

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias Sociales y Humanidades

Los fallos de encuestas electorales en América Latina desde el año 2000 y las condiciones en las que las encuestadoras pueden pronosticar el resultado de una elección. Los casos de estudio de Argentina, Brasil y Ecuador

Felipe David Grijalva Espín

Relaciones Internacionales

Trabajo de fin de carrera presentado como requisito para la obtención del título de Licenciado en Relaciones Internacionales

Quito, 15 de mayo de 2024

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias Sociales y Humanidades

HOJA DE CALIFICACIÓN DE TRABAJO DE FIN DE CARRERA

Los fallos de encuestas electorales en América Latina desde el año 2000 y las condiciones en las que las encuestadoras pueden pronosticar el resultado de una elección. Los casos de estudio de Argentina, Brasil y Ecuador

Felipe David Grijalva Espín

Nombre del profesor, Título académico

Sofia Zaragocín Carvajal, PhD.

Quito, 15 de mayo de 2024

© DERECHOS DE AUTOR

Por medio del presente documento certifico que he leído todas las Políticas y Manuales de la Universidad San Francisco de Quito USFQ, incluyendo la Política de Propiedad Intelectual USFQ, y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo quedan sujetos a lo dispuesto en esas Políticas.

Asimismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo en el repositorio virtual, de conformidad a lo dispuesto en la Ley Orgánica de Educación Superior del Ecuador.

Nombres y apellidos: Felipe David Grijalva Espín

Código: 00212800

Cédula de identidad: 1725862138

Lugar y fecha: Quito, 15 de mayo de 2024

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en <http://bit.ly/COPETHeses>.

UNPUBLISHED DOCUMENT

Note: The following capstone project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on <http://bit.ly/COPETHeses>.

AGRADECIMIENTOS

Agradezco a mis padres, Iván y Neyda, por todo el apoyo que me han dado durante toda mi estadía en la universidad. A mi hermano, Daniel, por ser una inspiración como persona y como profesional. A mi tío, Diego, por siempre estar conmigo y enseñarme tantas cosas dentro y fuera de la universidad. A mis abuelos, Jaime, Yolanda y Neida, por su cariño infinito. A mis profesores, Paolo, Raúl, Pedro, Daniela y Hugo por ser excelentes maestros y mentores. A mis amigos Daniel, Macarena, Michelle, Mateo, Nicolás, Noah y Dharma por su cariño y su conocimiento.

RESUMEN

Las encuestas de opinión sobre la intención de voto son herramientas muy utilizadas durante las jornadas electorales en los países democráticos. Desde la consolidación democrática en América Latina las encuestas han crecido en popularidad para tratar de pronosticar los resultados de las elecciones. Durante los últimos años ha surgido una narrativa dentro de los medios de comunicación, como en ciertos sectores de la academia que trata de desmeritar a las encuestas electorales, argumentando que fallan con mayor frecuencia que antes en representar adecuadamente la intención de voto previo a cada elección. Esta investigación tiene como objetivo identificar si las acusaciones de una mayor cantidad de fallos en contra de las encuestas son correctas, cuales variables pueden influenciar si una encuesta predice con mayor o menor exactitud el resultado final de la elección y en qué condiciones pueden predecir las encuestas más acertadamente los resultados electorales.

Palabras clave: encuesta, elección, desviación, candidatos, Argentina, Brasil, Ecuador, muestra, método de recolección

ABSTRACT

Opinion polls regarding voting intentions are immensely popular tools in democratic countries during the electoral period. Since the consolidation of democracy in Latin America, polls have grown in popularity during campaigns to try to predict election results. A narrative has emerged during the last few within the media, as well as in certain sectors of academia, which tries to discredit electoral polls, arguing that they fail more frequently than before to adequately represent the voting intention prior to each election. This research aims to identify whether the accusations of a greater number of failures within polls are correct, which variables could influence whether a poll accurately predicts the outcome of the election, and under what conditions can polls predict election results more accurately.

Key words: survey, election, deviation, candidates, Argentina, Brazil, Ecuador, sample, collection method

TABLA DE CONTENIDOS

Introducción.....	11
Contexto	12
Justificación	18
Preguntas y objetivos de investigación	21
Revisión de Literatura	22
Figura 1, Elaborado por Felipe Grijalva, 2024	¡Error! Marcador no definido.
Posicionalidad	31
Reflexibilidad	32
Discusión	33
Conclusiones	50
Bibliografía	52
Anexos	55

ÍNDICE DE TABLAS

Tabla 1	¡Error! Marcador no definido.
Tabla 2	¡Error! Marcador no definido.
Tabla 3	¡Error! Marcador no definido.
Tabla 4	¡Error! Marcador no definido.

ÍNDICE DE FIGURAS

Figura 1	25
Figura 2	41
Figura 3	42
Figura 5	44
Figura 6	45
Figura 7	46
Figura 8	47
Figura 9	48
Figura 10	49

Introducción

Las encuestas de opinión sobre la intención de voto son herramientas muy utilizadas durante las jornadas electorales en los países democráticos. Los conocidos como encuestas de seguimiento (tracking polls), es decir las encuestas que se realizan durante un periodo de tiempo específico antes de las elecciones, buscan pronosticar los resultados dentro de un margen de error determinado (Lewis-Beck, 2005). Desde que se popularizó el uso de encuestas electorales durante épocas de campaña existe un enorme interés por parte de periodistas, académicos, políticos y el público general por analizar la intención de voto y tratar de predecir el resultado que se obtendrá en las urnas. Durante varios años ha existido una narrativa, tanto dentro de los medios de comunicación como en ciertos sectores de la academia que trata de desmeritar las encuestas, argumentando que fallan con mayor frecuencia que antes, a la vez que han surgido múltiples hipótesis en base a esta afirmación (Prosser, 2018). El análisis de encuestas en América Latina es más limitada que en otras regiones del mundo, tanto por desarrollos históricos de la democracia en la región como por las propias limitaciones de la investigación en los países. Este estudio analiza el nivel de fallo de las encuestas electorales en Argentina, Brasil y Ecuador, tres países en donde las encuestas han sido especialmente controversiales en los ciclos electorales más recientes.

El propósito de esta investigación es para comprobar o refutar hipótesis preexistentes a cerca de los fallos de encuestas electorales. Se ha escogido a estos países de la región debido a que hay una brecha en la literatura de encuestas. Para este estudio se creó una base de datos con 375 encuestas durante múltiples periodos electorales y se realizó una serie de regresiones múltiples para determinar cuáles variables independientes tienen un impacto dentro de la precisión de la encuesta. Se mide la precisión de la encuesta como la suma de la desviación ponderada del porcentaje obtenido por cada candidato en comparación al resultado real obtenido en la elección. Para mejorar la validez del estudio se miden como variables

dependientes a la desviación de votos válidos y votos totales. Solo se está midiendo la precisión de las encuestas para las elecciones presidenciales; no se toman en cuenta elecciones legislativas, elecciones municipales, elecciones estatales, etc. La similitud del sistema electoral y la forma del conteo de votos entre los tres países permite comparar los resultados obtenidos.

Los resultados muestran que la precisión de las encuestas no ha empeorado durante los últimos años y que el impacto de ciertas variables específicas es menor al esperado. La variable que tiene un mayor impacto es la ronda electoral para la que se realizó la encuesta, seguido por la cantidad de candidatos que participaron en la elección y el país para el que se realizó la encuesta.

Las conclusiones destacan cuales variables deberían ser analizadas con mayor profundidad para un próximo estudio, como el método de recolección de datos. Así mismo se puede notar como los resultados refuerzan algunos aspectos de la literatura sobre encuestas electorales mientras que contradice otros.

Contexto

Las encuestas no son herramientas nuevas ni necesariamente complejas; se pueden definir como un procedimiento de recopilación de información de forma sistemática, mediante las respuestas a un cuestionario, sin modificar el entorno o el fenómeno donde se recoge la información. Los datos consisten en una o más variables observables (por ejemplo, edad u opinión) que se obtienen de una muestra representativa de la población a la cual se quiere investigar. Una encuesta no es solo una técnica de colección de información; los cuestionarios son la forma más común de llenar encuestas, pero también se pueden utilizar entrevistas, observaciones análisis de contenido, etc.... Lo que distingue a las encuestas de otras formas de investigación social es la forma de la estructura de los datos y el método de análisis (Vaus, 2013).

Las encuestas en su estado más simple existen desde que se tiene registros de la existencia de las primeras civilizaciones humanas, aunque no se puede determinar el punto de partida en el que surgen; lo que sí se sabe que desde hace milenios ya se recolectaba información mediante cuestionarios para entender la realidad de una situación y ejecutar políticas públicas en diferentes estados. Los egipcios enviaban con frecuencia a lo que actualmente podemos considerar encuestadores para recolectar información sobre el territorio que gobernaban, y los datos recogidos se usaban en conjunto con otras ciencias como astronomía, arquitectura y matemáticas para diseñar sistemas de irrigación, edificios públicos, templos, etc. Los romanos contribuyeron enormemente en el avance del diseño y uso de encuestas debido a que las utilizaban a lo largo y ancho de su imperio para censar a su población, llevar registros financieros y decidir dónde invertir los recursos del estado en base a esa información (Rossi, Wright & Anderson, 1983). Así mismo, la historia de las encuestas electorales tampoco es muy clara, la pérdida de registros históricos a lo largo del tiempo dificulta la capacidad de poner una fecha exacta a cuando se comenzaron a utilizar encuestas para tratar de medir las preferencias de un padrón electoral y predecir el resultado de una elección. Sin embargo, existen ciertos puntos de referencia que los académicos citan para marcar el punto de inicio. Algunos mencionan a las elecciones presidenciales de Estados Unidos en 1936, cuando Gallup, Roper y Crossley realizaron encuestas de preferencia entre los candidatos Franklin Delano Roosevelt y Alf Landon. Otros se marcan como referencia al año 1916, cuando la revista estadounidense *The Literary Digest* envió cuestionarios por todo el país y predijo la victoria del candidato presidencial Woodrow Wilson. Entre los académicos actuales el consenso general es que la primera encuesta electoral, o proto-encuesta como refieren algunos analistas, apareció en las elecciones presidenciales de Estados Unidos en 1824, donde se predijo la victoria de Andrew Jackson sobre John Quincy Adams en el periódico *Raleigh Star and North Carolina State Gazette* (Smith, 1990).

A partir de la Segunda Guerra mundial se generó un interés por las encuestas alrededor del mundo, tanto de parte de los gobiernos como de la academia. Debido a las olas de democratización el uso de encuestas se fue insertando dentro de los ciclos electorales y los partidos políticos empezaron a basarse en sus datos para sus diseñar sus estrategias de campaña y corregir el rumbo de esta en caso de ser necesario. John F Kennedy fue de los primeros candidatos presidenciales que utilizó encuestas para estudiar las opiniones del electorado y definir sus puntos fuertes y débiles con sus posibles votantes (Compilación, C. E. O, 2011). Pero no fue solo en Estados Unidos donde las encuestas llegaron a tener mucha importancia en el sistema democrático, el Partido Laborista británico utilizó encuestas para evaluar las necesidades más urgentes de la gente en cara a las elecciones de 1945, el líder del partido Clement Altee ajustó su estrategia de campaña en base a los resultados y fue en gran medida por esto que obtuvo una de las mayores victorias en la historia de la democracia británica (Fielding, 1992).

Fueron estos y otros ejemplos donde las encuestas probaron su valor, lo que llevó a que su uso se masifique y ganen un amplio prestigio y popularidad en los círculos políticos, académicos y civiles. A partir de los años 70 las encuestas electorales habían llegado a todas las regiones del mundo y se tomaban muy en serio para cada ciclo electoral. Sin embargo, las encuestas no son un método de análisis cuantitativo perfecto y encuestadoras no siempre han logrado a predecir a los resultados finales de una elección. Existen casos emblemáticos de fallos de las encuestas electorales, como en 1992, cuando Partido Conservador en el Reino Unido logró una inesperada victoria cuando las predicciones mostraban que los laboristas obtendrían una amplia mayoría. Durante la década de 2010 hubo un punto de inflexión en la reputación de las encuestas a nivel global debido a múltiples fallos en distintas elecciones alrededor del mundo y Latinoamérica no fue una excepción a esa tendencia.

Las encuestas electorales llegaron relativamente tarde a América Latina debido a que no se consolidó la democracia en la región hasta hace poco. De todas formas, desde hace cuatro décadas que se puede hablar de ciclo de encuestas acompañando las campañas electorales en todos los países democráticos de la región. Los tres países que se van a estudiar en esta tesis tienen varias características particulares que hacen que su relación con las encuestas sea bastante compleja y distinta a los otros, pese a que existan ciertos rasgos similares. Para comprender como han sido históricamente la dinámica de las encuestas electorales en estos países hay que también comprender la historia y evolución de sus sistemas políticos.

Ecuador se ha caracterizado desde su independencia por su constante inestabilidad política a causa de su debilidad institucional y a la cultura política de la nación. Durante el primer siglo de vida republicana se consolidaron como las principales fuerzas políticas las fuerzas armadas, la iglesia católica y las élites económicas latifundistas (Ayala Mora, 2008). El desarrollo del estado fue lento y sin una visión consistente en la etapa de formación; los intentos de presidentes como Gabriel García Moreno y Eloy Alfaro de crear un estado moderno, no existió un modelo unificado de política en el país. La democracia, si bien existía *de iure* dentro del régimen político, en la realidad era altamente restringida y solo podía acceder a ella un segmento reducido de la población; como consecuencia el joven país no era verdaderamente democrático y no fue hasta inicios del siglo XX que el sistema electoral se fue volviendo más accesible para la mayoría de la población. Sin embargo, las fuerzas políticas antiguas todavía mantenían gran influencia sobre el país, lo que los llevaba a intervenir en contra del sistema democrático; las fuerzas armadas organizaban en promedio un golpe de estado cada 6 años. El último golpe de estado en el que asumió el poder una dictadura militar fue en 1972, dirigida por el general Guillermo Rodríguez Lara cuando derrocó al quinto gobierno del presidente José María Velasco Ibarra. El gobierno militar fue menos represor que sus contemporáneos de la región y gobernó con políticas económicas estatistas, contrarias a la ortodoxia de la época; de

hecho, fue en esta época en la que se fundó el Centro de Estudios y Datos (CEDATOS), el primer centro de estudios de mercado y opinión pública en el Ecuador, que inició sus operaciones en 1974. La democracia ecuatoriana fue restaurada en 1979, con la llegada de Jaime Roldós a la presidencia tras ganar las elecciones generales, en las cuales se utilizaron encuestas electorales por primera vez de forma masiva para todo el país (Córdoba, 2019). La inestabilidad persistió en el sistema político del país tras el retorno a la democracia, aunque se llegó a un gran acuerdo nacional, en el que acordaba que las fuerzas armadas quedaron bajo control del gobierno civil de forma permanente y ya no habría más gobiernos militares. Esta consolidación democrática permitió que se creen cada vez más encuestadoras a nivel nacional y estas se metan de lleno en la dinámica de los ciclos electorales. Las encuestadoras solían gozar de un alto prestigio en el país, pero varios resultados en las últimas elecciones nacionales y seccionales han puesto en duda su credibilidad debido a que hay predicciones que están lejos del margen de error de los resultados finales. El paso del candidato Daniel Noboa a la segunda vuelta, junto con la derrota de la alcaldesa Cynthia Viteri no se encontraban dentro de los parámetros esperados por ninguna encuestadora oficial.

Brasil tiene una historia igualmente compleja, llena de inestabilidad política y golpes de estado que han dificultado la consolidación democrática en el gigante sudamericano. A pesar de esto, el país se fundó con una idea de crear una sociedad moderna a partir de un estado fuerte, cuya influencia se extiende por toda la sociedad, una visión heredada de Portugal (Telles, 2008). Como resultado de esto, una diferencia fundamental es que Brasil históricamente ha tenido una solidez institucional mayor a la del resto de países de la región, en la que muchas instituciones y costumbres se mantienen pese a los cambios de gobierno y de régimen. Brasil originalmente era un imperio, en el que la monarquía mandaba por encima de todos los demás poderes del estado, por lo que no existía la democracia hasta la creación de los Estados Unidos de Brasil (predecesor de la actual República Federativa de Brasil), que derrocó al emperador Pedro II,

apodado “el magnánimo”. La democracia se extendió por todo el país, al punto que la última dictadura que llegó al poder gracias al golpe de estado en 1964 no podía evitar que se celebraran elecciones a nivel local e incluso “elecciones” presidenciales. Entre 1984 y 1985 se da una apertura democrática a partir de la redacción de una nueva constitución y en 1989 regresa la democracia a Brasil al convocarse elecciones libres y directas por primera vez desde el golpe de estado. En las elecciones generales de 1989 se utilizaron encuestas electorales de forma masiva en por primera vez en la historia de Brasil por medio de *Datafohla*, la cual encuestadora del *Grupo Fohla*, segundo conglomerado de medios de comunicación más grande de Brasil; la encuestadora fue fundada en 1983. Como en el resto de los países de la región, Brasil vivió una fiebre de encuestas que se ha mantenido hasta sus más recientes elecciones generales en 2022, el año donde las encuestas han sido más cuestionadas en el país hasta la fecha.

Argentina nació tras el triunfo de la Revolución de mayo de 1810, cuando las Provincias Unidas del Río de la Plata (el estado predecesor de la actual República Argentina) ganaron la guerra de independencia y se declararon estado soberano; pero el dictamen de 1819, en el que se impone una constitución centralista generó un conflicto entre los partidarios del federalismo contra los partidarios de la centralización. A partir de esta pugna de poder la historia de la República Argentina se llenó de guerras civiles, inestabilidad política y golpes de estado militares. Como en el resto de los países de la región, las fuerzas armadas desempeñaron un papel político muy importante durante la mayoría de la historia del país debido a sus vínculos con las élites económicas y a los grupos políticos gobernantes (Romero, 2022). Durante el siglo XX las fuerzas armadas realizaron 6 golpes de estado en contra de los gobiernos civiles, aunque cabe resaltar que varios de los gobiernos civiles en aquella época también fueron autoritarios, lo que impidió la consolidación democrática en el país y la difusión de las encuestas electorales en la vida política. El último golpe de estado militar, que se llevó a cabo en 1976 bajo el contexto de la lucha comunismo y anticomunismo de la Guerra Fría, trajo al poder a una cúpula militar

auto dominada como “Proceso de Reorganización Nacional” y fue más represiva que los gobiernos militares de Ecuador y Brasil. Tras la salida del poder de los militares en 1983, provocada por la derrota en la Guerra de las Malvinas y a una enorme crisis económica, se llevó a cabo un proceso de consolidación democrática y ciclos electorales predecibles, pero a diferencia de otros países de la región no hubo encuestadoras creíbles en el primer ciclo electoral post dictadura. A pesar de esto, las encuestadoras se volvieron parte de la vida política en por otros medios y llegaron a las mismas tendencias que el resto del mundo; así como en el resto de América Latina, el auge de los estudios de opinión pública es paralelo a la restauración de la democracia a principios de los 80s (Adrogué & Adrogué, 1998). A partir de 1989 se regularizan las encuestas electorales en el país y esto se mantiene hasta el día de hoy. Siguiendo los mismos patrones que en otros países, las encuestadoras argentinas se han llevado varios shocks en algunas predicciones electorales durante los últimos años, especialmente en las elecciones generales de 2023.

Justificación

Las encuestas electorales son herramientas de recolección de datos muy poderosas que utilizan los partidos políticos, candidatos, empresas privadas y organizaciones sociales dentro de un régimen democrático (Kennedy, 2018), debido a que permiten evaluar las opiniones de la sociedad a cerca de uno o varios temas específicos y tomar decisiones en base a las preferencias que posee la ciudadanía en ese momento. Las encuestadoras miden la intención de voto de un partido o candidato de cara a unas elecciones para determinar su lugar en la preferencia del padrón electoral (Instituto de la Democracia, 2021). De esta manera, los candidatos pueden evaluar su nivel de popularidad en su campaña y modificar sus estrategias en caso de ser necesario, lo que les permite llegar con ciertos mensajes o estrategias de campaña a ciertos segmentos de la población y potencialmente revertir el resultado previsto de una elección.

Las encuestas despiertan un alto interés dentro de los círculos políticos, académicos y mediáticos debido a que muchas encuestadoras si aciertan a los resultados finales de una elección; es importante mencionar que la calidad de las encuestadoras varía entre si e históricamente no todas tenían un alto grado de precisión por diferentes que podían deberse a errores de muestreo, sesgos en las preguntas o errores de codificación. Sin embargo, tanto los políticos como la ciudadanía en general tomaban muy enserio a las encuestas y casi siempre esperaban resultados similares en días de elecciones. Esto ha cambiado durante los últimos años, a causa de que múltiples encuestadoras, incluyendo a varias de renombre, no han logrado predecir los resultados electorales en múltiples países, y en algunos casos, los resultados de los conteos de finales han estado lejos del margen de error de las predicciones electorales, lo que ha manchado el prestigio de todas las encuestadoras en general. Las consecuencias de estos errores no son marginales, ya que un resultado con un error lejos del margen de error puede quebrar a una encuestadora, o en el peor de los casos llegar a poner en duda la legitimidad de los resultados de una elección, aunque esto último es poco probable.

Las encuestas electorales continúan siendo herramientas muy comunes dentro de las comunidades académicas de relaciones internacionales ciencias políticas las pero la aparición nuevas dinámicas sociales entre diferentes segmentos de los electorados, junto con el uso de nuevas tecnologías durante campañas electorales, distorsionan la validez y la fiabilidad de los resultados de las encuestas. Se han realizado múltiples estudios académicos sobre fallos de encuestas electorales alrededor del mundo en los últimos años, sobre todo en elecciones que han sido consideradas de alta relevancia geopolítica como las presidenciales de Estados Unidos en 2016 y el referéndum sobre el Brexit en el Reino Unido (Kennedy,2018). En todos estos casos los sondeos presencian resultados diferentes a los que salieron finalmente de cada instituto nacional electoral y se han llevado a cabo varios análisis exhaustivos de que salió mal, con diferentes conclusiones para cada encuesta y cada país. No se trata de un tema nuevo, uno

de los casos más emblemáticos de este tipo de fenómenos data de 1992, cuando el partido conservador derroto de manera sorpresiva al partido laborista en las elecciones generales del Reino Unido, lo que iba en contra de las predicciones de la época. Sin embargo, no cabe duda de que actualmente este fenómeno ha vuelto más común, lo que de manera paradójica da mayor cobertura mediática a las encuestadoras durante épocas electorales.

El propósito de este estudio es verificar si las teorías ortodoxas que explican fallos de encuestas electorales en lugares como Europa, Estados Unidos y Asia aplican para los países de América Latina. La región, a partir de la tercera ola de democratización en los años 80s y 90s, es una de las más democráticas del mundo y el uso y circulación de sondeos son igual de comunes. Existe una cantidad no marginal de estudios sobre este tema en Latinoamérica, especialmente proviniendo de Brasil, pero sigue siendo bastante inferior a la cantidad de estudios que se realizan en Estados Unidos o España debido a que no se ha convertido en un tema tan grande dentro de los círculos académicos.

Este ensayo utiliza una investigación cuantitativa, los datos recolectados permiten entender a profundidad el fenómeno; aumenta la validez y fiabilidad de la investigación; da una mayor contextualización del fenómeno, lo cual es especialmente importante debido a que cada país tiene sistemas y leyes electorales diferentes, junto a una cultura democrática única.

Se escogió a Argentina, Brasil y Ecuador como casos específicos de estudio debido que los tres países tienen un sistema electoral similar, existe un extenso uso de encuestas en los procesos electorales y hubo una desviación significativa entre los sondeos y los resultados de sus más recientes elecciones. Entender como han evolucionado la capacidad de predicción de las encuestadoras y por qué han fallado ciertas ocasiones permite no solo ampliar la literatura sobre este tema dentro de América Latina, sino que permite entender mejor un fenómeno global e incluso permitir remediar el problema para todos aquellos que trabajen en encuestas electorales en un futuro. Al entender la causa de este fenómeno es posible desarrollar estrategias

de resolución que pueden requerir cambiar de metodologías, distintas formas de muestreo y ponderación, adaptarse a nuevas dinámicas sociales del electorado o en el peor de los casos cuestionar la utilidad de las encuestas electorales en su totalidad. Si las encuestadoras pueden recuperar su credibilidad en los ojos de la academia y del público en general esto potencialmente puede fortalecer la democracia y disminuir la desconfianza que se tiene a las instituciones de los estados democráticos.

Preguntas y objetivos de investigación

Pregunta central de investigación

¿En qué condiciones una encuestadora puede pronosticar más acertadamente el resultado de una elección general en Argentina, Ecuador y Brasil?

Preguntas auxiliares

1. ¿Cómo ha cambiado la dinámica de encuestas electorales en los últimos años?
2. ¿Cuáles variables afectan a la predicción de resultados durante las elecciones?
3. ¿Cómo pueden corregir las encuestadoras los errores cometidos los sondeos?

Objetivo general de investigación

Analizar las encuestas electorales para las elecciones generales en Argentina, Ecuador y Brasil desde el año 2000, identificando los momentos en los cuales las encuestas fallaron en predecir correctamente los resultados, para poder determinar las condiciones específicas en las que una encuesta puede predecir más acertadamente los resultados electorales.

Objetivos auxiliares

1. Analizar como los cambios tecnológicos, políticos y sociales de los últimos años han afectado la fiabilidad y validez de las encuestas
2. Analizar cuáles fueron las variables específicas que llevaron a las encuestas a fallar en ciertas predicciones electorales

3. Analizar cuáles pueden ser unas posibles soluciones para corregir los errores de predicción electoral

Revisión de Literatura

Cada jornada electoral es un periodo de tiempo donde hay un mayor interés por la política a nivel general dentro de la población; incluso puede llegar a generar interés en poblaciones extranjeras que no participan en las elecciones de un país, ya sea por mero interés o por el impacto que puede causar el resultado electoral afuera de la nación donde se llevan a cabo las elecciones. En varios países con democracias consolidadas existe una tradición de pronósticos electorales que lleva a que relacionistas internacionales, politólogos, antropólogos, economistas, periodistas y la población general traten de predecir los resultados de los comicios, aunque no siempre de manera acertada. La herramienta más usada para llevar a cabo estas predicciones son las encuestas de opinión, las cuales son diseñadas e interpretadas de diferentes maneras por diferentes grupos de expertos. El área de pronóstico electoral ha crecido mucho en las últimas décadas gracias a la democratización de varias regiones en el mundo, pero una región en la que existe una relativa falta de avances en el área de pronósticos electorales es América Latina. Sin embargo, aunque no haya tantos estudios hasta la fecha, existe una investigación emergente sobre las elecciones latinoamericanas (Lewis-Beck, 2022). Para comprender como se realizan las encuestas electorales en América Latina, y porque no siempre logran pronosticar adecuadamente el resultado de los comicios, existen múltiples ángulos posibles para un análisis. Primero se debe entender de manera general cómo funcionan las encuestas electorales. Segundo, que estudios existen sobre los fallos de encuestas electorales. Tercero, que estudios existen en América Latina sobre pronósticos electorales y que nos dicen sobre la situación actual de las encuestas electorales. Esta revisión de literatura se enfoca en estos tres elementos para su análisis y se basa los autores con mayor experiencia en cada área.

Como funcionan las encuestas electorales

Para comprender como funciona una encuesta electoral primero hay que definir con precisión que es una encuesta, que tipo de información recolectan y como se deben interpretar de forma estadística los datos. Una encuesta no es solamente una técnica específica para recolectar información. En la literatura algunos autores comienzan por describir los elementos que componen a una encuesta en lugar de dar una definición, mientras que otros tratan de definir que es una encuesta dando definiciones concisas hasta criterios más elaborados (de Leeuw, 2008). Basándonos estrictamente en la definición, una encuesta es un método sistemático de recolección de información (mediante una muestra) sobre entidades específicas, con el propósito de construir descriptores cuantitativos sobre los atributos de la población a la que pertenecen las entidades (Groves, 2004). Las encuestas se caracterizan por tener variables por caso sistematizadas que se representan en una cuadrícula de datos. Esto significa que colectamos información sobre las mismas variables o características a cerca dos o más casos para visualizar la cuadrícula de datos (de Vaus, 2014). Para recolectar la información se utiliza un cuestionario estandarizado para que todas las personas que forman parte de una muestra respondan. En la siguiente tabla se muestra cómo se construye una visualización de datos a partir de la recolección de información en una cuadrícula de datos.

		VARIABLES			
		<i>Sexo</i>	<i>Edad</i>	<i>Clase social</i>	<i>Ideología</i>
<i>Casos</i>	<i>Persona 1</i>	<i>Hombre</i>	<i>25 años</i>	<i>Media-alta</i>	<i>Derecha</i>
	<i>Persona 2</i>	<i>Mujer</i>	<i>27 años</i>	<i>Media</i>	<i>Centro</i>
	<i>Persona 3</i>	<i>Mujer</i>	<i>20 años</i>	<i>Media-baja</i>	<i>Izquierda</i>

Tabla 1, elaborado por Felipe Grijalva

Manejar encuestas es un trabajo complejo, ya que para que tengan información relevante y de valor científico el proceso de diseño de la encuesta y el posterior análisis debe manejarse

con mucha rigurosidad. Los cuatro pilares de la investigación mediante encuestas son cobertura, muestreo, respuesta y medición (de Leeuw, 2008).

1. Cobertura se refiere al alcance o extensión de la población o grupo de interés que se desea estudiar con la encuesta; es indispensable para determinar que la cantidad y las características de las personas encuestadas sea representativa de la población general que se quiere encuestar.
2. Muestro es el proceso de selección de una muestra representativa de una población objetivo para ser encuestada; la calidad de la información depende de si la técnica de muestreo (muestreo por conveniencia, muestreo aleatorio simple, muestreo estratificado, muestreo por conglomerados, etc) realmente alcanza a describir a la población objetivo.
3. Respuesta es tanto la cantidad de encuestador que resolvieron el cuestionario o la entrevista, como la calidad de las respuestas obtenidas. Es deseable que se dé una alta participación de respuestas válidas; y dependiendo de diferentes factores como el diseño de las preguntas (que sean claras y concisas), el nivel de compromiso de los participantes y la honestidad con la que estén respondiendo pueden cambiar bastante la calidad de la información.
4. Medición es el proceso de asignar valores numéricos o categóricos a las respuestas de los encuestados, normalmente dentro de una escala de valores (ya sea escalas de Likert, escalas de frecuencias, escalas de clasificación) de menor a mayor, o viceversa, dependiendo de la naturaleza de la pregunta y del criterio personal del encuestador. (de Leeuw, 2008)

Lo que busca una encuesta con la ayuda de estos cuatro pilares es medir conceptos o constructos, los cuales pueden ser simples o abstractos. Para ello se requiere en proceso de tres

pasos: aclarar el concepto, desarrollar indicadores y evaluar los indicadores. Esto se define como bajar en una escalera de abstracción y se visualiza de la siguiente forma.



Figura 1, Elaborado por Felipe Grijalva, 2024

Un concepto en este contexto es solo una herramienta que cuya función es resumir una serie de comportamientos, actitudes o características que tienen algo en común; por lo tanto, un concepto es lo que el investigador define como tal para dar consistencia a su trabajo, teniendo en cuenta que el uso de ciertas palabras no necesariamente significa lo mismo para todos, por lo que es necesario justificar porque se está usando tal definición de tal concepto para el diseño de la investigación (de Vaus, 2014). El desarrollo de indicadores es el proceso de mover estos conceptos abstractos a algo concreto que pueda ser observable. Un indicador se refiere a una variable o conjunto de variables que se utilizan para medir o evaluar un fenómeno, comportamiento, actitud, opinión, o cualquier otra característica de interés dentro de la población de estudio. Los indicadores más comunes en la investigación por encuestas son los indicadores de actitud, indicadores de comportamiento, indicadores sociodemográficos, indicadores de satisfacción e indicadores de percepción. Por último, evaluar los indicadores es el proceso en el que el investigador se asegura que estos tengan validez y fiabilidad. La fiabilidad se refiere a la extensión en la que la varianza de una variable x se deba a fuentes no aleatorias de la varianza total, es decir, la variable x (el indicador) siempre va a producir resultados consistentes en diferentes momentos, entornos y poblaciones (Rossi, 2013).

La fiabilidad de x , designada como ρ_x , es la relación entre la varianza de la puntuación real y la varianza observada.

$$\rho_x = \frac{\sigma_x^2}{\sigma_T^2}$$

La validez de la medición se refiere a que tan acertada es la respuesta a una pregunta en relación con el valor verdadero del constructo que se está midiendo (de Leeuw, 2008). Esto nos

indica si los valores de los indicadores tienen sentido dentro del diseño de la investigación. La correlación entre T y x , $P_{T,x}$, se conoce como la validez teórica de una variable x porque mide que tan bien se correlaciona un elemento observado con un constructo de interés latente (Rossi, 2013).

Todos estos son los elementos que tiene que tomar en cuenta el investigador al momento de diseñar y analizar la encuesta, sin embargo, otro aspecto fundamental es asegurar que tras haber diseñado la encuesta y seleccionado el muestreo, las personas respondan al cuestionario o las entrevistas. Para garantizar que los encuestados respondan bien a las preguntas se espera que la mayor cantidad posible realice un proceso de 4 pasos cognitivos.

1. Comprensión: atiende a las preguntas e instrucciones, identifica la información buscada y relaciona los términos utilizados a conceptos relevantes
2. Recuperación: recupera memorias genéricas o específicas para llenar cualquier detalle faltante
3. Juicio: evalúa la relevancia de las memorias recuperadas e integra el material recuperado
4. Respuesta: Decide cuál es su respuesta o edita su respuesta

Los encuestados no siempre realizan estos pasos cognitivos, ni tampoco es necesario en ciertas ocasiones, pero si permiten que entiendan bien lo que se les está preguntando y puedan dar una respuesta verdadera para la recolección de datos (Tourangeau, 2000). En muchas ocasiones esto ya no se encuentra dentro del control de los investigadores, algo que deben tomar en cuenta al hacer la medición de los indicadores. No obstante, existen maneras de optimizar el proceso cognitivo de los encuestados; esto se da en el diseño de las preguntas como en el método de recolección de datos que se utiliza. En los últimos 25 años se han inventado múltiples métodos de recolección de datos de encuestas. Existen entrevistas cara a cara, entrevistas por

teléfono, cuestionarios en línea, entrevistas personales con asistencia de computadoras, cuestionarios a mano, respuestas de voz interactivas, entrevistar por teléfono con asistencia de computadoras, etc. Desde una perspectiva analítica, entender como fueron recolectados los datos es vital para evaluar la calidad de la encuesta (Groves, 2004).

Una vez que se diseñó la encuesta, se decidió la muestra y se recopiló la información de los encuestados se procede al análisis de resultados. Existen cuatro factores que determinan como son analizados los datos: el número de variables que están siendo analizadas, el nivel de medición de las variables, si se quiere utilizar los datos para un análisis descriptivo o un análisis diferencial y las responsabilidades éticas de la encuesta (de Vaus, 2014). Si se desea describir una sola característica al momento se utiliza un análisis univariable (por ejemplo, ingresos, o ideología). Si se quiere analizar dos variables simultaneas se utiliza un análisis bivariable, que tiene como objetivo determinar si dos variables están relacionadas o no, y si los están, describir como se relacionan (Grasso, 2016); por ejemplo, si la variable ingresos está relacionada a la variable ideología. Para analizar tres o más variables simultaneas se utiliza el método multivariable, para determinar cuáles variables independientes afectan a una variable dependiente y en qué medida.

El último paso del análisis, después de escoger cual método es el más adecuado, es escoger el tipo de estadística que se va a usar para el caso, si estadística descriptiva o estadística inferencial. La primera consiste en describir y resumir las características principales de un conjunto de datos. La segunda consiste hacer inferencias o predicciones acerca de una población basándose en una muestra de datos recopilados de esa población (Gravetter, 2021). Independientemente del método, en toda descripción estadística existe un intervalo de confianza que estima un margen de error, esto es debido a que una muestra nunca va a ser lo suficientemente grande para medir a una población completa, por lo que nunca va a ser una representación exacta de la realidad.

El diseño y análisis de una encuesta electoral debe seguir todos estos principios para ser efectiva. Una característica especial de las encuestas electorales es su objetivo, ya que son un intento de previsión del futuro en base a la información del presente. La unidad agregada (normalmente la nación) es la unidad de análisis de la variable dependiente, más que la unidad individual (el votante) (Lewis-Beck, 2005). Se recoge una muestra que intenta ser representativa de la población, se pregunta por cual candidato o partido planea votar, se pondera por el margen de error y en base a eso se predice los resultados finales en las urnas.

Fallos de encuestas electorales

Que una encuesta no logre acertar en su predicción del resultado electoral no es un fenómeno nuevo, ya ha habido casos celebres de encuestas fallando en sus resultados que todavía sirven como referencia para la investigación actual. Sin embargo, el consenso general dentro de la academia es que hay más encuestas que están fallando en su objetivo de prever el futuro en base a los resultados obtenidos en el presente y el pasado. Se cree que existen muchas razones por las que los encuestadores no pueden predecir las elecciones. En primer lugar, el porcentaje de respuestas a encuestas realizadas tradicionalmente ha disminuido y cada vez es más difícil obtener la opinión de la gente. Las tasas de respuesta en las encuestas telefónicas con entrevistadores en vivo continúan disminuyendo y han alcanzado recientemente un límite inferior del 6%. Las tasas de respuesta podrían ser incluso más bajas para otras metodologías, como encuestas por cuestionarios en línea o respuestas de voz interactivas. Agravado con la disminución de las tasas de respuesta son el problema concomitante de la tergiversación de las muestras de la encuesta. Es decir, la muestra encuestada por los encuestadores no representa la población demográfica distribuciones de la población general (Zhou, 2021).

Como se mencionó anteriormente, la inferencia estadística permite extrapolar características de una muestra y que sea representativa de toda la población objetivo (en este

caso la nación). Hay dos pasos inferenciales que son centrales para las características de una encuesta.

1. Las respuestas dadas por la gente deben describir con precisión las características de los encuestados
2. El subconjunto de personas participando en la encuesta debe tener características similares a las de la población general.

Si es que no se cumplen ambas condiciones las estadísticas de la encuesta están sujetas a error (Groves, 2014).

Estas son explicaciones de los posibles fallos desde un punto de vista metodológico; otros autores argumentan que la capacidad de predicción se basa en la propia personalidad de los investigadores. La idea básica es que los investigadores y pensadores se pueden dividir en dos grandes categorías.

1. Los erizos son personalidades de tipo A, que creen en las grandes ideas, en principios rectores del mundo que se comportan como si fueran leyes físicas y sustentan prácticamente todas las interacciones en la sociedad. Pensadores como Karl Marx y la lucha de clases, o Sigmund Freud y el inconsciente.
2. Los zorros, por otro lado, son criaturas que creen en una gran cantidad de pequeñas ideas y en la adopción de una multitud de enfoques hacia un problema. Ellos tienden a ser más tolerantes con los matices, la incertidumbre, la complejidad y la disensión opinión. Si los erizos son cazadores y siempre buscan la gran presa, entonces los zorros son recolectores.

Los zorros son mucho mejores en realizar predicciones que los erizos. La política puede ser especialmente susceptible a malas predicciones precisamente por su elemento humano: una buena jornada electoral compromete nuestra sensibilidad dramática y crea sesgos en los criterios de los encuestadores. Esto no significa que uno deba sentirse totalmente ser imparcial

sobre un acontecimiento político para poder hacer una buena predicción al respecto. Pero sí significa que la actitud distante de un zorro puede dar sus frutos (Silver, 2012).

Sin embargo, existen factores que están fuera del control y de la personalidad de los investigadores que puede llegar a influenciar la calidad de los datos. El mayor problema es el sesgo de selección de los encuestados. La gente que responde a las encuestas electorales son gente que por su propia voluntad está dispuesta a responder; esto significa que tratar de crear una muestra representativa es mucho más difícil debido a que personas que no tienen tiempo o voluntad de responder a las encuestas de todas maneras vota y su opinión no se encuentra representada en ninguna base de datos (Jowell, 1993). Esta es una explicación que es independiente de lo que haga el propio investigador y no tiene una solución aparente; pero no implica que no se puedan realizar predicciones adecuadas.

Estudios de pronósticos electorales en América Latina

En América Latina el estudio de pronósticos electorales no está tan desarrollado como en Estados Unidos, Francia o Reino Unido. Esto se debe a diferentes factores como la relativa novedad de esta área de investigación en la región, la falta de investigadores calificados en el tema, etc. Sin embargo, se han realizado múltiples avances en ampliar la investigación en casi todos los países de América Latina. Algunos artículos han propuesto métodos para minimizar la diferencia entre predicciones y resultados electorales. Se basan en hallazgos que surgen de democracias establecidas, donde se llevado a cabo la mayor parte de la investigación, pero se ha centrado en llenar el vacío para las nuevas democracias, que hasta ahora han sido en gran medida ignoradas por la literatura. Han propuesto un TSM y ha calculado estimaciones con algoritmos bayesianos y cadenas de Markov. Han probado el modelo utilizando datos provenientes de 11 países y 26 elecciones en América Latina y produjo resultados notablemente precisos. En comparación con las encuestas, los modelos no sólo producen un pronóstico más preciso para cada elección de la muestra, sino que también lo hizo para nueve de cada diez

candidatos. Su principal contribución ha sido ampliar una línea de literatura que se ha desarrollado casi exclusivamente en las democracias establecidas. No sólo hay pocos estudios que hayan analizado en general las previsiones en las nuevas democracias (Bunker, 2020).

Con lo cual se puede concluir que, si bien no hay tantos estudios, existe un emergente grupo de trabajos que se enfocan en nuevas formas de interpretar y analizar las predicciones electorales en América Latina (Lewis-Beck, 2022).

Conclusiones

Diseñar y analizar una encuesta es un proceso largo y bastante complejo, que debe garantizar múltiples procesos metodológicos para garantizar que sea un trabajo verdaderamente científico al mismo tiempo que se trata de minimizar su error. La naturaleza de las encuestas electorales de tratar de prever el futuro en baso a la información recolectada en el presente y en el pasado hacen que su trabajo sea aún más complejo. Existen múltiples explicaciones de porque pueden fallar las encuestas en su objetivo de pronosticar los resultados de una elección, que pueden ser fallas metodológicas, el tipo de personalidad del encuestador o características de los encuestados que se encuentran fuera del control de los diseñadores de encuestas. No existen tantos estudios en América Latina a comparación de otras regiones, pero hay una incipiente cantidad de trabajos que están aportando enormemente a la literatura de encuestas electorales.

Posicionalidad

El estudio que estoy realizando se basa en el análisis cuantitativo y cualitativo de las encuestas electorales de Argentina, Brasil y Ecuador, por lo tanto, el alcance de la investigación se basa únicamente en las bases de datos empíricos que son posibles obtener. Este estudio busca saber cuándo se han equivocado las encuestas, cual fue el margen de equivocación y posibles hipótesis de porque se dieron los fallos. Mi enfoque de investigación es positivista dentro de la corriente epistemológica del empirismo lógico, por lo que cualquier conclusión que se llegue a obtener de la investigación debe ser tener una base metódica e intentar ser lo más objetiva

posible. Sin embargo, reconozco que, debido a las propias limitaciones de esta investigación en particular, los posibles resultados deben ser vistos como unas observaciones que pueden someterse a un proceso de falsabilidad por otros investigadores que estén analizando o llevando a cabo experimentos sobre el mismo tema.

Concibo a la ciencia dentro de su definición más ortodoxa; en la cual la describe como el conjunto de conocimientos obtenidos mediante la observación y el razonamiento, sistemáticamente estructurados y de los que se deducen principios y leyes generales con capacidad predictiva y comprobables experimentalmente (Etecé, 2022). Desde mi perspectiva, la ciencia es objetiva, esto incluye a las ciencias sociales, por lo cual cualquier razonamiento debe basarse en datos precisos y lo más libres de sesgos posibles. Si bien considero que existen dimensiones en las ciencias sociales que pueden potencialmente influenciar los argumentos de una investigación, dependiendo de las características del investigador o del grupo observado, tales como rasgos culturales, clase social, género, edad, educación, religión, orientación sexual, o posturas políticas. No considero que los investigadores deberían tener un posicionamiento personal en base a sus propias características al momento de hacer ciencia, ya que inevitablemente va a llevar a sesgos en el manejo de información o la formulación de conclusiones. Por eso mismo, yo no considero ninguno de estas posibles características relevantes para mi investigación.

Reflexibilidad

El proceso de investigación de esta tesis no es diferente a otras investigaciones que he realizado sobre diferentes temas. La estructura y los métodos son familiares para mí y para todos aquellos que creemos en una manera positivista de hacer ciencia. El círculo de investigadores en los que me inspiro no comparte la idea de que debe haber una posicionalidad para realizar investigación, por las mismas razones que argumento yo mismo. Al momento de realizar esta tesis algo que si me llamó la atención y me invitó a pensar de una manera diferente

sobre como concibo yo la epistemología, fue durante la revisión de literatura. El estadístico y escritor estadounidense Nate Silver, fundador de *FiveThirtyEight* y uno de los académicos más importantes en la historia de investigación mediante encuestas, argumenta en su libro *The Signal and the Noise* que una de las principales razones por las que no se cumplen las predicciones de las encuestas se debe a los rasgos de personalidad de los investigadores y los divide en dos grandes grupos, uno de los cuales puede predecir con mayor precisión el resultado (año). Sigo con la creencia que la ciencia debe practicarse de una manera específica, pero el argumento de una autoridad sobre mi tema de investigación sobre como las características personales pueden afectar los resultados de una investigación me lleva a considerar nuevas posibilidades sobre cómo llevar a cabo una investigación. Cabe resaltar que las características de personalidad a los que se refiere Nate Silver son puramente psicológicas, a diferencia de otras características clase social o género, lo cual genera otra perspectiva de cómo abordar las ciencias sociales, a mi parecer más precisa.

Discusión

Criterios generales del estudio de encuestas electorales

Tras haber comprendido cómo funciona una encuesta, las posibles teorías de por qué fallan las encuestas electorales y qué tipo de estudios se han realizado en América Latina, se puede explicar el propósito de esta investigación. Esta tesis va a medir el nivel de fallo de varias encuestas de elecciones generales en Argentina, Brasil y Ecuador durante los últimos años: desde 2014 al 2023; y va a analizar cuáles variables y en qué medida, estas pueden influenciar sobre el nivel la desviación de una encuesta respecto a los resultados finales obtenidos en la elección general.

Como se mencionó anteriormente, una encuesta es un método sistemático de recolección de información, el cual recoge datos mediante la construcción de medidores cuantitativos a los que se les asigna un valor, dependiendo de la cantidad de personas que voten,

por cada posible respuesta (Groves, 2004). Las encuestas electorales son más simples y a la vez más complejas que otras encuestas. Lo que busca toda encuesta es medir un concepto o un constructo y en el caso de una encuesta electoral ese constructo es relativamente sencillo de cuantificar: intención de voto. Para medir este indicador se utiliza siempre la siguiente pregunta “¿si la elección fuera hoy, por cual candidato votaría?”. Esta pregunta plantea una imagen temporal sobre un posible resultado en el momento que fue hecha. Con base en los resultados se realizan análisis por parte de politólogos, relacionistas internacionales, periodistas y las propias campañas políticas para tratar de predecir el resultado final.

Un ejemplo de una encuesta electoral se ve así:

¿Si la elección fuera hoy, a cuál candidato le daría su voto?	
Candidato A	33%
Candidato B	30%
Candidato C	26%
Nulo	8%
No sabe	3%

Tabla 2, Elaborado por Felipe Grijalva, 2024

En este caso la encuesta mide la cantidad de votos totales. En toda elección existe una cantidad del electorado que no da su voto a ningún candidato, lo que se considera un voto no válido. Estos votos si bien son contabilizados por las instituciones electorales de cada país estos no se ven reflejados en los porcentajes finales, que en este caso representan al 11% en conjunto. Sin embargo, es importante para las encuestadoras no excluir este segmento, debido a que se puede tener una idea más precisa de la intención de voto de todo el país. Con este mismo ejemplo se pueden obtener los siguientes resultados:

¿Si la elección fuera hoy, a cuál candidato le daría su voto?	
Candidato A	37%

Candidato B	34%
Candidato C	29%

Tabla 3, Elaborado por Felipe Grijalva, 2024

Este es el porcentaje de votos válidos que obtiene cada candidato cuando se excluye votos nulos e indecisos. En ambos casos quien resulta ganador es el candidato A. Toda encuesta electoral siempre hace una serie de preguntas demográficas a la población antes de preguntar por cual candidato va a dar su voto, esto para recolectar otras variables como sexo, etnicidad o clase social con el fin de visualizar alguna posible correlación. Algunas encuestadoras hacen esto con fines académicos, o para transferir esa información para la campaña de algún candidato; es muy importante mencionar que, si bien toda encuestadora trata de acertar con precisión los resultados, algunas de ellas pueden tener preferencias hacia ciertos partidos políticos o candidatos, lo que aumenta las posibilidades de dar información parcializada.

Si bien este es un factor que se debe tomar en cuenta al momento de analizar los resultados que muestran las encuestadoras, el propósito de este estudio no es determinar cuáles encuestas están sesgadas o si los sesgos por sí solos son los que generan los fallos. Esta investigación va a analizar los criterios técnicos y ciertos factores externos que pueden influenciar la cercanía entre los resultados de la encuesta y los resultados finales de la elección.

Características relevantes de los países investigados

Un aspecto importante que puede influir en el diseño y análisis de encuestas electorales es el propio sistema político del país, debido a que el tipo de gobierno y la forma de elegir a los representantes cambia la dinámica de los electores, por lo tanto, una encuesta hecha en Alemania puede realizarse o interpretarse de la misma manera que una encuesta hecha en Estados Unidos. Argentina, Brasil y Ecuador tienen semejanzas importantes que permiten analizar las encuestas de sus elecciones en conjunto. Los tres países son repúblicas

presidencialistas, celebran elecciones cada 4 años y se permite la reelección inmediata una sola vez.

El Código Electoral Nacional de la República Argentina, según la Ley N° 19.945 indica que son electores nacionales los ciudadanos de ambos sexos nativos, por opción y naturalizados, desde los dieciocho años cumplidos de edad, que no tengan ninguna de las inhabilitaciones previstas en esta ley. Existe un sistema de representación proporcional, por lo que el valor de todos los votos es igual. Según la Ley N.º 25.611, en caso de que ningún candidato obtenga más del 45% de los votos válidos o supere el 40% con una diferencia mayor al 10% sobre el segundo candidato más votado habrá una segunda ronda electoral y ganará el candidato más votado.

Según el CÓDIGO ELECTORAL de Brasil, en la Ley N.º 4.737, todos los ciudadanos mayores de 18 años tienen el derecho al voto, siempre que estén inscritos en el registro electoral. Brasil también utiliza un sistema de representación proporcional para asignar escaños en sus cámaras legislativas. Esto significa que los escaños se distribuyen entre los partidos de acuerdo con el número de votos que reciben.

De acuerdo con el CÓDIGO DE LA DEMOCRACIA de Ecuador, el voto es obligatorio para todos los ciudadanos mayores de 18 años, excepto para aquellos que se encuentren en algunas situaciones específicas establecidas por la ley. Ecuador utiliza un sistema de representación proporcional para la elección de cargos legislativos, como asambleístas y concejales. Los escaños se distribuyen entre los partidos políticos en proporción al número de votos que reciben. La única diferencia significativa entre Ecuador y los otros dos países es que dentro del Art.283 de la CONSTITUCIÓN DE LA REPÚBLICA DEL ECUADOR, ninguna persona puede publicar a menos de diez días de la elección.

Estas características de los sistemas electorales permiten una comparación directa entre los tres países, pese a que tengan diferencias en tamaño de población y demografía.

Metodología

Este es un estudio primero de estadística descriptiva y luego estadística inferencial, donde se van a presentar y analizar las variables de una muestra de datos. Se mide el nivel de error que tuvo una encuesta electoral en resultado a las elecciones generales. Para esto se recolectó una muestra de 375 encuestas que se realizaron en 3 países (Argentina, Brasil y Ecuador), durante los periodos electorales de 2014, 2015, 2017, 2018, 2019, 2021, 2022 y 2023.

En total la base de datos que se encuentra en el anexo 1 cuenta con 10 indicadores, 8 de los cuales son las variables independientes de tipo cuantitativo y cualitativo. En ejemplo limitado de los valores posibles de cada variable dentro de la matriz que se encuentran en la base de datos utilizada para esta investigación, el Anexo 1, se ven en la siguiente tabla.

País	Argentina
	Brasil
	Ecuador
Año	2015
	2017
	2018
	...
Fecha	1 día antes de la elección
	7 días antes de la elección
	14 días antes de la elección
	...
	CEDATOS
	Datafolha
	Atlas Intel

Encuestadora	CELAG
	...
Ronda electoral	Primera vuelta
	Segunda vuelta
Muestra	1000
	2780
	4500
	...
Cantidad de candidatos	2
	5
	8
	...
Método de recolección	Entrevista personal
	Teléfono
	Online

Tabla 4, Elaborado por Felipe Grijalva, 2024

Para garantizar que los resultados de esta investigación sean más representativos y precisos se utilizaron dos variables dependientes para medir el tamaño del error de una encuesta respecto a los resultados de la elección. Se definió a los fallos de encuestas electorales de en las variables dependientes como una serie de cálculos que llevan a dos resultados:

1. La suma de la desviación de los resultados de la encuesta contando votos ponderados válidos
2. La suma de la desviación de los resultados de la encuesta contando votos ponderados totales

La suma de la desviación de los resultados de la encuesta contando votos ponderados válidos, para propósitos de este estudio se va a llamar “variable dependiente 1”, y la suma de la desviación de los resultados de la encuesta contando votos ponderados totales se va a llamar “variable dependiente 2”.

La variable dependiente 1 se calcula como la suma de la desviación ponderada total de cada candidato es igual al valor absoluto de la diferencia entre el porcentaje del resultado real total y el porcentaje de resultado pronosticado total, multiplicado por el porcentaje del resultado real total.

$$VD1 = \sum_{i=1}^N |R_{rvi} - R_{pvi}| \times R_{rvi}$$

- VD1 representa la variable dependiente 1
- N es el número total de candidatos
- R_{rvi} es el porcentaje del resultado real total para el candidato i.
- R_{pvi} es el porcentaje del resultado pronosticado total para el candidato i.

La variable dependiente 2 se calcula como la suma de la desviación ponderada válida de cada candidato es igual al valor absoluto de la diferencia entre el porcentaje del resultado real válido y el porcentaje de resultado pronosticado válido, multiplicado por el porcentaje del resultado real válido.

$$VD2 = \sum_{i=1}^N |R_{rti} - R_{pvti}| \times R_{rti}$$

- VD2 representa la variable dependiente 2
- N es el número total de candidatos
- R_{rti} es el porcentaje del resultado real válido para el candidato i.
- R_{pvti} es el porcentaje del resultado pronosticado válido para el candidato i.

Se utiliza la suma de los votos ponderados en lugar de un promedio de desviación debido a que el peso del porcentaje es diferente para cada candidato. Una encuesta que le asigne un porcentaje de votos, ya sean totales o válidos, a un candidato todavía está expuesta a ciertos rangos de imprecisión. Contar el voto ponderado ayuda a corregir estos fallos gracias a que da una mayor representatividad de la muestra, al asignarle cierto peso a candidatos que están sobrerrepresentados o subrepresentados. Esto permite mejorar la representatividad del porcentaje obtenido por un candidato en una encuesta y corregir sesgos que se obtiene de la muestra.

La investigación no va a tomar en cuenta variables como el margen de error de cada encuesta o su intervalo de confianza debido a que la variabilidad entre cada una de las encuestas es demasiado baja para describir algún efecto dentro de las variables dependientes. Así mismo solo se recogieron encuestas realizadas dentro de la jornada electoral oficial de la primera y segunda ronda para candidatos oficialmente registrados por el respectivo organismo electoral regulatorio de cada país; por lo que encuestas fuera del periodo de tiempo en el que inician la campaña, encuestas para elecciones primarias y encuestas que especulan cual puede ser el candidato de algún partido en específico no son tomadas en cuenta.

Durante ciertas jornadas electorales ocurrieron circunstancias extraordinarias sobre los candidatos registrados oficialmente que deben ser mencionadas; la primera es sobre las encuestas que tomaban en cuenta al candidato del Movimiento Construye, Fernando Villavicencio, el cual fue víctima de un asesinato político el 9 de agosto de 2023, tan solo a 11 días de las elecciones presidenciales de Ecuador. Fernando Villavicencio fue reemplazado por Christian Zurita para la candidatura del movimiento Construye; para propósitos de este estudio se toman los resultados en las encuestas de la intención de voto de Christian Zurita como una continuación de los resultados obtenidos hasta entonces por Fernando Villavicencio. La segunda excepción es sobre las encuestas que tomaban en cuenta al candidato por el Partido

Socialista Brasileño, Eduardo Campos, quien falleció en un accidente aéreo el 13 de agosto de 2014, a 53 días de las elecciones presidenciales de Brasil. Eduardo Campos fue reemplazado por Marina Silva, por lo que se toman los resultados en las encuestas de la intención de voto de Marina Silva como una continuación de los resultados obtenidos hasta entonces por Eduardo Campos.

Limitaciones del estudio

Con base en los datos obtenidos se realizó una serie de regresiones múltiples para determinar la correlación que existe entre las variables independientes y las variables dependientes. Si bien este análisis no es capaz de determinar por sí solo si una, o varias de las variables independientes llegan a tener una relación causal con el valor de las variables dependientes, esta tesis puede dar evidencia para afirmar o falsear varias hipótesis a cerca de cuales características influyen en los fallos de las encuestas electorales en América Latina. La siguiente gráfica muestra la relación entre variables.

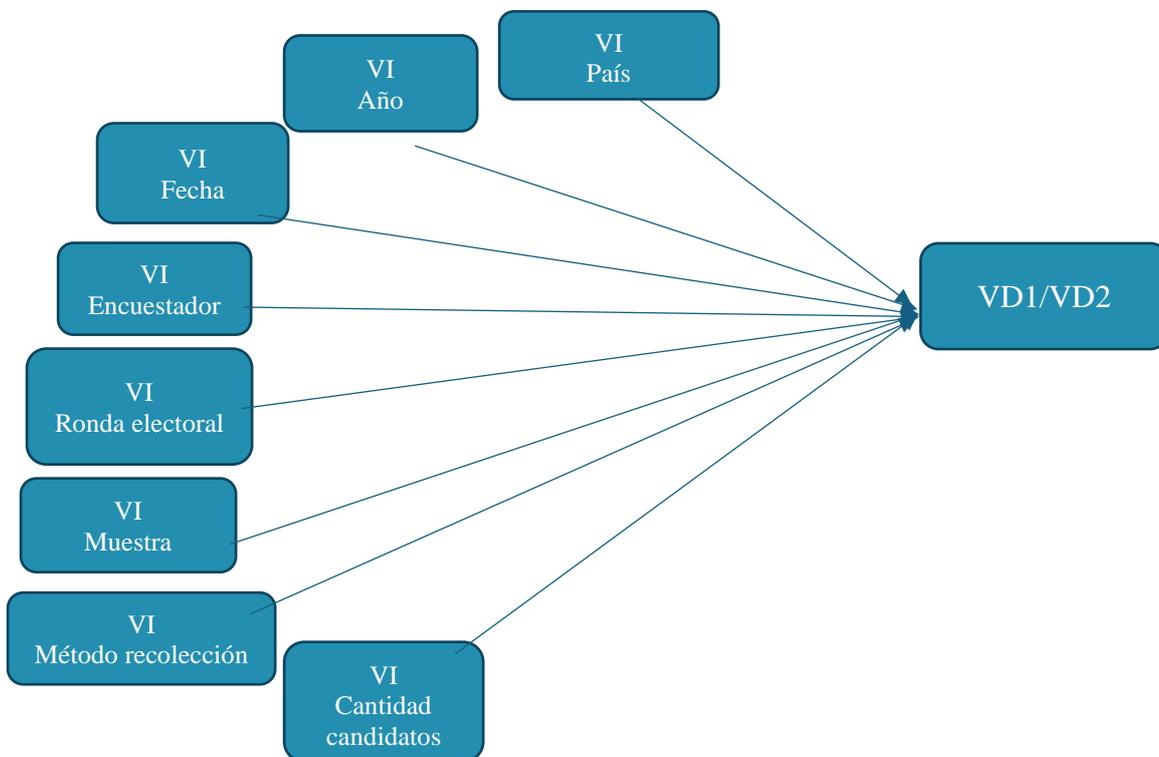


Figura 2, Elaborada por Felipe Grijalva, 2024

Resultados

Se realizó una regresión multivariable tomando para analizar a la variable dependiente 1 tomando como variables todas las que se encuentra en dentro de la figura 1. Dentro del programa se renombró a varias variables y se cambió todos los formatos *strings* a *números*. Las variables “numeric_Pas” es la variable “Países”; la variable “numeric_RD” es la variable “Ronda electoral”; la variable “numeric_MR” es la variable “Método de recolección”; la variable “numeric_Encuestadora” es la variable “Encuestadora”; la variable “Ao” es la variable “Año”; y la variable “numeric_Fecha” es la variable Fecha. En base a la regresión se pueden observar los siguientes resultados.

```
. regress VD1 numeric_Pas numeric_RD numeric_MR numeric_Encuestadora numeric_Fecha Muestra Ao NC
```

Source	SS	df	MS			
Model	587.988579	8	73.4985724	Number of obs =	375	
Residual	2413.49578	366	6.59425076	F(8, 366) =	11.15	
				Prob > F	= 0.0000	
				R-squared	= 0.1959	
				Adj R-squared	= 0.1783	
Total	3001.48436	374	8.02535925	Root MSE	= 2.5679	

VD1	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
numeric_Pas	.5440226	.2022328	2.69	0.007	.1463386	.9417066
numeric_RD	-1.136872	.5199152	-2.19	0.029	-2.159268	-.1144765
numeric_MR	-.109293	.1549653	-0.71	0.481	-.414027	.1954411
numeric_Encuestadora	.0050866	.004203	1.21	0.227	-.0031784	.0133517
numeric_Fecha	.0010745	.0075199	0.14	0.886	-.0137132	.0158622
Muestra	-.0000377	.0000397	-0.95	0.342	-.0001157	.0000403
Ao	.043885	.0426091	1.03	0.304	-.0399043	.1276744
NC	.1358472	.0591281	2.30	0.022	.0195737	.2521207
_cons	-84.26781	86.08675	-0.98	0.328	-253.5545	85.01893

Figura 3, Elaborado por Felipe Grijalva, 2024

El valor de la prueba F (evaluación estadística de la igualdad de las medias) indica que el modelo en su conjunto es estadísticamente significativo (p-valor = 0.0000), lo que significa que una o más de las variables independientes tienen una correlación significativa con la variable dependiente 1. R-cuadrado ajustado tiene un valor de 0.1783, lo que significa que un 17.83% de variabilidad de la variable dependiente 1 puede ser explicada por el modelo. La variable “numeric_Pas” tiene un coeficiente positivo de 0.544, con un error estándar de 0.202.

Con un p-valor = 0.007 el resultado es estadísticamente significativo. Lo que indica que el país donde se realizó la encuesta tiene una correlación con el nivel de desviación entre los resultados de la encuesta y el resultado final. El valor numérico de los países está ordenado por su posición en el orden alfabético, por lo que el valor de Argentina es 1, Brasil es 2, y Ecuador es 3. Lo que indica que entre los tres países las encuestas realizadas en Ecuador tienen una mayor probabilidad de tener una desviación del porcentaje de votos válidos más grande, mientras que las encuestas realizadas en Argentina tienen una menor probabilidad de tener una desviación del porcentaje de votos válidos más grande. Brasil tiene una probabilidad mayor de tener una desviación de porcentaje de votos válidos más grande a comparación de Argentina, pero menor a Ecuador.

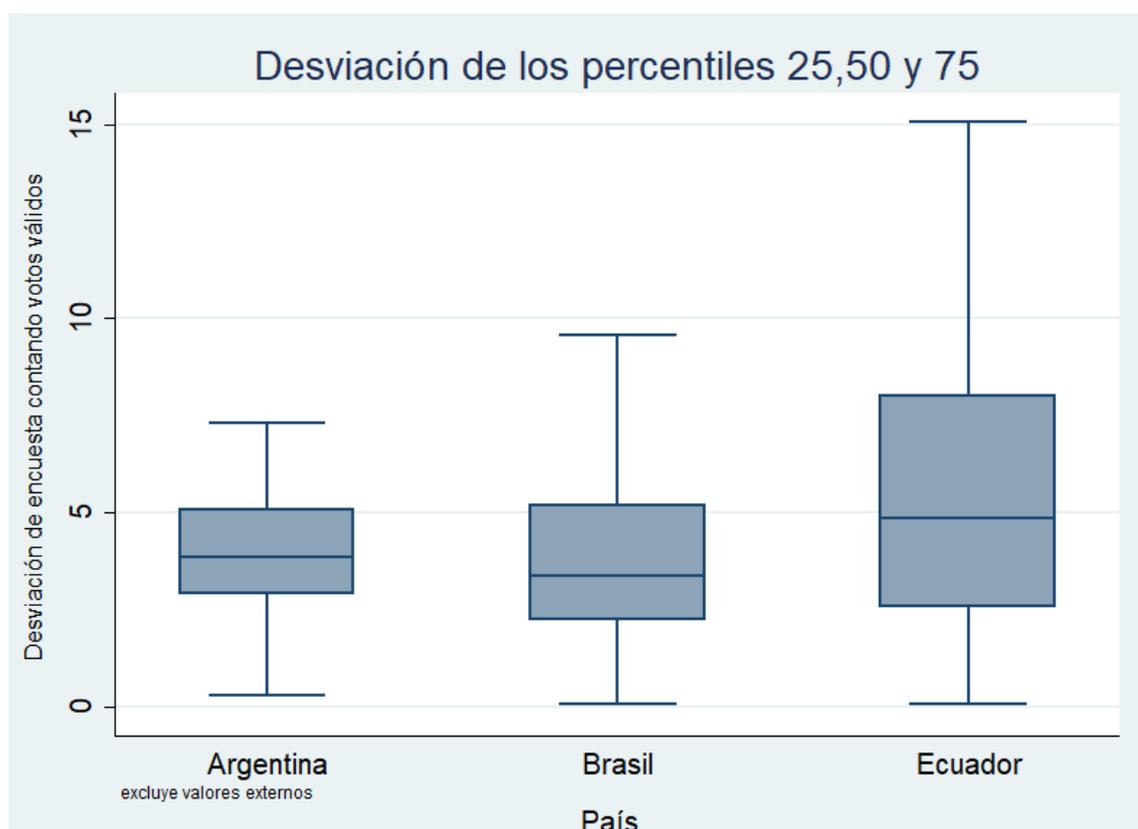


Figura 4, elaborado por Felipe Grijalva, 2024 año

La variable `numeric_RD` tiene un coeficiente negativo de -1.137, con un error estándar de 0.520. Con un p-valor = 0.029 el resultado es estadísticamente significativo. Esto indica que la ronda electoral tiene una correlación con el nivel de desviación entre los resultados totales

de la encuesta y el resultado válido final. El valor numérico de la ronda electoral está asignado dependiendo de su orden, por lo que el valor de “encuesta de primera vuelta” es 1 y “encuesta de segunda vuelta” es 2. Lo que indica que una encuesta realizada para la segunda vuelta electoral tiene una menor probabilidad de tener una desviación de porcentaje de votos válidos más grande a comparación de una encuesta realizada en primera vuelta electoral.

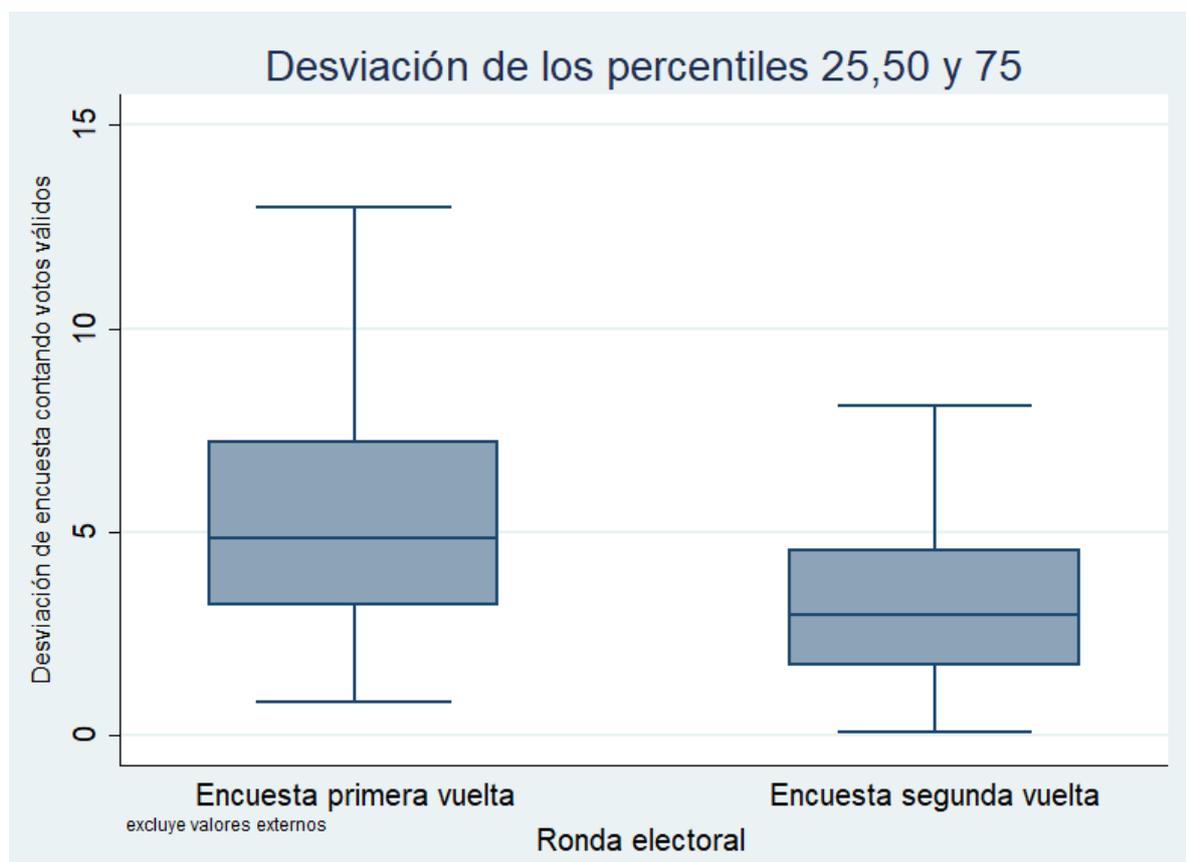


Figura 5 elaborado por Felipe Grijalva, 2024

La variable NC tiene un coeficiente positivo de 0.136, con un error estándar de 0.0591. Con un p-valor = 0.022 el resultado es estadísticamente significativo. Lo que indica que la variable “número de candidatos” tiene una correlación con el nivel de desviación entre los resultados totales de las encuestas y el resultado total final. Esto indica que a mayor el número de candidatos que participen en una ronda electoral, es mayor la probabilidad de que la encuesta

llegue a tener una desviación de porcentaje de votos válidos más grande a si es que participa una menor cantidad de candidatos.

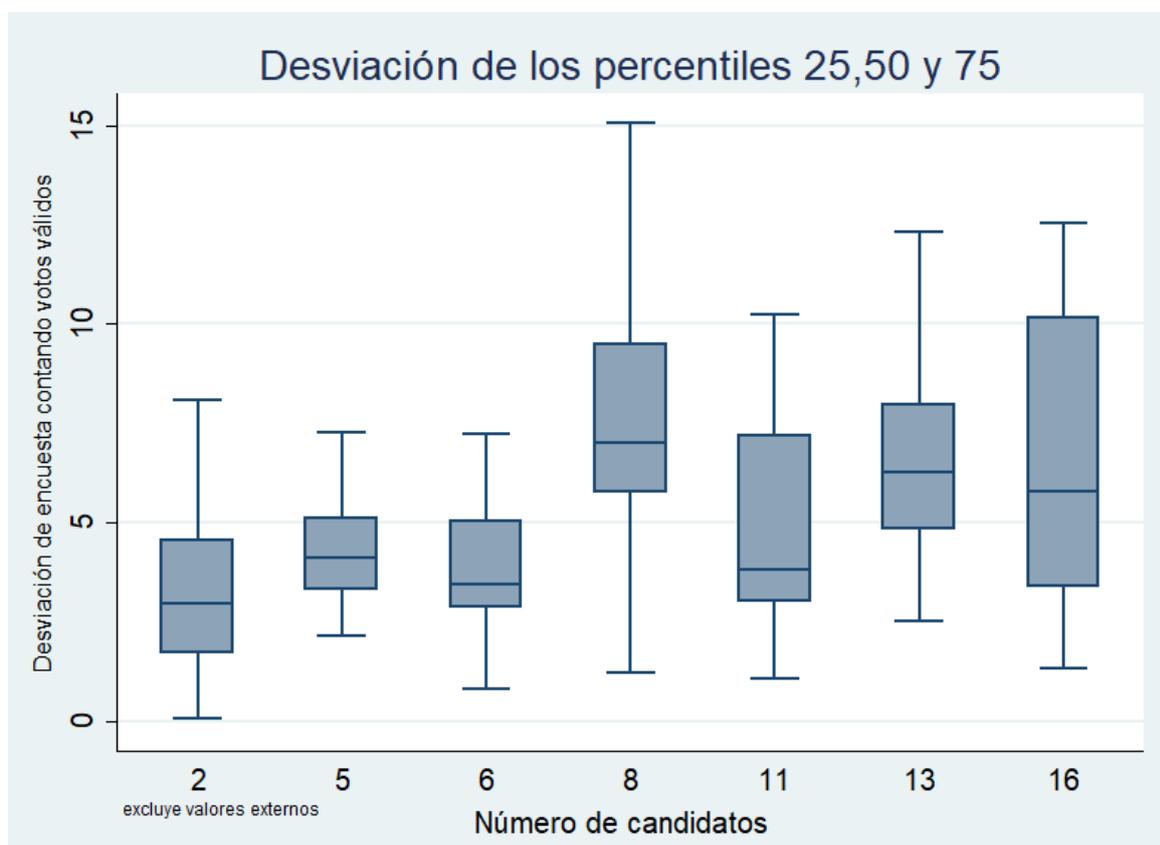


Figura 6, elaborado por Felipe Grijalva, 2024

Las variables “numeric_MR”, “numeric_Encuestadora”, “numeric_Fecha”, “Muestra” y “Ao” tienen todas un p-valor mayor 0.05, por lo que ninguna es estadísticamente significativa y por ende no tienen correlación con la variable dependiente 1. Esto significa el método de recolección, la cantidad de días en las que se realizó la encuesta antes de la elección, la propia encuestadora, el año de la encuesta y el tamaño de la muestra no tienen un efecto sobre si una encuesta tiene una mayor o menor desviación a comparación del resultado de votos válido real.

Utilizando las mismas variables independientes se corrió una segunda regresión multivariable para determinar el impacto sobre la variable dependiente 2. La hipótesis dicta que los resultados de ambas regresiones son compatibles entonces el modelo tiene una mayor validez. A continuación, se pueden observar los resultados.

```
. regress VD2 numeric_Pas numeric_RD numeric_MR numeric_Encuestadora numeric_Fecha Muestra Ao NC
```

Source	SS	df	MS			
Model	468.215428	8	58.5269285	Number of obs =	375	
Residual	2335.93448	366	6.38233465	F(8, 366) =	9.17	
				Prob > F =	0.0000	
				R-squared =	0.1670	
				Adj R-squared =	0.1488	
Total	2804.14991	374	7.49772703	Root MSE =	2.5263	

VD2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
numeric_Pas	.6471706	.1989567	3.25	0.001	.2559289	1.038412
numeric_RD	.2235922	.5114928	0.44	0.662	-.7822414	1.229426
numeric_MR	-.334569	.1524549	-2.19	0.029	-.6343665	-.0347714
numeric_Encuestadora	.003457	.0041349	0.84	0.404	-.0046742	.0115882
numeric_Fecha	-.009043	.0073981	-1.22	0.222	-.0235911	.0055051
Muestra	-.0000569	.000039	-1.46	0.146	-.0001337	.0000198
Ao	.037366	.0419188	0.89	0.373	-.045066	.119798
NC	.1980179	.0581703	3.40	0.001	.083628	.3124079
_cons	-72.14881	84.6922	-0.85	0.395	-238.6932	94.39557

Figura 7, Elaborado por Felipe Grijalva, 2024

El valor de la prueba F también indica que el modelo en su conjunto es estadísticamente significativo (p-valor = 0.0000), lo que sugiere que una o más de las variables independientes tienen una correlación significativa con la variable dependiente 2. El R-cuadrado ajustado del modelo es 0.1488, lo que indica que el 14.88% de la variabilidad en la variable dependiente VD2 es explicada por el modelo.

La variable “numeric_Pas” tiene un coeficiente positivo de 0.647, con un error estándar de 0.199. Con un p-valor = 0.001, el resultado es estadísticamente significativo, lo que indica que existe una correlación entre el país donde se realizó la encuesta y la desviación entre el resultado de votos totales que predice una encuesta y el resultado de real de votos totales. Entre los tres países las encuestas realizadas en Ecuador tienen una mayor probabilidad de tener una desviación del porcentaje de votos totales más grande, mientras que las encuestas realizadas en Argentina tienen una menor probabilidad de tener una desviación del porcentaje de votos totales más grande. Brasil tiene una probabilidad mayor de tener una desviación de porcentaje de votos

totales más grande a comparación de Argentina, pero menor a Ecuador.

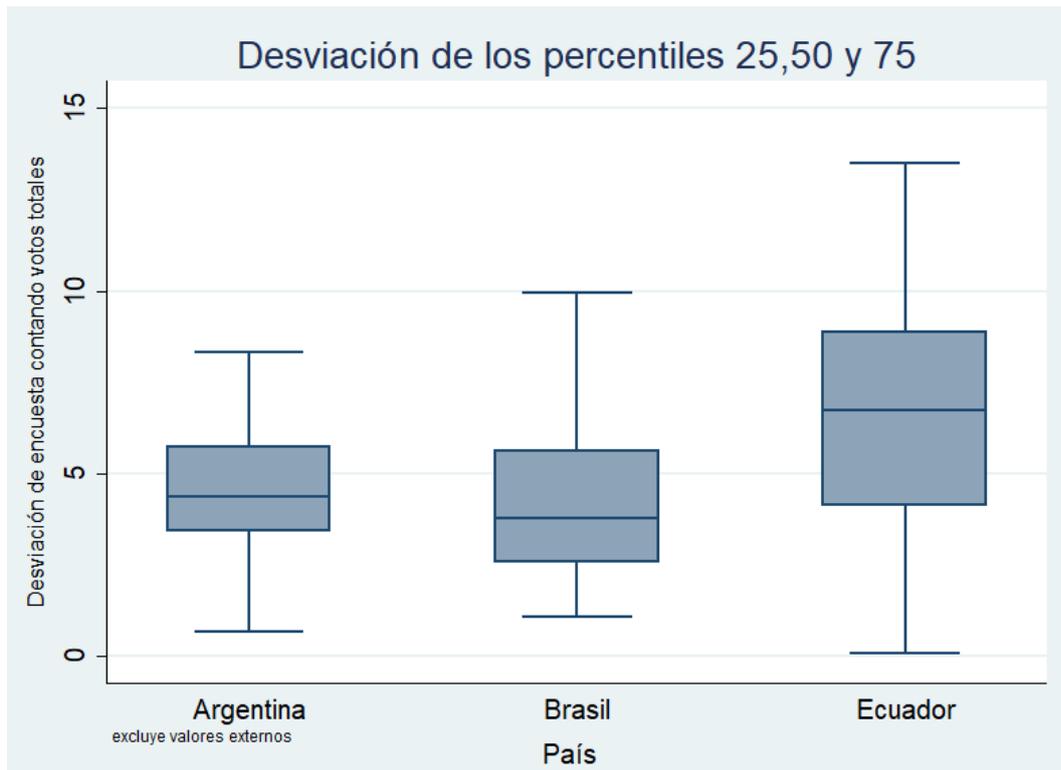


Figura 8, elaborado por Felipe Grijalva, 2024

La variable “numeric_MR” tiene un coeficiente negativo de -0.335, con un error estándar 0.1524. Con un p-valor = 0.029, el resultado es estadísticamente significativo, lo que indica que existe una correlación entre el método de recolección de datos y la desviación entre el resultado de votos válidos que predice una encuesta y el resultado de real de votos totales. El valor numérico de la variable “numeric_MR” está asignado por orden alfabético, por lo que el valor de “Entrevista personal” es 1, “Online” es 2 y “Teléfono” es 3. Lo que indica que entre los tres métodos de recolección de datos las encuestas hechas por teléfono tienen una probabilidad menor de tener una la suma de la desviación de los resultados de la encuesta contando votos ponderados válidos más grande, mientras que las encuestas realizadas mediante una entrevista presencial tienen una mayor probabilidad de tener la suma de la desviación de los resultados de la encuesta contando votos ponderados totales más grande. Las encuestas hechas en línea tienen una probabilidad menor de tener la suma de la desviación de los

resultados de la encuesta contando votos ponderados totales más grande en comparación al resultado de las encuestas realizadas mediante una entrevista presencial, pero mayor a las encuestas hechas por teléfono.

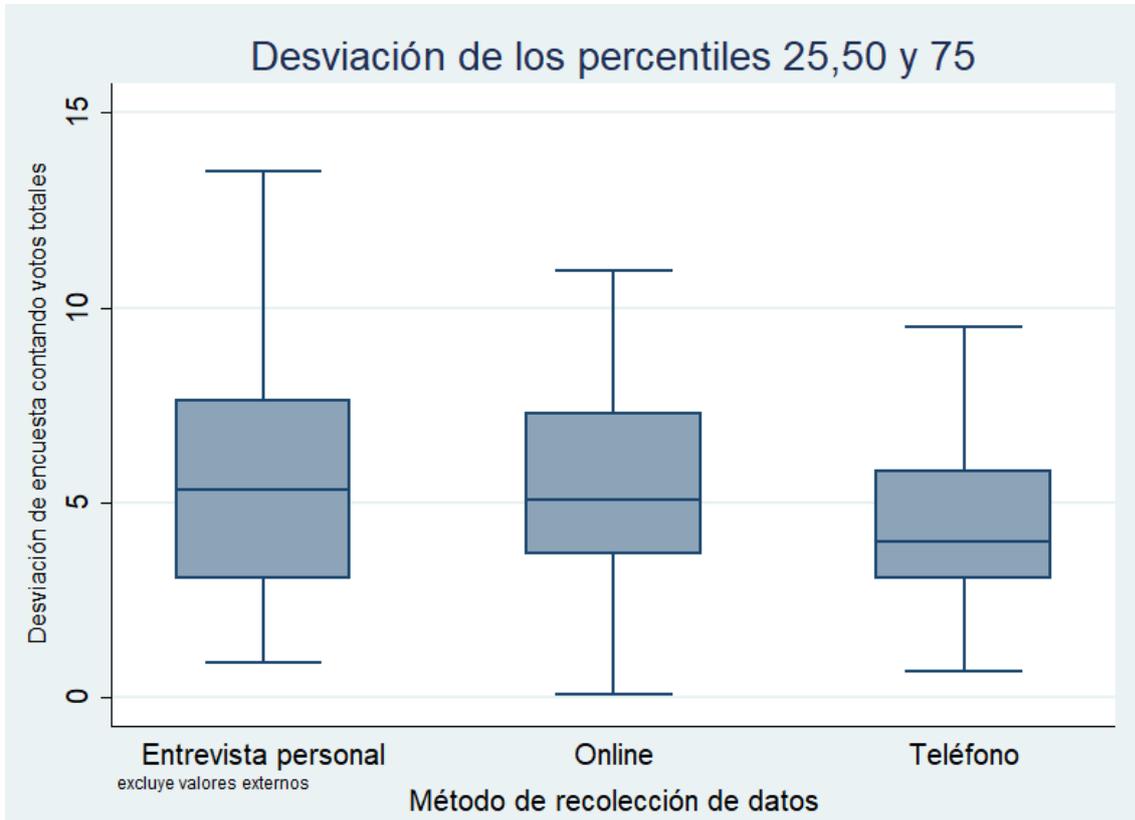


Figura 9, elaborado por Felipe Grijalva, 2024

Las variables “numeric_RD”, “numeric_Encuestadora”, “numeric_Fecha”, “Muestra” y “Ao” tienen un p-valor superior a 0.05, por lo que ninguna es estadísticamente significativa y por ende no tienen correlación con la variable dependiente 2. Esto indica que a la suma de la desviación de los resultados de la encuesta contando votos ponderados totales no afecta la ronda electoral, la encuestadora, la cantidad de días en los que se realizó la encuesta antes de la elección, el tamaño de la muestra ni el año en el que se realizó.

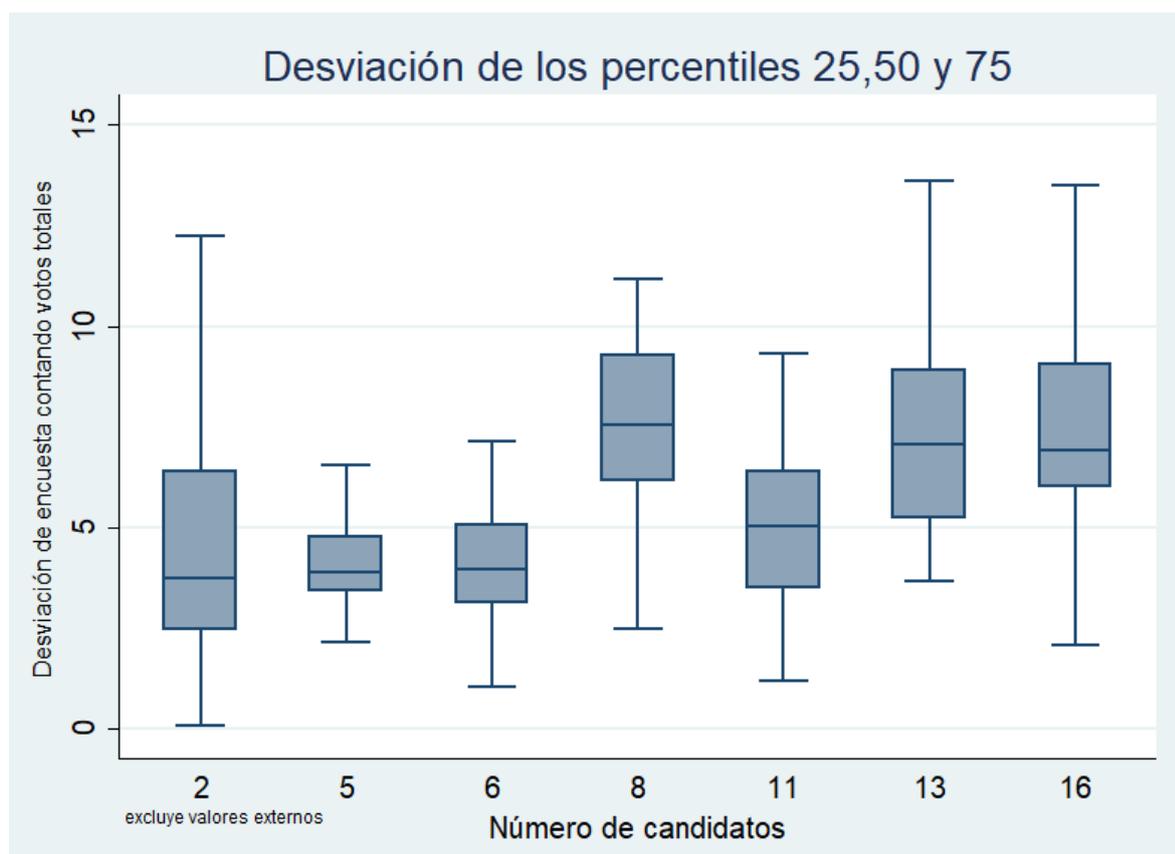


Figura 10, elaborado por Felipe Grijalva, 2024

La variable “NC” tiene un coeficiente positivo de 0.198, con un error estándar de 0.0581. Con un p-valor = 0.001, el resultado es estadísticamente significativo, lo que indica que existe una correlación entre el número de candidatos y la variable dependiente 2. A mayor el número de candidatos, existe una mayor probabilidad de que la suma de la desviación de los resultados de la encuesta contando votos ponderados totales sea más grande.

En base a los resultados obtenidos en ambas regresiones podemos afirmar que tan solo 3 de las 8 variables independientes tienen un efecto significativo sobre las variables dependientes. Es posible que existan otras variables fuera de la metodología de estudio de esta tesis que también puedan afectar el nivel de desviación de las encuestas y cuya inclusión dentro del modelo de regresión afecte el nivel de correlación entre las variables independientes. Sin embargo, los datos obtenidos actualmente indican que el número de candidatos, la ronda electoral y el país afectan a probabilidad de que tan precisa va a ser una encuesta, y la variable de método de recolección de datos dando resultados poco claros.

Conclusiones

Ambas regresiones tienen resultados similares respecto a cuál es el efecto de las variables independientes sobre las variables dependientes. El análisis de ambas regresiones indica en este estudio que, durante los últimos años, el nivel de precisión de las encuestas electorales en América Latina no ha mejorado ni empeorado. También se puede concluir que ciertas características específicas como el tamaño de la muestra, la fecha en la que se realizó la encuesta e incluso la propia encuestadora que realizó el estudio, no tienen un impacto significativo sobre la precisión de los resultados de la encuesta. Mientras que las variables que si tienen un impacto significativo sobre la precisión de la encuesta cuando mide tanto votos válidos como votos totales son el número de candidatos y el país en el que se realizó las encuestas. La variable de método de recolección de datos tiene un impacto significativo sobre la precisión de las encuestas, pero solo cuando están midiendo el porcentaje de votos totales de cada candidato, mientras que la variable de número de candidatos parece solo tener un impacto significativo sobre la precisión de las encuestas que miden el porcentaje de votos totales.

Como se mencionó anteriormente, estas no son las únicas variables que pueden llegar a afectar los resultados de las variables dependientes, el propio modelo de regresión múltiple indica que el mismo solo puede explicar un determinado porcentaje de la variabilidad de resultados. Se recomienda continuar con estudios sobre las variables que no son estadísticamente significativas dentro del modelo de regresión, especialmente las variables del tamaño de muestra, encuestadora y fecha de realización de la encuesta.

Los resultados están de acuerdo con aspectos de la literatura sobre fallos de encuestas electorales. Prosser resalta que en general, y alrededor del mundo, la precisión de las encuestas no ha empeorado con el paso del tiempo y que siempre ha habido márgenes de error similares durante décadas, independientemente del método de recolección de datos (2018). Groves, por otra parte, argumenta que la muestra es el principal factor que determina la precisión de una

encuesta, debido a que la representatividad de la población estudiada es el factor determinante (2009). Sin embargo, la representatividad de la muestra para una elección no depende tanto del tamaño del N en sí mismo, más que asegurarse que la cantidad de personas encuestadas tengan valor proporcional adecuado para la cantidad de personas encuestadas.

Zhenkun Zhou, en un estudio de caso realizado sobre las propias elecciones argentinas de 2019 argumenta que la principal causa de la falla de las encuestas para esa elección en particular fue el uso de la tecnología, ya que métodos de recolección de datos más modernos pueden captar mejores votos “volátiles” a comparación de encuestas tradicionales (2021). Los resultados de esta investigación no sustentan aquella conclusión, debido a que el método de recolección no tiene una correlación con la variable dependiente 1; mientras que la correlación con la variable dependiente 2 es menor al resto de variables independientes, además de que dentro del modelo el método de recolección probabilísticamente tiene un menor grado de desviación son las encuestas por teléfono, no las encuestas realizadas de forma online, aunque estas si tienen una mayor probabilidad de tener una desviación menor a comparación de encuestas realizadas mediante entrevistas presenciales.

En conclusión, existe una mayor probabilidad de que las encuestas se alejen más de los resultados finales dependiendo del país, de la ronda electoral, de la cantidad de candidatos que se encuentran participando en la elección y posiblemente del método de recolección de información, aunque esta última variable merece una especial atención debido a que tiene significancia en solo uno de los dos modelos de regresión. El tamaño de la muestra, la fecha de la encuesta, la encuestadora y el año no parecen ser lo suficientemente relevantes.

Bibliografía

- Adrogué, G., & Adrogué, G. (1998). Estudiar la opinión pública. Teoría y datos sobre la opinión pública argentina. *Desarrollo Económico*, 38(149), 387.
<https://doi.org/10.2307/3467385>
- Ayala Mora, M. E. (2008). *Resumen de historia del Ecuador*. Corporación Editora Nacional.
- Bunker, K. (2020). A two-stage model to forecast elections in new democracies. *International Journal of Forecasting*, 36(4), 1407–1419.
<https://doi.org/10.1016/j.ijforecast.2020.02.004>
- "Ciencia". Autor: Equipo editorial, Etecé. De: Argentina. Para: Concepto.de. Disponible en: <https://concepto.de/ciencia/>. Última edición: 1 de junio de 2022. Consultado: 15 de mayo de 2024 Fuente: <https://concepto.de/ciencia/#ixzz8aJc0tTPO>
- CODIGO NACIONAL ELECTORAL Ley N° 19.945 Argentina.gob.ar. (1972, 19 diciembre). Argentina.gob.ar. <https://www.argentina.gob.ar/normativa/nacional/ley-19945-19442/actualizacion>
- CODIGO ELECTORAL Ley N° 9.504/97 (1965, 15 julio) Brasil
<https://www.te.gob.mx/juriselectoral/OEA/content/codigo-electoral-de-brasil-y-leyes-complementarias>
- CONSTITUCIÓN DE LA REPÚBLICA DEL ECUADOR Art. 283 (2008, 9 abril) Ecuador
<https://www.igualdadgenero.gob.ec/wp-content/uploads/2020/02/Ley-Org%C3%A1nica-Electoral-C%C3%B3digo-de-la-Democracia.pdf>
- Córdoba, Ángel Polibio. (2019). *La fuerza de la opinión pública. Ecuador en perspectiva*.
<https://cedatos.com/wpcontent/uploads/2021/12/LibroLaFuerzaDeLaOpinionPublicaA-PC.pdf>
- Compilación, C. E. O. "Historia de las encuestas en el mundo." *La Sociología en sus Escenarios* 24 (2011).

- De Leeuw, E. D., Hox, J., & Dillman, D. (2012). *International handbook of survey methodology*. Routledge.
- De Vaus, D., & de Vaus, D. (2014). *Surveys in social research*. Routledge.
- Fielding, S. (1992). *What did 'the people' want?: the meaning of the 1945 general election*. *The Historical Journal*, 35(3), 623-639.
- Grasso, L. (2006). *Encuestas: elementos para su diseño y análisis*: (ed.). Editorial Brujas. <https://elibro.net/es/lc/usfq/titulos/77141>
- Verdesoto, L. Soldevilla, F. Tamayo. (2021). *A OPINIÓN PÚBLICA Y PRONÓSTICOS ELECTORALES: EL "TRAS CÁMARAS" DE LOS DATOS Y SU INCIDENCIA EN EL ELECTORADO*" Instituto de la Democracia.
- Kennedy, C., Blumenthal, M., Clement, S., Clinton, J. D., Durand, C., Franklin, C., ... & Wlezien, C. (2018). *An evaluation of the 2016 election polls in the United States*. *Public Opinion Quarterly*, 82(1), 1-33.
- Lewis-Beck, M. S. (2005). *Election forecasting: Principles and practice*. *The British Journal of Politics and International Relations*, 7(2), 145-164.
- Jowell, R., Hedges, B., Lynn, P., Farrant, G., & Heath, A. (1993). *Review: The 1992 British election: The failure of the polls*. *Public opinion quarterly*, 57(2), 238-263. <http://www.jstor.org/stable/2749470>
- Prosser, C., & 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), 757-790. doi:10.1017/gov.2018.7
- Ratto, M. C., Lewis-Beck, M. S., & Bélanger, É. (2022). *Forecasting elections in Latin America: An overview*. *Revista Latinoamericana de Opinión Pública*, 11(1), 5-13. <https://revistas.usal.es/cuatro/index.php/1852-9003/article/view/29171/28270>

- Rossi, P. H., Wright, J. D., & Anderson, A. B. (Eds.). (2013). *Handbook of survey research*. Academic press.
- Silver, N. (2020). *The signal and the noise: Why so many predictions fail-- but some don't*. Penguin Books.
- Smith, T. W. (1990). The First Straw?: A Study of the Origins of Election Polls. *The Public Opinion Quarterly*, 54(1), 21–36. <http://www.jstor.org/stable/2749389>
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*.
- Telles, H. D. S. (2008). Estado y representación política en Brasil. *Amérique latine histoire et mémoire*, 16. <https://doi.org/10.4000/alhim.3071>
- Zhou, Z., Serafino, M., Cohan, L., Caldarelli, G., & Makse, H. A. (2021). Why polls fail to predict elections. *Journal of Big Data*, 8(1), 1-28.

Anexos

Anexo 1, *Base de datos fallos de encuestas electorales (1)*, URL

https://docs.google.com/spreadsheets/d/1nv_AI_jwgB67t3_7KfuSFWOjorv0wX92/edit?usp=drive_web&ouid=116534448100967669801&rtpof=true

Anexo 2, Do file, *Proyecto integrador*, URL

https://drive.google.com/file/d/1LcPsLU7fZ0UW3xIXQzgUV_pElQk9dBAb/view?usp=drive_link

Anexo 3, Imagen, *Graph 1*, URL https://drive.google.com/file/d/1bl6XIqyfs3FX7TgT2rgz-sh93wRVBabs/view?usp=drive_link

Anexo 4, Imagen, *Graph 2*, URL

https://drive.google.com/file/d/12h4qczPsQMzgX_4SXOTf_SxOKd81_qjP/view?usp=sharing

Anexo 5, Imagen, *Graph 3*, URL

https://drive.google.com/file/d/1iYb7EnIdgLwS5SciTQT9spGHVksSz6y3_/view?usp=sharing

Anexo 6, Imagen, *Graph 4*, URL

https://drive.google.com/file/d/1GPqRPGvNy8nMKPDR4VZ_JBT2o_zQ0sUi/view?usp=sharing

Anexo 7, Imagen, *Graph 5*, URL

<https://drive.google.com/file/d/1IE6NN6GmkqfxKGClr1AnyRC9GRPmkIW-/view?usp=sharing>

Anexo 8, Imagen, *Graph 6*, URL <https://drive.google.com/file/d/16YbFH5xIv3riTYmZq-ySo0DD5Ss2n9VB/view?usp=sharing>