

EL ALGORITMO IMPROBABLE

Javier Aparicio -- javier.aparicio@cide.edu
Profesor-Investigador, División de Estudios Políticos, CIDE
Una versión más amplia de este análisis está en:
<http://www.cide.edu/investigadores/aparicio/elecciones>

En las últimas semanas han circulado diversos análisis—basados en datos divulgados por el IFE—sobre supuestas anomalías del PREP y del conteo distrital, mismas que parecen apoyar la hipótesis de un “fraude cibernético”. Ya sea por errores de interpretación o por partir de supuestos incorrectos, algunos de estos análisis han confundido más que aclarado la situación.

Algunos expertos afirman que debió haber habido múltiples cruces en el PREP, dada la aleatoriedad del proceso. Sin embargo, el PREP y el cómputo distrital son parecidos a un censo: no son procesos muestrales ni 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 el flujo de datos del PREP. Otros factores son menos predecibles: la tasa de participación, si una casilla cerró a tiempo, o si llovía esa noche—pero tales factores difícilmente dejarán huella en un conteo a nivel nacional.

Aún en una elección tan reñida como ésta, 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% de casillas fueron instaladas en zonas urbanas y sólo 30% en rurales. Si uno de los candidatos tenía un voto urbano suficientemente elevado, el PREP le favorecería aún si el orden de captura fuese aleatorio. Pero si las casillas urbanas llegan más rápido a los centros distritales, el sesgo urbano simplemente se refuerza.

Esto fue lo que sucedió con el PREP. Felipe Calderón consiguió 11.8 millones de votos en casillas urbanas (38.2% del total urbano), frente a 11.1 millones de votos de AMLO (35.9%). Por otro lado, AMLO consiguió 3.5 millones de votos rurales (33.5% del total rural), frente a 3 millones de votos de Calderón (29.1%). Como porcentaje, el margen rural de AMLO fue mayor al margen urbano de Calderón, pero no así en votos absolutos: 30% de casillas rurales no fueron suficientes para superar a Calderón en el voto nacional.

¿A qué ritmo se procesaron las actas del PREP? Por causas naturales (no aleatorias) las casillas urbanas llegan más rápido a los centros de cómputo: hacia la medianoche del lunes 2 de julio, alrededor del 50% de las casillas urbanas habían sido capturadas, mientras que las rurales sólo llevaban un avance del 25%. Así, la ventaja inicial de Calderón se vio favorecida por el rezago rural. El voto urbano de AMLO no fue nada despreciable, pero aún los 4.7 millones de votos urbanos que obtuvo en el DF y EDOMEX juntos sólo representan el 15.3% del voto urbano nacional, que en general favoreció a Calderón. Inclusive, viendo estado por estado, el margen urbano promedio de Calderón fue mayor que el de AMLO.

¿Por qué disminuyó el margen de Calderón cuando se incorporaron las 11,184 “actas inconsistentes”? El 65% de estas actas provenían de casillas rurales que en general favorecieron a Roberto Madrazo y a AMLO, disminuyendo el margen del puntero panista. Siendo rurales, en su caso se hubieran capturado poco después que las urbanas. 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 error doloso en zonas rurales.

Al no ser un proceso aleatorio, los resultados del PREP favorecieron al candidato con ventaja en las zonas urbanas. Otro proceso de conteo aún menos aleatorio fue el cómputo que inició a las 8am del 5 de julio en los 300 distritos del país. El orden de conteo de cada acta estaba preestablecido pero el ritmo de avance en cada distrito dependía de cuánto se discutía cada una de las casillas. Si los distritos que favorecían a un candidato eran consistentemente más debatidos, sus resultados demorarían más en ser computados. Este fenómeno produciría un sesgo partidista o regional en el flujo de datos de los cómputos distritales, tal como ocurrió. Si AMLO tenía una ventaja inicial en los primeros distritos, ésta disminuiría—hasta revertirse—conforme llegaban los distritos más debatidos y donde Calderón tenía ventaja.

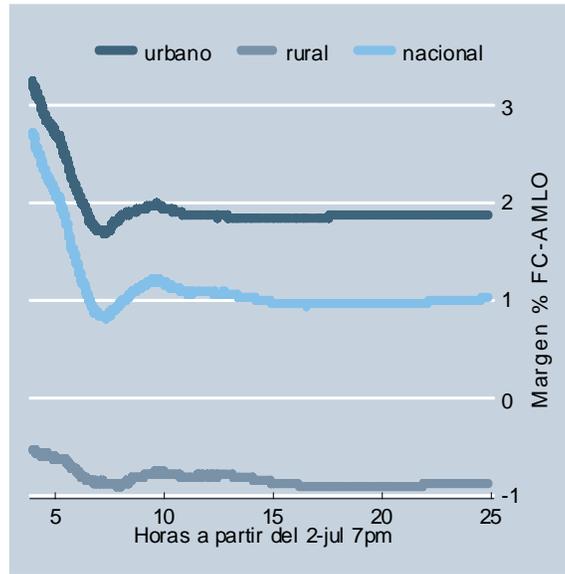
Una vez más el ritmo de avance en los cómputos no fue constante: a las 6pm del miércoles 5 de julio se habían computado 75% de las actas pero el 25% restante tomaría 20 horas más. En vez de urbano, ahora el sesgo fue partidista: a las 2pm del miércoles se habían computado 50% de las actas de estados donde AMLO ganó, y más del 75% para las 4pm. Por otro lado, a las 4pm el avance de las casillas de estados panistas era menor al 50%. Ante ese rezago el cruce esperado ocurrió hasta las 4am del jueves 6 de julio.

Durante las últimas horas del cómputo distrital se observó un efecto espejo entre el porcentaje de votos de AMLO y Calderón: lo que uno perdía el otro lo ganaba. Este fue un fenómeno curioso para una elección multipartidista, pero no inexplicable. El voto relativo de Roberto Madrazo se estabilizó alrededor de las 2am del jueves, cuando ya se habían computado más del 90% de las actas. En las casillas restantes Calderón y AMLO eran primera y tercera fuerza, respectivamente, de modo que sus votos adicionales se contraponían mientras que los de Madrazo no tenían un efecto significativo. La lentitud de los últimos cómputos alargó un proceso que bajo otras condiciones hubiera producido un cruce la misma noche del miércoles.

Finalmente, poco se ha dicho del conteo rápido—instrumento muestral que, a diferencia del PREP, estaba diseñado para pronosticar el resultado. Éste conteo sí era una muestra aleatoria y representativa de casillas que, en efecto, mostró muchos cruces conforme acumulaba datos. El conteo rápido hubiera podido detectar un "vencedor probable" el mismo domingo si el margen era mayor a 0.6%. Pero los ciudadanos decidieron dar una ventaja ligeramente menor: 0.58% que dejó mudo al conteo.

Entre el conteo rápido, PREP y cómputo distrital hay muy pocas diferencias. Tenemos dos explicaciones posibles: Una, que se fraguó un complejísimo algoritmo para manipular los tres procesos sin producir una contradicción entre ellos. Dos, que la elección fue muy reñida pero tuvo un ganador favorecido por el sesgo urbano del PREP y perjudicado, temporalmente, por el sesgo partidista del cómputo distrital. A menudo la explicación más simple es la mejor, independientemente de las convicciones iniciales de los expertos o los candidatos.

PREP: Margen % entre FC y AMLO



Cómputo distrital: % votos FC y AMLO

