

Banco de México
Documentos de Investigación

Banco de México
Working Papers

N° 2011-05

Una Prueba Simple para Regresiones Espurias

Antonio E. Noriega
Banco de México
Universidad de Guanajuato

Daniel Ventosa-Santaulària
Universidad de Guanajuato

Agosto 2011

La serie de Documentos de Investigación del Banco de México divulga resultados preliminares de trabajos de investigación económica realizados en el Banco de México con la finalidad de propiciar el intercambio y debate de ideas. El contenido de los Documentos de Investigación, así como las conclusiones que de ellos se derivan, son responsabilidad exclusiva de los autores y no reflejan necesariamente las del Banco de México.

The Working Papers series of Banco de México disseminates preliminary results of economic research conducted at Banco de México in order to promote the exchange and debate of ideas. The views and conclusions presented in the Working Papers are exclusively of the authors and do not necessarily reflect those of Banco de México.

Una Prueba Simple para Regresiones Espurias*

Antonio E. Noriega[†]
Banco de México
Universidad de Guanajuato

Daniel Ventosa-Santaulària[‡]
Universidad de Guanajuato

Resumen: Se ha encontrado que el estadístico t para probar la hipótesis nula de no relación entre dos variables independientes diverge asintóticamente bajo una variedad amplia de procesos generadores de datos no estacionarios. En este artículo se introduce un método simple que garantiza la convergencia de este estadístico t hacia una distribución límite pivotal, cuando existen derivas en los procesos integrados generadores de datos, permitiendo así inferencia asintótica. Este método se puede utilizar para distinguir una relación genuina de una espuria entre procesos integrados (I(1) e I(2)). Simulaciones muestran que la prueba tiene buenas propiedades en muestras pequeñas. Al aplicar el procedimiento a datos reales (incluyendo datos sobre bodas y mortalidad en Yule), no se encuentran relaciones (espurias) significativas entre las variables.

Palabras Clave: Regresión Espuria, Procesos Integrados, Eliminación de Tendencia, Teoría Asintótica, Cointegración, Experimentos Monte Carlo.

Abstract: It has been found that the t -statistic for testing the null of no relationship between two independent variables diverges asymptotically under a wide variety of nonstationary data generating processes. This paper introduces a simple method which guarantees convergence of this t -statistic to a pivotal limit distribution, when there are drifts in the integrated processes generating the data, thus allowing asymptotic inference. This method can be used to distinguish a genuine relationship from a spurious one among integrated (I(1) and I(2)) processes. Simulation experiments show that the test has good properties in small samples. When applying the proposed procedure to real data (including the marriages and mortality data of Yule), we do not find (spurious) significant relationships between the variables.

Keywords: Spurious Regression, Integrated Process, Detrending, Asymptotic Theory, Cointegration, Monte Carlo Experiments.

JEL Classification: C12, C15, C22, C46.

*Agradecemos a Carlos Capistrán, Daniel Chiquiar, Russell Davidson, Adrian de la Garza, Niels Haldrup, Søren Johansen, Robinson Kruse, Carlos Lever, Ignacio Lobato, y participantes de los seminarios en Banco de México y CREATES, por sus valiosos comentarios con respecto a la versión anterior de este documento. La versión final de este documento se escribió durante la visita del segundo autor en el Center for Econometric Analysis of Time Series en la Universidad Aarhus, Dinamarca. El agradece a Niels Haldrup y al equipo de CREATES por su hospitalidad y el apoyo financiero del CREATES financiado por Danish National Research Foundation, así como CONACyT, México. Las opiniones en este documento corresponden a los autores y no presentan necesariamente el punto de vista de Banco de México.

[†]Dirección General de Investigación Económica. Email: anoriega@banxico.org.mx.

[‡]Departamento de Economía y Finanzas, Universidad de Guanajuato. Email: daniel@ventosa-santaularia.com.

1. Introducción

Durante muchos años, la literatura de estadística y econometría ha estudiado el fenómeno de relaciones espurias entre variables independientes bajo una variedad amplia de procesos generadores de datos (*PGDs*). Una de las primeras referencias es la de Yule (1926) que encuentra una correlación por arriba de 95 % entre la proporción de bodas en la Iglesia de Inglaterra del total de bodas, y la tasa de mortalidad para los años 1866-1911. Mas recientemente, utilizando métodos computacionales de simulación, Granger y Newbold (1974) encuentran un estadístico t significativo para el parámetro de la pendiente en un modelo de regresión lineal simple, asumiendo caminatas aleatorias independientes sin deriva para las variables explicativas y dependientes. Luego, la teoría asintótica desarrollada por Phillips (1986) proporcionó una explicación teórica de los resultados en el estudio experimental de Granger y Newbold (1974): el estadístico t no tiene una distribución límite, más bien diverge a infinito cuando el tamaño de muestra aumenta, implicando que, asintóticamente, el estadístico t siempre rechazaría la (verdadera) hipótesis nula de no relación. La tasa a la cual el estadístico diverge es $T^{1/2}$, de acuerdo con Phillips (1986).¹ Permitiendo derivas en la representación de caminata aleatoria para las variables dependiente y explicativa en un modelo lineal de regresión, Entorf (1997) muestra que la divergencia ocurre a una tasa T (más rápida). En un artículo reciente, Noriega y Ventosa-Santaulària (2007) muestran que el fenómeno de regresión espuria es omnipresente bajo un rango amplio de combinaciones de *PGDs* para las variables dependiente y explicativa.²

Para fijar ideas, asuma que el interés se central en probar la hipótesis de no relación entre dos variables aleatorias y y x , esto es, $H_0 : \delta = 0$, en el modelo de regresión lineal:

$$y_t = \alpha + \delta x_t + \eta_t \tag{1}$$

en donde el parámetro (desconocido) de la pendiente, δ , mide el impacto instantáneo sobre y de un cambio unitario en x . La ecuación (1) representa un ejemplo para dos variables de una relación de largo plazo (posiblemente cointegrada) que es ampliamente utilizada en estudios empíricos en macroeconomía. Para relaciones cointegradas, δ es interpretado como la constante de equilibrio que liga a y con x en el largo plazo.

¹En un artículo subsecuente, la teoría de representación de Phillips (1998) muestra que un proceso (determinista) estocástico con tendencia se puede representar como una combinación lineal infinita de funciones determinista (estocásticas) con tendencia con coeficientes aleatorios. En tal entorno asintótico, los estadísticos t de regresión de los coeficientes ajustados divergen a una tasa $T^{1/2}$.

²Otros documentos que estudian el fenómeno de regresión espuria bajo distintos supuestos son: Marmol (1995, 1996, 1998), Cappuccio y Lubian (1997), Granger et al. (1998), Tsay y Chung (1999), Hassler (1996, 2000, 2003), Kim, Lee y Newbold (2004), Noriega y Ventosa-Santaulària (2006), y Stewart (2006). Ventosa-Santaulària (2009) presenta una revisión de la literatura.

Los estudios citados arriba indican que el estadístico t para probar una relación lineal entre y y x en el modelo (1), $t_{\hat{\delta}}$, crece con el tamaño de muestra, rechazando así la hipótesis nula asintóticamente. El rechazo de esta hipótesis cuando las variables son independientes es lo que se conoce como regresión espuria.

En busca de un estadístico t convergente en regresiones espurias, Sun (2004) reconoce que la "divergencia del estadístico t usual parece ser una característica definitoria de una regresión espuria" (p.943). Este autor muestra que tal divergencia surge de una subestimación del error estándar del estimador de mínimos cuadrados ordinarios (MCO), y propone un estimador alternativo, basado en el estimador del error estándar HAC con un ancho de banda que es proporcional al tamaño de la muestra. Con este estimador, el estadístico t resultante ya no diverge. Sun (2004) encuentra, sin embargo, que la distribución límite del estadístico t depende de parámetros; en particular, parámetros de memoria de los procesos fraccionales subyacentes que se asumen para las variables dependiente y explicativa. Sun (2004) argumenta que, si bien de interés teórico, estos resultados tienen poca importancia práctica, dado que, en general, los parámetros en un PGD son desconocidos, y, por lo tanto, no es factible llevar a cabo inferencia.

El presente documento aplica un enfoque distinto. Se propone filtrar los parámetros mediante MCO, así se elimina la tendencia lineal de cada variables. Después, se utilizan los residuales de las regresiones para verificar la significancia de la relación a través de una versión reescalada del estadístico t estándar. Granger y Newbold (1974) presentan un enfoque análogo, argumentando, en un contexto de estimar ecuaciones en econometría, que "un método que estamos considerando recientemente es la construcción de modelos de series individuales para cada variable, utilizando los métodos de Box y Jenkins (1970), por ejemplo, y después buscar relaciones entre las series al relacionar los residuales de estos modelos individuales." (pp. 117-118). Además, argumenta que en la construcción de modelos de regresión, la cantidad a explicar no es la variación en la serie original, sino la variación en los residuales.

Utilizando teoría asintótica, en el presente documento se muestra que, cuando ambas variables dependiente y explicativa siguen un proceso integrado, el estadístico t propuesto no dependerá de parámetros y no divergirá, salvo en el caso cuando las variables dependiente y explicativa no sean independientes, eliminando así el problema de regresión espuria. También se calculan valores críticos asintótico y para muestras finitas, los cuales se pueden utilizar para distinguir una relación genuina de una espuria. Los experimentos de simulación revelan que este procedimiento funciona bien en muestras finitas.

La siguiente sección introduce brevemente los tipos de $PGDs$ analizados en el presente documento y que son ampliamente utilizados en estudios empíricos en econometría. La Sección 3 muestra cómo la presencia de los parámetros complica la inferencia estándar, incluso para estadísticos t convergentes (adecuadamente reescalados). Para superar dicho problema, en la Sección

4 se introduce el enfoque descrito antes para probar si existe una relación estadística en un modelo de regresión simple, que resulta en distribuciones convergentes y pivotaes del estadístico t , permitiendo así inferencia asintótica. La Sección 5 presenta experimentos Monte Carlo que reportan las propiedades de nivel y de potencia del procedimiento de prueba propuesto. En la Sección 6, se presentan varias aplicaciones empíricas de dicho procedimiento. Los autores encuentran que, por ejemplo, la alta correlación estadística entre bodas y mortalidad encontrada por Yule (1926) es, en efecto, espuria, una vez que se aplique el procedimiento de filtrar propuesto. La última sección presenta las conclusiones.

2. Mecanismos de tendencia en el PGD

Se considera el siguiente modelo de regresión, estimado por MCO:

$$y_t = \hat{\alpha} + \hat{\delta}x_t + \hat{\eta}_t, \quad (2)$$

utilizado como vehículo para probar la hipótesis nula $H_0 : \delta = 0$. Note que la naturaleza del mecanismo de tendencia en las variables dependiente y explicativa no es conocida a priori. El siguiente Cuadro resume los *PGDs* considerados a continuación para ambas las variables dependiente y explicativa en el modelo (2). Los *PGDs* en el Cuadro 1 incluyen mecanismos de tendencia estocástica, los cuales se usan ampliamente en trabajos aplicados en economía, para modelar variables tales como el producto nominal y real, el consumo, el dinero, los precios, entre otras.

Cuadro 1

SUPUESTO. Los *PGDs* para $z = y, x$ son los siguientes.

<i>PGD</i>	Nombre	Modelo
1.	I(1)	$\Delta z_t = u_{zt}$
2.	I(1)+ <i>deriva</i>	$\Delta z_t = \mu_z + u_{zt}$
3.	I(2)	$\Delta^2 z_t = u_{zt}$
4.	CI(1,1)	$y_t = \alpha_1 + \beta_1 x_t + u_{yt}$
5.	CI(2,1)	$y_t = \alpha_2 + \beta_2 x_t + \xi_{yt}$
6.	Corr(1)	$y_t = \alpha_3 + \beta_3 x_t + \xi_{yt}$

En el Cuadro 1, u_{yt} y u_{xt} son innovaciones independientes que obedecen del Supuesto 1 en Phillips (1986), y $\xi_{yt} = \sum_{i=1}^t u_{yi}$, es decir ξ_{yt} sigue un proceso $I(1)$. El *PGD* 1 es una caminata aleatoria sin deriva, mientras que *PGD* 2 es una caminata aleatoria con deriva, μ_z . *PGD* 3 representa un procesos integrado con doble raíz unitaria, es decir, la serie se tiene que diferenciar dos veces para volverla estacionaria. Bajo *PGD* 4, se asume que x_t sigue *PGDs* 1 o 2 y, si $\beta_1 \neq 0$, entonces

y_t y x_t están cointegradas $CI(1, 1)$, según la notación de Engle y Granger (1987). En este caso, aunque ambas variables son $I(1)$, la combinación lineal produce errores estacionarios. Bajo *PGD 5* se asume que x_t sigue un proceso $I(2)$ y si $\beta_2 \neq 0$, entonces y_t y x_t están cointegradas $CI(2, 1)$. En este caso, una combinación lineal de procesos $I(2)$ reduce el orden de integración a $I(1)$. Finalmente, *PGD 6* corresponde al caso de $x_t \sim I(1)$ con deriva, correlacionada con y_t (asumiendo $\beta_3 \neq 0$), pero no están cointegradas, ya que $\xi_{yt} \sim I(1)$. Este caso se denomina *Corr(1)*, lo que significa que las variables $I(1)$ están (solamente) correlacionadas .

3. La naturaleza divergente del estadístico t

En un artículo reciente, Noriega y Ventosa-Santaulària (2007, NVS de aquí en adelante) mostraron que el estadístico t ($t_{\hat{\delta}}$) en una regresión espuria no tiene una distribución asintótica bajo una variedad amplia de Procesos Generadores de Datos, incluyendo procesos estacionarios en tendencia, procesos con una o dos raíces unitarias, procesos estacionarios en media quebrada o tendencia quebrada, y combinaciones de estos procesos, sino diverge a infinito cuando el tamaño de muestra aumenta.³

Para obtener un estadístico t convergente, se debería reescalarlo por T^κ . NVS encuentran que κ es, en general, $1/2$, pero en algunos casos $\kappa = 1$, o $\kappa = 3/2$, dependiendo del comportamiento de la tendencia de las variables dependiente y explicativa.

El Teorema 1 presenta el comportamiento asintótico del estadístico t ($t_{\hat{\delta}}$) de (2) para cuatro combinaciones de *PGDs* en el Supuesto: 1) una caminata aleatoria sin deriva sobre una caminata aleatoria sin deriva (llamado $t_{\hat{\delta}}^1$), una caminata aleatoria con deriva sobre una caminata aleatoria con deriva ($t_{\hat{\delta}}^2$), 3) un proceso con doble raíz unitaria sobre un proceso con doble raíz unitaria ($t_{\hat{\delta}}^3$), y 4) un proceso con doble raíz unitaria sobre un proceso con doble raíz unitaria bajo cointegración [$CI(2,1)$] ($t_{\hat{\delta}}^4$).⁴ En el Teorema, convergencia en distribución y en probabilidad se denotan como \xrightarrow{D} y \xrightarrow{P} , respectivamente, y W_z es un proceso Wiener estándar, es decir, $W_z(r)$ se distribuye normalmente para cada r en $[0, 1]$; es decir $W_z(r) \sim N(0, r)$. Para simplificar la notación, todos los integrales se calcularon sobre el intervalo $[0, 1]$, con respecto a la medida Lebesgue, es decir, integrales tales como $\int W_z$, $\int W_z^2$, $\int rW_z$, y $\int W_x W_y$ abrevian $\int_0^1 W_z(r)dr$, $\int_0^1 W_z^2(r)dr$, $\int_0^1 rW_z(r)dr$, y $\int_0^1 W_x W_y(r)dr$, respectivamente. También, $\overline{W}_z = \int_0^r W_z(s)ds$, para $s \in [0, 1]$, $s < r$.

La demostración del Teorema 1 se muestra en el Apéndice.⁵

³Estos resultados se obtuvieron por NVS calculando el orden en probabilidad del estadístico t para todas las combinaciones de los *PGDs* considerados.

⁴Los resultados para el caso $CI(1, 1)$ son bien conocidos (Stock, 1987) y, por lo tanto, no se reportan en este documento.

⁵Los resultados en las partes *a*), *b*), y *c*) del Teorema 1 confirman y extienden los resultados en Noriega y Ventosa-Santaulària (2007) (algunos de los cuales ya se habían obtenido en Phillips (1986), Park y Phillips (1989), Marmol (1995) y Entorf (1997)), ya que NVS solamente derivaron el orden en probabilidad del estadístico t y no la correspon-

TEOREMA 1. Se considera probar la hipótesis nula $H_0 : \delta = 0$ en el modelo de regresión (2).

a) $\hat{\delta}^1$ y $t_{\hat{\delta}}^1$ denotan la estimación de MCO de δ , y de estadístico t para probar H_0 , respectivamente, cuando ambas y y x siguen un proceso $I(1)$. Entonces, cuando $T \rightarrow \infty$,

$$\begin{aligned}\hat{\delta}^1 &\xrightarrow{D} (S_{xy} - S_x S_y) (S_{x2} - S_x^2)^{-1} \\ T^{-1/2} t_{\hat{\delta}}^1 &\xrightarrow{D} (S_{xy} - S_y S_x) (S_1)^{-1/2}\end{aligned}$$

b) $\hat{\delta}^2$ y $t_{\hat{\delta}}^2$ denotan la estimación de MCO de δ , y de estadístico t para probar H_0 , respectivamente, cuando ambas y y x siguen un proceso $I(1)$ con deriva. Entonces, cuando $T \rightarrow \infty$,

$$\begin{aligned}\hat{\delta}^2 &\xrightarrow{P} \frac{\mu_y}{\mu_x} \\ T^{-1} t_{\hat{\delta}}^2 &\xrightarrow{D} \mu_x \mu_y (12S_2)^{-1/2}\end{aligned}$$

c) $\hat{\delta}^3$ y $t_{\hat{\delta}}^3$ denotan la estimación de MCO de δ , y de estadístico t para probar H_0 , respectivamente, cuando ambas y y x siguen un proceso $I(2)$ con deriva. Entonces, cuando $T \rightarrow \infty$,

$$\begin{aligned}\hat{\delta}^3 &\xrightarrow{D} \frac{\sigma_y}{\sigma_x} S_3 S_4^{-1} \\ T^{-1/2} t_{\hat{\delta}}^3 &\xrightarrow{D} S_3 S_5^{-1/2}\end{aligned}$$

d) $\hat{\delta}^4$ y $t_{\hat{\delta}}^4$ denotan la estimación de MCO de δ , y de estadístico t para probar H_0 , respectivamente, cuando x_t se genera por un proceso $I(2)$ y y_t por un proceso $CI(2, 1)$, como en el PGD 5. Entonces, cuando $T \rightarrow \infty$,

$$\begin{aligned}\hat{\delta}^4 &\xrightarrow{D} \beta_{y2} S_4 \left(\int \overline{W}_x \right)^{-2} \\ T^{-1/2} t_{\hat{\delta}}^4 &\xrightarrow{D} S_4^{1/2} \int \overline{W}_x \left(\int \overline{W}_x^2 \right)^{-1}\end{aligned}$$

donde $S_z, S_{z2}, S_{xy}, S t_z$, y S_i para $i = 1, 2, \dots, 5$, son funciones del proceso Wiener definidas en el Apéndice A.3.

Como se puede apreciar en el Teorema 1, el parámetro de la pendiente no converge hacia su verdadero valor de cero en las primeras tres combinaciones de PGDs. Para confirmar la naturaleza espuria de la relación, note que el estadístico t diverge en todos los casos, indicando así que la hipótesis nula de no relación se rechazará en muestras grandes. El caso $CI(2, 1)$ de la parte d muestra que la estimación no converge hacia su verdadero valor, β_2 , y su estadístico t asociado diverge a tasa $T^{1/2}$. Además, note que cuando las variables siguen el PGD 2 ($I(1) + deriva$), típicamente el caso destacado en macroeconomía, la distribución asintótica normalizada no es pivotal: ésta depende de los parámetros de deriva deterministas. Por tanto, incluso después de utilizar un reescalamiento adecuado, no es posible realizar inferencia en este caso debido a la presencia de parámetros de perturbación.

diente distribución asintótica, como se hace en el Teorema 1.

4. Una prueba simple para regresión espuria

Al observar los resultados de la sección anterior, es evidente que para el caso del proceso de raíz unitaria con deriva, incluso cuando es conocido el factor de escalamiento necesario para que el estadístico alcance un límite bien definido, la correspondiente distribución asintótica no es pivotal (no está libre de parámetros). Proponemos en esta sección un método simple que filtra los parámetros de perturbación, permitiendo así inferencia asintótica.

El procedimiento empieza con la eliminación de la tendencia lineal de cada variable mediante la siguiente regresión de MCO:

$$z_t = c_z + b_z t + \varepsilon_{zt}, \quad t = 1, 2, \dots, T. \quad (3)$$

para $z = x, y$. Los residuales se definen como:

$$\hat{\varepsilon}_{zt} = z_t - \hat{c}_z - \hat{b}_z t,$$

utilizados para estimar la siguiente ecuación

$$\hat{\varepsilon}_{yt} = c_f + \beta_f \hat{\varepsilon}_{xt} + \nu_t. \quad (4)$$

Como se puede observar, la ecuación (4) utiliza variables generadas: residuales obtenidos de la primera ronda de estimación. Pagan (1984) muestra que, cuando los regresores son residuales de otro modelo, un estimador en dos etapas será consistente y eficiente, e “inferencias validas se pueden realizar con los errores estándar obtenidos como resultado de una regresión de segunda etapa” (p.242).⁶ Adicionalmente, como comprobaron Frisch y Waugh (1933), se hubieran obtenidos resultados idénticos para la estimación de β_f y su estadístico t de (4) si se hubiera utilizado el modelo de regresión (2) con un término adicional de tendencia en el tiempo.⁷

Note que el interés se centra ahora en probar la hipótesis nula de no relación entre dos variables aleatorias sin tendencia determinista, $\hat{\varepsilon}_{yt}$ y $\hat{\varepsilon}_{xt}$, $H_0 : \beta_f = 0$. La relación entre esta hipótesis y la original, $H_0 : \delta = 0$, puede ser entendida a partir de los argumentos de Granger y Newbold (1974), discutidos arriba, en el sentido que en la construcción modelos de regresión, la cantidad a ser explicada no es la variación en la serie original (z_t), sino la variación en la parte residual

⁶Sin embargo, nótese que este resultado se refiere a estimadores de modelos de regresión donde solamente el regresor es una variable generada. Mas aún, no está claro si los resultados en Pagan se mantienen cuando las variables se comportan de manera no estacionaria. En cualquier caso, los resultados asintóticos presentados abajo representan una buena señal, ya que muestran que el estimador del parámetro basado en un modelo de variables generadas converge, bajo cointegración, a su valor de equilibrio de largo plazo, conforme la muestra crece a infinito, mientras que el correspondiente estadístico t diverge, indicando correctamente la presencia de una relación de largo plazo.

⁷También, véase Lovell (2008) o Greene (1997, pp. 246-247).

($\hat{\varepsilon}_{zt}$). En cualquier caso, los resultados presentados abajo indican que $\hat{\beta}_f \xrightarrow{p} \beta_1$ bajo cointegración entre variables $I(1)$, y $\hat{\beta}_f \xrightarrow{p} \beta_2$ bajo cointegración entre variables $I(2)$. En este sentido, $\hat{\beta}_f$ es un estimador consistente del parámetro de pendiente de equilibrio de largo plazo, δ .

El siguiente teorema proporciona la teoría asintótica relacionada con el estimador de MCO $\hat{\beta}_f$, y una versión reescalada de su estadístico t asociado, $T^{-1/2}t_{\beta_f}$, denominado τ , en la ecuación (4). También se reporta el comportamiento asintótico del estadístico R^2 . Note finalmente, que en el Teorema 2 las partes a y b (c y d) se refieren a procesos $I(1)$ ($I(2)$). La demostración se presenta en el Apéndice.

TEOREMA 2. Considere probar la hipótesis nula $H_0 : \beta_f = 0$ en el modelo de regresión (4). El comportamiento asintótico ($T \rightarrow \infty$) del estimador de MCO $\hat{\beta}_f$, de su estadístico t asociado, reescalado por $T^{1/2}$, τ , y del estadístico R^2 se puede representar de la siguiente manera:

a) Cuando x_t se genera por un proceso $I(1)$ o $I(1)+deriva$, y y_t se genera por un proceso $CI(1, 1)$, como en el PGD 4:

$$\begin{aligned}\hat{\beta}_f &\xrightarrow{p} \beta_1 \\ T^{-1/2}\tau &\xrightarrow{D} \lambda_x \sigma_y^{-1} \beta_1 (A_x)^{-1/2} \\ (1 - R^2) &= O_p(T^{-1})\end{aligned}$$

b) Cuando x_t y y_t son independientes entre sí y se generan por procesos $I(1)$ o $I(1) + deriva$:

$$\begin{aligned}\hat{\beta}_f &\xrightarrow{D} (\lambda_y/\lambda_x)w_4A_x \\ \tau &\xrightarrow{D} ND^{-1/2} \\ (1 - R^2) &= O_p(1)\end{aligned}$$

c) Cuando x_t se genera con un proceso $I(2)$, y y_t se genera por un proceso $CI(2, 1)$, como en el PGD 5:

$$\begin{aligned}\hat{\beta}_f &\xrightarrow{p} \beta_2 \\ T^{-1}\tau &\xrightarrow{D} \beta_2(\lambda_x/\lambda_y)(w_5/Q_4^2)^{-1/2} \\ (1 - R^2) &= O_p(T^{-1})\end{aligned}$$

d) Cuando x_t y y_t son independientes y se generaron por procesos $I(2)$:

$$\begin{aligned}\hat{\beta}_f &\xrightarrow{D} (\lambda_y/\lambda_x)Q_5Q_4^{-1} \\ \tau &\xrightarrow{D} Q_5Q_6^{-1/2} \\ (1 - R^2) &= O_p(1)\end{aligned}$$

e) Cuando x_t se genera con un proceso $I(1) + deriva$, y y_t se genera por un proceso $Corr(1)$, como en el PGD 6:

$$\begin{aligned}\hat{\beta}_f &\xrightarrow{D} Q_1/\sigma_xQ_2 \\ \tau &\xrightarrow{D} Q_1(-\sigma_y^2Q_3)^{-1/2}\end{aligned}$$

$$(1 - R^2) = O_p(1)$$

donde N , D , y Q_i para $i = 1, 2, \dots, 6$ son funciones de procesos Wiener definidas en el Apéndice A.4.

El Teorema 2 proporciona resultados útiles. Primero, el parámetro estimado de la pendiente en el modelo de regresión (4), donde se eliminó la tendencia, $\hat{\beta}_f$, converge hacia el parámetro de cointegración del *PGD* 4 en el Supuesto, β_1 , cuando las variables están cointegradas, como se muestra en la parte *a*, implicando que el parámetro de cointegración será estimado consistentemente con el modelo de regresión (4).⁸ Esto implica que la prueba τ propuesta es consistente ya que está construida a partir de un estimador consistente, esto es, $\hat{\beta}_f \xrightarrow{p} \beta_i, i = 1, 2$ bajo la hipótesis alternativa.

Mas aún, bajo cointegración, el estadístico t reescalado diverge, indicando correctamente una relación de largo plazo, como se muestra en la parte *a*. Segundo, como se muestra en la parte *b*, el estadístico t reescalado no diverge para procesos integrados independientes, evitando así el problema de regresión espuria (asintótica). En adición, el estadístico t converge hacia una distribución límite pivotal. Nótese que este resultado se mantiene sin o con deriva en el *PGD*. Conclusiones similares se obtienen para procesos $I(2)$, como se muestra en las partes *c* y *d*. La parte *e* indica que, cuando las variables están correlacionadas pero no están cointegradas, la prueba no tiene potencia, ya que τ no diverge, sino converge hacia una distribución no pivotal. Finalmente, la R^2 converge en probabilidad hacia 1, solamente cuando existe cointegración entre las variables.

En resumen, el estadístico t reescalado (τ) divergirá solamente cuando exista una relación de cointegración de largo plazo entre las variables, de otro modo no aumentará con el tamaño de muestra.

Con base en los resultados anteriores, se propone una prueba simple que permite distinguir entre una relación lineal verdadera entre dos variables aleatorias integradas y una relación espuria. La prueba se basa en τ , el estadístico t reescalado por $T^{1/2}$ de β_f en el modelo de regresión (4) para probar la hipótesis nula $H_0 : \beta_f = 0$. Bajo la hipótesis nula, las variables filtradas son asintóticamente linealmente independientes. Una relación verdadera ocurre cuando se rechaza la hipótesis nula.

En el caso cuando las variables son independientes y siguen una combinación de procesos $I(1)$ e $I(1) + deriva$, la formula resultante (Teorema 2, parte *b*) muestra que la distribución asintótica es pivotal, es decir, no depende de ningún parámetro. Eso implica que este procedimiento permite inferencia a través de un estadístico pivotal adecuadamente reescalado, cuya distribución se puede

⁸Nelson y Kang (1981) argumentan que la dinámica de modelos econométricos estimados con variables integradas, con tendencia eliminada inapropiadamente, puede ser un artefacto del procedimiento de eliminación de la tendencia. Nótese que este fenómeno no parece afectar la consistencia con la cual se estima el parámetro de la pendiente, como se muestra en las partes *a*) y *c*) del Teorema 2.

tabular.

Se simuló la expresión límite para τ en el Teorema 2 (partes b y d) y se generaron valores críticos asintóticos, los cuales se reportan en el Cuadro 2 en el renglón indicado por el símbolo ∞ .⁹

Cuadro 2
VALORES CRÍTICOS PARA LOS ESTADÍSTICOS t REESCALADOS

T	τ [caso I(1)-I(1) con y sin deriva]				τ [caso I(2)-I(2)]			
	1 %	5 %	10 %	20 %	1 %	5 %	10 %	20 %
25	± 1.28	± 0.92	± 0.76	± 0.57	± 5.47	± 3.49	± 2.70	± 1.92
50	± 1.28	± 0.92	± 0.76	± 0.58	± 5.95	± 3.67	± 2.82	± 2.00
100	± 1.28	± 0.92	± 0.76	± 0.58	± 5.87	± 3.73	± 2.86	± 2.01
200	± 1.28	± 0.92	± 0.76	± 0.58	± 5.90	± 3.74	± 2.85	± 2.03
500	± 1.28	± 0.92	± 0.76	± 0.58	± 5.95	± 3.74	± 2.85	± 2.03
1,000	± 1.28	± 0.93	± 0.76	± 0.58	± 6.04	± 3.78	± 2.90	± 2.04
∞	± 1.29	± 0.93	± 0.76	± 0.58	± 6.04	± 3.79	± 2.92	± 2.06

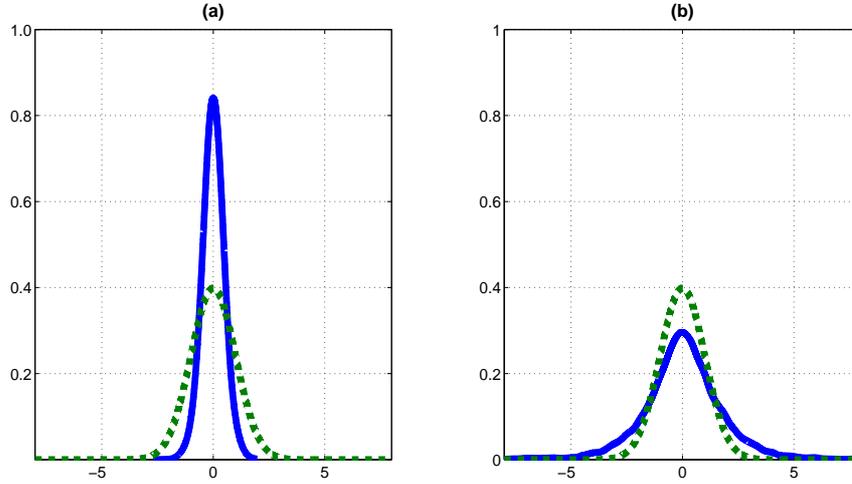
El Cuadro 2 reporta también valores críticos basados en datos simulados para muestras de tamaño $T = 25, 50, 100, 200, 500, 1000$. La parte izquierda del Cuadro 2 muestra los valores críticos para el estadístico t normalizado, τ , para el caso $I(1) - I(1)$ (sin importar si las variables tienen deriva o no), mientras que la parte derecha muestra valores críticos para el caso $I(2) - I(2)$, también basados en el Teorema 2 (parte d). Cabe destacar que los valores críticos asintóticos y de muestras finitas se encuentran muy cercanos uno al otro.

Como una guía para el uso de los valores críticos presentados en el Cuadro 2, supongamos que pruebas de raíz unitaria (tales como las pruebas Dicker-Fuller o Ng-Perron) indujeron al investigador a concluir que ambas y y x son $I(1)$, con un tamaño de muestra $T = 100$. El uso de τ , en conjunto con los valores críticos en el Cuadro 2, permiten probar si existe una relación entre estas dos variables integradas. Un valor (bajo) del estadístico que no rechaza la hipótesis nula (menor que, por ejemplo, 0.92, el valor crítico al nivel de significancia de 5 %), indicará que las variables son dos caminatas aleatorias independientes. Por otro lado, un valor alto del estadístico (mayor que 0.92) indicará que las variables están cointegradas.

La Gráfica 1 muestra la distribución asintótica de τ para el caso $I(1) + deriva$ (panel izquierdo) y para el caso $I(2)$ (panel derecho) bajo la hipótesis nula. Como se puede apreciar, ambas distribuciones muestran una marcada desviación de una distribución normal.

⁹El número de replicaciones es 10,000 y la simulación de los movimientos Brownianos siguen a Perron (1989, p. 375). Un código para Matlab está disponible mediante una solicitud a los autores.

Gráfica 1
Distribución Asintótica del Estadístico τ



Panel (a): τ [caso I(1)+deriva-I(1)+deriva]; Panel (b): τ [caso I(2)-I(2)]. $\sigma_z^2 = 1$ para $z = x, y$; $T = 1,000$, $R = 10,000$. Se representa una distribución normal estándar para la comparación (línea punteada).

4.1. Algunas extensiones

También, se examinó el caso de una regresión espuria entre dos procesos estacionarios en tendencia (*ET*) (véase Kim, Lee y Newbold, 2004) así como combinaciones entre procesos *ET* e $I(d)$, para $d = 1, 2$. Los resultados indican que $t_{\hat{\beta}_f}$ converge hacia una distribución normal estándar bajo la hipótesis nula cuando innovaciones en el *PGD* para ambas variables dependiente y explicativa son *iid*. Sin embargo, si las innovaciones están autocorrelacionadas, la distribución límite de $t_{\hat{\beta}_f}$ bajo la hipótesis nula no está libre de parámetros. Estos resultados, los cuales todavía son preliminares y fuera del alcance del presente documento, se reportaran mas adelante.

5. Propiedades de muestras finitas

5.1. Nivel y potencia de la prueba τ

Se calcularon tasas de rechazo del estadístico τ propuesto para probar la hipótesis nula $H_0 : \beta_f = 0$ en la ecuación (4), utilizando los valores críticos del Cuadro 2 con un nivel nominal de significancia de 5%. Las tasas de rechazo se computaron con base en datos simulados para muestras de tamaño $T = 50, 100, 200, 300$ y 500 , utilizando los modelos del Supuesto (véase el Cuadro 1), y 10,000 replicaciones.

El Cuadro 3 muestra las tasas de rechazo de la prueba τ para distintas combinaciones de valores de los parámetros, asumiendo que x_t y y_t se generaron como procesos independientes $I(1)$ (Panel

Cuadro 3
TASAS DE RECHAZO DEL ESTADÍSTICO DE PRUEBA τ .

		Panel (a)									
Relación	Parámetros					Tamaño de Muestra					
	μ_x	μ_y	β_{y1}	$\rho_{x,1}$	$\rho_{y,1}$	50	100	200	300	500	
PGDs Independientes	0.00	0.00	-	0.00	0.00	0.02	0.02	0.02	0.03	0.02	
			-	0.50	0.50	0.07	0.04	0.04	0.03	0.03	
	0.50	0.30	-	0.00	0.00	0.02	0.03	0.02	0.02	0.02	
			-	0.50	0.50	0.07	0.05	0.04	0.03	0.03	
		1.25	-	0.00	0.00	0.02	0.02	0.02	0.02	0.02	
			-	0.50	0.50	0.07	0.04	0.04	0.03	0.03	
	1.50	1.50	-	0.00	0.00	0.02	0.02	0.02	0.02	0.02	
			-	0.50	0.50	0.07	0.05	0.04	0.03	0.03	
		0.30	-	0.00	0.00	0.02	0.02	0.03	0.03	0.02	
			-	0.50	0.50	0.07	0.05	0.03	0.03	0.03	
	1.50	0.50	-	0.00	0.00	0.02	0.02	0.02	0.02	0.02	
			-	0.50	0.50	0.07	0.04	0.03	0.03	0.03	
1.25		-	0.00	0.00	0.03	0.02	0.03	0.02	0.03		
		-	0.50	0.50	0.07	0.04	0.03	0.03	0.03		
		Panel (b)									
Relación	Parámetros					Tamaño de Muestra					
	μ_x	α_y	β_{y1}	$\rho_{x,1}$	$\rho_{y,1}$	50	100	200	300	500	
Cointegrados	0.00	0.30	0.30	0.00	0.00	0.01	0.10	0.41	0.65	0.91	
			0.30	0.50	0.50	0.03	0.62	0.92	0.99	1.00	
	0.30	0.50	0.30	0.00	0.00	0.02	0.10	0.40	0.65	0.91	
				0.50	0.50	0.28	0.61	0.93	0.99	1.00	
			0.50	0.00	0.00	0.24	0.62	0.94	0.99	1.00	
				0.50	0.50	0.74	0.96	1.00	1.00	1.00	
			1.25	0.00	0.00	0.99	1.00	1.00	1.00	1.00	
				0.50	0.50	1.00	1.00	1.00	1.00	1.00	
			1.50	0.00	0.00	1.00	1.00	1.00	1.00	1.00	
				0.50	0.50	1.00	1.00	1.00	1.00	1.00	
			1.75	0.00	0.00	1.00	1.00	1.00	1.00	1.00	
				0.50	0.50	1.00	1.00	1.00	1.00	1.00	
2.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00				
	0.50	0.50	1.00	1.00	1.00	1.00	1.00				

Panel (a): x_t y y_t independientemente generadas por PGD (2); panel (b): x_t y y_t se generaron por PGD (2), y (4), respectivamente ($x_t, y_t \sim CI(1, 1)$).

a), o procesos cointegrados (Panel b). Para estudiar el efecto sobre el nivel y la potencia de autocorrelación en los disturbios de los procesos, u_{zt} , se permite $u_{zt} = \rho_z u_{zt-1} + \eta_{zt}$, para $\rho_z = 0, 0, 0, 5$ con $z = y, x$, y $\eta \sim iid\mathcal{N}(0, 1)$.¹⁰ Como se puede apreciar, el nivel tiende a ser conservador para distintos valores de T y de los parámetros (con excepción de $T = 50$ y $\rho_y = \rho_x = 0, 5$, en cuyo caso el nivel es 0.07). En el Panel b del Cuadro 3, se observa que la potencia, en general, es alta, salvo para tamaños de muestra menores que $T = 200$ y valores pequeños de los parámetros.

El Cuadro 4 muestra los resultados para el nivel y la potencia bajo el supuesto de que las variables se generaron por un proceso $I(2)$. El Panel a muestra que la prueba tiene el nivel correcto para todos los valores de parámetros y tamaños de muestra. El Panel b corresponde al caso cuando las variables están cointegradas de tal manera que una combinación lineal de dos variables $I(2)$

¹⁰Se asume $\sigma_y^2 = \sigma_x^2 = 1$ en todas las simulaciones.

Cuadro 4
TASAS DE RECHAZO DE τ .

		Panel (a)									
Relación	Parámetros					Tamaño de Muestra					
	σ_{ux}^2	σ_{uy}^2	β_{y2}	$\rho_{x,1}$	$\rho_{y,1}$	50	100	200	300	500	
PGDs Independientes	0.50	0.30	–	0.00	0.00	0.05	0.05	0.05	0.05	0.05	
			–	0.50	0.50	0.06	0.06	0.04	0.06	0.05	
		1.25	–	0.00	0.00	0.04	0.04	0.05	0.05	0.05	
			–	0.50	0.50	0.05	0.05	0.05	0.05	0.04	
		1.50	–	0.00	0.00	0.05	0.05	0.05	0.05	0.04	
			–	0.50	0.50	0.06	0.05	0.05	0.05	0.05	
	1.50	0.30	–	0.00	0.00	0.05	0.05	0.05	0.05	0.06	
			–	0.50	0.50	0.06	0.05	0.05	0.05	0.05	
		0.50	–	0.00	0.00	0.05	0.05	0.05	0.05	0.05	
			–	0.50	0.50	0.06	0.05	0.04	0.04	0.05	
		1.25	–	0.00	0.00	0.05	0.05	0.05	0.05	0.05	
			–	0.50	0.50	0.06	0.05	0.05	0.05	0.06	
		Panel (b)									
Relación	σ_{ux}^2	σ_{uy}^2	β_{y2}	$\rho_{x,1}$	$\rho_{y,1}$	50	100	200	300	500	
Cointegrados	1.00	1.00	0.30	0.00	0.00	0.32	0.68	0.92	0.98	0.99	
				0.50	0.50	0.42	0.72	0.94	0.98	0.99	
			0.50	0.00	0.00	0.58	0.88	0.98	0.99	0.99	
				0.50	0.50	0.66	0.90	0.99	0.99	1.00	
			1.25	0.00	0.00	0.93	0.99	1.00	1.00	1.00	
				0.50	0.50	0.95	0.99	1.00	1.00	1.00	
			1.50	0.00	0.00	0.96	0.99	1.00	1.00	1.00	
				0.50	0.50	0.97	0.99	1.00	1.00	1.00	
			1.75	0.00	0.00	0.98	0.99	1.00	1.00	1.00	
				0.50	0.50	0.98	0.99	1.00	1.00	1.00	
			2.00	0.00	0.00	0.99	0.99	1.00	1.00	1.00	
				0.50	0.50	0.99	0.99	1.00	1.00	1.00	

Panel (a): x_t y y_t se generaron independientemente con PGD (3); panel (b): x_t y y_t se generaron con PGDs (3), y (5), respectivamente ($x_t, y_t \sim CI(2, 1)$).

es $I(1)$, es decir, las variables están $CI(2, 1)$. Como se puede apreciar, la potencia es, en general, muy alta.¹¹

Note que la prueba propuesta se basa en inferencia estadística preliminar utilizada para determinar si existe (o no) una raíz unitaria en cada variable. Dado que esta pre-prueba puede inducir sesgos en el nivel, se aplicó la corrección Bonferroni. Se realizaron experimentos Monte Carlo adicionales donde la prueba τ depende de inferencia deducida de las pruebas Dickey-Fuller aplicadas a las series individuales. Los resultados (no se reportan, pero están disponibles mediante una solicitud a los autores) revelan que la corrección Bonferroni no parece ser necesaria, ya que las propiedades de nivel de la prueba son casi idénticas que las reportadas en los Cuadros 3 y 4.

¹¹También, se estudió la potencia de la prueba, asumiendo que las variables son $I(2)$ y $CI(2, 2)$, es decir, una combinación lineal de dos variables $I(2)$ es $I(0)$. En este caso, la potencia también es muy alta.

5.2. Regresión espuria y cointegración

Mostramos arriba que la prueba propuesta tiene potencia para distinguir entre procesos independientes (I) y cointegrados (CI). Para investigar este asunto más a fondo, se diseñó un experimento Monte Carlo, a través del cual se compara el desempeño en muestras finitas de nuestro estadístico τ con el desempeño de pruebas estándar de cointegración, tales como la prueba basada en residuales de Engle y Granger (1987), y la prueba de Johansen (1988). Para este experimento se generaron 10,000 muestras de tamaños $T = 50, 100, 200, 300$ y 500 de procesos integrados (con deriva) para y_t y x_t , bajo dos hipótesis: las variables son independientes una de la otra, y las variables están cointegradas. Calculamos entonces la proporción de veces que se rechaza la hipótesis nula $H_0 : \beta_f = 0$ en el modelo de regresión (4) a un nivel nominal de significancia de 5 %, bajo cada hipótesis.

La columna 3 en el Cuadro 5 muestra que la prueba τ arroja buenos resultados al distinguir entre procesos independientes y cointegrados: bajo la hipótesis de independencia, la tasa de rechazo es igual al nivel nominal (5 %), mientras que la potencia es casi 80 %, con una muestra pequeña de sólo 50 observaciones. En el Cuadro 5, *EG*, *Tr* y *Eig* se refieren a la prueba de Engle-Granger (1987), la prueba de la traza de Johansen (1988) y la prueba de eigenvalores de Johansen (1988), respectivamente. Nótese que el desempeño de las cuatro pruebas es muy similar para tamaños de muestra mayores a 100 observaciones. Si las innovaciones u_{zt} en el *PGD* están correlacionadas, la columna 4 del Cuadro 5 muestra algunas distorsiones de nivel (la tasa de rechazo llega a 10 % para $T = 50$) y menor potencia (52 %). Sin embargo, estos problemas desaparecen rápidamente cuando aumenta el tamaño de la muestra.

Por lo tanto, se puede utilizar la prueba τ no solamente para distinguir una relación genuina de una espuria, sino también para distinguir entre procesos independientes y cointegrados. Puede haber beneficios al utilizar la prueba τ como prueba de cointegración, dada su relativa simplicidad: por un lado, la prueba de Engle-Granger tiene que controlar por autorcorrelación mediante términos aumentados; por otro lado, la prueba de Johansen requiere una especificación del modelo VAR inicial, así como una decisión con respecto a la inclusión de componentes deterministas.

En general, consideramos que el investigador podría beneficiar al aplicar ambas pruebas, dada su distinta naturaleza: el procedimiento *EG* se basa en las propiedades de los residuales, mientras que la prueba del presente documento se basa en la estimación e inferencia sobre un parámetro. Si ambas pruebas encuentran evidencia de una relación genuina, el investigador debería estar más seguro acerca de la validez de tal inferencia. Por otro lado, si los resultados de las pruebas no coinciden, entonces, el investigador podría considerarlo como evidencia de una inferencia potencialmente errónea y, por lo tanto, debería revisar el ejercicio empírico. Por lo tanto, el uso del estadístico τ se podría considerar como un prueba acompañante en un análisis de cointegración, capaz de confirmar la inferencia derivada de otras pruebas o de poner en duda su validez.

Cuadro 5
TASAS DE RECHAZO DE LAS PRUEBAS τ , ENGLE-GRANGER Y JOHANSEN.

T	Caso	τ	τ^a	EG	EG ^a	1 rezago; parte no det.			
						Tr.	Tr. ^a	Eig.	Eig. ^a
50	I	0.05	0.10	0.05	0.02	0.07	0.09	0.07	0.09
	CI	0.78	0.52	1.00	0.33	0.99	0.28	0.99	0.26
100	I	0.05	0.09	0.05	0.02	0.06	0.07	0.06	0.07
	CI	0.98	0.77	1.00	0.89	1.00	0.77	1.00	0.78
200	I	0.05	0.07	0.05	0.02	0.06	0.06	0.06	0.06
	CI	0.99	0.96	1.00	0.99	1.00	0.99	1.00	0.99
300	I	0.05	0.06	0.04	0.02	0.07	0.06	0.06	0.06
	CI	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00
500	I	0.05	0.06	0.05	0.02	0.08	0.06	0.07	0.06
	CI	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00

EG: Prueba Engle-Granger; Tr/Eig: Pruebas de Traza y Valores Propios (Eigenvalues) de Johansen; ^a Innovaciones siguen un proceso estacionario $AR(1)$: $u_{zt} = 0,75u_{zt-1} + \epsilon_{zt}$; $\epsilon_{zt} \sim \mathcal{N}(0, 1)$; $\mu_x = 0,03$, $\mu_y = 0,04$, $\beta_y = 0,7$ (para relaciones cointegradas), y $R = 1,000$.

Para ilustrar estos argumentos, se realizó un experimento Monte Carlo adicional. Se generaron 10,000 muestras de dos procesos $I(1)$ independientes de tamaños de muestra reportados en la primera columna del Cuadro 6, para estudiar el comportamiento conjunto de la prueba τ y la prueba EG . Dado que los procesos se generan independientemente uno del otro, se espera que las pruebas no rechacen su hipótesis nula, es decir, no relación, y no cointegración, respectivamente. Para verificar lo anterior, se contó el porcentaje de veces que la hipótesis nula fue rechazada, y estos porcentajes se presentan en la columna 2 del Cuadro 6. A nivel nominal de 5 %, y asumiendo que las pruebas son independientes, se debería esperar que resulte algo similar a la siguiente probabilidad

$$\Pr(\tau \text{ no rechaza } H_0 \cap EG \text{ no rechaza } H_0 \mid H_0 \text{ es verdad}) = (1 - \alpha)^2 = (0,95)^2 = 0,9025$$

Como se puede observar a partir de los valores reportados en la columna del Cuadro 6, al parecer, las pruebas son, en efecto, independientes. Las columnas 3 y 4 reportan las estimaciones de la siguiente probabilidad:

$$\begin{aligned} \Pr(\text{una prueba no rechaza } H_0 \cap \text{ la otra prueba rechaza } H_0 \mid H_0 \text{ es verdad}) &= (1 - \alpha)\alpha \\ &= (0,95)0,05 \\ &= 0,0475 \end{aligned}$$

Según los valores reportados, se puede concluir que la probabilidad de que al menos una de las

Cuadro 6

TASAS DE RECHAZO ERROR TIPO-I DE LAS PRUEBAS τ Y EG.				
T	Ambas aceptan	Sólo τ acepta	Sólo EG acepta	Ambas rechazan
50	0.9061	0.0478	0.0418	0.0043
100	0.9110	0.0446	0.0405	0.0039
150	0.9047	0.0420	0.0479	0.0054
200	0.9065	0.0446	0.0441	0.0048
300	0.9048	0.0419	0.0480	0.0053
500	0.9072	0.0438	0.0431	0.0059
1,000	0.9058	0.0424	0.0473	0.0045

Nivel de las pruebas: $\alpha = 0,05$. Innovaciones son $\sim iid\mathcal{N}(0, 1)$; $\mu_x = 0,03$, $\mu_y = 0,04$, y $R = 10,000$.

pruebas no rechace la (verdadera) hipótesis nula es alrededor del 95 % (calculado como la suma de las columnas 2 y 3, o 2 y 4). Finalmente, la columna 5 estima la siguiente probabilidad:

$$\Pr(\tau \text{ rechaza } H_0 \cap EG \text{ rechaza } H_0 \mid H_0 \text{ es verdad}) = \alpha^2 = 0,0025$$

que significa que es muy poco probable que ambas pruebas consigan el resultado equivocado de rechazar la verdadera hipótesis nula.

Con respecto a la potencia, presentamos experimentos Monte Carlo que permiten la presencia de autocorrelación en las innovaciones: en el Cuadro 7, las innovaciones siguen un proceso $AR(1)$ con el parámetro autorregresivo igual a 0.75. En este caso, la prueba EG tiene que incluir un rezago de la variable dependiente como regresor adicional en la regresión auxiliar. Se incluyó tal rezago en la prueba EG y se comparan las tasas de rechazo de la aplicación conjunta de la prueba τ y EG . El Cuadro 7, construido bajo la hipótesis alternativa de cointegración, muestra que cuando el investigador usa ambas pruebas con un tamaño de la muestra $T = 100$, existe una probabilidad de 93 % de rechazar la hipótesis nula con al menos una de las pruebas (0.613+0.148+0.167), que resulta mejor que 76 % (0.613+0.148) o 78 % (0.613+0.167), que son las probabilidades de rechazar la hipótesis nula utilizando solamente el estadístico τ o la prueba EG . Al parecer, para muestras relativamente pequeñas (menores a 100 observaciones), la aplicación de ambas pruebas asegura mayor potencia. Por lo tanto, parecen existir beneficios potenciales en el uso de ambas pruebas, en lugar de solamente una de las dos pruebas.

6. Ilustraciones empíricas

En un documento frecuentemente citado, pero insuficientemente leído, (como argumenta Granger (2001), p.557), Yule (1926) discute por primera vez las correlaciones sin sentido que se encuentran “entre cantidades que cambian con el tiempo, a las cuales no se les puede dar ninguna

Cuadro 7

TASAS DE RECHAZO ERROR TIPO-II DE LAS PRUEBAS τ Y EG.				
T	Ambas rechazan	Sólo τ rechaza	Sólo EG rechaza	Ambas aceptan
50	0.1662	0.3593	0.0977	0.3768
100	0.6131	0.1480	0.1665	0.0724
150	0.8911	0.0116	0.0951	0.0022
200	0.9627	0.0001	0.0372	0.0000
300	0.9952	0.0000	0.0048	0.0000
500	1.0000	0.0000	0.0000	0.0000
1,000	1.0000	0.0000	0.0000	0.0000

Nivel de las pruebas: $\alpha = 0,05$. Innovaciones son $\sim AR(1)$ con $\rho = 0,75$; $\mu_x = 0,03$, $\mu_y = 0,04$, $\beta_y = 0,7$ (para relaciones cointegradas), y $R = 10,000$.

significancia física, . . .” (p.2). Yule (1926) ilustró lo anterior, utilizando datos anuales de la proporción de las bodas en la Iglesia de Inglaterra del total de las bodas (por 1,000 personas) y la tasa de mortalidad (por 1,000 personas) en Inglaterra y Gales para los años 1866-1911.

Yule (1926) encontró un coeficiente de correlación entre estas dos variables de 0.9512, y argumentó que, aunque podría ser posible que la difusión del pensamiento científico y el progreso de la ciencia pueden estar detrás de la caída de bodas y mortalidad, respectivamente, y, por lo tanto, ser un factor común que influye en ambas series, es sin embargo obvio que “la correlación simplemente no tiene sentido, que no tiene significado, que es absurdo suponer que las dos variables en cuestión están causalmente, aunque indirectamente, relacionadas entre sí de cualquier forma” (p.2).

La presente sección tiene dos propósitos. Primero, se utilizan datos reales para ilustrar la posibilidad de encontrar relaciones estadísticas espurias entre las variables, las cuales, a priori, no deberían tener una relación entre sí. Las variables a estudiar son las siguientes:

1. Datos anuales (1866-1911) sobre la proporción de bodas en la Iglesia de Inglaterra respecto del total de las bodas (por 1,000 personas) en Inglaterra y Gales (*bodas*, de aquí en adelante).¹²
2. Datos anuales (1866-1911) sobre la tasa de mortalidad (por 1,000 personas) en Inglaterra y Gales (*mortalidad*, de aquí en adelante).¹³
3. Datos mensuales (1991:1-2005:12) sobre el número total de ventas de carros en los EE.UU. (*carros*, de aquí en adelante).¹⁴

¹²Datos de las *bodas en la Iglesia* son de *Office for National Statistics, U.K.* (<http://www.statistics.gov.uk/cci/nscl.asp?ID=7537>).

¹³Datos para la serie de *tasa de mortalidad* provienen de Mitchell (1988).

¹⁴La fuente es *Bureau of Economic Analysis*: Ventas, Producción, Exportaciones y Inventarios de Carros y Camio-

4. Datos mensuales (1991:1-2005:12) sobre el número de asesinatos en los EE.UU. (*asesinatos*, de aquí en adelante).¹⁵
5. Datos mensuales (1991:1-2005:12) sobre la tasa de inflación en Brasil (*infbrasil*, de aquí en adelante).¹⁶
6. Datos mensuales (1991:1-2005:12) sobre dinero en reserva de bancos comerciales y de desarrollo en México (*BCbancos*, aquí en adelante).¹⁷

Segundo, después de mostrar que las técnicas usuales de regresión de MCO indican la presencia de una relación lineal (espuria) entre las combinaciones de las variables, se muestra que una vez que se implemente el procedimiento propuesto, la relación estadística desaparece, llegando al resultado que, en nuestra opinión, debería ser la inferencia correcta a priori.

Empezamos el análisis empírico estudiando el orden de integración de las variables. Primero, se sigue a Dickey y Pantula (1987) que observaron empíricamente que la probabilidad de rechazar la hipótesis nula de una raíz unitaria (denotada H_1) contra la hipótesis alternativa de estacionariedad (H_0) aumenta con el número de raíces unitarias presentes. Para superar esta posibilidad, utilizamos la metodología propuesta por Pantula (1989) que consiste en un procedimiento secuencial asintóticamente consistente para probar la hipótesis nula H_r : exactamente r raíces unitarias, contra la hipótesis alternativa H_{r-1} : exactamente $(r - 1)$ raíces unitarias, con $r = m, \dots, d + 1, d$, donde m es un número máximo asumido de raíces unitarias presentes en los datos, y d el verdadero número de raíces unitarias presentes en los datos. Pantula sugiere que se tienen que probar las hipótesis secuencialmente en el orden H_m, H_{m-1}, \dots, H_d . En el presente documento, se asume que a priori es conocido que el número máximo posible de raíces unitarias presentes en los datos es 2. Basados en los resultados de Pantula, se realizan pruebas de raíz unitaria "hacia abajo", iniciando con una prueba de hipótesis nula H_2 : exactamente dos raíces unitarias (o una raíz unitaria en las primeras diferencias de los datos). Si se rechaza la hipótesis nula H_2 , entonces se prueba la hipótesis nula H_1 : una raíz unitaria, contra la hipótesis alternativa de estacionariedad, si no, se concluye que existen dos raíces unitarias en las series.

Este procedimiento se implementa utilizando siete pruebas: la prueba de raíz unitaria de Dickey-Fuller Aumentada (véase Said y Dickey (1984)), la prueba de estacionariedad KPSS¹⁸ (véase

netas (miles) tomados de www.FreeLunch.com - <http://www.economy.com/freelunch>. El periodo de muestra (1991:1-2005:12) abarca 180 observaciones.

¹⁵La fuente es *FBI: Crime in the United States; BOC: County City Data Book*, tomados de www.FreeLunch.com - <http://www.economy.com/freelunch>.

¹⁶Calculado como el cambio porcentual del índice de precios al consumidor. Banco Central do Brasil. <https://www3.bcb.gov.br/sgspub/localizarseries/localizarSeries.do?method=prepararTelaLocalizarSeries>

¹⁷Miles de pesos, saldo nominal. La fuente es Banco de México. <http://www.banxico.org.mx/sitioingles/billetesymonedas/estadisticas/banknotescoins/banknotescoinsincircu.htm>

¹⁸En el caso de la prueba de estacionariedad KPSS, primero se prueba la hipótesis nula H_1 contra H_2 , es decir, la

Kwiatkowski, et al. (1992)), la prueba de raíz unitaria de punto óptimo de ERS (véase Elliott, et al. (1996)) y las cuatro pruebas de raíz unitaria con buenas propiedades de nivel y potencia de Ng-Perron (véase Ng y Perron (2001)). El Cuadro 8 resume las propiedades de series de tiempo de las variables.

Como se puede concluir de este Cuadro, la mayoría de las variables siguen, al parecer, un proceso de raíz unitaria, con la excepción de *asesinatos*, que parece ser un proceso $I(2)$, y *bodas* que algunas pruebas describen como $I(1)$ y otras como $I(2)$. El Cuadro 9 resume la inferencia que resulta de aplicar las pruebas.

Utilizando varias combinaciones de estas series integradas, se examinan los resultados de regresión bajo dos enfoques: mínimos cuadrados ordinarios lineales usando (1), y el procedimiento propuesto en la Sección 4 que utiliza $\tau = T^{-1/2}t_{\beta_f}$ de la estimación de MCO de la ecuación (4) como estadístico de prueba. El Cuadro 10 presenta los resultados.

Cuadro 8
RESULTADOS DE LAS PRUEBAS DE RAÍZ UNITARIA

Series	ADF	KPSS [†]	ERS [†]	Ng-Perron			
				MZa	MZt	MSB	MPT
Δ bodas	-2.03	0.83***	3.61*	-5.68	-1.67*	0.29	4.36*
bodas	—	—	15.10	—	-1.60	—	16.11
Δ mortalidad	-9.48***	0.02	3.30*	-14.91***	-2.61***	0.17**	2.11**
mortalidad	-1.19	0.43***	12.02	-7.49	-1.93	0.26	12.16
Δ carros	-19.41***	0.013	0.48***	-122.74***	-7.80***	0.06***	0.25***
carros	-1.40	75.90***	319.00	-0.15	-0.20	1.35	338.37
Δ asesinatos	-0.81	147.42***	38.05	-1.07	-0.72	0.68	22.69
asesinatos	—	—	—	—	—	—	—
Δ Infbrasil	-12.92***	0.048	0.28***	-149.93***	-8.66***	0.06***	0.16***
Infbrasil	-2.48	11.86***	7.68	-11.93	-2.42	0.20	7.75
Δ BCbancos	-19.12***	0.01	0.342***	-122.39***	-7.81***	0.06***	0.22***
BCbancos	0.22	119.45***	132.67	-1.77	-0.93	0.53	50.86

***, ** y * denotan significancia a nivel de 1%, 5% y 10%, respectivamente.

[†] Los estadísticos reportados son los que estiman el espectro de residuos en frecuencia cero mediante MCO (los resultados no cambian al utilizar métodos MCO sin tendencia).

El valor máximo de longitud de rezagos es $k_{\text{máx}} = \text{int}(12(T/100)^{1/4})$, véase Ng y Perron (2001).

La longitud de rezagos se seleccionó mediante MBIC, MAIC y MHQ; salvo en la prueba de Ng-Perron donde el rezago se seleccionó como en Perron y Qu (2007).

hipótesis nula de estacionariedad en primeras diferencias de los datos, contra la alternativa de dos raíces unitarias. Si la hipótesis nula H_1 se rechaza, entonces la prueba termina y se concluye que la serie tiene dos raíces unitarias. Si no se rechaza la hipótesis nula, se continúa probando la hipótesis nula de estacionariedad, H_0 , contra la hipótesis alternativa de una raíz unitaria, H_1 .

Cuadro 9
ORDEN DE INTEGRACIÓN SEGÚN LAS PRUEBAS DE RAÍZ UNITARIA.

Serie	ADF	KPSS	ERS	Ng-Perron			
				MZa	MZt	MSB	MPT
<i>bodas</i>	<i>I(2)</i>	<i>I(2)</i>	<i>I(1)</i>	<i>I(2)</i>	<i>I(1)</i>	<i>I(2)</i>	<i>I(1)</i>
<i>mortalidad</i>	<i>I(1)</i>						
<i>carros</i>	<i>I(1)</i>						
<i>asesinatos</i>	<i>I(2)</i>						
<i>Infbrasil</i>	<i>I(1)</i>						
<i>BCbancos</i>	<i>I(1)</i>						

Cuadro 10

y_t	x_t	t_{δ}	τ
<i>mortalidad</i>	<i>bodas</i>	23.542***	0.477
<i>carros</i>	<i>Infbrasil</i>	-6.390***	-0.062
<i>carros</i>	<i>BCbancos</i>	10.108***	-0.006
<i>Infbrasil</i>	<i>BCbancos</i>	-11.258***	-0.118
<i>mortalidad</i>	Δ <i>bodas</i>	0.688	0.010
<i>carros</i>	Δ <i>asesinatos</i>	2.047**	-0.068
<i>Infbrasil</i>	Δ <i>asesinatos</i>	-0.393	0.258
<i>BCbancos</i>	Δ <i>asesinatos</i>	4.424***	0.063

***, ** y * denotan significancia a nivel de 1 %, 5 % y 10 %, respectivamente.

Los resultados del Cuadro 10 indican que, en casi todos los casos, la regresión de MCO para variables aparentemente independientes resultará en el rechazo de la hipótesis nula de no relación, llevando a uno a concluir a favor de una relación espuria. El caso de *mortalidad* sobre Δ *bodas* sugiere que *bodas* es, en efecto, $I(1)$, no $I(2)$. Si ese fuera el caso, la aplicación del operador en diferencias eliminaría la tendencia estocástica en esta variable, es decir, Δ *bodas* $\sim I(0)$. Como se mostró en Noriega y Ventosa-Santaulària (2007), el estadístico t no diverge si una de las variables carece de tendencia. Por lo tanto, en este caso se concluye que la inferencia estándar a través de t_{δ} en el modelo de regresión (2) para estas variables indicaría un rechazo espurio de la hipótesis nula.

La última columna del Cuadro 10 muestra que el procedimiento propuesto indica, como era de esperarse, que las variables no están relacionadas estadísticamente, dado que el estadístico τ no es significativo a niveles convencionales de significancia, utilizando valores críticos del panel

izquierdo del Cuadro 2.

7. Conclusiones

El presente documento ha propuesto un procedimiento simple para superar el problema de regresión espuria en un modelo de regresión lineal simple cuando las variables son procesos integrados. Se estudian dos casos, uno en el cual las variables dependiente y explicativa son procesos integrados del orden uno (con y sin deriva), que es el caso más común en muchos estudios empíricos en macroeconomía, y el otro caso donde las variables son integradas del orden dos.

En el contexto de un modelo de regresión lineal simple, se conoce bien que, cuando las variables dependiente y explicativa siguen un proceso $I(1)$ con deriva, el estadístico t del parámetro de la pendiente diverge, mientras que el estadístico reescalado correspondiente converge hacia una distribución bien definida, expresada en términos de procesos Wiener, pero dependiente de parámetros (los parámetros de deriva). En este documento se introduce un enfoque simple basado en filtrar linealmente los datos, que resulta en un estadístico t con una distribución asintótica bien definida y libre de parámetros (pivotal). Se tabularon la distribución asintótica de este estadístico y su contraparte de muestras finitas, y se reportan los valores críticos para varios tamaños de muestra y niveles de significancia. Como se menciona en el artículo, el procedimiento se puede aplicar en un solo paso, simplemente estimando el modelo de regresión incluyendo una tendencia lineal como regresor extra.

La teoría asintótica detrás del procedimiento propuesto en este documento implica que, cuando las variables están cointegradas, la prueba rechazará la hipótesis nula de no correlación. Por otro lado, cuando las variables son independientes, entonces, la prueba no rechazara asintóticamente. Un experimento Monte Carlo revela que el estadístico de prueba propuesto en este artículo hace un muy buen trabajo en distinguir entre variables independientes y cointegradas y, por lo tanto, se podría considerar como “prueba acompañante” en un análisis de cointegración.

Finalmente, se aplicó el procedimiento propuesto al conjunto de datos de Yule (1926) sobre tasas de bodas y mortalidad, y se encontró que, al aplicar el método propuesto en el presente documento, ya no se encuentra una relación (espuria) significativa entre estas variables. Algunos ejercicios empíricos adicionales confirman que este procedimiento parece funcionar en la práctica.

Referencias

Box, G. y G. Jenkins (1970), *Time Series Analysis, Forecasting, and Control*, San Francisco: Holden Day.

- Cappuccio, N. y D. Lubian (1997), "Spurious Regressions Between $I(1)$ Processes with Long Memory Errors", *Journal of Time Series Analysis*, 18, 341-354.
- Dickey, D. A. y S. G. Pantula (1987), "Determining the Order of Differencing in Autoregressive Processes", *Journal of Business and Economic Statistics*, 5(4), 455-461.
- Entorf, H. (1997), "Random Walks with Drifts: Nonsense Regression and Spurious Fixed-Effect Estimation", *Journal of Econometrics*, 80, 287-296.
- Engle, R.F. y C.W.J. Granger (1987), "Cointegration and Error Correction: Representation, Estimation and Testing", *Econometrica*, 55, 251-76.
- Elliott, G., T.J. Rothenberg y J.H. Stock (1996), "Efficient Tests for an Autoregressive Unit Root", *Econometrica* 64, 813-836.
- Frisch, R. y F. V. Waugh (1933), "Partial Time Regressions as Compared with Individual Trends", *Econometrica*, 1(4), 387-401.
- Granger, C.W.J. y P. Newbold (1974), "Spurious Regression in Econometrics", *Journal of Econometrics*, 2, 111-120.
- Granger, C.W.J., N. Hyung, y Y. Jeon (1998), "Spurious Regressions with Stationary Series", Mimeo.
- Granger, C.W.J (2001), "Spurious Regressions in Econometrics", Capítulo 26 en *A Companion to Theoretical Econometrics*, editado por Badi H. Baltagi, Blackwell Publishing.
- Greene, W. (1997), *Econometric Analysis*, third edition, Prentice Hall.
- Hassler, U. (1996), "Spurious Regressions when Stationary Regressors are Included", *Economics Letters*, 50, 25-31.
- Hassler, U. (2000), "Simple Regressions with Linear Time Trends", *Journal of Time Series Analysis*, 21, 27-33.
- Hassler, U. (2003), "Nonsense Regressions Due to Neglected Time-Varying Means", *Statistical Papers*, 44, 169-182.
- Johansen, S. (1988), "Statistical Analysis of Cointegrated Vectors", *Journal of Economic Dynamics and Control*, 12, 231-254.
- Kim, T.-H., Y.-S. Lee y P. Newbold (2004), "Spurious Regressions with Stationary Processes Around Linear Trends", *Economics Letters*, 83, 257-262.
- Kwiatkowski, D., P.C.B. Phillips, P. Schmidt y Y. Shin (1992), "Testing the Null Hypothesis of Stationary against the Alternative of a Unit Root", *Journal of Econometrics*, 54, 159-178.
- Lovell, M. C. (2008), "A Simple Proof of the FWL (Frisch-Waugh-Lovell) Theorem", *The Journal of Economic Education*, 39(1), 88-91.
- Marmol, F. (1995), "Spurious Regressions for $I(d)$ Processes", *Journal of Time Series Analysis*, 16, 313-321.
- Marmol, F. (1996), "Nonsense Regressions Between Integrated Processes of Different Orders",

Oxford Bulletin of Economics and Statistics, 58(3), 525-36.

Marmol, F. (1998), "Spurious Regression Theory with Nonstationary Fractionally Integrated Processes", *Journal of Econometrics*, 84, 233-50.

Mitchell B.R. (1988), *British Historical Statistics*, Cambridge University Press.

Nelson, C. R., y H. Kang (1981), "Spurious Periodicity in Inappropriately Detrended Time Series", *Econometrica*, 49(3), 741-751.

Ng, S. y P. Perron (2001), "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power", *Econometrica* 69(6), 1519-54.

Noriega, A. y D. Ventosa-Santaulària (2006), "Spurious Regression Under Broken-Trend Stationarity", *Journal of Time Series Analysis*, 27(5), 671-684.

Noriega, A. y D. Ventosa-Santaulària (2007), "Spurious Regression and Trending Variables", *Oxford Bulletin of Economics and Statistics*, 69(3), 439-444.

Pagan, A. (1984), "Econometric Issues in the Analysis of Regressions with Generated Regressors", *International Economic Review*, 25(1), 221-247.

Pantula, S. G. (1989), "Testing for Unit Roots in Time Series Data", *Econometric Theory*, 5, 256-271.

Park, J. Y. y P. C. B. Phillips (1989), "Statistical Inference in Regressions with Integrated Processes: Part 2", *Econometric Theory*, 5, 95-131.

Perron, P. y Z. Qu (2007), "A Simple Modification to Improve the Finite Sample Properties of Ng and Perron's Unit Root Tests", *Economics Letters*, 94, 12-19.

Phillips, P.C.B. (1986), "Understanding Spurious Regressions in Econometrics", *Journal of Econometrics*, 33, 311-340.

Phillips, P.C.B. (1998), "New Tools for Understanding Spurious Regressions", *Econometrica*, 66(6), 1299-1325.

Said, S. E. y D.A. Dickey (1984), "Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order", *Biometrika*, 71, 599-607.

Sargan, J. D. (1980), "Some Tests of Dynamic Specification for a Single Equation", *Econometrica*, 48, 879-897.

Stewart, C. (2006), "Spurious Correlation of $I(0)$ Regressors in Models With an $I(1)$ Dependent Variable", *Economics Letters*, 91(2), 184-189.

Stock, J.H. (1987), "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors", *Econometrica*, 55(5), 1035-1056.

Sun, Y. (2004), "A Convergent t -statistic in Spurious Regressions", *Econometric Theory*, 20, 943-962.

Tsay, W. J. y C. F. Chung (2000), "The Spurious Regression of Fractionally Integrated Processes", *Journal of Econometrics*, 96(1), 155-82.

Ventosa-Santaulària, D. (2009), “Spurious Regression”, *Journal of Probability and Statistics*, article ID 802975, 27 pages, doi:10.1155/2009/802975.

Yule, G. U. (1926), “Why Do We Sometimes Get Nonsense-Correlations Between Time-Series?—A Study in Sampling and the Nature of Time-Series”, *Journal of the Royal Statistical Society*, 89(1), pp. 1-63.

APÉNDICE A.1

La ecuación (1) se puede escribir en forma matricial $y = X\beta + \varepsilon$, donde y es un vector $T \times 1$ de datos y_t , X es una matriz $T \times 2$ incluyendo una constante y datos sobre x_t , y ε es un vector $T \times 1$ de residuales con media cero. El vector de estimadores de *MCO* se define como:

$$\hat{\beta} = \begin{bmatrix} \hat{\alpha} \\ \hat{\delta} \end{bmatrix} = (X'X)^{-1}X'y,$$

donde (todas las sumas van desde $t = 1$ a T) $X'X = \begin{bmatrix} T & \Sigma x_t \\ \Sigma x_t & \Sigma x_t^2 \end{bmatrix}$, y $X'y = \begin{bmatrix} \Sigma y_t \\ \Sigma x_t y_t \end{bmatrix}$. El estadístico t se define por

$$t_{\hat{\delta}} = \hat{\delta} [\hat{\sigma}_{\varepsilon}^2 (X'X)^{-1}]^{-1/2},$$

donde $\hat{\sigma}_{\varepsilon}^2$ es la varianza de la regresión estimada,

$$\hat{\sigma}_{\varepsilon}^2 = \frac{\Sigma (y_t - \hat{\alpha} - \hat{\delta} x_t)^2}{T}$$

y $(X'X)^{-1}_{22}$ denota el segundo elemento diagonal en $(X'X)^{-1}$. Mediante un código de *Mathematica 7.0*, para cada combinación de *PGDs* en el Supuesto para y y x , se calcula el orden de magnitud de $\hat{\delta}$, $\hat{\sigma}_{\varepsilon}^2$, y $(X'X)^{-1}_{22}$, y, por lo tanto, se deriva el orden de magnitud de $t_{\hat{\delta}}$. El código también permite derivar una expresión para la distribución asintótica de $t_{\hat{\delta}}$. Este código está disponible en <http://dl.dropbox.com/u/1307356/SpRegTest/NVSSpRegTest1.zip>

APÉNDICE A.2

De igual manera como en el Apéndice A.1, la ecuación (3) se puede escribir en forma matricial $z = X\beta + \varepsilon$, donde z es un vector $T \times 1$ de datos ($z_t = y_t, x_t$), X es una matriz $T \times 2$ que incluye una constante y una tendencia lineal t , y ε es un vector $T \times 1$ de residuales con media cero. El vector de estimadores de *MCO* se define como:

$$\hat{\beta} = \begin{bmatrix} \hat{c}_z \\ \hat{b}_z \end{bmatrix} = (X'X)^{-1}X'z,$$

donde $X'X = \begin{bmatrix} T & \Sigma t \\ \Sigma t & \Sigma t^2 \end{bmatrix}$, y $X'z = \begin{bmatrix} \Sigma z_t \\ \Sigma z_t t \end{bmatrix}$. Mediante un código de *Mathematica 7.0* (disponible en <http://dl.dropbox.com/u/1307356/SpRegTest/NVSSpRegTest2.zip>), se pueden calcular expresiones analíticas para $\hat{\beta}$. Los residuales resultantes, $\hat{\varepsilon}_{yt}$ y $\hat{\varepsilon}_{xt}$ se utilizan después para estimar

por MCO el modelo de regresión (4):

$$\widehat{\beta}_f = \Sigma \widehat{\varepsilon}_{yt} \widehat{\varepsilon}_{xt} (\Sigma \widehat{\varepsilon}_{xt}^2)^{-1}$$

Nótese que en este caso $\Sigma \widehat{\varepsilon}_{zt} = \widehat{c}_f = 0$, por construcción. El código de *Mathematica 7.0* resuelve lo demás al derivar una expresión límite para $\widehat{\beta}_f$ (y también una expresión límite para t_{β_f}). El comportamiento de la R^2 se obtiene al estudiar el comportamiento asintótico de la suma de cuadrados de errores (SCE) y la suma de cuadrados totales (SCT) de $R^2 = 1 - SCE/SCT$, donde $SCE \sum \widehat{\nu}_t^2 = T \widehat{\sigma}_\nu^2$ y $SCT = \sum \widehat{\varepsilon}_{yt}^2 = T \widehat{\sigma}_{\varepsilon_y}^2$, las cuales se pueden obtener a partir del código de *Mathematica 7.0*.

APÉNDICE A.3

Las definiciones para las expresiones utilizadas en el Teorema 1 son las siguientes. Para $z = x, y$:

$$\begin{aligned} S_z &= \sigma_z \int W_z, S_{z2} = \sigma_z^2 \int W_z^2, S_{xy} = \sigma_x \sigma_y \int W_x W_y, St_z = \sigma_z \int r W_z, y \\ S_1 &= S_{x2} S_{y2} - S_{y2} S_x^2 - S_{x2} S_y^2 + 2 S_x S_y S_{xy} - S_y S_x^2 \\ S_2 &= (\mu_x^2)^{-1} [\mu_y^2 A_x + \mu_x^2 A_y + 2 \mu_x \mu_y (12 St_x St_y - 6 St_y S_x - 6 St_x S_y + 4 S_x S_y - 6 S_{xy})] \\ S_3 &= \int \overline{W}_x \int \overline{W}_y - \int \overline{W}_x \overline{W}_y, S_4 = (\int \overline{W}_x)^2 - \int \overline{W}_x^2 \\ S_5 &= 2 \int \overline{W}_x \int \overline{W}_y \int \overline{W}_x \overline{W}_y - (\int \overline{W}_x \overline{W}_y)^2 - \int \overline{W}_x^2 (\int \overline{W}_y)^2 - (\int \overline{W}_x)^2 \int \overline{W}_y^2 + \int \overline{W}_x^2 \int \overline{W}_y^2 \\ \text{con } A_z &= S_{z2} - 4(3 St_z^2 - 3 St_z S_z + S_z^2), \text{ y } \sigma_z^2 = \lim_{T \rightarrow \infty} (T^{-1} \sum_{t=1}^T u_{zt}^2). \end{aligned}$$

APÉNDICE A.4

Las definiciones para las expresiones utilizadas en el Teorema 2 son las siguientes.

$$\begin{aligned} N &= 6 \int r W_x (\int W_y - 2 \int r W_y) + \int W_x (6 \int r W_y - 4 \int W_y) + \int W_x W_y \\ D &= 4 (\int W_x)^2 \int W_y^2 - 12 (\int r W_x)^2 [(\int W_y)^2 - \int W_y^2] + 12 (\int W_x)^2 (\int r W_y)^2 \\ &\quad - 12 \int W_x \int r W_x [(\int W_y)^2 - 2 \int W_y \int r W_y] + w_{y1} \int W_x^2 + 4(w_2 + w_3) \int W_x W_y + (\int W_x W_y)^2 \\ Q_1 &= \beta_{y3} \sigma_x w_{x1} + \sigma_y [2(w_2 - w_3) - \int W_x W_y] \\ Q_2 &= 4 (\int W_x)^2 - \int W_x^2 - 12 \int r W_x (\int W_x - \int r W_x) \\ Q_3 &= 12 \int r W_y [2 \int W_x \int r W_x \int W_y - \int W_x^2 \int W_y - (\int W_x)^2 \int W_y + \int W_x^2 \int r W_y] \\ &\quad + (\int W_y)^2 [4 \int W_x^2 - 12 (\int r W_x)^2] + w_{x1} \int W_y^2 + \int W_x W_y [4(w_2 + w_3) + \int W_x W_y] \end{aligned}$$

$$Q_4 = 4 (\int \bar{W}_x)^2 - 12 \int \bar{W}_x \int r\bar{W}_x + 12 (\int r\bar{W}_x)^2 - \int \bar{W}_x^2$$

$$Q_5 = 4 \int \bar{W}_x \int \bar{W}_y - 6 \int r\bar{W}_x \int \bar{W}_y - 6 \int \bar{W}_x \int r\bar{W}_y + 12 \int r\bar{W}_x \int r\bar{W}_y - \int \bar{W}_x \bar{W}_y$$

$$\begin{aligned} Q_6 = & (\int \bar{W}_x \bar{W}_y)^2 - 4 \left(3 \int r\bar{W}_x^2 - \int \bar{W}_x^2 \right) (\int \bar{W}_y)^2 - 12 \left(\int \bar{W}_x^2 - 2 \int \bar{W}_x \int r\bar{W}_x \right) \int \bar{W}_y \int r\bar{W}_y \\ & + 12 \left[(\int \bar{W}_x^2 - (\int \bar{W}_x)^2) (\int r\bar{W}_y)^2 + Q_4 \int \bar{W}_y^2 \right. \\ & \left. - 4 \int \bar{W}_x \bar{W}_y (2 \int \bar{W}_x \int \bar{W}_y - 3 \int r\bar{W}_x \int \bar{W}_y - 3 \int \bar{W}_x \int r\bar{W}_y + 6 \int r\bar{W}_x \int r\bar{W}_y) \right] \end{aligned}$$

$$w_{z1} = 4 (\int W_z)^2 - 12 \int W_z \int rW_z + 12 (\int rW_z)^2 - \int W_z^2$$

$$w_2 = 3 \int rW_x \int W_y - 2 \int W_x \int W_y$$

$$w_3 = 3 \int W_x \int rW_y - 6 \int rW_x \int rW_y$$