

CUANDO LA DISYUNTIVA ELECTORAL TRANSGREDIÓ LOS ESPACIOS Y LÍMITES DE LA CIENCIA Y DE LA METODOLOGÍA

Oscar Viloría H.¹

ESCUELA DE ESTUDIOS INTERNACIONALES, UCV

... cuando los hechos son modelados de forma tal que la hipótesis se la hace aparecer como tesis, el método pasa a ser un instrumento perverso, amoral, de dominación ideológica.

Resumen:

El artículo resume un conjunto de observaciones hechas, de agosto de 2004, al informe "En busca del cisne negro". En aquel momento, tales observaciones tenían como objetivo único objetivar y ponderar metodológicamente las denuncias y presunciones de fraude electrónico apoyadas en el análisis estadístico, además de otros métodos cuantitativos. A su vez, estas observaciones tuvieron su origen en los juicios y las apreciaciones hechos al proceso electoral, y en la metodología aplicada en el informe señalado. Fue así entonces, cuando el trabajo que antecede al presente artículo, pretendió aislar toda influencia y ruido que, sobre la población venezolana, introducían las debilidades *post hoc* y de *composición*, envueltas en el informe en cuestión. Luego, siendo este último el objeto de análisis, el proceso electoral de agosto de 2004 viene a representar, a penas, un elemento que subyace dentro del objeto de análisis; limitando, en todo momento, la posibilidad de inferir o presumir la ausencia o presencia de ilícitos electrónicos durante el referendo revocatorio presidencial, de la fecha señalada. Por otro lado, el artículo no advierte ni compromete, la capacidad y experticia de sus autores. De hecho, presupone todo origen de las debilidades metodológicas, a la urgencia propia de la coyuntura política, al momento de su publicación.

Palabras claves: Cisne negro, referendo revocatorio, fraude electoral y fraude electrónico.

INTRODUCCIÓN

El trabajo que antecede al presente artículo incluyó, en su *primera parte*, el texto original del informe "En busca del cisne negro", combinado con una secuencia de notas al pie de página donde se reflejaron las observaciones originadas a partir de los juicios y apreciaciones envueltos en el informe señalado. A su vez, en la *segunda parte*, se registró un conjunto de observaciones que emanaron de la metodología aplicada.

Debido al calor político de la coyuntura, el trabajo fue titulado: "... y resultó que al cisne lo tiñeron de negro"; transmitiendo la idea de que los objetivos perseguidos por el informe en cuestión, no se alcanzaron. Hoy en día, ante una apa-

¹ oscarviloriah@yahoo.es

rente calma y madurez política, el presente artículo se publica con fines exclusivamente académicos.

En aquel momento, las observaciones no fueron más que la respuesta natural, académica, al informe presentado por los profesores Ricardo Hausmann y Roberto Rigobon, quienes, en voz del primero, restringieron toda opinión al plano académico.

Ahora bien, en la primera sección del informe "*En busca del cisne negro*" se establecen los motivos, premisas e hipótesis que soportan, la especificación de un modelo lineal de dos ecuaciones; pretendiendo, a partir de su estimación, explicar las diferencias observadas entre los votos SI y las intenciones de voto correspondientes, para una muestra de tamaño igual a 342 centros de votación.

La primera ecuación, sujeta los resultados que favorecieron a la opción SI en agosto de 2004, a los registros de intención de voto que se recolectaron en el *reafirmazo* de noviembre de 2003, conjuntamente con los registros de nuevos votantes y de participación electoral; en todo centro de votación perteneciente a la muestra.

Por su parte, la segunda ecuación sintetiza la probabilidad condicional entre la variable endógena de la primera ecuación y las variables que registran las intenciones de voto a partir de las encuestas de salida (*exit polls*), los nuevos votantes y la participación electoral; también, para todo centro de votación que pertenece a la muestra.

En el primer capítulo de la versión anterior a éste, se enumeró y argumentó el conjunto de observaciones que originaba la dialéctica presente en cada apreciación o juicio relacionado a los sucesos político-electorales del *firmazo*, *reafirmazo*, *referendo revocatorio* y los *exit_polls*. Lamentablemente, las restricciones de espacio en esta publicación, obligaron a omitir ese primer capítulo.

Entretanto, las observaciones que se exponen a continuación están dirigidas, exclusivamente, a la metodología e instrumentación aplicada en la especificación y estimación de las ecuaciones lineales, a la interpretación de los resultados y, como una consecuencia directa, al alcance y limitaciones técnicas que pretenden fundamentar las evidencias que presumen un fraude electrónico en el proceso electoral de agosto de 2004.

Por último, resulta pertinente, más no limitante, tener a la mano una copia del informe "*En busca del cisne negro*". De esta forma, el lector podrá aclarar toda duda, además de verificar la autenticidad de las citas que incorpora el presente artículo.

1.- DEBILIDADES POST HOC, DE DIVISIÓN Y DE COMPOSICIÓN

1.1.- Sobre el origen de la confusión

Como es del conocimiento de todos, el resultado electoral del 15 de agosto de 2004, desencadenó un debate que en algunos casos se apoyo en argumentos estadísticos y probabilísticos, a favor o en contra de la legalidad y transparencia de los procedimientos aplicados antes, durante y después del *referendo revocatorio presidencial*.

Cualquiera que hallan sido las explicaciones y métodos² utilizados para fortalecer científicamente la presunción de fraude electoral, todas partieron de un factor base que denominamos *originario*: *Los resultados del referendo revocatorio, como un fenómeno aleatorio, se presentan como una distribución de probabilidad binomial con parámetros n , ρ y $(1-\rho)$; donde ρ representa la probabilidad de ocurrencia una de las alternativas y $(1-\rho)$, la probabilidad de ocurrencia de la otra. Por su parte n , el número de repeticiones, corresponde a la participación electoral.*

El uso e interpretación generalmente dado, a la afirmación anterior, degenero en la falaz idea que afirma: *si los resultados totales siguen esta distribución, es porque los resultados parciales³ así lo hicieron*; aseveración que constituye una clara *falacia de división*, cuando se trata de una posibilidad *equis probable* como cualquier otra donde se observen distribuciones parciales que se desvían de la total. Luego, que la afirmación no sea axiomática; y todo estudio que se sujete a una incompleta visión teórica y metodológica del problema, condenará, de hecho, los resultados de la investigación; esto, sin importar el método de análisis que haya sido aplicado.

Ahora bien, para despejar toda confusión subyacente en el problema, resulta útil, como isomorfismo, la idea del *universo* y de la materia y energía que envuelve, en un momento determinado. *El universo* se encuentra constituido por toda la energía que encierra, no importa si ésta llega a ser percibida, o no, en su totalidad⁴. Al mismo tiempo, según el principio de conservación, la energía total que envuelve el universo debe ser la misma en todo momento. Se acepta entonces la existencia de los confines del universo, porque percibimos una parte de

² Para consultar sobre modelos y distribuciones de probabilidad, léase Chou (1991).

³ Por centro electoral y también por mesa de votación.

⁴ Se existe así no se perciba, "ser, es ser en sí".

éste, mientras que el resto, lo que no alcanzamos a percibir, lo aceptamos por *ser en si mismo*.

No obstante, dentro del contexto electoral, es imposible afirmar que los resultados totales existen. Lo que se observa son los resultados o distribuciones parciales, por centro de votación; siendo su agregado, el que viene a constituir una medida que puede desviarse de algunas de las distribuciones individuales.

Será entonces la distribución de los totales, la que puede desviarse de las distribuciones parciales, y toda desviación que presenten las distribuciones parciales con respecto a los resultados oficiales "*afirmativos*", a las distribuciones parciales y a la identidad que relaciona, por hipótesis, las intenciones de voto y los votos *SI*, no representarán en momento alguno, una evidencia probabilística que permita presumir un fraude electrónico sobre los votos *SI*; ya sea directamente o no proporcional. Al mismo tiempo, las mismas observaciones no implican presunción alguna de ausencia de fraude electrónico.

Un ejemplo de lo anterior se verifica en el Gráfico 3c de la página 14. La distribución observada en el gráfico permite, solamente, inferir que el mejor ajuste lineal no viene dado por una sola recta.

En síntesis, la naturaleza aleatoria del voto viene determinada por las opciones electorales, mientras que la desviación estadística del total con respecto a las distribuciones parciales, además de las que observen estas últimas entre sí, determinará la naturaleza aleatoria del proceso electoral.

La distribución de probabilidad que siguen las opciones electorales en proporción al total de votos, es el resultado de agregar las 4580 distribuciones parciales, cada una como porcentaje de su *padrón electoral*; siendo estas últimas las que constituyen variables aleatorias, y su agregado, no tendrá porque reflejar el comportamiento parcial de algún centro en particular.

A su vez, las distribuciones parciales podrán presentar desviaciones de un centro de votación a otro. Suponer lo contrario, sería negar la naturaleza estocástica del proceso electoral.

Es así entonces, cuando las desviaciones observadas de un centro a otro, no pueden ser interpretadas como una evidencia para presunción el fraude electrónico. Pero, lo anterior tampoco implica la ausencia del ilícito electoral.

Por último, supondremos que el problema se origina de una confusión, y no, de una intención para que los lectores acepten como axiomático que las desviaciones observadas por los resultados, de un centro de votación a otro, de una

máquina de votación a otra dentro de un mismo centro y, por último, entre las intenciones de voto “afirmativo” y los resultados oficiales que favorecieron la opción del SI, representen una anomalía que sólo pueda ser la consecuencia de una manipulación de los resultados; premisas que niegan la naturaleza aleatoria de un proceso que contraponen dos opciones electorales.

1.2. - El manejo metodológico de los datos

Los resultados aleatorios que origina un fenómeno aleatorio pueden no seguir un orden esperado o predeterminado. Además, los resultados que aleatoriamente genera un proceso vendrán a ser solamente eso, resultados.

Un número aleatorio, por ejemplo, puede ser el resultado de un proceso que genera, de una base aleatoria, números. Posteriormente, una vez generado, el número aleatorio pasará a ser simplemente un número; pero su generación aleatoria ha de determinar, y así se espera, una secuencia numérica que no presente comportamientos predeterminados. Supongamos que un proceso genera aleatoriamente la siguiente secuencia:

$$5 - 1 - 3 - 21 - 0 - 1 - 8 - 5 - 2$$

Posteriormente, la secuencia es ordenada ascendentemente:

$$0 - 1 - 1 - 2 - 3 - 5 - 8 - 13 - 21$$

Se observa entonces, que la secuencia resultante contiene los nueve primeros términos de la *Serie de Fibonacci*, sin que estos hallan sido generados en ese orden.

En síntesis, cuando una secuencia de números aleatorios se ordena bajo un criterio predeterminado, se corre el riesgo de que la misma quede sujeta a una *invariante* o tendencia particular, perdiendo así, su naturaleza aleatoria y su capacidad para modelar fenómenos aleatorios.

Luego, si aceptamos el “*comportamiento caótico de un fenómeno*”, como una realidad aún no pensada e interpretada, entonces será posible entender que *el orden* se inicia, también, cuando la observación sigue algún criterio predeterminado. En consecuencia, el *desorden*, también podrá tener su origen en el desconocimiento de los mecanismos que determinan el patrón de comportamiento.

Es así entonces cuando advertimos, que el sorteo de las variables aleatorias *Votos_SI*, *Firmas* y *exit_polls*, bajo un criterio común, puede desembocar en una correlación inducida para las tres variables aleatorias y, a su vez, en la imposibi-

alidad para inferir o presumir, la presencia o ausencia de fraude electrónico, en el proceso electoral del 15 de agosto de 2004.

En efecto, cuando analizamos la estructura AR1 de la variable *FRAUDE* (la diferencia entre los *exit_polls* y los votos SI) se observa que la misma sigue proceso "auto-regresivo no-estacionario"⁵ que traduce la sobrestimación que las encuestas de salida tienen sobre los votos S_i^6 (Tabla No. 1).

Tabla No. 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	1.004989	0.000398	2527.825	0.000000
R-squared	0.999796	Mean dependent var		24751.98
Adjusted R-squared	0.999796	S.D. dependent var		15460.55
S.E. of regression	220.7776	Akaike info criterion		13.63492
Sum squared resid	17791.106	Schwarz criterion		13.64558
Log likelihood	-2494.190	Durbin-Watson stat		1.586856
Inverted AR Roots	1.00			
Estimated AR process is nonstationary				

Fuente: Elaboración propia.

La ecuación del cuadro anterior en su forma estructural será:

$$\text{Fraude}_i = \beta \cdot \text{Fraude}_{i-1} + \mu_i$$

Luego,

$$\text{Fraude}_i = 1,0049 \cdot \text{Fraude}_{i-1} + \mu_i$$

Y, por definición,

$$(\text{Intención} - \text{Votos_SI})_i = 1,0049 \cdot (\text{Intención} - \text{Votos_SI})_{i-1} + \mu_i$$

⁵ Sobre los modelos *auto-regresivos*, $AR(q)$, se sugiere consultar Gujarati (1997), Maddala (1996), Sánchez (1999), Novales (1993) e Intriligator (1990). Entretanto, para consultar sobre modelos ARMA, léase Álvarez (2002) y Pérez (2006).

⁶ Contrario al argumento expresado en la página 17 del informe "En busca del cisne negro": "...en aquellos sitios en los que las firmas se equivocan proporcionalmente más en el sentido de predecir más votos SI que los obtenidos, las encuestas de salida también sobre-estiman relativamente más los votos alcanzados. Dado que ambas medidas son independientes, la implicación es que lo que tienen en común es el fraude".

$$\text{Intención}_i = 1,0049 \cdot \text{Intención}_{i-1} + (\text{Votos_SI}_i - 1,0049 \cdot \text{Votos_SI}_{i-1}) + \mu_i$$

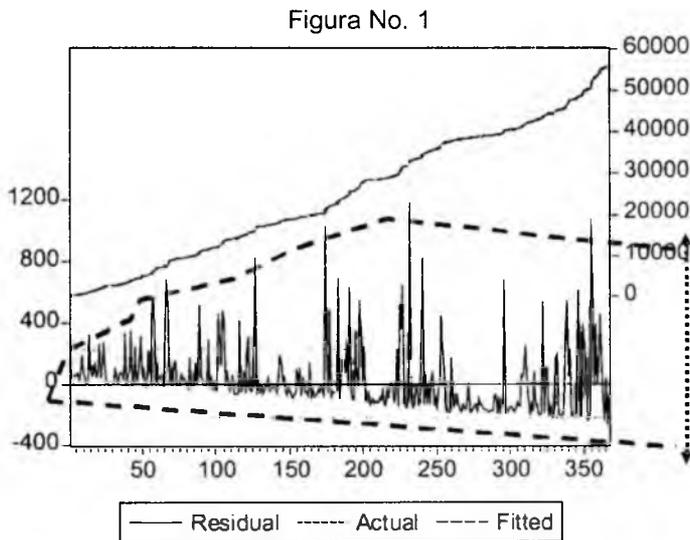
Al mismo tiempo,

$$\text{Votos_SI}_i = 1,0049 \cdot \text{Votos_SI}_{i-1} + v_i;$$

Donde, v_i sigue un proceso "auto-regresivo de orden 1, no-estacionario".

$$v_i = \text{Intención}_i - 1,0049 \cdot \text{Intención}_{i-1} + \mu_i$$

Este fenómeno de inestable y divergente que experimentan los residuos de la ecuación, es verificable, gráficamente, a partir del comportamiento "explosivo" que exhiben los residuos en siguiente gráfico.



Fuente: Elaboración propia

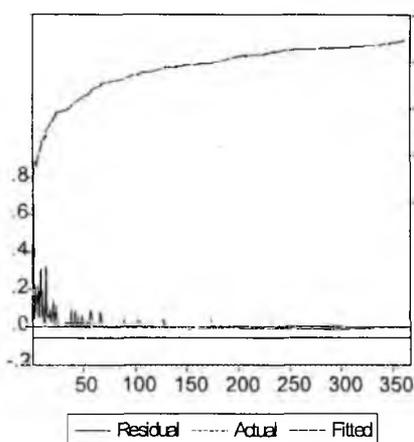
En el siguiente resumen estadístico (E-Views) se verifica que el logaritmo de la variable *FRAUDE*, también sigue un proceso *auto-regresivo, no-estacionario* (Tabla No. 2, Figura No. 2 y 3).

Tabla No. 2

Dependent Variable: LOG(FRAUDE)				
Method: Least Squares				
Date: 11/03/07 Time: 22:27				
Sample(adjusted): 2 367				
Included observations: 366 after adjusting endpoints				
Convergence achieved after 2 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	1.001383	0.000279	3590.25	0.000000
R-squared	0.997741	Mean dependent var		9.746613
Adjusted R-squared	0.997741	S.D. dependent var		1.099609
S.E. of regression	0.052264	Akaike info criterion		-3.062291
Sum squared resid	0.997006	Schwarz criterion		-3.051628
Log likelihood	561.3992	Durbin-Watson stat		0.900374
Inverted AR Roots	1.00			
Estimated AR process is nonstationary				

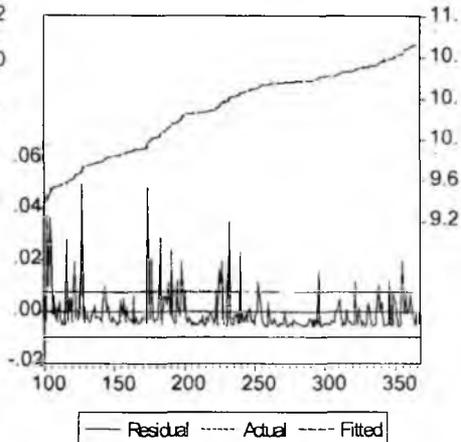
Fuente: Elaboración propia.

Figura No. 2



Fuente: Elaboración propia.

Figura No. 3



Fuente: Elaboración propia.

Ahora bien, cuando la muestra anterior es sorteada aleatoriamente, se observa que tanto el problema de *heterocedasticidad*, como la *varianza no estacionaria*, se corrigen; al menos en los tres casos siguientes (Tablas 3, 4 y 5; Figuras 4, 5 y 6).

Tabla No. 3

Dependent Variable: LOG(FRAUDE2)				
Method: Least Squares				
Date: 10/25/07 Time: 00:13				
Sample(adjusted): 2 366				
Included observations: 365 after adjusting endpoints				
Convergence achieved after 2 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.997522	0.003149	316.7475	0.000000
R-squared	0.728915	Mean dependent var	9.733627	
Adjusted R-squared	0.728915	S.D. dependent var	1.133140	
S.E. of regression	0.589979	Akaike info criterion	1.785276	
Sum squared resid	126.6594	Schwarz criterion	1.795961	
Log likelihood	-324.8129	Durbin-Watson stat	2.061069	
Inverted AR Roots	1.00			

Fuente: Elaboración propia.

Figura No. 4

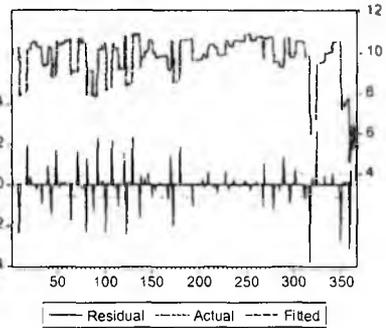


Tabla No. 4

Dependent Variable: LOG(FRAUDE3)				
Method: Least Squares				
Date: 10/25/07 Time: 00:16				
Sample(adjusted): 2 366				
Included observations: 365 after adjusting endpoints				
Convergence achieved after 2 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.997685	0.00316	315.6827	0.000000
R-squared	0.726731	Mean dependent var	9.731558	
Adjusted R-squared	0.726731	S.D. dependent var	1.132140	
S.E. of regression	0.591831	Akaike info criterion	1.791544	
Sum squared resid	127.4960	Schwarz criterion	1.802229	
Log likelihood	-325.9569	Durbin-Watson stat	2.130095	
Inverted AR Roots	1.00			

Fuente: Elaboración propia.

Figura No. 5

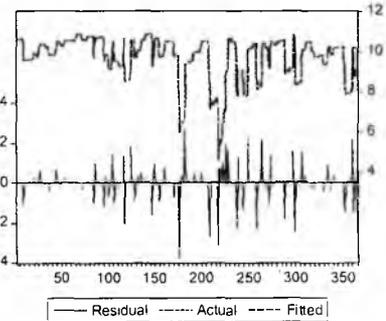
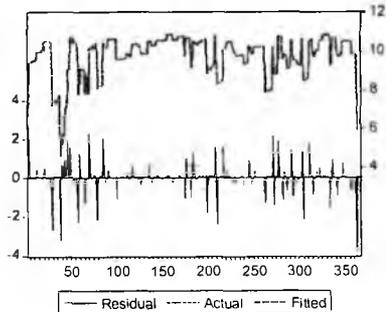


Tabla No. 5

Dependent Variable: LOG(FRAUDE4)				
Method: Least Squares				
Date: 10/25/07 Time: 00:15				
Sample(adjusted): 2 366				
Included observations: 365 after adjusting endpoints				
Convergence achieved after 2 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.998113	0.003208	311.1024	0.000000
R-squared	0.71294	Mean dependent var	9.742394	
Adjusted R-squared	0.71294	S.D. dependent var	1.121916	
S.E. of regression	0.60110	Akaike info criterion	1.822624	
Sum squared resid	131.5208	Schwarz criterion	1.833309	
Log likelihood	-331.629	Durbin-Watson stat	2.117396	
Inverted AR Roots	1.00			

Fuente: Elaboración propia.

Figura No. 6



Nótese que en estos tres casos, la variable $LOG(FRAUDE)$ conserva su estructura *auto regresiva*, pero esta vez, los residuos no presentan una tendencia

en particular, y su varianza no diverge. En otras palabras, la distribución aleatoria de las observaciones es una condición necesaria, pero no suficiente, para obtener estimadores MELI (*Mejor Estimador Linealmente Insesgado*).

Cuando los residuos no siguen una tendencia particular y su varianza es estable, entonces la magnitud en que difieren los votos *SI* con respecto a la intención de voto, será independiente del tamaño de los centros y cuadernos de votación; desapareciendo la sobrestimación inducida sobre los votos *SI*.

2.- OBSERVACIONES METODOLÓGICAS

2.1.- A la especificación

En la página 15 del informe "En busca del cisne negro" se especifica la ecuación que relaciona los votos oficiales a favor de la opción *SI* con los resultados del *re_afirmazo*, para una muestra de 342 centros electorales:

... Por razones técnicas, preferimos la forma logarítmica², sin embargo realizamos el análisis de las tres maneras para ver si nuestros resultados dependen de la forma funcional que adoptamos.

Un ejemplo de las estimaciones realizadas utilizando la forma logarítmica se presenta en el Cuadro 4. La ecuación estimada es:

$$LSI = a + b \cdot LFIRMA + c \cdot elc_now + d \cdot VEL + error$$

donde LSI es el logaritmo del número de votos *SI*, LFIRMA es el logaritmo del número de firmas en cada centro, *elc_now* es el porcentaje de nuevos votantes VEL es el porcentaje de participación electoral, y donde a, b, c y d son parámetros a ser estimados.

Más adelante, continúa:

... El parámetro b es la elasticidad entre firmas y votos y está estimado en casi 1 (en realidad es 0.994). Ello implica que si un centro tiene el doble de firmas que otro centro, obtiene en promedio el doble de votos.

A su vez, en la página 16 del mismo informe, se establece la relación⁷ entre los votos oficiales a favor de la opción *SI* y los *exit_polls*:

⁷ El presente artículo supone un error de transcripción dentro del informe evaluado. Según el "Cuadro 5" de la página 17, la ecuación debió ser expresada como sigue: $LSI = f + g \cdot lex_si + h \cdot VEL + j \cdot elc_now + error$

La segunda ecuación que estimamos es la relación entre votos y exit polls también para los 342 centros para los que tenemos datos de éstos. La ecuación que estimamos es:

$$LSI = f + g \cdot \text{lex_si} + h \cdot \text{VEL} + j + \text{error}$$

Donde *lex_si* es el número de votos *SI* que predice la encuesta para ese centro. Las letras *f*, *g*, *h* y *j* son parámetros mientras que *LSI* y *VEL* ya han sido definidos.

Y más adelante, en la página 17:

... La elasticidad estimada de las intenciones de acuerdo a las encuestas con los votos es de 0.97.

Se considera, entonces, de extrema importancia, definir el *coeficiente de elasticidad* entre el logaritmo de los votos *SI* y el logaritmo de los exit_polls, como la proporción de reacción o grado de sensibilidad de las ocurrencias del fenómeno electoral (logaritmo de los votos *SI*, *LSI*), cuando el logaritmo de los exit_polls (la variable *lex_si*) varía entre centros de votación contiguos, dentro de la muestra seleccionada.

En otras palabras, el coeficiente que mide el cambio porcentual experimentado por los votos *SI* de dos centros de votación contiguos, por cada uno por ciento de diferencia en la intención de voto *SI*, para el mismo par de centros dentro de la muestra.

Ahora bien, para que el coeficiente de regresión parcial de las variables logarítmicas pueda ser considerado como una medida de elasticidad, todas las variables o regresores del modelo deben estar expresados en las mismas unidades, en este caso, en término de sus logaritmos.

De hecho, la forma funcional y unidades que observan las dos ecuaciones simula⁸ un *modelo auto-regresivo de orden p*, donde la variable regresada (votos *SI*) y los regresores (nuevos votantes y participación electoral) se especifican como la primera diferencia de sus logaritmos. En ese caso, la especificación debió ser:

$$LSI_i = \beta_1 \cdot LSI_{i-1} + \beta_2 \cdot (\text{lex_si}_i - \text{lex_si}_{i-1}) + \beta_3 \cdot \text{VEL}_i + \beta_4 \cdot \text{elc_now}_i + \mu_1$$

Donde, las variables *VEL_i* y *elc_now_i* registran la diferencia porcentual entre centros de votación contiguos, de la participación electoral y de los nuevos electores, respectivamente. Por su parte, las variables *LSI_i* y *LSI_{i-1}*, miden el logarit-

⁸ Sin que así sean.

mo de los resultados oficiales a favor de la opción *SI*, para el mismo par de centros de votación.

Se reconoce entonces, la habilidad y la maestría metodológica que hace falta para simular la forma funcional de un *proceso auto-regresivo AR(1)* a partir de las desviaciones entre las variables *LSI* y *lex_si* de cada centro de votación de la muestra, y el ordenamiento predeterminado, no aleatorio, de estas variables.

No obstante, cuando la ecuación no incorpora el componente *auto regresivo* $\beta.LSI_{i-1}$, significa que los resultados electorales de cada centro de votación están desvinculados o no están sujetos al comportamiento observado en otros centros de votación de la muestra y, en consecuencia, al *universo electoral*. Luego, será imposible cualquier inferencia (presunción) sobre los resultados generales del proceso electoral, fundamentada a partir de los resultados o distribuciones individuales.

En síntesis, las debilidades metodológicas observadas en la especificación de las ecuaciones abren toda posibilidad para rechazar el modelo, los resultados⁹ y, por último, la causalidad¹⁰ inferida en el informe.

2.2.- A la estimación

Resulta obvio que las observaciones a la especificación de las ecuaciones, condenan los valores esperados de la estimación y toda interpretación derivada a partir de estos. No obstante, la posibilidad de encontrar lectores que den menor importancia a los errores de especificación¹¹, en esta sección se resume y analizan los resultados de la estimación como un ejercicio práctico y metodológico.

La Tabla No. 6, transcribe el "Cuadro 4" (página 16 del informe) de donde se extraen los valores esperados de los coeficientes de la ecuación:

⁹ *Bondad de ajuste, significancia, estabilidad y consistencia*, entre otras estadísticas de los coeficientes.

¹⁰ Si los datos no son homogéneos, sólo será posible aceptar la correlación entre las variables, pero en ningún momento, relaciones de determinación; aún, cuando los coeficientes resultasen significativos. Si este es el caso, la correlación vendrá dada por las tendencias que envuelva cada variable del modelo, expresadas éstas en sus unidades de medida correspondientes.

¹¹ Siendo la especificación correcta de las ecuaciones, algo fundamental para todo modelo.

$$LSI = a + b*LFIRMA + c*elc_now + d*VEL + error$$

a = +0,3059669; b = +0,9942821; c = +0,4604462; d = +0,3311808;

Tabla No. 6

Cuadro 4. Estimación de la ecuación entre votos y firmas, electores nuevos y participación electoral						
Source	SS	Df	MS			
Model	185.800888	3	61.9336295	Numbers of obs	=	342
Residual	5.84296339	338	.017286874	F(3,338)	=	3582.70
				Prob > F	=	0.0000
				R-squared	=	0.9695
				Adj. R-squared	=	0.9692
				Root MSE	=	.13148
Total	191.643852	341	.56200543			
LSI	Coef.	Std. Err.	t	P> t	[95% Conf.	Interval]
LFIRMA	.9942821	.0099034	100.40	0.000	.974802	1.013762
elc_now	.4604462	.0375	12.28	0.000	.3866834	.5342089
VEL	.3311808	.0813913	4.07	0.000	.1710835	.4912781
cons	.3059669	.0782436	3.91	0.000	.1520611	.4598727

Fuente: Informe "En busca del cisne negro", (2004: 16).

Se lee entonces que, a partir de una muestra de 342 centros de votación (Number of obs = 342):

- La varianza conjunta de los regresores {LFIRMA, elc_now, VEL} se aproxima a la varianza muestral del Logaritmo de los votos oficiales a favor de la opción SI {LSI}, con una bondad de ajuste R-squared igual a 0.9695.
- La significancia conjunta de los coeficientes de regresión parcial presenta un alto nivel de aceptación f(3, 338) igual a 3582.7
- Individualmente, el valor esperado de los coeficientes de la regresión presenta intervalos de confianza que excluyen el cero como valor probable¹², siendo consistente la dirección y el sentido de la relación entre la variable regresada LSI y sus regresores.

Por su parte, el "Cuadro 5" (página 17 del informe) resume los resultados de la estimación para la ecuación (para efectos del presente artículo-Tabla No. 7):

$$LSI = f + g*lex_si + h*VEL + j*elc_now + error$$

f = +0,0722736; g = +0,9701892; h = +0,4244489; j = -0,6612884;

¹² P > |t| = 0.000 para cada uno de los coeficientes y (t_a = 3.91), (t_b = 100.4), (t_c = 12.28) y (t_d = 4.07);

Tabla No. 7

Cuadro 5. Estimación de la ecuación entre votos y firmas, electores nuevos y participación electoral						
Source	SS	Df	MS			
Model	157.862978	3	52.6209927	Numbers of obs	=	342
Residual	33.7808737	338	.099943413	F(3,338)	=	526.51
				Prob > F	=	0.0000
				R-squared	=	0.8237
				Adj. R-squared	=	0.8222
Total	191.643852	341	.56200543	Root MSE	=	.31614
LSI	Coef.	Std. Err.	t	P> t	[95% Conf.	Interval]
lex_si	.9701892	.025357	38.26	0.000	.9203118	1.020067
elc_now	-.6612884	.0868377	-7.62	0.000	-.8320987	-.490478
VEL	.4244489	.1957766	2.17	0.031	.0393549	.8095429
cons	.0722736	.2086177	0.35	0.729	-.3380789	.4826261

Fuente: Informe "En busca del cisne negro" (2004: 17).

Los resultados señalan que a partir de una muestra de 342 centros de votación (Number of obs = 342):

- La varianza de las variables explicativas: *lex_si*, *elc_now* y *VEL*, que representan el logaritmo de los exit_polls a favor de la opción electoral SI en cada centro de votación, la variación porcentual inter-centros de los nuevos electores y la variación porcentual inter-centros de la no abstención, respectivamente, explican la varianza de la variable *LSI* (logaritmo de los votos oficiales SI por centro de votación) con una *bondad de ajuste* $R^2 = 0.8237$.
- La significancia conjunta de los coeficientes de regresión parcial presenta un alto nivel de aceptación $f(3, 338) = 526.51$.
- Con excepción de *f*, el término constante, el valor esperado de los coeficientes de la regresión presentan intervalos de confianza que excluyen el cero como valor probable¹³, siendo consistente la dirección y el sentido de la relación entre la variable regresada *LSI* y sus regresores.

Se evidencia que las cualidades del modelo que relaciona los votos *SI* con las firmas superan las correspondientes a la relación entre los resultados oficiales a favor del *SI* y los exit_polls.

Esta diferencia verifica que las firmas, como medida de intención de voto, cometen menos errores de predicción de los votos *SI*, cuando observamos que su intervalo de confianza es de menor amplitud y, por lo tanto, de mayor precisión. En consecuencia, las cualidades significativamente diferentes de ambas medidas de intención contradicen las bondades relativas establecidas en el informe.

¹³ $t(a) = 0.35$, $t(b) = 38.26$, $t(c) = -7.62$ y $t(d) = 2.17$, donde *a*, *b*, *c* y *d* son los coeficientes definidos en el modelo estructural (2).

A su vez, la inconsistencia que observan los coeficientes de regresión parcial de la variable **elc_now** (+0,4604462 en la primera ecuación y -0,6612884 en la correspondiente a los exit_polls) también sugiere contradicciones en los resultados y juicios que se exponen en el informe. En la ecuación que relaciona el logaritmo de los votos SI con el logaritmo las firmas, el coeficiente de la variación porcentual de los nuevos electores es positivo. Entretanto, en la segunda ecuación, la relación es inversa.

Por último, la relación entre la variación porcentual inter-centros de la no-abstención y el logaritmo de los votos SI, en ambas ecuaciones, también presentan bondades apreciablemente diferentes. En el "Cuadro 4" la proporción es igual a 0.3311808, mientras que en el "Cuadro 5", la relación es igual a 0.4244489.

El análisis técnico presentado en esta sección incorpora elementos adicionales que hacen posible cuestionar los resultados estimados e inferencias derivadas de los mismos.

2.3.- A la interpretación

El uso de métodos cuantitativos permite, dentro de sus límites, reducir¹⁴ toda posibilidad o intento de tergiversar la información que reúnen los resultados, a partir de una interpretación equivocada o incompleta de los coeficientes y los valores esperados dentro del modelo. En otras palabras, los coeficientes son medidas de aproximación que logran expresarse por sí solas.

No obstante, en algunos casos, también será posible tropezarse con interpretaciones que no guardan vínculo alguno con la información concentrada dentro de los resultados.

En tal sentido, surge la necesidad de fortalecer la "libre expresión" de los coeficientes. La interpretación de estos últimos debe reducirse a la simple "traducción simultánea" del discurso envuelto dentro de cada ecuación.

Las firmas

La primera ecuación sujeta las ocurrencias observadas por la variable *LSI*, "logaritmo de los votos SI", en 342 centros de votación, a los cambios conjuntos experimentados dentro de la misma muestra, por las variables regresoras

¹⁴ Siempre que no sean violados los supuestos y fundamentos metodológicos.

LFIRMA, *elc_now* y *VEL*, variables que miden el “logaritmo de las firmas”, la “variación porcentual de nuevos electores” y la “no abstención”, respectivamente.

$$LSI = + 0.3059669 + 0.9942821*LFIRMA + 0.4604462*elc_now + 0.3311808*VEL + \mu$$

La ecuación revela que el logaritmo de los votos *SI* (*LSI*) de un centro de votación es aproximadamente igual a la sumatoria de 0.9942821 del “logaritmo de las firmas”¹⁵, 0.4604462 de la variación porcentual inter-centros de los “nuevos electores”, 0.3311808 de la variación porcentual inter-centros de la “no abstención”, un término independiente igual a 0.3059669 y una perturbación aleatoria μ .

Supongamos entonces que C_i define el *i*-ésimo centro de votación dentro de la muestra; y C_{i-1} , el centro de votación inmediato anterior.

Cœteris paribus, la relación entre el logaritmo de los votos *SI* y el logaritmo de las firmas señala que por cada 1% en que las firmas de C_i exceden las de C_{i-1} , el centro C_i tendrá 0.994% más votos *SI* que los observados en el centro de votación inmediato anterior, C_{i-1} . Esto es, por ejemplo, que si el 1% de diferencia viene representado por 100 firmas, entonces la opción *SI* en el *i*-ésimo centro de votación (C_i) deberá tener 99.4 votos más que en el centro de votación inmediato inferior.

Por su parte, el coeficiente que mide la proporción de reacción de los votos *SI* ante una diferencia del 1% de nuevos electores (*elc_now*) entre centros de votación contiguos, C_i y C_{i-1} , define que la relación inelástica es igual a 0.460. Luego, si el *i*-ésimo centro de votación presenta 100 nuevos electores más que su inmediato anterior C_{i-1} , la opción del *SI* atraería en C_i , 46 de estos 100 nuevos electores.

Por último, el grado de sensibilidad que muestran los votos *SI* ante las diferencias del 1% en participación electoral (*VEL*) de dos centros de votación contiguos, C_i y C_{i-1} , define también una relación inelástica e igual a 0.331. En otras palabras, si la diferencia del 1% en participación electoral de dos centros de votación contiguos es igual a 100 electores, treinta y tres (33) de estos votarían por la opción del *SI*.

¹⁵ Recordando siempre que el proceso *no-estacionario* que sigue la variable *LFIRMA* y el factor auto-regresivo de los votos *SI*, LSI_{i-1} omitido en la ecuación; determinan que 0.9942821 es un sobrestimado de la capacidad que tienen las firmas para predecir los votos *SI*, y, por lo tanto, los errores que comete la medida de intención es significativamente mayor que lo establecido en el informe. Por definición en el informe evaluado, una firma como medida imperfecta de la intención del elector genera 0.7 votos efectivos por el *SI*.

Los exit polls

La segunda ecuación sujeta las ocurrencias observadas por la variable LSI , "logaritmo de los votos SI", en 342 centros de votación, a los cambios conjuntos experimentados dentro de la misma muestra por las variables regresoras lex_si , elc_now y VEL , variables que miden el "logaritmo de los exit polls", la "variación porcentual de nuevos electores" y la "no abstención", respectivamente.

$$LSI = + 0.0722736 + 0.9701892 * lex_si - 0.6612884 * elc_now + 0.4244489 * VEL + \mu$$

La ecuación revela que el logaritmo de los votos SI (LSI) de un centro de votación es aproximadamente igual a la sumatoria de 0.9701892 del "logaritmo de los exit_polls"¹⁶, -0.6612884 de la variación porcentual inter-centros de los "nuevos electores" (elc_now), 0.4244489 de la variación porcentual inter-centros de la "no abstención" (VEL), un término independiente igual a 0.0722736 y una perturbación aleatoria μ .

En este caso, la interpretación técnica de los coeficientes de regresión parcial seguirá el mismo orden utilizado para la primera ecuación.

Destaca entonces la inconsistencia observada por el coeficiente de impacto que tienen sobre el total de votos SI , las diferencias en la cantidad de nuevos electores, entre centros de votación contiguos. En la primera ecuación, el valor de este coeficiente muestra una relación directa con una magnitud igual a 0.4604462; mientras que en esta última ecuación, la relación es inversa con una magnitud igual a -0.6612884.

La relación entre el logaritmo de las firmas y el de los exit_polls, señala que por cada 1% de diferencia en las intenciones de voto a favor de SI de dos centros de votación contiguos, la diferencia en votos SI será igual a 0.97 votos. Y que si este 1% estuviese representado, por ejemplo, por 100 intenciones a favor de la opción SI , el i -ésimo centro de votación (C_i) deberá presentar 97 votos más por el SI que en el centro C_{i-1} .

¹⁶ Recordando siempre que el proceso *no-estacionario* que sigue la variable LFIRMA y el factor auto-regresivo de los votos SI, LSI_{i-1} , omitido en la ecuación; determinan que 0.9942821 es un sobrestimado de la capacidad que tienen las firmas para predecir los votos SI, y, por lo tanto, los errores que comete la medida de intención es significativamente mayor que lo establecido en el informe. Por definición en el informe evaluado, una firma como medida imperfecta de la intención del elector genera 0.7 votos efectivos por el SI.

Por su parte, la relación entre el logaritmo de los votos *SI* y la diferencia porcentual de nuevos electores que presenten dos centros de votación contiguos, C_i y C_{i-1} , establece que, en promedio, por cada 100 nuevos electores de diferencia la opción del *SI* perdería 66 votos.

Por último, la sensibilidad experimentada por los votos *SI* ante las diferencias del 1% en participación electoral de dos centros de votación contiguos, C_i y C_{i-1} , define una relación inelástica e igual a 0.424. Esto es, que para una diferencia de 100 electores en la participación de dos centros de votación contiguos, la opción del *SI* habría captado 42 votos.

La regresión en dos etapas(2SLS)

Las regresiones presentadas¹⁷ en el “Cuadro 7” y el “Cuadro 8”, prácticamente observan las mismas anomalías y características de las estimaciones expuestas en los Cuadros 4 y 5 (Para efectos de este artículo, tablas 6 y 7) (Cuadros impresos en el libro objeto de análisis que no son incluidos en este artículo).

A su vez, la sobrestimación de los votos *SI*, hecha tanto por las firmas como los *exit_polls*, aumenta de tal forma que los coeficientes de ambas variables regresoras superan la unidad. Esta anomalía trae como consecuencia inmediata la varianza explosiva de los residuos en ambas ecuaciones y una evidente correlación entre sus tendencias inducida, artificialmente, por la varianza *no estacionaria*.

3.- OBSERVACIONES FINALES

- Las observaciones presentadas en este artículo son estrictamente técnicas y carecen de toda valoración o juicio, sobre el objeto subyacente (el proceso electoral). Cualquier juicio será tarea exclusiva del lector, si éste, así lo cree necesario.
- A su vez, las observaciones se desvinculan de todo intento, hecho por terceros, para negar las presunciones que se derivan en el informe “*En busca del cisne negro*”. Más aún, cuando estas se apoyan en la descalificación de los autores, alejándose de todo carácter cívico y académico.
- Si bien es cierto que las observaciones apuntan hacia a la metodología aplicada y resultados del informe, al mismo tiempo, reconocen la habilidad y maestría metodológica necesaria para simular el comportamiento y la forma

¹⁷ Página 19 del informe “*En busca del cisne negro*”.

funcional de un *proceso auto-regresivo, no estacionario*, capaz de degenerar las perturbaciones de ambas ecuaciones, e inducir una correlación artificial entre las dos variables residuales a partir de la “explosión” de las varianzas no estacionarias y la sobrestimación que origina la intercepción no vacía entre las variables o medidas donde subyace la dinámica de los procesos de: *recolección de firmas, intención del voto SI, incorporación de nuevos electores y la participación electoral*; para todo centro de votación que pertenece a la muestra seleccionada.

- El método viene a ser entonces, el instrumento que hace comprensible los mecanismos de “lo concreto” en relación a sus “múltiples determinaciones”, en general, una parte a la vez. Luego, al exceder los límites de la metodología, vendrá a ser el objeto quien, probablemente, termine condenando al sujeto.
- Cuando los hechos son modelados de forma tal que la hipótesis se la hace aparecer como tesis, el método pasa a ser un instrumento perverso, amoral, de dominación ideológica.
- Es probable que la ausencia de una respuesta científica, por parte del sujeto indiciado en el informe “*En busca del cisne negro*”, a las interrogantes y colaterales electorales originados de los resultados electrónicos de agosto de 2004, haya contribuido para que las dudas continúen vigentes en un porcentaje, no despreciable, de la población venezolana.

4.- REFERENCIAS BIBLIOGRÁFICAS

- Álvarez V, Nelson (2002), *Econometría II. Análisis de modelos econométricos de series temporales*, 1ra. Edición, Thomson, Madrid.
- Chou, Ya-Lun (1991), *Análisis Estadístico*, 2da. Edición, Mc Graw Hill, México, D.F.
- Gujarati, Damodar N. (1997), *Econometría*, 3ra. Edición, Mc Graw Hill, Bogotá, Colombia.
- Hausmann, Ricardo y Rigobon, Roberto (2004), *En busca del cisne negro*, Edición digital, <http://web.sumate.org/archivos.asp>
- Intriligator, Michael D. (1990), *Modelos econométricos, técnicas y aplicaciones*, Fondo de Cultura Económica, México, D. F.
- Maddala, G. S. (1996), *Introducción a la Econometría*, 2da. Edición, Prentice-Hall Hispanoamericana.
- Novalés, Alfonso (1993), *Econometría*, 2da. Edición, Mc Graw Hill, Madrid.
- Pérez L, César (2006), *Econometría de las Series Temporales*, 1ra. Edición, Pearson Educación, Madrid.
- Sánchez G, Carlos (1999), *Métodos Econométricos*, 1ra. Edición, Ariel Economía, Barcelona.