



El papel de las encuestas en el proceso electoral



Javier Márquez Peña

Retos al INTERPRETAR encuestas electorales en México

- **Autocrítica:** La mayoría de las encuestadoras reporta márgenes de error asumiendo muestreo aleatorio simple

El margen de error MUESTRAL es mayor de lo que creemos.

- **Muestreo por conglomerados:**

- 10 ciudadanos que viven en la misma sección electoral no representan la misma diversidad que 10 individuos seleccionados aleatoriamente.
- Ajustes con ponderadores también aumentan el error
- El tamaño de muestra “real” de una encuesta de **1,000** casos es de entre **500** y **600** para propósitos de margen de error

Ejemplo: encuestas en abril

(Sólo efectos por conglomeración y ponderación; no incluye estratos)

(No todas las encuestadoras reportan al INE la información para poder calcular apropiadamente el margen de error).

Estas encuestas de abril destacan por su transparencia:

	Proporción	ME Supuesto	ME Real	Diferencia
Berumen (5 de abril)				
R. Anaya	26	2.4	3.8	-1.4
J. A. Meade	18	2.1	3.8	-1.7
A. M. López Obrador	35	2.6	4	-1.4
M. Zavala	4	1.1	1	0.1
No declara	17	2.1	3.6	-1.5
Consulta (15 de abril)				
R. Anaya	21	2.5	5.7	-3.2
J. A. Meade	17	2.3	7.7	-5.4
A. M. López Obrador	32	2.9	10	-7.1
J. Rodríguez	2	0.9	1.6	-0.7
M. Zavala	4	1.2	4.2	-3
No declara	25	2.7	9.2	-6.5
Reforma (30 de abril)				
R. Anaya	25	2.5	4.8	-2.3
J. A. Meade	14	2	3	-1
A. M. López Obrador	39	2.8	5	-2.2
J. Rodríguez	1	0.6	0.8	-0.2
M. Zavala	2	0.8	1.2	-0.4
No declara	18	2.2	3	-0.8

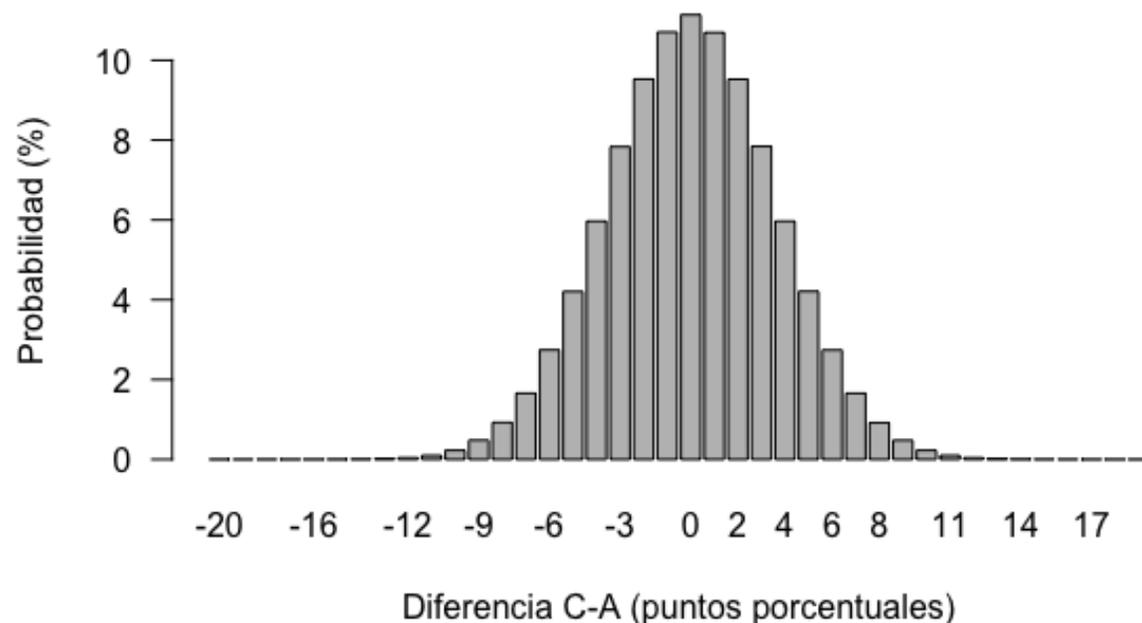
Consecuencias: Ilusión de la brecha entre dos candidatos

Problemas de las “brechas”: su error muestral es **aproximadamente el doble** del error al que estamos acostumbrados

Ejemplo: si el error muestral es ± 3.5 pts, el error muestral de la brecha es aproximadamente ± 7 pts.

Ejemplo: **A=20%, B=20%, C=20%**, D=10%, E=5%, F=5%, No Declara=20%

$n = 1,000$; $d_{eff} = 1.8$



La probabilidad de obtener a través de una encuesta la diferencia real (cero) es menor a 12%.

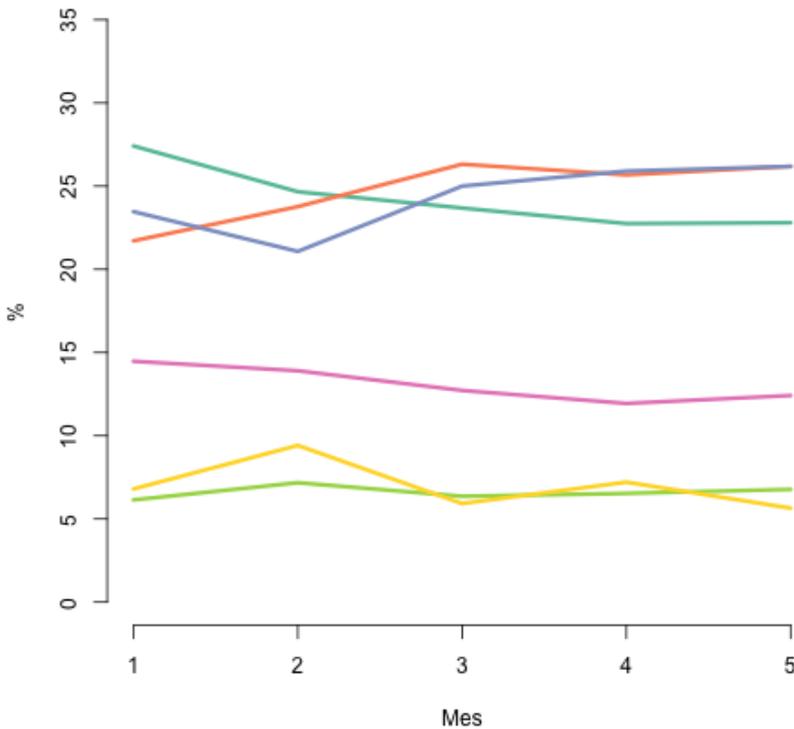
7 pts a favor de **A** o 7 pts a favor de **C** son igual de probables!!!

Consecuencias: interpretación errónea de “tendencias”

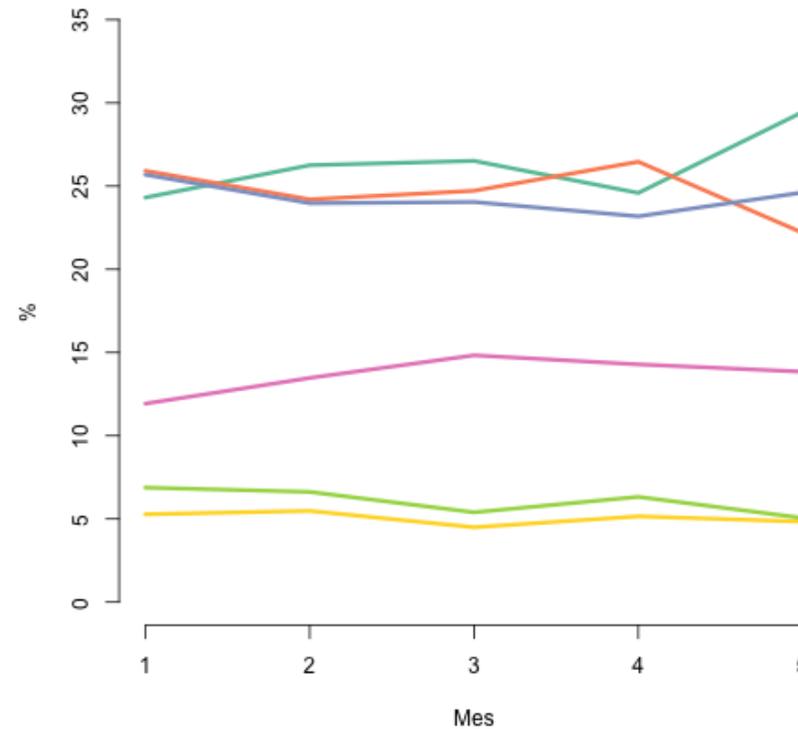
A todas estas encuestas sintéticas les subyace una preferencia (**efectiva**) verdadera de 25% para los principales tres candidatos.

Las variaciones se deben a error muestral (la variación sería mayor si consideramos el error no muestral)

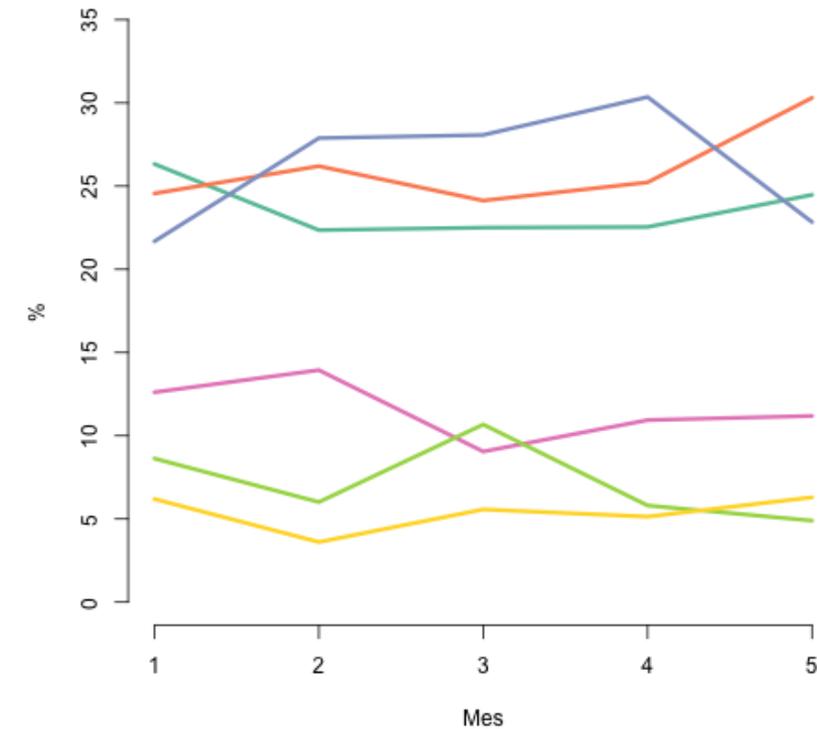
Encuestadora 1



Encuestadora 2



Encuestadora 3



No respuesta

- **Autocrítica:** Las encuestas reportan tasa de **RECHAZO**.
- La tasa de **NO-RESPUESTA** comprende tasas de CONTACTO, RECHAZO y COOPERACIÓN (*Disposition Rates*, AAPOR)
- La tasa de no-respuesta en México suele ser **mayor a 50%**.
- **Hemos podido replicar variables de control** de encuestas oficiales que tienen tasas de respuesta superiores a 80% y que tratan sobre percepciones:
 - Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública (ENVIPE)
 - Encuesta Nacional sobre Confianza del Consumidor

What's the Matter With Polling?

Cliff Zukin, NYT, Junio 20, 2015



Esta disminución (de la tasa de respuesta) es preocupante por dos razones. Primero, por supuesto, es la representatividad. Extrañamente, **por alguna razón que nadie realmente entiende**, las muestras probabilísticas bien hechas parecen haber conservado su carácter representativo a pesar de la escasa tasa de respuesta.

Sabemos esto porque podemos comparar los resultados que obtenemos de nuestras encuestas con los puntos de referencia del estándar de oro del gobierno, como la Encuesta de la Comunidad Estadounidense del Censo, donde la participación es obligatoria. Aun así, Robert M. Groves, ex director de la Oficina del Censo, advierte:

“El riesgo de que las encuestas fallen para reflejar los hechos aumenta con tasas de respuesta decrecientes. **El riesgo no siempre se cumple, pero con las muy bajas tasas de respuesta ahora comunes, deberíamos esperar más predicciones fallidas basadas en encuestas**”.



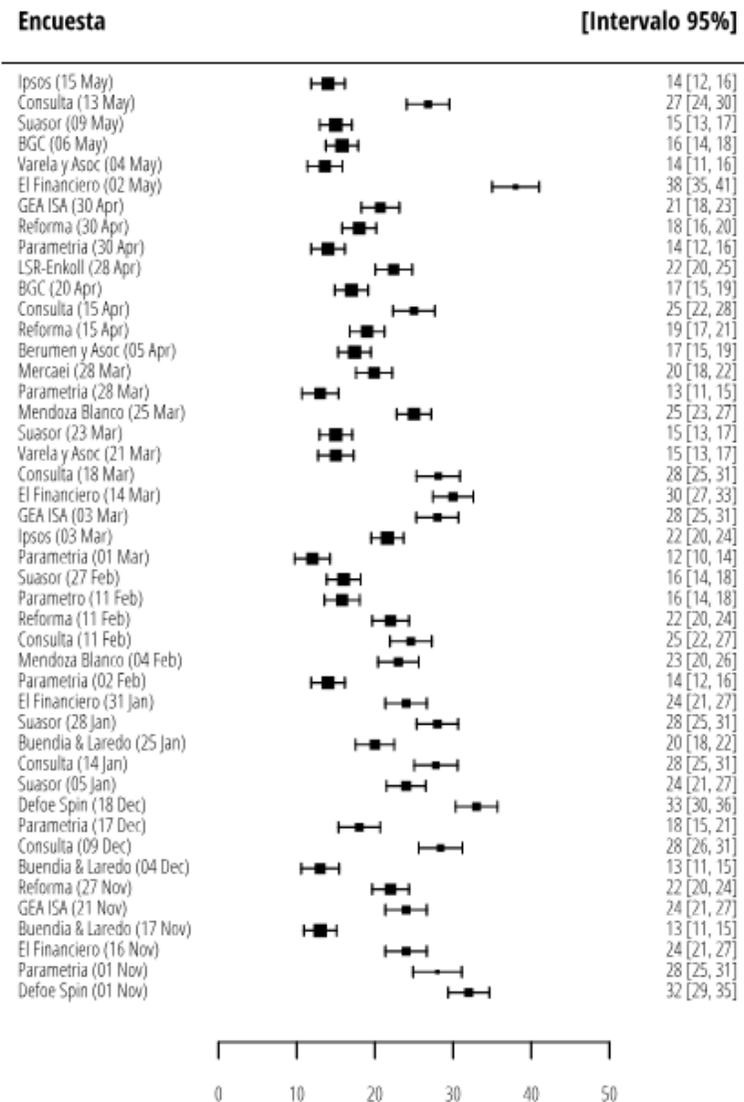
No respuesta al ítem

- **Autocrítica:** Los encuestadores usan términos diferenciados para referirse al mismo concepto.
- Cuando se pregunta a los entrevistados por cuál partido votarían si el día de hoy fuera la elección, un porcentaje importante ofrece respuestas como “por ninguno”, “no sé”, “no estoy seguro”, “no pienso ir a votar”, “el voto es secreto”, “no quiero decirle”, “todos son iguales”, “no creo en las encuestas”, y otras similares.

Términos (intercambiables) que los encuestadores usan para englobar a este grupo:

- No-respuesta (al ítem)
- No declara
- Indefinidos
- Indecisos

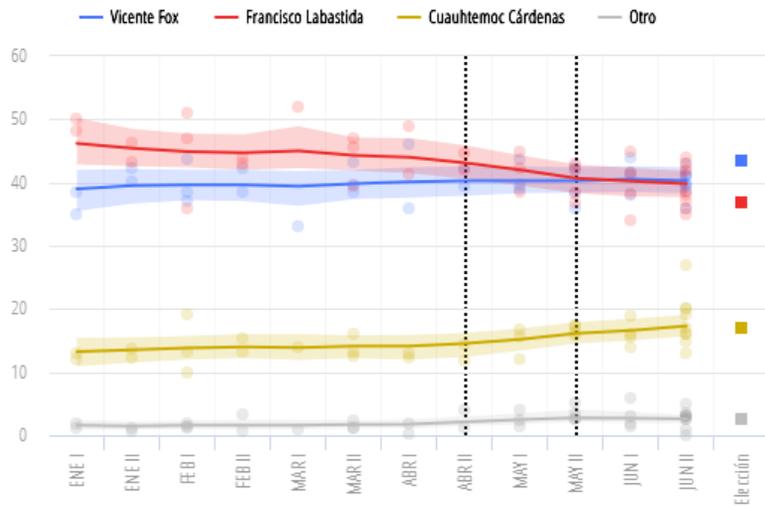
Tratamiento de la no-respuesta al ítem



- Los encuestadores **no “reasignan”** la no-respuesta en la preferencia efectiva.
- La denominada “preferencia efectiva” se trata de una **estimación para subpoblación** (la de quienes sí declararon preferencias).
- Aunque **aritméticamente son equivalentes** (pueden variar ligeramente por cuestiones de ponderación), estadística y conceptualmente son diferentes (e.g. cálculo de márgenes de error).
- En encuestas preelectorales, la **preferencia efectiva se emplea para hacer comparables los resultados de encuestas con distinta no-respuesta al ítem.**
- La no respuesta al ítem tiene un fuerte componente de “**efecto de casa**”.
- La preferencia efectiva **puede servir como estimación del resultado electoral.** Si es el caso, no(necesariamente) es porque el encuestador crea que la no-respuesta se distribuye aleatoriamente, sino porque parece ser un buen predictor del resultado electoral.
- Pueden existir otros modelos de reasignación de preferencias

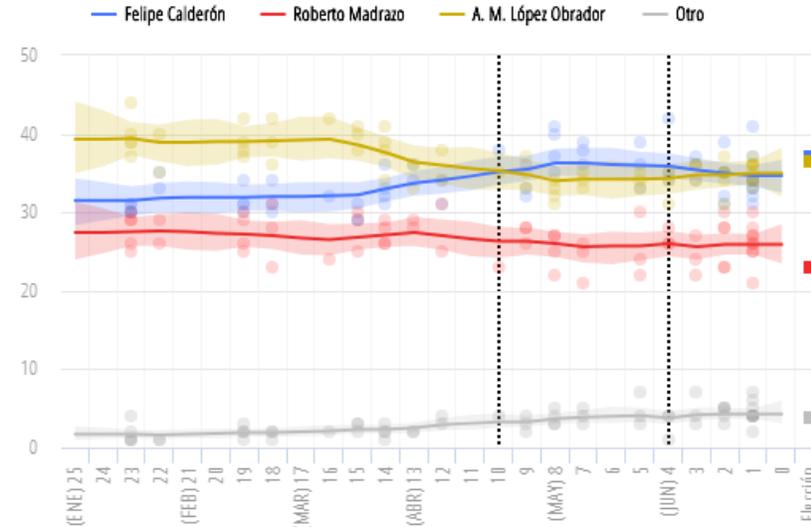
Poder predictivo de la preferencia efectiva

2000



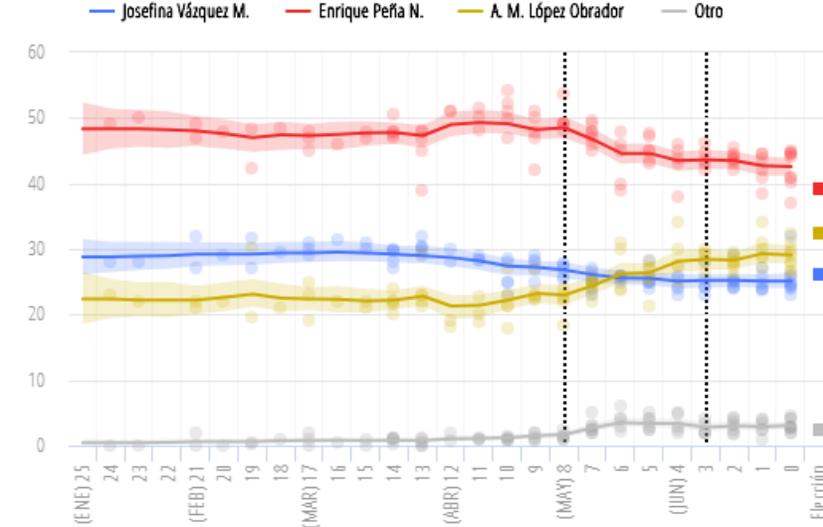
VFQ: -3.2
FLO: +2.9
CCS: +0.3

2006



FCH: -2.3
RMP: +3.0
AMLO: -1.4

2012



JVM: -1.0
EPN: +3.4
AMLO: +3.3

Población imaginaria antes de la elección



Población real

Ejemplo de población real:

Abstención	40%
Candidato A	20%
Candidato B	20%
Candidato C	15%
Candidato D	5%

- Ninguna encuesta en México captura intenciones de abstención en la magnitud que se observa el día de la elección.
- No contamos con **modelos validados de votantes probables** (no hay registros públicos)
- Con 20% de no respuesta al ítem vs 40% de abstención, podemos concluir que el problema **no sólo son los “indecisos”, sino los “decisos”** (*Molinar Horcasitas*)

¿Las encuestas son fotografías?

Si son fotografías de la población, parecen muy distorsionadas

No son pronósticos, pero sí buenos predictores (error promedio: 3.5 pts porcentuales en elecciones presidenciales, 2.4% en intermedias).

Hipótesis:

- **Sí son fotografías, pero existe problema de medición** (que sin pretenderlo, ayuda al encuestador). Preguntas de control se parecen a mediciones del INEGI.
- **No son fotografías, pero sin pretenderlo, el sesgo de selección se centra entre quienes sí van a votar**
- **No, pero sin pretenderlo, el sesgo de selección se centra entre una subpoblación similar a la de los votantes**

Cómo contrarrestar la imprecisión de las encuestas electorales para que los ejercicios demoscópicos sean útiles para la ciudadanía en la decisión de su voto.

“Cómo contrarrestar la imprecisión de las encuestas electorales”

- Voto obligatorio elimina problema de población imaginaria de votantes
- Registros públicos de quién votó y quién no

“..para que los ejercicios demoscópicos sean útiles para la ciudadanía en la decisión de su voto”

- Reportar mejores márgenes de error muestral y no muestral
- Estandarizar cálculos de no-respuesta
- Presentación de preferencias brutas y efectivas
- No sobredimensionar cambios/brechas
- Ofrecer panorama completo de los resultados de distintas encuestas