Ensayos–Volumen XXVIII, núm.1, mayo 2009, pp. 1-20

Regresión espuria en especificaciones dinámicas

Manuel Gómez Zaldivar[1]
Oscar Manjarrez Castro[2]
Daniel Ventosa-Santaulària[3]

Fecha de recepción: 17 XII 2008 Fecha de aceptación: 17 III 2009

Resumen

La regresión espuria ha sido documentada en econometría desde el trabajo de Granger y Newbold (1974). Dicho fenómeno ha sido identificado usando una vasta diversidad de Procesos Generadores de Datos que van desde una simple raíz unitaria sin deriva (*unit root without drift*), hasta una serie estacionaria en tendencia con rompimientos estructurales (*broken trend stationary*). No obstante, la especificación bajo la cual se han realizado estos trabajos es la regresión simple con una sola variable explicativa. En este documento se demuestra que, usando series estacionarias en tendencia independientes entre sí, la regresión espuria también ocurre cuando se estima una especificación dinámica. Dicha especificación dinámica es empleada frecuentemente en el estudio de las expectativas. Los resultados amplían y complementan dos cuestiones que con frecuencia son sugeridas en los manuales de econometría: cuando el proceso es estacionario en tendencia con quiebres estructurales, (i) en especificaciones dinámicas, el estadístico Durbin-Watson no tiende a cero, por lo que no es una prueba fiable de regresión espuria, y (ii) la inclusión de la variable dependiente rezagada en el conjunto de explicativas no siempre corrige el problema de la regresión espuria.

***Palabras clave*:** Regresión Espuria, procesos estacionarios en tendencia, especificación dinámica.

---

[1]Departamento de Economía y Finanzas, Universidad de Guanajuato.
[2]Departamento de Economía y Finanzas, Universidad de Guanajuato.
[3]Autor responsable de la correspondencia. Departamento de Economía y Finanzas, Universidad de Guanajuato. DCEA-Campus Marfil, Fracc. I, El Establo, Guanajuato, Gto. CP 36250, México. Teléfono y fax: 52 (473) 735-2925; correo electrónico: daniel@ventosa-santaularia.com.

Abstract

The spurious regression phenomenon, identified by Granger and Newbold (1974) is well known in econometrics. In fact, spurious regression occurs under a wide variety of Data Generating Processes: driftless unit root, unit root with drift, trend stationarity, broken-trend stationarity,… However, the phenomenon has been solely studied under the assumption that the specification to be estimated is a simple linear regression with a single regressand. We prove in this article that the spurious regression phenomenon also occurs when a dynamic specification is estimated. Dynamic specifications are commonly employed to model expectations.

Our results extend the common knowledge concerning spurious regression usually found in popular textbooks: when the variables are trend stationary (i) using them in dynamic specification does not preclude the Durbin-Watson statistic to collapse so the latter is not a reliable tool in the identification of the spurious regression, and (ii) including the lagged value of the dependent variable as a regressand does not always solve the problem of spurious regression.

***Keywords*:** Spurious Regression, Trend Stationarity, Dynamic Specification.
***JEL classification*:** C12, C13, C22.

Introducción

El estudio de la regresión espuria -en economía- tiene su origen en un estudio de Monte Carlo realizado por Granger y Newbold (1974), mismo que la pone en evidencia mediante la simulación de procesos no estacionarios e independientes entre sí. Con dichas variables simuladas, Granger y Newbold estimaron la especificación $y_t = \alpha_0 + \beta_0 x_t + u_t$ por el método de Mínimos Cuadrados Ordinarios (MCO). Los resultados de sus simulaciones arrojaron lo que hoy se conoce como regresión sin sentido: se detectó muy frecuentemente una relación estadística inexistente entre variables independientes. El fenómeno de regresión espuria puede ser definido formalmente, como en seguida se expone. Considere dos procesos generados como caminatas aleatorias independientes entre sí:

$$x_t = x_{t-1} + u_t, \qquad u_t \sim iid(0, \sigma_u^2) \tag{1}$$

$$y_t = y_{t-1} + v_t, \qquad v_t \sim iid(0, \sigma_v^2) \tag{2}$$

Estime por MCO la siguiente regresión:

$$y_t = \alpha_0 + \beta_0 x_t + \varepsilon_t \tag{3}$$

Sabiendo que las variables x_t y y_t son independientes, sería de esperar que el estimador del parámetro β_0 así como la bondad del ajuste R^2 del ejercicio de estimación, fueran cercanos a cero. Sin embargo, éste no es siempre el caso. Esta relación estadística sin sentido, resultante de estimar una regresión usando series no estacionarias, es conocida como regresión espuria. La explicación teórica de este fenómeno fue expuesta por Phillips (1986), quien, definiendo las variables x_t y y_t como en (1) y (2) respectivamente, demostró que el estimador por MCO de β_0 no tiende a cero como debería; mientras que su estadístico t asociado, diverge a tasa $\sqrt{T}$ (con lo que la hipótesis nula de no significancia acaba siendo rechazada, conforme aumente el tamaño de muestra). Así mismo, Phillips demostró que, en esas condiciones, la R^2 converge en probabilidad a uno. Cabe mencionar que los Procesos Generadores de Datos usados por Phillips (1986) son en extremo sencillos: raíz unitaria sin deriva.

En estudios posteriores, diversos autores han demostrado que el fenómeno también ocurre usando procesos más complejos que los empleados por Phillips: caminata aleatoria con deriva, procesos integrados fraccionalmente, procesos integrados de órdenes superiores, etc.

No obstante, la literatura no incluye en su haber especificaciones dinámicas usando procesos con tendencia determinista. En este trabajo, se emplea justamente esa especificación y se demuestra que el fenómeno de regresión espuria también se gesta. Al estimar por MCO la relación $y_t = \alpha + \beta x_t + \delta y_{t-1} + \varepsilon_t$, siendo ambas variables estacionarias en tendencia e independientes entre sí, las herramientas de inferencia estadística -y en particular, el estadístico t asociado a $\hat{\beta}$- indican erróneamente la existencia de una relación estadística entre x_t y y_t. La importancia de este resultado radica en que complementa y extiende lo señalado en algunos libros de econometría, que frecuentemente son utilizados (Maddala, 1992; Griffiths, Hill y Judge, 1993; Hamilton, 1994; Johnston y Dinardo, 1998; Wooldridge, 2000; Davidson y Mackinnon, 2004).

Este trabajo se estructura de la siguiente manera: en la primera sección (1), se ofrece una breve revisión de la literatura concerniente a la regresión espuria; en la segunda sección (2), se presenta la importancia de las especificaciones dinámicas en economía aplicada junto con los resultados asintóticos; y en la tercera sección (3), se encuentran los experimentos de muestra finita. Por último, se incluyen los comentarios finales y tres apéndices, en los que: (i) se sitúan los resultados del artículo sobre la literatura concerniente a regresión espuria; (ii) se explica la metodología para realizar los cálculos asintóticos y (iii) se presenta el código informático que efectúa tales cálculos. Cabe mencionar que los últimos dos apéndices corresponden a documentos anexos que están disponibles en línea.

1. Breve revisión de la literatura

Al estimar una regresión lineal simple por el método de MCO utilizando dos variables -la explicada y la explicativa-, ambas independientes y con un fuerte componente tendencial, existe la posibilidad de que al estudiar la significancia estadística del parámetro asociado al regresor, se infiera una relación -inexistente- entre tales variables. A lo anterior se le denomina "Error tipo I" y suele minimizarse estableciendo un nivel de, por ejemplo, 5%. Cuando la frecuencia del error tipo I es mucho mayor a la que uno fija con el nivel, se dice que ese estadístico arroja resultados espurios. La relación sin sentido resultante es conocida en la literatura económica y econométrica como regresión espuria. Como ya se mencionó anteriormente, las consecuencias de las regresiones espurias entre caminatas aleatorias independientes fueron señaladas, inicialmente, por Granger y Newbold (1974), usando simulaciones de Monte Carlo. Los autores enfatizaron lo inapropiado que podría llegar a ser el uso de las pruebas estadísticas tradicionales, en presencia de series no estacionarias.

1.1 Regresión espuria en procesos con tendencia estocástica

Phillips (1986) propuso la utilización de una teoría asintótica no estándar que explica en el ámbito teórico dicho fenómeno. Los procesos generadores de datos (PGDs) empleados por este último autor son las sencillas caminatas aleatorias sin deriva. Phillips demostró que las distribuciones tanto de los parámetros como de los estadísticos, difieren de las distribuciones estándar, es decir, las que se obtienen bajo los supuestos clásicos del modelo de regresión; encuentra, además, que la distribución del intercepto diverge y que el coeficiente de la pendiente así como el coeficiente de determinación, R^2, no convergen en su verdadero valor, que es cero. De igual forma, las expresiones de los estadísticos t asociados a los parámetros estimados divergen a tasa $\sqrt{T}$, siendo T el tamaño de la muestra. Lo anterior implica, asintóticamente, un rechazo sistemático de la hipótesis nula, $H_0 : \beta = 0$, con base en los valores críticos convencionales. Por último, Phillips encuentra que la prueba Durbin-Watson (DW) de autocorrelación de primer orden converge a cero. Marmol (1995) estudia el fenómeno usando series que son generadas por procesos de integrados de orden d, donde $d \in \{1,2,...\}$. Marmol encuentra que el estimador del intercepto, diverge; la tasa de divergencia depende del orden de integración de las variables. Por otra parte, tanto el estimador de β como la R^2 tienen una distribución límite no-degenerada, independientemente del orden de los procesos. Al igual que Phillips (1986), Marmol demuestra que el estadístico t asociado a la pendiente diverge a tasa $\sqrt{T}$. Además, el estadístico DW converge a cero independientemente del orden de integración.

Otra extensión importante la llevan a cabo Cappuccio y Lubian (1997), quienes estudian las propiedades asintóticas de la estimación de la regresión (3), cuando ambas variables son procesos de memoria larga. Al estimar por MCO el coeficiente $\hat{\beta}$, los autores demuestran que éste tiene una distribución límite no-degenerada. Con respecto al estadístico t, Cappuccio y Lubian obtienen la misma tasa de divergencia que Phillips. La R^2 tiene también una distribución límite no-degenerada y el estadístico DW, como en los casos previos, converge en probabilidad a cero. Marmol (1996b) demuestra que la regresión espuria se da también cuando las variables empleadas tienen distintos órdenes de integración, no importando cual sea éste en cada una. Los estadísticos t, tanto $t_{\hat{\alpha}}$ como $t_{\hat{\beta}}$, divergen a tasa $T^{1/2}$. Con respecto a la DW, ésta converge a cero a una tasa que, esa sí, depende del orden de integración de las variables. En lo que respecta a la R^2, se obtiene una distribución límite no-degenerada, no importando el orden de integración de las series. Entorf (1997) realiza un estudio similar al de Phillips (1986); sólo que los PGDs que él usa son raíces unitarias con deriva. Entorf encuentra que la estimación por MCO de β converge en probabilidad, a la razón de las derivas de. y_t y de x_t. Además, el estadístico $t_{\hat{\beta}}$ diverge a tasa T. En el caso del coeficiente de determinación, R^2, éste converge a 1.

1.2. Regresión espuria en procesos con tendencia determinista

La regresión espuria también ocurre cuando el componente de tendencia de las series es de naturaleza determinista (ver ecuación 5 en la próxima sección). Uno de los primeros resultados en este sentido fue provisto por Hasseler (2000), en cuyo trabajo se demuestra que el fenómeno de regresión espuria está presente en series estacionarias con respecto a una tendencia lineal. Este mismo resultado fue obtenido, de manera independiente, por Kim, Lee y Newbold (2004). En forma análoga a Entorf (1997), Kim, Lee y Newbold encuentran que el estimador del coeficiente de la pendiente converge en probabilidad a la razón de las pendientes de y_t y de x_t; en Entorf, era la razón de derivas. De igual manera, el estadístico $t_{\hat{\beta}}$ crece a tasa $T^{3/2}$. Extendiendo los resultados de Kim, Lee y Newbold (2004), Noriega y Ventosa-Santaulària (2006) estudian el fenómeno de la regresión espuria donde las tendencias deterministas de los PGDs están sujetas a cambios estructurales. Noriega y Ventosa-Santaulària demuestran que, cuando existen múltiples rompimientos estructurales en las series, la regresión espuria se presenta, como en los demás casos mencionados, asintóticamente. Sin embargo, en ese trabajo la tasa de divergencia del estadístico t asociado a $\hat{\beta}$ vuelve a ser $\sqrt{T}$, como en la mayoría de los casos en los que el componente de tendencia es estocástico. Recientemente Noriega y Ventosa-Santaulària (2007) realizan un estudio del fenómeno de regresión espuria, donde las variables usadas en la especificación (3) resultan de combinaciones de los PGDs que antes se mencionan. Destacan en particular las combinaciones de procesos deterministas y estocásticos en los que, como anteriormente sucede, también se da la regresión espuria.

2. Regresión dinámica: resultados asintóticos

Para cualquier disciplina que aborde el estudio de la naturaleza, resulta particularmente relevante conocer las consecuencias de la realización de un evento (variaciones en el precio de petróleo, aumento de la oferta monetaria, crecimiento del ingreso, recaudación fiscal,...), ya que es necesario, por ejemplo, reducir o amplificar los efectos negativos (positivos), a través de políticas económicas. Dicho conocimiento es también necesario si la finalidad del estudio es la elaboración de pronósticos. No obstante, raras son las ocasiones en que los efectos de dicho fenómeno son exclusivamente instantáneos; por lo general, repercuten en periodos posteriores. En otras palabras, es importante estudiar la dinámica de los procesos; los modelos econométricos dinámicos constituyen, de hecho, una herramienta práctica para hacerlo. En ese sentido, conviene saber si las especificaciones dinámicas tradicionalmente usadas en econometría son susceptibles de arrojar inferencia espuria, y bajo qué circunstancias lo harían. En este trabajo, se aborda la cuestión desde el ángulo de la estacionariedad. ¿Qué ocurre con la estimación de especificaciones dinámicas, cuando las series empleadas contienen un componente de tendencia determinista? En esta sección, se presentan las propiedades asintóticas de los estimadores en una especificación dinámica. En una primera instancia, definimos los Procesos Generadores de Datos con los que se trabaja. Como ya se señaló, la dependencia de una variable, y_t, (dependiente) respecto a otra u otras variables x_t (explicativas) raramente es instantánea. Es común constatar que los cambios en esta última tienen efecto no sólo en la realización contemporánea de la variable dependiente, sino también en las realizaciones futuras. Semejante dinámica puede ser modelada, por ejemplo, mediante especificaciones que reciben el nombre de Modelos de Rezagos Distribuidos (MRD). El ejemplo más simple de un MRD, un MRD (1,0), es:

$$y_t = \alpha + \beta x_t + \delta y_{t-1} + \varepsilon_t \tag{4}$$

El rezago de la variable dependiente en tanto que explicativa en la ecuación (4), implica una serie de repercusiones dinámicas en y_t, que se dirimen en el tiempo ante un cambio en x_t. Hay un efecto inmediato, seguido además de respuestas en el corto, mediano y largo plazo. En esta sección se presenta el resultado teórico principal del artículo. La especificación estudiada consiste en un modelo dinámico simple, un MRD(1,0), como el estipulado en (4). El PGD de ambas variables, $\{w_t\}_{t=1}^{\infty}$ para $w = x, y$, es un proceso estacionario en tendencia, que denotaremos TS (por sus siglas en inglés, *Trend Stationary*):

$$PGDA\begin{cases} y_t = \mu_y + \beta_y t + \varepsilon_{yt} \\ x_t = \mu_x + \beta_x t + \varepsilon_{xt} \end{cases} \tag{5}$$

donde ε_{yt} y ε_{xt} son, o bien (i) variables independiente e idénticamente distribuidas (iid) o, si no, (ii) procesos que satisfacen las condiciones generales establecidas en la proposición 17.3 de Hamilton (1994, p. 505): $u_t = \psi_w(L)\varepsilon_{wt} = \sum_{j=0}^{\infty} \psi_{w,j}\varepsilon_{wt-1,}$

para w=x,y

donde $\sum_{j=0}^{\infty} j\left|\psi_{w,j}\right| < \infty$ y la secuencia $\{\varepsilon_{wt}\}$ con media cero, varianza σ_w^2 y cuarto momento finito [lo cual autoriza que haya autocorrelación y heteroscedasticidad]. Se estudia en primera instancia el caso en el que las innovaciones, ε_{yt} y ε_{xt} son iid. Si al investigar una posible relación dinámica entre y_t y x_t se estima la regresión (4), entonces, dicha estimación posee las siguientes propiedades asintóticas:

Teorema 1 Sean y_t y x_t dos variables independientes generadas de acuerdo al PGD A, donde ε_{yt} y ε_{xt} son independientes entre sí y también son variables aleatorias (v. a.), independiente e idénticamente distribuidas con esperanza cero y varianza constante, σ_y^2 y σ_x^2, respectivamente. Estime por MCO la regresión (4) y denote $\hat{\alpha}, \hat{\beta}$ y $\hat{\delta}$ los estimadores de α, β y δ, respectivamente. Entonces, cuando $T \to \infty$:

$$\hat{\alpha} \xrightarrow{p} \frac{\sigma_x^2 + \sigma_y^2 \beta_x (\mu_y \beta_x - \mu_x \beta_y)}{\sigma_x^2 \beta_y^2 + \sigma_y^2 \beta_x^2}$$

$$\hat{\beta} \xrightarrow{p} \frac{\sigma_y^2 \beta_x \beta_y}{\sigma_x^2 \beta_y^2 + \sigma_y^2 \beta_x^2}$$

$$\hat{\delta} \xrightarrow{p} \frac{\sigma_x^2 \beta_y^2}{\sigma_x^2 \beta_y^2 + \sigma_y^2 \beta_x^2} < 1$$

$$T^{-1/2} t_{\hat{\alpha}} = O_p(1)$$

$$T^{-1/2} t_{\hat{\beta}} = O_p(1)$$

$$T^{-1/2} t_{\hat{\delta}} = O_p(1)$$

$$R^2 = 1 - O_p(T^{-2})$$

$$F = O_p(T^3)$$

$$DW \xrightarrow{p} 2\left[\frac{\sigma_y^2\beta_x^2 + 3\sigma_x^2\beta_y^2}{\sigma_y^2\beta_x^2 + 2\sigma_x^2\beta_y^2}\right]$$

donde $\xrightarrow{p}$ denota convergencia en probabilidad y $O_p(\cdot)$ indica el orden de convergencia.

Prueba: ver apéndices B y C disponibles en: http://www.economia.uanl.mx/

Es fundamental destacar que los tres coeficientes estimados, $\hat{\alpha}, \hat{\beta}$ y $\hat{\delta}$, convergen en probabilidad, sin necesidad de escalar por una potencia de T, a una expresión distinta de cero. Note que el estimador de δ es asintóticamente positivo y menor a la unidad. Más importante aún es el comportamiento de los estadísticos t asociados a dichos parámetros. Éstos divergen a tasa $\sqrt{T}$. Debido a lo anterior, conforme el tamaño de la muestra aumente, habrá un rechazo sistemático de la hipótesis nula. En otras palabras, se obtendrá evidencia de que los parámetros no son nulos. Resalta en particular el estimador $\hat{\beta}$, pues al concluir que es distinto de cero, se considerará que sí existe una relación estadística entre las variables x_t y y_t, aun cuando ésta es inexistente. Además, la medida de bondad del ajuste, la R^2, converge en probabilidad con la unidad, lo que indica un "ajuste perfecto"; mientras que la prueba de significancia conjunta, la prueba de F, diverge en extremo rápidamente, a tasa T^3. Ello indica que, asintóticamente, siempre se rechaza la hipótesis nula de que todos los parámetros son iguales a cero. Todo lo anterior inducirá a una conclusión errónea, al existir, (i) un aparente buen ajuste del modelo, y (ii) una fuerte significancia estadística de todos los parámetros. En lo que concierne al estadístico DW, éste también converge en probabilidad a un valor no nulo. Lo anterior contrasta con la creencia de que una regresión espuria arroja un DW cercano a cero. Ahora, presentamos los resultados asintóticos de la estimación de esta misma especificación [ecuación (4)], suponiendo que las innovaciones de los PGDs no son iid sino que más bien son procesos -estacionarios- con una cierta estructura de autocorrelación:

Teorema 2 Sean x_t y y_t dos variables independientes generadas de acuerdo al PGD A, donde ε_{yt} y ε_{xt} son independientes entre sí y obedecen a las condiciones establecidas por Hamilton (1994, p.505). Estime por MCO la regresión (4) y denote $\hat{\alpha}, \hat{\beta}$ y $\hat{\delta}$ los estimadores de α, β y σ, respectivamente. Entonces, cuando $T \to \infty$:

$$\hat{\alpha} \xrightarrow{p} \frac{\rho_{y1}\beta_x(\beta_x\beta_y + \mu_x\beta_y - \mu_y\beta_x) + \sigma_x^2\beta_y^3 + \sigma_y^2\beta_x(\mu_y\beta_x - \mu_x\beta_y)}{\sigma_x^2\beta_y^2 + \sigma_y^2\beta_x^2}$$

$$\hat{\beta} \xrightarrow{p} \frac{(\sigma_y^2\rho_{y1})\beta_x\beta_y}{\sigma_x^2\beta_y^2 + \sigma_y^2\beta_x^2}$$

$$\hat{\delta} \xrightarrow{p} \frac{\rho_{y1}\beta_x^2 + \sigma_x^2\beta_y^2}{\sigma_x^2\beta_y^2 + \sigma_y^2\beta_x^2} < 1$$

$$T^{-1/2}t_{\hat{\alpha}} = O_p(1)$$

$$T^{-1/2}t_{\hat{\beta}} = O_p(1)$$

$$T^{-1/2}t_{\hat{\delta}} = O_p(1)$$

$$R^2 = 1 - O_p(T^{-2})$$

$$F = O_p(T^3)$$

$$DW = O_p(1)$$

donde ρ_{y1} denota la primera autocovarianza de la v.a. ε_{yt} .

Prueba: ver apéndices B y C disponibles en: www.economia.uanl.mx/

Los resultados asintóticos en presencia de autocorrelación en las innovaciones de los PGDs no son muy diferentes a los anteriores. Las expresiones son más extensas, pero los órdenes de divergencia se mantienen intactos. Ello implica que el fenómeno de la regresión espuria sigue manifestándose en este caso: estimadores estadísticamente significativos individualmente y en conjunto y bondad del ajuste alta. En lo que concierne al estadístico DW, en este caso tampoco se colapsa; por lo que contrariamente a la mayor parte de los resultados de regresión espuria, no es posible identificarla mediante la observación de un estadístico DW cercano a cero.

Los resultados obtenidos en este estudio complementan a los ya existentes en especificaciones estáticas. Dado que los procedimientos para detectar y corregir regresión espuria en especificaciones dinámicas difieren de su contraparte estática, incluimos una breve discusión en el apéndice A, donde se explica más detalladamente los contrastes.

3. Resultados experimentales

Los resultados hasta ahora obtenidos son de naturaleza asintótica, por lo que sólo son válidos, en principio, en muestras de gran tamaño. Existe la posibilidad de que estos últimos provean una aproximación muy pobre a lo que se observa en muestras de tamaño menor, como las que usualmente están disponibles para información macroeconómica. Con objeto de estudiar las propiedades de los estimadores en muestras finitas, conviene realizar experimentos de Monte Carlo. En esta sección, se presentan los resultados de estimaciones hechas con datos simulados así como con las fórmulas asintóticas obtenidas anteriormente, con objeto de evaluar qué tan rápido son estas últimas una representación adecuada de los estimadores de MCO, cuando se usan muestras relativamente pequeñas. Para generar los datos, se emplea el PGD A. Se retiene -en cada replicación de la simulación- el valor de los estadísticos t asociados a los parámetros que acompañan a las variables x_t y y_{t-1} en la regresión (4); con base en éstos, se decide si hay evidencia suficiente para rechazar la hipótesis nula[4], $H_0 : \hat{\beta}$ ó $\hat{\delta} = 0$. Lo anterior se repite 10,000 veces y la proporción o tasa de rechazo es la que se reporta en el cuadro (1). Dichas tasas de rechazo se calculan para muestras de tamaño T=50, 100, 200, 500, 1,000 y 10,000.

Los parámetros usados en los PGDs de x_t y y_t son los siguientes: $\mu_y = 0.5$, $\mu_x = 0.8$, $\beta_x = 0.35$ y $\beta_y = 0.8$. La magnitud de éstos corresponde aproximadamente a los obtenidos en el estudio empírico de Noriega (2004). Los resultados simulados muestran que, en un 99% de los casos, cuando $T \leq 50$, el estadístico $t_{\hat{\beta}}$ señala que existe una falsa relación entre las series. Estos resultados tienen aún mayor contundencia con muestras mayores a 100 (donde el rechazo es sistemático).

[4] Nivel de confianza 5%; valor crítico $|1.96|$.

Cuadro 1

Tasa de rechazo usando los estadísticos $t_{\hat{\beta}}$ y $t_{\hat{\delta}}$: variables generadas con base en el PGD A. Las innovaciones son iid

T	$t_{\hat{\beta}}$		$t_{\hat{\delta}}$	
	Teórico	**Simulado**	**Teórico**	**Simulado**
50	1.00	0.99	1.00	0.83
100	1.00	0.99	1.00	0.98
200	1.00	1.00	1.00	1.00
500	1.00	1.00	1.00	1.00
1,000	1.00	1.00	1.00	1.00
10,000	1.00	1.00	1.00	1.00

La importancia de este cuadro radica en la similitud existente entre resultados asintóticos y de muestra finita. En efecto, todo parece indicar que los primeros constituyen una buena aproximación a los segundos, aun en muestras tan pequeñas como las de 50 observaciones. La segunda simulación que se presenta en esta sección, detalla la observación del fenómeno de la regresión espuria bajo diversos escenarios, concernientes a la naturaleza de las innovaciones en el **PGD A**. Se incorporan a éste:

(i) Innovaciones $e_z \sim iidN(0,1)$

(ii) procesos autorregresivos

$\left[AR(1): u_{zt} = \varphi_z u_{zt-1} + e_{zt}\right.$, donde $z = x, y; \varphi_y = 0.4; \varphi_x = 0.7y; e_{zt} \sim iidN(0,1)$];

(iii) procesos de promedios móviles

$\left[MA(1): u\_zt = \theta_z \theta_{zt-1} + e_{zt}\right.$, donde $z = x, y; \theta_y = 0.4; \theta_x = 0.7y; e_{zt} \sim iidN(0,1)$]

(iv) Ruidos blancos heteroscedásticos

$\lfloor e_{yt}$ ~iid *N(0,1)* para *t=1,2,…,T/2 ;* e_{yt} ~ iid *N(0,16)* para *t=(T/2)+1, (T/2)+2,…,T°,* $e_{xt} \sim iidN(0{,}9)$ para *t=1,2,…T/2;* $e_{xt} \sim iidN(0{,}16)$ *para* $t = (T/2) + 1, (t72) + 2, \ldots T° \rfloor$

El cuadro (2) arroja resultados interesantes. Por una parte, confirma el experimento anterior de Monte Carlo, en tanto indica que los resultados asintóticos son buenas aproximaciones a los estimadores por MCO, en muestras finitas; la tasa de rechazo de H_0 usando el estadístico t asociado a $\hat{\beta}$ es, para todo fin práctico, igual al 100%; la medida de bondad del ajuste, la R^2, es en todos los casos superior a 0.9 para muestras con 100 observaciones o más; el estadístico de prueba F se dispara desde *T=50,* con lo que la significancia conjunta de los parámetros queda "respaldada" por éste, y el estadístico *DW* es $O_p(1)$, lo que deja patente que no en toda regresión espuria, dicho estadístico tiende a cero. Todos los apuntes anteriores se mantienen bajo distintos tipo de innovaciones en ambas variables; éstas pueden ser ruido blanco homoscedástico o heteroscedástico, procesos autorregresivos o de promedios móviles. Destaca el hecho de que, bajo heteroscedasticidad, los resultados de muestras finitas parecen requerir de "más observaciones" para equipararse a los asintóticos.

Cuadro 2

Inferencia en regresiones dinámicas espurias: (i) Tasa de rechazo del estadístico $t_{\hat{\beta}}$; (ii) Bondad de Ajuste R^2; (iii) Estadístico F; (iv) Estadístico DW: Variables generadas con base en el PGD A. Las innovaciones son iid, procesos autoregresivos, procesos de promedios móviles y procesos independiente y no-idénticamente distribuidos (heteroscedásticos).

T	Estadístico	Innovaciones			
		Ruido Blanco	AR(1)	MA(1)	Heterosc.
50	Tasa de rechazo	1.00	0.99	0.99	0.99
	R^2	0.95	0.96	0.96	0.73
	F	545.25	586.09	565.97	68.76
	DW	2.23	2.06	1.99	2.13
100	Tasa de rechazo	1.00	1.00	1.00	1.00
	R^2	0.98	0.98	099	0.91
	F	$4.4\ 10^3$	$4.6\ 10^3$	$4.5\ 10^3$	$5.4\ 10^2$
	DW	2.26	2.10	2.01	2.15
250	Tasa de rechazo	1.00	1.00	1.00	1.00
	R^2	0.99	0.99	099	0.98
	F	$6.9\ 10^4$	$7.2\ 10^4$	$7.1\ 10^4$	$8.6\ 10^3$
	DW	2.27	2.13	2.03	2.17
500	Tasa de rechazo	1.00	1.00	1.00	1.00
	R^2	0.99	0.99	099	0.99
	F	$5.5\ 10^5$	$5.8\ 10^5$	$5.7\ 10^5$	$6.8\ 10^4$
	DW	2.27	2.13	2.03	2.17

Conclusiones

En este trabajo, se ha presentado la teoría asintótica de los estimadores de Mínimos Cuadrados Ordinarios usando una especificación dinámica simple y asumiendo que las variables empleadas son generadas como Procesos Estacionarios en Tendencia independientes entre sí. Con base en lo anterior fue posible extender los resultados de Kim, Lee y Newbold (2004) y Noriega y Ventosa-Santaulària (2006) en una dirección original, puesto que el fenómeno de regresión espuria suele ser estudiado únicamente en el marco de una especificación sencilla con ordenada en el origen y una única variable explicativa. El resultado obtenido muestra que la regresión espuria también se da en especificaciones dinámicas, cuando la naturaleza de la tendencia en las variables es determinista. Se demostró que este fenómeno se da, tanto asintóticamente como en muestras finitas en extremo pequeñas, es decir, menores a 100 observaciones. El estadístico t asociado al estimador de β diverge a tasa $\sqrt{T}$, mientras que la R^2 converge a 1 y el estadístico de significancia conjunto, F , diverge a tasa T^3. Cabe resaltar que el coeficiente de determinación, R^2, así como la prueba F , indican -erróneamente- que el modelo está correctamente especificado, por lo que existe un gran riesgo en el supuesto de llevar a cabo inferencia sin sentido, al estimar una especificación dinámica. Por otra parte, se encontró que el estadístico *DW*, contrario a lo que usualmente se obtiene en esta literatura, no se colapsa; por lo cual, la regla heurística para decidir si hay regresión espuria- $R^2 > DW$ - definitivamente no aplica en este caso. Finalmente, se presentó evidencia con datos simulados, que corrobora la pertinencia de los resultados asintóticos en muestras finitas.

Apéndice A. Complemento de resultados conocidos en la literatura

Cuando se trata de estimar una especificación estática, como la representada en la ecuación (3) y se considera que las variables tienen un componente de tendencia estocástico[5], el trabajo de Phillips (1986) permite vislumbrar la posibilidad de identificar, de manera heurística, la regresión espuria mediante el estadístico *DW*; la "regla" señala que, cuando $DW < R^2$, hay regresión espuria. En efecto, cuando la regresión es espuria, dicho estadístico tiende a cero. Lo anterior quedó confirmado en casos en los que las variables se construyen con PGDs más elaborados, que los que empleó Phillips (1986).[6] La sugerencia de revisar el estadístico *DW* con objeto de identificar una posible regresión espuria puede encontrarse en muchos manuales de econometría; a manera de ejemplo, mostramos los siguientes casos:

- **Johnston y Dinardo; Econometric Methods, p.261**: "One should not overreact to these seemingly alarming results **[Se refieren al fenómeno de regresión espuria]**. First, low *DW* statistics are now usually taken as indications of seriously misspecified relations, so that one abandons the relation and works on a re-specification..."

- **Griffiths, Hill y Judge; Learning and Practicing Econometrics, pp. 696-697**: "Granger and Newbold suggested that when a Least Squares regression leads to a high R^2 but a low Durbin-Watson statistic, then the relationship should be estimated in first differences..."

- **Gujarati; Basic Econometrics, pp. 806-807**: "That there is something wrong in the preceding regression is suggested by the extremely low Durbin−Watson [DW] value, which suggests very strong first-order autocorrelation..."

- **Maddala; Introduction to Econometrics, pp. 235-236**: "However, if the DW statistic is very low, it often implies a misspecified equation, no matter what the value of the R^2 is. In such case..."

Los anteriores extractos son válidos en los escenarios bajo los cuales estuvieron escritos. En este artículo, probamos que si modificamos tales escenarios, es decir que cuando se estima una regresión dinámica con

---

[5] Concretamente, consideran el caso en que ambas variables contienen una raíz unitaria.

[6] Ver Marmol (1995), Cappuccio y Lubian (1997) y Marmol (1996b), entre otros.

procesos TS dicha regla definitivamente no aplica, puesto que el estadístico *DW* no sólo no converge a cero, sino que de hecho se mantiene relativamente cercano a "2". En otras palabras, los resultados aquí expuestos complementan dicha regla heurística al señalar que, justamente, no aplica en especificaciones dinámicas donde los procesos son TS. También es relevante mencionar que hay otros casos en los que la regresión es espuria y la *DW* no tiende a cero. Tsay y Chung (2000), quienes generalizaron el trabajo de regresión espuria usando procesos –independientes- de memoria larga de Cappuccio y Lubian (1997), encontraron combinaciones de procesos de esta naturaleza que, al ser utilizadas en la estimación por MCO de la especificación (3), arrojan estadísticos *DW* distintos a cero.[7]

Otra de las consecuencias de los resultados de los **Teoremas 1 y 2**, que complementa modestamente algunos consejos esbozados en los manuales de econometría, tiene que ver con el uso de las especificaciones dinámicas como solución a la regresión espuria. Se sugiere ocasionalmente, que: incluir rezagos de la variable dependiente del lado de los regresores puede resolver el problema de la regresión espuria. Lo anterior, reiteramos, es una práctica válida si el componente de tendencia de las series es estocástico, tal y como suponen los autores de los siguientes extractos de libros de econometría:

- **Johnston y Dinardo; Econometric Methods, p. 261**: "Second, these results [de regresión espuria] are all derived from regressions involving only current variables. If y_t y x_t are independent random walks and y_t is regressed on x_t, then a double specification error has been committed; a relevant variable y_{t-1} has been excluded and..."

- **Davidson y MacKinnon; Econometric Theory and Methods, pp. 611-612**: "Upon reflection, it is not entirely surprising that tests based on the spurious regression model [ver especificación (3): $y_t = \alpha_0 + \beta_0 x_t + \varepsilon_t$]do not yield sensible result [...] It might seem that we could obtain sensible results by running the regression [ver especificación (4): $y_t = \alpha + \beta x_t + \delta y_{t-1} + \varepsilon_t$]"

[7] Conviene aclarar, en este caso, que es relativamente conocido en econometría que la prueba Durbin-Watson de autocorrelación de primero orden, no funciona en especificaciones dinámicas; es necesario emplear la prueba *h* de Durbin. No obstante, recalcamos que la DW también se emplea para evaluar si la regresión es espuria o no. Es en este sentido, que nosotros la estudiamos.

- **Hamilton; Time Series Analysis, pp. 561-562**: "[Cures for Spurious Regressions]...The first approach is to include lagged values of both the dependent and independent variable in the regression..."

Sobra decir que las tres afirmaciones son ciertas, bajo los condiciones dispuestas por los autores. Cuando la naturaleza de la tendencia de las series es estocástica, el incluir variables rezagadas aminora o corrige parcialmente los efectos de la regresión espuria. No obstante, conviene notar que la regresión espuria no sólo se gesta entre variables con tendencias estocásticas, sino también con su equivalente determinista (Kim, Lee y Newbold, 2003; Noriega y Ventosa-Santaulària, 2006 y Noriega y Ventosa-Santaulària, 2007). Cuando los PGD son de este tipo, como se demuestra aquí, incluir un rezago de la variable dependiente entre los regresores no evita el problema ni mucho menos permite identificarlo. [8]

Referencias

Cappuccio, N. y D. Lubian, (1997). Spurious Regressions Between I(1) Processes with Long Memory Errors, *Journal of Time Series Analysis*, 18, 341–354.

Davidson, R. y J. Mackinnon, (2004). Econometric Theory and Methods, Oxford University Press.

Entorf, H. (1997). Random Walks Wiht Drifts: Nonsense Regression and Spurious Fixed-Effect Estimation, *Journal of Econometrics*, 80, 287-296.

Granger, C. W. J. y P. Newbold, (1974). Spurious Regressions in Econometrics, *Journal of Econometrics*, 2, 11–20.

Griffiths, W. E., R. C. Hill, y G. Judge, (1993). *Learning and Practicing Econometrics*, John Wiley & Sons Inc.

Hamilton, J. (1994). *Time Series Analisys*. Princeton.

[8] En este sentido, es importante agregar que la "cura" de la regresión espuria propuesta por Wooldridge (2000) -consistente en incluir una tendencia determinista en la especificación- tampoco funciona cuando las series involucradas en la regresión espuria son estacionarias en tendencia con rompimientos estructurales. Ello está demostrado en Noriega y Ventosa-Santaulària (2006).

Hasseler, U. (2000). Simple Regressions with Linear Time Trends, *Journal of Time Series Analysis*, 21, 27–32.

Hendry, D. F. (1995). *Dynamic Econometrics*, Oxford University Press, Advanced Texts in Econometrics.

Johnston, J. y J. Dinardo (1998). *Econometric Methods*. McGraw Hill.

Kim, T. H., Y. S. Lee, y P. Newbold (2003). Spurious RegressionsWith Processes Around Linear Trends or Drifts, *Discussion Papers in Economics*.

______ (2004). Spurious regressions with stationary processes around linear trenes, *Economics Letters*, 83,2, 257–262.

Maddala, G. (1992). Introduction to Econometrics (ed.), *New York, NY*: *McMillan.*

Marmol, F. (1995). Spurious Regressions Between I(d) Processes, *Journal of Time Series Analysis*, 16, 313–321.

______ (1996a). Correlation Theory of Spuriously Related Higher Order Integrated Processes, *Economics Letter*, 50, 169–173.

______ (1996b). Nonsense Regressions Between Integrated Processes of Different Orders, *Oxford Bulletin of Economics and Statistics*, 58, 525–536.

______ (1998). Spurious Regression Theory with Nonstationary Fractionally Integrated Processes, *Journal of Econometrics*, 84, 233–250.

Noriega, A. E. (2004). Unit Root Testing under Multiple Structural Breaks: a Monte Carlo Study, Escuela de Economía, Universidad de Guanajuato, [Mimeo].

Noriega, A. E. y D. Ventosa-Santaularia, (2006). Spurious Regression Under Broken Trend Stationarity, *Journal of Time Series Analysis*, 27, 671–684.

______ (2007). Spurious Regression and Trending Variables, *Oxford Bulletin of Economics and Statistics*, 69(3), 439-444.

Phillips, P. (1986). Understanding Spurious Regressions in Econometrics, *Journal of Econometrics*, 33, 311–340.

______ (1998). New Tools for Understanding Spurious Regressions, *Econometrica*, 66, 1299–1325.

Tsay, W.J. y C. Chung, (2000). The spurious regression of fractionally integrated processes, *Journal of Econometrics*, 96(1), 155–182.

Wooldridge, J. M. (2000). *Introductory Econometrics. A Modern Approach*. Western College Publishing.