

Análisis estadístico de la elección presidencial: ¿fraude o error humano?

Javier Aparicio*

javier.aparicio@cide.edu

(Por publicarse en *Reportes de Investigación*, CIDE, 2º semestre 2006)

La más reciente elección presidencial puso a prueba tanto a los jugadores como a las instituciones del sistema electoral. Después del 2 de julio, uno de los candidatos perdedores y otros observadores más acusaron la posibilidad de un fraude de diversa índole ejecutado durante la jornada electoral. Más allá de las estrategias políticas de los actores ¿qué hay detrás de estas acusaciones? No es fácil encontrar evidencia de un fraude, pero el análisis estadístico de los resultados de las más de 130 mil casillas ayuda a descartar o apoyar algunas posibilidades.

La concordancia entre el conteo rápido, el PREP y el cómputo distrital, todos disponibles la misma semana de la elección, fueron una señal temprana de una elección confiable. ¿Por qué? Si bien el conteo rápido no apuntaba un claro ganador fuera del margen 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 victoria. Días más tarde el cómputo distrital confirmó esta predicción: el margen victoria fue de sólo 0.58%. Si el cómputo distrital hubiera dado un resultado muy diferente al del conteo rápido, indicaría una posible irregularidad en uno u otro procedimiento.

Ante la escasa información la noche de la elección, hubo quienes leyeron de más en el PREP. Sin embargo, la alteración del flujo de datos mediante un “fraude cibernético” es inútil cuando lo importante es la suma final de votos. Se dijo que el PREP debió haber mostrado múltiples cruces, dada la aleatoriedad inherente del proceso. Sin embargo, el PREP y el cómputo distrital no son procesos muestrales ni completamente aleatorios, como sí lo son los conteos rápidos y las encuestas de salida.

Factores no aleatorios afectan el ritmo de captura de las casillas. Alrededor de 70% de las casillas 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. 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, las cuales fueron capturadas con rezago. El cómputo distrital fue aún menos aleatorio. 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. A las 4pm del 5 de julio se habían computado más del 75% de las actas de estados perredistas, pero sólo el 50% en estados panistas. Por este rezago el cruce ocurrió hasta la madrugada del 6 de julio.

* Doctor en economía, Universidad George Mason, candidato a investigador del SNI, profesor-investigador de la División de Estudios Políticos del CIDE. Nota: Una versión más amplia de esta investigación está en: <http://www.cide.edu/investigadores/aparicio/elecciones/>

Más adelante, los llamados errores aritméticos en actas de escrutinio causaron sospechas. 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 menos sobrantes. Hasta ahora sabíamos muy poco sobre 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) que las del 2000 (51.4%). Sin embargo, la magnitud promedio de estos errores fue similar en ambas elecciones.

Los errores aritméticos no son de gran magnitud: el error promedio por acta fue de ± 4.36 votos. Para ser determinantes, tendrían que estar sesgados a favor de un candidato. Sin embargo, la magnitud y frecuencia de los errores es relativamente similar en casillas ganadas por uno u otro candidato así como en zonas rurales y urbanas, lo cual sugiere errores aleatorios que afectan por igual a los candidatos sin ser determinantes en el resultado final.

Otra fuente de incertidumbre fue si las actas coincidían o no con el contenido de los paquetes electorales. Un análisis estadístico formal requeriría del recuento de una muestra aleatoria y representativa de paquetes. La lógica jurídica casuística del TEPJF no produjo tal muestra. No obstante, la evidencia disponible 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, se trató de una muestra sesgada. 66.4% de los 2,864 paquetes recontados durante el cómputo distrital provenían de distritos panistas. Al recontar casillas panistas, el margen FC-AMLO se redujo en 2.9 votos promedio. Por otro lado, al recontar casillas perredistas, el margen aumentó en 7.5 votos. Es decir, al recontar casillas el candidato con más votos en la casilla pierde más votos tras el recuento. Además, 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 eran una muestra aún más sesgada: 91.4% de las casillas provenían de distritos panistas. Si consideramos que Calderón aventajó a AMLO por 76.5 votos promedio en estados panistas, 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.

En suma, la evidencia sugiere que los errores aritméticos o de conteo no fueron determinantes en el resultado electoral. Al delegar el conteo de votos a ciudadanos elegidos al azar, el sistema electoral garantiza cierta imparcialidad. Como todos, estos ciudadanos cometerán errores e imprecisiones. Es posible diseñar mecanismos con mayor precisión pero difícilmente encontraremos uno más imparcial.