



Caracas, 17 de septiembre de 2004

NPS_ 79

Informe Carter sobre Análisis Hausmann & Rigobón no tiene base estadística

El Centro Carter emitió ayer un Informe sobre el Análisis de la representatividad en la Segunda Auditoría y la correlación entre firmantes y votos SI en el Referendo Revocatorio Presidencial del 15 de agosto, en el cual fija su posición sobre el Informe Hausman & Rigobón.

El Informe del Centro Carter rechaza las conclusiones de los profesores Ricardo Hausman y Roberto Rigobón quienes plantean que (1) la auditoría posterior a los resultados del Referendo Revocatorio Presidencial en Venezuela, observada por la OEA y el Centro Carter, no fue hecha sobre una muestra representativa del universo de todos los centros de votación automatizados ni la muestra fue escogida al azar y (2) la prueba estadística que desarrollaron para probar la manipulación de los resultados del Referendo dio positiva. En ambos casos, H&R desarrollaron los análisis estadísticos que sustentan su posición y que concluyen que hubo una probabilidad superior a 99% de manipulación tanto de la muestra aleatoria de la auditoría como de los resultados del RRP.

Para fijar su posición, el Centro Carter presenta un Informe Estadístico que no está avalado por la firma de ningún especialista, pero que argumenta: (1) la correlación de las firmas y los votos, tanto en la muestra general como en la auditoría son idénticos; y (2) la media de los votos en la muestra auditada es similar a la media de la votación.

Ante estos argumentos, el profesor del MIT y uno de los autores del Informe evaluado por el Centro Carter, emitió la siguiente respuesta:

Primer argumento de Carter: "Correlación idéntica de firmas y votos en universo y auditoría"

"El primero de los argumentos es estadísticamente incorrecto. Déjenme explicarles por qué: (1) la correlación de una variable consigo misma es exactamente igual a uno; (2) la correlación de una variable con el 10 por ciento de sí misma también es uno; (3) la correlación de dos variables no cambia si multiplicas las variables por números positivos. En otras palabras, la correlación entre las firmas y los votos por el SI en los datos reales es exactamente la misma que la correlación entre las firmas y los votos fraudulentos, considerando que los votos fraudulentos son iguales a los votos SI menos 10 o 20 o 90 por ciento. Ahora bien, a pesar de que las correlaciones son idénticas el número total de votos no lo es.

Veamos un ejemplo: supongamos que en una empresa hay 2 personas que ganan el mismo salario, digamos Bs. 1.000.000 y le aumentan anualmente los salarios en 20%. Estas dos personas siempre van a ganar lo mismo porque sus sueldos están correlacionados en aumentos de 20% anuales. Pero si en esa oficina hay una persona que gana la mitad, Bs. 500.000, y también le dan aumentos anuales de 20%, su sueldo también está correlacionado con el de las otras dos personas porque siempre le aumentan 20% aunque gana la mitad de los otros...Es decir, las correlaciones son idénticas pero los sueldos no.

Segundo argumento: "media de votos en muestra auditada es similar a la media de la votación"... Un sofisma estadístico

Este argumento es también, estadísticamente, incorrecto. Déjenme explicarles cómo se pudo haber hecho un fraude que mantenga la misma media (o parecida) entre la muestra auditada y la no auditada. Ahora bien, como esto es un caso ficticio no vamos a usar como ejemplo el Referéndum en Venezuela, sino usaremos una supuesta



elección presidencial en Florida.

Supongamos que en Florida existe la siguiente distribución de votantes: la mitad de los centros de votación automatizados tiene una mayoría demócrata, y la otra mitad tiene una mayoría republicana... fácil, mitad y mitad. ¿Cómo sabemos que hay unos centros republicanos y otros demócratas? Muy fácil, conocemos el resultado de votaciones previas y, además, sabemos cuántos demócratas y republicanos están inscritos en cada centro de votación automatizado. Ahora, supongamos que se va a hacer un fraude en la mitad de los centros demócratas, es decir en una cuarta parte del total los centros. En estos centros las máquinas de votación convierten 10 por ciento de los votos demócratas en republicanos. Y supongan, solo para simplificar, que este cambio de votos demócratas en republicanos es suficientemente grande como para convertir el centro de demócrata a republicano. Esto es lo que podríamos llamar a esto fraude.

El resultado oficial de esta elección es que $\frac{3}{4}$ de los centros son republicanos y $\frac{1}{4}$ son demócratas. Ahora bien, el problema radica en cómo generar la muestra aleatoria para que no sea detectada en la auditoria de los observadores internacionales. Para entender esto es importante saber que los programas que generan números aleatorios necesitan un número inicial que se introduce en la computadora y, a partir del cual, el programa genera el resto de la secuencia aleatoria. Este número se llama la semilla o seed. Como indica el Centro Carter, el programa usado por el CNE genera exactamente la misma muestra **si parte de la misma semilla**. Por lo tanto, es fácil que alguien genere un set de números aleatorios antes de las elecciones con una mayor representatividad de centros donde naturalmente ganan los republicanos, en otras palabras que haya $\frac{3}{4}$ partes republicanos y $\frac{1}{4}$ parte demócrata. Simplemente tengo que acordarme cuál es ese "número semilla" y lo introduzco para que escoja los centros de votación donde hay que hacer la auditoría. Conclusión: en los centros con mayoría demócrata no incluidos en la auditoría puedo alterar los resultados disminuyendo los votos de los demócratas. ¿Resultado final? En estas elecciones, 75 por ciento de las mesas son republicanas y 75 por ciento en la auditoria. Lo interesante



de este ejemplo es que el Centro Carter estaría dispuesto a decir que la muestra de la auditoría de esta elección es representativa.

Pensemos en el Kino como ejemplo: todos los domingos introducen 25 bolitas en la máquina para escoger 15. Pero si "cargan 15 números", es decir las bolitas son más pesadas, esas serán las que saldrán siempre, pero si la mano peluda que las cargó sabe dónde enviaron los cartones con esos números va y los compra y se gana todos los premios!

No se puede demostrar el fraude comparando medias ni comparando correlaciones. ¡Por favor! Eso no pasa el examen más elemental de estadística. Se deben comparar medias y correlaciones CONDICIONALES. Lo que demuestra el fraude en el ejemplo anterior no es que las medias son diferentes, sino que son condicionales de la información que se tiene sobre la muestra - cuántos centros de votación son demócratas y cuántos republicanos- por lo tanto, la muestra auditada tiene una correspondencia más cercana con esa información que la muestra no auditada. Esto es exactamente lo que demostramos Ricardo Hausmann y yo.

Una manipulación mal hecha posiblemente falle en las dimensiones en las cuales el Centro Carter se está basando, pero no subestimemos a quienes hicieron la manipulación" concluye Rigobón.