

MESA 1

.....

ENCUESTAS PREELECTORALES

Serie incluyendo últimas encuestas

⋮

RELATORÍA

Ana Cristina Covarrubias

PONENCIAS

Ricardo de La Peña

Alejandro Moreno

Manuel Barberena

Rolando Ocampo

Eduardo Ragasol

Ana Cristina Covarrubias

RELATORÍA

.....
Ana Cristina Covarrubias

Esta relatoría presenta en forma sintética lo expresado por cada ponente de la Mesa 1, y la última ponencia incluye, además de las principales conclusiones metodológicas derivadas de estas ponencias, las de los comentaristas participantes en el debate que se realizó al final de la sesión.

Un objetivo central del Taller de Sumiya, fue el de analizar y discutir franca y abiertamente, en un clima de diálogo y confianza, qué fue lo que pasó con las encuestas preelectorales ya que, retomando las palabras del maestro Woldenberg, “....en el trayecto previo al 2 de julio nos dieron un retrato grueso, a brochazos de la competencia electoral... dejaron de ser instrumentos de confianza... y al contrario, muchas encuestas pasaron a formar parte del conflicto político e incluso se degradaron para ser un mero instrumento de propaganda...”

A la mesa, objeto de la presente relatoría, fueron invitados seis expositores que tenían en común el haber publicado sus estimaciones de las elecciones presidenciales de 2000 y el haberlo hecho con una periodicidad por lo menos mensual. Todas eran agencias reconocidas por su experiencia en el campo de la investigación de mercados y/o de la opinión pública. Este perfil de expositores marcó a esta mesa, como una de las más relevantes para cumplir con el objetivo central del taller.

Resumen de ponencias

.....
Ricardo de la Peña

ISA-GEA

El autor inicia su presentación aclarando que el programa de encuestas que llevó a cabo estaba inserto dentro de un proyecto de construcción de escenarios político electorales que tenía por tarea seguir la evolución política del país en el mediano plazo. En total realizó 12 levantamientos que se iniciaron desde junio de 1998. En el año 2000 llevó a cabo cinco.

En cuanto a los aspectos metodológicos que rigieron a sus encuestas, el autor puntualizó lo siguiente:

- 1) Definió como población objetivo a los mexicanos nacidos antes de julio de 1982, residentes en el territorio nacional, con credencial de elector.

- 2) Para estimar el segmento de probables votantes utilizó la siguiente pregunta: ¿Qué tan factible es que vaya a ir a votar en las próximas elecciones federales? La respuesta se registró en una escala de cuatro puntos: Seguro sí, Probable sí, Probable no, Seguro no. Se consideraron votantes probables a los que declararon ir a votar con seguridad. En las encuestas realizadas entre febrero y junio de 2000, la proporción de votantes que respondió seguro que sí se elevó de 58 por ciento a 75 por ciento.
- 3) En sus distintas evaluaciones utilizó siempre muestras independientes. Aclaró que no empleó la alternativa de panel por los problemas que conlleva de mortalidad y sensibilización.
- 4) El tamaño de muestra para cada levantamiento (a excepción de la encuesta final en donde empleó dos mil 400 casos) fue de mil 200 casos distribuidos con igual proporción en los 30 distritos electorales que previamente seleccionó con probabilidad proporcional al tamaño (PPT). Dentro de cada distrito electoral seleccionó ocho secciones mediante muestreo aleatorio simple. Las unidades primarias de muestreo (distritos electorales) fueron las mismas en todas las muestras a excepción de la última encuesta en la que se duplicó el número de distritos.

El autor aclaró que ante diversos cuestionamientos provenientes de los propios clientes del proyecto, en el sentido de que podría haber un sesgo en las estimaciones debido a la específica selección de distritos en muestra, en el mes de marzo llevó a cabo otra encuesta en la que siguió los mismos procedimientos de selección, pero cambiando los distritos. El resultado fue que no encontró diferencias significativas entre la muestra original y la alterna en las estimaciones de voto por candidato.

También señaló que ante la posibilidad de que sus encuestas tuvieran un sesgo debido al orden de presentación de la pregunta de intención de voto, realizó un experimento consistente en aplicar en muestras paralelas y bajo condiciones equivalentes, dos cuestionarios diferentes en cuanto al orden de presentación de la citada pregunta. Uno tenía la pregunta al principio y el otro casi al final. Los resultados demostraron que sí hay un efecto asociado al distinto orden de presentación de la pregunta y este efecto es que la colocación al principio del cuestionario, eleva la proporción de indefinidos. Sin embargo, ello no produce cambios significativos en las estimaciones de preferencias entre candidatos.

En la segunda parte de su exposición el autor presentó su serie histórica en la que dio cuenta de que en lo tocante a las evaluaciones del año 2000, él encontró a Vicente Fox Quesada (VFQ) aventajando a Francisco Labastida Ochoa

(FLO) desde el mes de febrero en intención manifiesta de voto por candidato. La tendencia de VFQ fue descendente hasta el punto de que se invirtió en su primera evaluación del mes de junio y finalmente en su segunda evaluación de junio, detectó un repunte de VFQ con el cual quedó con un punto porcentual por encima de FLO.

El autor terminó remarcando la importancia del registro de la variable de “probables votantes” porque en su opinión el decantar la muestra en función de la seguridad de ir a votar, sí mejora las estimaciones de las proporciones de intención de voto. Deja como reto para el futuro de la investigación el mejorar la forma de “medir” dicha variable.

.....

Alejandro Moreno

DIARIO REFORMA

La ponencia de Alejandro Moreno se inició con la presentación de la vitrina metodológica que fue publicada por el diario Reforma acompañando a las siete encuestas realizadas de enero a junio del 2000. Siendo los datos principales de dicha vitrina metodológica, los siguientes:

Marco muestral: Lista de secciones electorales estratificadas por tipo de sección.

Tamaño de muestra: Entre mil 500 y mil 650 casos y error teórico de estimación del 2.5 por ciento al 95 por ciento de confianza.

Población objetivo: Mexicanos de 18 años o más de las 32 entidades federativas del país.

Procedimiento de muestreo: Probabilístico con selección en tres etapas que incluyen: a) puntos de levantamiento, b) selección de viviendas y c) selección de las personas entrevistadas.

Técnica de recolección de información: Entrevista cara a cara en vivienda.

Pregunta sobre intención de voto: Utiliza urna con boleta que contiene nombres de candidatos y logos de partidos y alianzas.

El autor hizo especial hincapié en el tema de control y supervisión del operativo de campo, puntualizando que se realizó en tres etapas complementarias:

- a) Supervisión directa durante la aplicación de cuestionarios en un recorrido aleatorio, en el que se contactan entre 25 y 30 por ciento de los encuestadores.

- b) Supervisión después de la aplicación de cuestionarios, regresando al 20 y hasta el 25 por ciento de los puntos elegidos al azar.
- c) Supervisión estadística por medio del análisis de resultados obtenidos, tomando como variable de control el número de encuestador.

Asimismo, el autor señaló que en sus publicaciones presentó las estimaciones de la intención de voto en porcentajes ajustados, esto es eliminando la “no respuesta” pero que siempre dio cuenta de la proporción de “no respuesta” hallada en cada evaluación.

En la segunda parte de su presentación, el autor discutió, analizó y presentó evidencias que apoyan la conveniencia de usar el indicador de “probables votantes” para mejorar la calidad de las estimaciones de las proporciones de voto. Con el fin de llegar a definir el indicador de “probables votantes” el autor utilizó un modelo de regresión LOGIT donde la variable dependiente es el probable votante y las independientes son múltiples de carácter político, geográfico y demográfico. Fueron tres las variables que el autor encontró como las que mejor permiten definir a los “probables votantes”: “interés en la campaña”, “probabilidad de ir a votar” y “probabilidad de éxito del candidato preferido”.

Para terminar, el autor presentó un cuadro en el que compara las cifras que publicó en junio, las que resultan de decantar la muestra en función del indicador de “probables votantes” y las de los resultados oficiales. El cuadro ilustra cómo las proporciones derivadas de la aplicación del indicador de “probables votantes” mejoran considerablemente en precisión. Sin decantar, el error promedio para los tres principales candidatos y resto es de 2.5 por ciento y con el uso del indicador se reduce a uno por ciento.

.....

Manuel Barberena

PEARSON/GREENBERG

nes federales de 1994, o que aún no tenían edad para votar en dichas elecciones, pero hoy sí están en posibilidad de votar; y el otro, el de personas que en una escala de 10 puntos sobre la seguridad de ir a votar, contestaron del siete al 10.

Marco de muestreo: listado de secciones electorales del IFE, estratificado, regiones y tipo de sección electoral.

Pregunta sobre intención de voto: Se utilizó boleta electoral simulada.

Técnica de recolección de información: Entrevista cara a cara en vivienda.

Tamaño de muestra: mil 700 entrevistas.

Selección de la muestra: Se consideraron 30 entidades federativas excluyendo Quintana Roo y Baja California. La totalidad de las secciones electorales se organizaron con respecto a su número de electores, de mayor a menor, al interior de cada estrato. Posteriormente, se hizo una selección sistemática de secciones. Se calculó el número de secciones necesarias para cubrir el total de entrevistas a razón de ocho por cada sección. En promedio se visitaron 200 secciones electorales por encuesta. Ubicados en la vivienda el muestreo fue por cuotas de edad y sexo.

a) Ponderación. Los datos brutos arrojados por las encuestas eran ponderados en función de los siguientes factores: región, edad, sexo, educación e identificación partidista. Los datos base provenían de las encuestas de salida de las elecciones federales de 1994 y 1997 y de los datos del Censo General de Población y Vivienda 1995 (INEGI). El autor puntualizó que los efectos de ponderación fueron mínimos, ya que en promedio la diferencia entre las estimaciones de voto derivadas de datos crudos y ponderados fue de aproximadamente 1.2 por ciento.

El autor mostró un gráfico con las proporciones de voto, decantadas por el indicador de “probables votantes”, derivadas de las siete encuestas que se realizaron de enero a junio del 2000. Los resultados muestran que en todo momento Labastida aventajó a Fox y nunca se cruzaron las tendencias. Asimismo, el gráfico muestra que la tendencia de Labastida era descendente (de 48 a 38 por ciento de febrero a junio) en tanto que la de Fox se mantenía más o menos constante.

En primer lugar se presentaron los aspectos metodológicos que caracterizaron a su serie de encuestas.

Selección de entrevistados: Se utilizó como filtro la credencial de elector.

Definición de “probables votantes”: Se utilizaron dos criterios. Uno, el de personas que votaron en las elecciones

.....
Rolando Ocampo

GAUSS

El autor inició su presentación mostrando las estimaciones de voto de seis encuestas realizadas de enero a junio del 2000. El análisis de su gráfica muestra un marcado crecimiento en las preferencias por Vicente Fox; en enero estaba cinco puntos debajo de Labastida y para abril ya tenía ventaja que mantuvo hasta junio. En las estimaciones del voto en la última encuesta se tuvo un crecimiento de los indecisos, aunado a una ligera baja de Fox.

En la segunda parte de su ponencia, el autor presentó una serie de indicadores que explican las tendencias detectadas en su serie histórica.

Con base en la pregunta sobre la probabilidad de ganar que tenían los diferentes candidatos, se concluyó que Vicente Fox fue convenciendo gradualmente al electorado de tener serias posibilidades de ganar, lo que pudo haberle valido muchos votos. Aunque las probabilidades de Labastida eran más altas al principio, al final fueron descendiendo. También se investigó la percepción de los electores con relación a si comenzaban a ver más como presidente a uno u otro candidato y allí se observó un escenario cada vez más favorable para Fox y menos para Labastida. Aunado a estas percepciones, también se comenzó a ver a un presidente proveniente del PRI como "más de lo mismo" y en consecuencia se fue incrementando el deseo de que perdiera el PRI a fin de sentar las bases para el cambio.

Finalmente, el autor presentó un análisis del perfil de los votantes por edad, sexo y tipo de sección en el cual dio cuenta de los avances que Fox fue registrando a lo largo de la campaña entre las mujeres, los electores de 35 a 44 años y las áreas rurales, avances que le dieron el triunfo.

.....
Eduardo Ragasol

AC NIELSEN-MILENIO

En la primera parte de su ponencia el autor presentó los aspectos metodológicos más relevantes de las dos encuestas que levantó para el diario Milenio en los meses de mayo y junio de 2000, los cuales son:

El tamaño de muestra que utilizó fue de mil 800 casos en la primera evaluación y dos mil 793 en la segunda y última.

El diseño de muestreo es probabilístico, no autoponderado y estratificado con selección de conglomerados en varias etapas con probabilidad proporcional al tamaño. Las etapas consideradas son a) localidades, b) AGEB's, c) manzanas, d) viviendas y e) ciudadanos. Para efectos de construcción de las estimaciones de ambas encuestas se agruparon los 32 estados en nueve regiones y las localidades del país en cuatro estratos a) localidades con menos de dos mil 500 habitantes (rurales), b) localidades con población entre dos mil 501 y 15 mil habitantes, c) localidades con población entre 15 mil uno y 100 mil habitantes y d) localidades con más de 100 mil habitantes. Dado el diseño no autoponderado, se utilizaron factores de expansión para construir las estimaciones.

La selección del respondiente no fue por cuotas sino en función del criterio del miembro del hogar que tuviera la fecha de cumpleaños más cercana al día del levantamiento de la encuesta. Si el sujeto seleccionado no se encontraba en el hogar se regresaba posteriormente, haciendo hasta dos visitas.

El respondiente debía tener 18 años o más y contar con credencial de elector.

Se utilizaron dos indicadores para la estimación de "probables votantes": Probabilidad de ir a votar, medida en una escala de cinco puntos (totalmente, bastante, medio, poco y nada) e interés en participar, empleando una escala de cuatro puntos. Para los fines de la estimación de "probables votantes" el criterio usado fue estricto, pues se tomaron a los sujetos que dieron la calificación máxima en ambas escalas.

La pregunta sobre preferencia electoral se realizó por medio de urna y boleta y se aplicó inmediatamente después de los filtros y el registro de los indicadores de "votantes probables".

En cuanto a los resultados hallados, AC Nielsen reportó ventaja de Labastida sobre Fox de siete puntos en el mes de mayo y en la evaluación de junio encontró a Labastida aventajando a Fox por una diferencia de tres puntos porcentuales con lo cual, acerca de esta última encuesta, concluyó que "...el ejercicio no arrojó evidencia significativa sobre quién estaba en primer lugar en la preferencia electoral"

La segunda parte de su ponencia la dedicó el autor a presentar una serie de propuestas para los ejercicios futuros, propuestas que se pueden sintetizar así:

- Buscar indicadores que permitan realmente estimar a los probables votantes y publicar resultados con población "filtrada" y "sin filtrar".

- Especificar los filtros y los indicadores empleados.
- Buscar elementos en los cuestionarios y en el trabajo de campo para reducir la no respuesta.
- Publicar la distribución del voto incluyendo la no respuesta.
- Especificar si se hizo y cómo se hizo una asignación de la no respuesta.
- Utilizar muestreo probabilístico. El tener control sobre la probabilidad de selección debe prevalecer sobre cualquier otro criterio.
- Especificar a qué variable corresponde el error publicado. Se deben publicar los estimadores, no nada más puntuales, sino interválicos.
- Especificar si el método para calcular el error muestral toma en cuenta el diseño de muestreo o supone un muestreo aleatorio simple.
- Acordar previamente con el cliente una supervisión cercana de la publicación de cuadros, gráficas, notas metodológicas, pies de cuadro, etc.
- Procurar siempre una transmisión correcta de los resultados en cualquier foro.
- Distinguir cuándo una afirmación es el resultado duro de una encuesta y cuándo es una interpretación.

CONCLUSIONES METODOLÓGICAS

1) Estimación de probables votantes: uso de "filtros"

En primer lugar es necesario precisar que es un error hablar genéricamente de "filtros" de probables votantes cuando en la realidad no se filtran a los sujetos contactados más que por tenencia de credencial de elector. En estricto sentido un filtro es un reactivo utilizado para determinar si una persona contactada para entrevista entra o no en muestra; su función es cribar, en el momento de iniciar la entrevista, a los entrevistados que no cumplen los requisitos preestablecidos para definir a un votante probable. De la aplicación del filtro se deriva la decisión de suspender la entrevista o continuarla.

En la realidad nadie emplea tal procedimiento por lo costoso que es y porque elimina la posibilidad de hacer análisis posteriores que permitan establecer con más claridad el segmento de probables votantes. Lo que en la práctica se hace es usar indicadores de control, esto es reactivos que sirven para analizar los resultados de votación considerando varios escenarios de participación ciudadana. Estos escenarios corresponden a segmentaciones de la población definidos por los diferentes patrones de respuesta en los reactivos de control.

Con relación a las encuestas preelectorales encontramos que se utilizaron reactivos conceptualmente diferentes:

probabilidad de ir a votar, seguridad de ir a votar, interés en las elecciones, información de las campañas y voto en elecciones previas.

Los reactivos se construyeron en preguntas dicotómicas y escalas diferentes, unas ordinales y otras de intervalo, y además las escalas eran de distintos tamaños, de tres, cuatro, cinco, siete y 10 puntos. Además, encontramos también la utilización de diferentes niveles de rigor en el establecimiento de las líneas de corte cuando se usan escalas de más de dos puntos. Para ilustrar este último punto se halló que cuando se usaron escalas de 10 puntos, hay quien estableció la línea de corte para definir a un probable votante en el punto nueve (alto rigor) y hay quien la estableció en el punto siete (menor rigor).

Las diferencias antes reportadas y particularmente las de rigor, condujeron a conclusiones encontradas ya que algunos investigadores dejaron asentado que los indicadores de probables votantes no les habían modificado las estimaciones originales y por el contrario otros hallaron que estos indicadores constituyen una herramienta básica para mejorar la calidad de las estimaciones sobre la intención de voto.

Finalmente, varios de los investigadores coincidieron en afirmar que el registro de indicadores para definir a los probables votantes sí es importante para mejorar significativamente la calidad de las estimaciones. Sin embargo, hoy por hoy no hay un consenso en cuanto a cuál o cuáles indicadores son los que mejor estiman la probabilidad de ir a votar. En consecuencia se requiere hacer más esfuerzos de investigación en este sentido.

2) La pregunta sobre intención de voto

La variable fundamental de una encuesta electoral es la pregunta sobre la intención de voto y al respecto son sorprendentes las diferencias que se encuentran entre las formulaciones de las diferentes agencias investigadoras, las cuales no sólo son de orden de colocación de la pregunta sino de la forma de frasearla e instrumentarla.

En cuanto al orden de presentación de la pregunta hay quien la coloca al principio, hay quien lo hace en medio y hay quien la formula al final de la entrevista. Hay un mayor consenso a ubicar la pregunta al principio de la entrevista, pero después de haber hecho cuatro o cinco preguntas que no tengan nada que ver con preferencias electorales y que sí ubiquen al entrevistado en el tema de las elecciones. Es cuando generalmente se introducen los indicadores sobre "probables votantes".

Sobre el fraseo de la pregunta se halló que hay quien pregunta sólo sobre candidatos y hay quien lo hace por partidos y candidatos, pero las diferencias más ostensibles están en la forma de instrumentar la pregunta. Al respecto de esto hallamos que hay quien al hacer la pregunta

sólo nombra a los partidos y candidatos, otros emplean una tarjeta de ayuda con los nombres impresos, otros más utilizan una réplica de boleta en la que aparecen los nombres y logos de los partidos y alianzas, así como de los candidatos; dentro de este grupo de investigadores hay quien presenta la réplica de boleta en blanco y negro y hay quien lo hace a todo color. Por si fueran pocas las diferencias hay quien le pide al entrevistado que simplemente escoja de la boleta presentada, a su candidato/partido, y hay quien le pide al respondiente marcar su boleta en privado y depositarla en la urna que porta el entrevistador.

Al respecto de esta variedad, en la forma de hacer la pregunta clave de la encuesta hay varios investigadores que estuvieron de acuerdo en cuanto al uso de la réplica de boleta a todo color junto con la urna, en tanto que con estos elementos se reproduce lo más cercanamente posible la situación real de votar y además se respeta la confidencialidad del voto.

3) Tamaño de muestra y procedimiento de muestreo

En esta materia se observó bastante homogeneidad en los tamaños de muestra empleados (mil 500 a dos mil casos, excepcionalmente más de dos mil casos) y los procedimientos de muestreo empleados por las diferentes empresas (muestras probabilísticas, polietápicas haciendo uso de estratos y conglomerados con selección aleatoria simple o sistemática de secciones electorales).

Sólo se detectaron algunas diferencias como:

En marco muestral, dos agencias de las seis participantes continúan empleando las áreas geoestadísticas básicas del INEGI. Las otras cuatro utilizaron las estadísticas de la lista nominal del IFE.

En procedimiento de muestreo hubo algunas excepciones a la regla:

- a) La primera es que una de cinco agencias no empleó un muestreo probabilístico porque utilizó como técnica de recolección de información la entrevista por intercepción en la vía pública. Esta decisión por un muestreo por intercepción se basó en el criterio de preferir perder el control sobre la probabilidad de selección de los entrevistados, a caer en el posible sesgo de que los entrevistados mientan debido al temor que les produce contestar una encuesta electoral en la puerta de su casa. Según este enfoque el miedo se minimiza cuando se responde en la vía pública, lejos de la propia vivienda.
- b) Otra consistió en que una de las cinco agencias dejó fija la selección de los 30 distritos de su primera muestra en las diferentes mediciones de su serie histórica, con lo cual sus muestras fueron parcialmente independientes. Esto es iguales en cuanto a distritos e

independientes en la selección de secciones, viviendas y sujetos.

- c) La última excepción consistió en que una de las agencias reportó haber hecho selección probabilística, no por cuotas, del entrevistado al interior de la vivienda.

En cuanto a la metodología del muestreo quedó establecida la conveniencia de utilizar el marco de secciones electorales del IFE cuya información permite el cálculo preciso de los ponderadores o factores de expansión atados al diseño muestral. Se observó también una clara preferencia por los muestreos probabilísticos polietápicas con entrevistas cara a cara en la vivienda, que hacen posible el cálculo de la ponderación correspondiente a cada entrevistado, permitiendo una construcción sólida de estimadores, así como de sus respectivos márgenes de error.

4) Uso de ponderadores

En general se utilizaron factores de expansión de dos tipos: intrínsecos al diseño de muestra (estadísticas de la lista nominal relevantes al diseño muestral como pueden ser: tipo de sección, edad, sexo y otros) y extrínsecos al diseño de muestra (escolaridad, preferencia partidaria y otras). Para ello se utilizaron al menos cuatro fuentes de información: las estadísticas de la lista nominal del IFE, las del censo de población del INEGI, los resultados electorales de elecciones pasadas y las encuestas de salida de elecciones pasadas. Cada investigador hizo su propia mezcla y no se especificaron los ingredientes de la misma. La pregunta que estuvo en el aire fue ¿se deben o no emplear ponderadores? Y en su caso ¿cuáles y de cuáles fuentes?

En la mesa no se llegó a precisar sobre estas materias, sin embargo es la opinión de la autora de la presente relatoría lo siguiente: Que sí se deben usar ponderadores porque la mayoría de las muestras no son estrictamente autoponderadas, aun cuando el diseño original así lo pretenda. En lo que toca a la mezcla de variables y fuentes lo crucial es lo segundo. En nuestra opinión la mejor fuente de información para el caso de factores de expansión intrínsecos al diseño de muestreo son las estadísticas de la lista nominal del IFE, porque están construidas para el propósito electoral y porque se actualizan con frecuencia. En lo que toca a factores de expansión extrínsecos al diseño de muestreo, como puede ser el caso de las preferencias electorales, la mejor fuente de información es la de los resultados definitivos de elecciones previas, del propio IFE, pero si el factor de expansión es la escolaridad, la mejor fuente de información es la encuesta de salida de la elección anterior y no el censo general de población, porque la distribución de escolaridad de los votantes es diferente a la distribución de la población en general.

5) La técnica de recolección de información

Todos los investigadores participantes, salvo el caso de Alduncin, emplean la técnica de entrevista cara a cara en vivienda. El fundamento de esta preferencia radica en la posibilidad de hacer muestreo probabilístico de viviendas, con base en los marcos muestrales del IFE y del INEGI, en el que todo ciudadano de 18 o más años tenga probabilidad positiva y conocida de pertenecer a la muestra.

Alduncin reportó haber utilizado en su primera encuesta entrevistas telefónicas y en el resto utilizó la técnica de entrevistas por intercepción en la calle. Prefirió no mantener el control del muestreo probabilístico en vivienda para no caer en el posible sesgo de obtener respuestas falsas inducidas por el temor a responder la verdad cuando se está en la puerta de la propia vivienda.

De lo anterior se deriva la conclusión de que la técnica preferida de recolección de información es la de entrevista cara a cara en vivienda. La alternativa de colecta probabilística de datos por medio de entrevistas telefónicas se rechazó por los posibles sesgos que genera la limitación de la disponibilidad de teléfono a nivel nacional.

6) El trabajo de campo y su supervisión

Este tema resulta importante de ser mencionado, precisamente por la falta de referencias al mismo que fue el común denominador de los presentadores. Sólo uno de los ponentes lo destacó y señaló el uso de tres diferentes métodos de control de campo: supervisión *in situ*, del 25 al 30 por ciento; supervisión *a posteriori*, con revisitas a hogares seleccionados al azar, del 20 al 25 por ciento; y supervisión estadística con base en el análisis de la información por clave de entrevistador.

Un tema relacionado con el trabajo de campo que afecta al diseño de muestra es el del registro de los rechazos en campo y la forma de sustitución de los casos que se niegan a contestar la entrevista. En general, los ponentes no dieron cuenta de sus experiencias en esta materia, por lo que no es posible establecer conclusiones sobre las prácticas más adecuadas.

7) Los modelos de tratamiento de la no respuesta

Ninguno de los ponentes utilizó un modelo especial para la asignación de indecisos; sencillamente hicieron asignaciones proporcionales a las cifras encontradas.

Sin embargo, sí se destacó la importancia de este tema y la necesidad de profundizar en la investigación de modelos que minimicen las tasas de error en la asignación de indecisos y consecuentemente aumenten la calidad de las estimaciones globales de voto.

Asimismo, se concluyó que cuando se publica una encuesta es importante dar cuenta de los datos crudos,

así como de los ajustados (asignación proporcional de la no respuesta), y que en el caso de usar algún modelo de asignación, se especifique y se publiquen tanto los datos crudos como los resultantes de la aplicación del modelo.

8) La relación con el cliente en el caso de publicaciones

Si el cliente quiere publicar una encuesta preelectoral es imperativo acordar, desde el principio, los términos en que se hará dicha publicación. En la publicación la agencia investigadora pone en riesgo su prestigio y por tanto debe asumir la responsabilidad y tener el control sobre los formatos de publicación y la interpretación de resultados.

9) Vitrinas metodológicas de las publicaciones

Todos los investigadores participantes publicaron sus encuestas incluyendo vitrinas metodológicas que contengan los siguientes datos: Cliente, fecha del levantamiento de la información, tamaño de la muestra, procedimiento de muestreo, técnica de recolección de información y error teórico de estimación.

Se mencionó como deseable que en el futuro la vitrina metodológica incluya otros aspectos: La forma en que se levantó la pregunta sobre la intención de voto, la duración del trabajo de campo, número de entrevistados y de supervisores empleados, procedimiento de control de campo, dispersión de la muestra y errores reales de estimación con intervalos de confianza, para cada uno de los partidos, especialmente para el primer y segundo lugar, incluyendo su diferencia.

10) Formatos de publicación

En lo referente a los formatos mismos de la publicación se habló de la conveniencia de acordar una homologación de la terminología en general.

Acordar una homologación de los grupos de edad y de las categorías de la variable educación.

Publicar siempre las estimaciones de voto crudas acompañadas de la pregunta con la que se exploró la intención de voto. Además, publicarlas en porcentajes ajustados (eliminando la no respuesta) a fin de poder compararlas fácilmente, pero siempre incluir la información sobre la proporción de no respuesta.

Publicar márgenes individuales de error para cada proporción alcanzada por un partido político, así como los intervalos de confianza, con el fin de transparentar el hecho de si la diferencia de puntos entre un candidato/partido y otro, es o no estadísticamente significativa.

En el caso de publicar estimaciones de voto derivadas de la aplicación de algún modelo, explicar que tal es el caso y siempre acompañar la publicación con los datos crudos, incluyendo la proporción de no respuesta. ■

PONENCIAS MESA 1

.....

Ricardo de la Peña

PRESIDENTE EJECUTIVO DE INVESTIGACIONES
SOCIALES APLICADAS, S.C.

SOBRE EL MÉTODO

Desde mediados de 1998 y hasta las elecciones federales de julio de 2000 operó un proyecto sindicado¹ para el seguimiento de las preferencias ciudadanas y la construcción de escenarios electorales y de gobernabilidad en México, que realizaron de manera conjunta Grupo de Economistas y Asociados (GEA) y la empresa Investigaciones Sociales Aplicadas, S.C. (ISA). Dicho proyecto consideró la realización de una serie de encuestas nacionales de opinión ciudadana específicamente orientadas a la determinación de las intenciones del voto del electorado. Este programa comprendió 12 mediciones entre junio de 1998 y junio de 2000, conforme al siguiente calendario:

ETAPA	ENCUESTA	LEVANTAMIENTO
1998-1999	Primera	6-9 de junio de 1998
	Segunda	24-27 de octubre de 1998
	Tercera	23-26 de enero de 1999
	Cuarta	22-25 de mayo de 1999
1999-2000	Quinta	19-22 de agosto de 1999
	Especial*	21-26 de octubre de 1999
	Sexta	25-28 de noviembre de 1999
	Séptima	3-6 de febrero de 2000
	Octava [†]	16-21 de marzo de 2000
	Novena	3-7 de mayo de 2000
	Décima	1-4 de junio de 2000
	Undécima*	15-18 de junio de 2000

* Por las características particulares de este estudio, sus resultados no resultan estrictamente seriales.

[†] Esta medición incluyó otras dos encuestas de contrastación: para el cuestionario y para la muestra.

* Esta encuesta consideró la duplicación de la muestra originalmente prevista para cada levantamiento.

OBJETIVOS DEL PROGRAMA

Los objetivos generales del modelo de escenarios políticos fueron: elaborar escenarios generales de la evolución política del país en el mediano plazo que permitiera identificar las principales áreas de conflictividad y estar en condiciones de determinar su aportación a la gobernabilidad y certidumbre; actualizar periódicamente los escenarios políticos, con base en el seguimiento y el análisis

¹ Que fuera patrocinado de manera conjunta por diversos partidos políticos (PRI, PAN, Convergencia por la Democracia, Democracia Social, PCD), así como por dependencias públicas (Gobierno del Distrito Federal, Secretaría de Gobernación, entre otras) y diversos representantes del sector privado nacional.

cualitativo de información y evidencia empírica producto de mediciones por encuesta; y elaborar prospectivas de corto plazo para identificar conflictos políticos y agendas de los principales actores.

Conforme a lo anterior, se definieron las encuestas nacionales de opinión ciudadana como un instrumento orientado no únicamente a la medición de intenciones de voto, sino a la atención de los siguientes aspectos:

- Conocer el estado de la opinión pública respecto a la situación económica y política nacional y respecto a acontecimientos coyunturalmente relevantes, la percepción sobre relaciones del Estado con los actores sociales fundamentales y puntos de vista en torno a los conflictos sociales y políticas más importantes, así como identificar las opiniones de la ciudadanía respecto a la gestión presidencial en general y en los principales ámbitos de acción gubernativa; las posiciones sobre los temas centrales en la agenda legislativa y la evaluación de la labor desarrollada por las fracciones parlamentarias; y los puntos de vista en torno a la seguridad pública y al aparato de justicia, así como sobre los principales problemas en este rubro.
- Detectar los medios y niveles de información de la ciudadanía respecto al proceso electoral federal del año 2000 y la presencia de partidos y candidatos entre el electorado y reconocer los elementos primordiales que conforman la imagen de los partidos políticos; determinar preocupaciones y demandas de la ciudadanía, así como sus posiciones en torno a los principales temas políticos y económicos que conformen la agenda de discusión pública entre los actores fundamentales.
- Medir las preferencias electorales tanto por partidos como por precandidatos y candidatos a la Presidencia de la República, determinar perfiles de los votantes por las diversas opciones políticas contendientes y detectar variaciones en las preferencias y factores que las propician.
- Precisar los niveles de confianza ciudadana en las instituciones electorales y en la limpieza de los comicios federales; la motivación a la participación activa en el proceso, así como las expectativas de conflictividad social y entre actores políticos fundamentales, de cara a las elecciones de 2000.

DISEÑO DEL PROGRAMA DE ENCUESTAS

Las decisiones básicas adoptadas respecto a las características de las mediciones a realizarse tomaron en consideración la pertinencia de disponer de secciones transversales sucesivas, referidas a un mismo universo de estudio, con muestras extraídas conforme a procedimientos homogéneos y aplicando de manera reiterada reactivos relevantes, todo ello para facilitar la comparabilidad de resultados y la construcción de series de tiempo referidas tanto para los agregados nacionales como para segmentaciones de interés.

El muestreo repetido a una misma población pudiera adoptar diversas características: entrevistas reiteradas a una misma población (tipo panel) o estudios de tendencia con muestras independientes. Para fines de este programa de observaciones, se decidió disponer de muestras independientes en cada sección que permitan el análisis de tendencias, pero reteniendo las unidades primarias de muestreo para maximizar la precisión al comparar entre las distintas encuestas sucesivas.²

La decisión de adoptar un diseño para el estudio de tendencias mediante secciones transversales sucesivas independientes se adoptó, asimismo, tomando en consideración la relativamente escasa relevancia que tendría disponer de estimaciones del cambio individual en respuestas a reactivos centrales para el estudio, recordando la existencia de evidencia de una sustancial inestabilidad en las respuestas individuales, lo que fortalece el postulado de que la opinión pública es un fenómeno colectivo, afectado por factores de información y comprensión y por la ocurrencia de sucesos modificadores, pudiendo detectarse tendencias comparativamente estables en los agregados.³

Con relación a la viabilidad de determinar factores causales de variaciones en las distribuciones al nivel de los agregados, al considerar una sucesión de diversas mediciones homogéneas en cuanto a sus características metodológicas, es factible disponer de observaciones seriadas que en conjunto conformen un diseño cuasiexperimental con validez interna y externa, referida respectivamente al diseño del instrumento y su operación y al método de selección de la muestra,⁴ donde la amplitud del espectro de cuestiones consideradas como asunto de atención favorece minimizar, sin suprimir, los riesgos de estable-

² La construcción de un panel nacional presentaba ventajas y riesgos. La principal ventaja es que posibilita un análisis de factores propiciadores de cambios en el ámbito individual, aumentando las posibilidades de intercorrelación de variables y de determinación de causalidad. Empero, dos factores de riesgo perjudican la efectividad de este diseño: la mortalidad inevitable que se presenta en la muestra (esperándose una tasa mínima próxima a 30 por ciento en el curso de dos años) y la incorporación de efectos de sensibilización en los sujetos en muestra, que afectaría las mediciones, disminuyendo la representatividad de la muestra respecto a la población de la que se extrajo, lo que tiene mayor importancia dada la relevancia de los niveles de información en los comportamientos electorales de la ciudadanía. La imposibilidad fáctica de eliminar o reducir de manera significativa estos factores de riesgo llevaron a eliminar esta opción.

³ Cfr.: Benjamin I. Page and Robert Y. Shapiro, *The Rational Public. Fifty Years of Trends in Americans' Policy Preferences*; The University of Chicago Press, Chicago, 1992.

⁴ Lo anterior, asumiendo los términos de Donald Campbell y Julian Stanley en: *Diseños experimentales y cuasiexperimentales en la investigación social*; Amorrortu Editores, Buenos Aires, 1973.

cer relaciones espurias que respondan a estímulos relevantes no considerados.

Asimismo, con miras a responder a los objetivos planteados para el programa de encuestas, dentro de la lógica del modelo de escenarios políticos 1998-2000, se diseñó un instrumento base de recuperación de información, que consideró una multiplicidad de reactivos sobre los diversos asuntos de interés para el proyecto. La disposición de un modelo básico de cuestionario para la serie de encuestas apoyó la comparabilidad de distribuciones de frecuencias que se observaron en diversas mediciones, permitiendo la construcción de series de tiempo que dieran cuenta de variaciones en las características, preferencias y opiniones de la ciudadanía.

En las diversas encuestas realizadas, se buscó que los cuestionarios guardaran similar orden e idéntica redacción y opciones de respuesta, evitando en lo posible efectos de diferencias en el diseño del instrumento que limiten la comparabilidad de las respuestas. Sin embargo, una de las críticas a los datos obtenidos en la serie fue precisamente el ordenamiento de las preguntas en el cuestionario, en particular la inclusión hacia el final de la entrevista de los reactivos correspondientes a la intención de voto de los entrevistados.

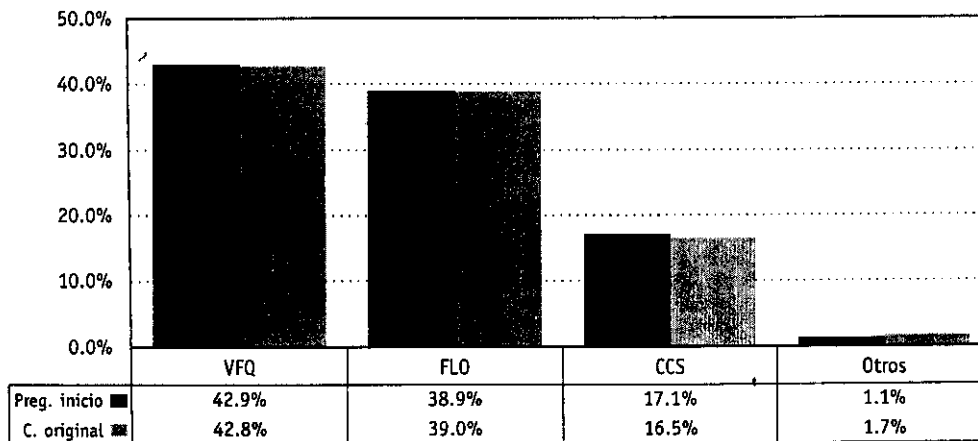
Por ello, con miras a determinar el efecto de la posición de la pregunta sobre intención de voto en el cuestionario, en paralelo a la aplicación de la octava encuesta de la serie, a mediados de marzo de 2000 se realizó otra encuesta con un cuestionario donde se puso al principio la pregunta sobre la intención de voto, aplicándose mil 200

entrevistas para esta versión alterna del cuestionario. Y para evitar la interferencia de otros factores en el experimento, las entrevistas fueron aplicadas los mismos días, por los mismos operadores, intercalando la aplicación de los cuestionarios en las mismas áreas de entrevista. Los resultados obtenidos constataron que la colocación de la pregunta sobre intención de voto impacta los niveles de definición que se obtengan: su colocación al principio del cuestionario eleva la proporción de indefinidos. Sin embargo, ello no produce cambios significativos en las estimaciones de preferencias entre candidatos. El giro medio entre ambas estimaciones es apenas de 0.7 por ciento.

Respecto al trabajo de campo, para el desarrollo de este proyecto se dispuso de un grupo relativamente estable, constituido por personal que forma parte regular de la planta de la empresa responsable del operativo, contando con experiencia en el desarrollo de trabajos similares. Además de la supervisión regular directa de las entrevistas, se efectuó una verificación selectiva de consultas con posterioridad a su aplicación. Asimismo, para posibilitar una adecuada distribución de la muestra conforme a criterios sociodemográficos, el trabajo de campo se efectuó a lo largo de cuatro jornadas, dos entre semana y dos en fin de semana, operando en horarios matutinos como vespertinos, para maximizar las posibilidades de encuentro con diferentes integrantes de las familias residentes en viviendas seleccionadas.

Es de mencionarse que el calendario de observaciones consideró la reducción de lapsos entre observaciones en el periodo próximo a la celebración de los comicios federales. Ello, en el entendido de que la fuente primor-

GRÁFICA 1
EFECTO DEL ORDEN DE PREGUNTA EN LA ESTIMACIÓN DE INTENCIONES DEFINIDAS DE VOTO
POR CANDIDATO PARA LA ELECCIÓN DE PRESIDENTE DE LA REPÚBLICA*



* Respuesta a la pregunta "Si en este momento se celebraran las elecciones para Presidente de la República, ¿por cuál candidato votaría usted?" Solamente considera ciudadanos con credencial de elector al momento de la entrevista.

dial de divergencia entre las estimaciones por encuesta de distribuciones de preferencias y resultados electorales efectivos ha sido, generalmente, la lejanía temporal con el momento de la votación.⁵

POBLACIÓN OBJETIVO DEL ESTUDIO

La población objetivo de las encuestas nacionales de opinión ciudadana para este proyecto se definió como las personas de origen mexicano nacidas antes del mes de julio de 1982, residentes en el territorio nacional al momento de la entrevista. En tal sentido, el universo del estudio comprende a la totalidad de la población que legal y procedimentalmente se encontró en posibilidades de participar en los comicios federales de 2000, al contar para entonces con 18 y más años de edad.⁶

Luego, el conjunto de entrevistados no corresponde estrictamente a un conjunto de ciudadanos actuales, sino a una muestra de entre quienes serían ciudadanos al momento de las elecciones federales. Igualmente, considera la restricción al territorio nacional, por lo que excluyó del universo de estudio a los ciudadanos mexicanos residentes en el extranjero, la mayoría de ellos imposibilitados de ejercer, en la práctica, su derecho de voto.

Al definir la población bajo estudio como la totalidad de los ciudadanos mexicanos al momento de la elección residentes en el país, resultaba obligado adoptar un método de aproximación mediante entrevista directa en hogares.⁷ Ello, para disponer de condiciones que permitieran recurrir a procedimientos aleatorios que garantizaran la representatividad de la muestra respecto al universo.⁸

Como un primer segmento relevante para el análisis de las intenciones de sufragio, se dispuso del indicador para generar, a partir de la totalidad de personas entre-

vistadas, la submuestra de ciudadanos con credencial de elector, universo al que refirieron los datos sobre distribución de intenciones de voto en su oportunidad.⁹

Un segundo segmento relevante para el análisis de las intenciones de sufragio fue la submuestra de ciudadanos con credencial de elector que pudieran ser considerados como votantes probables, que conforme experiencia empírica no debieran presentar diferencias mayores respecto al conjunto de los electores.¹⁰

Basados en la experiencia lograda en la misma serie con la elección interna del PRI, el procedimiento adoptado fue la aplicación de una pregunta sobre la factibilidad del entrevistado de concurrir a votar.¹¹

Para fines de estimación de los "votantes probables", se consideraron entonces sólo a quienes, teniendo credencial de elector, respondieron que seguramente sí acudirían a votar, lo que constituyó un filtro aplicado *a posteriori* que dejó un segmento que permitía estimar una participación electoral potencial y un reparto esperable de votos, en términos absolutos y relativos.

CÁLCULO DEL TAMAÑO DE LA MUESTRA

El tamaño de muestra para la encuesta se determinó tomando en cuenta que se requería disponer de estimaciones a nivel nacional, con un margen de precisión tolerado de seis por ciento a un 95 por ciento de confianza. La variable relevante a estimar para los fines de la serie de encuestas nacionales es la proporción de la población con la intención manifiesta de sufragar a favor de alguno de los partidos políticos nacionales contendientes en las próximas elecciones federales.

De manera acertada, Page y Shapiro¹² recuerdan que el tradicional reporte de un margen de error de +/- tres

⁵ Como se demuestra en el estudio clásico de Irving Crespi, *Pre-election polling: Sources of Accuracy and Error*, Russell Sage Foundation, 1988.

⁶ Esta definición no fue utilizada para la encuesta especial realizada en octubre de 1999, previa a la elección del candidato del PRI a la Presidencia de la República, cuyo universo de estudio comprendió exclusivamente a los ciudadanos mexicanos al momento de la elección (nacidos antes de noviembre de 1991). Como segmentos relevantes de esta población, mediante la aplicación *a posteriori* de filtros, se contó con datos referidos a la población que disponía de credencial de elector y que, por ende, se encontraba en condiciones fácticas de poder participar en la elección (población credencializada) y de datos relativos a la población que, disponiendo de su credencial de elector, declaraba tener seguridad de acudir a votar en la consulta (considerados como votantes potenciales).

⁷ Respecto al debate calle-casa en México, entre otros textos puede verse: Raúl Trejo, *The Worst Opinions. Public Opinion Polls, Elections and the Media in Mexico, 1994*, en: Roderic Ai Camp (editor), *Polling for Democracy. Public Opinion and Political Liberalization in Mexico*, Scholarly Resources, Wilmington, 1996.

⁸ Asimismo, dada la reducida cobertura del servicio telefónico en hogares, la opción de operativos por este medio fue excluida. Series de encuestas telefónicas nacionales pueden no sólo dar cuenta del comportamiento en el segmento correspondiente, sino permitir detectar el sentido y magnitud de cambios en preferencias del electorado, contrastándolas con mediciones en domicilio y asumiendo ciertos supuestos; ello, sin menoscabo del hecho constatado de que arrojan muestras sesgadas conforme a criterios socioeconómicos y de que los respondientes por teléfono tienden a adoptar una actitud más complaciente con el entrevistador (cfr.: Alicia de la Macorra y Jorge Iván Castro, *En el caso de México, ¿es posible reemplazar una entrevista personal por una telefónica o viceversa*, disponible en el site: www.bimsa.com.mx).

⁹ Dado que la población bajo observación correspondió al total de personas que estarían en posibilidad de cumplir con las condiciones formales para ejercer su sufragio el 2 de julio de 2000 y en razón a que el proceso para obtener la credencial de elector concluyó hasta fines de marzo de 2000, no resultaba necesariamente pertinente, desde el principio de la serie, la aplicación de filtros en consideración a la tenencia de credencial.

¹⁰ Ello, dado que no existen procedimientos validados que permitan identificar eficazmente a los votantes probables y qué ejercicios de decantación en encuestas cercanas a elecciones dan estimaciones de preferencias para «votantes probables» similares a las correspondientes al total de electores, como se demuestra en: Ulises Beltrán and Marcos Valdivia, *Accuracy and Error in Electoral Forecasts: The Case of Mexico*, International Journal of Public Opinion Research, Vol. II, No. 2, 1999.

¹¹ Así, la batería de reactivos que permitieron detectar los segmentos relevantes para el análisis de la intención de voto fue: ¿Está usted empadronado? y si sí, ¿tiene credencial de elector? (respuesta espontánea); ¿qué tan factible es que vaya usted a ir a votar en las próximas elecciones federales? (presentando tarjeta con cuatro opciones de respuesta: seguramente sí, probablemente sí, probablemente no y seguramente no); si tuviera que votar en este momento para elegir Presidente de la República, ¿por cuál partido o coalición política votaría? o, en su caso, ¿por qué candidato votaría usted? (en ambos casos, mostrando tarjeta con opciones rotadas).

¹² Benjamin Page y Robert Shapiro, *op. cit.*

por ciento, a un nivel de confianza de 95 por ciento, para encuestas en torno a mil casos, corresponde al margen para muestreos aleatorios simples, siendo que en realidad las encuestas adoptan procedimientos de selección más complejos, por lo que un criterio de +/- 6 por ciento resulta por lo general más adecuado.

Lo anterior debido a que la ecuación empleada de forma convencional para el cálculo del tamaño de muestra, que parte de la proporción esperada de ocurrencia de un evento en el conjunto de la población bajo estudio, no resulta formalmente pertinente si se adopta un muestreo por conglomerados, como es el caso. De hecho, la ecuación pertinente en estos casos debe considerar el cociente de la varianza dividida por la cantidad de unidades independientes de muestreo, conforme al procedimiento que se detalla enseguida:

Primer paso: estimación de la varianza de la proporción relativa de votos por partido político a nivel distrital. Conforme a los resultados de las elecciones para diputados federales de 1997 y suponiendo la máxima varianza en la votación (correspondiente a una proporción para algún partido de $p = 0.5$), se tendría una estimación máxima para la varianza distrital de $s_c^2 = 0.0276$.

PARTIDO	VARIANZA DE VOTACIÓN RELATIVA NACIONAL	VARIANZA DISTRITAL DE VOTACIÓN RELATIVA	VARIANZA DISTRITAL / VARIANZA TOTAL
PRI	.2381	.0118	.0496
PAN	.1953	.0211	.1080
PRD	.1910	.0211	.1105
Resto	.0779	.0024	.0308

Segundo paso: se estima el número de conglomerados requeridos para satisfacer el margen de precisión estadística deseada, conforme a la relación¹³ $m = s_c^2 / e^2$ (donde "m" es igual al número de conglomerados a incluir en la muestra; " s_c^2 " es la varianza distrital máxima estimada; en este caso, 0.027625; " e^2 " es el cuadrado del error estándar tolerado; de 0.000936). No se incluye corrector finito en esta estimación. Conforme a lo anterior, se requirió disponer de 30 conglomerados en muestra, para satisfacer la condición de precisión establecida.

Tercer paso: se calculan los casos requeridos para cada conglomerado en muestra, conforme a la relación $A = (s^2 / s_c^2) + m + 1$, donde "A" corresponde al número de

casos en muestra por conglomerado. Sustituyendo, se tiene que el tamaño de muestra por conglomerado puede ser redondeado al valor de 40 casos y que el tamaño de muestra total requerido es luego de mil 200 casos.¹⁴

Derivado de lo anterior es posible estimar el efecto del diseño por conglomerados adoptado, a partir de la relación: $ed = r(A - 1) = (n s_c^2 / m s^2) - 1$, que da por resultado un efecto de diseño de 3.4 y un coeficiente de correlación intraclase (r) de + 0.087.

Es de mencionarse que conforme experiencias directas de estimación de propensiones al voto por partido, se considera que los diversos factores subjetivos involucrados en el cuestionamiento y verbalización de la intención de voto tienden a contraponerse, nulificando el efecto de estos errores en la estimación. De igual manera, la evidencia empírica disponible permite afirmar que si bien la propensión a manifestar una respuesta definida a favor de algún partido político no es homogénea en la población —estando afectada por diferencias socioeconómicas, informativas e ideológicas—, la distribución de las respuestas definidas tiende a ser un estimador eficiente de la distribución efectiva de sufragios por partido.

Lo anterior, en el entendido de que el segmento de no definidos tiende a corresponder bien a población abstencionista, bien a votantes que se definen tardíamente, pero que se distribuyen conforme a patrones relativamente coherentes a los del resto de la población, por lo que en mediciones muy próximas a los comicios provocan un cambio marginal en las distribuciones de preferencias.

Para este proyecto se prefirió ser conservadores y considerar que se ejercita un procedimiento de exclusión¹⁵ y no de reasignación. Entenderlo de una u otra forma no tiene impacto en las proporciones y, dado que el margen de error tolerado se estimó a partir de la relación entre la varianza y el número de unidades independientes en muestra, tampoco debe tenerlo en el nivel de precisión tolerado para las estimaciones.

De igual manera no se afecta el error tolerado para las estimaciones cuando se retiene solamente el segmento correspondiente a los "votantes probables", aunque únicamente las proporciones correspondientes a los "votantes probables" respecto al total de electores pueden ser consideradas como pertinentes para fines de una "expansión" de la muestra que permita estimar volúmenes esperables de votantes por contendiente.

¹³ Ver: L. Festinier y D. Katz, *Los métodos de investigación en ciencias sociales*, Paidós, México, 1993.

¹⁴ Con miras a establecer un criterio para la comparabilidad entre el margen de error tolerado efectivamente para una encuesta por conglomerados y el margen de error calculado como si fuese un muestreo aleatorio simple, puede estimarse este último a partir de la relación: $E^2 = Z^2 pq / n$; donde "E" corresponde al margen de error en un muestreo aleatorio simple con el tamaño de muestra indicado y "Z" al número de unidades de desviación estándar de la media a la que se estima el error (que en este caso asume un valor de 1.96, donde el área bajo la curva normal ubicada entre la media y las ordenadas al valor referido a ambos lados es de 0.95, que es el nivel de confianza definido para la estimación). Sustituyendo, se tiene que el margen de error calculado para una muestra aleatoria simple del tamaño elegido es de 2.8 por ciento para la población observada (tres por ciento si se considera solamente a la población a entrevistar que se asume dispondrá de credencial de elector). Este dato es el que corresponde al margen de error convencionalmente proyectado en los reportes de encuestas con fines de medición de intenciones de sufragio, que suele corresponder al criterio adoptado de seis por ciento de margen de error correctamente estimado para un muestreo por conglomerados.

Es de mencionarse que en las ocasiones en que se realizaron encuestas próximas a elecciones (en octubre de 1999, antes de la elección interna del candidato del PRI; y en la segunda quincena de junio de 2000, antes de las elecciones constitucionales), se incrementó el tamaño de la muestra con miras a reducir el margen de error tolerado.

Así, para la encuesta especial previa a la elección interna del candidato del PRI a la Presidencia de la República, con la intención de minimizar los márgenes de error sin incrementar sustancialmente costos de operación, se adoptó un método de muestreo doble con un ajuste de la mitad de las unidades primarias en muestra entre ambas mediciones.¹⁶

En el caso de la encuesta inmediata previa a las elecciones federales, al tomarse el doble de unidades primarias en muestra,¹⁷ se toleró un margen de error de 4.2 por ciento a un nivel de confianza de 95 por ciento. En total, se dispuso de dos mil 400 entrevistas a personas de 18 y más años de edad residentes en el territorio nacional, dos mil 287 de las cuales disponían de su credencial de elector. Dicha muestra fue levantada en 60 distritos ubicados en 30 entidades federativas.

DISEÑO DE LA MUESTRA

Un muestreo para un estudio previo a elecciones que tenga por intención estimar las proporciones de votación esperables para cada contendiente deberá buscar una representatividad del electorado nacional y, de considerarse pertinente, del segmento relevante para los fines de estimación, debido a su propensión a participar en el pro-

ceso. Ello obliga a un muestreo probabilístico, al menos en las primeras etapas de selección, aunque en la fase de elección de vivienda y respondente pueda adoptar procedimientos relativamente más laxos.¹⁸

El cálculo anterior del tamaño de muestra supone la adopción de un procedimiento de muestreo por conglomerados, que asuma como una primera etapa la selección de distritos electorales federales. El diseño de muestra adoptado de acuerdo a lo anterior, que parte del Listado Nominal de Electores para 1997, cuyos resultados agrupados a nivel de distrito están públicamente disponibles, es el siguiente:

Primera etapa: selección de unidades primarias de muestreo, en este caso distritos electorales federales. Se seleccionaron de manera sistemática tanto distritos electorales federales como se requiere conforme al nivel de precisión establecido, con probabilidad proporcional al tamaño (PPT), tomando como medida del tamaño el número de ciudadanos inscritos en el listado nominal del distrito electoral federal para el proceso electoral 1997.¹⁹

Con miras a maximizar la comparabilidad de resultados y reducir las variaciones temporales en estimaciones diversas, se decidió proceder a un único ejercicio de selección de unidades primarias de muestreo, válido para el conjunto de encuestas nacionales incluidas en este programa.

La retención de las unidades primarias en muestra, aun y cuando no permitía la contrastación de estimaciones producto de las diversas mediciones efectuadas a lo largo de la serie, sí posibilitaba una mayor y más precisa comparación entre las propias estimaciones, lo que era objetivo central en el proyecto.²⁰

Conforme al procedimiento indicado, los distritos seleccionados en muestra fueron los siguientes:

¹⁶ Que supone asumir a los indefinidos en encuestas como abstencionistas en elecciones, *cf.*: R. de la Peña, *Sobre la dinámica de las preferencias electorales y el problema de los indefinidos*, ponencia presentada en el Décimo Congreso Nacional en Estudios Electorales, Sociedad Mexicana de Estudios Electorales; La Paz, Baja California Sur, diciembre de 1998.

¹⁷ Para contar con estimaciones para la población credencializada con un margen de precisión tolerado de 4.3 por ciento, a un nivel de confianza de 95 por ciento, se calculó que se requerían 30 conglomerados en cada muestra, con ajuste de la mitad de unidades primarias, partiendo de un coeficiente de correlación intraclase de 0.087, empíricamente observado, y considerando para el cálculo una relación propia para muestreo doble (basada en: Des Raj, *Teoría del muestreo*; Fondo de Cultura Económica, México, 1980), que establece una $m = [s^2(1 + \rho) / 2(1 - \rho)] / \rho^2$ (donde "m" es el número de conglomerados a incluir en muestra; " ρ " es la proporción no retenida o ajustada de la muestra; y " ρ^2 " es el error estándar tolerado, en este caso 0.022).

¹⁸ Para fines de ampliación de las unidades primarias en muestra, se adoptó un procedimiento sistemático no independiente del muestreo original, al incluir en muestra los distritos ubicados a la mitad del intervalo previo considerado.

¹⁹ Al respecto, cabe recordar la experiencia de la encuesta de Belden y Alagón en ocasión de las elecciones federales de 1994, que llevó la selección aleatoria hasta el respondente, pero que arrojara estimaciones muy próximas a las logradas al mismo tiempo con encuestas que adoptaron procedimientos menos rigurosos (al respecto, véase: Ricardo de la Peña, *Las encuestas después de 1994*, en: M. Larrosa y L. Valdés, *Elecciones y partidos políticos en México, 1995*, Universidad Autónoma Metropolitana Iztapalapa, México, 1998).

²⁰ Para ello, tomando el listado de distritos electorales federales, que sigue un orden alfabético por entidad federativa y numérico por distrito electoral, se forman las sumas acumuladas de ciudadanos comprendidos en cada distrito electoral federal y el acumulado total de ciudadanos listados en el país. El intervalo de selección I se obtiene dividiendo el acumulado total de ciudadanos entre el número de distritos a seleccionar (en este caso, considerando 30 distritos electorales federales a elegir y un número total de ciudadanos en listado de 52 millones 208 mil 966 personas, se tiene un intervalo de selección de un millón 740 mil 299). Luego, se selecciona un número aleatorio r entre 1 e I para disponer del primer número aleatorio, y luego se adiciona I a r tantas veces como se requiere para generar los 30 números aleatorios requeridos (en este caso, el primer número aleatorio generado fue un millón 473). Los números aleatorios generados en forma sistemática sirven para identificar en el listado de sumas acumuladas las unidades primarias correspondientes, que serán las seleccionadas en muestra. Formalmente, este muestreo sería con reemplazo, pues si la medida del tamaño de alguna de las unidades primarias resulta mayor que el intervalo de selección, queda seleccionada más de una vez, tomándose tantas submuestras como veces fue seleccionada; empero, dado la relativa homogeneidad de tamaño de las unidades primarias, tal eventualidad no se presenta en la práctica.

²¹ De hecho, este diseño permite considerar que se disponía solamente de un segmento del total de secciones electorales del país, que incluía las seis mil 859 secciones ubicadas en los 30 distritos considerados. En cada ocasión se eligió una muestra aleatoria de 240 secciones, estratificadas por distrito, lo que (dado un coeficiente de correlación intraclase negativo), arrojaría un margen de error tolerado para las estimaciones en este segmento de ± 2.5 por ciento a un nivel de confianza de 95 por ciento, que permitiría una adecuada observación de cambios a su interior.

DISTRITOS ELECTORALES FEDERALES
EN LA MUESTRA BÁSICA DE LA SERIE GEA-ISA

Baja California	04	Hidalgo	05	Oaxaca	08
Coahuila	04	Jalisco	08	Puebla	09
Chiapas	06	Jalisco	18	Querétaro	04
Chihuahua	05	México	08	Sinaloa	01
Distrito Federal	05	México	17	Sonora	03
Distrito Federal	14	México	26	Tamaulipas	01
Distrito Federal	23	México	36	Tlaxcala	02
Durango	02	Michoacán	11	Veracruz	10
Guanajuato	07	Nayarit	03	Veracruz	20
Guerrero	03	Nuevo León	09	Zacatecas	02

Es de mencionar que esta muestra de distritos presentaba un comportamiento electoral muy similar al total nacional, con una desviación media por partido de 0.0027, como se muestra en la siguiente tabla:

PARTIDO	VOTACIÓN RELATIVA NACIONAL	VOTACIÓN EN DISTRITOS EN MUESTRA	DIFERENCIA DE LA MUESTRA AL TOTAL
PRI	.391	.387	- .004
PAN	.266	.269	+ .003
PRD	.257	.259	+ .002
Resto	.086	.085	- .001

Al efectuar similar ejercicio respecto a elecciones federales anteriores, efectuando el ajuste correspondiente de las secciones para ubicarlas en la actual división distrital, se encontraría una diferencia media entre la votación por partidos en los distritos en muestra y el resultado nacional en 1994, de 0.0104, y para 1991, de 0.0171.

Las secciones electorales en esta muestra de 30 distritos se distribuían según su carácter urbano o rural de manera similar al promedio nacional. Si a nivel nacional y conforme al Registro Federal de Electores de 1997, 67.94 por ciento de las secciones son urbanas, en los 30 distritos en muestra 67.08 de las secciones son urbanas, lo que supuso un ligero sesgo (de 0.86 por ciento) hacia lo rural. Dado que se selecciona una cantidad similar de secciones en cada distrito, es importante mencionar que los distritos en muestra cuentan con un promedio de 28 por ciento de secciones rurales, contra 29 por ciento en el total nacional, siendo ésta la fracción de secciones rurales en muestra esperable.

Como dato complementario, disponible una vez celebradas las elecciones federales de 2000, la muestra de 30 distritos electorales tomada regularmente para la serie, efectivamente reprodujo el comportamiento nacional, confirmando lo atinado de decisiones de método: la diferencia entre los resultados en estos distritos y el resul-

tado del PREP difiere en menos de medio punto para cada contendiente.²¹

CONTENDIENTE	VOTACIÓN RELATIVA NACIONAL	VOTACIÓN EN DISTRITOS EN MUESTRA	DIFERENCIA DE LA MUESTRA AL TOTAL
Alianza por el Cambio	.438	.445	+ .007
PRI	.367	.361	- .006
Alianza por México	.169	.166	- .003
Resto	.026	.028	+ .002

Empero, frente a cuestionamientos expresados respecto a la posible existencia de sesgos derivados de la muestra de distritos elegida y retenida en las diversas mediciones, en la segunda quincena de marzo se efectuó un ejercicio de contrastación de muestra, dejando constantes otros elementos como el personal y controles de campo y el instrumento de recuperación.

Si de los días del 16 al 21 de marzo se había llevado a cabo el levantamiento regular de la octava encuesta de la serie, de los días del 25 al 28 de marzo se levantó otra encuesta con una muestra nacional similar a la utilizada en esta serie, pero en una selección distinta de distritos, aplicándose mil 200 entrevistas de un cuestionario donde se ubicó la pregunta sobre intención de voto por candidato inmediatamente después de los datos de identificación del entrevistado, confrontables con los resultados obtenidos en el ejercicio previo de contrastación de cuestionario.

Los resultados obtenidos permitieron constatar que no existía una diferencia significativa entre ambas mediciones, por lo que se excluía la existencia de un sesgo derivado de la específica selección de distritos en muestra. De hecho, la diferencia entre ambas mediciones fue de 2.1 puntos sobre la distribución de respuestas observada y de apenas 0.9 puntos al quedarse con las intenciones definidas de voto por candidato (Ver Gráfica 2).

Segunda etapa: selección de las unidades secundarias de muestreo, en este caso secciones electorales. En cada una de las unidades primarias de muestreo elegidas, se seleccionaron en cada ocasión, de manera aleatoria, ocho secciones electorales con probabilidad proporcional a tamaño, medido por el número de votantes reportado en la elección federal anterior. De esta forma, en total se dispuso de 240 unidades secundarias en muestra.

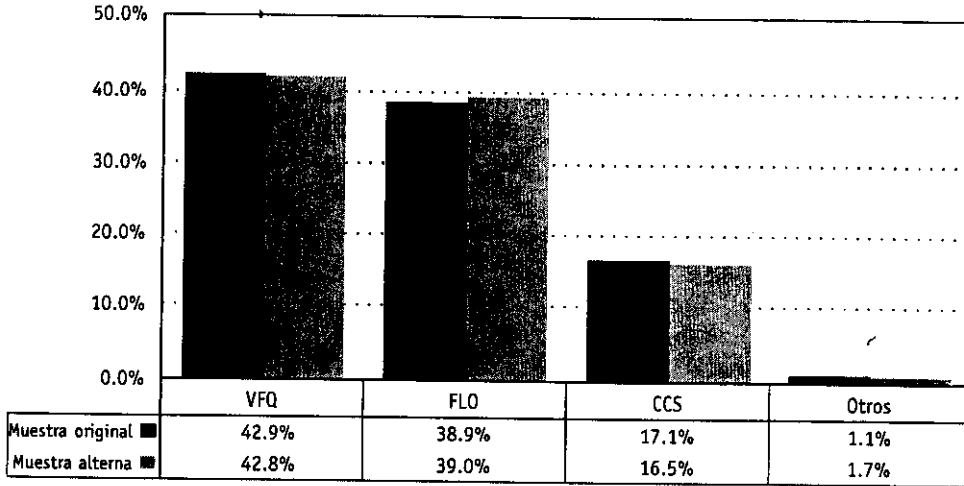
Tercera etapa: en cada una de las unidades secundarias de muestreo seleccionadas y mediante un procedimiento de recorrido aleatorio,²² se eligieron cinco viviendas a ser incluidas en muestra, para alcanzar el volumen de mil 200 casos en muestra requerido. En cada una de estas viviendas se selecciona a la persona que,

²¹ Que se reduce a 0.2 por ciento cuando se consideran los sesenta distritos incluidos en la muestra completa para la última encuesta nacional de la serie.

²² Que tomó como punto de arranque la ubicación de la casilla básica en el proceso electoral federal 1997, información disponible mediante consulta de fuentes hemerográficas diversas.

GRÁFICA 2

EFFECTO DEL ORDEN DE PREGUNTA EN LA ESTIMACIÓN DE INTENCIONES DEFINIDAS DE VOTO POR CANDIDATO PARA LA ELECCIÓN DE PRESIDENTE DE LA REPÚBLICA*



* Respuesta a la pregunta "Si en este momento se celebraran las elecciones para Presidente de la República, ¿por cuál candidato votaría usted?" Solamente considera ciudadanos con credencial de elector al momento de la entrevista.

formando parte de la población objetivo, acudía a la puerta en primer término; en el caso de no poderse efectuar la entrevista en dicha vivienda, por estar ausentes sus habitantes, no cumplir condiciones para inclusión en muestra o rechazar la entrevista, se procedió a entrevistar en la vivienda contigua.

La fracción de muestreo global estuvo dada por el producto de las fracciones de muestreo calculadas en cada una de las tres etapas.²³ Al efectuar los cálculos, se estableció que esta fracción de muestreo era igual al número de casos totales en muestra entre la población total observada. Por tanto, con el procedimiento adoptado se conseguía que cualquier ciudadano en muestra tenga la misma probabilidad de ser seleccionado; es decir, este diseño de muestra es autoponderado. Luego, el factor de expansión muestral correspondiente vendrá dado como el cociente de la población total en el listado entre el número de casos en muestra, por lo que es factible una expansión simple al multiplicar las respuestas por el inverso de la fracción de muestreo (equivalente para el caso a 43 mil 507), lo que no afecta las proporciones estimadas con las que se ha de trabajar.

Este criterio de expansión pudo ser aplicado para la estimación no solamente de proporciones de intención de voto por contendiente, sino para el cálculo de la votación esperada, que correspondería específicamente a las proporciones del electorado que declara tener la intención de votar por alguno de los contendientes y que

afirma tener la seguridad de acudir a votar ("votantes seguros").

RESULTADOS DE LA SERIE GEA-ISA

Dos momentos fueron claves para validar esta serie: la elección del candidato del PRI a la Presidencia de la República, en noviembre de 1999, y las elecciones federales del 2 de julio de 2000, por lo que vale la pena analizar por separado cada una de ellas.

Estimaciones para la elección del candidato del PRI

Un procedimiento de validación de una metodología es su contrastación con evidencia empírica. En el caso de la serie de encuestas nacionales que es materia de este ensayo, un recurso preliminarmente disponible para validarla es confrontar las estimaciones de intención de voto que arrojara la medición especial previa a la elección del candidato del PRI a la Presidencia de la República, contra los resultados oficiales de esta consulta.

Efectuando el ejercicio de comparación de la muestra doble correspondiente a la encuesta de la serie que nos ocupa contra los resultados de la consulta, la encuesta final GEA-ISA habría presentado una diferencia media en la estimación por aspirante de 3.8 puntos, subestimando el margen de victoria en 11 puntos. La precisión de la encuesta mejora de considerarse la muestra original seriada: la diferencia media entre estimación y resultado por

²³ La fracción de muestreo para cada distrito electoral federal seleccionado en la primera etapa era igual al cociente de la población bajo muestreo que reside en tal distrito electoral federal seleccionado entre el total de los ciudadanos en el listado; la fracción de muestreo para cada sección electoral seleccionada en la segunda etapa era igual al cociente del número de secciones electorales elegidas por distrito seleccionado entre el total de secciones electorales existentes en el distrito en cuestión; y la fracción de muestreo para cada ciudadano seleccionado en la tercera etapa era igual al cociente del número de casos en muestra en la sección electoral entre el total de residentes en la sección electoral específica.

aspirante erraría en 3.0 por ciento, siendo de seis puntos la subestimación del margen de victoria.

Lo anterior sustenta la confianza en que la serie de encuestas nacionales aporta estimaciones con suficiente precisión para los fines de estudio, puesto que ante el problema añadido de extraer una submuestra pertinente de votantes potenciales para la estimación de intenciones de voto, la medición se ubicó dentro de los márgenes de error esperados (Ver Gráficas 3 y 4).

Estimaciones para la elección presidencial de 2000

La presentación de los resultados de estimaciones de intención de voto para la elección presidencial de 2000 pudiera dividirse en dos grandes paquetes informativos:

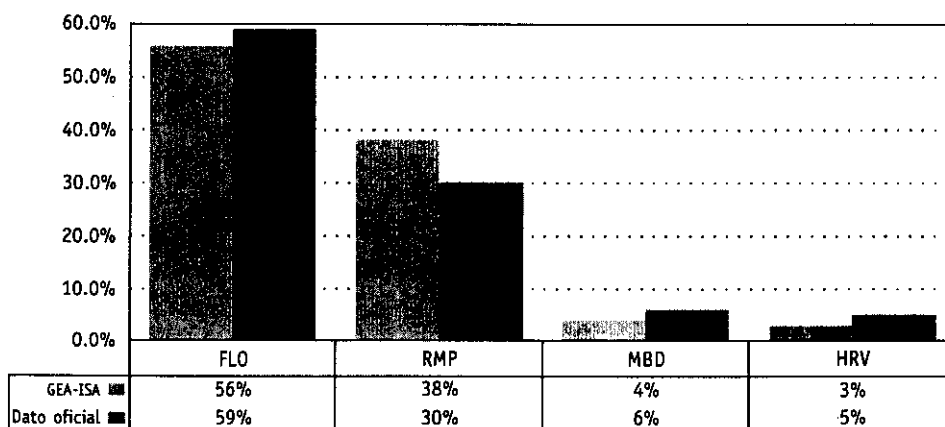
lo correspondiente a las preferencias electorales por partido y lo relativo a las preferencias por candidato.

Respecto a las preferencias por partido, la serie observa los cambios a lo largo de dos años, entre junio de 1998 y junio de 2000, detectando los siguientes movimientos:

El PRI pasaría de 28 a casi 39 por ciento de las preferencias en el periodo junio de 1998-enero de 1999, estabilizándose en torno a 35 por ciento a lo largo de 1999, para posteriormente descender a niveles de alrededor de 30 puntos en 2000. Ello significaba que, a nivel de intenciones definidas de voto, el PRI ascendería de 41 a 48 por ciento a enero de 1999, situándose próximo a 45 por ciento el resto del año, mientras que durante 2000 se ubicaría regularmente por debajo de 40 por ciento.

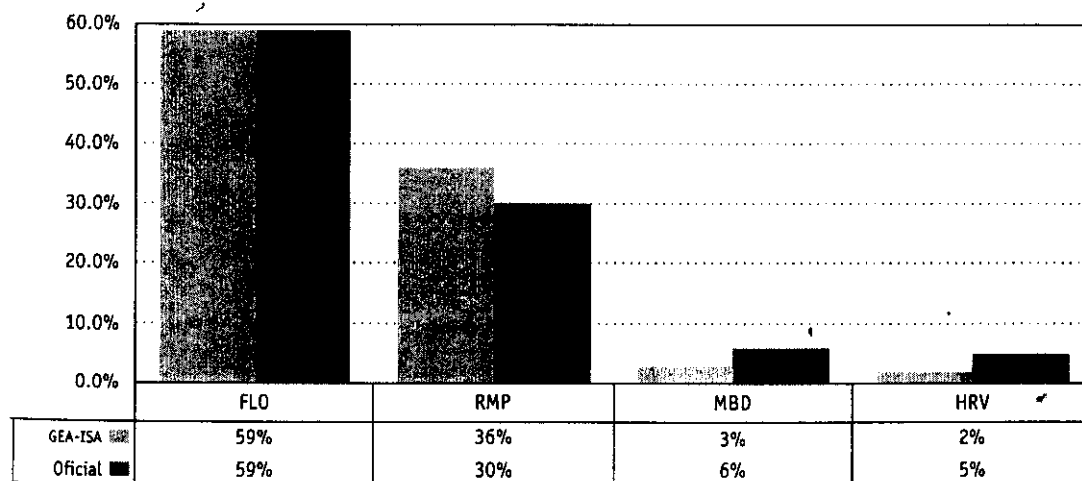
GRÁFICA 3

INTENCIÓN DE VOTO DE VOTANTES SEGUROS EN LA ENCUESTA GEA-ISA (MUESTRA DOBLE)
Y RESULTADO DE LA ELECCIÓN DEL CANDIDATO DEL PRI A PRESIDENCIA DE LA REPÚBLICA



GRÁFICA 4

INTENCIÓN DE VOTO DE VOTANTES SEGUROS EN LA ENCUESTA GEA-ISA (MUESTRA BÁSICA)
Y RESULTADO DE LA ELECCIÓN DEL CANDIDATO DEL PRI A PRESIDENTE DE LA REPÚBLICA

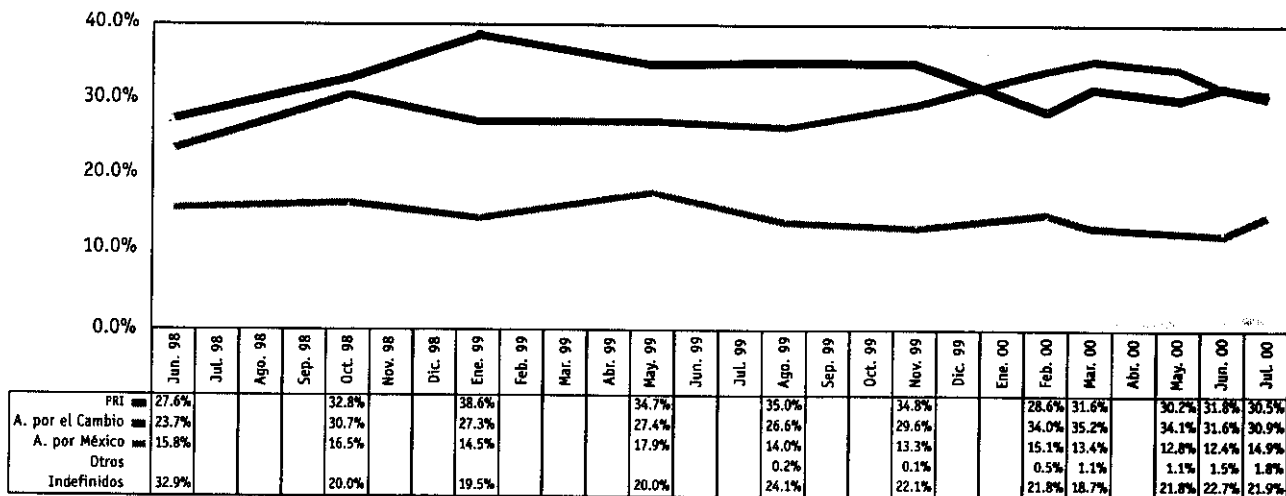


Los partidos que conformarían la Alianza por el Cambio mantenían una estrecha cercanía con el PRI en 1998, que se ve afectada a principios de 1999 cuando se ubican por debajo de 30 por ciento, entre cinco y nueve puntos por debajo del PRI, aunque en 2000 la Alianza por el Cambio superaría el nivel de 30 por ciento en forma sostenida, rebasando en muchos momentos al PRI. A nivel de intenciones definidas, lo anterior significa que los partidos de la Alianza por el Cambio estarían rondando nive-

les de entre 34 y 39 por ciento a lo largo de 1998 y 1999, rebasando la marca de 40 por ciento durante 2000, aunque con un relativo declive en el mes de junio.

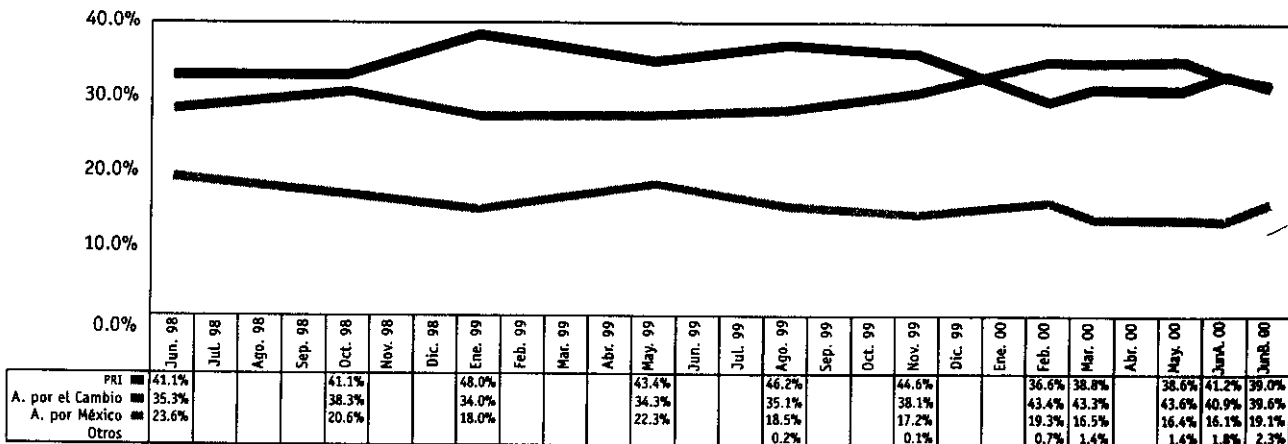
En el caso de los partidos que integrarían la Alianza por México, se situarían a lo largo de los dos años bajo observación en torno a 15 por ciento de las intenciones manifiestas de voto, aunque a partir de agosto de 1999 no alcanzarían a superar la barrera de 20 por ciento de las intenciones definidas de sufragio.

GRÁFICA 5
INTENCIÓN MANIFIESTA DE VOTO PARA PRESIDENTE DE LA REPÚBLICA POR PARTIDO O COALICIÓN*



* Respuesta a la pregunta "Si tuviera usted que votar en este momento para elegir Presidente de la República, ¿por qué partido (o coalición) votaría usted?". Solamente considera ciudadanos con credencial de elector al momento de la entrevista.

GRÁFICA 6
INTENCIÓN DEFINIDA DE VOTO PARA PRESIDENTE DE LA REPÚBLICA POR PARTIDO O COALICIÓN*



* Respuesta a la pregunta "Si tuviera usted que votar en este momento para elegir Presidente de la República, ¿por qué partido (o coalición) votaría usted?". Solamente considera ciudadanos con credencial de elector al momento de la entrevista.

En cuanto a las intenciones de voto por candidato, éstas pueden ser diferenciadas en dos momentos: antes de que se definiera la lista efectiva de contendientes y después de que ésta estuvo definida.

Durante 1999 se hizo el seguimiento de las preferencias por candidatos mediante preguntas específicas que presentaban opciones únicas por partido, siendo una de ellas la que consideraba a Labastida, por el PRI; Fox, por el PAN; y Cárdenas, por el PRD, aunque aún sin atender posibles coaliciones. En dichos "careos" se observó durante 1999 una situación de práctico empate entre Labastida y Fox, que solamente se rompió en la medición más próxima posterior a la elección interna del PRI, cuando Labastida superó en tres puntos a Fox. La altura alcanzada durante 1999 por cada uno de los dos principales contendientes rondaba en los cuarenta puntos.

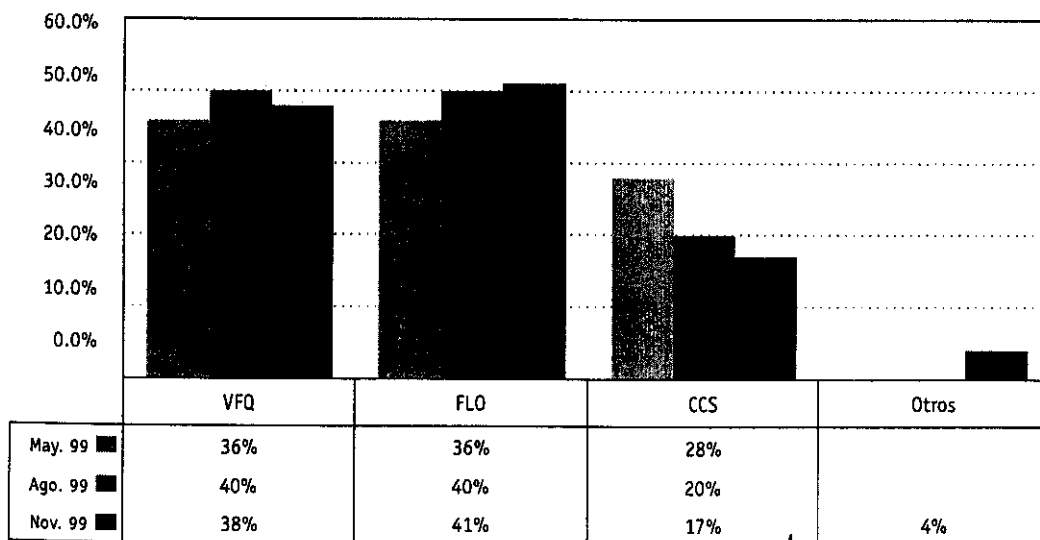
Estas distribuciones parecían mostrar desde entonces el eventual arribo a un "equilibrio duvergeriano"²⁴ no sólo como una propuesta teórica, sino como propensión efectiva: los dos primeros lugares, Fox y Labastida, pasaban de reunir 72 por ciento de intenciones a prácticamente 80 por ciento y mantenían una condición de empate en la disputa por la Presidencia de la República, mientras que el tercer lugar, Cárdenas, tendía a rezagarse, reduciendo su participación en más de diez puntos por debajo del nivel de 20 por ciento.

En apoyo a esta hipótesis estaba la elevada estabilidad mostrada en la serie de mediciones a lo largo de 1999, con giros implícitos del orden de tres puntos cada trimestre. En segundo lugar era interesante constatar que la serie mostró un impacto reducido de la campaña por la candidatura del PRI en la distribución de preferencias (Ver Gráfica 7).

Lo anterior hacía suponer que se estaba ante un escenario de competencia tendencialmente bipartidista, no como consecuencia de cambios en la decisión de voto entre partido en el gobierno y opciones de oposición, sino por desplazamientos en el sentido del voto al interior del segmento de votantes opositores. La hipótesis anterior, aventurada desde fines de 1999,²⁵ se vio confirmada durante los meses de campaña: un formato bipartidista, donde la disputa por la Presidencia de la República se dio efectivamente entre el candidato del partido en el gobierno y el candidato opositor que logró concentrar mayor respaldo desde el arranque.

Es en este marco de "equilibrio duvergeriano" que la serie GEA-ISA detecta un ligero giro en las preferencias entre noviembre de 1999 y febrero de 2000, de apenas cinco puntos en dos y medio meses (dos puntos por mes), pero que fuera suficiente para invertir el liderato en la carrera presidencial, pasando Fox a ocuparlo y desatando las críticas y cuestionamientos a la serie.

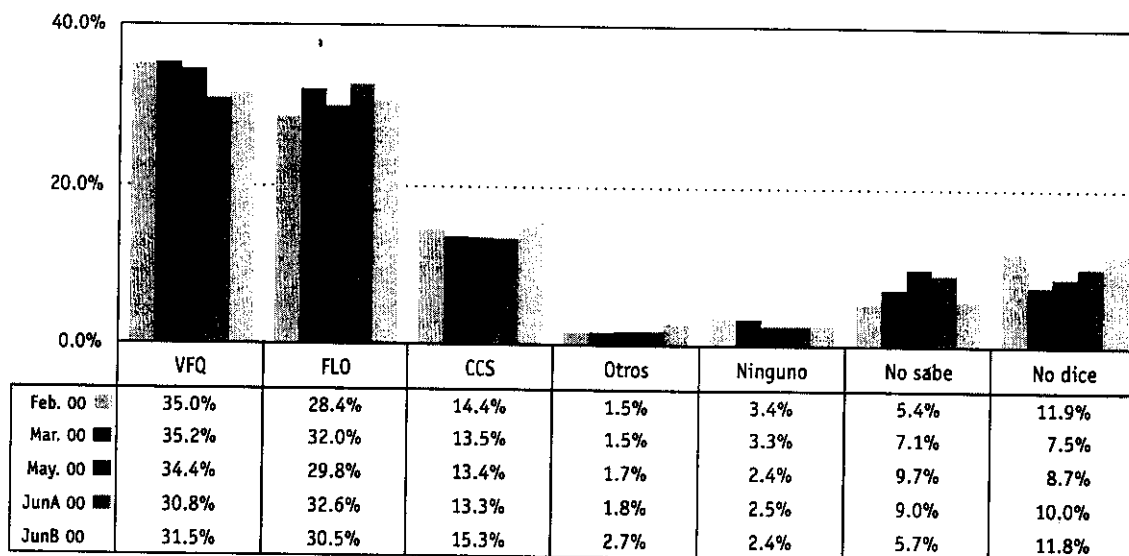
GRÁFICA 7
INTENCIÓN DEFINIDA DE VOTO POR PRECANDIDATOS
PARA LA ELECCIÓN DE PRESIDENTE DE LA REPÚBLICA



²⁴ En el sentido planteado por Gary W. Cox, en: *Making Votes Count*; Cambridge University Press, 1997.

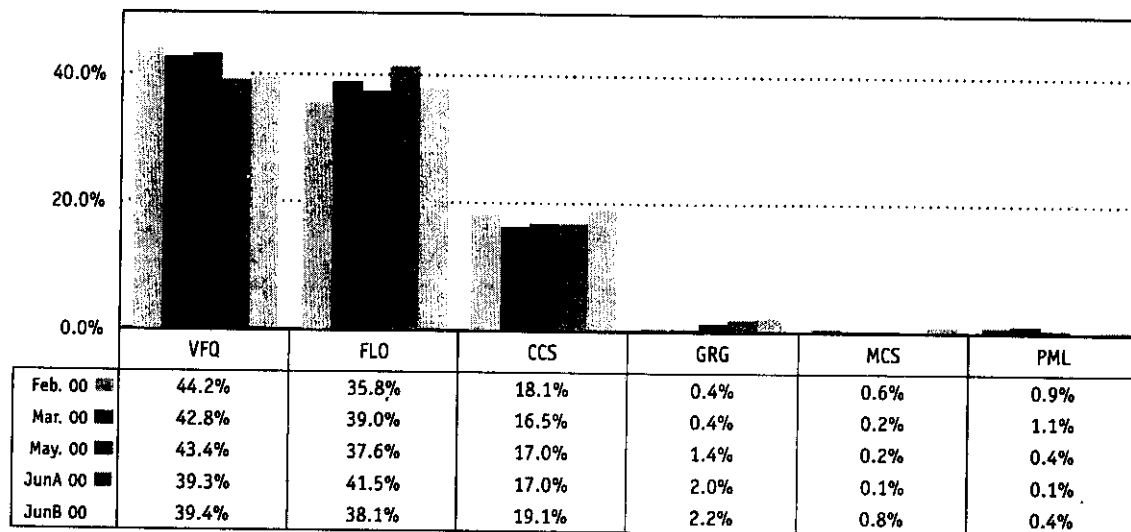
²⁵ Al respecto, véase: Ricardo de la Peña y Rosario Toledo, *Escenarios para las elecciones federales del 2000*, ponencia presentada en el Décimo primer Congreso Nacional de Estudios Electorales; Universidad Autónoma de Puebla y Sociedad Mexicana de Estudios Electorales; Puebla, 1-4 de diciembre de 1999.

GRÁFICA 8
INTENCIÓN MANIFIESTA DE VOTO POR CANDIDATOS
PARA LA ELECCIÓN DE PRESIDENTE DE LA REPÚBLICA*



* Respuesta a la pregunta "Si en este momento se celebraran las elecciones para Presidente de la República, ¿por qué candidato votaría usted?". Solamente considera ciudadanos con credencial de elector al momento de la entrevista.

GRÁFICA 9
INTENCIÓN MANIFIESTA DE VOTO POR CANDIDATOS PARA LA ELECCIÓN DE PRESIDENTE DE LA REPÚBLICA*



* Respuesta a la pregunta "Si en este momento se celebraran las elecciones para Presidente de la República, ¿por qué candidato votaría usted?". Solamente considera ciudadanos con credencial de elector al momento de la entrevista.

A lo largo de 2000 las encuestas de GEA-ISA mostraron a Fox siempre por encima de 30 por ciento de las intenciones manifiestas de voto, que representaban entre un máximo de 44 y mínimos apenas por debajo de 40 puntos. A diferencia, salvo en una observación, Labastida siempre se mantuvo debajo de 40 por ciento de las intenciones definidas, mientras que Cárdenas se ubicaba en un distante tercer lugar, entre 16 y 19 por ciento de las intenciones definidas.

La brecha más amplia medida, de ocho puntos a favor de Fox, fue en febrero en un levantamiento efectuado justamente entre el estallido de los hechos de violencia más importantes ocurridos en la Universidad Nacional y la ocupación de la Ciudad Universitaria. Luego, Fox se situaría entre marzo y mayo con una ventaja de cerca de cinco puntos, que no se afectaría de manera detectable por el primer debate, pero que sí se revertiría como consecuencia de los eventos relacionados con el segundo

debate. Así, una observación hecha a principios de junio mostró a Labastida adelante, pero como un fenómeno de “burbuja” que se desactivaría 15 días después, cuando nuevamente se detecta una ventaja para Fox, que se ratificaría en la elección.

Los datos anteriores corresponderían a las preferencias de la totalidad de entrevistados que disponían de su credencial de elector. Pero un segundo nivel de análisis, validado por la experiencia previa en la elección interna del PRI, era la observación del segmento de “votantes probables”, entendiendo por tales a los ciudadanos con credencial de elector que manifestaban estar seguros de ir a votar en los comicios federales de 2000.

La proporción del electorado que se manifestaba segura de ir a sufragar fue creciendo a lo largo de la campaña. Así, si en febrero 58 por ciento de los consultados afirmaban que lo haría, entre marzo y principios de junio la proporción se eleva a casi 70 por ciento, llegando a 75 por ciento en la medición de mediados del mes de junio. A estas proporciones habría que restarle el segmento de electores que, asegurando que irían a votar, no manifestaron una preferencia definida por alguno de los contendientes.

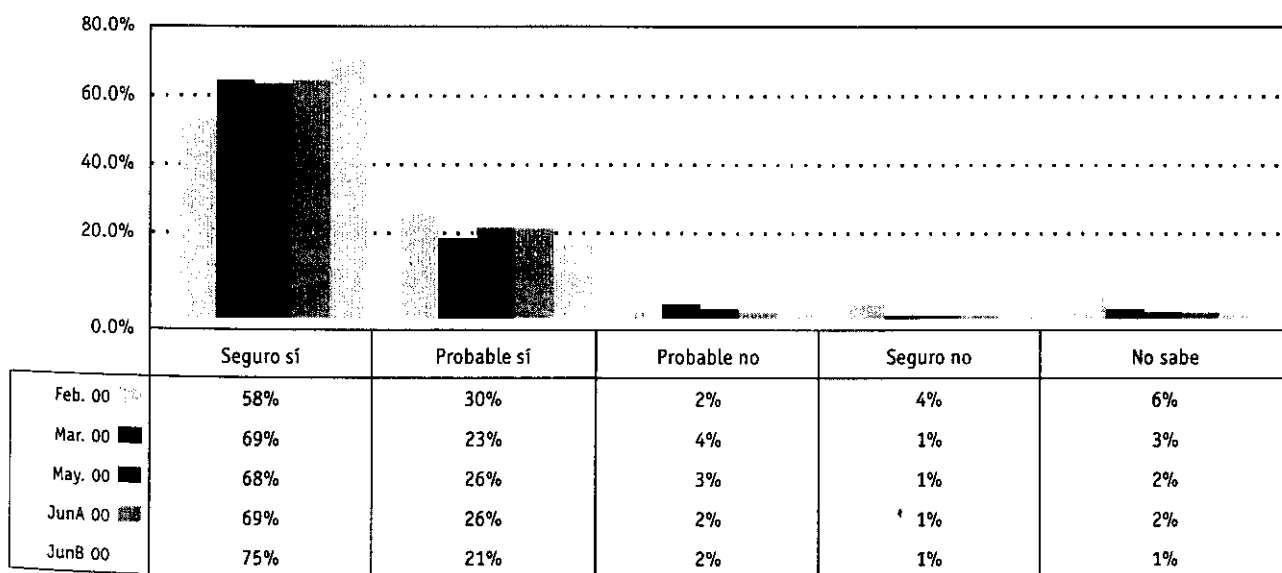
Efectuando este ajuste, se tendría que en febrero los votantes probables se ubicaban en torno a 49 por ciento, aproximadamente 29 millones de electores. Para el mes de marzo, la proporción había subido a 59 por ciento, que representaban casi 35 millones de electores. Para

principios de junio, luego de los debates y eventos que los rodearon, había bajado a 56 por ciento, disminuyendo en dos millones la votación esperable. Finalmente, para mediados de junio, había subido a 62 por ciento, equivalente a 36 millones de votantes probables, cifra muy cercana al volumen real de votos emitidos por candidatos en la jornada electoral (Ver Gráfica 10).

El segmento de votantes probables mostraría a lo largo de 2000 una mayor estabilidad en sus inclinaciones de voto que el conjunto de ciudadanos con credencial. Así, Fox llegaría a 38 por ciento de intenciones manifiestas y 45 por ciento de las definidas dentro de este segmento en el mes de marzo, para descender hasta 33 por ciento de manifiestas y 41 por ciento de definidas para mediados de junio. Su principal contrincante, Labastida, rondaría en torno a 32 por ciento de preferencias manifiestas en este segmento a lo largo del año, entre 38 y 41 por ciento de las definidas. Cárdenas, por su parte, rondaría entre 13 y 18 por ciento de las intenciones definidas de voto a su favor.

El margen de ventaja de Fox sobre Labastida sería de cinco puntos en febrero y marzo y de cuatro puntos en mayo, que se reducirían a medio punto para principios de junio, para nuevamente abrirse a más de dos puntos para mediados de junio, terminando en seis y medio puntos al día de la elección. A lo largo de la campaña, Fox habría tenido ventaja en todo momento.

GRÁFICA 10
¿QUÉ TAN FACTIBLE ES QUE VAYA A IR A VOTAR EN LAS PRÓXIMAS ELECCIONES FEDERALES?*



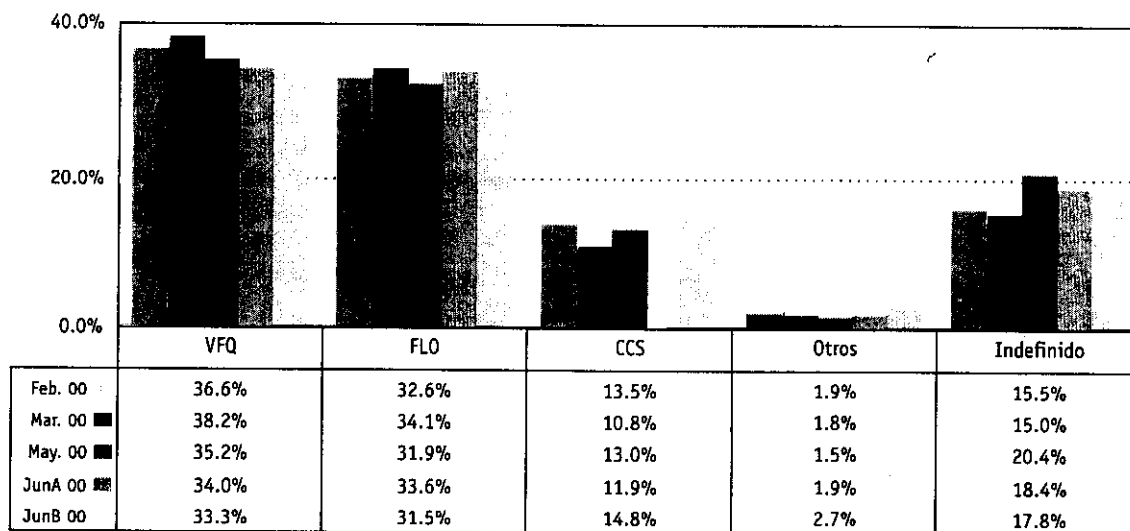
* Solamente considera ciudadanos con credencial de elector al momento de la entrevista.

La estimación de intención de voto de los votantes seguros de la encuesta final de la serie resultó sumamente próxima con el resultado de la elección presidencial. De hecho, el giro implicado entre esta estimación y el dato oficial es de tres por ciento (lo que es lo mismo, un error medio por contendiente de 1.5 por ciento).

En todos los casos, la proporción real de votos para cada uno de los tres principales contendientes se ubicaría dentro del margen de error estadístico calculado para la estimación (Ver Gráficas 11, 12 y 13).

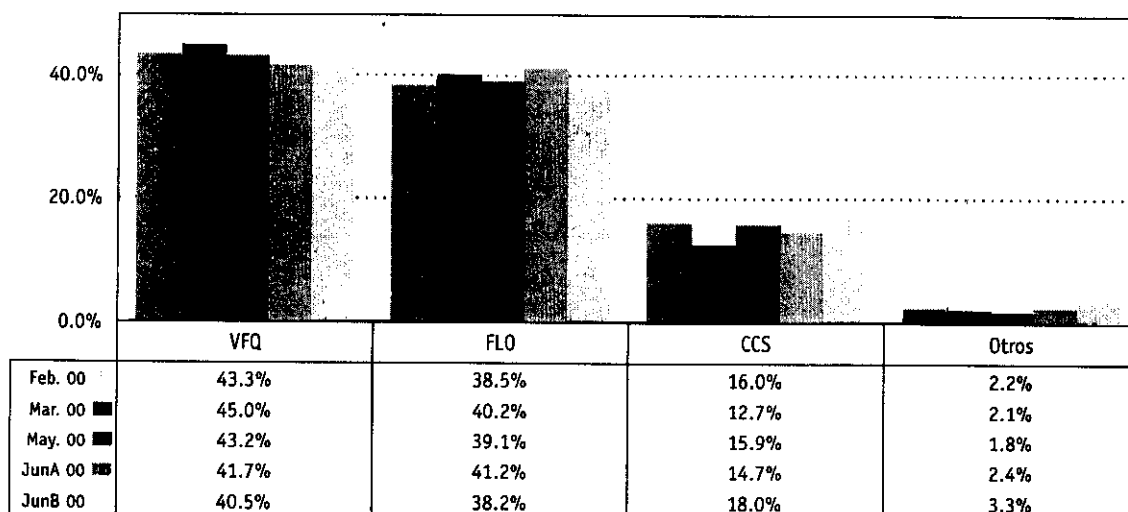
Más correcto es efectuar la comparación entre las proporciones de votantes seguros por contendiente res-

GRÁFICA 11
INTENCIÓN MANIFIESTA DE VOTO POR CANDIDATOS PARA LA ELECCIÓN DE PRESIDENTE DE LA REPÚBLICA (VOTANTES SEGUROS)*



* Respuesta a la pregunta "Si en este momento se celebraran las elecciones para Presidente de la República, ¿por qué candidato votaría usted?" Solamente considera ciudadanos con credencial de elector al momento de la entrevista que se declaran seguros de ir a votar.

GRÁFICA 12
INTENCIÓN DEFINIDA DE VOTO POR CANDIDATOS PARA LA ELECCIÓN DE PRESIDENTE DE LA REPÚBLICA (VOTANTES SEGUROS)*

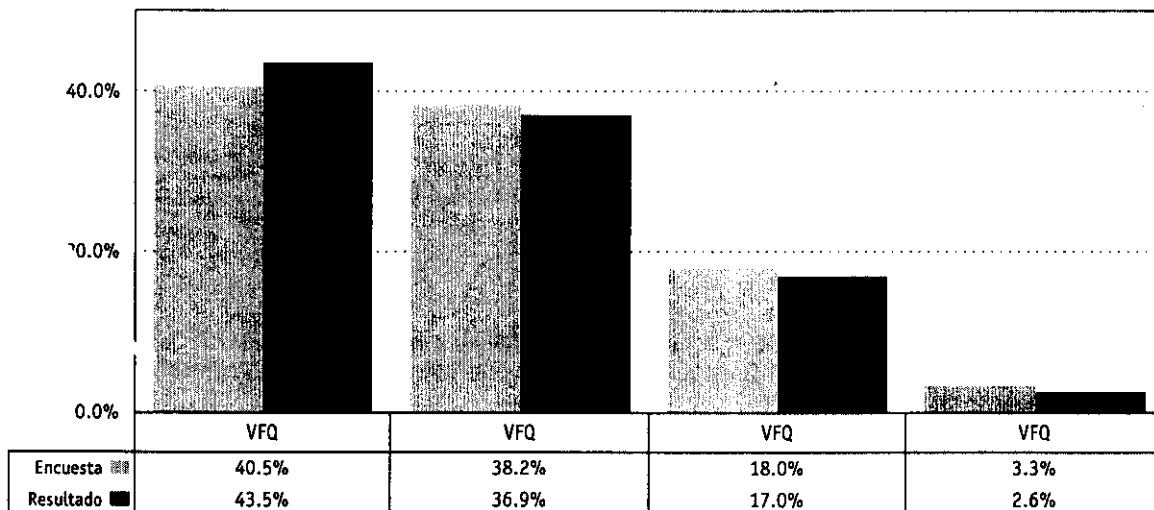


* Respuesta a la pregunta "Si en este momento se celebraran las elecciones para Presidente de la República, ¿por qué candidato votaría usted?" Solamente considera ciudadanos con credencial de elector al momento de la entrevista que se declaran seguros de ir a votar.

pecto al total de credencializados en la muestra con la proporción de votos obtenidos por dicho contendiente respecto al total de ciudadanos empadronados. Al efectuar este ejercicio, se encuentra que el giro medio entre estimaciones y resultado oficial es de sólo 2.3 por ciento (un error medio por grupo de 0.9 por ciento) (Ver Gráficas 13 y 14).

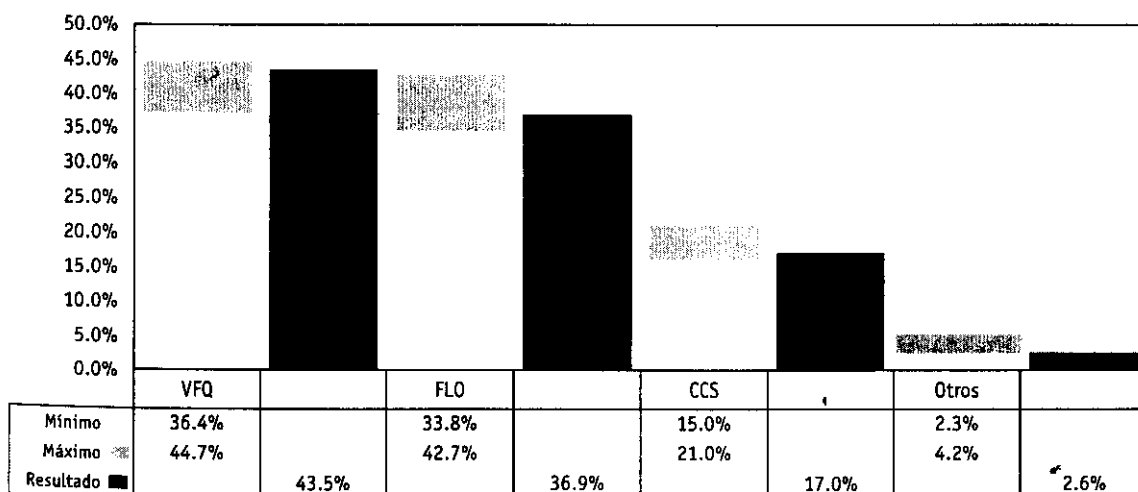
Si se revisa el intervalo de las estimaciones, se encuentra que los datos reales efectivamente se ubicaron dentro de estos intervalos tanto en el caso de cada uno de los principales contendientes como de la proporción de abstencionistas, lo que permite sustentar la hipótesis de la pertinencia de efectuar una decantación para ubicar a los votantes probables dentro del universo de los electores entrevistados.

GRÁFICA 13
INTENCIÓN DEFINIDA DE VOTO POR CANDIDATOS DE LOS VOTANTES SEGUROS*
Y RESULTADO OFICIAL DE LA ELECCIÓN PARA PRESIDENTE DE LA REPÚBLICA (RESPECTO A VOTANTES)



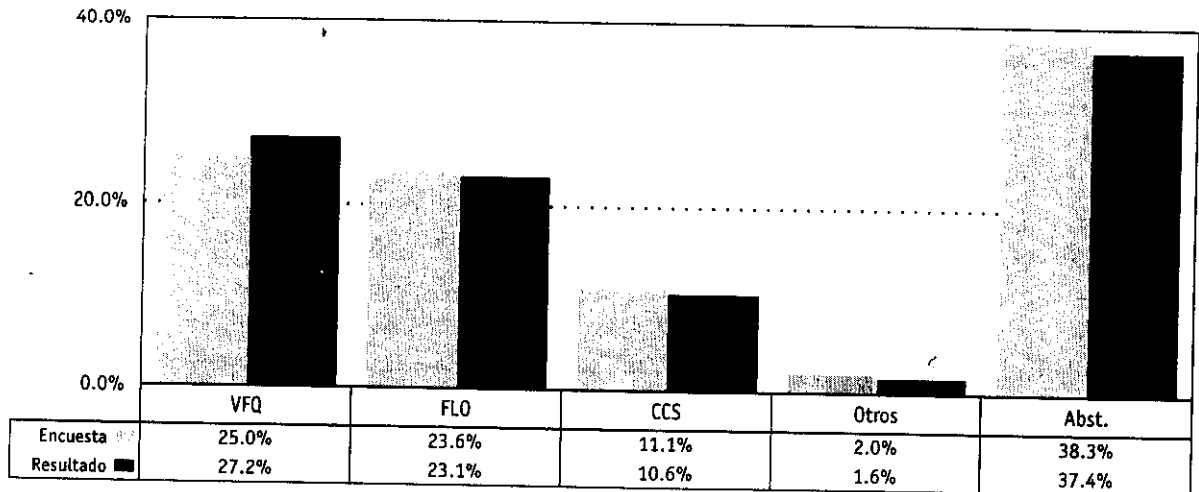
* Respuesta a la pregunta "Si en este momento se celebraran las elecciones para Presidente de la República, ¿por qué candidato votaría usted?" Solamente considera ciudadanos con credencial de elector al momento de la entrevista que se declaran seguros de ir a votar.

GRÁFICA 14
RANGOS EN LA ESTIMACIÓN DE LA PROPORCIÓN DE VOTANTES SEGUROS POR CANDIDATO*
Y RESULTADO OFICIAL DE LA ELECCIÓN PARA PRESIDENTE DE LA REPÚBLICA (RESPECTO A VOTANTES)



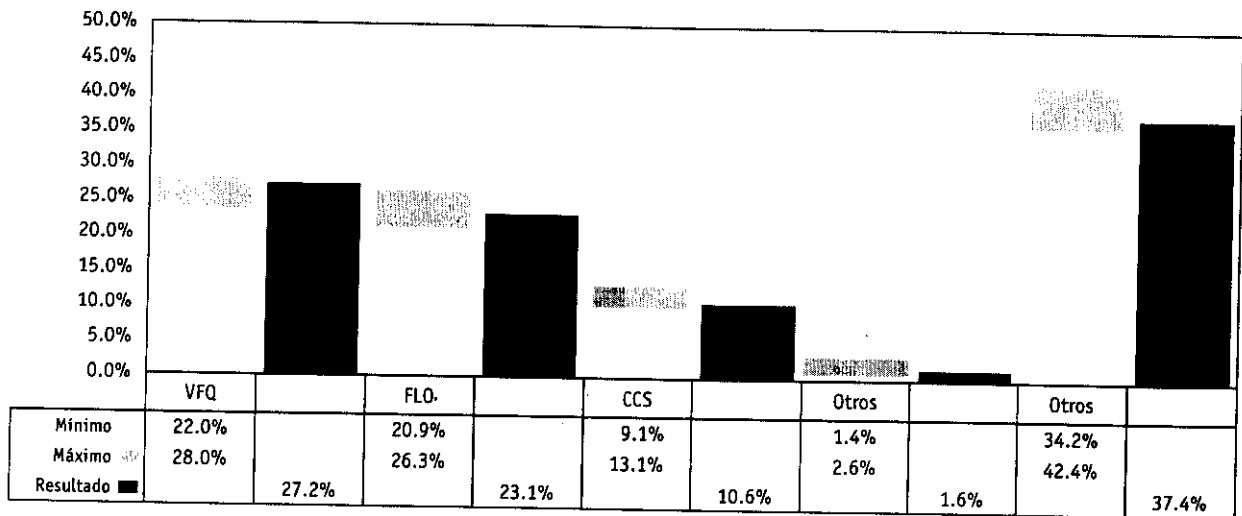
* Respuesta a la pregunta "Si en este momento se celebraran las elecciones para Presidente de la República, ¿por qué candidato votaría usted?" Solamente considera ciudadanos con credencial de elector al momento de la entrevista que se declaran seguros de ir a votar.

GRÁFICA 15
INTENCIÓN DEFINIDA DE VOTO POR CANDIDATOS DE LOS VOTANTES SEGUROS* Y RESULTADO OFICIAL DE LA ELECCIÓN PARA PRESIDENTE DE LA REPÚBLICA (RESPECTO A ELECTORES)



* Respuesta a la pregunta "Si en este momento se celebraran las elecciones para Presidente de la República, ¿por qué candidato votaría usted?" Solamente considera ciudadanos con credencial de elector al momento de la entrevista que se declaran seguros de ir a votar.

GRÁFICA 16
RANGOS EN LA ESTIMACIÓN DE LA PROPORCIÓN DE VOTANTES SEGUROS POR CANDIDATO* Y RESULTADO OFICIAL DE LA ELECCIÓN PARA PRESIDENTE DE LA REPÚBLICA (RESPECTO A ELECTORES)



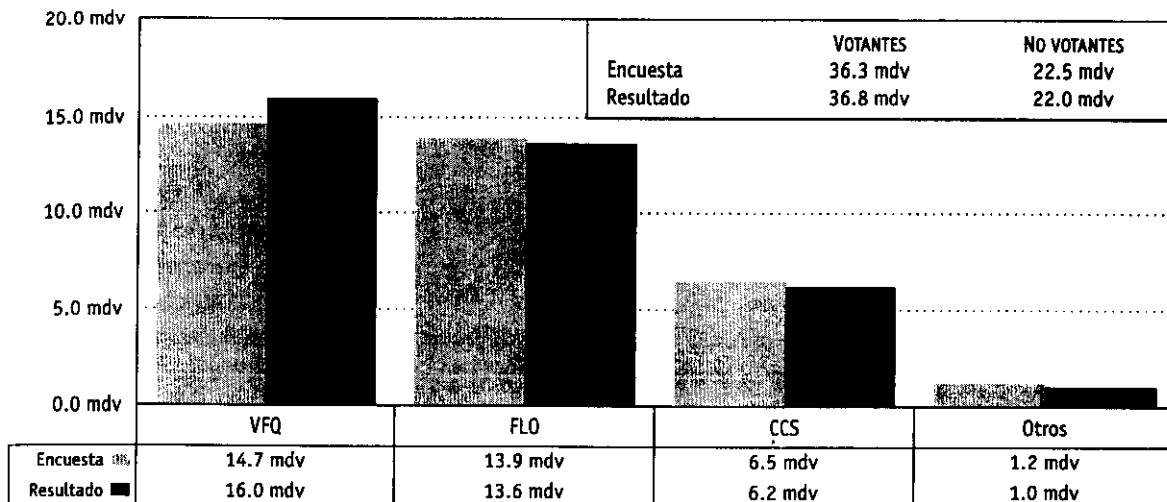
* Respuesta a la pregunta "Si en este momento se celebraran las elecciones para Presidente de la República, ¿por qué candidato votaría usted?" Solamente considera como votantes a los ciudadanos con credencial de elector que se declaran seguros de ir a votar.

Conforme lo previamente indicado, la proporción de votantes seguros por contendientes y de abstencionistas probables calculada con respecto al total de electores es utilizable para calcular volúmenes de votación esperable por contendiente. Este ejercicio redundaría en la correcta asignación de 35.5 millones de votantes, ubicando equívocamente como abstencionistas a medio millón de electores y existiendo un error en ubicación de 0.8 millones de votantes (Ver Gráficas 17 y 18).

Consecuentes con todo lo anterior, los volúmenes de votación obtenidos por cada uno de los contendientes se ubicaron, efectivamente, dentro de los rangos de votación absoluta estimados por la encuesta final de la serie, y la votación válida efectivamente emitida se sitúa dentro del rango estimado para la votación conforme la encuesta.

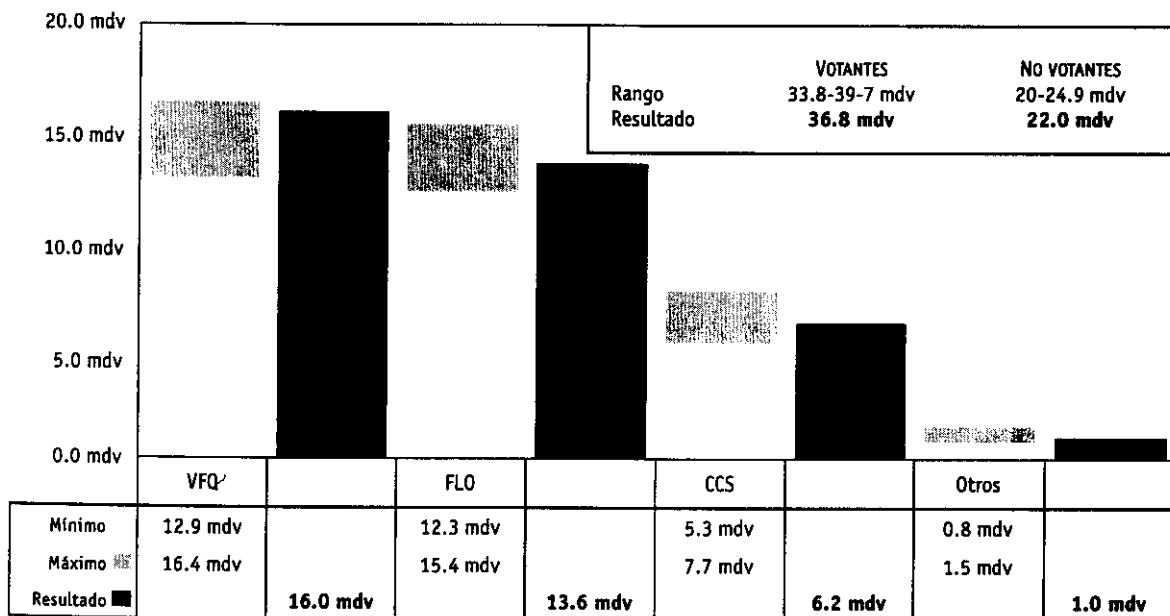
Si se traducen las proporciones de respaldo entre los votantes seguros a votos probables, Fox habría pasado de

GRÁFICA 17
ESTIMACIÓN DEL NÚMERO DE VOTANTES SEGUROS POR CANDIDATO* Y RESULTADO
OFICIAL DE LA ELECCIÓN PARA PRESIDENTE DE LA REPÚBLICA



* Respuesta a la pregunta "Si en este momento se celebraran las elecciones para Presidente de la República, ¿por qué candidato votaría usted?" Solamente considera como votantes a los ciudadanos con credencial de elector que se declararon seguros de ir a votar.

GRÁFICA 18
RANGOS EN LA ESTIMACIÓN DEL NÚMERO DE VOTANTES SEGUROS POR CANDIDATO*
Y RESULTADO OFICIAL DE LA ELECCIÓN PARA PRESIDENTE DE LA REPÚBLICA



* Respuesta a la pregunta "Si en este momento se celebraran las elecciones para Presidente de la República, ¿por qué candidato votaría usted?" Solamente considera como votantes a los ciudadanos con credencial de elector que se declararon seguros de ir a votar.

12.5 millones de votos posibles en febrero, a 15.5 millones para marzo, bajando a 14 millones los primeros días de junio, para alcanzar nuevamente los 15 millones a mediados de junio, ubicándose en la última medición de la serie poco más de un millón de votos por debajo del volumen efectivamente alcanzado en las urnas.

Ello implica la obtención por Fox de un respaldo adicional en los días previos a los comicios, de alrededor de tres por ciento de electores, la mitad provenientes de quienes decidieron votar de última hora y la otra mitad por giros en las preferencias, en detrimento de los niveles de otros candidatos.

Por su parte, Labastida habría pasado de 11 millones de votos probables en febrero, hasta alcanzar casi 14 millones a mediados de junio, perdiendo en los días finales alrededor de 300 mil votos potenciales, proporción similar a la que perdería Cárdenas entre el momento de la última medición y la jornada electoral.

De hecho, Fox habría logrado entonces el respaldo de alrededor de 44 por ciento de los electores que se definirían por votar entre febrero y el día de la elección, contra 31 por ciento que se habría inclinado por Labastida; 22 por ciento que lo habría hecho a favor de Cárdenas y tres por ciento por otros candidatos.

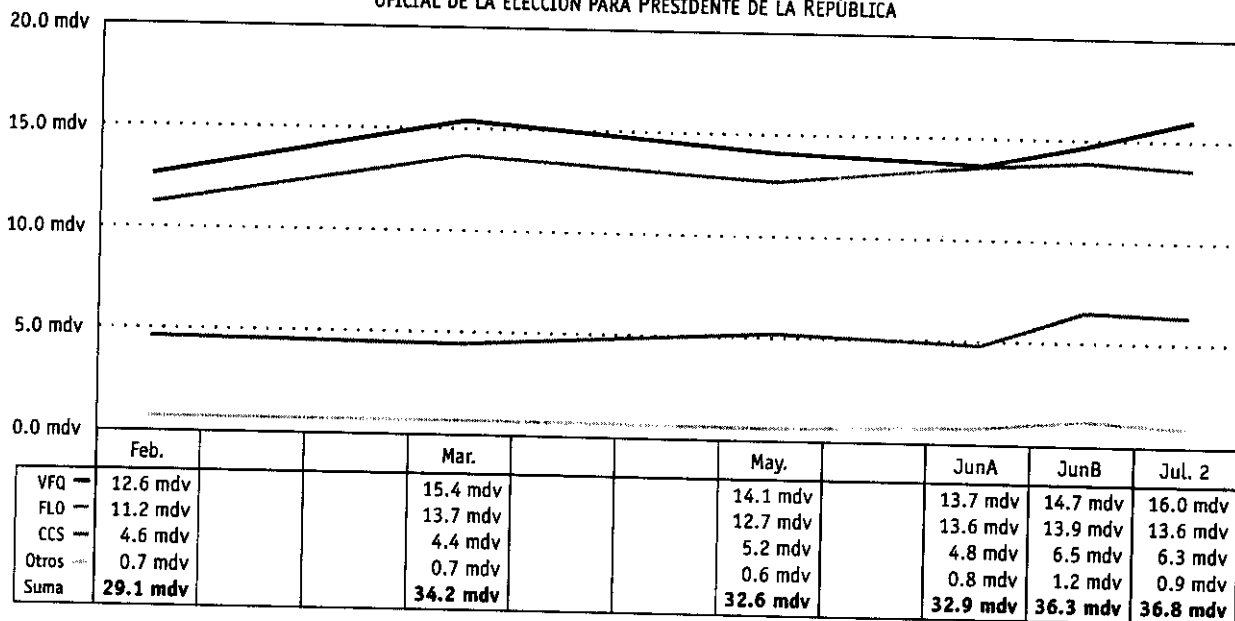
Así, Fox habría ido adelante desde el principio de la campaña oficial, en una contienda relativamente cerrada, donde prevalecieron tendencias históricas del electorado mexicano, reflejándose cambios en las preferencias que habían ocurrido desde el estallido de la crisis económica, en diciembre de 1994, cuando el PRI perdió más de diez

puntos de respaldo,²⁶ teniendo en julio de 2000 el electorado su primera oportunidad de votar para un relevo en el Ejecutivo Federal luego de esta crisis.

Cabe recordar que al cotejar datos de las elecciones intermedias con los resultados de la reciente contienda presidencial, se advierte que el PRI no perdió en estos tres años una proporción significativa de votos, simplemente, como Cárdenas lo había hecho en la capital; Fox logró concentrar una proporción importante del voto opositor, lo que hizo posible la derrota del partido gobernante.

El 2 de julio de 2000 fue así una postrer manifestación del reclamo ciudadano ante la frustración de sus expectativas de crecimiento sostenido, que en la elección presidencial de 1994 le habían permitido al PRI obtener un voto mayoritario. Estas expectativas se derrumbaron al inicio del actual sexenio, cuando estalló la peor crisis económica en siete décadas. Después, para muchos ciudadanos sólo fue cuestión de esperar una oportunidad para el cambio. ■

GRÁFICA 19
VOTANTES SEGUROS POR CANDIDATO DURANTE LA CAMPAÑA* Y RESULTADO OFICIAL DE LA ELECCIÓN PARA PRESIDENTE DE LA REPÚBLICA



* Respuesta a la pregunta "Si en este momento se celebraran las elecciones para Presidente de la República, ¿por qué candidato votaría usted?" Solamente considera como votantes a los ciudadanos con credencial de elector que se declararon seguros de ir a votar.

²⁶ Lo que es detectado oportunamente por la serie de encuestas que el Gabinete de Estudios de Opinión (GEO) realizara entonces para el semanario *Ércetera*, que detectarían una brusca caída de las preferencias por el PRI entre noviembre de 1994 y enero de 1995.

PONENCIAS MESA 1

Alejandro Moreno

PERIÓDICO REFORMA

Me alegra estar aquí por varios motivos. Primero, por compartir con todos los encuestadores y los interesados en este oficio las experiencias del trabajo que se hizo con motivo de las elecciones del 2 de julio. Segundo, me da gusto saber que buena parte de la discusión que ha generado la pregunta de si fallaron o no fallaron las encuestas gira en torno a una hipótesis que he propuesto y que me gustaría profundizar el día de hoy.

Ulises Beltrán mencionaba un artículo que publiqué hace un par de semanas en el semanario “Enfoque” del periódico Reforma, bajo el título “Un electorado imprevisible”. El artículo presenta las ideas generales acerca de la hipótesis de los votantes probables, bajo un punto de vista periodístico más que académico. El día de hoy me gustaría darle un poco más de evidencia dura a esa hipótesis, es decir, al argumento de los votantes probables.

Aprovecho también para mencionar que en la edición del artículo se editó también el título. Mi sugerencia original era titularlo “En busca del votante del 2 de julio”. No me gusta la idea de “un electorado imprevisible”, porque me da la impresión de que nos deja a los encuestadores como en una posición pasiva, como si algo sucede y no podemos hacer nada al respecto. Creo que eso no es exactamente lo que voy a mostrar hoy, que los encuestadores estamos en busca de un electorado, en busca del votante, de los individuos que definen las elecciones. Me parece que eso refleja más un esfuerzo que nos reta hacia adelante.

Quiero de antemano advertir que la búsqueda de los votantes probables en nuestras estimaciones de encuestas nos puede ofrecer mucho, pero ésta no es la solución a todos nuestros problemas. Quisiera dar evidencia en favor de la necesidad de dar estimaciones con base en votantes probables, pero también hacer ver algunos problemas.

Antes de entrar de lleno a la evidencia de los votantes probables, quisiera nada más poner algunas transparencias sobre el trabajo que hizo Reforma en cuanto a las encuestas electorales, sus detalles metodológicos y seguir, así, con el espíritu de la reunión. Este no es más que un cuadro que resume las vitrinas metodológicas que se fueron presentando en las publicaciones de Reforma, quizás con algunos datos adicionales. Estoy de acuerdo con Ulises Beltrán en que a veces las vitrinas metodológicas de las publicaciones de encuestas no nos dan suficiente información (Ver cuadro 1).

Creo que sería importante llegar a un punto de acuerdo entre nosotros para saber qué información adicional se necesita en las publicaciones de metodologías. Algunas de las cosas que estoy poniendo extra en este cuadro es, por ejemplo, la tasa de rechazos, mes con mes, que obtuvimos en nuestras encuestas.

Las encuestas que realizó Reforma son representativas de los adultos del país, por lo que sus resultados se pueden generalizar a nivel nacional. De manera muy rápida, pero estoy abierto a volver a este punto si fuese necesario, sobre todo para darle más tiempo al asunto de los votantes probables, voy a poner un mapa, sólo para darles una idea acerca de la dispersión de los puntos en donde se levantan las encuestas, y de dónde provienen nuestras estimaciones nacionales.

En Reforma en todo momento realizamos una simulación del voto –si me permiten llamarla así– a través del uso de urnas y boletas para tratar de dar no solamente un mayor realismo a esa acción de decir “¿por quién va a votar usted?”, sino también de garantizar el anonimato del entrevistado al indicar su intención de voto. Esta es una copia de la boleta que utilizamos, que es más o menos una réplica de la que utilizó el IFE el día de las elecciones, con la diferencia de que la nuestra lleva el logotipo de Reforma y la leyenda “no oficial”.

La boleta la utilizamos en las encuestas preelectorales y la usamos también en nuestra *exit poll*, o encuesta de salida, con la diferencia de que en esta última preguntamos en pasado: “¿por quién votó usted?” Otra diferencia es que las boletas preelectorales estaban impresas en color, mientras que las de la *exit poll* estaban en blanco y negro para no causar confusión o problemas con alguna autoridad electoral, con alguien que pudiera decir que nuestros encuestadores estaban manejando boletas extras. También utilizamos el mismo tipo de boletas en las encuestas para Jefe de Gobierno del Distrito Federal y en cualquier otra elección a nivel estatal. Todas las encuestas se hicieron, como se indica en el cuadro, de manera personal y en el domicilio de los entrevistados seleccionados. (Ver cuadro 2).

El tamaño de la muestra en cada una de las siete encuestas nacionales realizadas de enero a junio, dos de ellas en mayo, fue de alrededor de mil 500 entrevistados. La segunda encuesta de mayo se hizo para cubrir la necesidad de información en torno al segundo debate presidencial, dado el vacío informativo que hubo en el primer debate, el cual no lo teníamos previsto nosotros ni ninguna otra empresa encuestadora.

El plan de las encuestas mensuales se hizo con anticipación. El contenido de las preguntas estaba definido de antemano, para dar un seguimiento riguroso a las opiniones, salvo por un espacio que dejábamos en el cuestionario para algunas preguntas coyunturales. En las siete encuestas tenemos un total aproximado de 250 variables. Obviamente no todas se preguntaban en un solo cuestionario, pero sí teníamos una batería de preguntas que se hacían a lo largo del tiempo, sin cambiarlas, y también había preguntas que teníamos contempladas cada tres meses, o un mes sí

y un mes no. En general, lo que tenemos es básicamente una recopilación de datos que, al final, parece ser bastante valiosa, sobre todo para tratar de darle evidencia a varias preguntas relacionadas con el comportamiento y las actitudes del electorado en general y, en particular, para probar la hipótesis de los votantes probables.

En el artículo al que se hacía referencia, mi argumento básico es que si filtramos por votantes probables, que en este caso son aquellos que nos daban un *score* alto, en una escala de 10 puntos de probabilidad de voto y que también se mostraban interesados en el proceso de las campañas políticas.

El seguimiento de la intención de voto por cada uno de los candidatos, como la publicamos, se basaba en la totalidad de la muestra, recalculando porcentajes sin considerar a quienes no nos respondían por quién votarían, quienes dijeron que no votarían y quienes no contestaron. Generalmente a este subgrupo se les llama “indecisos”, pero una y otra vez he argumentado que no son necesariamente indecisos, solamente son esa proporción que dice “no sabe, no contestó”. Puede ser que sean indecisos pero también puede ser que sean personas que no vayan a votar. La idea de “indecisos” es alguien que creemos que va a votar, pero no nos está diciendo por quién, porque aún no lo sabe. No todo el que nos dice que “no sabe” o “no contestó”, o no nos marca la boleta, es necesariamente un “indeciso”.

Si comparamos nuestros resultados de encuesta con los resultados oficiales para comparar la efectividad observamos que la submuestra de votantes probables –con base a los criterios antes mencionados– ajusta el resultado de la encuesta mucho más a lo que fue el resultado de la elección que la muestra total. El error promedio de la encuesta llega a bajar hasta 1.5 por ciento. Ese fue el caso también de las encuestas a nivel estatal que hicimos en Guanajuato y Morelos.

Otra cosa que argumentaba en el artículo tan mencionado, es que, si comparamos la submuestra de los votantes probables en las encuestas preelectorales con las características de los votantes reales de acuerdo con el *exit poll*, la composición demográfica, política, e incluso ideológica en algunos casos, es mucho más similar que si comparamos el *exit poll* con la muestra preelectoral total. En otras palabras, el parecido entre la submuestra de votantes probables y el *exit poll* es mucho mayor que entre la submuestra de votantes probables y la muestra total de esa misma encuesta. Esto nos da un poco más de evidencia de que efectivamente los votantes probables con una personas mucho más similares a las que de hecho votan en la elección.

Es el mismo caso para las estimaciones estatales. El error, por ejemplo, en el caso de Morelos, que fue nues-

tra encuesta más certera del proceso del 2 de julio, también baja si se toman en cuenta los votantes probables. Lo mismo en el caso de Guanajuato. El caso del Distrito Federal es un poco más complejo, de un mayor análisis. Tanto en Morelos como en Guanajuato, el resultado es mucho más cercano al oficial filtrando por estos los criterios de votante probable. No realizamos *exit polls* en Morelos y Guanajuato, desafortunadamente, para poder hacer la comparación de perfiles.

A propósito de esta reunión, desarrollamos un modelo de regresión logística con distintas variables independientes, en donde la variable dependiente es precisamente el votante probable, es decir, este filtro, este estimador que estoy haciendo y que aparentemente por los puros cruces de variables me da una mejor estimación de la elección.

El modelo es muy exploratorio, es bastante simple. Incluye variables sociodemográficas como la región del país donde se hizo la encuesta, algunas otras demográficas, tales como ocupación, si es profesionista, gerente, oficinista, trabajador manual especializado, urbano, agrícola, ama de casa, estudiante, también tengo algunas variables de identificación partidista. De hecho resulta que los muy partidistas son más probables de votar que quienes no son tan partidistas, y mucho más probables de votar que quienes no tienen una identificación partidaria.

También incluimos algunos termómetros de opinión para ver qué tanto la opinión acerca de los candidatos afecta o deja de afectar las probabilidades de voto; la exposición a medios de comunicación, qué tanto se siguen las noticias en radio, en televisión, en los periódicos, platicando con la gente. Si ven más o menos el seguimiento de todos los meses: desde enero hasta junio, ésta es una de las variables consistentemente más significativas en explicar o predecir quiénes son los votantes probables.

Me parece que aquí hay un hallazgo muy importante también: qué tanto la gente percibía que su candidato iba a ganar o iba a perder. Teníamos a lo largo de la serie excepto en enero, una escala del 1 al 10, en donde el 10 significaba que la gente creía que su candidato definitivamente va a ganar, es decir, lo veían como el ganador de la elección y 0 que definitivamente no va a ganar, es decir, que no tienes chance de nada. Esta es la llamada probabilidad subjetiva de triunfo.

Entre otras variables, en la regresión también incluimos la escolaridad y una variable que indica que las entrevistas se hicieron en zonas urbanas con referencia a las rurales. Básicamente ese es el modelo.

Corrimos la regresión mes por mes y también una a nivel agregado con todas las variables, un *pooled dataset*, en el cual tenemos casi 10 mil 500. En el análisis quedan, para el caso de la regresión con los datos juntos, alrededor de unos seis mil 500 más o menos para quienes estén

interesados en todo esto y de lo más interesante que resulta aquí, es que no solamente el interés o el seguimiento de las campañas explican una mayor propensión a votar, sino que también la probabilidad de triunfo: 1 quienes más probables eran de salir a votar con un signo claramente positivo, eran quienes creían que Fox iba a ganar; y quienes menos probabilidad tenían eran quienes creían que Labastida iba a ganar. ¿Cómo se puede interpretar esto? Muchos de quienes estaban seguros del triunfo de Labastida, simplemente no eran probables de salir a votar.

Quisiera agregar dos cosas más. Si calculamos la probabilidad promedio para los grupos partidistas, es evidente que, al menos en junio (a lo largo de la serie se cumple, pero en junio es mucho más evidente) que aquellos que se consideraban como muy panistas en la encuesta, tenían 0.7 de probabilidad de salir a votar contra 0.6 de los muy priístas, es decir, una probabilidad yo no diría nada más ligeramente más alta, sino mucho más alta. Incluso los muy perredistas tenían 0.7 de probabilidad y los algo perredistas 0.6. Lo interesante es que aquellos que no son partidistas tenían una probabilidad promedio de 0.4 y los indecisos de 0.36, aquellos otra vez mal llamados "indecisos" que nos dicen "no sabe, no contestó". Entonces sí hay, de hecho, una propensión mayor o menor a votar de acuerdo con las características de las personas. Las más partidistas son los que más votan.

Concluyo con un par de cosas. ¿A qué nos lleva todo esto? Bueno, si es cierto este argumento de los votantes probables, entonces qué pasa con el resto de la serie de Reforma, es decir, cómo se hubiera visto la elección si aplicamos estos filtros a lo largo de la serie de Reforma que fue publicada sin filtros. ¿Hubo cruzamiento de preferencias? ¿Cuándo se dio el cruzamiento de primer lugar? ¿Hasta cuándo fue ganando Labastida y cuándo comenzó a ir perdiendo? Creo que esta es una pregunta bastante sensata, aunque sí con la precaución de que no necesariamente los filtros de probabilidad de voto del último mes, son aplicables a los meses anteriores. Conforme vamos avanzando de enero a junio, la proporción de votantes probables va aumentando y eso se debe básicamente a la variable de interés que incrustamos no necesariamente de probabilidad de voto y esto nos causa un poquito de ruido. Pero parece normal que aumente el interés a lo largo de una campaña.

Entonces ¿cuál es la imagen de la elección si hubiéramos ido con votantes probables desde un principio? En la serie publicada por Reforma que ustedes conocen todo el movimiento de preferencias hasta llegar a nuestra última estimación, la de 42 por ciento Labastida contra 39 por ciento Fox, y 16 por ciento Cárdenas, filtrando por probables votantes es otra historia. En enero, Labastida 50, Fox 35; febrero, Labastida 44, Fox 42; mar-

zo, Labastida 44, Fox 42; abril, Labastida 42, Fox 45, esta es la encuesta donde nuestra publicación se redujo a 3 puntos, de hecho nuestra primera encuesta de mayo, Labastida 37, Fox 43, un resultado muy parecido al final; en la segunda de mayo, 39 Labastida, 43 Fox; y en la última de junio, 39 Labastida, 43 Fox. Estos datos reflejan mucho más lo que al final fue el resultado del 2 de julio.

Ahora, con esto concluyo, nada más quisiera enfatizar el parecido con algo que no se publicó en Reforma,

pero que sí seguimos como control nuestro, un control que finalmente no utilizamos adecuadamente en la estimación final, encuestas nacionales telefónicas en las que Fox, siempre desde enero viene arriba y en cuya última estimación tenemos 48 por ciento para Fox contra 35 por ciento de Labastida en encuestas telefónicas nacionales. Eso nos da una idea también del sesgo socioeconómico y de la composición de quienes finalmente decidieron el 2 de julio. Gracias. ■

CUADRO 1
RESUMEN METODOLÓGICO DE LA SERIE DE ENCUESTAS PREELECTORALES NACIONALES DE GRUPO REFORMA

Encuesta	Enero	Febrero	Marzo	Abril	Mayo 1	Mayo 2	Junio
Fecha de levantamiento	Ene 7-11	Feb 11-14	Mar 17-20	Abr 7-10/14-17	May 12-14	May 27-28	Jun 16-18
Fecha de publicación	Ene 17	Feb 21	Mar 27	Abr 24	May 18	Jun 5	Jun 22
Puntos de levantamiento	155	155	155	155	155	155	155
Número de ciudades (15,000 o más habitantes)*	75	90	73	77	83	83	81
Número de localidades (menos 15 mil habitantes)*	40	38	45	46	47	47	42
Municipios	-	117	117	121	121	122	118
Delegaciones	-	8	8	7	9	9	11
Tamaño de muestra	1,544	1,510	1,533	1,647	1,547	1,543	1,545
Número de entrevistas por punto	10	10	10	10	10	10	10
Margen de error teórico	+/-2.5%	+/-2.5%	+/-2.5%	+/-2.5%	+/-2.5%	+/-2.5%	+/-2.5%
Nivel de confianza	95%	95%	95%	95%	95%	95%	95%
Tasa de rechazos	-	28%	26%	27%	29%	22%	33%
Número de entrevistadores	-	111	116	102	110	116	118
Número de supervisores	-	14	15	20	17	15	15

* En enero y febrero se consideraron ciudades de más de 10 mil habitantes y localidades de menos de 10 mil habitantes.
Fuente: Grupo REFORMA, Encuestas Nacionales Preelectorales de enero a junio de 2000.

CUADRO 2
ASPECTOS METODOLÓGICOS DE LA SERIE DE ENCUESTAS PREELECTORALES NACIONALES DE GRUPO REFORMA

Población de estudio	Mexicanos de 18 años o más en las 32 entidades federativas del país
Marco muestral	Lista de secciones electorales estratificadas por criterios urbano/rural/mixto del IFE
Método de muestreo	Probabilístico con selección en etapas que incluyen: 1) Puntos de levantamiento 2) Selección de viviendas 3) Selección de las personas entrevistadas
Tipo de entrevista	Personal en vivienda
Supervisión del trabajo de campo	En 3 etapas complementarias: 1) Supervisión directa durante la aplicación de cuestionarios en un recorrido aleatorio, en el que se contactan entre 25 y 30 por ciento de los encuestadores. 2) Supervisión después de la aplicación de cuestionarios, regresando al 20 y hasta el 25 por ciento de los puntos elegidos al azar. 3) Supervisión estadística por medio del análisis de resultados obtenidos, tomando como variable de control el número de encuestador.

Fuente: Grupo REFORMA, Encuestas Nacionales Preelectorales de enero a junio de 2000.

CUADRO 3

COMPARACIÓN DE LA MUESTRA TOTAL CON LA SUBMUESTRA DE VOTANTES PROBABLES Y LOS RESULTADOS OFICIALES

Si hoy fueran las elecciones, ¿por quién votaría usted?

	<i>Resultados brutos</i>	<i>Resultados efectivos</i>	<i>Resultados oficiales</i>	<i>Diferencia con el resultado oficial</i>
Muestra Total				
Fox (PAN-PVEM)	32	39	43	4
Labastida (PRI)	34	42	37	5
Cárdenas (PRD-PT-CD-PAS-PSN)	13	17	17	0
Otro	2	2	3	1
No contestó	19	-		
Error promedio				2.5
Submuestra de votantes probables*				
Fox (PAN-PVEM)	37	43	43	0
Labastida (PRI)	34	39	37	2
Cárdenas (PRD-PT-CD-PAS-PSN)	14	16	17	1
Otro	2	2	3	1
No contestó	13			
Error promedio	1			

Fuente: Grupo REFORMA. Encuesta Nacional Preelectoral de junio, n=1,545

*Filtrados por respuestas altas en la escala de probabilidad de voto e interés en las campañas.

CUADRO 4

ENCUESTA ELECTORAL DE MORELOS

COMPARACIÓN DE LA MUESTRA TOTAL CON LA SUBMUESTRA DE VOTANTES PROBABLES Y LOS RESULTADOS OFICIALES

SI HOY FUERAN LAS ELECCIONES, ¿POR QUIÉN VOTARÍA USTED?

	<i>Resultados brutos</i>	<i>Resultados efectivos</i>	<i>Resultados oficiales</i>	<i>Diferencia con el resultado oficial</i>
Muestra Total				
Sergio Estrada (PAN)	40	53	55	2
Juan Salgado (PRI)	24	32	27	5
Raúl Iragorri (PRD)	10	13	13	0
Otro	1	2	5	3
No contestó	25			
Error promedio				2.5
Submuestra de votantes probables*				
Sergio Estrada (PAN)	46	56	55	1
Juan Salgado (PRI)	24	29	27	2
Raúl Iragorri (PRD)	10	13	13	0
Otro	1	2	5	3
No contestó	19			
Error promedio				1.5

Fuente: Grupo REFORMA. Encuesta Preelectoral de junio, n=845

*Filtrados por respuestas altas en la escala de probabilidad de voto e interés en las campañas.

CUADRO 5
ENCUESTA ELECTORAL DE GUANAJUATO
COMPARACIÓN DE LA MUESTRA TOTAL CON LA SUBMUESTRA DE VOTANTES PROBABLES Y LOS RESULTADOS OFICIALES

Si hoy fueran las elecciones, ¿por quién votaría usted?

	<i>Resultados brutos</i>	<i>Resultados efectivos</i>	<i>Resultados oficiales</i>	<i>Diferencia con el resultado oficial</i>
Muestra Total				
Juan Carlos Romero Hicks (PAN)	40	51	57	6
Juan Ignacio Torres Landa (PRI)	32	42	34	8
Miguel Alonso Raya (PRD)	5	6	6	0
Otro	2	1	3	2
No contestó	21			
Error promedio				4
Submuestra de votantes probables*				
Juan Carlos Romero Hicks (PAN)	49	60	57	3
Juan Ignacio Torres Landa (PRI)	27	34	34	0
Miguel Alonso Raya (PRD)	4	4	6	2
Otro	1	2	3	1
No contestó	19			
Error promedio				1.5

Fuente: Grupo REFORMA. Encuesta Preelectoral de junio, n=796

*Filtrados por respuestas altas en la escala de probabilidad de voto e interés en las campañas.

CUADRO 6
COMPOSICIÓN PARTIDISTA DE LA MUESTRA TOTAL Y DE LA SUBMUESTRA DE VOTANTES PROBABLES COMPARADA CON LA ENCUESTA DE SALIDA

En general, ¿usted se considera priísta, panista o perredista?

	<i>Muestra total</i>	<i>Submuestra de quienes expresaron intención de voto</i>	<i>Votantes probables</i>	<i>Exit Poll</i>
Priísta	30	33	34	33
Panista	19	23	27	27
Perredista	9	10	12	12
Otro partido	1	1	1	1
De ningún partido	41	33	26	27

Fuente: Grupo REFORMA. Encuesta Nacional Preelectoral de junio n=1,545. Grupo REFORMA, Encuesta Nacional de Salida, Julio 2, 2000; n=3,380

CUADRO 7
COMPARACIÓN DE ENCUESTAS DE SALIDA NACIONAL CON LOS RESULTADOS OFICIALES

	<i>Encuestas de salida</i>	<i>Resultados Oficiales</i>	<i>Diferencia con el resultado oficial</i>
Vicente Fox	45	43	2
Francisco Labastida	36	37	1
Cuauhtémoc Cárdenas	17	17	0
Otro	2	3	1
Error promedio			1

Fuente: Grupo REFORMA, Encuesta Nacional de Salida. Julio 2, 2000; n=3,313.

CUADRO 8
ENCUESTA DE SALIDA DEL GRUPO REFORMA
 ¿Podría marcar por quién votó usted para Presidente de la República?

		<i>Fox</i>	<i>Labastida</i>	<i>Cárdenas</i>	<i>Otros</i>	<i>Porcentaje de la submuestra</i>
TOTAL		43	37	17	3	100
Tipo de localidad	Urbano	49	32	16	3	71
	Rural	30	50	19	1	20
	Mixto	35	45	19	1	9
Región	Norte	48	39	12	1	23
	Centro-occidente	46	39	12	3	17
	Centro	42	35	20	3	36
	Sur	42	26	20	2	24
Género	Hombre	46	33	19	2	52
	Mujer	42	41	14	3	48
Edad	18-24	48	34	17	1	18
	25-29	46	35	16	3	16
	30-34	48	34	16	2	15
	35-39	47	28	12	3	14
	40-44	40	37	20	3	11
	45-49	42	38	19	1	8
	50-54	44	41	13	2	6
	55-59	28	47	24	1	5
	60 o más	33	42	24	1	7
Escolaridad	No tiene estudios	28	48	22	2	8
	Primaria	32	48	19	1	28
	Secundaria/Técnica/equivalente	46	36	15	3	23
	Preparatoria/equivalente	50	30	17	3	18
	Universidad o más	58	24	15	3	23
Lugar de trabajo	Sector Público	40	28	19	3	19
	Sector Privado	52	32	14	2	27
	Por cuenta propia	41	37	20	2	23
	No trabaja	41	41	16	2	28
	No sabe/no contestó	48	34	16	2	3
Interés en las campañas	Mucho	49	33	16	1	48
	Algo	42	39	16	3	30
	Poco	36	41	20	3	17
	Nada	34	46	18	3	4
	No sabe/no contestó	18	54	28	0	1
Razón principal del voto	Por un cambio	66	15	18	1	42
	Por el candidato	27	50	19	4	9
	Por obligación	28	60	12	0	3
	Por costumbre	11	84	5	0	7
	Por ser el menos malo	35	41	21	3	4
	Por lealtad al partido	9	79	10	2	5
	Por las propuestas	36	43	17	4	23
Otra	33	44	21	2	7	
Cuándo decidió su voto	Hoy	43	37	18	2	6
	Última semana	44	35	15	6	5
	Último mes	45	35	16	4	8
	Últimos 3 meses	53	27	15	5	14
	Desde antes	42	39	18	1	66
No sabe/no contestó	33	53	11	3	1	

Fuente: Grupo REFORMA, Encuesta Nacional de Salida del 2 de Julio, n=3,313.

Nota: Los datos en esta tabla están ponderados por el resultado oficial de la elección.

		<i>Fox</i>	<i>Labastida</i>	<i>Cárdenas</i>	<i>Otros</i>	<i>Porcentaje de la submuestra</i>
Aprobación al Presidente Zedillo	Aprueba mucho	22	70	6	1	21
	Aprueba algo	44	39	15	2	45
	Desaprueba algo	59	14	24	3	16
	Desaprueba mucho	61	8	29	2	13
	No sabe/no contestó	44	30	22	4	5
Evaluación de la situación económica del país	Ha mejorado mucho	24	65	8	3	4
	Ha mejorado algo	31	57	11	1	20
	Sigue igual	43	36	18	3	45
	Ha empeorado algo	57	21	20	2	15
	Ha empeorado mucho	60	15	24	1	13
	No sabe/no contestó	39	45	11	5	2
Cree que habrá crisis económica	Sí	50	28	20	2	51
	No	38	47	13	2	33
Se ha beneficiado del "Progresá"	Sí	26	58	16	0	14
	No	47	34	17	2	84
	No está seguro	49	33	16	2	2
Identificación partidista	Muy priísta	8	91	1	0	20
	Algo priísta	11	86	2	1	13
	Muy panista	96	2	2	0	14
	Algo panista	93	5	2	0	13
	Muy perredista	10	1	87	2	7
	Algo perredista	12	6	80	2	6
	Otro	35	13	22	30	1
	Ninguno	54	20	20	6	23
No sabe/no contestó	47	32	18	3	3	
Auto-ubicación ideológica	Izquierda	50	17	32	1	21
	Centro-izquierda	54	19	24	3	11
	Centro	51	32	15	2	23
	Centro-derecha	59	34	6	1	9
	Derecha	30	57	12	1	36
Por quién votó en 1994	Ernesto Zedillo	24	67	7	2	53
	Diego Fernández	88	5	5	2	27
	Cuauhtémoc Cárdenas	29	6	62	3	19
	Otro	50	22	16	12	1
Por quién voto en las elecciones internas del PRI en 1999	Francisco Labastida	14	80	5	1	24
	Roberto Madrazo	52	32	14	2	7
	Manuel Bartlett	55	35	10	0	1
	Humberto Roque	72	20	8	0	1
	No votó	56	19	22	3	59
	No sabe/no contestó	29	49	19	3	8
Pertenece a un sindicato	Sí	40	39	18	3	14
	No	45	36	17	2	85
	No sabe	36	38	13	13	1
Religión	Católica	45	37	16	2	86
	Cristiana, no católica	34	40	23	3	6
	Otra	42	38	19	1	3
	Ninguna/no contestó	35	31	30	4	5
Frecuencia con la que asiste a servicios religiosos	Más de 1 vez por semana	41	42	16	1	13
	Una vez por semana	45	38	15	2	36
	Una vez al mes	44	36	18	2	20
	Sólo ocasiones especiales	46	35	16	3	21
	Nunca	40	26	30	4	8
	No sabe/no contestó	25	47	19	9	2