

Análisis estadístico de la elección presidencial de 2006 ¿Fraude o errores aleatorios?

Javier Aparicio*

Resumen: Este artículo confronta algunas de las principales acusaciones de fraude electoral con la evidencia disponible mediante un análisis estadístico de los datos a nivel casilla del PREP y del cómputo distrital de la elección presidencial de 2006. En primer lugar se analiza el conteo rápido, el flujo de datos del PREP y del cómputo distrital; en segundo lugar, los errores aritméticos en las actas de casilla y el recuento de paquetes electorales, y por último el papel de las casillas atípicas y los representantes de casilla. La conclusión principal es que, a pesar de los errores en actas y los sesgos propios del flujo de datos, el resultado electoral del 2 de julio resulta estadísticamente confiable.

Palabras clave: errores en actas, fraude electoral.

Statistical Analysis of the Presidential Election in 2006: Fraud or Random Errors?

Abstract: This paper tests some of the main fraud allegations of the 2006 presidential election in Mexico with the evidence from a statistical analysis of precinct-level data. First, I analyze the data flow from the quick count, the preliminary results (PREP) and the official district tally (*cómputo distrital*). Secondly, I analyze the size and distribution of errors in the polling-station acts and the partial recount. Lastly, I consider atypical precincts and the role of party representatives. My main conclusion is that, since most fraud allegations are not supported by the evidence, the presidential election outcome is statistically reliable.

Keywords: election forensics, fraud.

La elección presidencial de 2006 fue la más reñida en la historia política reciente de México. La noche del 2 de julio el consejero presidente del IFE, Luis Carlos Ugalde, anunció en cadena nacional que los resultados del conteo rápido de casillas electorales no permitían anticipar un claro ganador de la contienda. A partir de entonces se inició una controversia que se extendería

* Javier Aparicio es profesor-investigador del Centro de Investigación y Docencia Económicas (CIDE). Carretera México-Toluca, 3655, Col. Lomas de Santa Fe, 01210, México, D.F. Tel. 52 (55) 57 27 98 00. Correo electrónico: javier.aparicio@cide.edu.

Artículo recibido en julio de 2008 y aceptado para su publicación en abril de 2009.

a lo largo de varios meses, detonando lo que muchos consideraron como una crisis postelectoral sin precedentes: en el año 2006 volvimos a escuchar acusaciones de fraude electoral que muchos creían superadas al menos desde la muy cuestionada elección de 1988.

En este artículo se confrontan algunas de las principales acusaciones de fraude electoral durante la jornada del 2 de julio de 2006 con la evidencia disponible a partir de un análisis estadístico de los datos del Programa de Resultados Electorales Preliminares (PREP)¹ y del cómputo distrital de las más de 130 788 casillas instaladas ese día. Estas acusaciones iban desde un novedoso “fraude cibernético” en el flujo de datos del PREP, irregularidades en el cómputo distrital, hasta acusaciones de fraude “a la antigua”, como el llenado incorrecto de actas de escrutinio y cómputo o la intervención indebida de los funcionarios de casilla o de los representantes de cada partido en las mismas.

El debate sobre la elección presidencial tiene múltiples aristas. Algunos autores han evaluado el proceso electoral desde las campañas hasta la resolución del conflicto postelectoral (Schedler, 2007; Lawson, 2008; Ugalde, 2008). Otros autores han analizado la cuestión del acceso a las actas y paquetes electorales, así como los límites o alcances de la resolución del Tribunal Electoral (Ackerman, 2007). Algunos de estos argumentos cuentan con cierto mérito, sobre todo los concernientes al estatus legal —y por lo tanto a la viabilidad de impugnar y a los remedios posibles— de los errores en actas de escrutinio y cómputo (Crespo, 2008). O bien podemos analizar la evidencia de las actas de casilla, como se hace aquí y en otros estudios (Poiré y Estrada, 2006; Mebane, 2007; Pliego, 2007).

Ante la complejidad de una elección cerrada y las dudas surgidas durante el proceso de cómputo y el recuento parcial ordenado por el Tribunal Electoral del Poder Judicial de la Federación (TEPJF), ¿cómo saber si el resultado electoral es confiable? Este trabajo presenta un análisis estadístico de la evidencia disponible en los datos a nivel casilla de la elección presidencial de 2006. La conclusión principal es que, a pesar de los posibles errores tanto de los funcionarios de casilla como de las autoridades electora-

¹ El PREP es un mecanismo de monitoreo, en tiempo real, del flujo de datos electorales de las casillas hacia los centros de conteo. No es un instrumento de medición que permita hacer pronósticos válidos o confiables, toda vez que no se basa en una muestra aleatoria de casillas, sino en el flujo, en tiempo real, de los datos. Como este flujo por naturaleza tiene una serie de sesgos, es muy difícil obtener resultados representativos a partir del PREP, sobre todo en una elección reñida.

les del IFE y el TEPJF, el resultado electoral del 2 de julio resulta estadísticamente confiable.

El análisis estadístico de los datos casilla por casilla proporcionados por el IFE desde el día siguiente a la elección nos permite responder, con diferentes grados de confianza estadística, algunas de estas interrogantes. Cabe aclarar de antemano que algunas acusaciones, como las referentes a la intervención del presidente Fox en la campaña o el efecto de los *spots* de grupos empresariales, tienen que ver con el proceso previo a la jornada electoral y no pueden atenderse en este análisis.

El artículo está organizado de la siguiente manera. En primer lugar se analizan el conteo rápido y su relación con los resultados definitivos conocidos posteriormente. Enseguida se presenta un análisis del flujo de datos tanto del PREP como del cómputo distrital. En el tercer apartado se describen los llamados “errores aritméticos” en las actas de escrutinio y cómputo. En cuarto lugar se estudia la relación entre representantes de casilla, sustitución de funcionarios y resultados. Por último, se discuten los resultados del recuento de casillas.

La consistencia del conteo rápido

Al igual que en otras elecciones recientes, desde los meses previos a la elección el IFE había diseñado un conteo rápido que le permitiera adelantar un resultado confiable pocas horas después de terminada la jornada electoral. El conteo rápido fue levantado en una muestra de 7 636 casillas, misma que fue diseñada para ser representativa tanto del resultado nacional como de los estratos regionales o partidistas más importantes. En su momento se cuestionó la incapacidad del conteo rápido para arrojar un ganador definitivo la misma noche de la elección. ¿Acaso el IFE pretendía ocultar un resultado que no le favorecía al PAN? ¿Hubo irregularidades en el levantamiento de los datos? ¿Se pretendía robar la elección?

Dado lo cerrado del resultado —los datos disponibles a las 11 de la noche del 2 de julio—, el IFE no pudo anticipar un candidato ganador. ¿Por qué? Como en todo ejercicio estadístico, el tamaño de la muestra determina cierto margen de error, a partir del cual pueden hacerse inferencias estadísticamente confiables (por ejemplo, una encuesta típica con una muestra de 1 200 entrevistados tiene un margen de error de +/- 3 por ciento). En este caso, el margen de error del conteo rápido era de 0.6 por ciento

pero como el resultado fue aún más cerrado (0.58%) no era posible declarar un ganador con los grados de confianza estadística comúnmente aceptados y que el mismo IFE había anunciado previamente.

La concordancia entre el resultado electoral estimado por el conteo rápido, el PREP y, finalmente, el cómputo distrital, todos disponibles la misma semana de la elección, proveen evidencia a favor 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 (es decir, fuera de los márgenes de error del instrumento), éste sí sugería una elección con un margen menor a 0.6 por ciento 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 observado al cierre del PREP la noche del 3 de julio, y que daba un margen de 1.04 por ciento a favor de Calderón, tenía que reducirse al llegar al cómputo distrital —tal como ocurrió: el cómputo distrital arrojó un margen de sólo 0.58 por ciento, el cual validó la estimación inicial del conteo rápido.

De haber existido una manipulación fraudulenta a favor de uno u otro candidato en ciertas zonas del país, ésta difícilmente hubiera sido capturada por un conteo rápido basado en una muestra aleatoria. Por eso, si el cómputo distrital hubiera dado un resultado muy diferente al del conteo rápido, indicaría una posible irregularidad en alguno de los dos procedimientos: o acusaría un error en el diseño o representatividad inicial de la muestra, o bien revelaría un resultado posiblemente alterado. Pero tanto la muestra del conteo rápido como el cómputo distrital fueron consistentes. Si el resultado del cómputo hubiera sido fruto de una alteración sistemática de actas posterior al 2 de julio, como se acusó, sería difícil explicar que el conteo rápido anticipara dicha manipulación. Del mismo modo, si el cómputo distrital hubiera dado un margen mayor a 0.6 por ciento, sería difícil explicar que el conteo rápido no lo hubiera detectado en su momento.

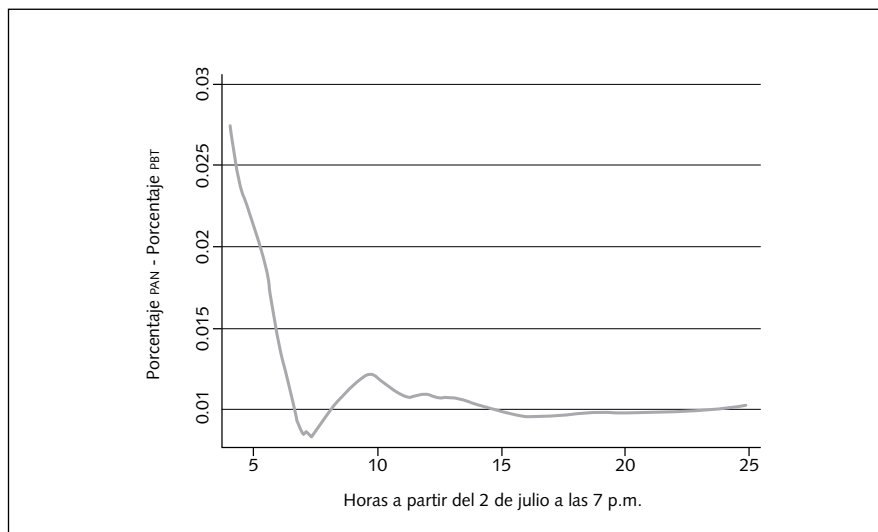
PREP y cómputo distrital: el algoritmo inexistente

El resultado del conteo rápido no permitía identificar con precisión estadística a un ganador; por eso el IFE anunció que, ante lo cerrado de la elección, habría que esperar los resultados del cómputo distrital. En vista de la incertidumbre y la confusión, propios y extraños leyeron de más los datos del PREP.

Se dijo, por ejemplo, que debido a lo reñido de los resultados y bajo una supuesta aleatoriedad del flujo de datos de casilla, el PREP debió haber mostrado múltiples cruces entre el primer y el segundo lugar. Sin embargo, esta afirmación partía de un supuesto erróneo. Ni el PREP ni el cómputo distrital son procesos 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 horarios, afectan (sesgan) el flujo de captura de las casillas. Aun en una elección reñida, si las condiciones iniciales del flujo de datos favorecen suficientemente a un candidato, no tendríamos que esperar los cruces múltiples. Recuérdese que una elección cerrada en el ámbito nacional no equivale a un resultado cerrado en cada región, estado o casilla (véase gráfica 1).

Durante las primeras cinco horas del PREP, Felipe Calderón tenía una ventaja de dos puntos porcentuales o más sobre López Obrador, pero esa ventaja parecía disminuir hora tras hora. Al tratarse de la tendencia inicial, era fácil imaginar que dicha superioridad se revertiría durante la madrugada. Sin embargo, se revirtió alrededor de las 3 de la mañana del lunes hasta estabilizarse en un ventaja de alrededor de uno por ciento para el PAN.

GRÁFICA 1. Margen porcentual entre PAN y CPBT durante el PREP, 2 de julio de 2006



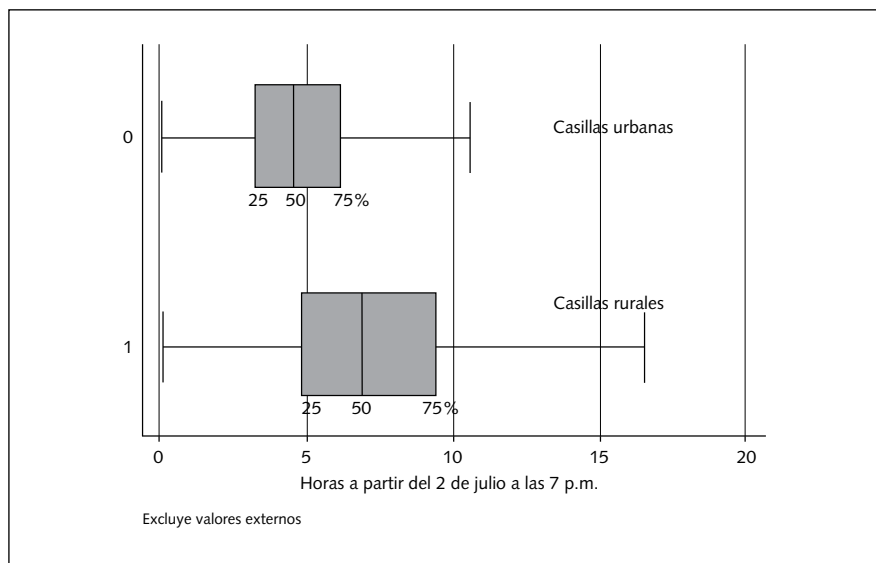
Fuente: Elaboración propia con base en datos del IFE.

¿Cómo pudo ocurrir esto? ¿Cómo fue posible que en una elección tan reñida como la que analizamos uno de los candidatos se mostrara consistentemente por encima del otro durante 24 horas consecutivas?

Sucede que, de las más de 130 000 casillas, alrededor de 70 por ciento fue instalado en zonas urbanas y sólo 30 por ciento 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 691 000 votos sobre Andrés Manuel López Obrador en casillas urbanas, mientras que éste superó por casi 450 000 votos a Calderón en casillas rurales. A la medianoche del lunes 3 de julio, alrededor de 50 por ciento de las casillas urbanas ya habían sido capturadas, mientras que las rurales llevaban un avance de 25 por ciento, de modo que la ventaja de Calderón se vio favorecida por el sesgo urbano del PREP.

Al finalizar el PREP, surgieron acusaciones contra el IFE por *ocultar* los resultados de más de 11 000 casillas, provenientes de *actas inconsistentes*. ¿Por qué no fueron reportadas junto con las demás casillas? Desde enero de 2006, el IFE y los partidos acordaron que se dejarían fuera del PREP las actas

GRÁFICA 2. Avance relativo de captura de casillas urbanas y rurales durante el PREP

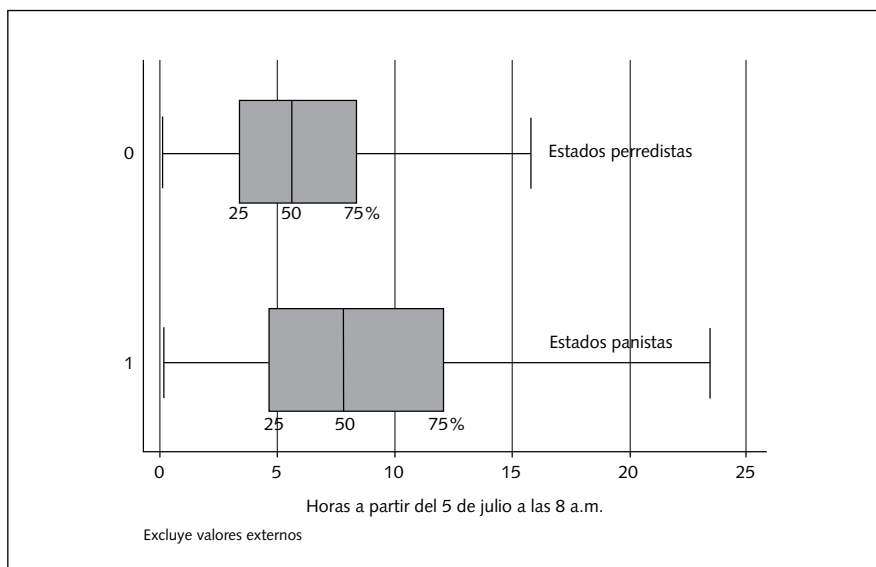


Fuente: Elaboración propia con base en datos del IFE a nivel casilla.

en las que uno o más rubros del número de votos fueran omitidos o ilegibles (Tello, 2007). Una vez incorporadas las actas inconsistentes en el cómputo distrital, el margen de Calderón disminuyó en casi 150 000 votos. ¿Por qué disminuyó el margen de Calderón cuando se incorporaron las 11 184 actas inconsistentes? Esto se debió a que 65 por ciento de estas actas provenía de casillas rurales que en general favorecieron a López Obrador, lo que disminuyó el margen del puntero panista. El error más común de tales actas, por cierto, fue que se dejó en blanco el número de votos para Alternativa y/o Nueva Alianza —difícilmente un acto doloso, dada la escasa presencia de dichos partidos en zonas rurales.

El cómputo distrital de las actas electorales es el conteo oficial realizado por parte del IFE unos días después de la elección. Durante el cómputo se revisan todas las actas y los paquetes electorales en cada una de las 300 juntas distritales del IFE; también se cotejan las actas originales con las de los partidos y es la etapa en la que los partidos se manifiestan de acuerdo o no con los resultados asentados en ellas. Así, conforme se aprobaba, uno a uno, el cómputo de cada casilla, se transmitían los datos en tiempo real al IFE. La

GRÁFICA 3. Avance relativo de captura de actas para presidente durante el cómputo distrital



Fuente: Elaboración propia con base en datos del IFE a nivel casilla.

duda principal en esta etapa fue: ¿por qué el flujo de datos de cómputo tenía un patrón tan distinto del observado durante el PREP?

La respuesta es que el cómputo distrital es un flujo de datos mucho menos aleatorio que el PREP. Aquí, el ritmo de avance en cada distrito dependía de cuánto se discutía cada una de las actas en la junta distrital. Y si los distritos que favorecían al PAN eran consistentemente más impugnados y debatidos que otros, estos resultados demorarían más en ser computados. Esta dinámica produjo un sesgo partidista en el flujo de datos del cómputo distrital. Si la Coalición tenía ventaja en los primeros distritos computados, ésta tendría que disminuir conforme llegaban los distritos más debatidos hasta revertirse. En vez de urbano, ahora el sesgo fue partidista: a las cuatro de la tarde del 5 de julio se había computado más de 75 por ciento de las actas de estados donde López Obrador ganó, mientras que el avance en estados panistas era menor a 50 por ciento. Este rezago políticamente inducido hizo que el cruce ocurriera en 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 las 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 ofrece la apariencia de tener votos de más o de menos respecto de los ciudadanos que votaron, o bien respecto de las boletas recibidas y/o 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 de manera forzosa realizar sumas o restas ni conciliación numérica alguna. Los estudios sobre el impacto del diseño de las boletas en los errores de los votantes (Kimball y Kropf 2005, Reynolds y Steenbergen, 2006) permiten suponer que el diseño de actas y procedimientos electorales también puede producir errores de escrutinio y cómputo.

Hasta la elección de 2006 lo poco que sabíamos sobre la magnitud y la frecuencia de los errores aritméticos parece sugerir que son un fenómeno

CUADRO 1. Votos promedio según tipo de casilla y afiliación política

	Votos promedio por casilla			
	PAN	CPBT	APM	Margen PAN-CPBT
Casillas urbanas	132.51 (37.79)	124.74 (35.52)	67.11 (20.09)	7.76 (2.27)
Casillas rurales	76.06 (28.49)	87.50 (33.12)	80.59 (31.97)	-11.44 (-4.63)
Estados panistas	145.34 (45.26)	68.82 (22.21)	75.88 (25.73)	76.51 (23.06)
Estados perredistas	89.28 (26.37)	149.36 (45.13)	67.18 (22.16)	-60.07 (-18.77)
Total	114.70 (34.94)	112.84 (34.74)	71.12 (23.78)	1.87 (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)

Fuente: Cálculos propios basados en datos del cómputo distrital del IFE. *Nota:* Porcentajes entre paréntesis.

nuevo, cuando no es así. De hecho, las actas de la elección presidencial de 2006 tuvieron menos errores (46.7 % del total) y menos datos omitidos que las de 2000 (51.4 %). Sin embargo, la magnitud promedio de estos errores es muy similar: 1.26 por ciento de los votos por casilla en 2000 frente a 1.35 por ciento en 2006. En ambas elecciones, los errores aritméticos se distribuyeron de manera casi uniforme o aleatoria en las casillas ganadas por uno u otro candidato, lo cual sugiere errores humanos aleatorios o no dolosos, no intentos de cometer fraude por parte de un partido u otro.

Los errores aritméticos son numerosos porque la consistencia aritmética de las actas es, por así decirlo, frágil: basta una entrada mal contada, regis-

trada u omitida para producir uno o más errores y que no cuadren las cifras. Basta que un funcionario de casilla omita marcar el nombre de un votante en la lista nominal o bien que marque por error un nombre de más en ella para producir la apariencia de votos faltantes o sobrantes. 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, pero no desaparecen del todo. Por ello es muy importante analizar la distribución y la magnitud de estos errores en diferentes tipos de casillas. Si los errores aritméticos de las actas de casilla fueran fruto de una manipulación dolosa de las actas es de esperarse encontrar un sesgo en la distribución de estos errores; ya sea una mayor proporción de errores en las casillas ganadas por el PAN o bien en las casillas ganadas por la Coalición; pero si los errores se distribuyen de manera uniforme o aleatoria, difícilmente pueden provenir de actos dolosos.

Aunque numerosos, los errores aritméticos no son de gran magnitud. Si calculamos las diferencias entre los distintos rubros de las actas de casilla podemos calcular el error promedio para diferentes tipos de error aritmético. El cuadro 2 presenta el error promedio de las actas de escrutinio y cómputo, calculado a partir de los datos del PREP, para tres tipos de errores aritméticos: cuando el voto total difiere del número de ciudadanos que votaron según la lista nominal (A), cuando el voto total difiere del número de boletas depositadas (B) y cuando las boletas depositadas difieren del número de ciudadanos que votaron según la lista nominal (C). Los errores positivos y negativos no se cancelan del todo: en promedio, las actas de casilla indican 4.33 votos menos que el número de ciudadanos que votaron en la lista nominal, mientras que en otros rubros el error típico es positivo (B). El error absoluto promedio por acta fue de 4.36 votos, equivalentes a 1.3 por ciento del voto promedio por casilla (328 votos).

Estos errores producen la impresión de que se trata de votos faltantes o sobrantes, sin embargo, existen múltiples causas de error aritmético que no requieren intervención dolosa. Para explorar un posible sesgo partidista o urbano en la magnitud de los errores promedio, las columnas 2 a 5 del cuadro 2 indican el error promedio en el subconjunto de casillas panistas *vs.* perredistas, y en casillas urbanas *vs.* rurales. 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.*

CUADRO 2. Errores aritméticos promedio en actas de escrutinio para presidente

Tipo de error	Total de actas	Afilación de casilla		Tipo de casilla	
		Panistas	Perredistas	Rurales	Urbanas
A. Voto total-ciudadanos que votaron	-4.33	-3.92	-4.78	-6.03	-3.69
B. Voto total-boletas depositadas	2.04	2.26	1.80	2.74	1.78
C. Boletas depositadas-ciudadanos que votaron	-5.57	-5.39	-5.77	-7.81	-4.75
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

Fuente: Cálculos propios basados en datos del PREP.

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 básicamente aleatorios.

De acuerdo con el cómputo distrital del IFE, antes del recuento del Tribunal Electoral, el margen de victoria para Calderón fue de 243 934 votos, equivalentes a 1.8 votos promedio por casilla (243 934 votos divididos entre 130 788 casillas). Se ha dicho que los errores aritméticos en actas, de alrededor de 4.3 votos por casilla, pueden resultar determinantes ante lo cerrado de un margen de victoria de 1.8 votos promedio (Crespo, 2008). En efecto, los errores aritméticos no son parte de un error muestral que desaparezca al acumular más y más casillas, sino que constituyen un error de medición al contar boletas y asentar los resultados en actas. Para atender esta inquietud hay que considerar, más allá de la magnitud de los errores, de qué manera están distribuidos los errores en diferentes tipos de casillas. 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 modificar el margen de victoria.

El cuadro 3 analiza un tipo de error particular, obtenido de los datos del PREP (única fuente de datos que incluye los llamados campos de referencia necesarios para calcular estos errores). El número de ciudadanos que vota-

CUADRO 3. Errores aritméticos en actas de escrutinio y cómputo de casilla, elección presidencial, 2006

		Casilla ganada por			
		PAN	Por el Bien de Todos	Alianza por México	Total
Actas sin error	Voto total IGUAL a núm. de ciudadanos que votaron	28 159 (47.6)	23 468 (39.7)	7 493 (12.7)	59 120 (100.0)
Actas con error negativo	Voto total MENOR a núm. de ciudadanos que votaron	13 469 (46.8)	12 702 (44.2)	2 590 (9.0)	28 761 (100.0)
Actas con error positivo	Voto total MAYOR a núm. de ciudadanos que votaron	11 780 (46.5)	11 348 (44.8)	2 222 (8.8)	25 350 (100.0)
Casillas totales		53 408 (47.2)	47 518 (42.0)	12 305 (10.9)	113 231 (100.0)

Fuente: Cálculos propios basados en datos del PREP 2006. *Nota:* No incluye casillas con inconsistencias ni actas donde el rubro “número de ciudadanos que votaron” estaba omitido. Porcentajes entre paréntesis.

ron, según la lista nominal, 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 haber quedado fuera del PREP). Como indica el cuadro 3, el PAN ganó en 47.2 por ciento de las casillas al tiempo que obtuvo 47.6 por ciento de las actas libres de error (28 159 de 53 408 casillas ganadas). Por otro lado, 46.8 por ciento de las actas del PAN tuvo error negativo y 46.5 por ciento de sus actas registró error positivo. Por su parte, la Coalición ganó 42 por ciento de las casillas consideradas y sólo 39.7 por ciento de las actas no tuvo error (23 468 de 47 518 ganadas) mientras que concentró 44.2 por ciento de los errores negativos y 44.8 por ciento de los positivos. Es decir, hubo ligeramente mayor número de errores en las casillas ganadas por la Coalición que en las del PAN, pero en general la proporción es muy similar —es decir, un patrón contrario al sugerido por la acusación de fraude— lo que sugiere, de nueva cuenta, errores aleatorios.

En conclusión, ¿qué tan seguros podemos estar de un resultado electoral cuando el proceso mismo de escrutinio y llenado de actas induce a una gran cantidad de errores aritméticos? Ningún argumento estadístico o sistema

electoral ofrece una certeza absoluta ni está exento de errores. La naturaleza misma del proceso electoral en México en el que, en aras de la imparcialidad, el escrutinio de votos se ha confiado a ciudadanos elegidos aleatoriamente, produce un número de errores considerable. Pero en la medida en que estos errores no tengan una distribución o sesgo claramente a favor o en contra de algún partido o coalición, difícilmente constituyen evidencia de una irregularidad dolosa.

Recuento de paquetes

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

Durante el cómputo distrital se recontaron 2 864 paquetes electorales y más adelante el Tribunal Electoral ordenó el recuento de los paquetes de 11 839 casillas más. 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 o inferencias directas, pero al menos ofrece indicios del tipo de ajustes que resultaron tras recontar diferentes tipos de casillas.

Los 2 864 paquetes recontados durante el cómputo distrital eran una muestra sesgada: 66.4 por ciento provenía de distritos panistas y 33.4 por ciento, 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 López Obrador (de 0.6 votos en promedio). Cabe resaltar que en 60 por ciento de los paquetes recontados ni el voto de Calderón ni el de López Obrador tuvieron 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 en distritos dominados por la Coalición, el ajuste promedio por casilla fue de 5.8 votos menos para Calderón y de 13.3 votos menos para López Obrador, lo que aumentó en 7.5 votos el margen del PAN. 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 esta muestra (66.4 por ciento de casillas en distritos panistas). Este es un punto que prácticamente quedó fuera de la controversia sobre el fraude electoral: cuando se recontaron paquetes en los distritos donde aventajaba la Coalición por el Bien de Todos, ésta perdía más votos que los que perdía el PAN en sus propios distritos.

Sin embargo, existe una razón natural por la cual un recuento afecta más al candidato con más votos en una casilla. Usemos un ejemplo: si los errores de conteo ocurren a una tasa más o menos constante, digamos 5 por ciento,

CUADRO 4. Recuento de paquetes durante el cómputo distrital, ajuste promedio entre los datos del PREP y el 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 anulados	-7.339	-7.326	-7.365
Voto total	-17.546	-12.268	-28.035
Margen PAN-CPBT	0.594	-2.864	7.551
Núm. de casillas	2 864	1 901	963
Porcentaje	100.0	66.4	33.6

Fuente: Cálculos propios basados en datos del IFE.

un recuento en una casilla con 200 votos para uno y 100 votos para otro le reducirá 10 votos al primer candidato, pero sólo cinco votos al segundo, lo cual disminuye el margen entre ambos en cinco votos. Para que estos errores aleatorios se cancelen habría que recontar casillas precisamente con un margen opuesto, es decir donde el segundo candidato tenga 100 votos más que el primero; por eso 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 tanto por el IFE como por el TEPJF.

Al ser fruto de la impugnación de la Coalición, las 11 839 casillas recontadas por el TEPJF resultaron ser una muestra aún más sesgada: 91.4 por ciento (10 818) de las casillas provenían de distritos panistas y sólo 8.6 por ciento (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—. Queda por preguntarnos por qué el PAN no impugnó los distritos donde ganó la Coalición. Una respuesta tentativa, y que escapa al análisis presente, estriba en que la estrategia del PAN era minimizar el número de casillas recontadas —quizá temiendo que el recuento cambiaría de manera importante el resultado o que condujera a la anulación de la elección (Crespo, 2008).

En cualquier caso, los ajustes encontrados en una muestra no representativa de casillas no pueden generalizarse para 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 sistemático en el mismo sentido e igual magnitud. Más bien, a la luz de la evidencia disponible, es plausible suponer que un recuento adicional en casillas donde AMLO aventajó habría recompensado la ventaja de Calderón en vez de disminuirla, toda vez que AMLO perdería más votos que el panista.

En todo caso, si uno u otro recuento hubieran arrojado modificaciones sistemáticas respecto al resultado del cómputo distrital respectivo, sugeriría irregularidades generalizadas en el proceso —lo cual no ocurrió—. Más bien, con la evidencia parcial disponible a la fecha, puede suponerse que los *errores de conteo* son de magnitud y distribución similares a los *errores aritméticos* reflejados en las actas, mismos que no parecieron favorecer al candidato ganador de manera sistemática.

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 algunas implicaciones de suponer que estos errores son evidencia de un fraude. Por un lado, alterar dolosamente las actas de casilla requiere contar con una inmensa capacidad de coordinación entre miles de funcionarios en miles de casillas, cuyo alcance llegaría incluso a los bastiones de la Coalición. Por otro lado, resulta ilógico que los mismos funcionarios que alteran la votación de una casilla simultáneamente acusen o pongan en evidencia 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

Aun cuando el recuento de casillas del Tribunal Electoral 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 el 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 por ciento del total), tenemos que Felipe Calderón ganaría por 494 000 votos —lo cual confirma que hubo un poco más de errores en las casillas ganadas por la Coalición, como se ha visto—. 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 por ciento del total).

Como sabemos, el nivel de participación promedio fue de 58 por ciento a lo largo del país. ¿Es posible que el resultado final haya estado determinado por casillas con participación inusualmente elevada? De acuerdo con el

cómputo distrital del IFE, si eliminamos todas las casillas con participación mayor a 75 por ciento de la lista nominal (alrededor de 4 555 casillas, 3.5 por ciento del total) Felipe Calderón mantiene el primer lugar. De este modo, el resultado electoral del 2 de julio es robusto con la eliminación de casillas que presentan errores aritméticos, así como con la eliminación de casillas con una alta tasa de participación. Resulta claro que el TEPJF no puede aplicar criterios de este tipo, pues al final de cuentas su misión es garantizar la integridad del sufragio de los ciudadanos en la medida de lo posible, a menos que cuente con pruebas para lo 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 en los distritos con los resultados electorales, podemos encontrar una respuesta.

En el ámbito nacional el PAN y la Coalición tuvieron representantes en 79.9 y 79.4 por ciento de las casillas, respectivamente, aunque su distribución no fue uniforme: el PAN tuvo representantes en 35 por ciento de las casillas del DF, bastión del PRD, mientras que la Coalición tuvo representantes en 33 por ciento 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 cuatro estados competidos: Puebla, Campeche, Veracruz y Zacatecas.

En efecto, la votación por AMLO en distritos donde no tuvo muchos representantes de casilla fue menor que la de los distritos donde tuvo una mejor cobertura. Sin embargo, la votación por Calderón también fue menor en los distritos donde no tuvo tantos representantes como en sus bastiones. En términos estadísticos, la magnitud de esta relación es similar en ambos sentidos: por cada 10 por ciento adicional de representantes, cada partido obtuvo alrededor de 4 por ciento más votos. Por lo tanto, leer esta relación como evidencia de irregularidades del PAN en contra de 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 que se observen correlaciones como las indicadas en las líneas precedentes.

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

El conteo de votos: imparcialidad vs. precisión

La evidencia de los datos a nivel casilla sugiere 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 estos ciudadanos cometerán errores e imprecisiones al contar votos y llenar actas con la misma aleatoriedad con la que todos lo haríamos. 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 cambios relativamente sencillos como rediseñar las actas para disminuir la frecuencia de los errores aritméticos. De manera más sustantiva, valdría la pena facilitar la realización de un recuento muestral en elecciones con márgenes reducidos, sin el innecesario desgaste de una impugnación generalizada como el visto en 2006. La reforma electoral de 2007 atiende en cierta medida algunas de estas consideraciones, pero de manera muy limitada.

Referencias bibliográficas

- Ackerman, John M. (2007), "The Limits of Transparency: The Case of Mexico's Electoral Ballots", *Mexican Law Review*, 8, julio-diciembre.
- Chappell, Lawson (2007), "How Did We Get Here? Mexican Democracy after the 2006 Elections", *PS: Political Science and Politics*, XL (1), pp. 45-48.
- Crespo, José Antonio (2008), *2006: hablan las actas: Las debilidades de la autoridad electoral mexicana*, México, Debate-Random House Mondadori.
- Estrada, Luis y Alejandro Poiré (2007), "Taught to Protest, Learning to Lose", *Journal of Democracy*, 18 (1), pp. 73-87.
- Kimball, David y Martha Kropf (2005), "Ballot Design and Unrecorded Votes on Paper-based Ballots", *Public Opinion Quarterly*, 69 (4), pp. 508-529.
- Mebane, Walter R., Jr. (2007), "Election Forensics: Statistics, Recounts and Fraud", ponencia preparada para la *Annual Meeting of the Midwest Political Science Association*, Chicago.
- Pliego, Fernando (2007), *El mito del fraude electoral en México*, México, Editorial Pax.
- Poiré Alejandro y Luis M. Estrada (2006), "Allegations of Fraud in Mexico's 2006 Presidential Election", ponencia presentada ante la *102nd Annual Meeting of the American Political Science Association*, Filadelfia.
- Reynolds, Andrew y Marco Steenbergen (2006), "How the World Votes: The Political Consequences of Ballot Design, Innovation and Manipulation", *Electoral Studies*, 25 (3), pp. 570-598.
- Schedler, Andreas (2007), "The Mobilization of Distrust", *Journal of Democracy*, 18 (1), pp. 88-102.
- Tello Díaz, Carlos (2007), *2 de julio: Crónica minuto a minuto del día más importante de nuestra historia contemporánea*, México, Planeta.
- Ugalde, Luis Carlos (2008), *Así lo viví: Testimonio de la elección presidencial de 2006*, México, Grijalbo-Random House Mondadori.