

La evidencia de una elección confiable

Javier Aparicio*

javier.aparicio@cide.edu

(Publicado en Nexos no. 346, octubre 2006)

Nunca como en esta elección presidencial se había prestado tanta atención a las pantallas del IFE, a los consejos distritales, a los posibles errores de una elección contada por ciudadanos, o a las sentencias del TEPJF. Asimismo, ningún candidato o partido había explotado este desconocimiento para sembrar tantas dudas con tan escaso sustento. Mediante un análisis estadístico de los datos de más de 130 mil casillas, este texto confronta algunas de las acusaciones de fraude durante la jornada electoral con la evidencia disponible a la fecha.

Consistencia estadística

La concordancia entre el resultado electoral estimado por el conteo rápido, el PREP y el cómputo distrital, todos disponibles la misma semana de la elección, son una evidencia clara de que la elección presidencial fue confiable. ¿Por qué? Si bien el conteo rápido levantado en 7,636 casillas la noche del 2 de julio no apuntaba a un claro ganador (i.e., fuera de los márgenes de error), éste sí sugería una elección con un margen menor al 0.6% de los votos, al tiempo que daba a Felipe Calderón una mayor probabilidad de aventajar en el resultado final. Con este dato era claro que el resultado al cierre del PREP, la noche del 3 de julio, que daba un margen de de 1.04% a Calderón tenía que reducirse al incorporar las actas con inconsistencias y llegar al cómputo distrital—tal y como ocurrió: el cómputo distrital dio un margen de sólo 0.58%, resultado que validó la estimación inicial del conteo rápido.

Si el cómputo distrital hubiera dado un resultado muy diferente al del conteo rápido, indicaría una posible irregularidad en uno u otro procedimiento: si el cómputo distrital hubiera dado un margen de más de 0.6%, sería difícil de explicar que el conteo rápido no lo hubiera detectado en su momento. Del mismo modo, si el resultado del cómputo hubiera sido fruto de una alteración de actas posterior al 2 de julio, como se acusó, sería difícil de explicar que el conteo rápido hubiera anticipado dicha manipulación.

PREP y cómputo distrital: el algoritmo inexistente

El resultado del conteo rápido no permitía identificar con precisión estadística a un ganador, y por ello el IFE anunció que la elección era muy cerrada y que había que esperar hasta el cómputo distrital. Ante la escasa información, hubo quienes erróneamente leyeron de más en los datos del PREP. En principio de cuentas, la alteración del flujo de datos mediante un “fraude cibernético” es fútil cuando lo importante es la suma final de votos y no la forma en que se acumula dicha suma; además, los actas de escrutinio son cotejadas y

* Doctor en economía, con especialidad en economía política, por la Universidad George Mason. Profesor Investigador de la División de Estudios Políticos del CIDE. Este artículo sintetiza un trabajo de investigación en proceso, disponible en: <http://www.cide.edu/investigadores/aparicio/elecciones/>

verificadas durante el cómputo distrital, de modo que cualquier manipulación del PREP no duraría más de 48 horas. Aún así, vale la pena explicar los flujos de datos del PREP y del cómputo distrital.

Se dijo que el PREP debió haber mostrado múltiples cruces, dada la aleatoriedad del proceso. Sin embargo, el PREP y el cómputo distrital no son procesos muestrales ni plenamente aleatorios, como sí lo son los conteos rápidos y las encuestas de salida. Factores no aleatorios como la urbanización de los distritos y los diferentes husos afectan el ritmo de captura de las casillas. Aún en una elección reñida, si las condiciones iniciales del flujo de datos favorecen suficientemente a un candidato no tiene por qué haber muchos cruces.

¿Cómo puede ocurrir esto? De las más de 130 mil casillas, alrededor de 70% fueron instaladas en zonas urbanas y sólo un 30% en áreas rurales. Si las casillas urbanas llegan más rápido a los consejos distritales, el PREP favorecería al candidato con mayor voto urbano—como fue el caso. Felipe Calderón tuvo una ventaja de 691 mil votos sobre AMLO en casillas urbanas, mientras que AMLO superó por casi 450 mil votos a Calderón en casillas rurales. A la medianoche del lunes 2 de julio, alrededor del 50% de las casillas urbanas ya habían sido capturadas, mientras que las rurales llevaban un avance del 25%. Así, la ventaja de Calderón se vio favorecida por el sesgo urbano del PREP.

¿Por qué disminuyó el margen de Calderón cuando se incorporaron las 11,184 “actas inconsistentes”? El 65% de éstas provenían de casillas rurales que en general favorecieron a AMLO, disminuyendo así el margen del puntero panista. El error más común de estas actas, por cierto, fue dejar en blanco el número de votos de Alternativa y/o Nueva Alianza—difícilmente un acto doloso.

El cómputo distrital del 5 y 6 de julio fue aún menos aleatorio que el PREP. Aquí, el ritmo de avance en cada distrito dependía de cuánto se discutía cada una de las actas. Si los distritos que favorecían a Calderón eran consistentemente más debatidos, sus resultados demorarían más en ser computados. Esta dinámica produjo un sesgo partidista en el flujo de datos de los cómputos distritales. Si AMLO tenía ventaja en los primeros distritos computados, ésta tenía que disminuir conforme llegaban los distritos más debatidos, hasta revertirse. En vez de urbano, ahora el sesgo fue partidista: a las 4pm del 5 de julio se habían computado más del 75% de las actas de estados donde AMLO ganó, mientras que el avance en estados panistas era menor al 50%. Por este rezago el cruce ocurrió hasta la madrugada del 6 de julio.

Errores aritméticos en actas

Durante la impugnación de los cómputos, una fuente de suspicacia adicional fueron los llamados errores aritméticos en actas de escrutinio. Estos errores surgen cuando alguno de los “campos de referencia” (boletas recibidas, depositadas, o sobrantes, y número de ciudadanos que votaron según la lista nominal) no es consistente con el voto total registrado. Por ejemplo, una casilla con 700 boletas recibidas y 500 votos emitidos debe tener 500 marcas en la lista nominal y 200 boletas sobrantes. Cuando un acta tiene errores

aritméticos da la apariencia de tener votos de más o de menos respecto de los ciudadanos que votaron, o bien de las boletas recibidas y sobrantes.

Una aclaración pertinente es que el acta de escrutinio está diseñada para que los funcionarios de casilla cuenten boletas y anoten los resultados, de manera que su llenado no requiere realizar sumas o restas ni conciliación numérica alguna.

Hasta ahora sabíamos muy poco sobre la magnitud y frecuencia de los errores aritméticos pero no son algo nuevo. De hecho, las actas de la elección presidencial de 2006 tuvieron menos errores (46.7% del total) y datos omitidos que las del 2000 (51.4%). Sin embargo, la magnitud promedio de estos errores es muy similar: 1.26% de los votos por casilla en 2000 frente a 1.35% en 2006. En ambas elecciones, los errores aritméticos se distribuyeron de manera casi uniforme en las casillas ganadas por uno u otro candidato, lo cual sugiere errores humanos aleatorios.

Los errores aritméticos son numerosos porque la consistencia aritmética de las actas es, por así decirlo, frágil: basta una entrada mal contada, registrada u omitida para producir uno o más errores y que no cuadren las cifras. Además, aunque los funcionarios de casilla hicieran su mejor esfuerzo por conciliar las actas, basta que un ciudadano deposite por error su voto en una casilla contigua a la suya para producir errores en dos actas: en una sobrarán votos mientras que en otra faltarán. Muchos de estos errores, sin embargo, se cancelan mutuamente al agregar votos a nivel sección o a nivel distrito.

Aunque numerosos, los errores aritméticos no son de gran magnitud. El error promedio por acta fue de ± 4.36 votos (1.3% de los 328 votos promedio por casilla). Las casillas ganadas por la Coalición muestran errores ligeramente mayores (4.54) que las panistas (4.20). Asimismo, las casillas urbanas tienen errores mayores que las rurales (4.48 vs. 4.21). Es decir, la magnitud de los errores es relativamente similar en casillas ganadas por uno u otro candidato así como en zonas rurales y urbanas, tal como se esperaría de errores aleatorios.

Se ha dicho que ante lo cerrado del resultado electoral, con un margen de 1.8 votos promedio por casilla, estos errores pueden ser relevantes. Para atender esta inquietud hay que considerar, más allá de su magnitud, de qué manera están distribuidos los errores. Si hay más errores en las casillas ganadas por el puntero, indicarían una posible irregularidad. Pero si los errores se distribuyen de manera aleatoria, es decir, en una proporción similar en las casillas ganadas por uno u otro candidato, sugerirían simples errores humanos que afectan por igual a los candidatos sin afectar el margen de victoria.

Considérese un error particular obtenido de los datos del PREP (única fuente que incluye los llamados “campos de referencia”). El número de ciudadanos que votaron según la lista debe ser igual a la votación total registrada: 59,120 actas reflejaban esta igualdad mientras que 54,111 tenían error de uno u otro signo (el resto de las actas no permiten el cálculo por tener datos omitidos o estar fuera del PREP). El PAN ganó en 52.7% de estas casillas y tuvo 51.6% de las actas con error (27,917 actas con error de 59,629 ganadas). Por su parte, la Coalición ganó 47.3% de las casillas consideradas y tuvo 48.4% de las actas con error (26,194 actas con error de 53,602 ganadas). Es decir, hubo un poco más de errores en las

casillas ganadas por la Coalición, pero en general la proporción es muy similar—lo que sugiere errores aleatorios.

Recuento de paquetes

El siguiente nivel de incertidumbre radicaba en si las actas coincidían o no con el contenido de los paquetes electorales. En este respecto, un análisis estadístico formal requeriría del recuento de una muestra aleatoria y representativa de paquetes. Sin embargo, las restricciones legales para la apertura de paquetes durante el cómputo distrital y la lógica jurídica, básicamente casuística, del TEPJF no produjeron una muestra representativa de paquetes recontados. No obstante, la evidencia disponible a la fecha permite hacer algunas inferencias parciales.

Durante el cómputo distrital se recontaron 2,864 paquetes electorales, y más adelante el TEPJF ordenó el recuento de 11,839 casillas. En ambos casos, las razones que produjeron el recuento de estas casillas, y no otras, no son aleatorias, de modo que ni el recuento distrital ni el del Tribunal se basaron en muestras representativas. Por ende, el resultado de estos recuentos no permite hacer extrapolaciones directas, pero al menos ofrece indicios del tipo de ajustes que resultaron de recontar diferentes tipos de casillas.

Los 2,864 paquetes recontados durante el cómputo distrital eran una muestra sesgada: 66.4% provenían de distritos panistas, y 33.4% de distritos ganados por la Coalición. En general, el recuento produjo ajustes a la baja en los votos de todos los candidatos y un ligero aumento del margen de Calderón sobre AMLO (0.6 votos en promedio). Cabe resaltar que en 60% de los casos el voto de Calderón o AMLO no tuvo cambio alguno.

Al dividir la muestra en distritos panistas y perredistas surgen asimetrías interesantes. Al recontar casillas panistas, el ajuste promedio por casilla fue de 4.7 votos menos para Calderón, y de 1.9 votos menos para AMLO, lo que redujo en 2.9 votos el margen entre los punteros. Por otro lado, al recontar casillas perredistas, el ajuste promedio por casilla fue de 5.8 votos menos para Calderón, y de 13.3 votos menos para AMLO, lo que aumentó en 7.5 votos el margen del primero. Es decir, al recontar casillas ambos candidatos pierden votos pero el candidato con más votos en la casilla pierde más votos tras el recuento. Además, AMLO perdió proporcionalmente más votos que Calderón al recontarse casillas en sus respectivos distritos, lo cual hizo que el resultado neto favoreciera a Calderón a pesar del sesgo de la muestra (66.4% casillas panistas).

Existe una razón natural por la cual un recuento afecta más al candidato con más votos en una casilla. Si el error de conteo es una tasa más o menos constante, digamos 5%, un recuento en una casilla con 200 votos para uno y 100 votos para otro le reducirá 10 votos al primer candidato pero sólo 5 votos al segundo, disminuyendo el margen entre ambos en 5 votos. Para que este error aleatorio se cancele habría que recontar casillas con el margen contrario, donde el segundo candidato tenga 100 votos más que el primero. Por ello, un recuento en una muestra sesgada de casillas produce un ajuste sesgado de votos—razón de más para interpretar con cautela los resultados de los recuentos realizados durante el proceso electoral.

Las 11,839 casillas recontadas por el TEPJF, al ser fruto de la impugnación de la Coalición, eran una muestra aún más sesgada: 91.4% (10,818) de las casillas provenían de distritos panistas y sólo 8.6% (1,021) eran perredistas. Si consideramos que Calderón aventajó a AMLO por 76.5 votos promedio en casillas de estados panistas (145.3 vs. 68.8 votos, respectivamente), era de esperarse que el recuento del Tribunal produjera una mayor disminución de votos para Calderón que para AMLO y una reducción del margen de victoria—tal como ocurrió en el cómputo definitivo del 5 de septiembre.

En cualquier caso, los ajustes encontrados en una muestra no representativa de casillas no pueden generalizarse hacia el resto del país: si el recuento parcial del TEPJF encontró ajustes de votos promedio de ± 2 votos a favor de AMLO, no se puede afirmar que un recuento más amplio hubiera encontrado un ajuste en el mismo sentido y magnitud. Más bien, a la luz de la evidencia disponible, es plausible suponer que un recuento adicional en casillas donde AMLO aventajó, en vez de disminuir o revertir la ventaja de Calderón, la habría recompensado toda vez que AMLO perdería más votos que el panista.

En todo caso, si ambos recuentos hubieran arrojado modificaciones sistemáticas respecto del resultado del cómputo distrital respectivo, indicarían alguna irregularidad más general en el proceso—lo cual no ocurrió. Con la evidencia parcial disponible a la fecha, es plausible suponer que los errores de conteo son de magnitud y distribución similar a los errores aritméticos reflejados en las actas, mismos que no parecieron favorecer al candidato ganador.

Más allá de la evidencia estadística, si los errores aritméticos o de conteo hallados a lo largo del país son indicativos de irregularidades dolosas, resulta difícil explicar algunos sinsentidos: Por un lado, contar con una inmensa capacidad de coordinación entre miles de funcionarios de miles de casillas, cuyo alcance llegaría incluso a los bastiones de AMLO. Por otro lado, resulta ilógico que los mismos funcionarios que alteran la votación de una casilla al mismo tiempo acusen sus “errores” en las actas—haciéndolas con ello más fáciles de impugnar. Una falsedad similar es suponer que todos los errores favorecen o afectan a un mismo candidato. Si los errores aritméticos dan la apariencia de votos sobrantes o faltantes, no hay razón para pensar que todos los votos que “faltan” son del candidato perdedor ni que todos los votos que “sobran” son del ganador.

Pruebas del ácido para casillas atípicas

Aún cuando el recuento de casillas del TEPJF produjo ajustes en los cómputos respectivos, se ha dicho que el criterio de “determinancia” usado por el Tribunal, consistente en anular sólo las casillas con errores mayores al margen entre el primero y segundo lugar en cada casilla, significó pasar por alto irregularidades importantes. ¿Qué pasaría si aplicamos, de manera hipotética, criterios más estrictos?

Como se dijo antes, los errores aritméticos están distribuidos de manera casi uniforme o aleatoria a lo largo del país. ¿Es posible que el resultado de la elección haya sido determinado por casillas con errores aritméticos? Partiendo del PREP, si eliminamos del cómputo todas las casillas con más o menos votos que el número de ciudadanos que votaron (54,111 actas con error, 41.3% del total), tenemos que Felipe Calderón ganaría por

494 mil votos—lo cual confirma que hubo un poco más de errores en las casillas ganadas por la Coalición, como vimos antes. Algo similar sucede si eliminamos del cómputo todas las casillas donde las boletas recibidas menos las boletas sobrantes es diferente del voto total (56,702 actas con error, 43.3% del total).

Como sabemos, el nivel de participación promedio fue de 58% a lo largo del país. ¿Es posible que el resultado final haya sido determinado por casillas con participación inusualmente elevada? Partiendo del cómputo distrital del IFE, si eliminamos todas las casillas con participación mayor al 75% de la lista nominal (alrededor de 4,555 casillas, 3.5% del total) Felipe Calderón continúa arriba.

De modo que el resultado electoral del 2 de julio es robusto a la eliminación de casillas con errores aritméticos, así como a la eliminación de casillas con una alta tasa de participación. Claramente, el TEPJF no puede aplicar criterios como éstos pues al final de cuentas su misión es garantizar la integridad del sufragio de los ciudadanos en la medida de lo posible, a menos de que cuente con pruebas en contrario.

Representantes de casilla

Otra presunta irregularidad fue que AMLO tuvo una votación inusualmente baja en zonas donde no tuvo suficientes representantes de casilla. Detrás de este cargo hay un supuesto implícito: que de no ser por los representantes de partido, los ciudadanos elegidos al azar como funcionarios de casilla alterarían las votaciones. Si comparamos la presencia de representantes de partido, a nivel distrito, con los resultados electorales podemos encontrar una respuesta.

A nivel nacional el PAN y la Coalición tuvieron representantes en 79.9% y 79.4% de las casillas, respectivamente, aunque su distribución no fue uniforme: el PAN tuvo representantes en 35% de las casillas del DF, bastión del PRD, mientras que la Coalición tuvo representantes en 33% de las casillas de Nuevo León, bastión del PAN. De modo similar, la reñida votación de ambos candidatos a nivel nacional fue producto de una votación sesgada regionalmente: 14 estados ganados con holgura por cada candidato, y 4 estados competidos: Puebla, Campeche, Veracruz y Zacatecas.

En efecto, la votación de AMLO en distritos donde no tuvo muchos representantes de casilla fue menor a la de los distritos donde tuvo una mejor cobertura. Sin embargo, la votación de Calderón también fue menor en los distritos donde no tuvo tantos representantes como en sus bastiones. Estadísticamente, la magnitud de esta relación es similar en ambos sentidos: por cada 10% adicional de representantes, cada partido obtuvo aproximadamente 4% más votos. Por lo tanto, leer esta relación como evidencia de irregularidades del PAN contra la Coalición implica por fuerza que ésta hizo lo mismo contra aquel partido.

Hay una alternativa simple a la explicación dolosa: La fuerza electoral de un partido puede medirse tanto por los votos obtenidos como por su capacidad local de organización. En distritos donde un partido no tiene presencia en votos, difícilmente tendrá cuadros

interesados en ser representantes de partido. Si esto es cierto, es de esperarse observar correlaciones como las indicadas arriba.

Otra cuestión relacionada fue si la presencia de representantes de casilla de uno u otro partido afectó el número de votos nulos. En total hubo 904,604 votos nulos (2.16% del voto total), una cifra similar al 2.10% de 2000. Aproximadamente 60% de los votos nulos se dieron en casillas urbanas y 40% en rurales. Sin embargo, los representantes de casilla del PAN o la Coalición no tienen una relación significativa con el porcentaje de votos nulos, como sí la tiene el porcentaje de casillas rurales.

Imparcialidad vs. precisión

La evidencia indica que el error humano no fue determinante en el resultado electoral que hoy conocemos. Al delegar la instalación de casillas y el conteo de los votos a ciudadanos elegidos al azar, el sistema electoral mexicano garantiza la imparcialidad de la jornada electoral como casi ningún otro país. Es obvio que éstos ciudadanos cometerán errores e imprecisiones al contar votos y llenar actas con la misma aleatoriedad con que todos lo hacemos. Como este sistema nunca había enfrentado una elección presidencial tan competitiva, no se había prestado atención a los posibles errores de una elección contada por ciudadanos. Quizá sea posible diseñar mecanismos para contar votos con mayor precisión pero difícilmente encontraremos uno más imparcial.

Entre las posibles mejoras del proceso electoral sin duda se encuentran aminorar la frecuencia de errores, e incluso facilitar un recuento muestral en elecciones con márgenes reducidos, sin el innecesario desgaste visto ahora. Tan cierto es esto como la necesidad de proteger al sistema electoral de una andanada irresponsable por parte de candidatos que no quieran o puedan aceptar la evidencia más allá de sus profundas o legítimas convicciones.

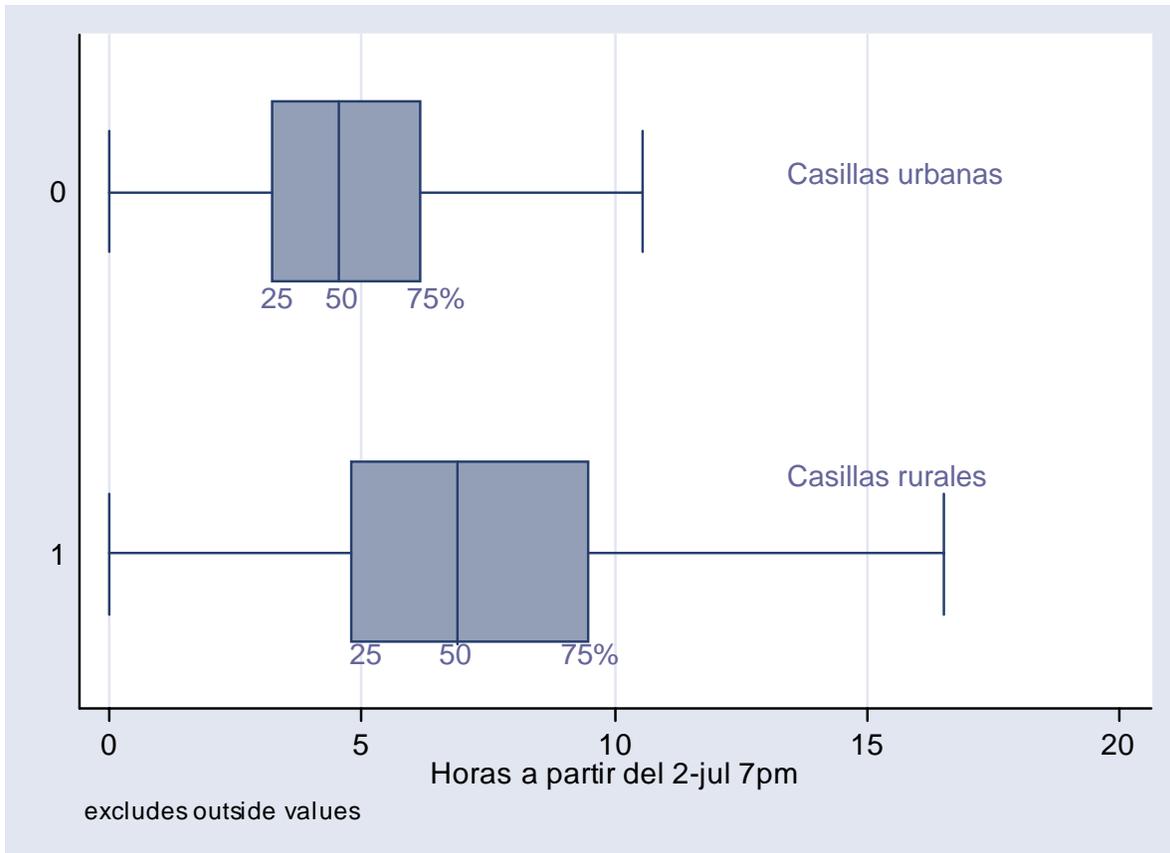


Figura 1.
Avance relativo de captura de casillas urbanas y rurales durante el PREP.
Nota: cálculos propios con base en datos del IFE a nivel casilla.

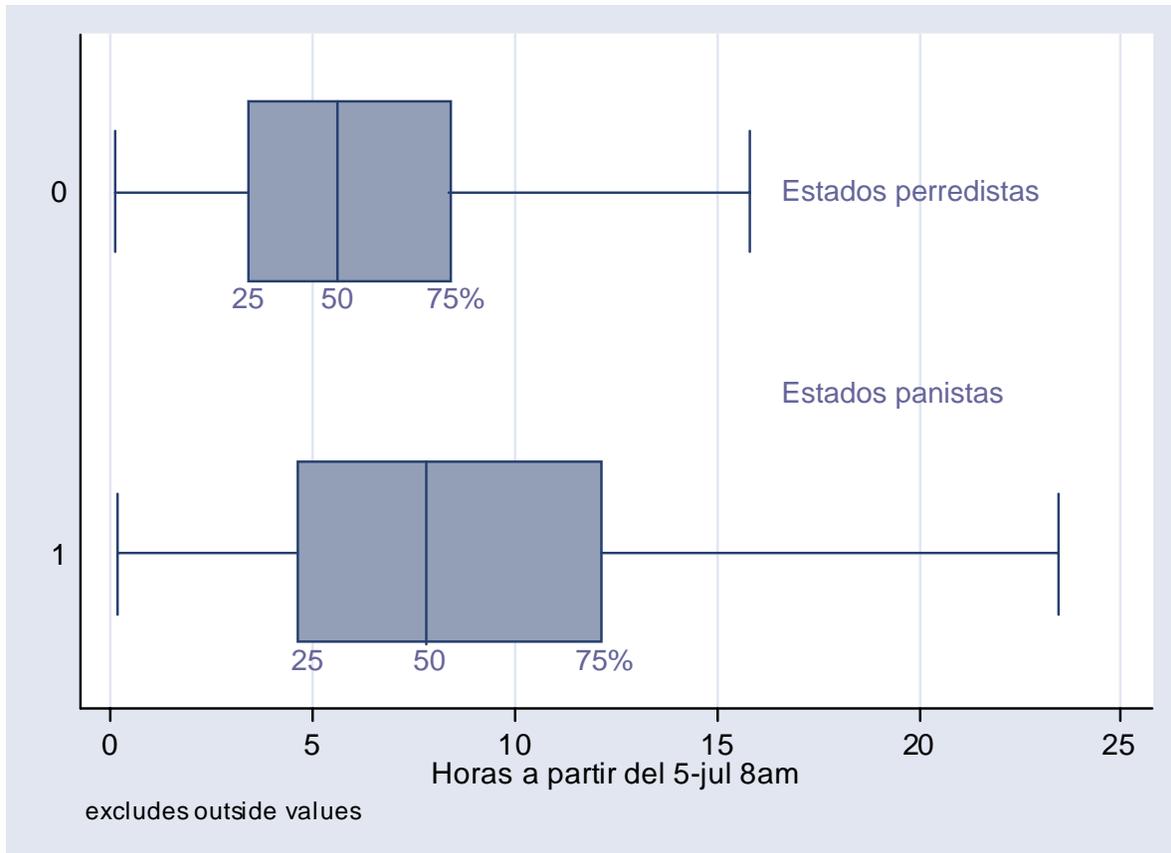


Figura 2.
Avance relativo de captura de actas para presidente durante cómputo distrital.
Nota: cálculos propios con base en datos del IFE a nivel casilla.

Tabla 1
Votos promedio según tipo de casilla y filiación política

	Votos promedio por casilla			
	PAN	CPBT	APM	Margen PAN-CPBT
Casillas Urbanas	132.51	124.74	67.11	7.76
%	37.79%	35.52%	20.09%	2.27%
Casillas Rurales	76.06	87.50	80.59	-11.44
%	28.49%	33.12%	31.97%	-4.63%
Estados Panistas	145.34	68.82	75.88	76.51
%	45.26%	22.21%	25.73%	23.06%
Estados Perredistas	89.28	149.36	67.18	-60.07
%	26.37%	45.13%	22.16%	-18.77%
Total	114.70	112.84	71.12	1.87
%	34.94%	34.74%	23.78%	0.20%
Total Nacional				
41,791,322	15,000,284	14,756,350	9,301,441	243,934
100.00%	35.89%	35.31%	22.26%	0.58%

Nota: Cálculos propios basados en datos del cómputo distrital del IFE.

Tabla 2
Errores aritméticos promedio en actas de escrutinio para Presidente

Tipo de error	Total de actas	Filiación de casilla		Tipo de casilla	
		Panistas	Perredistas	Rurales	Urbanas
Voto total - boletas depositadas	2.04	2.26	1.80	2.74	1.78
Boletas depositadas - ciudadanos que votaron	-5.57	-5.39	-5.77	-7.81	-4.75
Voto total - ciudadanos que votaron	-4.33	-3.92	-4.78	-6.03	-3.69
Boletas recibidas - sobrantes - depositadas	-0.66	-0.15	-1.22	0.65	-1.13
Boletas recibidas - sobrantes - voto total	-3.46	-3.18	-3.78	-2.98	-3.65
Boletas recibidas - sobrantes - ciudadanos que votaron	-7.79	-7.05	-8.62	-8.97	-7.35
Boletas recibidas - (lista nominal + 10)	5.29	5.87	4.65	2.60	6.31
Boletas sobrantes + depositadas - (lista nominal + 10)	5.75	5.78	5.71	1.87	7.17
Error promedio	-1.09	-0.72	-1.50	-2.24	-0.66
Error absoluto promedio	4.36	4.20	4.54	4.21	4.48

Nota: Cálculos propios basados en datos del PREP.

Tabla 3
Distribución de errores aritméticos por tipo de casilla
Voto total = número de ciudadanos que votaron según lista nominal

	Voto PAN <= Voto PBT	Voto PAN > Voto CPBT	Total casillas
Actas sin error	27,408 46.36%	31,712 53.64%	59,120 100.00%
Actas con error	26,194 48.41%	27,917 51.59%	54,111 100.00%
Total casillas	53,602	59,629	113,231
%	47.34%	52.66%	100%

Nota: Cálculos propios basados en datos del PREP.

Tabla 4
Recuento de paquetes durante cómputo distrital
Ajuste promedio entre dato PREP y recuento

	Muestra completa	Distritos panistas	Distritos perredistas
PAN	-5.051	-4.678	-5.804
CPBT	-5.665	-1.859	-13.310
APM	-6.510	-4.244	-11.060
PANAL	-1.387	-1.017	-2.144
ASDC	-0.306	-0.024	-0.873
Cand. no registrados	-0.395	-0.170	-0.838
Votos nulos	-7.339	-7.326	-7.365
Voto total	-17.546	-12.268	-28.035
Margen PAN - CPBT	0.594	-2.864	7.551
No. de casillas	2,864	1,901	963
Porcentaje	100.0%	66.4%	33.6%

Nota: Cálculos propios basados en datos del IFE.