Conceptos básicos y aspectos matemáticos sobre el análisis de raíces unitarias y cointegración

Carlos A. Rodríguez Ramos *

"La unidad de la teoría y la práctica no se da sólo en la teoría, sino que subsiste también en la práctica".

Lukacs (1969)

I. Introducción

El inicio de la metodología econométrica tradicional, la cual sienta las bases de la econometría moderna, puede ubicarse próximo a la fundación de la Sociedad de Econometría en 1930 y posteriormente a los trabajos de la Fundación de Cowles (Landreth y Colander, 1998; Galindo, 1995). Esta última se basó en el planteamiento de Trygve Haavelmo en cuanto al supuesto de que el mejor enfoque de la econometría es el probabilístico, en el cual las ecuaciones estructurales tienen una mejor distribución del término de error (Landreth y Colander, 1998).

El desarrollo de la econometría en los años cuarenta y cincuenta estuvo orientado, en gran parte, a la creación de nuevas técnicas y al inicio del uso de las computadoras (Galindo, 1995). Se desarrollaron modelos de ecuaciones simultáneas con el propósito de simular y proyectar diversas series económicas y para análisis de la estructura de la economía. Lo que se pretendía era generar un "afinamiento preciso". Es decir, ante cualquier eventualidad de tipo económico, el modelo indica cuál deberá ser el tipo de política económica a realizarse (Landreth y Colander, 1998; Mankiw, 1990).

A partir de la década de los setenta, estos modelos fueron altamente criticado ya que no pudieron explicar, de manera precisa, los choques económicos ocurridos (Landreth y

Colander, 1998; Bashkara, 1994). Esto generó un gran escepticismo y falta de credibilidad en los resultados econométricos. Lo anterior se tradujo en un desencanto, especialmente para propósitos de política económica (Galindo, 1995). Se señalaba que la mayoría de los investigadores se basaban sólo en buscar el mejor ajuste estadístico. Además, ante la disponibilidad limitada de datos, se crea la necesidad de utilizar sustitutos, los cuales pueden ser o no los apropiados. Por último, se ha señalado que rara vez en economía se realizan experimentos controlados, lo que ocasiona que cualquier confianza en el resultado sea desconocida y dependa del análisis subjetivo.

Pero la crítica más contundente fue la "Crítica de Lucas", desarrollada por el economista Robert Lucas (1972), de la Universidad de Chicago. Este argumentó que las acciones individuales son función de las políticas esperadas; por lo que, la estructura del modelo va a cambiar con las políticas empleadas. Pero, si la estructura fundamental del modelo cambia, también lo hará la política a utilizarse y el modelo ya no será el apropiado. Quiere decir que es inadecuado utilizar los modelos econométricos para pronosticar los efectos de la política económica (Landreth y Colander, 1998, Lucas, 1972).

Como respuesta a esta situación, desde el punto de vista econométrico, se empezaron a resolver estos problemas mediante el uso de modelos uniecuacionales.[1] Luego, a principio de la década de los ochenta, se introdujo el uso de modelos de vectores autorregresivos (VAR), los cuales pretendían no imponer restricciones, a priori, a los datos (Sims, 1980; Rodríguez, 2002; Rodríguez, 2001b).

Más tarde, Nelson y Plosser (1982) demostraron que una gran cantidad de variables de la economía de Estados Unidos muestran variaciones en su media o en su varianza y algunas en ambas. Es decir, no presentan momentos de primer y segundo orden constantes, sino que frecuentemente éstos son función del tiempo (Rodríguez, 2001). Así, se observa que estas variables presentan una tendencia a aumentar a través del tiempo y a acentuarse su variabilidad.

---

[1] Antes del uso de estos modelos, la metodología econométrica se basaba en el método de ecuaciones simultáneas (Charemza y Deadman, 1993). La utilización de este tipo de modelos tiene una gran tradición en la econometría.

Si el investigador no considera este fenómeno puede cometer diversos errores, entre ellos el de tipo espurio. El análisis de estacionariedad, por lo tanto, es clave para todo el análisis posterior. La presencia de no estacionariedad en la media puede recogerse si se introducen elementos deterministas en la especificación del proceso (Rodríguez, 2001; Suriñach, Artís, López y Sansó, 1995). Si la introducción de estos elementos deterministas captura la no estacionariedad en medio del proceso, la inferencia estándar es aplicable bajo los supuestos clásicos (Suriñach, Artís, López y Sansó, 1995). Por su parte, cuando la varianza es función del tiempo, puede deberse a la existencia de una raíz unitaria en el polinomio de la representación autorregresiva del proceso (Suriñach, Artís, López y Sansó, 1995; Bhaskara, 1994; Enders, 1995).[2] Este tipo de tendencia se conoce como estocástica.

La importancia que, para el análisis de un sistema económico dado y en la toma de decisiones de política económica, tiene el determinar la existencia de una raíz unitaria en el proceso autorregresivo y, dado esto, determinar su orden de integración, se pone de manifiesto en las distintas respuestas de las variables ante choques no anticipados (Suriñach, Artís, López y Sansó, 1995). No considerar este análisis puede conducir a serios errores de especificación. También surge el problema de la sobre identificación, la cual ocasiona una pérdida de eficiencia e invalidación de las pruebas al incluir un esquema de media móvil no invertible en los errores (Suriñach, Artís, López y Sansó, 1995; Novales, 1997; Maddala y Kim, 1998, Patterson, 2000).

El estudio de variables no estacionarias se puede analizar en un contexto multivariable. Ya que la existencia de una similitud en el orden de integración de las series puede mostrar una relación estable a través del tiempo, lo que sugiere la posibilidad de que también se cumpla a largo plazo (Novales, 1997; Bhargava, 1996; Rodríguez, 2001b).

[2] Esta tendencia en varianza que se analiza es la que es provocada por la existencia de una raíz unitaria en el polinomio autorregresivo y no por la presencia de raíces en el polinomio autorregresivo dentro del círculo unidad. A diferencia de las raíces unitarias éstas no desaparecen al aplicar el operador diferencia (1-L).

Este punto fue en el que se basó Granger (1981) para demostrar el concepto de cointegración y su equivalencia con el modelo de corrección de errores.[3]

El análisis de cointegración es esencial cuando se tiene una combinación de variables que presenten una similitud en el orden de integración. Una similitud en el orden de integración sugiere la necesidad de utilizar series que cointegren para obtener estimadores insesgados y consistentes y resolver el problema de regresiones espurias (Rodríguez, 2001; Rodríguez y Luciano, 2001).

En el caso en que exista una relación de cointegración entre las series, se minimiza la varianza del residual en el espacio paramétrico y los estimadores resultan también ser superconsistentes, ya que convergen a su verdadero valor (Rodríguez, 2001a; Novales, 1997; Galindo y Cardero, 1998). Si la especificación de la existencia de este fenómeno es incorrecta, se pueden cometer errores en la modelación económica, al aceptar como válidas relaciones de tipo espurio, cuando se analizan las características de las estimaciones obtenidas en el proceso de inferencia (Bhargava, 1986; Maddala, 1996; Maddala y Kim, 1998; Enders, 1995). Es decir que, no llevar dicho análisis correctamente, en términos de política económica, puede conducir a conclusiones erróneas en términos de la toma de decisiones (Rodríguez, 2002). Este planteamiento es clave en cualquier modelo econométrico con series de tiempo.

El objetivo del presente trabajo es presentar el concepto de raíces unitarias y el de cointegración y cuáles son las pruebas econométricas más utilizadas, para corroborar la presencia de estos fenómenos en el análisis de series de tiempo. El trabajo se divide en cinco partes. En estas se presentan, de manera general, los conceptos de estacionariedad e integrabilidad de la serie así como algunas pruebas de raíces unitarias. Por último, se presenta el concepto de cointegración.

II. Estacionariedad e integrabilidad

[3] Este tipo de modelo analiza cómo los desajustes en el corto plazo se ajustan a la dinámica de largo plazo.

La mayoría de los trabajos econométricos están basados en el supuesto de estacionariedad en las variables lo que quiere decir que su distribución de probabilidad no es función del tiempo (Rodríguez, 2002). Pero se puede observar que, en prácticamente todos los países, la mayoría de las variables han sufrido variaciones tanto en su media como en su varianza (Rodríguez, 2001; Rodríguez, 2002). Esto quiere decir que los momentos de primer y segundo orden no son constantes sino que, en muchas ocasiones, son función del tiempo. Estas variables presentan una tendencia a aumentar a través del tiempo, acentuándose su variabilidad.

Si el investigador no considera este fenómeno puede cometer diversos errores, entre ellos el de tipo espurio. El análisis de estacionariedad, por lo tanto, es clave para todo el análisis posterior. En el análisis de series de tiempo se utiliza el concepto de estacionariedad en sentido débil, o de segundo orden, y se refieren al mismo como estacionariedad (Bhaskara, 1994). Esto es, al considerar una serie de tiempo como la realización de un proceso estocástico, ésta se considera estacionaria en el sentido débil si los momentos de primer y segundo orden no son función del tiempo. Es decir:

(1) $E\left[x(t_i)\right] = E\left[x(t_i + h)\right] = \square_1 \square$

(2) $E\left[x(t_i)\right]^2 = E\left[x(t_i + h)\right]^2 = \square_2 \square$

(3) $E\left[x(t_i)x(t_j)\right] = E\left[x(t_i + h) + x(t_j + h)\right] = \square_{ij} \square$

Son constantes y no- función del tiempo

Una serie puede ser no estacionaria de acuerdo al comportamiento de su media y/o varianza. En el caso de la presencia de no estacionariedad en la media, ésta es función del tiempo (Bhaskara, 1994; Suriñach, Artís, López y Sansó, 1995). Este comportamiento puede recogerse si se introducen elementos deterministas: tendencias lineales; tendencias polinómicas; tendencias segmentadas; y variables mudas en la especificación del proceso generador de datos (PGD). Si la introducción de estos elementos deterministas capturan la no estacionariedad en medio del proceso, la inferencia estándar es aplicable bajo los supuestos clásicos. En este caso, los estimadores de mínimos cuadrados ordinarios (MCO) tendrán distribuciones asintóticas normales.

Por su parte, cuando la varianza es función del tiempo, se genera un tipo de tendencia conocida como estocástica. Esto origina que las distribuciones utilizadas en la inferencia estándar no sean aplicables. Estas pueden ser generadas por:

1. La presencia de raíces en el polinomio autorregresivo dentro del círculo unitario. Estas no desaparecen al aplicar el operador de diferencias (1-L)
2. La presencia de una raíz unitaria.

El segundo punto puede verse mediante un paseo aleatorio:

(4)
$$x_t = \rho x_{t-1} + \varepsilon_t$$
$$x_t - \rho x_{t-1} = \varepsilon_t$$
$$(1 - L)x_t = \varepsilon_t$$

Para que exista una raíz unitaria, el parámetro de la ecuación (4) debe ser igual a uno. En este caso, cuando:

$t = 0$

$$x_0 = constante$$

$t = 1$

$$x_1 = \rho x_0 + \varepsilon_1$$

$t = 2$

$$x_2 = \rho x_1 + \varepsilon_2$$
$$x_2 = \rho\left[\rho x_0 + \varepsilon_1\right] + \varepsilon_2$$
$$x_2 = \rho^2 x_0 + \rho\varepsilon_1 + \varepsilon_2$$

$t = 3$

$$x_3 = \rho x_2 + \varepsilon_3$$
$$x_3 = \rho^3 x_0 + \rho^2\varepsilon_1 + \rho\varepsilon_2 + \varepsilon_3$$

En términos generales:

(5) $$x_t = \rho^t x_0 + \sum_{i=0}^{t-1} \rho^i \varepsilon_{t-i}$$

Analizando el término de error en (5), así como la varianza de x_t, se corrobora que lo que hace que cambie la varianza es el tiempo. Por ejemplo, evaluando (5) en t=2 y t=3 e i desde 0 a 2 se obtiene:

$t = 2$

$$x_t = \rho^2 x_0 + \rho^0 \varepsilon_{2-0} + \rho^1 \varepsilon_{2-1} + \rho^2 \varepsilon_{2-2}$$

$$x_t = \rho^2 x_0 + \varepsilon_2 + \rho\varepsilon_1$$

$t = 3$

$$x_t = \rho^3 x_0 + \rho^0 \varepsilon_{3-0} + \rho^1 \varepsilon_{3-1} + \rho^2 \varepsilon_{3-2}$$

$$x_t = \rho^3 x_0 + \varepsilon_3 + \rho\varepsilon_2 + \rho^2 \varepsilon_1$$

Dado que el parámetro es igual a uno y el término de error cambia a través del tiempo:

(6) $$Var(x_t) = t\sigma^2$$

En este caso lo que hace que varíe la varianza es el tiempo. Este tipo de tendencia se conoce como estocástica. Por lo que, retomando la ecuación (4), se obtiene una nueva serie estacionaria en varianza:

(7) $$x_t = \rho x_{t-1} + \varepsilon_t \Rightarrow x_t - \rho x_{t-1} = \varepsilon_t \Rightarrow (1-\rho L)x_t = \varepsilon_t \Rightarrow x_t = \frac{\varepsilon_t}{(1-\rho L)}$$

Como pudo verse en los párrafos anteriores, una tendencia determinística es una cuya media es función del tiempo y una estocástica es cuando lo es la varianza. Las propiedades que presentan las variables con un tipo u otro de tendencia difieren. Existen cuatro procesos generadores de datos (PGD) como ejemplo para cada una de las posibles combinaciones:

1. Ausencia de tendencias - Suponga que el PGD de x_t es generado por el proceso autorregresivo estacionario; $x_t =$ _$+ \rho x_{t-1} + e_t$ con e_t ruido blanco [$e_t \sim N(0,\sigma^2)$] y

$(\square\ \square\square<1)$.

(7) $$x_t \square \square x_{t\square 1} = \square + e_t \quad \square \quad (1\square\square L)x_t = \square + e_t \quad \square \quad x_t = \frac{\square}{(1\square\square)} + \frac{e_t}{(1\square\square)}$$

Su media y varianza estarán dados por:

(8) $$E(x_t) = E\square\frac{1}{1\square\square}(\square + e_t)\square = \frac{\square}{1\square\square}$$

(9) $$Var(x_t) = Var\square\frac{1}{1\square\square}(\square + e_t)\square = \frac{\square_e^2}{1\square\square}$$

Se puede ver que ambas son finitas y no son función del tiempo. En este caso, la estimación de mínimos cuadrados para el estimador será consistente pero sesgada por la presencia como regresor de la variable endógena retardada, siendo su distribución asintótica normal.

2. Tendencia determinística - Sí, por ejemplo, el PGD de x_t está dado por un proceso autorregresivo estacionario sobre una tendencia determinística; x_t = _ + $\square$t + _x_{t-1} + $\__t$ con e_t ruido blanco y ($\square$_$\square$< 1).

(10) $$(1\square\square x_{t\square 1}) = \square + \square t + \square_t$$

$$(1\square\square L)x_t = \square + \square t + \square$$

$$x_t = \frac{\square}{(1\square\square L)} + \frac{\square t}{(1\square\square L)} + \frac{\square_t}{(1\square\square L)}$$

Su media y varianza estarán dados por:

(11) $$E(x_t) = \frac{\square}{(1\square\square L)} + \frac{\square}{(1\square\square L)}E(t)$$

(12) $$Var(x_t) = \square vat(t) + \frac{\square_\square^2}{(1\square\square L)} = \frac{\square_\square^2}{(1\square\square L)}$$

En este caso la varianza es independiente del tiempo pero no así la media. La especificación de la tendencia determinística para la estimación del modelo elimina completamente el problema de la no estacionariedad en la media. Los parámetros estimados por MCO dan estimadores con las propiedades adecuadas (Suriñach, Artís, López y Sansó, 1995). A las variables que siguen este proceso se les conoce como *estacionarias sobre una tendencia*, por lo que no es necesario diferenciar la variable para llegar a la estacionariedad. Si se diferenciara la serie se obtendrá un proceso con media móvil no invertible.[4] En este caso la estimación por MCO de los parámetros será consistente pero no eficiente.

3. Tendencia estocástica - Si el PGD de x_t es la suma del valor inicial mas todos los choques aleatorios, como en (4) y suponiendo que $x_0 = 0$, su varianza está dada por la ecuación (6) y el valor esperado es igual a cero. En este caso la varianza es función del tiempo. La estimación por MCO será consistente pero sesgada tendiendo a subestimar el valor poblacional de parámetro. Este estimador no se distribuye asintóticamente como una normal por lo que hay una discontinuidad en la distribución asintótica del parámetro de MCO cuando vale 1. Esto impide la aplicación de la inferencia estándar sobre el parámetro.

4. Tendencia determinística junto a la estocástica - Si el PGD de x_t es $x_t = \_ + x_{t-1} + \__t$ con $\__t$ ruido blanco. Al sustituirse recursivamente, y suponiendo que $x_t = 0$, se obtiene:

(13)
$$x_0 = 0$$
$$x_1 = \square + x_0 + \square_1$$
$$x_2 = \square + x_1 + \square_2$$
$$x_2 = \square + (\square + x_0 + \square_1) + \square_2$$
$$x_2 = 2\square + x_0 + \square_1 + \square_2$$
$$x_3 = \square + x_2 + \square_3$$
$$x_3 = \square + (2\square + x_0 + \square_1 + \square_2) + \square_3$$
$$x_3 = 3\square + x_0 + \square_1 + \square_2 + \square_3$$

[4] Cuando un proceso de medias móviles es invertible, existe una representación autorregresiva, en la cual los valores pasados de la variable dependiente reciben un peso menor cuando más alejados estén del tiempo. Este proceso es fundamental, en especial para procesos de proyección. Por lo que se tiene que invertir primero el proceso de medias móviles para obtener su representación autorregresiva.

En términos generales:

(14) $$x_t = t\alpha + \sum_{i=0}^{t-1} \varepsilon_{t-i}$$

Analizando (14) en t=2 y t=3 e i desde cero hasta dos en t=2 y desde cero hasta 3 tres en t=3 se obtiene:

$t = 2$

$$x_t = 2\alpha + \varepsilon_{2-0} + \varepsilon_{2-1} + \varepsilon_{2-2}$$

$$x_t = 2\alpha + \varepsilon_2 + \varepsilon_1 + \varepsilon_0$$

$t = 3$

$$x_t = 3\alpha + \varepsilon_{3-0} + \varepsilon_{3-1} + \varepsilon_{3-2} + \varepsilon_{3-3}$$

$$x_t = 3\alpha + \varepsilon_3 + \varepsilon_2 + \varepsilon_1 + \varepsilon_0$$

La media y la varianza son:

(15) $$E(x_t) = t\alpha$$

(16) $$Var(x_t) = t\sigma^2$$

Tanto la media como la varianza son función del tiempo, por lo que el proceso presenta tendencia en éstas. A las variables que siguen un comportamiento similar se les denomina *estacionarias por diferenciación*.

II.1. Implicaciones de política económica

Utilizando, como ejemplo, un paseo aleatorio como en (4)

(4) $$x_t = \rho x_{t-1} + \varepsilon_t$$

el cual al sustituirse recursivamente se obtiene:

(5) $$x_t = \rho^t x_0 + \sum_{i=0}^{t-1} \rho^i \varepsilon_{t-i}$$

Se pueden analizar las implicaciones de política económica de acuerdo al valor absoluto del parámetro de la ecuación. En el caso en el que el valor absoluto del parámetro se encuentre entre 0 y uno (caso de estabilidad), la influencia del valor inicial (x_0) y de los choques pasados decaen a medida que aumenta el tamaño de la muestra. Por lo que el presente es más importante que el pasado y las medidas no anticipadas de política económica (choques aleatorios) tienden a perder efecto, a través del tiempo. Al ser el valor absoluto de la ecuación igual a uno (caso de una raíz unitaria), la influencia del valor inicial (x_0) y de los choques pasados y presentes tienen la misma importancia, teniendo efectos permanentes en el nivel de la variable. Todas las medidas no anticipadas de política económica afectan la evolución presente y futura de la variable.

Cuando el valor absoluto del parámetro es mayor a uno (caso explosivo), la influencia del valor inicial (x_0) y de los choques pasados se vuelve cada vez más importante a medida que el tiempo pasa, implicando que el pasado es más importante que el presente. Esta situación no suele observarse en series económicas.

II.2. Procesos I(0), I(1) e I(2)

Algunos modelos teóricos, como el de Nelson y Plosser (1982), sugieren la existencia de raíces unitarias en las variables (Suriñach, Artís, López y Sansó, 1995; Enders, 1995; Rodríguez, 2002). Por esto, muchas series macroeconómicas, en términos reales, son I(1) y nominales son I(2) (Suriñach, Artís, López y Sansó, 1995; Enders, 1995; Patterson, 2000; Enders, 1995; Bhaskara, 1994).

Un proceso I(0) tiene las siguientes características:

1. Una media constante y una tendencia de la serie a volver ante cualquier desviación. Es decir, tiende a fluctuar sobre la media.
2. Una varianza finita e independiente del tiempo.
3. Los efectos de los choques son transitorios y van decreciendo en el tiempo.

Un proceso I(1) se caracteriza por:

1. La serie no se mantiene sobre un valor medio a través del tiempo.
2. La varianza depende del tiempo y tiende a infinito cuando la variable tienda a infinito.
3. Los choques aleatorios tienen efectos permanentes en el proceso.

Una serie I(2) presenta básicamente las mismas características que una I(1). La presencia de raíces unitarias en la representación autorregresiva del proceso provoca momentos de primer y segundo orden que cambian a través del tiempo. Esto hace que la inferencia clásica no sea utilizable ya que ésta se basa en el supuesto de estacionariedad, por lo que se necesita probar la existencia de una raíz unitaria en el PGD.

III. Pruebas de raíz unitaria

Una prueba alternativa sobre estacionariedad que se ha empleado con frecuencia en los últimos años se conoce como la prueba de raíz unitaria (Gujarati, 1997). Esta prueba es sumamente importante ya que el rechazo de la hipótesis nula de raíz unitaria en favor de alternativas estacionarias tiene interpretaciones económicas importantes, admitiendo la posibilidad de relaciones a largo plazo entre variables económicas. Además, en algunas aplicaciones, es deseable probar no estacionariedad vs. alternativas explosivas (Bhargava, 1986). La forma más fácil de introducir esta prueba es considerando el siguiente modelo:

(17) $$Y_t = Y_{t-1} + u_t$$

donde u_t es el término de error estocástico que sigue los supuestos clásicos, a saber: tiene media cero, varianza constante, y no está autocorrelacionado. Un término de error con tales propiedades es conocido también como ruido blanco.

La ecuación es una regresión de primer orden, o AR(I), en la cual se relaciona el valor de Y en t sobre t - l. Si el coeficiente de Y_{t-1}, es en realidad igual a 1, surge lo que se conoce como raíz unitaria (una situación de no estacionariedad). Por consiguiente, si se efectúa la regresión (caminata aleatoria):

(17') $$Y_t = \rho Y_{t-1} + u_t$$

y se encuentra que ρ=1, entonces se dice que la variable estocástica Y tiene una raíz unitaria.

La ecuación puede rescribirse como:

(18) $\Delta Y_t = (\rho-1)Y_{t-1} + u_t$

(18') $= \delta Y_{t-1} + u_t$

donde $\delta=(\rho-1)$, Δ es el operador de primera diferencia estacionaria.

Si δ es igual a 0, se puede escribir (18) como:

(19) $\Delta Y_t = (Yt-Y_{t-1}) = u_{t.}$

Esta ecuación indica que la primera diferencia de una serie de tiempo de caminata aleatoria es una estacionaria porque u_t es puramente aleatorio.

Desde el punto de vista estadístico, existen dos problemas (Maddala, 1996; Novales, 1997): el primero es con respecto a los métodos de eliminación de tendencia que se emplean (regresión o diferencias). Según Maddala (1996), los resultados de autocorrelación son espurios siempre que se elimine la tendencia de una serie en diferencia estacionaria o se diferencie una serie de tendencia estacionaria. El otro problema es que la distribución del estimado de mínimos cuadrados del parámetro autorregresivo tiene una distribución no estacionaria cuando existe una raíz unitaria. Para esto es preciso calcular la distribución en forma numérica en cada caso, dependiendo de las demás variables que se incluyen en la regresión. Esto representa la proliferación de las pruebas de raíces unitarias y las tablas asociadas (Maddala, 1996).

III.a. Dickey-Fuller (DF) y Dickey-Fuller aumentada(ADF)

Para analizar si una serie de tiempo "Y" es no estacionaria, se efectúa la regresión (17') y se determina si "ρ" es estadísticamente igual a 1 o, en forma equivalente, se estima e investiga sí δ=0 (Gujarati, 1997; Johnston y DiNardo, 1997; Pindyck y Rubinfeld, 1997). El valor t así obtenido no sigue la distribución t de "Student" aun en muestras grandes (Gujarati, 1997; Johnston y DiNardo, 1997; Pindyck y Rubinfeld, 1997; Pankratz, 1995).

Bajo la hipótesis nula de que □=1, el estadístico "t" calculado convencionalmente se conoce como "□", cuyos valores críticos han sido tabulados por Dickey y Fuller con base en simulaciones de Monte Carlo (Gujarati, 1997; Johnston y DiNardo, 1997; Pindyck y Rubinfeld, 1997; Patterson, 2000). Esta prueba se conoce como la prueba Dickey-Fuller (DF). Si la hipótesis nula de que □=1 es rechazada (la serie de tiempo es estacionaria), se puede utilizar la prueba "t" usual (de Student) (Gujarati, 1997; Johnston y DiNardo, 1997; Pindyck y Rubinfeld, 1997, Bahskara, 1994).

Sin embargo, estas tablas no son totalmente adecuadas y han sido ampliadas por MacKinnon a través de simulaciones de Monte Carlo (Gujarati, 1997; Johnston y DiNardo, 1997). Si el valor absoluto calculado del estadístico □ (es decir, |□|) excede los valores absolutos críticos de (DF) o de MacKinnon, DF, entonces no se rechaza la hipótesis de que la serie de tiempo dada es estacionaria. Por el contrario, si éste es menor que el valor crítico, la serie de tiempo es no estacionaria.

Además del estadístico "t" existen otras dos estadísticas de prueba (Maddala,1996):

1. K(1) = T(□ - 1), F(0, 1);
2. F(0, 1) es la estadística F usual bajo la prueba de hipótesis conjuntas de que el parámetro de la tendencia es cero y □* es igual a uno.

La prueba Dickey-Fuller se aplica a regresiones efectuadas en las siguientes formas:

(18') $Yt = □Y_{t-1} + u_t$

(19) $□Y_t = □_1 + □Y_{t-1} + u_t$

(20) $□Y_t = □_1 + □_2\, t + □Y_{t-1} + u_t$

donde "t" es la variable de tiempo o tendencia. En cada caso, la hipótesis nula es que □ = 1, es decir, que hay una raíz unitaria. La diferencia entre (18') y las otras dos regresiones se encuentra en la inclusión del intercepto y el término de tendencia.

Si el término de error está autocorrelacionado, se modifica (20) como sigue (Gujarati, 1997; Johnston y DiNardo, 1997):

$$(21) \qquad \square Y_t = \square_1 + \square_2 t + \square_1 Y_{t\square 1} + \sum_{i=1}^{n} \square_i \square Y_{t\square i} + \square_t$$

Se utilizan términos en diferencia rezagados y se incluirán hasta que el término de error no tenga autocorrelación. Esta prueba se conoce como Dickey-Fuller aumentada (ADF). Las pruebas de hipótesis serán las mismas que en la prueba DF.

Dado que (21) incluye una constante, su significancia estadística puede corroborarse al analizar el valor calculado del estadístico t con los valores críticos de la distribución empírica $t_{\square'\square}$, presentados en el trabajo original de Dickey y Fuller (1981). Por otro lado, la prueba de hipótesis conjunta en (21), que indica que el PGD es un paseo aleatorio, se puede probar a través del estadístico F presentado también en el trabajo de Dickey y Fuller(1981) y así, sucesivamente, para la significancia de los diversos parámetros. En este caso, para tener conocimiento del PGD y probar de manera más concisa la existencia de una raíz unitaria, se recomienda seguir los siguientes pasos (Bhaskara, 1994) tomando en consideración el cuadro esquemático para las pruebas de raíces unitarias:

1. Estime la ecuación que aparece en la ecuación (21).
2. Utilice el estadístico $\square_3$, el cual aparece en el trabajo de Dickey y Fuller (1981) para corroborar; $H_o:(\square_1, \square_2, \square_1) = (\square_1, 0, 1)$ vs. $H_A:(\square_1, \square_2, \square_1) \neq (\square_1, 0, 1)$. En el caso de que la hipótesis sea rechazada siga el paso 3. De no ser así vaya al paso 5.
3. Pruebe que $\square_1 = 1$, utilizando el estadístico t del primer paso, con los valores críticos de la tabla normal estándar. En el caso de que la hipótesis no sea rechazada, se concluye que $\square_2 \neq 0$ y $\square_1 = 1$. Si se rechaza la hipótesis nula vaya al paso 4.
4. Utilice la prueba t convencional para decidir si se debe tomar a $\square_2$ como cero o no. En el caso en que $\square_1 \neq 1$, y la hipótesis nula para $\square_2$ sea aceptada, se concluye que la serie es estacionaria sin tendencia. Al ser la hipótesis nula rechazada se concluye que la serie es estacionaria con una tendencia lineal. En ambos casos, se puede probar la hipótesis concerniente al parámetro $\square_1$ de manera convencional.

5. Dado que se acepta la hipótesis nula en el paso 2, se sabe que la serie tiene una raíz unitaria, sin tendencia pero con una posible constante. Para apoyar esta decisión utilice el estadístico t para probar que $\square_1$=1, suponiendo que $\square_2$=0. El estadístico t requerido es el mismo que en (3), pero se requieren valores críticos no-estándar. Estos son invariantes con respecto al valor de $\square_1$. Suponiendo que este estadístico provee la verificación que se busca, se procede al paso 6.
6. Realice la prueba $\square_2$ para ($\square_1$, $\square_2$, $\square_1$)= (0, 0, 1). Si esta prueba permite concluir que $\square_1$ =0, se concluye que la serie es un paseo aleatorio sin tendencia. De otra forma será un paseo aleatorio con tendencia. En ambos casos vaya al paso #7.
7. Imponga la restricción de que $\square_2$=0 y trate de restablecer la conclusión concerniente a $\square_1$ y $\square_1$ usando (21), pero sin tendencia, y el estadístico $\square_1$, para probar la hipótesis nula de una raíz unitaria y cero tendencia.

Cuadro 1
Cuadro esquemático para pruebas de raíces unitarias

Estadístico	Ecuación	H$_o$:	H$_A$:	Distribución	Supuestos requeridos para derivar los valores críticos
t_1	$y_t = \square_1 y_{t-1} + e_t$	$\square$=1	$\square$<1	F(8.5.2)* arriba	$\square$=1
t_2	$y_t = \square_1 y_{t-1} + \square_1 + e_t$	$\square$=1	$\square$<1	F(8.5.2) centro	($\square_1$, $\square_1$)= (0, 1)
t_3	$y_t = \square_1 y_{t-1} + \square_1 + \square_2 t + e_t$	$\square$=1	$\square$<1	F(8.5.2) abajo	($\square_1$, $\square_2$, $\square_1$)= (0, 0, 1)
$\square_1$	$y_t = \square_1 y_{t-1} + \square_1 + \square_2 t + e_t$	($\square_1$, $\square_2$, $\square_1$)= ($\square_1$, 0, 1)	($\square_1$, $\square_2$, $\square_1$)≠ ($\square_1$, 0, 1)	DF(VI)**	($\square_1$, $\square_2$, $\square_1$)= (0, 0, 1)
$\square_2$	$y_t = \square_1 y_{t-1} + \square_1 + \square_2 t + e_t$	($\square_1$, $\square_2$, $\square_1$)= (0, 0, 1)	($\square_1$, $\square_2$, $\square_1$) ≠ (0, 0, 1)	DF(V)	($\square_1$, $\square_2$, $\square_1$)= (0, 0, 1)
$\square_3$	$y_t = \square_1 y_{t-1} + \square_1 + e_t$	($\square_1$, $\square_1$)= (0, 1)	($\square_1$, $\square_1$) ≠ (0, 1)	DF(IV)	($\square_1$, $\square_1$)= (0, 1)

*Indica que los valores críticos se encuentran en la parte de arriba del Cuadro 8.5.2 en Fuller (1976).
**Indica que los valores críticos se encuentran en el Cuadro VI en Dickey y Fuller (1981). Las otras observaciones en estas columnas deben interpretarse de la misma manera.

III.b. Philips - Perron

Suponga que hay 1,2,□ , T observaciones de una serie y_t y se estima una ecuación de regresión de la forma (Enders, 1995; Phillips y Perron, 1988):

(22) $y_t = □ + □ (t\text{-}T/2) + □\, y_{t-1} + u_t$

donde □, □ y □ son los coeficientes convencionales de mínimos cuadrados ordinarios.

En 1988 Peter Phillips y Pierre Perron derivaron una prueba estadística para los coeficientes de la regresión bajo la hipótesis nula de que los datos son generados por:

(23) $y_t = y_{t-1} + u_t$

donde u_t es tal que E $u_t = 0$.

Según Phillips y Perron, no es un requisito que el término de error no esté correlacionado o sea homogéneo. La prueba Phillips - Perron permite que los errores sean débilmente dependientes y distribuidos heterogéneamente.

El estadístico de Phillips - Perron modifica las pruebas "t" de Dickey - Fuller ya que permite un ajuste que contabiliza la heterogeneidad en el proceso de error. Usando los estadísticos $t_□$, $t_□$ y $t_□$, las pruebas "t" usuales para la hipótesis nula de que □ = 0 , □ = 0 y □= 1, respectivamente, los estadísticos Phillips - Perron son (Enders, 1995):

(24) $Z(t_□) = (S^*/□^*_{T□})t_□ - (T^3 / 43^{0.5} D_x^{0.5} □^*_{T□})(□^{*2}_{T□} - S^{*2})$

(25) $Z(t_□) = (S^*/□^*_{T□})t_□ - (T^3 / 24D_x^{0.5} E_x □^*_{T□})(□^{*2}_{T□} - S^{*2})(T^{-3/2} □y_{t-1})$

(26) $Z(t_□) = (S^*/□^*_{T□})t_□ - (T^3 / 2 D_x^{0.5} □^*_{T□})[T^{-2} □ (y_{t-1} - y_{t-1}^*)^2]^{-1/2} (□^{*2}_{T□} - S^{*2}$
$[(1/2)T^{-3/2} □y_{t-1} - T^{-5/2} □ty_{t-1})$

donde D_x es el determinante de la matriz del regresor (X'X); $E_x = [T^{-6} D_x + (1/2)(T^{-3/2} □y_{t-1})^2]^{1/2}$; S* es el error estándar de la regresión; $□^{*2}_{T□} = T^{-1} □_1^T u_1^2 + 2 T^{-1} □^1_{s=1} □^T_{t=s+1} u_t u_{t-s}$; □ es el número de las autocorrelaciones estimadas.

Obsérvese que S^{*2} y $□^{*2}_{T□}$ son estimadores consistentes de $□_u^2 = \lim_{T\_}$ $E(u_T^2)$ y $□^2 = \lim_{T\_}$ $E(T^{-1}S_T^2)$, donde $S_T = □_{ut}$ y todas las sumatorias corren sobre t.

Para la hipótesis conjunta de □ = 0 y □= 1, se une el estadístico $Z(□_3)$:

(27) $Z(\Phi_3) = (S^*/\sigma^2_{T\ell})\Phi_3 - (1/2\,\sigma^2_{T\ell})(\sigma^2_{T\ell} - S^2)[T(\beta - 1) - (T^6 / 48\, D_x)(\sigma^2_{T\ell} - S^2)]$.

Si no se incluye una tendencia determinística en la ecuación de regresión, la hipótesis de que $\alpha=1$ es probada usando:

(28) $Z(t_{\alpha^*}) = (S/\sigma_{T\ell})t_{\alpha^*} - (1/2\,\sigma_{T\ell})(\sigma^2_{T\ell} - S^2)[T^{-6} \sum (y_{t-1} - y_{-1}^*)^2]^{-1/2}$

donde:

$$Y_{-1} = T^{t-1}\sum_{i=1}^{n} y_{t-i}$$

III.c. Caso de más de una raíz unitaria

Al analizar series trimestrales puede encontrarse más de una raíz unitaria, debido a la presencia de características estacionales. Estas características se deben a que, por ejemplo, las decisiones tomadas por los agentes económicos en determinado trimestre pueden estar correlacionadas con las decisiones tomadas en el mismo trimestre de años anteriores (Rodríguez, 2002). En el caso en que los componentes estacionales de una variable sean estocásticos, existe la probabilidad de la presencia de raíces unitarias tanto en su comportamiento a largo plazo como en los efectos estacionales (Rodríguez, 2002).

La presencia de raíces unitarias regulares y estacionales en las series a estudiar pueden observarse en la prueba HEGY desarrollada por Hylleberg (1990). Esta prueba se basa en que el crecimiento anual de cualquier serie de tiempo con frecuencia trimestral puede ser descompuesta según el siguiente operador de rezagos:

(29) $(1 - L^4) = (1 - L)(1 + L)(1 - iL)(1 + iL)$

El término izquierdo de la ecuación es la diferencia en logaritmos en cuatro periodos. Los términos del lado derecho representan la diferencia anual, la diferencia semestral y, los últimos dos términos, el componente trimestral. La ventaja de usar esta relación estriba en que permite la estimación de una prueba de manera general con la capacidad de evaluar de manera simultánea varias hipótesis sobre la evolución de la variable. La ecuación anterior se puede expresar en su forma general como:

(30) $(1 - L^4) = (1 - \alpha_1 L)(1 + \alpha_2 L)(1 - \alpha_3 iL)(1 + \alpha_4 iL)$

en la cual los alfas representan parámetros. De acuerdo al valor de estos parámetros se puede determinar la frecuencia de las raíces unitarias. En el caso en que α_1 sea igual a uno la variable tiene una raíz unitaria no estacional. Al ser α_2 igual a uno existe una raíz unitaria semestral en la serie. Cuando α_3 y α_4 son iguales a uno la variable tiene una raíz unitaria trimestral.

La implementación de esta prueba requiere definir las siguientes variables:

(31) $$x_{1t} = (1 + L + L^2 + L^3)x_{t-1} = x_{t-1} + x_{t-2} + x_{t-3} + x_{t-4}$$

(32) $$x_{2t} = (1 - L + L^2 - L^3)x_{t-1} = x_{t-1} - x_{t-2} + x_{t-3} - x_{t-4}$$

(33) $$x_{3t} = x_{t-1} - x_{t-3}$$

Estas variables se usan para estimar la siguiente ecuación por MCO:

(34) $$(1 - L^4)x_t = \pi_1 x_{1t-1} - \pi_2 x_{2t-1} + \pi_3 x_{3t-1} - \pi_4 x_{3t-2} + \varepsilon_t$$

En la misma se acepta la hipótesis nula de la presencia de raíces unitarias estacionales y regulares. Si no se rechaza la hipótesis nula de que π_1 es igual a cero existe una raíz unitaria no estacional en la variable. Al no rechazarse la hipótesis nula de que π_2 es igual a cero existe una raíz unitaria semestral en la variable. Sin embargo, si no se puede rechazar la hipótesis nula conjunta de que π_3 y π_4 son iguales a cero existe una raíz unitaria trimestral en la variable

La prueba HEGY se analiza sin intercepto ni tendencia (None); con intercepto (I only); con intercepto y variables mudas estacionales (I,SD), con intercepto y tendencia (I,Tr) y con intercepto, variables mudas estacionales y tendencia (I,SD,Tr). El propósito de este tipo de estimación es para capturar la presencia de estacionalidad no estocástica. Dado que la prueba está basada en la hipótesis de no estacionariedad, su capacidad es limitada para distinguir procesos estacionarios con autocorrelación cercana a uno. Sin embargo, al incluir rezagos de la variable dependiente permite controlar la posible autocorrelación de los residuos. La presencia de una tendencia, variables mudas estacionales y una constante enriquece la hipótesis alternativa, la cual, en ausencia de estos, indica que la serie es un ruido blanco.

IV. Cointegración

Los supuestos econométricos indican que las variables tienen que ser estacionarias. Esto hace que los errores presenten ciertas características específicas. Pero, se ha demostrado que una combinación lineal de dos variables, donde ambas son procesos estocásticos no estacionarios o caminatas aleatorias, podría ser estacionaria. Quiere decir que la relación entre éstas no es espuria, pueden ser estimadas por MCO y no hay que diferenciarlas (evitándose pérdida de información relevante). Es decir, estas variables tienen una relación de cointegración. Por ejemplo, al tener una ecuación lineal de dos variables no estacionarias:

(35) $Y_t = \beta_1 + \beta_2 X_t + u_t$

la cual se puede rescribir como:

(35') $u_t = Y_t - \beta_1 - \beta_2 X_t$

y si se encuentra que la combinación lineal ($Y_t - \beta_1 - \beta_2 X_t$) es estacionaria, entonces se dice que las variables Y_t y X_t están cointegradas (Enders, 1995; Gujarati, 1998; Maddala y Kim, 1998). Intuitivamente se observa que cuando u_t es I(0), las tendencias de ambas variables se eliminan al estar integradas del mismo orden. Al ser una serie Y, I(1) y otra X también I(1), ambas pueden estar cointegradas. En general, si Y es I(d) y X es también I(d), donde d es el mismo valor, estas dos series pueden estar cointegradas. Si éste es el caso, la regresión de las dos variables es significativa (es decir, no es espuria); y no se pierde información valiosa de largo plazo, lo cual sucedería si se utilizaran sus primeras diferencias.

IV.a. Orígenes del concepto de cointegración

El concepto de cointegración fue introducido por Engel y Granger y su análisis formal estriba en que habrá un equilibrio a largo plazo entre un conjunto de variables cuando:

(36) $\beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_n x_{nt} = 0.$

Esta ecuación se puede rescribir como: $\beta = (\beta_1, \beta_2, \dots, \beta_n)$ y $x_t = (x_{1t}, x_{2t}, \dots, x_{nt})'$, y el sistema estará en un equilibrio a largo plazo cuando $\beta x_t = 0$. La desviación del equilibrio a largo plazo se conoce como el error (e_t), así que:

(37) $e_t = \beta x_t.$

Si el equilibrio es significativo en la relación de las variables, entonces el error es estacionario. Engel y Granger proveen la siguiente definición de cointegración: Los componentes del vector $x_t = (x_{1t}, x_{2t}, \square, x_{nt})$ se dicen que están cointegrados de orden d, b, denotado por $x_t \sim CI(d,b)$, en el caso en que:

1. Todos los componentes de x_t son integrados de orden d.
2. Existe un vector $\square = (\square_1, \square_2, \square, \square_n)$ en el cual la combinación linear $\square_1 x_{1t} + \square_2 x_{2t} + \square + \square_n x_{nt}$ es integrada de orden (d-b), donde b>0.

IV.b. Aspectos importantes sobre la cointegración:

1. Si $(\square_1, \square_2, \square, \square_n)$ es un vector de cointegración, cualquier escalar no igual a cero puede ser multiplicado por $(\square_1, \square_2, \square, \square_n)$ y también lo es.
2. Los parámetros de largo plazo ($\square$'s) se analizan al ser normalizados. Para normalizar el vector de cointegración se multiplica la $\square$ cualquiera que sea por un número que haga que sea igual a uno. Es decir, para normalizar el vector de cointegración con respecto a x_1, se selecciona un número ($\square$) tal que: $\square = 1/\square_1$;
3. Todas las variables deben estar integradas del mismo orden. Aunque puede surgir el caso de que, por ejemplo; x_{1t} y x_{2t} son I(2) y x_{3t} es I(1). Si x_{1t} y x_{2t} son C(2,1) existe una combinación lineal de la forma $\square_1 x_{1t} + \square_2 x_{2t}$ que es I(1). Es posible que éstas estén cointegradas con x_{3t} y que la combinación lineal entre $\square_1 x_{1t} + \square_2 x_{2t} + \square_3 x_{3t}$ sea estacionaria.
4. No todas las variables similarmente integradas cointegran. Esa falta de cointegración implica que no hay una relación a largo plazo entre las variables.
5. En el caso en que las variables estén integradas en un orden diferente, éstas no pueden estar cointegradas.
6. El término equilibrio es usado de forma diferentes por los teóricos y por los econometristas. Los teóricos económicos emplean el término al referirse a una igualdad entre las transacciones deseadas y observadas. El uso econométrico del término hace referencia a cualquier relación a largo plazo entre variables no estacionarias.

7. La existencia del fenómeno de cointegración entre las variables de estudio no requiere que la relación a largo plazo sea generada por las fuerzas del mercado o por el comportamiento de los individuos. Esta puede estar dada por causalidad, comportamientos de las variables o simplemente una relación de forma reducida entre éstas con similar tendencia.
8. Si x_t tiene n componentes, debe haber n-1 vectores de cointegración linealmente independientes. El número de vectores cointegrados se llama el rango de cointegración.
9. Mucha de la literatura de cointegración se enfoca en el caso en el cual cada variable es I(1). Pocas variables económicas son integradas de orden mayor de uno.
10. Muchos autores utilizan el término cointegración para referirse al caso cuando las variables son CI(1,1). Pero pueden surgir muchas otras posibilidades. Por ejemplo, un conjunto de variables I(2) pueden estar cointegradas de orden CI(2,1), así que existe una combinación lineal que es I(1).

IV.c. Metodologías de cointegración

IV.c.1. Método de dos pasos Engle-Granger

Partiendo de la teoría económica se obtiene la siguiente relación:

(38) $Y=b'X$; donde Y y X son series I(1)

Paso(1) Mínimos Cuadrados Ordinarios:

(39) $Y_t=b'X_t+e_t$

(40) Residuales de mínimos cuadrados: $e_t= Y_t-b'X_t$

Se realiza prueba de estacionariedad (gráfica + ADF)

Paso(2) Dinámica y modelo de corrección de errores(MCE)

(41) _Y_t=_$_0$+_(L)_X_t+_(L) _Y_t+ _$_1$$e_t$+$v_t$

Simplifique (Wald, Prueba F) + pruebas de diagnóstico en ECM (AIC, SCH, especificación incorrecta)

IV.c.1.a. Puntos importantes

1. No hay dinámica en el modelo (los efectos de X_t sobre Y_t son contemporáneos).
2. Lo que prueba es posible cointegración.
3. La prueba ADF sobre los errores es incorrecta.

Una forma práctica de resolver el problema es con mínimos cuadrados no lineales y realizar cualquier prueba de cointegración que se desee.

IV.c.1.b Dinámica: Modelo de corrección de errores (MCE)

1. MCE analiza (de la relación a largo plazo) la dinámica de corto plazo y cuán rápido es el ajuste al largo plazo ante cualquier desequilibrio provocado por choques no esperados.
2. Esta "dinámica" puede darse por relaciones económicas o por el manejo y problemas de estimación de datos.
3. En el largo plazo, las variables son de órdenes de integración similares (por lo general I(1)) y se espera que exista, por lo menos, un vector de cointegración.
4. Es por todo esto que se debe partir de un modelo general (metodología de lo general a lo específico) el cual:
    a. Contenga fundamento teórico.
    b. Tenga suficientes rezagos para respuestas flexibles de Y_t a X_t.
    c. No exista correlación serial entre los errores.
    d. Reparametrize en un modelo sensible de corrección de errores.

IV.c.1.c Reparametrización

Una ecuación en niveles puede reparametrizarse en diferencias (que resulten en series I(0)) y en el error de cointegración, el cual es una combinación lineal de variables I(1) (la cual se espera que sea una relación de cointegración).

Ecuación en niveles:

(42) $Y_t=a_1Y_{t-1}+a_2X_t+a_3X_{t-1}+\_{t}$

IV.c.1.d. Reparametrización en un MCE

Se resta Y_{t-1} y se suma $(a_2X_{t-1}- a_2X_{t-1})$

(43) $Y_t-Y_{t-1}= a_1Y_{t-1}-Y_{t-1}+ a_2X_t+a_3X_{t-1}+\_{}_t + (a_2X_{t-1}- a_2X_{t-1})$

(44) $\_Y_t= (a_1-1)Y_{t-1}+ a_2 (X_t- X_{t-1})+(a_2+ a_3) X_{t-1}+\_{}_t$

(45) $\_Y_t=\pi_1 Y_{t-1}+ \pi_2 \_X_t+ \pi_3X_{t-1}+\_{}_t$

$\pi_3= -\pi_1(-\pi_3/ \pi_1)$

(46) $\_Y_t=\pi_1 Y_{t-1}+ \pi_2 \_X_t- \pi_1(-\pi_3/ \pi_1)X_{t-1}+\_{}_t$

Reagrupando:

(47) $\_Y_t= \pi_2 \_X_t + \pi_1[Y-(-\pi_3/ \pi_1)X]_{t-1} +\_{}_t$

MCE (estimación por mínimos cuadrados no lineales (MCNL)):

(48) $\_Y_t= \pi_2 \_X_t+ \pi_1[Y-\_X]_{t-1}+\_{}_t$

donde:

$\pi_1= a_1-1$

$\pi_2 = a_2$

$\_ = (-\pi_3/ \pi_1) = -(a_2+ a_3)/ (a_1-1)$

IV.c.1.e. Revisión de la reparametrización

1. No cambia el número de parámetros ni de variables.
2. Los parámetros estimados por MCO deben ser igual que los obtenidos por MCNL.

Ej. $\pi_2 = a_2$

3. $\_{}_t$ será I(0) si las variables están cointegradas.

IV.d. Procedimiento de Johansen

El procedimiento de Johansen permite contrastar simultáneamente el orden de integración de las series, así como la existencia de vectores de cointegración, el cálculo de todos los vectores de cointegración, sin imponer que sólo existe un vector de cointegración, y no verse afectado por las condiciones de endogeneidad de las variables envueltas en la relación de cointegración (Suriñach, Artís, López y Sansó, 1995; Enders, 1995). El procedimiento parte de un modelo de vectores autorregresivos VAR(p) sin restricciones.

(49) $$Y_T = \mu + \Pi_1 Y_{t-1} + \ldots + \Pi_{t-p} + \varepsilon_t$$

donde:

Y_t es un vector columna (mx1); “m” representa el número de variables en el modelo; _ es un vector de constantes; y $\_ _t$ es un vector de errores aleatorios con media cero y varianza constante.

Este modelo de vectores autorregresivos puede rescribirse de la siguiente manera:

(50) $$\Delta Y_t = \mu + \Gamma_1 \Delta Y_{t-1} + \ldots + \Gamma_{p-1} \Delta Y_{t-p+1} + \Pi Y_{t-p} + \varepsilon_t$$

donde:

$$\Gamma_i = -I + \Pi_1 + \ldots + \Pi_i \qquad i = 1,2,\ldots,p-1$$

$$\Pi = -I + \Pi_1 + \ldots + \Pi_p$$

La matriz $\_ _{mxm}$ contiene la información a largo plazo sobre la relación entre las variables. El modelo en (50) en un modelo de corrección de errores (MCE) en forma matricial. Este MCE es muy parecido a la ecuación que se utiliza para la prueba ADF. No obstante, para que exista un equilibrio, el término $\_Y_{t-p}$ en (50) debe ser I(0), implicando que la matriz _ recoge las relaciones de cointegración.

Lo anterior puede verse en detalle considerando un VAR(1):

(51) $x_t=A_1x_{t-1}+\_ _t$

donde:

x_t= un vector nx1;

$\_ _t$ = un vector nx1;

A_1= una matriz de parámetros (nxn).

Restando x_{t-1} a ambos lados de la ecuación se obtiene:

(52) $x_t - x_{t-1} = A_1x_{t-1}-x_{t-1}+\_ _t$

(53) $\_x_t= -(I-A_1)\, x_{t-1}+\_ _t$

(54) $\_x_t = \pi\, x_{t-1} + \_{}_t$

donde:

$\pi = -(I-A_1)$.

Para determinar la existencia de una o varias relaciones de cointegración hay que analizar el rango de la matriz π.

1. Si el rango de la matriz es igual a cero:

(55) $\_x_t = \_{}_t$

quiere decir que las variables son paseos aleatorios independientes y aunque presenten el mismo orden de integración no están cointegradas.

2. Si el rango es completo, la solución a largo plazo se da por n ecuaciones independientes, ya que cada ecuación es una restricción independiente de la solución a largo plazo de las variables. En este caso las variables son estacionarias.

3. En casos intermedios, en los cuales el rango de π = r, existen r vectores de cointegración. En el caso en que r = 1, existe un solo vector de cointegración, dado por cualquier fila de π. Cada fila en la matriz π puede rescribirse como un MCE. La primera fila del modelo es:

(56) $\_x_{1t} = \pi_{11}x_{1t-1} + \pi_{12}x_{2t-2} + \pi_{13}x_{3t-3} \ldots + \pi_{1n}x_{nt-1} + \_{}_{1t.}$

Si de esta fila se obtiene π_{11} como factor común se puede obtener una ecuación de la forma de un MCE:

(57) $\_x_{1t} = \_{}_1(x_{1t-1} + \_{}_{12}x_{2t-2} + \_{}_{13}x_{3t-3} \ldots + \_{}_{1n}x_{nt-1}) + \_{}_{1t}$

donde:

$\_{}_1 = \pi_{11}$

$\_{}_{ij} = \pi_{1j}/\pi_{11.}$

`

Dada la existencia de una relación de cointegración, en el largo plazo, se cumple:

(58) $x_{1t-1} + \_{}_{12}x_{2t-2} + \_{}_{13}x_{3t-3} \ldots + \_{}_{1n}x_{nt-1} = 0$

En esta ecuación el vector normalizado de cointegración es:

(59) $(1, \_{}_{12}, \_{}_{13}, \ldots, \_{}_{1n})$

y la velocidad de ajuste viene dado por $\_{}_1$. Por lo que, en términos generales, π= __'(se demostrará más adelante)

.

IV.d.1. Aplicación

1. Teoría económica

y_t = logaritmo del nivel de salarios

x_t = logaritmo del nivel de precios

En el largo plazo, los salarios son proporcionales al nivel de precios.

(60) $y_t - \_x_t = 0$ y posiblemente que _ = 1 (el salario real es constante).

2. Econometría (respuestas rezagadas)

IV.d.1.a. VAR y VEC[5]

(61) $y_t = \_{11} y_{t-1} + \_{12} x_{t-1} + \_{1t}$

(62) $x_t = \_{21} y_{t-1} + \_{22} x_{t-1} + \_{2t}$

Restando y_{t-1} en la primera ecuación del VAR y x_{t-1} en la segunda:

(63) $y_t - y_{t-1} = \_{11} y_{t-1} - y_{t-1} + \_{12} x_{t-1} + \_{1t}$

(64) $x_t - x_{t-1} = \_{21} y_{t-1} + \_{22} x_{t-1} - x_{t-1} + \_{2t}$

(65) $\_y_t = -(1 - \_{11})\, y_{t-1} + \_{12} x_{t-1} + \_{1t}$

(66) $\_x_t = \_{21} y_{t-1} - (1 - \_{22})\, x_{t-1} + \_{2t}$

(67) $\_y_t = \pi_{11} y_{t-1} + \pi_{12} x_{t-1} + \_{1t}$

(68) $\_x_t = \pi_{21} y_{t-1} + \pi_{22} x_{t-1} + \_{2t}$

donde:

$\pi_{11} = -(1 - \_{11})$

$\pi_{12} = \_{12}$

$\pi_{21} = \_{21}$

$\pi_{22} = -(1 - \_{22})$

[5] VEC es un modelo de vectores autorregrevivos, en el cual cada ecuación es un MCE. Se conoce como "Vector Error Correction Model" y es equivalente a la ecuación (50).

IV.d.1.b. Matriz de parámetros _

$$\Pi = \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix}$$

IV.d.1.c. Prueba de la traza del procedimiento de Johansen y obtención de los parámetros de largo plazo y de la velocidad de ajuste mediante máxima verosimilitud

Mediante máxima verosimilitud, se genera una prueba para el rango de esta matriz (=r) a través la prueba de la traza. El estadístico es distribuido $\_^2$.[6] Además, el procedimiento de Johansen estima los parámetros de largo plazo _, así como los de la velocidad de ajuste _ y la matriz de varianza-covarianza _.

A continuación se presenta el caso más sencillo en el cual no se incluyen términos deterministas. Este se basa en los siguientes supuestos:

4. La ecuación (50) es el PGD
5. El vector de constantes es cero (_=0)
6. $Y_{i-p} \ldots Y_0$ son dados
7. $\_{}_t \sim N(0,\_^2)$

La estimación de máxima verosimilitud de _ puede obtenerse a través de los siguientes pasos:

a. Realice las siguientes dos ecuaciones por medio de mínimos cuadrados ordinarios:

(69) $$\Delta Y_t = \Pi_{01}\Delta Y_{t-1},\cdots,\Pi_{0p-1}\Delta Y_{t-p+1} + R_{ot}$$

(70) $$Y_{t-p} = \Pi_{11}\Delta Y_{t-1},\cdots,\Pi_{1p-1}\Delta Y_{1-p+1} + R_{pt}$$

2. Mediante R_{ot} y R_{pt} genere el producto residual S_{ij} (mxm):

(71) $$S_{ij} = (1/T)\sum_{t=1}^{n} R_{it}R_{jt}^{'}$$

donde: i,j=0,p

[6] Sin embargo, como se verá más adelante, los valores críticos no son estándar.

Por ejemplo:

(82) $$S_{oo} = 1/n\sum_{t=1}^{n} R_{ot}R_{ot}^{'}$$

obteniendo así S_{pp}, S_{oo}, S_{po}

3. Según Johansen (1988), la estimación máximo verosímil de la matriz de vectores de cointegración (_) se obtiene al imponerse la restricción: _'S_{pp}_=I, la cual se obtiene a partir de los valores propios de ($S_{p0}S_{00}^{-1}S_{0p}$) respecto a S_{pp}[7]. Quiere decir, los valores propios, $\__{i}$=1,…m, tal que:

(83) $$\left|\lambda S_{pp} - S_{po}(S_{oo})^{-1}S_{op}\right| = 0$$

Como S_{pp} es una matriz simétrica y definida positiva para T finito, puede rescribirse como ___', donde _ es una matriz diagonal con los valores propios de S_{pp} y _ es una matriz ortogonal la cual tiene por columnas los vectores propios estandarizados de S_{pp}. Por lo tanto:

(84) $$S_{pp}^{-1} = \square\square^{-1}\square' = PP'$$

donde:

$$P = \square\square^{-0.5}$$

$$P'S_{pp}P = I$$

En el caso que P sea no singular:

(85) $$P'(\lambda S_{pp} - S_{p0}S_{00}^{-1}S_{0p})P = (\lambda I - P'S_{p0}S_{00}^{-1}S_{0p}P)$$

dado que el determinante de un producto es el producto de los determinantes:

(86) $$|P'| = \left|\lambda S_{pp} - S_{p0}S_{00}^{-1}S_{0p}\right||P| = \left|\lambda I - P'S_{p0}S_{00}^{-1}S_{0p}P\right|$$

Por lo tanto, la solución de (83) puede obtenerse calculando los valores propios dado:

(87) $$|\lambda I - A| = 0 \qquad con \qquad A = P'S_{p0}S_{00}^{-1}S_{0p}P$$

[7] El problema de maximizar la función de verosimilitud concentrada en _ se reduce a encontrar las correlaciones canónicas entre $\_Y_t$ y $\_Y_{t-p}$ las cuales son corregidas por los efectos de la presencia de los valores rezagados de $\_Y_t$ (Suriñach, Artís, López y Sansó, 1995; Patterson, 2000). Las relaciones canónicas entre R_0 y R_p. Es decir, es un problema de valores propios (Suriñach, Artís, López y Sansó, 1995; Bhaskara, 1994).

Una vez se obtienen los valores propios se ordenan de mayor a menor. De forma tal que: _1>_2>…_n

4. Para corroborar la hipótesis nula de que existe como máximo r vectores de cointegración frente a que existen m (r≤m), la prueba de razón de verosimilitud viene dada por:

$$(88) \qquad -2\ln Q = -T\sum_{i=r+1}^{m}(1-\lambda_i)$$

la cual sigue una distribución que puede aproximarse por c_^{2}(f) donde c=0.85-0.58/f, y _^{2}(f) tiene una distribución _^{2} con f=2(m-r)^2 grados de libertad (Johansen, 1988; Suriñach, Artís, López y Sansó, 1995; Bhaskara, 1994). Este estadístico se conoce como el estadístico de la traza (Johansen, 1998; Suriñach, Artís, López y Sansó, 1995; Patterson, 2000; Bhaskara, 1994; Enders, 1995).

El siguiente estadístico es uno alternativo para corroborar la significancia del r-ésimo valor propio mayor (_r) (Johansen, 1998; Suriñach, Artís, López y Sansó, 1995; Patterson, 2000; Bhaskara, 1994):

$$(89) \qquad \lambda_r^{max} = -T\ln(1-\lambda_r)$$

Los valores críticos de ambos estadísticos se consiguen en los libros especializados de series de tiempo. Cabe señalar que las distribuciones de los estadísticos son función de la cantidad de relaciones de cointegración.

5. Una vez corroborado el rango de cointegración se obtendrá una estimación de los parámetros de largo plazo, con los valores característicos _1>_2>_3…_m y los siguientes vectores característicos:

$$V^{\square} = (V_1^{\square}, V_2^{\square}, \ldots, V_m^{\square})$$

Los vectores característicos son normalizados usando:

$$(90) \qquad V' S_{pp} V = I_m$$

Los estimadores de máxima verosimilitud pueden obtenerse directamente del problema de los valores característicos y son:

(91) $\underset{(mxr)}{\square^{\square}} = \left\{ V_1^{\square}, V_2^{\square} \ldots V_m^{\square} \right\}$

(92) $\underset{(mxr)}{\square^{\square}} = S_{op}\square^{\square}$

(93) $\hat{\square} = S_{00} \; \square\square\square\square$

Es decir, de acuerdo a (91), las columnas de □ son los vectores propios asociados a cada valor propio. De esta manera, la matriz de □'s se estima a partir de:

(94) $\left(\square S_{pp} \; \square \; S_{p0} S_{00}^{\square 1} S_{0p}\right)\square = 0$

En este caso lo que se intenta es encontrar las combinaciones lineales del vector Y_t que están correlacionados al máximo con $\_Y_t$.

IV.d.1.c.1. Posibles resultados de la prueba de la traza del procedimiento de Johansen (Punto 4)

1. Si el rango de _ es igual a cero ambas variables son paseos aleatorios independientes y, a pesar de que sean variables con el mismo orden de integración, no están cointegradas.

8. Esto se puede ver analizando la existencia de paseos aleatorios en las series (raíces unitarias en el polinomio autorregresivo que determina el PGD);
9. Examinando, mediante la prueba de la traza del procedimiento de Johansen, que todos los elementos de la matriz _ sean iguales a cero.

2. Si el rango de _ es igual a 2 ambas variables son estacionarias.
    a. La ecuación en el MCE está compuesta por series I(0).
    b. Examinando, mediante la prueba de la traza del procedimiento de Johansen, que todos los elementos de la matriz _ no son iguales a cero.
    c. Si el rango de cointegración es igual 1, existe un solo vector de cointegración entre ambas variables y, en el largo plazo, se cumple: $y_t - \_x_t=0$. Quiere decir que ambas variables están cointegradas.

IV.d.1.c.1.a. Demostración del punto c

El punto c es el más difícil de demostrar

(95) $\_y_t = \pi_{11}y_{t-1} + \pi_{12}x_{t-1} + \_{}_{1t}$

(96) $\_x_t = \pi_{21}y_{t-1} + \pi_{22}x_{t-1} + \_{}_{2t}$

(97) $\_y_t = \_{}_{11}[y + (\pi_{12}/ \pi_{11})x\,]_{t-1} + \_{}_{1t}$

(98) $\_y_t = \_{}_{11}[y + (\pi_{22}/ \pi_{21})x\,]_{t-1} + \_{}_{2t}$

IV.d.1.c.1.b Prueba de la existencia de un solo vector de cointegración

(99) $(\pi_{12}/ \pi_{11}) = (\pi_{22}/ \pi_{21})$

(100) $\pi_{12} = (\pi_{11}\pi_{22}/ \pi_{21})$

Sustituye π_{12} en la matriz _:

$$\square = \begin{bmatrix} \square_{11} & \square_{11}\square_{22} / \square_{21} \\ \square_{21} & \square_{22} \end{bmatrix}$$

$\_\_\ \_ = \pi_{11}\, \pi_{22} - (\pi_{11}\pi_{22} / \pi_{21})\, \pi_{21} = 0$; rank(_) = 1

Quiere decir que cuando el rank(_) = 1, la relación entre y y x es única.

IV.d.1.c.1.c Demostración del VEC dada la existencia de un solo vector de cointegración

Puede determinarse el VEC con la relación de cointegración entre y y x de la forma: y – _x = 0

Sustituye π_{12}:

(101) $\_y_t = \pi_{11}y_{t-1} + (\pi_{11}\pi_{22} / \pi_{21})x_{t-1} + \_{}_{1t}$

(102) $\_x_t = \pi_{21}y_{t-1} + \pi_{22}x_{t-1} + \_{}_{2t}$

(103) $\_y_t = \pi_{11}[y + (\pi_{22} / \pi_{21})x]_{t-1} + \_{}_{1t}$

(104) $\_x_t = \pi_{21}[y + (\pi_{22}/\pi_{21})x]_{t-1} + \_{}_{2t}$

$\_ = (\pi_{22}/\pi_{21})$

(105) $\_y_t = \pi_{11}[y + \_x]_{t-1} + \_{}_{1t}$

(106) $\_x_t = \pi_{21}[y + \_x]_{t-1} + \_{}_{2t}$

IV.d.1.c.1.d. _ = __' , con un solo vector de cointegración

Demostración:

$$\alpha = \begin{bmatrix} \pi_{11} \\ \pi_{21} \end{bmatrix}$$

$$2x1$$

$$\beta' = \begin{bmatrix} 1 & (\pi_{22} / \pi_{21}) \end{bmatrix}$$

$$1x2$$

$$\alpha\beta' = \begin{bmatrix} \pi_{11} \\ \pi_{21} \end{bmatrix} \begin{bmatrix} 1 & (\pi_{22} / \pi_{21}) \end{bmatrix}$$

$$\Pi = \begin{bmatrix} \pi_{11} & \pi_{11}\pi_{22} / \pi_{21} \\ \pi_{21} & \pi_{22} \end{bmatrix}$$

IV.d.1.c.1.e. Consecuencias de la existencia de una relación de cointegración en el VAR

1. Existencia de parámetros estables y errores "ruido blanco".
2. Los estimadores son eficientes.
3. Si y y x están cointegradas, el resto de las pruebas son estándar.

IV.d.1.c.1.f. Problemas

A base de simulaciones de Monte Carlo, los siguientes puntos pueden generar problemas en los resultados de las pruebas:

1. Problemas en la selección de variables y de rezagos
2. Problemas de especificación
3. La presencia de heterocedasticidad y no-normalidad

También existe otro problema significativo el cual es asociado a la existencia de múltiples vectores de cointegración. Por ejemplo, suponga que tiene 3 variables: salarios, precios y bienes importados, y que el rank(_) = 2. Entonces, existen dos vectores de cointegración diferentes. En este caso, existe un problema de interpretación: cualquier combinación lineal de vectores pueden estar cointegradas.

IV.d.1.c.1.g. Restricciones a los vectores de cointegración

Dada la posibilidad de la existencia de varios vectores de cointegración, una forma para identificar si una o varias de las ecuaciones que se quieren estudiar en el sistema de ecuaciones tiene una relación estable en el largo plazo es imponiendo la misma restricción lineal a cada vector de cointegración. Por ejemplo, suponga que existen cuatro variables (k=4) y tres relaciones de equilibrio (r=3). En esta caso, la matriz que expresa las relaciones de equilibrio puede expresarse como:

$$\square = \begin{bmatrix} \square_{1,1} & \square_{2,1} & \square_{3,1} \\ \square_{1,2} & \square_{2,2} & \square_{3,2} \\ \square_{1,3} & \square_{2,3} & \square_{3,3} \\ \square_{1,4} & \square_{2,4} & \square_{3,4} \end{bmatrix}$$

de tal forma que las tres relaciones de equilibrio están dadas por:

(107)
$$\begin{aligned} &\square_{1,1} y_{1t} + \square_{1,2} y_{2t} + \square_{1,3} y_{3t} + \square_{1,4} y_{4t} = \\ &\square_{2,1} y_{1t} + \square_{2,2} y_{2t} + \square_{2,3} y_{3t} + \square_{2,4} y_{4t} = \\ &\square_{3,1} y_{1t} + \square_{3,2} y_{2t} + \square_{3,3} y_{3t} + \square_{3,4} y_{4t} = 0 \end{aligned}$$

Por ejemplo, suponga que se quiere evaluar la validez de un coeficiente unitario en y_2 en una relación normalizada de tal forma que $y_1$aparezca con un coeficiente igual a uno. Por tanto, se desea considerar si es aceptable una relación de equilibrio de la forma:

(108)
$$y_{1t} = \square_1 y_{2t} + \square_2 y_{3t} + \square_3 y_{4t}$$

para la cual _=1. No obstante, siempre es posible seleccionar una combinación lineal de las tres columnas de la matriz anterior la cual resulta una relación de equilibrio de este tipo. Esto requiere que se imponga una restricción conveniente en cada uno de los tres vectores de cointegración. La restricción establecerá que el coeficiente multiplicado por y_2 igual al negativo del coeficiente multiplicado por y_1. Esta restricción, impuesta en cada vector de cointegración, resulta en un total de tres restricciones:

(109)
$$\square_{1,1} = \square\square_{1,2} \qquad \square_{2,1} = \square\square_{1,2} \qquad \square_{3,1} = \square\square_{1,2}$$

Estas restricciones implicarán que cualquier vector de cointegración tiene la propiedad de que el coeficiente multiplicado por y_1 es igual al negativo del coeficiente multiplicado por y_2, de tal forma que todas las relaciones de cointegración tienen la forma de (109) con _=1, cuando se normaliza de forma tal que el coeficiente de y_1 es uno.

La prueba de máxima verosimilitud, dado que la misma restricción es impuesta en cada uno de los vectores de cointegración, tendrá tres grados de libertad. En términos generales, si un conjunto de q restricciones lineales son impuestas en cada uno de los r vectores de cointegración, para estar seguros de que cada vector de cointegración satisface las mismas q restricciones lineales, entonces la prueba de máxima verosimilitud tendrá qr grados de libertad.

Sin embargo, se puede también probar la hipótesis envolviendo la especificación completa del número de vectores de cointegración o la imposición de restricciones en un solo vector de cointegración. El requisito de que un vector de cointegración estimado sea de la forma propuesta puede ser impuesto en la estimación y se obtendrá un estadístico de máxima verosimilitud para contrastar la validez de imponer la restricción.

V. Comentarios finales

El debate teórico y metodológico en la economía no es reciente. A partir de la Segunda Guerra Mundial, la economía ha pasado por la revolución keynesiana, neoclásica, monetarista y nueva clásica. Uno de los efectos mayores en cada una de estas revoluciones es el énfasis en cambiar la forma de modelar los fenómenos económicos. Esto genera una gran controversia entre economistas, ya que surgen nuevas propuestas de política económica y se sugiere abandonar otras por, supuestamente, no ser las apropiadas.

Los modelos estructurales basados en la estimación de ecuaciones simultáneas han sido de gran tradición en la econometría y fueron de los primeros en utilizarse. Sin embargo, dados los problemas presentados por estos modelos a principios de los setenta y las críticas contundentes recibidas por diferentes escuelas de pensamiento, en especial la nueva escuela clásica, surgen diversas metodologías que intentan resolver los problemas asociados a este tipo de modelos. Se pueden mencionar los modelos uniecuacionales de series de tiempo basados en ecuaciones en diferencias. Luego, a principio de los ochenta, surgen los modelos multiecuacionales de series de tiempo, basados en un sistema de ecuaciones en diferencias.

Sin embargo, el análisis de series de tiempo señalaba que las variables tienen que ser estacionarias. Pero, la mayoría de las variables económicas no son estacionarias ya que tanto su media, su varianza o ambas son función del tiempo. A principios, se propuso que para eliminar este problema se debe diferenciar las series. El problema con esto es que se pierde información relevante. Por lo tanto, el análisis de estacionariedad se convierte en pieza clave en el análisis de series de tiempo económicas.

Lo anterior motiva a un gran grupo de matemáticos, estadísticos y economistas a desarrollar pruebas para contrastar la presencia de una raíz unitaria en el proceso autorregresivo de la serie, el cual se define como el proceso generador de datos (PGD). Este tiene forma de los modelos uniecuacionales de series de tiempo tradicionales.

Una de las primeras pruebas que, a pesar de las críticas recibidas es la más usada e importante, es la prueba DF y en una versión aumentada, la prueba ADF. Además de DF y ADF, existen un sinnúmero de pruebas las cuales intentan corroborar la existencia de una o varias raíces unitarias. Entre estas, las más utilizadas son la prueba PP y la prueba HEGY, la cual se utiliza para corroborar la existencia de más de una raíz unitaria en series de periodicidad menor a un año.

La contrastación de la existencia de una raíz unitaria toma gran importancia en un contexto multivariable, dada la posibilidad de relaciones espurias. Sin embargo, en 1987, Engle y Granger probaron que dos series no estacionarias pueden estar relacionadas en el largo plazo si éstas están cointegradas y cualquier desviación en el corto plazo será transitoria debido a la relación estable entre las series. Este concepto de cointegración, junto al de raíces unitarias, genera una nueva revolución en términos de los métodos de estimación de las relaciones económicas, así como de las fluctuaciones de la actividad económica. Actualmente, los debates respecto a la significancia de esta nueva revolución se llevan en un marco positivo. Por lo que esta nueva revolución ha atraído la atención de econometristas aplicados, matemáticos especializados en ecuaciones en diferencias y diferenciales, estadísticos, economistas teóricos y economistas orientados en el análisis de política económica.

VI. Bibliografía

Bhargara, A. (1986). "On the Theory of Testing Unit Roots in Observed Time Series." Review of Economic Studies, núm. 53., pp. 369-384.

Bhaskara Rao, B. (1994). Cointegration: for the Applied Economist. Palgrave: Macmillan, pp. 1-85.

Charemza, W. y Deadman, D.F. (1993). "New Directions in Econometric Practice." General to Specific Modeling, Cointegration and Vector Autorregresion, editor Edward Elgar. Publicación limitada, University Press, Cambridge.

Dickey, D. y Fuller, W.A. (1981). "Likelihood Ratio Statistics for Autorregresive Time Series with a Unit Root". Econometrica, núm, 49, pp.427-431.

Enders, W. (1995). Applied Econometrics Time Series, pp.63-128, 294-319. New York, Wiley.

Fuller, W. (1976). Introduction to Statistical Time Series. John Wiley & Sons.

Hylleberg, S, Engle, R.F., Granger, C.W.J. y Yoo, B.S. (1990). "Seasonal Integration and Co-Integration". Journal of Econometrics, núm. 44, pp. 215-228.

Galindo, L.M. (1995). "La metodología econométrica moderna: una versión aplicada". Economía moderna aplicada: Cuadernos de trabajo, núm. 18. UACPYP, CCH, UNAM, pp. 1-30.

________. (1997). "El concepto de exogeneidad en la econometría moderna." Investigación económica, núm. 220, pp. 97-111.

Galindo, L.M. y Cardero, M.E. (1998), "Modelo de vectores autorregresivos con cointegración para la economía mexicana: 1980-1996." Economía Mexicana, vol. VI, núm. 2, pp. 223-246.

Gujarati, D. (1998). Econometría. Mc Graw Hill, pp. 694-716.

Johnston, J. y DiNardio, J. (1997). Econometric Methods. McGraw – Hill, pp. 57-59.

Johansen, S. (1988). "Statistical Analysis of Cointegration Vectors." Journal of Economic Dynamic and Control, núm. 12, pp. 231-54.

Landreth, H. y Colander, D.C. (1998). Historia del Pensamiento Económico. Compañía Editorial Continental, SA. DE C.V. México, pp. 478-505.

Lucas, R.E. (1972). “Expectations and the Neutrality of Money.” Journal of Economic Theory, núm. 4, pp. 103-124.

Maddala, G.S. (1996). Introducción a la Econometría. Prentice Hall, pp.653-582.

Maddala, G.S. y Kim, I.M. (1998). Unit Roots, Cointegration and Structural Change. Cambridge University Press, pp. 15, 50.

Mankiw, G. (1990). “A Quick Refreshner Course in Macroeconomics.” Journal of Economic Literature, vol. XXVIII, pp. 1645-1660.

Nelson, C. y Plosser, C. (1988). “Trends and Random Walks in Macroeconomics Time Series: Some Evidence and implications”. Journal of Monetary Economics, núm. 10, pp. 130-162.

Novales, A. (1993). Econometría. McGraw-Hill, 477-528.

Pankratz, A. (1995). Forecasting With Univariate Box-Jenkins Models: Concepts and Cases. John Wiley & Sons,Inc, .

Patterson, K. (2000). An Introduction to Applied Econometrics: A Time Series Approach. St. Martin’s Press, pp. 208-372, 597-693.

Phillips, P. y Perron, P. (1988). “Testing for Unit Root in Time Series Regression”. Biometrica, núm. 75, pp. 335-346.

Pyndick, R. y Rubinfeld, D. (1997). Econometric Models and Economic Forecasts. Mc Graw Hill, pp. 514-540.

Rodríguez, C. (2002). “Análisis dinámico de la economía de Puerto Rico con un modelo de vectores autorregresivos y cointegración.” Serie de Ensayos y Monografías. Unidad de Investigaciones Económicas, Universidad de Puerto Rico, núm. 110.

Rodríguez, C. (2001). La hipótesis de la neutralidad del dinero en México: un análisis de series de tiempo para el período, 1980-1994. Tesis sometida para obtener el grado de Doctor en Economía. Universidad Nacional Autónoma de México.

Rodríguez, C. y Luciano, I. (2001) “La propensión marginal al consumo en Puerto Rico (1954 -1995): un análisis de cointegración.” Carta de políticas públicas, núm.16, FE-UNAM.

Sims, C. (1980). “Macroeconomics and Reality.” Econometrica, vol. 48, núm. 1, pp.1-48.

Surinach, J., Artís, M. López, E. y Sansó, A. (1995). Análisis económico regional: nociones básicas de la teoría de cointegración. Primera edición. Antoni Bosh. Barcelona, España.

---

* Catedrático Auxiliar, Departamento de Economía, Facultad de Ciencias Sociales, Recinto de Río Piedras, Universidad de Puerto Rico.