

# The Utility of Unpacking Survey Bias in Multiparty Elections: Mexican Polling Firms in the 2006 and 2012 Presidential Elections

Francisco Cantú (U of Houston)  
Verónica Hoyo (UC San Diego)  
Marco Morales (NYU)

# Evaluando la precisión de las encuestas

## Contexto

- ▶ Global:
  - ▶ Ejemplos recientes muestran la necesidad de estimar y explicar la precisión de las encuestas electorales
- ▶ Local:
  - ▶ 2012: "crisis de las encuestas"
  - ▶ No existe una medición sistemática para evaluar a las encuestadoras
    - ▶ por elección
    - ▶ por candidato

# Estimando la precisión de las encuestas

- ▶ La mayoría de los métodos existentes producen una estimación única
  - ▶ Mosteller (1949): **1-8**
  - ▶ Martin, Traugott & Kennedy (2005):  $A, A'_i$
  - ▶ Arzheimer & Evans (2014):  $B, B_w$

Year	Mosteller (1949)			Martin et al. (2005)			Arzheimer et al. (2014)	
	1	3	5	$\bar{A}'_{PRI}$	$\bar{A}'_{PAN}$	$\bar{A}'_{PRD}$	$\bar{B}$	$\bar{B}_w$
2006	-2.37	2.73	-3.88	.181	-.095	-.056	.129	.119
2012	4.05	3.07	1.76	.128	-.103	-.053	.140	.145

# Estimando la precisión de las encuestas

- ▶ Necesitamos desempacar las evaluaciones de precisión
  - ▶ Diferentes sesgos para cada candidato
  - ▶ Diferentes sesgos por encuestadora

# "Efectos de casa" 2006-2012

## Modelo

- ▶ **Propósito:** estimar los sesgos de cada casa encuestadora para cada candidato y elección
- ▶ Usamos todas las encuestas disponibles para distinguir el “ruido” de cada “señal”
  - i) “Señal”: la intención de voto para cada candidato durante la campaña electoral
  - ii) “Ruido”: el error sistemático de cada encuestadora por candidato
- ▶ **Datos:** Todas las encuestas publicadas durante las campañas presidenciales de 2006 y 2012
  - ▶ Información obtenida de los repositorios públicos del INE y de la AMAI

# "Efectos de casa" 2006-2012

## Modelo

- ▶ Modelo de filtros de Kalman (Jackman, 2005)
- ▶ Para cada encuesta  $k$ :

$$y_{i[k]} \sim \text{Binomial}(\mu_{ij[k]t[k]}, n_k) \quad (1)$$

- $y_{i[k]}$  número de encuestados apoyando al candidato  $i$
- $\mu_{ij[k]t[k]}$  apoyo reportado para el candidato  $i$  por la encuestadora  $j$  al momento  $t$
- $n_k$  tamaño de muestra

# "Efectos de casa" 2006-2012

## Modelo

### ► Estimación

$$\mu_{ijt} = \text{logit}^{-1}(\alpha_{it} + \delta_{ij}) \quad (2)$$

donde:

$\mu_{ijt}$  apoyo reportado para el candidato  $i$  por la encuestadora  $j$   
al momento  $t$

$\alpha_{it}$  el apoyo real para el candidato  $i$  al momento  $t$

$\delta_{ij}$  "efectos de casa" de la encuestadora  $j$  para el candidato  $i$

# "Efectos de casa" 2006-2012

## Modelo

- ▶ Estimación aleatoria inversa

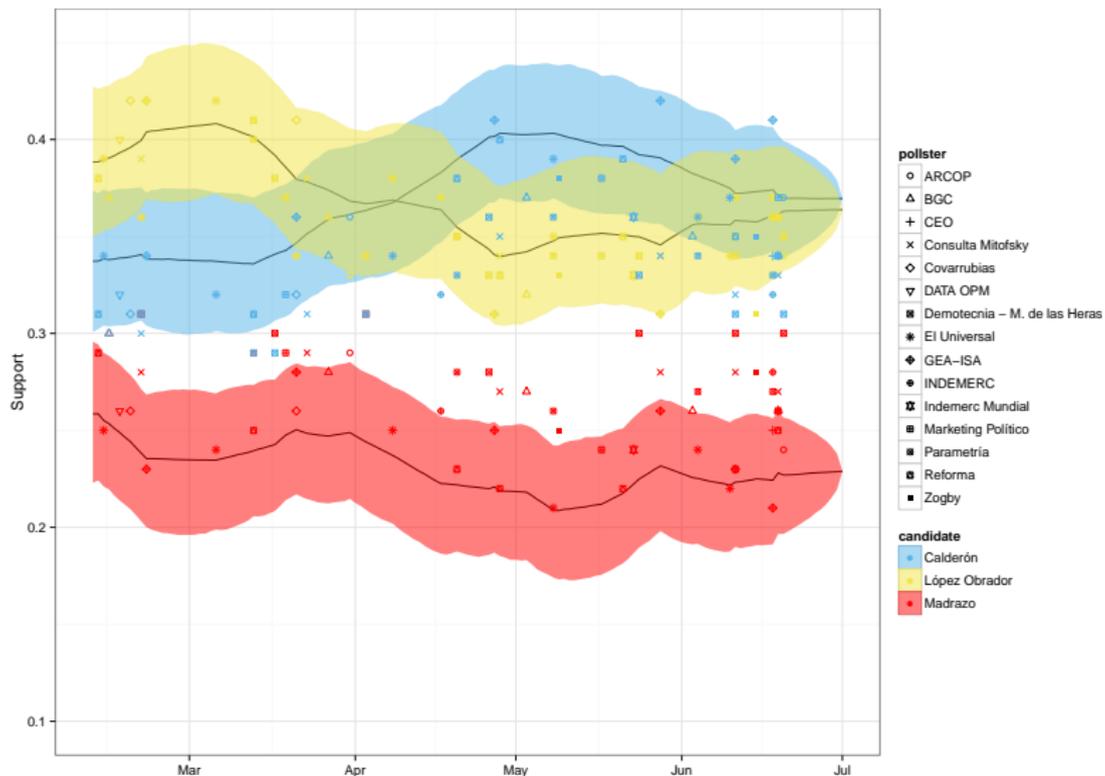
$$\alpha_{i,t} | \alpha_{i,t+1} \sim N(\alpha_{i,t+1}, \mathbf{s}_i^2) \quad (3)$$

- ▶ Supuestos

- ▶  $\alpha_{it}$  el día de la elección es la proporción de votos para cada candidato  $i$
- ▶  $\delta_{ij} \sim N(0, 0.05)$

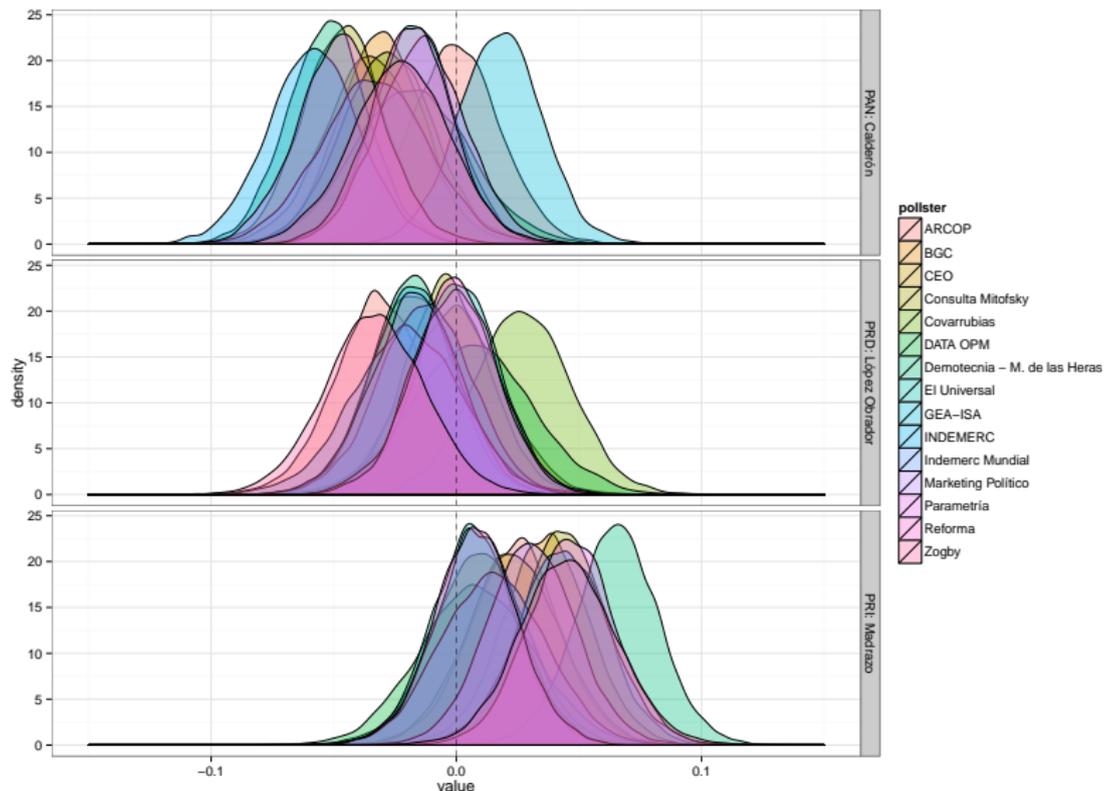
# Elección Presidencial 2006

Intención del voto estimado



# Elección Presidencial 2006

Estimación de "los efectos de casa" - intención de voto



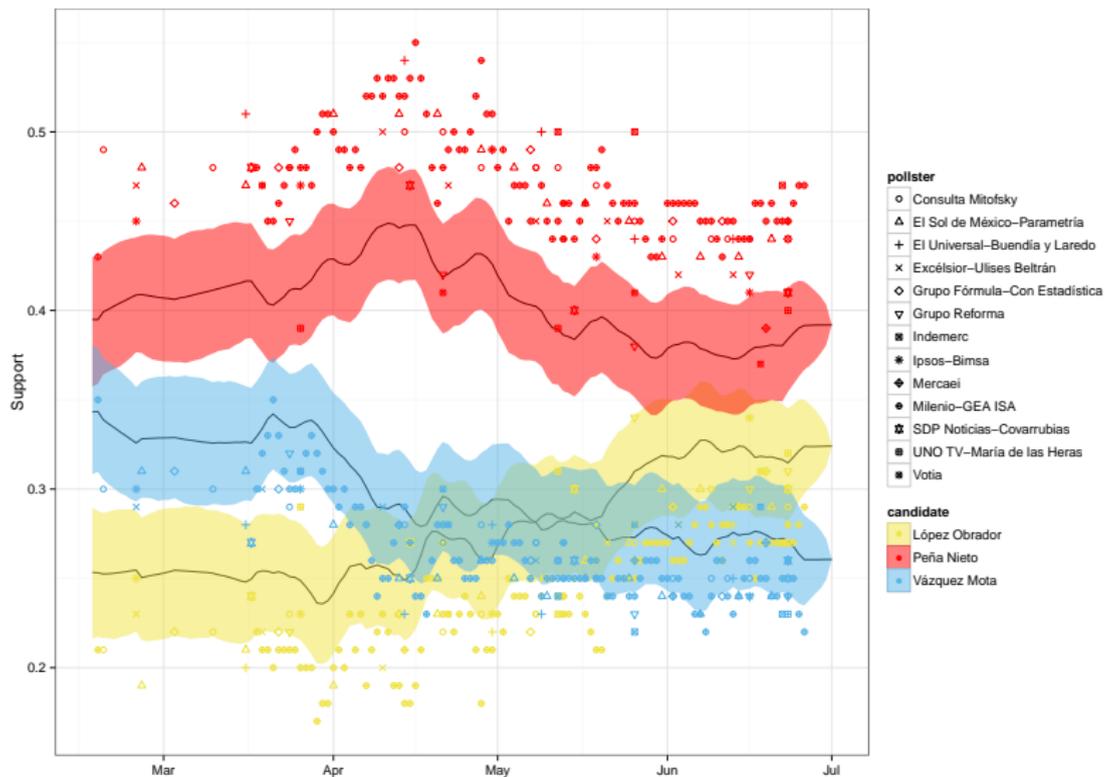
# Elección Presidencial 2006

Estimación de "los efectos de casa" - intención de voto

- ▶ 16 empresas publicaron 58 encuestas
- ▶ Todas las empresas **estimaron adecuadamente** a AMLO (16)
- ▶ La cuarta parte de las empresas **subestimaron** a FCH (4)
- ▶ Algunas empresas **sobreestimaron** a RMP (6)
- ▶ La mayoría de las empresas estimaron bien a todos los candidatos (9)

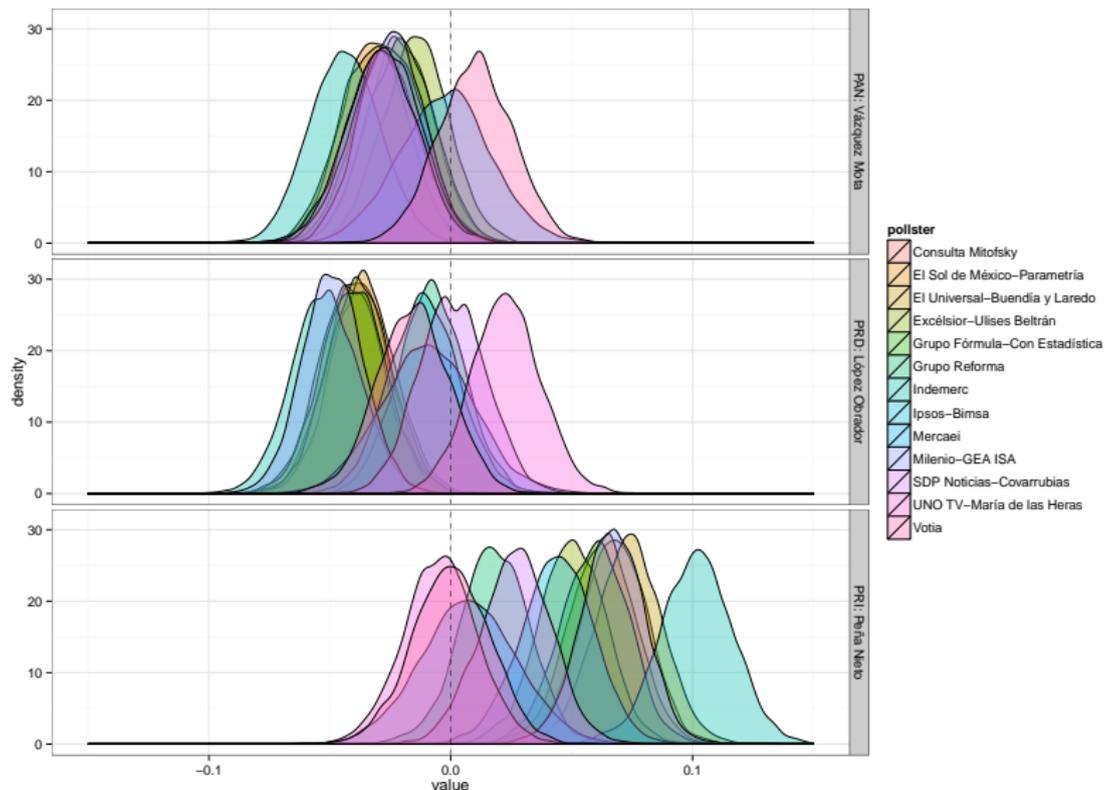
# Elección Presidencial 2012

## Intención del voto estimado



# Elección Presidencial 2012

Estimación de "los efectos de casa" - intención de voto



# Elección Presidencial 2012

Estimación de "los efectos de casa" - intención de voto

- ▶ 13 empresas publicaron 183 encuestas
- ▶ La mayoría de las empresas **sobreestimaron** a EPN (8) y **subestimaron** a AMLO (7)
- ▶ Algunas empresas **subestimaron** a JVM (3)
- ▶ Cinco empresas estimaron a los tres candidatos adecuadamente

# "Efectos de casa" 2006-2012

## Resumen de resultados

- ▶ Sesgos **siempre** se mueven colectivamente en la misma dirección
- ▶ "Efecto del líder"? - no hay evidencia
  - ▶ subestimación en el 2006
  - ▶ sobreestimación en el 2012
- ▶ "Efecto del caballo negro"? - no hay evidencia
  - ▶ sobreestimación en el 2006
  - ▶ subestimación en el 2012
- ▶ "Efectos de partido"? - evidencia más sugerente
  - ▶ PRI es **sobreestimado**
  - ▶ PAN es **subestimado**
  - ▶ PRD efectos ambiguos

# "Efectos de casa" 2006-2012

## Resumen de resultados

- ▶ Sesgos **siempre** se mueven colectivamente en la misma dirección
- ▶ "Efecto del líder"? - no hay evidencia
  - ▶ subestimación en el 2006
  - ▶ sobreestimación en el 2012
- ▶ "Efecto del caballo negro"? - no hay evidencia
  - ▶ sobreestimación en el 2006
  - ▶ subestimación en el 2012
- ▶ "Efectos de partido"? - evidencia más sugerente
  - ▶ PRI es **sobreestimado**
  - ▶ PAN es **subestimado**
  - ▶ PRD efectos ambiguos

# "Efectos de casa" 2006-2012

## Resumen de resultados

- ▶ Sesgos **siempre** se mueven colectivamente en la misma dirección
- ▶ "Efecto del líder"? - no hay evidencia
  - ▶ subestimación en el 2006
  - ▶ sobreestimación en el 2012
- ▶ "Efecto del caballo negro"? - no hay evidencia
  - ▶ sobreestimación en el 2006
  - ▶ subestimación en el 2012
- ▶ "Efectos de partido"? - evidencia más sugerente
  - ▶ PRI es **sobreestimado**
  - ▶ PAN es **subestimado**
  - ▶ PRD efectos ambiguos

# "Efectos de casa" 2006-2012

## Resumen de resultados

- ▶ Sesgos **siempre** se mueven colectivamente en la misma dirección
- ▶ "Efecto del líder"? - no hay evidencia
  - ▶ subestimación en el 2006
  - ▶ sobreestimación en el 2012
- ▶ "Efecto del caballo negro"? - no hay evidencia
  - ▶ sobreestimación en el 2006
  - ▶ subestimación en el 2012
- ▶ "Efectos de partido"? - evidencia más sugerente
  - ▶ PRI es **sobreestimado**
  - ▶ PAN es **subestimado**
  - ▶ PRD efectos ambiguos

# Conclusión

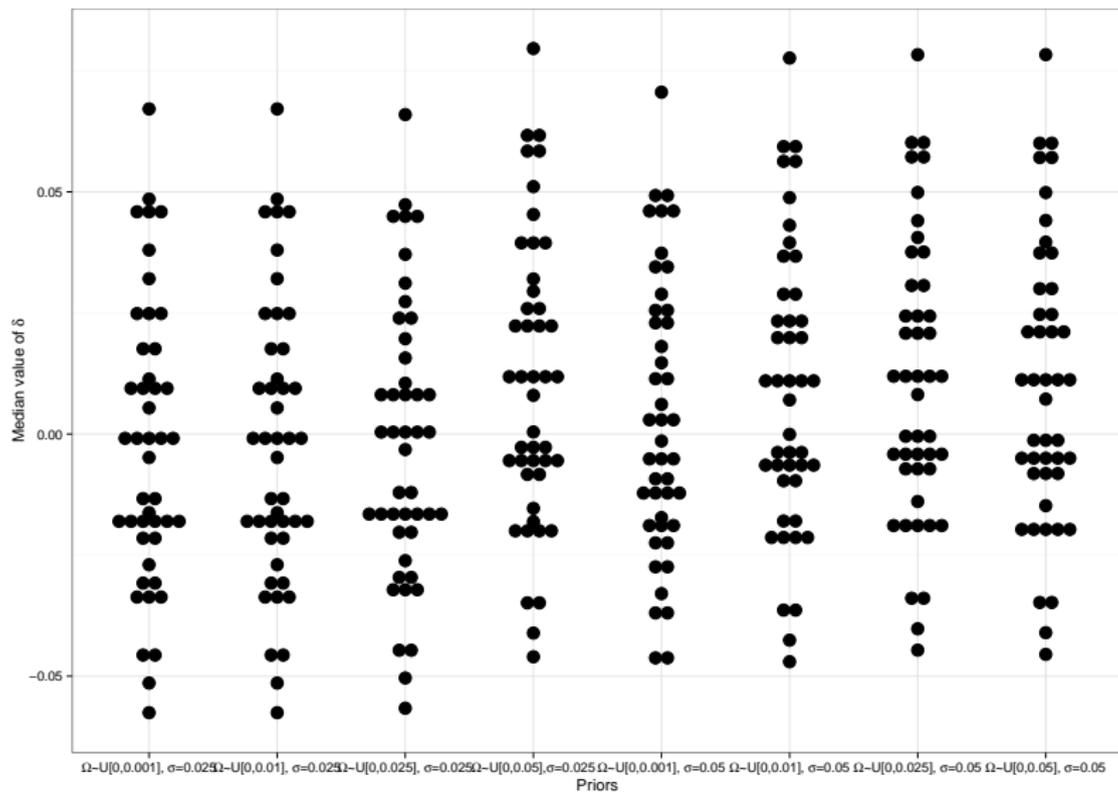
- ▶ Utilidad de la investigación
  - ▶ Evaluación de las firmas encuestadoras por elección y candidato
  - ▶ Los “efectos de casa” nos ayudarán a calibrar la interpretación de las encuestas en elecciones futuras
- ▶ Agenda de investigación
  - ▶ Las razones de los sesgos - imposible ir más allá de su identificación con los datos disponibles

# The Utility of Unpacking Survey Bias in Multiparty Elections: Mexican Polling Firms in the 2006 and 2012 Presidential Elections

Francisco Cantú (U of Houston)  
Verónica Hoyo (UC San Diego)  
Marco Morales (NYU)

# 2006 Presidential election

## Sensitivity analyses



# 2012 Presidential election

## Sensitivity analyses

