

PONENCIAS MESA 1

Manuel Barberena

PEARSON PARA PRI

Me permito hacer algunas reflexiones con respecto al proceso que vivimos todos nosotros en los últimos meses.

Es ya la segunda vez que la autoridad máxima del gobierno en México acepta el triunfo del candidato opositor, basándose en resultados de ejercicios probabilísticos. La primera de estas fue en 1997, reconociendo que había ganado Cárdenas en la ciudad de México y, obviamente, la segunda fue esta última elección presidencial.

No parece que haya vuelta de hoja en esto y que en México pasará lo mismo que en otros países del mundo, donde se confían encuestas de salida y conteos rápidos para determinar el ganador de una elección aún antes que se recopilen los datos de todos los votos. Esto, aunque no le guste a algunos miembros del PRI, el que el Presidente salga diciendo que ganó uno u otro, gracias a los estudios.

Los pronósticos de las encuestas se quedaron cortos contra algunos resultados finales, como el triunfo o el amplio triunfo de Fox y que el PAN vencería en la Cámara de Diputados o la victoria por tan pocos puntos de López Obrador en el DF.

Ahora empiezan a manejarse varias hipótesis. Algunos hablan del efecto Nicaragua que supone que la gente ocultó sus preferencias a los encuestadores bajo una espiral de silencio. Otras hipótesis aseguran que los indecisos pesaron más que nunca y definieron, finalmente, la elección. Una tercera hipótesis asegura que existe un voto fluctuante que pasó del PRI al PRD en 1994 y que en la elección de 2000 pasó al PAN. Por último, algunos analistas aseguran que el voto duro del PRI suponía que ganaría su candidato y dejó de participar. Todas estas aseveraciones son simplemente ideas.

En cualquier caso es previsible que se hagan ataques a las encuestas por haber fallado y para eso estamos nosotros aquí. Pienso que este tipo de foros son valiosísimos para estudiar qué fue lo que pudo haber sucedido y cómo hacer mejor las cosas.

Habría que evitar en el futuro usos malsanos de encuestas como, por ejemplo, suponer que el ganador de una elección sea el que tenga el liderazgo en más encuestas a lo largo de la campaña. Es decir, esas famosas publicaciones en las cuales se aseguraba que un candidato llevaba la delantera sólo porque más encuestas le daban el triunfo.

Otro error que se cometió con bastante frecuencia fue el calcular promedios de diferentes encuestas como si ésta se hubiese levantado con la misma metodología.

Un tema fundamental en el análisis de las diferentes encuestas es el del trato que se le dio a los llamados indecisos o personas que no contestaron. Creo que este es un tema en el que se deberá profundizar.

Si como algunas encuestas lo mostraban, la imagen de Labastida era mejor que la de Fox (tanto Reforma como Pearson publicaron algo al respecto basán-

dose en diferentes variables). Convendría repensar los atributos de las escalas que se están midiendo, quizá no sea necesariamente negativo calificar a alguien como terco o inflexible.

Bueno, por último ¿cuál es el balance? Yo creo que el balance todavía se está midiendo, aunque el hecho de que las encuestas se utilicen cada vez más para conocer las preferencias electorales ya es positivo.

Pasemos ahora a comentar acerca de la metodología utilizada por Pearson durante el pasado proceso electoral.

En el caso de Pearson, empezamos a participar desde el proceso preelectoral del PRI, en el cual, nuestras predicciones de los resultados de la elección primaria fueron de los más acertados. Una vez iniciada la carrera presidencial, en enero de 2000, se nos invitó a seguir participando en el proceso, trabajando para el Comité de campaña del Lic. Labastida. Es importante mencionar que desde que empezamos a trabajar para el PRI, ya se nos criticaba por ese simple hecho. En varios foros comenté que yo no estaba de acuerdo con los comentarios de varios periodistas que descalificaban nuestras encuestas por el simple hecho de ser financiadas por el PRI, sin tomar en cuenta la larga trayectoria de la empresa.

En nuestro caso, fuimos muy transparentes durante todo el proceso. La metodología aplicada en todas y cada una de las encuestas levantadas por Pearson fueron entregadas al IFE en los tiempos establecidos por la ley electoral. Asimismo, fuimos más que transparentes en la relación profesional que se tuvo con la agencia norteamericana Greenberg.

Es decir, Greenberg fue contratada por el PRI no solamente para estudios de opinión pública. Greenberg participó en muchísimas otras tareas, como fueron publicidad, asesoría en los debates, discurso político, etc.

Esta relación con Greenberg fue muy estrecha en términos metodológicos. Las decisiones de muestra, cuestionario, etc. se hacían por mutuo acuerdo, aunque Greenberg llevaría la batuta por ser la empresa que integraba toda la información.

Los cuestionarios eran bastante sólidos, metodológicamente hablando. Eran cuestionarios de aproximadamente 25-30 minutos que incluían una serie de variables. Varias de las preguntas que se incluían en el cuestionario no eran publicadas ya que eran para uso interno, en términos de lo que ya platicábamos, discurso, publicidad, etc. Obviamente, la pregunta de intención del voto siempre se incluía.

Los filtros eran los siguientes (esta información se encuentra en el documento ya presentado): personas con credencial de elector del estado que se estaba evaluando.

Se les aplicaba una escala para conocer la probabilidad de que el entrevistado fuese a votar. La escala era

la siguiente: era una escala de uno a 10, en la cual el 10 era segurísimo voy a votar o participar en la elección del 2 de julio y uno, seguramente no voy a votar. Y tomábamos para el análisis solamente a la gente de siete a 10, así como a aquellas que votaron en las elecciones federales de 1994 o que aún no tenían edad para votar en aquella época.

Es decir, aquí había todo un mito de que Pearson quitaba a los jóvenes. Esta aseveración era falsa, la gente que no tenía la edad en 1994 pasaba automáticamente el filtro y era considerada para todo el análisis

La unidad de muestreo que se utilizó en todos los levantamientos fue la sección electoral. Se definía una manzana semilla y usábamos la típica ruta de dar la vuelta a la manzana, hacíamos un salto sistemático de tres casas. Era una muestra por cuotas y siempre se buscaba a la persona para cubrir las cuotas.

Se tocaba la primera vivienda, si no se encontraba a la persona con las características que estábamos buscando íbamos a la siguiente vivienda y suplíamos a la persona con otra de las mismas características.

Se utilizaba la última versión del Listado Nominal del IFE disponible al público. El diseño muestral era el siguiente: del marco muestral se obtenían siete estratos. Esto se estudiaba desde antes junto con Greenberg.

Dividíamos al país en diferentes regiones: Pacífico, Norte, Golfo, Sureste, Sur, Centro y Metropolitano. En cada uno de estos estratos se dividía por tipo de secciones que había: urbana, mixta, rural.

Las muestras eran aproximadamente de mil 600 a mil 800.

A cada estrato, es decir, urbano, mixto, rural, se le asignaba una cantidad de entrevistas proporcional al número de electores que existían en el Listado Nominal.

La totalidad de las secciones electorales se organizaban con respecto a su número de electores de cada estrato y, posteriormente, se hacía una selección sistemática de sección totalmente aleatorio. Se obtenía un número de secciones necesario para cubrir las encuestas y se levantaban aproximadamente ocho a 10 entrevistas en cada sección.

Esto nos daba entre 180 a 200 secciones electorales que levantábamos en cada estudio.

Este método se utilizó en todas las encuestas que realizamos a partir de enero.

Después de realizar el proceso aleatorio se generaba una muestra en 30 entidades federativas (por razones de tiempo casi siempre quedaba fuera BCS). Una vez que la muestra estaba lista, aproximadamente el 79 por ciento de los casos quedaba en secciones urbanas y/o mixtas y el resto de los casos totalmente rurales.

Hay que recordar que sólo se tenían cuatro días para levantar el campo y a pesar de esta limitante, puedo afir-

mar que se hizo un excelente trabajo de campo. Teníamos a cerca de 100 gentes en campo y en cuatro o cinco días teníamos que estar entregando resultados.

Los resultados eran ligeramente ponderados en las siguientes variables: edad, sexo, educación e identificación partidista, con el fin de ajustar la muestra de la encuesta de forma que reflejara con mayor precisión el perfil de los votantes.

La empresa Greenberg utilizaba, con base en algunas de las preguntas, un ponderador, el cual no afectaba de manera fundamental los resultados. A diferencia de lo que se pensaba el ponderador Greenberg afectaba al PRI, beneficiaba al PRD y mantenía estable al PAN.

¿Por qué?

Se utilizaban los resultados de las encuestas de salida (exit polls) realizados en 1997, en términos de qué tipo de gente fue a votar. Es decir, sexo, edad y clase social y de cómo votó la gente ese año. Con base en estos resultados, Greenberg aplicaba su modelo, que vuelvo repetir, no afectaba de manera importante los resultados puros.

Las gráficas (con y sin indecisos) que se presentan al final muestran claramente como de enero a junio existió una importante caída de las preferencias por Labastida. Estoy seguro que la ventaja de más de 14 puntos que se reportaban entre enero y marzo eran reales debido a la influencia de la elección interna del PRI. Sin embargo, a partir del mes de abril vino el declive priísta.

Un tema que poco se ha mencionado en esta mesa, es la importancia que tuvo en esta elección la investigación cualitativa. Se realizaron decenas de grupos de enfoque en el ámbito cualitativo en toda la República, los cuales concordaban perfectamente con los resultados cuantitativos. Es decir, un Labastida con una cómoda ventaja entre enero y marzo. Sin embargo, la idea de cambio que el PRI quiso comunicar con su elección interna se fue diluyendo a través del tiempo y vemos cómo a partir del mes de abril, mayo, junio el PAN se apropia de esta idea de cambio vía una excelente campaña de comunicación y las preferencias electorales se empiezan a cerrar hasta llegar al último

levantamiento en la cual, sin indecisos, teníamos todavía a Labastida arriba dos puntos. (Eliminando indecisos).

El último dato publicado por Pearson fue la ventaja 45-38, 10 días antes de la elección. Todavía la última semana se levantó una encuesta nacional en la cual esta ventaja se redujo 43-41 a favor del PRI.

Estos resultados, inclusive, concordaban con los resultados publicados por Reforma, Milenio-Nielsen, UdeG. Es decir, la mayoría todavía daban entre tres y cinco puntos de diferencia a favor de Labastida.

En la presentación me permito incluir una gráfica la cual incluye un resumen de las cerca de 50 encuestas publicadas desde enero a junio. Lo que se puede observar es la gran disparidad de resultados muchas veces explicado por las diferentes metodologías y, en algunos casos, se puede deber a la inexistencia de muchas de las “empresas” encuestadoras.

Existe, asimismo, un tema al que se le ha restado importancia debido a la importancia de la elección presidencial. Me refiero a los casos de las elecciones del DF y Guanajuato. Estos dos casos son también dignos de análisis.

En el caso DF la gran mayoría de las encuestas, si no es que todas, le daban una amplia ventaja de 14 a 17 puntos al PRD unas semanas antes de la elección. Sin embargo, los resultados finales no dejan de sorprender. El PRD ganó por una ventaja de apenas 3-4 puntos.

El caso de Guanajuato también es digno de mencionarse, ya que la gran mayoría de las encuestas le daban una ventaja de entre 13 a 18 puntos al PAN y la ventaja real fue de cerca de 30 puntos.

Estos resultados sorprendivos no creo que sean porque se hayan equivocado las encuestas. Habrá que estudiar este fenómeno. Puede ser un fenómeno de última semana que nadie detectó. Tengo la impresión que no se puede dar una respuesta clara a estas interrogantes.

Lo importante es que como investigadores tenemos que aprender mucho de este proceso, para que en próximas elecciones seamos más precisos en nuestras predicciones. ■

1. FILTROS

- Personas con credencial de elector.
- Personas con mayor probabilidad de votar en las elecciones de julio próximo, según lo siguiente:
 - a) Aquellas que votaron en las elecciones federales de 1994, o que aún no tenían edad para votar en dichas elecciones pero hoy sí están en posibilidades de votar y
 - b) Aquellas que en una escala de probabilidad de votar del 1 al 10, donde 1 es “Poco probable que vaya a votar” y 10 “Segurísimo que vaya a votar”, contestan del 7 al 10.

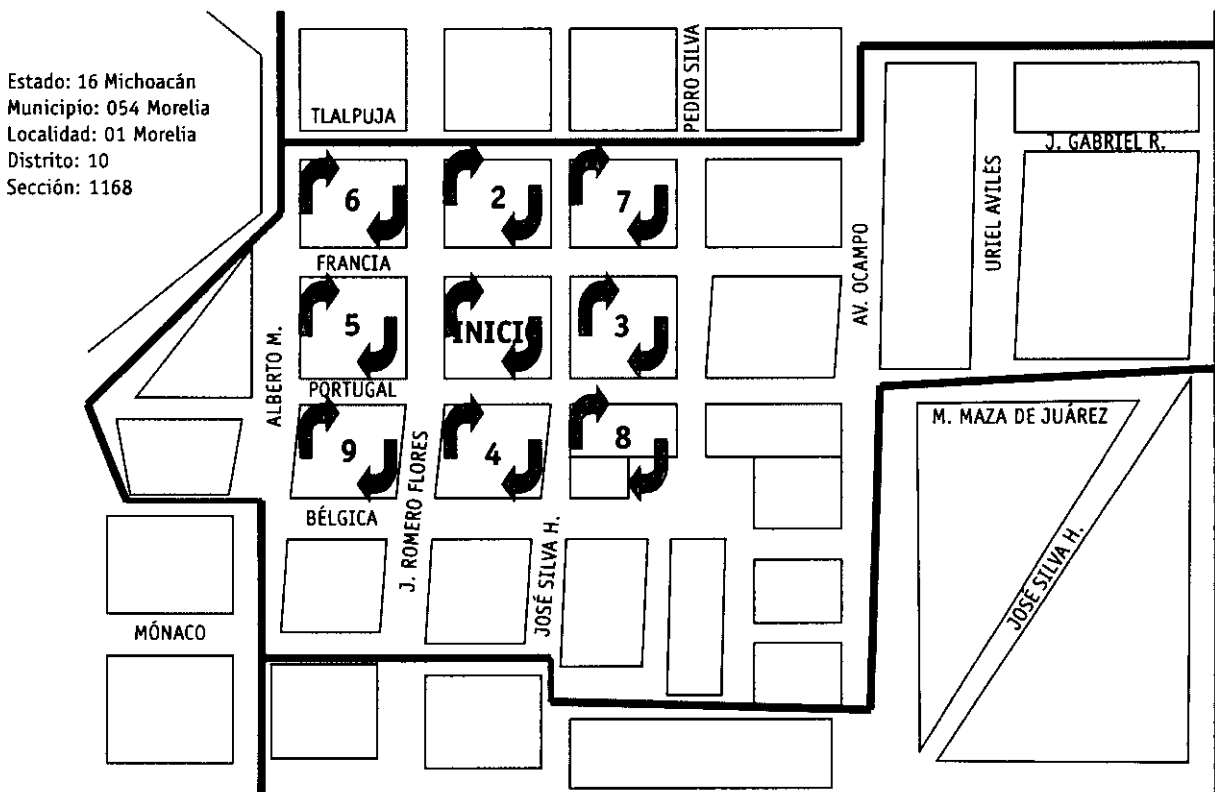
2. Método para recopilar información:

- En el mapa de la sección electoral seleccionada se designa una manzana como la “manzana semilla” (aquella donde se inicia el recorrido de campo), la que usualmente está en el centro de la sección

electoral. El recorrido se hace de izquierda a derecha empezando por la esquina superior izquierda de la “manzana semilla” y caminando en espiral hacia las siguientes manzanas, en caso de que no se hubiesen completado las entrevistas que se necesitaban.

- La primer vivienda que se toca es la que se encuentra donde se inicia el recorrido. Después se hace un salto sistemático de tres viviendas cada vez que se obtiene una entrevista efectiva.
- Para seleccionar el sujeto a entrevistar, debido a que se cuenta con cuotas por edad y sexo, se pide a quien abre la puerta diga si en la vivienda vive una persona que estuviera dentro de las cuotas asignadas, para posteriormente corroborar que cumple con los requisitos previamente fijados. Sólo se entrevista a un adulto que hubiera pasado el filtro por cada vivienda.
- Todas las entrevistas se hacen cara a cara, en el domicilio de los entrevistados.

MAPA DE SECCIÓN ELECTORAL



METODOLOGÍA

3. Método de muestreo.

- Marco Muestral.
 - Se utiliza la última versión del Listado Nominal del IFE disponible al público.
- Diseño Muestral.
 - Del marco muestral se obtienen siete estratos, los cuales corresponden a las regiones: pacífico, norte, golfo, sureste, sur, centro, y metropolitana. En una segunda etapa, cada uno de estos estratos se divide por tipo de sección: urbana, mixta y rural, de acuerdo con los criterios del IFE.
 - A cada estrato obtenido se le asigna una cantidad de entrevistas proporcional al número de electores que existen en el listado nominal.
 - La totalidad de las secciones electorales se organizan con respecto a su número de electores, de mayor a menor dentro de cada estrato. Posteriormente, se hace una selección sistemática de secciones, la que inicia en una sección escogida aleatoriamente. Se obtiene el número de secciones necesarias para cubrir el total de entrevistas de la muestra tomando como base 8 por cada sección.
 - Después de realizar el proceso aleatorio se genera una muestra repartida en 30 entidades federativas* diferentes.
 - La muestra se distribuye en secciones urbanas/mixtas y rurales, de tal manera que se conserva la misma proporción existente en el Listado Nominal, que cuenta con aproximadamente un 79 por ciento de casos urbanos/mixtos, y el resto de casos rurales.

1 AGUASCALIENTES
2 BAJA CALIFORNIA SUR
3 CAMPECHE
4 COAHUILA
5 COLIMA
6 CHIAPAS
7 CHIHUAHUA
8 DISTRITO FEDERAL
9 DURANGO
10 GUANAJUATO

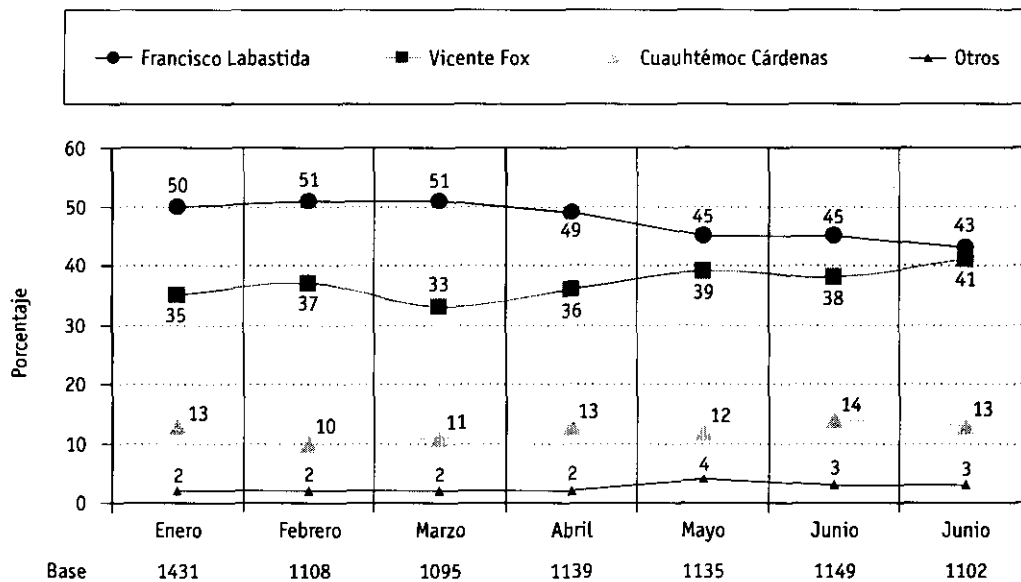
11 GUERRERO
12 HIDALGO
13 JALISCO
14 MÉXICO
15 MICHOACÁN
16 MORELOS
17 NAYARIT
18 NUEVO LEÓN
19 OAXACA
20 PUEBLA
21 QUERÉTARO
22 SAN LUIS POTOSÍ
23 SINALOA
24 SONORA
25 TABASCO
26 TAMAULIPAS
27 TLAXCALA
28 VERACRUZ
29 YUCATÁN
30 ZACATECAS

4. RESULTADOS

Los datos brutos arrojados por las encuestas eran ponderados de acuerdo con la siguiente metodología:

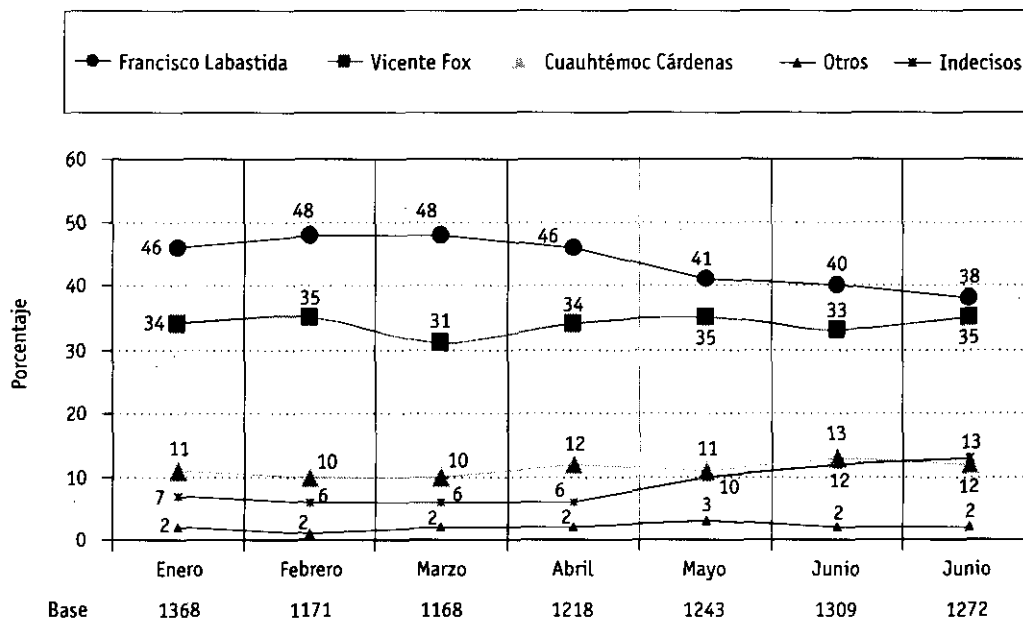
- Cinco factores eran ligeramente ponderados: región (siete regiones); edad, sexo, educación e identificación partidista, con el fin de ajustar la muestra de la encuesta de forma que reflejara con mayor precisión el perfil de los votantes que con mayor probabilidad participarían en la elección del 2 de julio.
- El perfil demográfico y partidista de los votantes probables, se basó en un análisis de los resultados electorales más recientes; en los datos de encuestas de salida de las elecciones constitucionales de 1994 y 1997, y en los datos del Conteo General de Vivienda 1995 (Greenberg).
- La ponderación realizada era mínima ya que el ajuste promedio era 1.2 puntos porcentuales en promedio en la mayoría de nuestras encuestas.
- Para cada levantamiento se utilizaron muestras nuevas e independientes.
- La pregunta de voto se realizó utilizando boleta electoral simulada.

EN ESTA HOJA ESTÁN LOS CANDIDATOS PARA LA ELECCIÓN PRESIDENCIAL DE JULIO PRÓXIMO.
 ¿SI LA ELECCIÓN PRESIDENCIAL FUERA HOY, USTED VOTARÍA POR?
 (Probables votantes y eliminando a los indecisos)



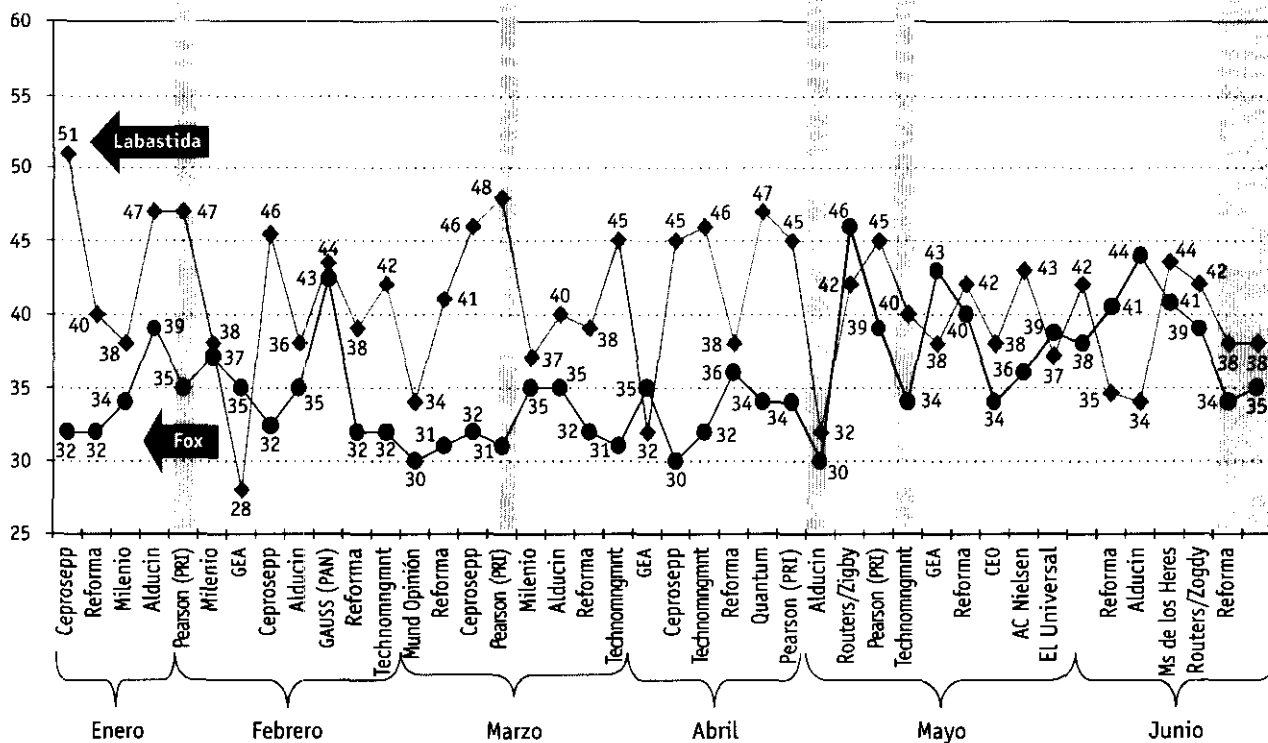
Datos sin ponderar.

EN ESTA HOJA ESTÁN LOS CANDIDATOS PARA LA ELECCIÓN PRESIDENCIAL DE JULIO PRÓXIMO.
 ¿SI LA ELECCIÓN PRESIDENCIAL FUERA HOY, USTED VOTARÍA POR?
 (Probables votantes)



Datos sin ponderar.

HISTÓRICO DE ENCUESTAS PUBLICADAS
Resultados: intención de voto para la presidencia



* Encuesta no publicada pero mencionada por el candidato del PRI.
** Encuesta ajustada a la base una vez eliminados los indecisos.