no probabilísticos. Esta práctica no ayuda a corregir la falsa idea por parte del público y los usuarios de los estudios de opinión de que el error muestral es la principal o única fuente de error. La investigación social ha estudiado desde hace décadas los distintos tipos de errores “no de muestreo” y los enfoques para minimizarlos (Biemer et al., 2017; Groves y Lyberg, 2010).

Los desvíos de los sondeos preelectorales han sido históricamente mayores que los márgenes de error informados. Un estudio realizado en EE.UU. a partir de 4.200 encuestas en más de 600 comicios de nivel estadual entre 1998 y 2014 arrojó que los errores de las predicciones duplicaron en promedio los márgenes reportados (Shirani-Mehr et al., 2017).

La conclusión es que si una encuesta declara, por ejemplo, un margen de +/- 3 puntos porcentuales –propio de una muestra típica en torno de los 1.000 casos-, el verdadero sería de alrededor de +/- 6 ó 7 puntos (Rothschild y Goel, 2016). Esta aproximación se basa en la observación empírica, pues en contraste con los errores de muestreo es difícil y quizá imposible construir una teoría estadística general sobre los otros componentes de error de las encuestas.

El promedio de los desvíos de las firmas argentinas en las Tablas 1 y 2 concuerda en buena medida con esta regla empírica. Una estimación del voto de 50% implica, según esto, que el porcentaje real rondaría, para tamaños típicos de muestra, entre un máximo de 56% y un mínimo de 44%. Esto indica que los sondeos no son lo bastante precisos para pronosticar el ganador en una competencia ajustada (Kennedy, 2020).

Examinando las fuentes de error no muestral en las encuestas electorales de EE.UU. y el Reino Unido, Prosser y Mellon (2018) hallan poca evidencia que apoye los argumentos sobre la influencia del voto vergonzante, “oculto” o “mentiroso”. Las causas más comunes han sido la falta de representatividad –en particular, la falla en representar adecuadamente a las personas no interesadas en la política-, votantes indecisos volcándose en una dirección dada y, a veces, modelos de participación electoral mal calibrados y cambios de último momento en la intención de voto.

Estas causas tienden a tomar la forma de “sesgo”, en la que los errores se producen en una sola dirección y afectan a todas las encuestas. También el estudio de Shirani-Mehr et al. (2017) observó que los sondeos en una elección dada tenían un componente de error que era común a todos ellos. La posible causa de ese error compartido era que todas las firmas enfrentaban dificultades similares para alcanzar a varios subgrupos de la población, además de aplicar procedimientos análogos para identificar a los probables votantes.

Un corolario es que, en una elección en la que se presenta este sesgo compartido, calcular un promedio de las distintas encuestas –una práctica popularizada por organizaciones como *FiveThirtyEight* y *RealClearPolitics*- no provee estimaciones más exactas, pues los errores no muestrales de cada relevamiento no se anulan entre sí y el error muestral real simplemente se aglomera alrededor de las estimaciones sesgadas (Prosser y Mellon, 2018, p. 767).

En EE.UU. las encuestas de 2020 -y las de 2016 en varios estados- tuvieron un sesgo a favor del partido demócrata, mientras que en 2000 y 2012 sobrestimaron el apoyo a los republicanos (Clinton et al., 2021; Kennedy et al., 2017, pp. 11-13).

En el Reino Unido los sondeos han mostrado una tendencia recurrente a subestimar el voto conservador desde 1992 (Sturgis et al., 2016, pp. 17-20). El gran yerro de ese año los movió a adoptar procedimientos para corregir el desbalance, que en buena medida –pero no totalmente- tuvieron éxito en las cuatro elecciones siguientes, hasta que la tendencia resurgió con fuerza en 2015. Una inclinación más variable a sobrestimar el voto laborista se ha observado desde 1979.

Tras el fracaso de 2015, las firmas británicas, siguiendo las recomendaciones del comité de expertos formado por el BPC, hicieron importantes cambios metodológicos para subsanar ambas tendencias (Ibíd., pp. 70-74). Pero en 2017, aún pronosticando correctamente el voto por los conservadores, se hizo evidente que habían compensado en exceso los fallos de la elección previa y subestimado esta vez el apoyo al laborismo (Curtice, 2017; Cowley y Kavanagh, 2018, pp. 259-83).

En las primarias de 2019 en Argentina, un sesgo favorable a Juntos por el Cambio afecta en mayor o menor grado a todos los sondeos de la Tabla 1. Es posible que los ajustes posteriores de las firmas hayan compensado esos errores en exceso, pues en la elección general, coincidiendo además con un aumento de la participación electoral, casi todas mostraron un sesgo a favor del Frente de Todos (Tabla 2).



### **La transformación de los estudios por encuesta**

La investigación de la opinión pública experimenta grandes cambios con impactos problemáticos. Hay una explosión de modos de recolección de datos más baratos, rápidos y fáciles de implementar, pero cuyos efectos no se conocen bien, como los que usan Internet y la Respuesta de Voz Interactiva (IVR), que administra un cuestionario por teléfono con una voz grabada.

Los paneles online se están convirtiendo en la técnica predominante, mientras persisten las dudas sobre su representatividad y exactitud (Clinton et al., 2021, p. 10; Fieldhouse y Prosser, 2018). Algunos están fundados en muestras probabilísticas de personas seleccionadas offline por métodos tradicionales, como el estudio piloto realizado en 2016 por los *American National Election Studies* (ANES) (DeBell et al., 2018). Pero la gran mayoría consiste en muestras no probabilísticas de individuos reclutados en línea a través de avisos, referencias de otros participantes y otros medios (Fielding, Lee y Blank, 2017; Callegaro et al., 2014; Tourangeau, Conrad y Couper, 2013; Baker et al., 2013). La base de miembros de estos paneles —o una sub-muestra tomada de ellos— es luego ponderada para hacerla corresponder con las características demográficas de la población.

Sin embargo, los encuestados que se auto-seleccionan para participar tienden a diferir de la población general en aspectos importantes, como la edad, la educación o los ingresos, variables asociadas con el uso de Internet (Mercer et al., 2017; Dutwin y Buskirk, 2017). Otro problema son los “profesionales” que se registran en múltiples paneles para acceder a las pequeñas compensaciones que se ofrecen como estímulo y que suelen llenar los cuestionarios con ligereza (Kennedy et al., 2020).

Hay una fuerte caída en las tasas de respuesta de las encuestas presenciales y las telefónicas realizadas en vivo por un entrevistador, en general por una combinación de fallas para hacer contacto con miembros de la muestra y de rechazos entre los que son contactados. La tendencia afecta también a los estudios académicos, que proveen este tipo de información. En la ANES, en 1952 respondió el 77% de los individuos seleccionados. En 2016 el porcentaje había descendido a 50% para el modo presencial y 44% para el estudio online (DeBell et al., 2018; Leeper, 2019). La tasa presencial del *British Election Study* (BES) cayó de 79% en 1963 a 56% en 2015 (Prosser y Mellon, 2018, p. 758). En los sondeos telefónicos del *Pew Research Center* este indicador bajó de 36% en 1997 a solo 9% desde 2012 (Keeter et al., 2017). Las encuestas cara-a-cara logran una participación mayor que los otros modos, en buena medida por la efectividad del contacto personal para persuadir a los individuos seleccionados (Groves et al., 2009).

El declive de las tasas de respuesta aumenta el riesgo de obtener resultados sesgados si las personas que contestan son diferentes de las que no lo hacen, fenómeno conocido como “sesgo de no-respuesta”. Por sus tasas sumamente bajas, este riesgo es mayor en el modo telefónico, al que ha migrado el grueso de los sondeos probabilísticos —usando generalmente la técnica de marcación aleatoria de números (*Random Digit Dialing* o RDD)- debido a los costos y el tiempo que insumen los relevamientos presenciales. A esto se suma el aumento de la

población que solo usa celular y no es cubierta por las encuestas telefónicas tradicionales basadas en líneas fijas (Lavrakas et al., 2017, 2007; Lepkowski et al., 2008; Brick et al., 2011). Aunque la incorporación de celulares en los marcos muestrales presenta desafíos, estudios del *Pew Research Center* muestran que los sondeos telefónicos con niveles muy bajos de respuesta no producen necesariamente bases de entrevistados poco representativas, en la medida que incluyan líneas fijas y móviles y se hagan ajustes para equipararlas al perfil demográfico de la población (Keeter et al., 2017).

Esto se debe a que la tasa de respuesta solo influye en el sesgo de no-respuesta cuando la probabilidad de contestar en un sondeo es más alta o más baja para ciertos segmentos de la población con características relevantes para el fenómeno que se mide. En los estudios electorales, un caso obvio es cuando los simpatizantes de un partido o candidato tienden en mayor medida que el resto a rechazar al encuestador o a colgar el teléfono tras recibir la llamada.

El tema de una encuesta suele afectar la probabilidad de tomar parte de ella. Las personas interesadas en la política son más proclives que las no interesadas a aceptar el requerimiento de contestar sobre ese tópico (Groves et al., 2000, 2004). Una muestra compuesta en forma desproporcionada por individuos políticamente implicados puede producir grandes errores en la estimación de los porcentajes de voto (Mellon y Prosser, 2017). A raíz de su proceso de auto-selección de los encuestados, los paneles online no probabilísticos están aún más expuestos a sufrir esta distorsión.

El análisis de las elecciones presidenciales de 2008, 2012, 2016 y 2020 en EE.UU. sugiere que el modo de encuesta –Internet, teléfono en vivo, IVR y combinaciones de estas tecnologías- no influyó de manera significativa en la exactitud de las estimaciones de voto (Clinton et al., 2021, pp. 25-28; Kennedy et al., 2017, pp. 14-18).

En EEUU y otros países muchas organizaciones han emprendido una transición a modos auto-administrados y/o mixtos de recolección de datos, combinando Internet, el teléfono, el correo postal y la entrevista presencial (Dillman, Smyth y Christian, 2014). La *American Association for Public Opinion Research* (AAPOR) elaboró informes que analizan las estrategias utilizadas en esa transición y ofrecen lineamientos para adoptar los nuevos modos (Olson et al., 2019; Lavrakas et al., 2017). Los objetivos que impulsan esta conversión son mejorar las tasas de respuesta, la cobertura de la población y la exactitud de las estimaciones, así como reducir costos.

La *European Social Survey* (ESS), un estudio académico transnacional con altos estándares metodológicos, lleva a cabo desde 2003 experimentos sistemáticos con modos mixtos de relevamiento basados en telefonía e Internet para investigar su efecto en la calidad de las mediciones entre y dentro de los países relevados (Villar y Fitzgerald, 2017). Los resultados muestran la mayor confiabilidad y validez de los datos generados por el modo presencial, que la ESS implementa con criterios muy rigurosos, utilizando el muestro probabilístico en cada una de las etapas –incluso dentro del hogar- y sin admitir el reemplazo de ninguna unidad de

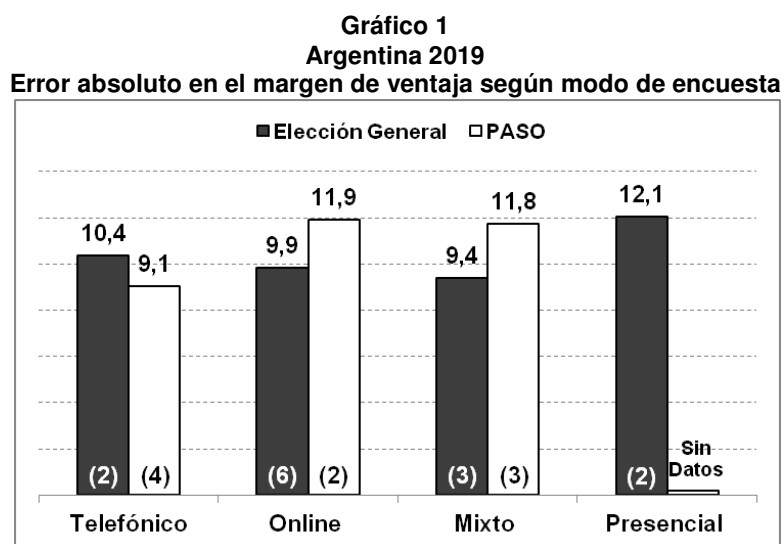
análisis seleccionada (Jowell et al., 2007). El estudio fija requisitos mínimos de duración del trabajo de campo y tasas de respuesta.

El cambio hacia una diversidad de modos en los países participantes de la ESS es un desafío complejo, pues las mediciones deben ser comparables entre las naciones y a lo largo del tiempo. Los experimentos arrojan que cada modalidad puede tener efectos específicos en las respuestas de los entrevistados. Esto hace necesario introducir correcciones para que los datos recogidos con distintas tecnologías sean comparables. La ESS ha optado por enfocarse en la posibilidad de crear un panel exclusivamente online, con miembros reclutados de una muestra probabilística de áreas.

La proliferación de modos de recolección de datos rápidos y de bajo costo ha conducido a un gran aumento del número de encuestas, realizadas a su vez por una pluralidad de organizaciones, pero esto no se ha traducido en un mejor retrato de las preferencias del público.

Este último fue el escenario de las elecciones de 2019 en la Argentina, en las que hubo una cantidad sin precedentes de encuestadoras. Con la información disponible de las Tablas 1 y 2 es posible obtener una apreciación genérica del impacto del modo de encuesta sobre la exactitud de las predicciones (Gráfico 1).

Los errores absolutos en el margen de ventaja son elevados para todos los modos. La escasa información sobre el diseño de los sondeos impide indagar si una o más de estas tecnologías tiene pese a todo un efecto diferencial estadísticamente significativo.



**Nota:** Número de encuestas entre paréntesis. Datos de Tablas 1 y 2

Una estrategia para hacerlo sería calcular una regresión con el error en el margen como variable dependiente, incluyendo entre los predictores a los modos discriminados de una manera más fina –IVR, teléfono en vivo, etc.–, junto a factores como el diseño muestral y la proximidad al día de la elección (Kennedy et al., 2017, pp.14-18). En el análisis bivariado del

Gráfico 1, el posible impacto específico de cada tecnología está confundido con el de estos y otros posibles factores adicionales.

### **Los errores no de muestreo en EE.UU. y el Reino Unido**

La investigación reciente ha aumentado en forma sustancial nuestra comprensión de las fuentes de error de las encuestas, aunque esas causas constituyan un blanco móvil (Prosser y Mellon, 2018, p.780). Estos mecanismos de error no son exclusivos de las metodologías actuales y es probable que hayan afectado a las mediciones de las actitudes y conductas políticas desde su surgimiento hace casi un siglo.

Por haber sido objeto de intenso escrutinio, los casos de EE.UU. y el Reino Unido son especialmente reveladores. Tras la fuerte reacción pública contra las encuestas que siguió al triunfo de Trump en 2016, la AAPOR creó un comité de expertos para investigar su desempeño (Kennedy et al., 2017, 2018). Trabajando con bases de datos de entrevistados provistas por encuestadoras seleccionadas y con información detallada de todos los sondeos proveniente de tres entidades que los compilaron –*FiveThirtyEight*, *RealClearPolitics* y *HuffPollster*–, el equipo contrastó empíricamente las principales teorías sobre las razones por las que gran parte de los estudios subestimaron el apoyo a Trump.

Según la evidencia reunida, las dos causas más relevantes fueron una falla generalizada en ajustar las muestras obtenidas para corregir una sobre-representación de entrevistados con educación universitaria –en su mayoría votantes de Clinton– y un movimiento de último momento a favor de Trump por parte de los indecisos y simpatizantes de terceras opciones.

La primera de las causas está relacionada con el sesgo de no-respuesta. Está documentado que la probabilidad de aceptar participar en un sondeo es menor para las personas con bajo nivel educativo, un rasgo coincidente con perfiles de los votantes de Trump. Este posible desbalance no produciría estimaciones sesgadas si al procesar los datos se incluye una variable de ponderación que otorgue el peso adecuado en el resultado final al segmento educativo sub-representado, siempre que las personas menos instruidas entrevistadas no difieran de las no entrevistadas en otros rasgos relevantes.

Muchas encuestas en EE.UU. han ponderado regularmente sus bases de entrevistados para equipararlas con diversos perfiles demográficos de la población, pero no con la distribución del nivel educativo. Históricamente esto no produjo errores porque la educación no tenía una relación lineal con el voto a presidente. En 2012, por ejemplo, Obama tuvo el apoyo mayoritario de los menos y los más instruidos, mientras aquellos con educación media se repartieron en partes iguales entre el entonces presidente y Mitt Romney. Aunque ese año los menos y los más educados hayan estado respectivamente sub y sobre-representados en las encuestas, el hecho de que ambos grupos votaran del mismo modo encubrió el desbalance.

En 2016, sin embargo, cuanto mayor era la educación del elector, más probable era que votara por Clinton. La falla en asignar el peso apropiado a los distintos segmentos educativos derivó así en importantes errores en los pronósticos de voto. Fue una omisión que cometieron la

mayoría de los sondeos en estados decisivos –Michigan, Wisconsin, Pensilvania, Florida, Ohio y Carolina del Norte- y la mitad de los de alcance nacional.

La segunda causa fundamental de error tiende en cambio a atenuar la responsabilidad de las encuestadoras. Hay evidencia de que en estados clave donde Trump ganó por diferencias mínimas –Michigan, Wisconsin, Pensilvania y Florida- y, en menor medida, en el nivel nacional, un alto porcentaje de los encuestados que no habían decidido o revelado su voto o se inclinaban por un tercer partido sufragaron finalmente por Trump. Esto sugiere que los sondeos reflejaban bien la realidad del momento en que se realizaron y que fallaron al proyectar el resultado final. Los datos que apoyan estas conclusiones proceden de una encuesta de boca de urna del *National Election Pool* –un consorcio formado por grandes cadenas y organizaciones de noticias- y un sondeo telefónico del *Pew Research Center* con re-llamadas posteriores a los comicios.

El equipo de expertos de AAPOR halló solo efectos pequeños e inconsistentes para una tercera hipótesis, referida al posible impacto de modelos defectuosos de participación electoral en la exactitud de las predicciones. Identificar a los votantes probables es un problema complejo, que las encuestadoras abordan con enfoques disímiles mediante una combinación de ciencia y arte (Kennedy et al., 2017, pp. 42-4).

En 2016 la participación electoral cambió respecto de 2012 en modos que beneficiaron a Trump. La concurrencia a las urnas de los afroamericanos, el grupo más inclinado por Clinton, cayó 5% a nivel nacional y más pronunciadamente en estados como Michigan y Wisconsin, que fueron críticos para el colegio electoral. Algunas encuestadoras, que construyen sus modelos de participación basándose en los patrones de la elección previa, podrían haber subestimado por ello el apoyo a Trump. Pero el ejercicio de comparar en estados clave los votos obtenidos por Trump y Clinton con las predicciones de los sondeos antes y después de ser modificadas por sus modelos de votantes probables no mostró que éstos generaran errores significativos.

La indagación halló aún menos respaldo para una hipótesis popular: la del voto vergonzante por Trump, cuya causa general –planteada hace décadas- es la posible renuencia sistemática de los entrevistados a expresar su apoyo por un candidato controvertido. Dado que la raíz de este sesgo es lo que el encuestado percibe como una respuesta socialmente indeseable, los sondeos telefónicos administrados en vivo por un entrevistador deberían mostrar una mayor subestimación del voto por Trump que los conducidos online en forma auto-administrada. Esta diferencia no pudo observarse en un análisis de regresión que usó como insumo todos los sondeos publicados desde dos meses antes de los comicios y que mantuvo constantes otras características de esos estudios. De modo más concluyente, las encuestas en estados competitivos no solo subestimaron el voto por Trump, sino también por los candidatos republicanos al senado y la cámara de representantes. Esto sugiere que el error de las predicciones no derivó de la deseabilidad social.

El informe de AAPOR subrayó que los principales errores observados en 2016 provinieron de encuestas estatales que contaban a menudo con bajos presupuestos. En 2020, la mayoría de estos estudios ponderaron sus datos por nivel educativo. Además, el caudal de indecisos, la

otra fuente importante de error en la contienda entre Trump y Clinton, fue esta vez mucho menor. Sin embargo, pese al esfuerzo invertido en dilucidar y corregir los problemas surgidos en 2016, en 2020 el desempeño de la industria empeoró (Clinton et al., 2021; Clinton, 2021; Keeter y Clinton, 2021; Merkle, 2021; Keeter et al., 2020). Los sondeos sobrestimaron mucho más la ventaja demócrata y lo hicieron de manera aún más acentuada en las elecciones de senadores y gobernadores que en la competencia presidencial.

Un conjunto de factores únicos contribuyó en esta ocasión a dificultar la tarea de las encuestadoras. Junto a la anomalía de la pandemia del Covid-19 y la alta proporción de votos anticipados y por correo, la participación electoral creció siete puntos respecto de 2016, alcanzando su máximo nivel en más de un siglo (Desilver, 2021).

La investigación de un nuevo comité de expertos de AAPOR arrojó que las dos principales causas de error en 2016 -una falla en ponderar por niveles educativos y un movimiento tardío de votantes a favor de Trump- no explican esta vez el fracaso de 2020 (Clinton et al., 2021). Tampoco influyeron los modelos de votantes probables, un presunto voto vergonzante o un desbalance en las muestras entre los encuestados que sufragaron tempranamente o el día de la elección.

Aunque los datos provenientes de más de 2.800 encuestas permitieron descartar estas hipótesis sobre las posibles fuentes de error, el estudio concluyó que era "imposible" determinar las causas precisas de los desvíos con la información disponible (Ibíd., p. 5). Planteó empero varias "suposiciones fundadas", a contrastar empíricamente en investigaciones ulteriores. Al menos una parte de los errores se originó en ciertas formas de sesgo de no-respuesta. Pero al carecer de datos sobre los votantes que se negaron a responder a los sondeos, no fue posible establecer en qué rasgos diferían de los que sí respondieron.

Si las encuestas sub-representaron a los simpatizantes más firmes de Trump, esto pudo ocurrir, por ejemplo, porque los republicanos más proclives a votar por sus propios candidatos estaban menos dispuestos a responder que los simpatizantes de ese partido más abiertos a votar por los demócratas. En este caso, aunque las encuestas hayan incluido la proporción correcta de personas que se identificaban como republicanos, las diferencias entre los que respondieron y los que no lo hicieron explicarían por qué la ventaja demócrata resultó sobrestimada.

Abona esta hipótesis el hecho de que, según estudios previos, la caída de la confianza pública en las instituciones y las encuestas ha sido más pronunciada entre los republicanos. Además, Trump criticó reiteradamente a los sondeos -que, con raras excepciones, le eran muy desfavorables-, calificándolos de "fake news" destinadas a desalentar a sus adherentes de concurrir a votar. Según el informe de AAPOR, para quienes apoyaban al presidente esto pudo haber convertido la decisión de no responder en un "acto político".

Es posible empero que el sesgo de no-respuesta se haya producido en varios modos distintos y no excluyentes entre sí. Uno es simplemente que los republicanos en general hayan sido menos proclives que los demócratas a participar en las encuestas. Otro es que en 2020 Trump haya atraído a las urnas a un grupo de nuevos votantes, sin adhesión partidaria y poco

dispuestos a ser entrevistados. El hecho de que los resultados de las elecciones de medio término de 2018 -en las que el entonces presidente no fue candidato- hayan sido anticipados por la industria con exactitud, sugiere que los desvíos de 2020 podrían ser un fenómeno pasajero, que no volvería a presentarse una vez que Trump deje de ocupar el centro de la escena (Clinton et al., 2021, p. 72; Merkle, 2021). Si, por el contrario, ha surgido una tendencia a no responder a los sondeos en ciertos tipos de conservadores, el problema sería más profundo y de compleja solución.

En el Reino Unido, el equipo de expertos creado por el BPC para investigar el fallo de 2015 concluyó que la “causa primaria” de error fueron las “muestras no representativas” (Sturgis et al., 2016, pp. 47-58). Debido a “los modos en que las muestras fueron recolectadas y ajustadas”, éstas “sobre-representaron en forma sistemática a los votantes laboristas y sub-representaron a los conservadores en sus estimaciones ponderadas”.

Los investigadores tuvieron a su disposición las bases de datos de los sondeos preelectorales de los miembros del BPC –y de los poselectorales de esas mismas firmas en los que se recontactaron a los encuestados previos-. También recurrieron a dos encuestas académicas posteriores a la elección –que arrojaron buenas estimaciones de la diferencia entre conservadores y laboristas-, realizadas por los estudios periódicos BES y *British Social Attitudes* (BSA) (Mellon y Prosser, 2017; Curtice, Phillips y Clery, 2015).

El BES y el BSA son sondeos presenciales basados en el muestreo probabilístico de áreas y su trabajo de campo se extiende por varios meses. Los hogares seleccionados que no logran ser relevados inicialmente son objeto de repetidos intentos posteriores y en ningún caso son reemplazados. Los datos recogidos se ponderan buscando corregir por la desigual selección de hogares e individuos y por posibles sesgos de no-respuesta (Ibíd., pp. 118-122).

Examinando estos dos estudios, Sturgis et al. (2016) comprobaron que, considerando solo a las personas entrevistadas al primer o segundo intento –y ajustando los datos con las ponderaciones demográficas estándar que usan las encuestas comerciales-, las estimaciones de voto en la elección de 2015 resultaban sesgadas a favor del laborismo. Esto ilustra la forma en que una muestra defectuosa –de personas dispuestas a responder en un trabajo de campo realizado en un breve lapso de tiempo- podía producir estimaciones sesgadas, aún haciendo ajustes para equiparar la base de datos al perfil demográfico de la población. Curtice (2016c) arribó a conclusiones similares con los datos del BSA, notando que los votantes laboristas parecían ser más accesibles para los encuestadores.

El informe del BPC subrayó además que en el Reino Unido muchas encuestas telefónicas que declaran usar métodos de marcación aleatoria de números (RDD) utilizan en realidad técnicas de muestreo no aleatorias, pues aún las que se basan en directorios combinados de teléfonos fijos y móviles no verifican que esas listas reflejen adecuadamente la distribución real en la población de los subgrupos que usan uno o ambos tipos de línea. Esto es importante porque diferentes grupos demográficos tienen distintos patrones de uso de estas tecnologías.

La investigación destacó las “muchas causas potenciales de no-representatividad en las muestras” y la dificultad de “identificar el ‘modelo’ correcto” que explicara lo sucedido en 2015.

Pero Mellon y Prosser (2017) mostraron que la principal causa de falta de representatividad –y del error general de los pronósticos- fue un tipo específico de sesgo de no-respuesta. Las personas que no tenían intención de votar rechazaron a los encuestadores en mayor medida que el resto y quedaron así sub-representadas en las muestras. Este problema alcanzó sobre todo a los segmentos de la población inclinados al laborismo, en especial los jóvenes. Muchos no votantes de estos grupos no respondían a las encuestas y fueron reemplazados en las muestras por personas de esos mismos grupos que votaban por los laboristas, inflando así el voto estimado para este partido. Los autores llegaron a este resultado comparando el sondeo poselectoral presencial del BES con la versión online del mismo estudio, en la que los no-votantes quedaron sub-representados.

### **Ahondando en las causas de error**

En su balance sobre el estado actual de las encuestas electorales, Prosser y Mellon (2018, pp. 771-4) destacan que las muestras no representativas son la causa más común de las fallas en los pronósticos. Los problemas de representatividad surgen en dos etapas de los estudios. Primero en la fase de muestreo, por un sesgo de no-respuesta –o por el que surge de una cobertura defectuosa de la población, por ejemplo de la que solo usa celular-. Después, en la fase de ponderación de los datos, si los pesos asignados a las distintas categorías de estratificación de las personas –edad, sexo, educación, voto anterior, etc.- fallan en corregir los sesgos de cobertura y no-respuesta o, en ciertas situaciones, los introducen.

A su vez, el sesgo de no-respuesta puede deberse a que la probabilidad de responder a un sondeo es menor: 1) para personas de ciertos grupos demográficos, como las del segmento de baja instrucción en EE.UU, que en 2016 apoyaba en su mayoría a Trump; 2) para ciertas personas dentro de un mismo grupo demográfico, como ocurrió en el Reino Unido en 2015 con los no votantes de segmentos proclives al laborismo e, hipotéticamente, en EE.UU. en 2020 con determinados tipos de republicanos.

El sesgo en el caso 1) es fácil de corregir ponderando las respuestas del grupo sub-representado de acuerdo con su peso en la población. Fue lo que hizo la mayoría de las encuestadoras con el grupo de educación baja en la elección de 2020 en EE.UU.

Pero en el caso 2) es difícil conocer el sesgo con antelación y disponer de información demográfica para ponderar los datos. Si en 2015 las encuestadoras británicas ignoraban que muchos jóvenes que no respondían no irían a votar y ponderaron las respuestas de los que sí contestaron aumentando su peso para equipararlas al tamaño de todo ese grupo etario en la población, lo que hicieron fue introducir sesgo al sobre-representar el voto joven –en su mayoría laborista- en relación al de otros segmentos de edad.

La capacidad de corregir el sesgo de las encuestas mediante la ponderación de los datos descansa en el supuesto de que, dentro de cada categoría que se pondera –edad, educación, identificación partidaria, etc.-, las personas que responden no difieren en aspectos importantes de las que no lo hacen. Solo recientemente se ha advertido en qué medida el sesgo de no-respuesta, especialmente el que se origina en los no-votantes y las personas con escaso



interés por la política o con baja confianza social o institucional, es causa de error en las estimaciones del voto.

Aunque este tipo de sesgo puede hacerse evidente después de la elección –una vez conocidos los niveles de participación electoral y de apoyo a cada partido-, no es fácil de solucionar con anticipación. Prosser y Mellon (2016) proponen incluir entre los criterios de ponderación la probabilidad de votar de cada entrevistado y el porcentaje esperado de participación electoral el día de los comicios, a fin de ajustar la muestra obtenida para asemejarla no a la población general, sino a la que concurre a votar. Algunas encuestadoras estadounidenses usan procedimientos de este tipo (Kennedy et al., 2017, pp. 42-44).

Las encuestadoras suelen atribuir sus errores de pronóstico a la llamada “participación electoral diferencial”, cuando los simpatizantes de distintos partidos concurren a votar en diferentes proporciones. Pero el problema parece originarse más a menudo en modelos de participación “mal calibrados” (Prosser y Mellon, 2018, p. 771). Algunos sondeos preguntan al entrevistado cuál es su probabilidad de votar y ponderan las intenciones de voto a partir de ese dato, o usando un esquema más sofisticado basado en información demográfica. Los pronósticos se verán afectados si estos ajustes predicen mal la participación de los seguidores de los diferentes partidos. Se ha observado que los encuestados tienden a exagerar tanto su posibilidad de ir a sufragar como, a posteriori, el hecho de haber votado.

Los modelos de participación no condujeron a errores significativos en las elecciones de 2016 y 2020 en EE.UU., pero fueron una causa fundamental del fallo de 2017 en el Reino Unido, cuando produjeron peores estimaciones que los datos originales antes de ser ajustados por estos criterios (Sturgis y Jennings, 2017).

Los votantes indecisos que terminan inclinándose en grandes números por un partido o candidato son quizá la segunda causa más frecuente de error de las predicciones, después de las muestras no representativas. Movimientos de este tipo a favor de Trump en 2016 y de los conservadores británicos en 1992 contribuyeron a los errores de los pronósticos (Kennedy et al., 2017; Jowell et al., 1993).

El término “indeciso” suele encerrar ambigüedad, pues los encuestados que dicen “no saber” a quién votarán pueden estar genuinamente indecisos u ocultar su preferencia. Para proyectar el voto de este grupo, el método usado por defecto es removerlo simplemente de la muestra e informar los porcentajes de intención de voto de quienes manifestaron una opción. Esto equivale a asumir que ese grupo votará de la misma manera que el de los ya decididos. Algunas encuestadoras usan procedimientos basados en el voto en la elección anterior, el partido de preferencia del entrevistado y otros criterios. El informe del BPC destacó que los métodos empleados en 2015 en el Reino Unido para proyectar el voto de quienes “no saben” o “no contestan” eran “*ad hoc* y carentes de fundamentos teóricos coherentes”.

El concepto de “cambio tardío” del voto (“*late swing*”) no está definido en forma consistente en la literatura, donde suele incluir los movimientos de última hora de los indecisos y otras conductas (Sturgis et al, 2016, p. 34). Aquí lo utilizamos en el sentido restringido de los encuestados que “cambian de opinión”, es decir, que votan por un partido distinto del que

habían declarado (Prosser y Mellon, 2018, p. 768). Aunque este fenómeno es citado con frecuencia como causa de los malos pronósticos, en la mayoría de los casos los estudios de panel –que encuestan a las mismas personas antes y después de una elección- no han encontrado evidencia que respalde la hipótesis. Sin embargo, este giro de los votantes ha tenido efectos considerables en algunos casos, como la elección general de 1970 en el Reino Unido y, en menor medida, la de 1992 (Abrams, 1970; Jowell et al., 1993).

Los encuestados que ocultarían su voto por vergüenza son a menudo, según Prosser y Mellon (2018, p. 769), “más acusados que culpables”, aunque hay evidencia que sugiere su presencia “al menos en algunas ocasiones”. La hipótesis para la que ninguno de los dos estudios de AAPOR halló respaldo en EE.UU. ha sido planteada también en el Reino Unido para explicar la tendencia a subestimar el voto a los conservadores. Según el argumento, muchos británicos piensan que las políticas de este partido benefician a los más favorecidos a expensas de los necesitados, de modo que algunas personas serían renuentes a decir que lo votan porque podrían ser vistas como más preocupadas por su ventaja personal que por el bien público (Sturgis et al., 2016, p. 39).

Ni la investigación del BPC ni el análisis de Mellon y Prosser (2017) con datos poselectorales del BES hallaron prueba de este fenómeno en la elección de 2015. Algunos sugerían que una pregunta de intención de voto realizada al final –no al principio- de un cuestionario sobre opiniones políticas obtendría una respuesta más consistente con las actitudes del encuestado, revelando mejor el voto conservador. Pero un experimento en la versión online del BES, que manipuló aleatoriamente la ubicación de esa pregunta, no arrojó diferencias significativas. Los sondeos comerciales mostraron poca variación tanto en el orden de la pregunta como en sus estimaciones del apoyo a los conservadores.

Si los encuestados ocultaran su voto conservador antes de la elección –no contestando, o diciendo que “no saben” o que votarán a otro partido-, pero dieran una respuesta honesta después, el voto vergonzante se manifestaría como un movimiento de último momento de los indecisos o un proceso de “cambio tardío”. Analizando los datos de los sondeos comerciales que re-contactaron a poco de la elección a sus entrevistados previos, los expertos del BPC no observaron ningún giro significativo de este tipo. El hecho de que los relevamientos cara-a-cara poselectorales del BES y el BSA –conducidos tras los comicios durante varios meses- arrojaran buenas estimaciones del voto, sugiere que los simpatizantes conservadores no eran renuentes a admitir su preferencia una vez conocido el resultado.

Los datos del BES mostraron además que en los distritos eminentemente laboristas las estimaciones del voto no se desviaban del resultado electoral en mayor medida que en los distritos conservadores, en contra de lo que esperaríamos si las respuestas de los encuestados proclives al conservadorismo hubieran estado condicionadas por el estigma social.

Aunque estos contrastes no permiten descartar en forma definitiva la existencia de un voto vergonzante, sugieren que es improbable que éste haya sido un factor significativo en los errores de pronóstico en 2015.

Un hecho denunciado a menudo es la conducta de “manada” (“*herding*”), cuando las encuestadoras ajustan sus sondeos u omiten difundirlos si sus resultados no coinciden con el consenso existente. Aunque no produce necesariamente un aumento del error –pues las firmas pueden aglomerarse alrededor del resultado correcto-, esta práctica tiende a reducir de modo artificial la variabilidad de las encuestas, creando así una falsa confianza en el desenlace electoral y amplificando la sorpresa si éste difiere del esperado. El signo de un posible movimiento de manada es que las estimaciones de las distintas encuestadoras discrepan cada vez menos –y por debajo de lo que cabría esperar- a medida que se acerca la fecha de los comicios. En términos estadísticos, su varianza disminuye más de lo que esperaríamos por un eventual aumento del tamaño de las muestras hacia el final de la campaña.

Los expertos del BPC hallaron evidencia consistente con este fenómeno en la elección de 2015, si bien la reducción de la varianza en la última semana no fue estadísticamente significativa debido al escaso número de observaciones (Sturgis et al, 2016, pp. 64-68). Entre los posibles ajustes efectuados por las encuestadoras para alinear sus resultados con las expectativas, el informe menciona cambios en las ponderaciones de post-estratificación y de participación electoral, así como aumentos adicionales del tamaño de la muestra.

Según el estudio, el tipo de conducta de manada que mejor explica los pronósticos sesgados de 2015 –tanto el error sistemático de los sondeos como la similitud de los márgenes de ventaja que reportaron- es una convergencia de las firmas hacia una diferencia estrecha entre conservadores y laboristas –buscando minimizar el riesgo de una “predicción equivocada” del partido ganador-, más que una confluencia hacia el promedio colectivo de los últimos sondeos publicados.

### **Hipótesis sobre el caso argentino**

Los hallazgos recientes sobre las fuentes de error de las encuestas proveen un marco para plantear hipótesis acerca de las posibles causas de la falla generalizada de los pronósticos en las elecciones argentinas de 2019. La revisión de investigaciones en las secciones previas deja empero al descubierto limitaciones fundamentales de la industria de las encuestas y los estudios académicos sobre comportamiento electoral en nuestro país. Al no disponer de un corpus suficiente de conocimiento acumulado sobre la conducta electoral de los argentinos, como tampoco de la evidencia empírica adecuada para contrastar hipótesis, el estudioso enfrenta serias restricciones para esclarecer aspectos como los móviles del voto y el desempeño de los sondeos en una elección determinada.

La información disponible sugiere que una causa principal del mal desempeño de los sondeos de 2019 han sido las muestras no representativas, debido tanto a una cobertura defectuosa del electorado como también, probablemente, al sesgo de no-respuesta.

Luego de las primarias de agosto, algunas encuestadoras, conscientes de sus problemas de cobertura –debido en especial a la exclusión de usuarios de telefonía móvil y a las limitaciones de sus paneles online-, decidieron combinar distintas tecnologías de administración del cuestionario, incluyendo entrevistas presenciales. Vimos sin embargo que los modos mixtos de

relevamiento tienen sus propios inconvenientes, pues las diversas modalidades pueden suscitar respuestas disímiles de los entrevistados. Además, al combinar marcos muestrales diferentes –por ejemplo, de usuarios de teléfonos fijos, celulares e Internet- es necesario tomar recaudos para que los grupos demográficos que difieren en el uso de esas tecnologías –en función de la edad, la educación, el nivel de ingreso y otros factores- queden representados en la encuesta según su distribución real en la población.

Varias firmas modificaron sus muestras sobre la base del voto en las primarias, aunque no es claro si esto envolvió un esquema de ponderaciones de post-estratificación. Un encuestador consultado por Clarín (27/9/2019) puntualizó que esas elecciones, previas a la general, “sirven para armar tu base de encuestados. Les preguntás cómo votaron en las PASO y buscás que tu base del sondeo sea proporcional en ese sentido”. Según La Nación (5/10/2019), un consultor “de alto perfil” que usaba la tecnología IVR incorporó a su muestra de teléfonos la localidad de Virrey del Pino –en el municipio de La Matanza-, donde Alberto Fernández había logrado en agosto una enorme ventaja sobre Macri. Este caso sugiere que por lo menos algunas muestras telefónicas no fueron aleatorias, aunque contuvieran un elemento aleatorio en la selección de números.

Al tratar de explicar el fallo de las PASO, hubo encuestadores que pusieron el foco en el sesgo de no-respuesta (Clarín, 28/8/2019). Uno apuntó específicamente a quienes “estaban descontentos con el gobierno de Macri pero tampoco estaban muy convencidos de votar al kirchnerismo” como el grupo que habría sido más renuente a participar de los sondeos y conducido al sesgo en las estimaciones.

Varias firmas se inclinaron por subrayar como fuente de error en las primarias un movimiento de último momento de los indecisos a favor del Frente de Todos (Ibídem; La Nación, 12/8/2019). Un argumento frecuente fue que el comportamiento de los indecisos era difícil de predecir, pero se tradujo mayoritariamente en un voto de “castigo al gobierno” por la mala situación económica. Otros alegaron que este grupo encerraba un alto porcentaje de voto vergonzante por el Frente de Todos –es decir, de personas que no estaban realmente indecisas, sino que ocultaban su preferencia por creerla socialmente indeseable.

Fueron recurrentes las versiones sobre encuestadoras que actuaron en “manada”. De acuerdo con una fuente de la industria citada por La Nación (23/8/2019), existió “una comunidad que consensuó que había paridad” y “cuando a una encuestadora los números le daban distinto, lo consideraba un error o no le creían”. Otra fuente declaró a Clarín (27/9/2019) que “como la mayoría de las encuestas y los medios hablaban de un brecha de 3 ó 4 puntos, era difícil abstraerse y empezás a dudar de tus propios números”.

Los responsables de las firmas coincidieron en que la calidad de los sondeos se había visto resentida por los bajos presupuestos y que muy pocos clientes estaban dispuestos a afrontar los altos costos de las encuestas presenciales. Sin embargo, esta modalidad tampoco es inmune a los problemas de representatividad, no solo debido a sesgos de no-respuesta, sino también a muestras defectuosas por la dificultad de contactar, en el corto espacio de tiempo de los relevamientos preelectorales, a muchos de los hogares e individuos seleccionados en el

diseño muestral original. De hecho, los trabajos presenciales no parecen haber tenido en 2019 un mejor desempeño que los telefónicos u online.

El estudio de Jennings y Wlezien (2018) indica que la idea de una “era dorada” de las encuestas dominada por los relevamientos cara-a-cara no se ajusta a los hechos. Los niveles de (in)exactitud de la industria no son más grandes hoy que en el pasado y es probable que muchas de las causas de error sean similares.

En la Argentina también encontramos un caso de yerro impactante de los pronósticos hace más de treinta años, cuando estas estimaciones procedían solo de sondeos presenciales. Las elecciones legislativas y de gobernadores del 6 de septiembre de 1987 estuvieron rodeadas de gran expectativa por sus implicaciones para la carrera presidencial de 1989. El presidente Raúl Alfonsín había logrado un fuerte respaldo en las legislativas de 1985, pero el apoyo a su gobierno estaba en declive a raíz de los problemas económicos. La contienda de mayor trascendencia en 1987 era la de gobernador de la provincia de Buenos Aires, no solo por el peso del distrito, sino porque el candidato del PJ, Antonio Cafiero, lideraba el movimiento de renovación peronista y era el principal aspirante presidencial de su partido de cara a 1989.

La mayoría de las encuestas –que en pocos años, desde 1983, habían adquirido gran prominencia pública- anticipaba el triunfo del postulante a gobernador de la UCR, Juan Manuel Casella, según algunas con una ventaja de dos dígitos. Cuando Cafiero ganó por casi 7 puntos, la industria se vio envuelta en una gran controversia. “¿Sirven o no las encuestas?”, tituló la revista *Mercado Publicitario*, en un extenso informe de tapa donde los responsables de firmas líderes dieron sus explicaciones (Ritacco, 1987).

Los argumentos apuntaron a dos posibles fuentes de error: las muestras imperfectas y las dificultades para proyectar el voto de los indecisos. Según el sociólogo Manuel Mora y Araujo, “hubo una cantidad de encuestas hechas sin muestras rigurosas y con métodos no científicos, que se han publicado profusamente”. Raúl Hernández –que en 1987 hizo estudios en la provincia de Río Negro- estimó que el 90% de las muestras realizadas en el país eran “baratas” y de “mala calidad”, en muchos casos “sin respetar el principio de aleatoriedad”. “Yo los he visto encuestar y noté que están siempre muy apurados, reemplazan todo muy rápidamente, demasiado fácilmente. Encuestadores apurados y mal pagos, entonces” (Ibíd., p. 13).

Guillermo Bravo y Esther Kaplan, cuyas firmas habían pronosticado en 1983 el triunfo de Alfonsín –en el que no muchos creían-, fallaron esta vez al vaticinar la victoria de Casella por 13 y 18 puntos respectivamente. Ambos atribuyeron su error a la proyección de los indecisos. Los procedimientos para adjudicar el voto a los miembros de este grupo fueron intuitivos y específicos del contexto. Bravo se basó en el puntaje que asignaban los indecisos a la gestión del gobierno nacional; Kaplan, en las razones que daban para dudar entre votar al PJ o la UCR. La segunda adjudicó más votos a la UCR porque este partido era objeto en menor proporción de razones negativas inmodificables, como los “malos candidatos” (Ritacco, 1987, p. 12).

Hay indicios de que en 2019 se seguían aplicando estos mismos enfoques *ad hoc*. Cuando Clarín (27/9/2019) pidió un pronóstico final al encuestador que a la postre anticipó mejor el

resultado de las PASO, el consultado, según el diario, “se la jugó: ‘Pasame todos los indecisos a Alberto’”. El responsable de una encuestadora establecida estimó en un programa de televisión que los indecisos se volcarían en las primarias hacia Juntos por el Cambio, debido a que, de todos los presidentes de la democracia, Alfonsín era el que ostentaba la mejor imagen entre los miembros de este grupo.

Pero el voto económico de castigo al gobierno, que en 2019 parece haber tomado por sorpresa a las encuestadoras, era el más probable de los escenarios, según todo lo que sabemos sobre las teorías del voto y los estudios empíricos sobre comportamiento electoral conducidos en la Argentina y América Latina, igual que en los países de la periferia de Europa que aplicaron políticas de austeridad tras la crisis financiera de 2008-2009.

Dentro de una teoría del voto económico que, desarrollada por más de cuatro décadas, ha ganado creciente complejidad (Lewis-Beck y Stegmaier, 2013; Anson y Hellwig, 2015), la hipótesis que ha recibido mayor apoyo empírico en los estudios internacionales –realizados en gran parte, pero no solamente, en países de altos ingresos, en especial EE.UU., Reino Unido y Francia- sostiene que los votantes actúan sobre todo de un modo “retrospectivo” –evaluando el desempeño pasado del gobierno- y “sociotrópico” –juzgando el estado de la economía nacional más que el de sus bolsillos.

Hay contextos en los que el voto puede ser “prospectivo” y/o “egotrópico”, es decir, basado en las expectativas sobre la evolución futura de la economía y/o en el estado de la propia billetera. Se sabe de factores capaces de afectar el impacto de la economía sobre las decisiones de voto. Uno es el grado en que resulta o no claro a los ojos de los ciudadanos si el gobierno es responsable por la situación económica. Administraciones a cargo de coaliciones de varios partidos, o condicionadas por organismos supranacionales –como los de la Unión Europea-, podrían reducir la “claridad de la responsabilidad” y debilitar el peso del voto económico. Además, sería posible que los simpatizantes del partido de gobierno evaluaran el estado de la economía de manera distinta que los de la oposición.

La teoría del voto económico tiene importantes consecuencias normativas, pues subraya el hecho de que los ciudadanos son efectivamente capaces, por medio del sufragio, de hacer responsable al gobierno por los resultados de su gestión. El principio de “rendición de cuentas” o “*accountability*” está en el corazón de la teoría de la democracia.

Las investigaciones sobre América Latina, enfocadas en la región en conjunto o en países particulares, arrojan que, por una confluencia de factores, la evaluación retrospectiva del desempeño económico del gobierno está al frente de las consideraciones de los latinoamericanos a la hora de decidir el voto (Bustikova y Zechmeister, 2017; Nadeau et al., 2015a; Lewis Beck y Ratto, 2013).

El voto basado en la ideología o en temas políticos, exigente en términos cognitivos –tanto que no es pronunciado aún en democracias desarrolladas-, se ve todavía más limitado en nuestra región por sistemas partidarios volátiles y fragmentados, aunque la polarización política actúa como contrapeso de esa tendencia menguante. Los frágiles sistemas de partidos generan adhesiones partidarias igualmente débiles, aumentando la susceptibilidad de los votantes a los

resultados económicos. El nivel relativamente bajo de desarrollo y la alta desigualdad del ingreso hacen muy visibles las vulnerabilidades económicas, de modo que en las campañas electorales la economía ocupa a menudo un lugar central. Estos dos factores –la visibilidad general y el énfasis en la campaña- son importantes para el nivel de voto económico. Las recesiones y la inestabilidad vuelven a la economía aún más saliente y agudizan la susceptibilidad del electorado.

Según Bustikova y Zechmeister (2017, p. 101), en América Latina “el espectro de la economía es tan significativo que, en ciertas circunstancias, puede inducir memorias de atípica duración en votantes que rechazan no solo al mandatario en ejercicio, sino a todos los partidos establecidos”. Su relevancia es tan grande que condiciona la de otros factores. El ciudadano promedio castiga al gobierno por los malos resultados económicos con independencia de si lo percibe o no como corrupto, pero es menos probable que le retire el apoyo a causa de la corrupción percibida si esos resultados son buenos (Zechmeister y Zizumbo-Colunga, 2013).

Los presidentes fuertes que caracterizan a la región -y el grado en que controlan el poder legislativo- aumentan la “claridad de la responsabilidad”, ayudando a atribuir al ejecutivo la culpa por los malos tiempos y acentuando el voto económico en general. Esta tendencia es atemperada por las restricciones que la globalización impone a los gobiernos –derivadas, entre otros factores, de la apertura comercial, los flujos financieros, la inversión extranjera directa y organismos como el Fondo Monetario Internacional (FMI)-. Sin embargo, se ha observado que los votantes latinoamericanos suelen no advertir el influjo de factores exógenos como los cambios en los precios de los commodities y en las tasas de interés internacionales, que impactan sobre las economías nacionales pero escapan al control del gobierno (Campello y Zucco, 2015; Carlin y Hellwig, 2019). En estos casos el electorado castigaría o premiaría a una administración atribuyéndole responsabilidad por efectos que no dependen mayormente de sus políticas.

Aunque el volumen de investigaciones específicas sobre Argentina no es grande –aun siendo uno de los países más estudiados de la región-, la fuerte incidencia del voto económico retrospectivo y sociotrópico puede considerarse suficientemente establecida (ver en especial Nadeau et al., 2015b; Tagina, 2012 y Ratto, 2013; Gervasoni y Tagina, 2019). La inflación, recesión y creciente pobreza que signaron la segunda mitad de la gestión de Macri, unidas a las duras políticas de austeridad implementadas en 2018 tras el masivo rescate del FMI, hacían de la reelección del presidente un objetivo improbable frente a un peronismo unificado.

Otra clara señal en la misma dirección era la experiencia reciente del sur de Europa con sus programas de ajuste tras la crisis de 2008-2009. Por exigencia del FMI y los organismos de la Unión Europea, o por presión de los mercados financieros internacionales, España, Portugal, Grecia e Italia –además de Irlanda- adoptaron medidas de austeridad muy impopulares que provocaron una “epidemia electoral” y creciente inestabilidad política (Bosco y Verney, 2012, 2016; Talving 2017). En 2011, los gobiernos de los cinco países fueron desalojados. Analizando estos casos, Lewis-Beck y Costa Lobo (2017, p. 607) concluyeron que, en comparación con tiempos normales, una crisis económica aguda “intensifica el voto económico

en el nivel micro del individuo y, finalmente, en el nivel macro del resultado de la elección nacional”.

Pese a estos antecedentes, camino a las PASO los grandes medios de comunicación de nuestro país instalaron en el público –a través de un consistente mensaje por parte de periodistas, analistas y encuestadores- la idea de una competencia equilibrada.

Las principales columnas políticas de los diarios de mayor difusión reflejaron ese consenso. En vísperas de las primarias, Carlos Pagni (2019a) subrayaba que “pocas veces un presidente al que se imputa una pésima situación económica está en condiciones de competir con una razonable posibilidad de ser reelegido”. Citando a una fuente de *Isonomía* –una encuestadora cercana a Macri-, destacaba que “quienes creen que un gobierno de Alberto Fernández y Cristina Kirchner les mejoraría mucho su nivel de vida son más numerosos que los votantes de esa fórmula”. Y atribuyendo al electorado una mirada prospectiva, antes que retrospectiva, concluía: “Para muchos ciudadanos, la amenaza del malestar político es más poderosa que la promesa del bienestar económico”. El día siguiente a los comicios, el mismo analista remarcaba que “el oficialismo vio desmentidas sus principales hipótesis” (Pagni, 2019b).

Joaquín Morales Solá (2019) interpretó que “un voto oculto” mostró “una realidad muy distinta de la que habían percibido las mediciones de opinión pública más prestigiosas”. Y agregaba: “El ajuste más severo que se haya hecho desde la crisis de 2001/2002 no podía pasar inadvertido en una elección presidencial sin cambiar todos los manuales políticos y electorales”. En un comentario titulado “El lapidario peso de la crisis económica”, Eduardo Van der Kooy (2019) puntualizó que “la estrategia de polarización que el macrismo ensayó obstinadamente contra Cristina Fernández” no había rendido los frutos esperados. “Fue, sin dudas, la crisis económico-social el detonante de la durísima derrota del gobierno en las PASO de este domingo”. La revista *The Economist* (2019) resumió: “La búsqueda de un segundo mandato por parte de Macri siempre apareció quijotesca luego de que la economía entrara en problemas el año pasado. Hoy los argentinos están peor que hace cuatro años”.

Los muchos encuestadores que frecuentaban los programas de televisión para interpretar los procesos políticos y de opinión pública de los que surgía la paridad electoral que arrojaban sus números tenían motivos para ser más cautelosos. En un artículo publicado en *Clarín* el 1 de agosto, donde explicaba la medición de *Management & Fit* que daba solo dos puntos de ventaja al Frente de Todos, la responsable de la firma afirmaba que “si bien la gente vota por preferencias, ideología y religión, la situación económica es relevante. Los números duros de la macro no son muy alentadores para el gobierno, pero la estabilidad financiera de los últimos meses ayudó en las encuestas” (Fornoni, 2019).

Parte del problema reside en que los sondeos utilizados como herramienta de pronóstico electoral “no se basan en ninguna teoría del voto” para predecir el resultado, sino que “proveen meramente estimaciones de una variable dependiente” (Lewis-Beck, 2005, p. 148). Los análisis fundados en el campo de estudio científico del comportamiento electoral estuvieron poco menos que ausentes en los comentarios y debates de los medios.



Con el desenlace de las PASO, los mismos analistas, dando un giro de 180 grados, arribaron precipitadamente a un nuevo consenso: Alberto Fernández estiraría en la elección de octubre su ya abrumadora ventaja sobre Macri. Pero si, como se esperaba, la participación electoral aumentaba en la elección general respecto de las primarias, no era tarea sencilla anticipar cuánto lo haría y cómo se comportarían los nuevos votantes. No solo la participación creció casi 5 puntos porcentuales –de 76% a 81%-, sino que los votos en blanco y anulados se redujeron sustancialmente en términos absolutos. Esto produjo un aumento de más de dos millones de votos por alguno de los candidatos a presidente, que en su gran mayoría optaron por Macri.

Es plausible que la estrategia de polarización del presidente y su campaña posterior a las PASO llamando a forzar una segunda vuelta –a la que sumó el énfasis en temas políticos salientes como la oposición al aborto y la “mano dura” contra el delito- haya contribuido a morigerar el peso del voto económico entre los nuevos votantes, así como a captar un voto “útil” o “estratégico” proveniente de los pequeños partidos de derecha, que vieron reducido su caudal. Indagar empíricamente la influencia relativa de estos y otros posibles factores requeriría de investigaciones por encuesta –preferiblemente de panel- que relevaran datos sobre las variables apropiadas.

Es claro que las encuestadoras no habían hecho un buen diagnóstico de sus problemas en las primarias. Veamos como ejemplo dos hipótesis sobre aspectos que quizá no recibieron suficiente atención. Por un lado, los modelos de participación electoral. Las firmas argentinas no detallaron los procedimientos que usaron en las dos elecciones para incorporar en sus pronósticos la probabilidad de votar de los entrevistados y/o el nivel de participación previsto. Y hemos observado que calcular las estimaciones de voto incluyendo a encuestados que no irán a votar puede ser una causa significativa de error.

Un segundo aspecto –con un impacto probablemente más grande que el anterior- es el sesgo de no-respuesta. Un encuestador especuló con la posibilidad de que un segmento de votantes insatisfecho con el gobierno pero poco atraído por el Frente de Todos haya sido reacio a participar en los sondeos previos a las PASO (Clarín, 28/8/2019). No es evidente sin embargo cuál es la posible conexión entre esta causa de indecisión del voto y la propensión a rechazar la entrevista. Si se está pensando en un hipotético estigma social asociado en algunos sectores con el voto por el kirchnerismo –que conduciría, más que a la negativa a ser entrevistado, al ocultamiento de la intención de voto-, una vía para indagar su existencia sería comprobar si en los distritos muy inclinados a Juntos por el Cambio –como la Ciudad Autónoma de Buenos Aires y Córdoba- los sondeos subestimaron el apoyo al Frente de Todos en mayor medida que en los distritos históricamente peronistas, como el Conurbano bonaerense.

Una hipótesis razonable de sesgo de no-respuesta es que al menos una parte de quienes no votaron en las PASO pero lo hicieron en octubre eran personas proclives a Juntos por el Cambio que tendían a no responder a las encuestas, tal vez por estar menos interesadas en la política. Si así fuera, hay que preguntarse además si se trataba de un grupo demográfico

específico, o de ciertos miembros de distintos grupos que diferían en su conducta de voto de los miembros que sí respondían.

Es muy probable que los errores en cada una de las dos elecciones hayan sido el producto de una combinación de causas. Pero las posibles explicaciones solo pueden dilucidarse a través de investigaciones independientes, como las examinadas en los casos de EE.UU y el Reino Unido.

Tres factores conspiran en la Argentina contra esa posibilidad. Uno es la falta de una asociación profesional que reúna a los expertos del campo de la opinión pública y la investigación por encuesta, como la AAPOR estadounidense y el BPC británico. Entidades de este tipo fijan estándares de calidad y difusión de datos de los sondeos y de las organizaciones que los realizan, para promover la transparencia y confiabilidad de los resultados. Los comités de expertos que establecen para producir informes sobre temas específicos o, en ciertas situaciones, para examinar el desempeño de los sondeos electorales, tienen acceso a los datos necesarios de los trabajos de las firmas comerciales.

El segundo factor que impide investigar en profundidad no solo los errores de esas mediciones, sino las actitudes, móviles y conductas de los ciudadanos durante los ciclos electorales, es la ausencia de estudios periódicos por encuesta con objetivos académicos que releven datos sistemáticos en cada uno de esos ciclos. Dos influyentes modelos son la ANES de EE.UU., que implementan las universidades de Michigan y Stanford, y el BES británico, de las universidades de Manchester y Oxford. Estos trabajos proveen datos comparables en el tiempo de variables políticas y electorales clave a lo largo de muchas décadas. En democracias maduras hay organizaciones independientes sin fines de lucro –como el *Pew Research Center* en EE.UU.- que también conducen investigaciones de tipo académico.

El tercer aspecto es el escaso desarrollo en nuestro país del campo de estudio científico de la opinión pública en general y del comportamiento electoral en particular. Es un déficit vinculado en parte con el del punto anterior. Los investigadores se ven forzados a recurrir a encuestas transnacionales que incluyen a la Argentina en sus ondas periódicas –como *Latinobarómetro*, *Barómetro de las Américas* y *World Values Survey* (Jorge, 2015, 2020)-, a realizar sus propios trabajos –una opción limitada por los problemas de financiamiento- e incluso a apelar a bases de datos de sondeos comerciales. Otros dos modelos de estudios académicos sobre las actitudes y conductas de la población en una variedad de tópicos, ampliamente utilizados en investigación, son la *General Social Survey* (GSS) de la Universidad de Chicago en EEUU y la ya mencionada BSA del *National Centre for Social Research* (NatCen) en el Reino Unido.

### **El impacto de las encuestas y la gestión de la incertidumbre**

En un panel de encuestadores organizado poco después de las PASO por un foro empresarial, el representante de una firma establecida pidió al periodismo: “Déjense de hinchar con las encuestas. Pueden generar un click pero alteran el escenario electoral” (Clarín, 28/8/2019). Otro subrayó: “La encuesta no es un oráculo. Y su lugar tiene que ser menos relevante”.

Son argumentos paradójicos en miembros de una industria donde la calidad se ve resentida por la falta de clientes dispuestos a pagar por un trabajo bien hecho. Perder relevancia no es la vía para desarrollar el mercado.

El fondo de la cuestión reside quizás en una peculiaridad de los sondeos de intención de voto que ha sido destacada por Jowell et al. (1993, p. 239). “En virtualmente ningún otro campo de la investigación por encuesta los profesionales deben pasar por un examen tan inmediato de su exactitud. La mayoría no pasa nunca por ninguna prueba. Si lo hicieran, sin duda muchas encuestas –aún las conducidas a un ritmo pausado, con más presupuesto y empleando métodos más rigurosos- también fallarían. Hay tantas cosas que pueden salir mal con las encuestas que es asombroso que alguna vez salgan bien, y sin embargo lo hacen”.

La solución de los problemas que los estudios de opinión defectuosos plantean para la democracia no pasa por eludir este examen o reducir su visibilidad. Todo sugiere que la baja calidad de las encuestas locales –que las elecciones de 2019 pusieron en evidencia- alcanza también a los estudios que proveen el mapa de las opiniones de los argentinos sobre los temas políticos, económicos y sociales más importantes, y que influyen por esa vía en las decisiones sobre políticas de gobierno y en la imagen que la sociedad tiene de sí misma (Jorge, 2018).

La posibilidad de tener un retrato distorsionado de las opiniones del público es un foco de preocupación en democracias donde ese riesgo es menor que en la Argentina. Un artículo del *Pew Research Center* tras la elección de 2020 en EE.UU. resaltó que si los sondeos “subrepresentan sistemáticamente a ciertos tipos de conservadores o republicanos, esto tiene ramificaciones para las mediciones de otras conductas y temas” (Keeter et al., 2020).

Pensemos, por ejemplo, en los estudios que nos muestran un ranking de los “problemas que más preocupan a los argentinos”, los porcentajes de la población que está a favor y en contra de una decisión de política económica, o una lista de políticos con “mejor imagen”. Como destaca el artículo citado, este tipo de encuestas no requiere el nivel de exactitud de 51%-49% necesario en una competencia presidencial, pero “ningún encuestador quiere un desvío sistemático de sus datos, incluso si es de ‘solo’ 5 puntos porcentuales”.

El *Pew Research Center* llevó a cabo un ejercicio de simulación para explorar el grado en que los resultados de los sondeos sobre temas de la agenda pública podrían verse distorsionados por errores “no de muestreo” como los registrados en 2020 en EE.UU. (Keeter et al., 2021). El estudio utilizó un conjunto de encuestas online de la propia organización, realizadas en 2020 entre más de 10 mil miembros de su *American Trends Panel*, una muestra probabilística de toda la población adulta del país basada en el listado nacional de direcciones residenciales. Cada uno de los integrantes de este panel fue entrevistado dos veces por mes a lo largo del año sobre una variedad de tópicos. Tras la elección presidencial, se le preguntó además si había votado y por quién.

Con esta base de encuestados, se empleó la técnica de ponderación de datos para generar dos muestras diferentes, que simulaban dos escenarios alternativos. En la muestra “equilibrada”, Biden tenía una ventaja en intención de voto sobre Trump de 4,4 puntos porcentuales -la diferencia en el voto popular que arrojó la elección-. En la muestra “sesgada”,

esa ventaja de estiraba a 12 puntos, la máxima observada entre todas las encuestas de las dos semanas previas a los comicios. Dado que el panel de entrevistados tenía además casi un 40% de personas que no votaron, el estudio construyó su muestra equilibrada incluyendo a no votantes proclives a los demócratas y republicanos en igual proporción, mientras que en su muestra sesgada dio una ventaja de 10 puntos porcentuales a los no votantes demócratas.

Una primera conclusión fue que, en un electorado dividido con poca diferencia entre dos candidatos, pequeños errores en la composición de la muestra pueden producir grandes cambios en el margen de ventaja estimado entre ambos contendientes. Partiendo, por ejemplo, de una muestra equilibrada de 1.000 entrevistados, basta que solo 38 votantes cambien su preferencia de Trump a Biden para que la ventaja de este último aumente de 4,4 a 12 puntos porcentuales, produciendo así un error en el margen de 7,6 puntos.

Pero el principal resultado fue que una distorsión de esta magnitud en la representación de los republicanos parece tener consecuencias menos severas al estimar las opiniones sobre temas de interés público. Comparando las mismas muestras equilibrada y sesgada, el error absoluto en el margen de "ventaja" entre los porcentajes de aprobación y desaprobación de la gestión de Trump fue de 5 puntos porcentuales, 2,6 puntos menos que entre los porcentajes de intención de voto de Biden y Trump. El error fue también de 5 puntos para opiniones sobre el control de la pandemia del Covid-19 y el problema del racismo, y de 4 puntos para las relacionadas con el tamaño y el rol deseado del Estado, el impacto de la inmigración y la influencia de la actividad humana sobre el cambio climático.

La mayoría de las actitudes analizadas arrojó empero desviaciones de menor magnitud. Sobre un total de 48 preguntas, el error absoluto en el margen de ventaja entre respuestas alternativas fue, en promedio, de 2 puntos porcentuales. El estudio subraya que un desvío de esta magnitud, que puede trastocar -como ocurrió en 2016- el pronóstico de una elección ajustada, no altera de un modo dramático, en la mayoría de las circunstancias, nuestra evaluación del estado de la opinión pública sobre un tema de la agenda.

En contraste con la medición de las intenciones de voto, es raro que podamos obtener una radiografía acabada de los puntos de vista sociales sobre un tema político con una sola pregunta. Normalmente debemos utilizar varias, que aborden distintos aspectos de la cuestión. Además, las respuestas en materia de opinión se ven más influidas que el voto por el modo como una pregunta es formulada. Más importante que determinar con gran exactitud el porcentaje de la población que está, por ejemplo, a favor del aborto legal, es saber si esa posición es sostenida por una mayoría grande o una minoría pequeña, si aumenta o disminuye con el tiempo o si exhibe una división generacional.

¿Por qué un sesgo en la representación de los adherentes de los distintos partidos tiene un impacto menor al estimar sus opiniones que al hacerlo con sus intención de voto? La razón es que las opiniones de los ciudadanos distan mucho de estar perfectamente alineadas con las posiciones de los partidos con los que simpatizan. Y entre los no votantes, esa falta de alineación puede ser aún mayor. En EE.UU., según el estudio del *Pew Research Center*, casi la cuarta parte de quienes apoyaban a Biden preferían un Estado más pequeño, una posición no

asociada tradicionalmente con el partido demócrata. A la inversa, un tercio de los simpatizantes de Trump tenía una visión positiva de los inmigrantes y un cuarto pensaba que el gobierno federal debía garantizar la cobertura de salud a todos los estadounidenses. Estas "inconsistencias" -que eran aún más grandes entre los no votantes con inclinaciones hacia uno u otro partido- debilitan el impacto de esta clase de sesgos de una muestra sobre la medición de las opiniones.

La mayor limitación este ejercicio de simulación es que descansa en el supuesto de que la base de encuestados del *American Trends Panel* representa adecuadamente al universo nacional de simpatizantes republicanos y demócratas y de votantes de Trump y Biden. El hecho de que esta base esté integrada por personas dispuestas a responder cada quince días, a lo largo de un año, un cuestionario sobre asuntos políticos, sugiere que la muestra puede estar expuesta a sesgos de no-respuesta como los antes analizados en nuestro trabajo. Si así fuera, es probable que los errores en la medición de las opiniones arrojados por la simulación estén subestimados.

Al evaluar las derivaciones de esta investigación para el caso de la Argentina, advertimos que el máximo error en el margen de ventaja observado en 2019 en nuestro país fue de 14 puntos porcentuales, casi el doble del que usó el ejercicio del *Pew Reserch Center* para construir su modelo de muestra sesgada. Más aún, la mitad de las encuestas de nuestras Tablas 1 y 2 tiene errores en el margen de más de 11 puntos. Es probable, por lo tanto, que las mediciones sobre las opiniones de la sociedad argentina sufran distorsiones mucho más grandes que las estimadas en la simulación descripta.

El impacto negativo para la democracia de una radiografía inexacta de las actitudes de los argentinos es potenciado en nuestro país por la falta de una asociación profesional, que fije no solo normas de calidad, sino además un código de ética (ver, por ejemplo, AAPOR, 2015). Las firmas comerciales realizan encuestas sobre asuntos públicos que conciernen a sus clientes y en la Argentina, como en otros países, algunas pueden difundir un cuadro de temas y opiniones del público sesgado hacia los intereses y la agenda de quienes contratan sus servicios. El riesgo de que el público sea engañado aumenta cuando estas encuestadoras divulgan sus datos e interpretaciones en los medios presentándose como analistas independientes. El proceso de deliberación colectiva sobre temas trascendentes para la sociedad puede verse así seriamente trastornado.

En casi todas las democracias las encuestas contribuyen en alto grado a conformar la narrativa de los medios sobre la marcha y los vaivenes de una elección. Concluido el ciclo electoral 2020, los grandes medios de EE.UU., que financian y promueven estos sondeos, volvieron a reevaluar el modo de presentarlos. Algunos editores del *New York Times*, por ejemplo, piensan que lo mejor podría ser darles menos prominencia, "a pesar del intenso interés de los lectores y del papel dominante que juegan en el diseño de las estrategias de campaña" (Leonhardt, 2020).

La prensa norteamericana ha sufrido iguales o peores fiascos en el pasado, cuando las mediciones preelectorales la indujeron a transmitir al público un relato distorsionado de lo que

estaba ocurriendo. En 1980 casi todos los sondeos pronosticaban una carrera muy cerrada entre Ronald Reagan y Jimmy Carter, quien buscaba su reelección. Las mediciones finales de Gallup y *CBS News / New York Times* daban solo un punto de ventaja para Reagan, que ganó por casi diez. El periodista Bill Green, que ejercía dentro del *Washington Post* el puesto – común por entonces- de defensor de los lectores del diario, se preguntó si la prensa no debía abandonar los sondeos (Campbell, 2020, pp. 108-30).

Al declarar márgenes de error inferiores a los reales, las encuestadoras no ayudan a que los periodistas y el público adviertan los niveles de incertidumbre que entrañan sus resultados. Lo sucedido en 2019 en la Argentina sugiere que incluso algunas firmas locales tampoco los advertían. Al cierre de los comicios de las PASO, encuestadores invitados a los programas de televisión restaban seriedad a los primeros datos de sondeos de boca de urna que anticipaban una importante ventaja de Alberto Fernández.

El creciente volumen de encuestas, así como la aparición de organizaciones que las compilan y promedian –“*polling aggregators*”- y de pronosticadores electorales, no han mejorado la exactitud de las estimaciones. Los pronósticos calculados a partir de un gran número de sondeos crean empero una falsa impresión de estabilidad y certidumbre sobre nuestro conocimiento de la opinión pública. En 2012, el modelo estadístico del fundador de *FiveThirtyEight*, Nate Silver, predijo correctamente el resultado presidencial en los 50 estados, pero en 2016 asignó a Hillary Clinton una chance de ganar de 71%. Para el influyente modelo del *Huffington Post*, las chances de Clinton crecían a 98,2% (Campbell, 2020, pp. 1-20).

Prosser y Mellon (2018, p. 781) notan que las encuestas cuyas estimaciones se desvían del punto hacia el que converge la mayoría pueden ser engañosas en términos de porcentajes de voto, pero son “indicadores importantes de la variación y la incertidumbre que es inherente” a este tipo de estudios. “Encuestadores, pronosticadores y periodistas necesitan lograr un equilibrio comunicando tanto la información real que proveen los sondeos como la incertidumbre que los rodea”.

Para algunos científicos, el “analfabetismo estadístico” es un problema social generalizado, que se vuelve peligroso en contextos como la pandemia del Covid-19, cuando gobiernos y ciudadanos deben usar el razonamiento probabilístico para sugerir conductas y evaluar la posibilidad de contagiarse en actividades ordinarias (Rovelli, 2020; Goodwin, 2020).

Silver (2017b) considera que los medios tienen “un problema de probabilidad”, pues a menudo “han sobrestimado la precisión de las encuestas, seleccionando los datos” que mejor se ajustaban a su narrativa y “presentando una elección como cosa segura, cuando esa conclusión no era respaldada por los sondeos u otra evidencia empírica”. Silver encuentra una similitud con los pronósticos meteorológicos sobre el impacto de los huracanes. “Antes de que llegue una tormenta, los medios demandan una precisión imposible, ignorando las incertidumbres de los pronósticos y magnificando ciertos escenarios (como el huracán azotando a Miami) a expensas de otros casi igualmente probables”.

Una práctica prevalente en los medios del Reino Unido son los titulares que exageran las consecuencias de pequeños cambios de los porcentajes de intención de voto –no significativos

estadísticamente- en el rumbo de una elección (UK House of Lords, 2018). Aquí la prensa interpreta en exceso un mero “ruido” estadístico de origen aleatorio, de un modo que coincida con su narrativa de la campaña.

La cobertura periodística de elecciones recientes –incluyendo la interpretación y selección de los datos de encuesta- parece haber sido afectada además por el fenómeno popularizado con el nombre de “burbuja mediática” (Silver, 2017a; Shafer y Doherty, 2017).

Según una hipótesis plausible, tras los cambios experimentados en años recientes, los medios tradicionales y digitales de EE.UU. han adquirido un fuerte sesgo progresista, al haberse concentrado a lo largo de las costas en distritos demócratas con ingresos superiores al promedio, formando redacciones con escasa diversidad de género y racial, en las que más de 90% de los reporteros son graduados universitarios y menos de 10% se identifican como republicanos. Estas características habrían confluído en 2016 para crear un “pensamiento de grupo” –la inclinación psicológica de los miembros de un grupo a conformarse al consenso colectivo-, sesgando la visión de los periodistas de un modo que los llevó a exagerar las chances de Clinton y minimizar las de Trump. Un proceso similar habría conducido a los medios británicos, basados principalmente en Londres, a subestimar las posibilidades de la opción por abandonar la Unión Europea en el referéndum del Brexit, que era apoyada en gran proporción por votantes de menores ingresos e instrucción –y de mayor edad- residentes fuera de la capital del Reino Unido.

En la Argentina, es probable que un fenómeno análogo haya sesgado la cobertura de la campaña previa a las primarias de 2019 por parte de los medios nacionales, que están concentrados en la Ciudad Autónoma de Buenos Aires. La tendencia política preponderante del distrito y sus altos niveles de ingreso, que influyen en la composición de los planteles de periodistas de mayor visibilidad, sumados a la línea editorial de las grandes empresas periodísticas, pueden haber contribuido a la creación de un “pensamiento de grupo” proclive a sobrestimar las posibilidades de Macri. Este consenso distorsionado alcanzó también a muchos encuestadores que, con su alta exposición pública, establecieron una relación casi simbiótica con los medios. La conducta de manada habría hecho el resto, ayudando a propagar al conjunto de la industria esta visión alterada de la competencia electoral.

Un último punto a considerar es el impacto de los sondeos electorales erróneos sobre las decisiones políticas y económicas. En 2020 Trump optó por reducir las visitas y la publicidad en Michigan y Wisconsin, dos estados donde las encuestas le eran muy desfavorables, pero en los que finalmente perdió por diferencias mínimas. En el Reino Unido, según los testimonios recibidos por el Comité Especial sobre Encuestas Políticas creado por la Cámara de los Lores para analizar los fallos recientes (UK House of Lords, 2018), tanto los sondeos privados como los publicados, igual que la narrativa de los medios, han tenido vastos efectos en las estrategias políticas. Los pronósticos errados contribuyeron a muy malas decisiones, como el llamado a elecciones anticipadas por parte de Theresa May en 2017. Pero equívocos análogos vinculados al cronograma de las elecciones generales se encuentran ya a fines de los 70 en la administración del primer ministro James Callaghan y en 2007 en la de Gordon Brown.

En nuestro país, las dos grandes coaliciones se vieron sorprendidas por el desenlace de las primarias de 2019, mientras el resultado de octubre estuvo muy por debajo de las expectativas del Frente de Todos.

Con el derrumbe de los mercados luego de las PASO, la agencia *Bloomberg* (17/8/2019) destacó que “el cambio tectónico en la evaluación del riesgo sobre la Argentina fue tan violento porque las encuestas inexactas significan que muy pocos lo vieron venir, y ahora son los inversores los que deben sopesar las ramificaciones”. La última rueda de la bolsa de Buenos Aires antes de las primarias había vivido un día de euforia, en lo que algunos vieron como una operación orquestada mediante la filtración de sondeos favorables al gobierno (Clarín, 9/8/2019; La Nación, 12/8/2019).

La campaña posterior a las primarias del Frente de Todos se vio afectada por la creencia en una victoria arrasadora en octubre, como anticipaban todas las encuestas. A días de la elección general, *Bloomberg* (16/10/2019) consideró que “la carrera presidencial ha perdido algo de intensidad, mientras un confiado Fernández dedica más tiempo a construir su reputación en el exterior en lugar de participar en eventos de campaña locales”. Esto pudo haber tenido consecuencias para la cosecha final de legisladores nacionales y otros cargos provinciales y municipales obtenidos por esta coalición.

### **Conclusiones**

Un corolario de nuestro análisis es que los problemas de las encuestas electorales y políticas en nuestro país, que el grueso fallo de 2019 ha mostrado en toda su magnitud, constituyen un aspecto disfuncional de la democracia argentina, que llama a una revisión en profundidad y una reforma del funcionamiento de la industria. En particular, si las firmas comerciales no se autorregulan por medio de una asociación profesional, el público, la política y otros usuarios tienen razones para dudar de la transparencia y confiabilidad de los datos y análisis que producen.

El argumento de que las encuestas “están en crisis” en todo el mundo elude la cuestión y no está respaldado por la evidencia empírica. Algunas democracias asisten en rigor a una pérdida de confianza pública en este tipo de estudios, originada en gran parte en la dificultad que encuentran las encuestadoras, medios y analistas para reconocer, comprender y comunicar la incertidumbre asociada a las estimaciones y pronósticos electorales.

Pese a lo que algunos comentaristas se han apresurado a afirmar, “la investigación sugiere que estamos lejos de asistir al ocaso de las encuestas” (Prosser y Mellon, 2018, p. 781). “Si encuestadores e investigadores siguen estudiando las fuentes de sesgo y proveen posibles soluciones, podemos en realidad estar en los albores de una nueva era en la exactitud de los sondeos”.

El uso de preguntas que midan la confianza interpersonal y política de las personas –que parece influir en la disposición a responder a las encuestas-, las entrevistas a través de mensajes de texto y la experimentación con nuevos métodos, como el de Regresión Multinivel y Post-estratificación (MRP) –que podría ser especialmente útil para las muestras no



probabilísticas- (Lax y Phillips, 2009), son solo algunos ejemplos de las ideas y líneas actuales de desarrollo.

Pero las falencias de nuestro conocimiento científico sobre las actitudes y la conducta política de los argentinos exceden el ámbito estrecho de la industria de las encuestas. El desarrollo académico de este campo de investigación –que incluye el área de los comportamientos electorales- es aún incipiente en nuestro país, no menos que el de los estudios de opinión pública en general. La implementación de encuestas académicas periódicas sería un paso decisivo para el desarrollo del campo, al proporcionar una fuente de datos que atraería el interés de los investigadores del país y del exterior. Avanzar en esta dirección forma parte de las tareas largamente demoradas en una democracia que entra ya en su madurez.

## Notas

(1) La diferencia entre dos proporciones sigue la ley de suma de las varianzas. La ecuación de cálculo del margen de error de la diferencia entre las proporciones de voto  $p_1$  menos  $p_2$  de dos candidatos es, con un nivel de confianza de 95%, la siguiente:

$$\text{Margen de error de la diferencia } p_1 - p_2 = \pm 1,96 \times \sqrt{\frac{[p_1 \times (1 - p_1)] + [p_2 \times (1 - p_2)]}{n}}$$

Siguiendo con el ejemplo de la última encuestadora de la Tabla 1, *Management & Fit* (Clarín, 1/8/2019), que asignó 41,2% de los votos al Frente de Todos y 39,2% a Juntos por el Cambio, el cómputo es:

$$= \pm 1,96 \times \sqrt{\frac{[0,412 \times (1 - 0,412)] + [0,392 \times (1 - 0,392)]}{2000}}$$

El resultado es un margen de +/- 0,03. La regla de decisión es que si el intervalo de confianza de la diferencia de proporciones contiene el valor cero, esa diferencia no es estadísticamente significativa –y, si no lo contiene, sí lo es-. Aquí,  $p_1$  menos  $p_2$  es igual a 0,02 (0,412 menos 0,392). El intervalo de confianza 0,02 +/- 0,03 se extiende desde un mínimo de -0,01 a un máximo de 0,05, incluyendo el 0. Concluimos que la ventaja del Frente de Todos sobre Juntos por el Cambio estimada por este sondeo no es estadísticamente significativa.

## Bibliografía citada

AAPOR (2015). *The Code of Professional Ethics and Practices*. American Association for Public Opinion Research, November 30. <https://www.aapor.org/Standards-Ethics/AAPOR-Code-of-Ethics.aspx>

- Abrams, M. (1970). The Opinion Polls and the 1970 British General Election. *Public Opinion Quarterly*, 34(3), pp. 317–324.
- Anson, I.G. y Hellwig, T. T. (2015). Economic Models of Voting. En Scott, R. A. y Kosslyn, S. (eds.) (pp. 1-14). *Emerging Trends in the Social and Behavioral Sciences*. Hoboken, NJ.: John Wiley.
- Baker, R., Brick, J.M., Bates, N.A., Battaglia, M., Couper, M.P., Dever, J.A., Gile, K.J. y Tourangeau, R. (2013). Summary Report of the AAPOR Task Force on Non-Probability Sampling. *Journal of Survey Statistics and Methodology*, 1(2), pp. 90–105.
- Biemer, P. P., de Leeuw, E., Eckman, S. Edwards, B., Kreuter, F., Lyberg, L. E., Tucker, N. C. y West, B. T. (eds.) (2017). *Total Survey Error in Practice*. Hoboken, NJ.: John Wiley.
- Bosco, A. y Verney, S. (2012). Electoral Epidemic: The Political Cost of Economic Crisis in Southern Europe, 2010–11. *South European Society and Politics*, 17(2), 129–54.
- Bosco, A. y Verney, S. (2016). From Electoral Epidemic to Government Epidemic: The Next Level of the Crisis in Southern Europe. *South European Society and Politics*, 21(4), 383–406.
- Brick J. M., Cervantes I. F., Lee S. y Norman G. (2011), Nonsampling Errors in Dual Frame Telephone Surveys. *Survey Methodology*, 37(1), pp. 1-12.
- Bustikova, L. y Zechmeister, E. J. (2017). Voting in New(er) Democracies. En Arzheimer, K., Evans, J. y Lewis-Beck, M. S. (eds.) (pp. 92-133). *The SAGE Handbook of Electoral Behaviour*. Thousand Oaks: SAGE.
- Callegaro, M., Baker, R., Bethlehem, J. Goritz, A., Krosnick, J. A. y Lavrakas, P.J. (eds.) (2014). *Online Panel Research. A Data Quality Perspective*. Hoboken, NJ: John Wiley.
- Campbell, W. J. (2020). *Lost in a Gallup. Polling Failure in US Presidential Elections*. Oakland, CA: University of California Press.
- Campello, C. y Zucco, C. (2016). Presidential Success and the World Economy. *Journal of Politics*, 78(2), pp. 590-602.
- Carlin, R. E. y Hellwig, T. (2019). Policy Regimes and Economic Accountability in Latin America. *Comparative Political Studies*, 52(13-14), pp. 2032-2060.

Cassino, D. (2016). Why Pollsters Were Completely and Utterly Wrong. *Harvard Business Review*, November 9. <https://hbr.org/2016/11/why-pollsters-were-completely-and-utterly-wrong>

Clinton, J. D., Agiesta, J., Brennan, M., Burge, C., Connelly, M., Edwards-Levy, A., Fraga, B., Guskin, E., Hillygus, D. S., Jackson, C., Jones, J., Keeter, S., Khanna, K., Lapinski, J., Saad, L., Shaw, D., Smith, A., Wilson, D. y Wlezien, C. (2021). *Task Force on 2020 Pre-Election Polling: An Evaluation of the 2020 General Election Polls*. American Association for Public Opinion Research. [https://www.aapor.org/AAPOR\\_Main/media/MainSiteFiles/AAPOR-Task-Force-on-2020-Pre-Election-Polling\\_Report-FNL.pdf](https://www.aapor.org/AAPOR_Main/media/MainSiteFiles/AAPOR-Task-Force-on-2020-Pre-Election-Polling_Report-FNL.pdf)

Clinton, J. D. (2021). *Why did Pollsters Blow the 2020 Election?* Entrevista de Michael Smerconish, May 17. <https://m.youtube.com/watch?v=2HGy9D0LiXI>

Cochran, W. G. (1977). *Sampling Techniques*. New York: John Wiley.

Cowley, P. y Kavanagh, D. (2018). *The British General Election of 2017*. Cham: Palgrave Macmillan.

Curtice, J., Phillips, M. y Clery E. (2015). *British Social Attitudes: the 33rd Report*. London: NatCen Social Research. Recuperado de <https://www.bsa.natcen.ac.uk>.

Curtice, J. (2016a). EU referendum – how the polls got it wrong again. *The Conversation*, June 26. <http://theconversation.com/eu-referendum-how-the-polls-got-it-wrong-again-61639>

Curtice, J. (2016b). *Polling in the EU Referendum: an overview*. Presentación en el Seminario Opinion Polling in the EU Referendum: Challenges and Lessons. London: Royal Statistical Society. December 8. <https://www.ncrm.ac.uk/polling/eu.php>

Curtice, J. (2016c). *The Benefits of Random Sampling. Lessons From the 2015 UK General Election*. London: NatCen Social Research. Recuperado de <https://www.bsa.natcen.ac.uk>.

Curtice, J. (2017). Why 2017 UK election polls underestimated Labour. *The Conversation*, June 15. <https://theconversation.com/amp/why-2017-uk-election-polls-underestimated-labour-79513>

DeBell, M., Amsbary, M., Meldener, V., Brock, S. y Maisel, N. (2018). *Methodology Report for the ANES 2016 Time Series Study*. Palo Alto, CA, and Ann Arbor, MI: Stanford University and the University of Michigan.

Desilver, D. (2021). Turnout soared in 2020 as nearly two-thirds of eligible U.S. voters cast ballots for president. *Pew Research Center*, January 2021. <https://www.pewresearch.org/fact-tank/2021/01/28/turnout-soared-in-2020-as-nearly-two-thirds-of-eligible-u-s-voters-castballots-for-president/>

Dillman, D.A., Smyth, J.D. y Christian, L.M. (2014). *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method* (4th edition). Hoboken, NJ: John Wiley.

Dutwin D., Buskirk T. D. (2017). Apples to Oranges or Gala versus Golden Delicious? Comparing Data Quality of Nonprobability Internet Samples to Low Response Rate Probability Samples. *Public Opinion Quarterly*, 81(S1), pp. 213-239.

Enten, H. (2017a). The French Election Is Way Too Close To Call. *FiveThirtyEight*, April 21. <https://fivethirtyeight.com/features/the-french-election-is-way-too-close-to-call/>

Enten, H. (2017b). Macron Won, But The French Polls Were Way Off. *FiveThirtyEight*. May 8. <https://fivethirtyeight.com/features/macron-won-but-the-french-polls-were-way-off/>

Fieldhouse, E. y Prosser, C. (2018). Horses for Courses: Using Internet Surveys for Researching Public Opinion and Voting Behavior. En Fisher, J., Fieldhouse, E., Franklin, M. N., Gibson, R., Cantijoch, M. y Wlezien, C. (eds) (pp. 472-483). *The Routledge Handbook of Elections, Voting Behavior and Public Opinion*. London: Routledge.

Fielding, N. G., Lee, R. M. y Blank, G. (eds.) (2017). *The SAGE Handbook of Online Research Methods* (2nd edition). Thousand Oaks: SAGE.

Fornoni, M. (2019). Análisis. Las proyecciones electorales para las PASO: una gran encuesta nacional rodeada de interrogantes. *Clarín*, 1 de agosto. [https://www.clarin.com/politica/proyecciones-electorales-paso-gran-encuesta-nacional-rodeada-interrogantes\\_0\\_fQR5COXfA.html](https://www.clarin.com/politica/proyecciones-electorales-paso-gran-encuesta-nacional-rodeada-interrogantes_0_fQR5COXfA.html)

Gervasoni, G. y Tagina, M. L. (2019). Explaining Support for the Incumbent in Presidential Elections. En Lupu, N., Oliveros, V. y Schiumerini, L. (eds.) (pp. 114-135). *Campaigns and Voters in Developing Democracies: Argentina in Comparative Perspective*. Ann Arbor: University of Michigan Press.

Goodwin, P. (2020). Without learning to think statistically, we'll never know when people are bending the truth. *The Guardian*, October 31. <https://theguardian.com/commentisfree/2020/oct/31/think-statistically-truth-falsehood>

Groves, R.M., Singer, E. y Corning, A. (2000). Leverage-Saliency Theory of Survey Participation: Description and an Illustration. *Public Opinion Quarterly*, 64(3), pp. 299–308.

Groves, R.M., Presser, S. y Dipko, S. (2004). The Role of Topic Interest in Survey Participation Decisions. *Public Opinion Quarterly*, 68(1), pp. 2–31.

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology* (2nd edition). Hoboken, NJ: John Wiley.

Groves, R. y Lyberg, L. (2010). Total survey error: Past, present and future. *Public Opinion Quarterly*, 74(5), pp. 849–879.

Henley, J. (2017). Pollsters breathe sigh of relief after calling French election right. *The Guardian*, April 24. <https://www.theguardian.com/world/2017/apr/24/french-pollsters-relief-after-calling-election-right>

Jennings, W. y Wlezien, C. (2018). Election polling errors across time and space. *Nature Human Behavior*, 2(4), pp. 276-283.

Jorge, J.E. (2015). La Cultura Política Argentina: una Radiografía. *Question*, 1(48), pp. 372-403.

Jorge, J. E. (2018). Ni “grieta” ni “degradación moral”: un contraste empírico del relato político en Argentina. *Question*, 1(59), 1-33. <https://doi.org/10.24215/16696581e074>

Jorge, J. E. (2020). Las causas de las protestas pacíficas en América Latina. *Question*, 2(67), 1-55. <https://doi.org/10.24215/16696581e425>

Jowell, R., Hedges, B., Lynn, P., Farrant, G. y Heath, A. (1993). The Polls - A Review: The 1992 British Election: The Failure of the Polls. *Public Opinion Quarterly*, 57(2), pp. 238–63.

Jowell, R., Roberts, C., Fitzgerald, R. y Eva, G. (2007). *Measuring Attitudes Cross-Nationally Lessons from the European Social Survey*. Thousand Oaks: SAGE.

Keeter, S., Hatley, N., Kennedy, C. y Lau, A. (2017). What Low Response Rates Mean for Telephone Surveys. *Pew Research Center*, May 15. <https://www.pewresearch.org/methods/2017/05/15/what-low-response-rates-mean-for-telephone-surveys/>

Keeter, S., Kennedy, C. y Deane, C. (2020). Understanding how 2020 election polls performed and what it might mean for other kinds of survey work. *Pew Research Center*, November 13. <https://www.pewresearch.org/fact-tank/2020/11/13/understanding-how-2020s-election-polls-performed-and-what-it-might-mean-for-other-kinds-of-survey-work/>

Keeter, S., Hatley, N., Lau, A. y Kennedy, C. (2021). What 2020’s Election Poll Errors Tell Us About the Accuracy of Issue Polling. *Pew Research Center*, March 2. <https://www.pewresearch.org/methods/2021/03/02/what-2020s-election-poll-errors-tell-us-about-the-accuracy-of-issue-polling/>

Keeter, S. y Clinton, J. D. (2021). Q&A: A conversation about U.S. election polling problems in 2020. *Pew Research Center*, July 21. <https://www.pewresearch.org/fact-tank/2021/07/21/a-conversation-about-u-s-election-polling-problems-in-2020/>

Kennedy, C., Blumenthal, M., Clement, S., Clinton, J., Durand, C., Franklin, C., McGeeney, K., Miringoff, L., Olson, K., Rivers, D., Saad, L., Witt, E. y Wlezien, C. (2017). *An Evaluation of 2016 Election Polls in the US*. American Association for Public Opinion Research. <https://www.aapor.org/Education-Resources/Reports/An-Evaluation-of-2016-Election-Polls-in-the-U-S.aspx>

Kennedy, C., Blumenthal, M., Clement, S., Clinton, J., Durand, C., Franklin, C., McGeeney, K., Miringoff, L., Olson, K., Rivers, D., Saad, L., Witt, E. y Wlezien, C. (2018). An Evaluation of 2016 Election Polls in the US. *Public Opinion Quarterly*, 82(1), pp. 1-33.

Kennedy, C., Hatley, N., Lau, A., Mercer, A., Keeter, S., Ferno, J. y Asare-Marfo, D. (2020). Assessing the Risks to Online Polls From Bogus Respondents. *Pew Research Center*, February 18. <https://www.pewresearch.org/methods/2020/02/18/assessing-the-risks-to-online-polls-from-bogus-respondents/>

Kennedy, C. (2020). Key things to know about election polling in the United States. *Pew Research Center*, August 5. <https://www.pewresearch.org/fact-tank/2020/08/05/key-things-to-know-about-election-polling-in-the-united-states/>

Lavrakas, P.J., Shuttles, C.D., Steeh, C. y Fienberg, H. (2007). The State of Surveying Cell Phone Numbers in the United States. *Public Opinion Quarterly*, 71(5), pp. 840–854.

Lavrakas, P. J., et al. (2017). *Report From the AAPOR Task Force on The Future of U.S. General Population Telephone Survey Research*. American Association for Public Opinion Research. <https://www.aapor.org/getattachment/Education-Resources/Reports/Future-of-Telephone-Survey-Research-Report.pdf.aspx>

Lax, J.R. y Phillips, J.H. (2009). How Should We Estimate Public Opinion in the States? *American Journal of Political Science*, 53(1), pp. 107–121.

Leeper, T.J. (2019). Where Have the Respondents Gone? Perhaps We Ate Them All. *Public Opinion Quarterly*, 83(S1), pp. 280-288.

Leonhardt, D. (2020). 'A Black Eye': Why Political Polling Missed the Mark. Again. *The New York Times*, November 12. <https://www.nytimes.com/2020/11/12/us/politics/election-polls-trump-biden.html>

Lepkowski, J. M. et al. (eds.) (2008). *Advances in Telephone Survey Methodology*. Hoboken, NJ: John Wiley.

Lewis-Beck, M. S. (2005). Election Forecasting: Principles and Practice. *British Journal of Politics and International Relations*, 7(2), pp. 145-164.

Lewis-Beck, M. S. y Stegmaier, M. (2013). The VP-function revisited: a survey of the literature on vote and popularity functions after over 40 years. *Public Choice*, 157(3-4), pp. 367-385.

Lewis-Beck, M. S. y Ratto, M. C. (2013). Economic Voting in Latin America: A General Model. *Electoral Studies*, 32(3), pp. 489-493.

Lewis-Beck, M. S. y Costa Lobo, M. (2017). The Economic Vote: Ordinary vs. Extraordinary Times. En Arzheimer, K., Evans, J. y Lewis-Beck, M. S. (eds.) (pp. 606-29). *The SAGE Handbook of Electoral Behaviour*. Thousand Oaks: SAGE.

Mellon, J. y Prosser, C. (2016). Correcting Polling Error in Over-Engaged Samples Using Representative Turnout Weighting, *Social Science Research Network*. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2720206](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2720206).

Mellon, J. y Prosser, C. (2017). Missing Nonvoters and Misweighted Samples: Explaining the 2015 Great British Polling Miss. *Public Opinion Quarterly*, 81(3), pp. 661-687.

Mercer, A. W., Kreuter, F., Keeter, S. y Stuart, E. A. (2017). Theory and practice in nonprobability surveys. Parallels between causal inference and survey inference. *Public Opinion Quarterly*, 81(Special Issue), pp. 250-279.

Merkle, D. (2021). *AAPOR 2021 Presidential Address*. AAPOR Headquarters, May 14. <https://m.youtube.com/watch?v=OM7AIsJ87qq>

Morales Solá, J. (2019). Una Argentina complicada e imprevisible. *La Nación*, 12 de agosto. <https://www.lanacion.com.ar/politica/una-argentina-complicada-e-imprevisible-nid2276604>

Nadeau, R., Ratto, M. C., Lewis-Beck, M. S., Bélanger, E., Gélinau, F. y Turgeon, M. (2015a). Rendición de cuentas en las democracias en desarrollo: El votante latinoamericano. *Revista de Ciencia Política*, 35(3), pp. 463-488.

Nadeau, R., Ratto, M. C., Lewis-Beck, M. S., Bélanger, E., Gélinau, F. y Turgeon, M. (2015b). Economía y elecciones en Argentina: las dimensiones clásica, posicional y patrimonial de la teoría del voto económico. *Revista SAAP*, 9(2), pp. 235-266.

Olson, K. et al. (2019). *Report of the AAPOR Task Force on Transitions from Telephone Surveys to Self-Administered and Mixed-Mode Surveys*. American Association for Public

Opinion Research. [https://www.aapor.org/getattachment/Education-Resources/ Reports/Report-of-the-Task-Force-on-Transitions-from-Telephone-Surveys-FULL-REPORT-FINAL.pdf.aspx](https://www.aapor.org/getattachment/Education-Resources/Reports/Report-of-the-Task-Force-on-Transitions-from-Telephone-Surveys-FULL-REPORT-FINAL.pdf.aspx)

Pagni, C. (2019a). Extravagancias y anomalías de las PASO. *La Nación*, 8 de agosto. <https://www.lanacion.com.ar/politica/extravagancias-y-anomalias-de-las-paso-nid2275334>

Pagni, C. (2019b). El desafío de garantizar la gobernabilidad. *La Nación*, 12 de agosto. <https://www.lanacion.com.ar/politica/el-desafio-de-garantizar-la-gobernabilidad-nid2276649>

Prosser, C. y Mellon, J. (2018). The Twilight of the Polls? A Review of Trends in Polling Accuracy and the Causes of Polling Misses. *Government and Opposition*, 53(4), pp. 757–790.

Ratto, M. C. (2013). ¿Rechazo a la democracia o al partido de gobierno? Despejando el camino para el estudio de la accountability electoral o algunas reflexiones sobre la relación entre crisis económica y voto en los últimos 30 años. *Revista SAAP*, 7(2), pp. 265-277.

Ritacco, E. (1987). Después del 9 de septiembre. ¿Sirven o no las encuestas? *Mercado Publicitario*, Año II, N° 15, 10 de octubre, pp. 6-15.

Rothschild, D. y Goel, S. (2016). When You Hear the Margin of Error Is Plus or Minus 3 Percent, Think 7 Instead. *The New York Times*, October 5. <https://www.nytimes.com/2016/10/06/upshot/when-you-hear-the-margin-of-error-is-plus-or-minus-3-percent-think-7-instead.html>

Rovelli, C. (2020). Statistical illiteracy isn't a niche problem. During a pandemic, it can be fatal. *The Guardian*, October 26. <https://theguardian.com/commentisfree/2020/oct/26/statistical-illiteracy-pandemic-numbers-interpret>

Schultheis, E. (2017). What Went Right With the French Campaign Polls? *The Atlantic*, May 13. <https://www.theatlantic.com/international/archive/2017/05/what-went-right-with-the-french-election-polls/526326/>

Shafer, J. y Doherty, T. (2017). The Media Bubble is Worse Than You Think. *Politico Magazine*, May/June. <https://www.politico.com/magazine/story/2017/04/25/media-bubble-real-journalism-jobs-east-coast-215048>

Shirani-Mehr, H., Rothschild, D., Goela, S. y Gelman, A. (2018). Disentangling Bias and Variance in Election Polls. *Journal of the American Statistical Association*, 113(522), pp. 607-614.

Silver, N. (2017a). There Really Was A Liberal Media Bubble. *FiveThirtyEight*, March 10. <https://fivethirtyeight.com/features/there-really-was-a-liberal-media-bubble/>



- Silver, N. (2017b). The Media Has A Probability Problem. *FiveThirtyEight*, September 21. <https://fivethirtyeight.com/features/the-media-has-a-probability-problem/>
- Silver, N. (2020). The Polls Weren't Great. But That's Pretty Normal. *FiveThirtyEight*, November 7. <https://fivethirtyeight.com/features/the-polls-werent-great-but-thats-pretty-normal/>
- Sturgis, P., Baker, N., Callegaro, M., Fisher, S., Green, J., Jennings, W., Kuha, J., Lauderdale, B. y Smith, P. (2016). *Report of the Inquiry into the 2015 British General Election Opinion Polls*. London: Market Research Society and British Polling Council.
- Sturgis, P. y Jennings, W. (2017). Will Turnout Weighting Prove to Be the Pollsters' Achilles Heel in #GE2017? *Politics Upside Down*, June 4. <https://sotonpolitics.org/2017/06/04/will-turnout-weighting-prove-to-be-the-pollsters-achilles-heel-in-ge2017/>
- Sturgis, P. (2019). After the failures of recent years, can the polls be trusted in 2019? *The Guardian*, November 12. <https://www.theguardian.com/commentisfree/2019/nov/12/polls-2019-election>
- Tagina, M. L. (2012). Controlando al gobierno a través de las urnas. Un análisis del caso argentino entre 1995 y 2005. *Política / Revista de Ciencia Política*, 50(1), pp. 111-144.
- Talving, L. (2017). The electoral consequences of austerity: economic policy voting in Europe in times of crisis. *West European Politics*, 40(3), pp. 560–583.
- The Economist (2019): Argentina's crisis shows the limits of technocracy, August 22. <https://www.economist.com/the-americas/2019/08/22/argentinas-crisis-shows-the-limits-of-technocracy>
- Tourangeau, R., Conrad, F. G. y Couper, M. P. (2013). *The Science of Web Surveys*. New York: Oxford University Press.
- UK House of Lords (2018). *The Politics of Polling. Report of Session 2017-19. HL Paper 106*. Select Committee on Political Polling and Digital Media, April 17. <https://publications.parliament.uk/pa/ld201719/ldselect/ldppdm/106/10602.htm>
- Van der Kooy, E. (2019). El lapidario peso de la crisis económica. *Clarín*, 12 de agosto. [https://www.clarin.com/opinion/lapidario-peso-crisis-economica\\_0\\_q7QwjD7Xr.html](https://www.clarin.com/opinion/lapidario-peso-crisis-economica_0_q7QwjD7Xr.html)
- Villar, A. y Fitzgerald, R. (2017). Using mixed modes in survey research: Evidence from six experiments in the ESS. En Breen, M. J. (ed.) (pp. 259-293). *Values and Identities in Europe. Evidence from de European Social Survey*. New York: Routledge.

Zechmeister, E. J., y Zizumbo-Colunga, D. (2013). The Varying Political Toll of Corruption in Good versus Bad Economic Times. *Comparative Political Studies*, 46(10), pp. 1190–1218.

Zukin, C. (2015). What's the Matter with Polling? *The New York Times*, June 20.

<https://www.nytimes.com/2015/06/21/opinion/sunday/whats-the-matter-with-polling.html>

### Fuentes de datos periodísticas

Bloomberg, 17/8/2019. “Wiped Out Once Again, Argentina Bulls Sift Through the Ashes”.

<https://www.bloomberg.com/news/articles/2019-08-17/burned-again-argentina-s-one-time-bulls-sift-through-the-ashes>

Bloomberg, 16/10/2019. “Polls in Argentina Nowhere to Be Found Two Weeks Before Election”.

<https://www.bloomberg.com/news/articles/2019-10-16/polls-in-argentina-nowhere-to-be-found-two-weeks-before-election>

Clarín, 1/8/2019. “Una nueva encuesta le da sólo dos puntos de diferencia a Alberto Fernández sobre Mauricio Macri”.

[https://www.clarin.com/politica/nueva-encuesta-da-solo-puntos-diferencia-alberto-fernandez-mauricio-macri\\_0\\_wLYYNW9LA.html](https://www.clarin.com/politica/nueva-encuesta-da-solo-puntos-diferencia-alberto-fernandez-mauricio-macri_0_wLYYNW9LA.html)

Clarín, 2/8/2019. “La guerra de encuestas, a pleno: aparecieron 12 pronósticos para la presidencial en solo 48 horas”.

[https://www.clarin.com/politica/guerra-encuestas-pleno-aparecieron-12-pronosticos-presidencial-solo-48-horas\\_0\\_xab-JeIrE.html](https://www.clarin.com/politica/guerra-encuestas-pleno-aparecieron-12-pronosticos-presidencial-solo-48-horas_0_xab-JeIrE.html)

Clarín, 9/8/2019. “Mercados y elecciones. Última rueda antes de las PASO: la Bolsa subió 7,8% y el riesgo país cayó 4% a 860.”

[https://www.clarin.com/economia/ultima-rueda-paso-bolsa-sube-5-riesgo-pais-cae-874\\_0\\_fiYbQbTR6.html](https://www.clarin.com/economia/ultima-rueda-paso-bolsa-sube-5-riesgo-pais-cae-874_0_fiYbQbTR6.html)

Clarín, 11/8/2019. “PASO 2019: ningún encuestador previó una brecha tan grande para los K”.

[https://www.clarin.com/politica/paso-2019-encuestador-previo-brecha-grande-k\\_0\\_TPRwIBckD.html](https://www.clarin.com/politica/paso-2019-encuestador-previo-brecha-grande-k_0_TPRwIBckD.html)

Clarín, 28/8/2019. “Tres conocidos encuestadores y una pregunta clave: ¿por qué erraron tanto sus pronósticos?”

[https://www.clarin.com/politica/conocidos-encuestadores-pregunta-clave-erraron-pronosticos\\_0\\_5dxvVpQdn.html](https://www.clarin.com/politica/conocidos-encuestadores-pregunta-clave-erraron-pronosticos_0_5dxvVpQdn.html)

Clarín, 27/9/2019. “Encuestadoras bajo fuego: por qué erraron en las PASO y qué dicen para octubre”.

[https://www.clarin.com/politica/encuestadoras-fuego-erraron-paso-dicen-octubre\\_0\\_T72H9hdl.html](https://www.clarin.com/politica/encuestadoras-fuego-erraron-paso-dicen-octubre_0_T72H9hdl.html)

Clarín, 9/10/2019. “Nuevas encuestas presidenciales: ¿hubo impacto del ‘Sí, se puede’?”

[https://www.clarin.com/politica/nuevas-encuestas-presidenciales-impacto-puede\\_0\\_S2mQRTlj.html](https://www.clarin.com/politica/nuevas-encuestas-presidenciales-impacto-puede_0_S2mQRTlj.html)

Clarín, 13/10/2019. “Nueva encuesta presidencial de uno de los consultores que mejor pronosticó las PASO”.

[https://www.clarin.com/politica/nueva-encuesta-presidencial-consultores-mejor-pronostico-paso\\_0\\_NEqC6L\\_D.html](https://www.clarin.com/politica/nueva-encuesta-presidencial-consultores-mejor-pronostico-paso_0_NEqC6L_D.html)

Clarín, 16/10/2019. "Dos nuevas encuestas presidenciales: cómo está la pelea entre Alberto Fernández y Mauricio Macri". [https://www.clarin.com/politica/nuevas-encuestas-presidenciales-pelea-alberto-fernandez-mauricio-macri\\_0\\_h3ohlzVE.html](https://www.clarin.com/politica/nuevas-encuestas-presidenciales-pelea-alberto-fernandez-mauricio-macri_0_h3ohlzVE.html)

Clarín, 18/10/2019. "La guerra de encuestas, a pleno: aparecieron 7 pronósticos para la presidencial al filo de la veda". [https://www.clarin.com/politica/guerra-encuestas-pleno-aparecieron-7-pronosticos-presidencial-filo-veda\\_0\\_vlhbUJcu.html](https://www.clarin.com/politica/guerra-encuestas-pleno-aparecieron-7-pronosticos-presidencial-filo-veda_0_vlhbUJcu.html)

Clarín, 27/10/2019. "Lo que dejó la presidencial. Elecciones 2019: cómo les fue esta vez a las encuestadoras". [https://www.clarin.com/politica/encuestas-elecciones-2019-solo-consultora-brasilena-acercor-resultado\\_0\\_uuM8n-jY.html](https://www.clarin.com/politica/encuestas-elecciones-2019-solo-consultora-brasilena-acercor-resultado_0_uuM8n-jY.html)

Financial Times, 2/9/2019. "Argentina: how IMF's biggest ever bailout crumbled under Macri". <https://www.ft.com/content/5cfe7c34-ca48-11e9-a1f4-3669401ba76f>

La Nación, 12/8/2019. "Las razones de los encuestadores tras haber fallado en los pronósticos de las PASO". <https://www.lanacion.com.ar/politica/las-razones-encuestadores-haber-fallado-pronosticos-paso-nid2276699>

La Nación, 23/8/2019. "Los encuestadores evalúan cambios metodológicos tras el papelón de las PASO". <https://www.lanacion.com.ar/politica/los-encuestadores-evaluan-cambios-superar-papelon-paso-nid2280291>

La Nación, 5/10/2019. "Las consultoras cambian la metodología para encuestar". <https://www.lanacion.com.ar/politica/las-consultoras-cambian-la-metodologia-para-encuestar-nid2294314>

Página 12, 2/8/2019. "Encuestas PASO 2019: esto arrojan los sondeos". <https://www.pagina12.com.ar/210003-encuestas-paso-2019-esto-arrojan-los-sondeos>

Página 12, 19/10/2019. "Casi 21 puntos de ventaja, llegando a 53". <https://www.pagina12.com.ar/226171-casi-21-puntos-de-ventaja-llegando-a-53>

Reuters, 9/9/2019. "Argentina pollsters emerge after primary fail, forecast huge win for Fernandez". <https://www.reuters.com/article/us-argentina-election-polls-idUSKCN1VU212>