

# **ECONOMETRÍA PRÁCTICA: Fundamentos de Series de Tiempo**

**Versión Preliminar**

**Agradecemos de Antemano Comentarios**

Ramón A. Castillo Ponce

Rogelio Varela Llamas

# ECONOMETRÍA PRÁCTICA: Fundamentos de Series de Tiempo

## INDICE

Introducción .....	7
<b>Parte I: Modelos Univariados .....</b>	<b>10</b>
Capítulo 1: Modelos ARMA .....	10
1.1 . Preliminares .....	10
1.2 . Modelos Básicos .....	12
1.3 . Relación entre los Modelos AR y MA .....	17
1.4 . Funciones de Autocovarianza y Autocorrelación .....	19
1.5 . Sugerencia de Procedimiento Práctico .....	23
Capítulo 2: Estacionariedad .....	34
2.1. Concepto .....	34
2.2 . Identificación de Estacionariedad: Primera Aproximación .....	39
2.3 . Ejemplo Numérico: el Producto Interno Bruto de México .....	44
Capítulo 3: Raíces Unitarias .....	50
3.1 . Concepto .....	50
3.2 . Pruebas Básicas de Raíz Unitaria .....	54
3.3 . Pruebas de Raíz Unitaria sobre Series que Presentan Corte Estructural .....	57
3.4. Ejercicio práctico de Raíces Unitarias .....	59
Capítulo 4: Series Integradas y los Modelos ARMA .....	66
4.1 . Definición .....	66
4.2 . Tratamiento de la Estacionalidad en los Modelos ARIMA .....	67
4.3 . Ejemplo para Modelos ARIMA: Índice e Volumen Físico de la Producción Manufacturera .....	70
Capítulo 5: Temas Selectos de Modelos Univariados .....	91
5.1 . Criterios para Evaluar la Capacidad Predictiva .....	91
5.2 . Promedio Simple y Promedio Móvil .....	93
5.3 . Suavizamiento Exponencial Simple y Doble .....	97
5.4 . Suavizamiento Exponencial Lineal Ajustado a la Tendencia .....	103
5.5 . Suavizamiento Exponencial para Tendencia y Estacionalidad Multiplicativa .....	106
5.6. Filtro Hodrick-Prescott .....	109
5.7. La Tendencia y Estacionalidad en los Modelos Clásicos de Regresión .....	112

<b>Parte II: Modelos Multivariados .....</b>	<b>117</b>
Capítulo 6: Introducción a los Sistemas Multivariados .....	117
6.1. Fundamentos .....	117
Capítulo 7: Cointegración .....	126
7.1. Concepto .....	126
7.2 . Metodologías de Cointegración .....	128
Capítulo 8: Introducción al Análisis VAR's y la Metodología de Cointegración de Johansen .....	159
8.1. Fundamentos .....	159
8.2 . Metodología de Johansen .....	164
8.3. Ejercicio Práctico: Función de Consumo .....	168
Capítulo 9: Temas Selectos en el Análisis de VAR's .....	179
9.1. Causalidad Según Granger .....	179
9.2. Función Impulso-Respuesta .....	186
9.3. Descomposición de la varianza .....	190
9.4. Descomposición tendencia – ciclo .....	193
Bibliografía .....	204

## Parte I: Modelos Univariados

### Capítulo 1: Modelos ARMA

En esta primera parte del libro el análisis se centra en los modelos univariados de series de tiempo, que de acuerdo con Box y Jenkins (1970), son representaciones muestrales de un proceso estocástico. Estos modelos presentan la característica de que la variable objeto de estudio depende de su propia evolución histórica. En el presente capítulo se puntualizarán conceptos básicos en la construcción de modelos de series de tiempo y se introducirán las especificaciones univariadas más comúnmente estudiadas.

#### 1.1. Preliminares

Definimos una serie de tiempo como un conjunto de observaciones repetidas de la misma variable tal como:

$$\{y_1, y_2, \dots, y_T\}$$

Donde los sufijos representan el periodo en el tiempo en el que se observa la variable  $y_t$ .

A partir de esta caracterización se desprenden dos conceptos fundamentales en el análisis de series de tiempo. Los mismos se definen mediante la imposición de algunos supuestos sobre el comportamiento de  $y_t$ .

1. Ruido blanco: el proceso  $y_t$  se define como ruido blanco si cumple con las siguientes condiciones:

- a) La esperanza de  $y_t$  es igual a 0 para todos los periodos  $t$ . Esto es  $E(y_t) = 0$  para toda  $t$ .
- b) La varianza de  $y_t$  es constante y por consiguiente independiente del tiempo. Esto es  $Var(y_t) = \sigma^2$ .
- c) Las autocorrelaciones son iguales a 0.

2. Independiente e idénticamente distribuido (iid): el proceso  $y_t$  se dice ser iid si cumple con las siguientes condiciones:

- a) La esperanza de  $y_t$  es constante pero no necesariamente igual a 0 para todos los periodos  $t$ . Esto es  $E(y_t) = \mu$ .
- b) La varianza de  $y_t$  es constante y por consiguiente independiente del tiempo. Esto es  $Var(y_t) = \sigma^2$ .
- c)  $y_t$  es independiente de  $y_k$  para todas las  $t$ 's y  $k$ 's con  $t \neq k$ .

Otro concepto útil en el entendimiento del comportamiento de series de tiempo se refiere al operador de rezagos. El mismo es básicamente una representación abreviada de la dinámica de una serie. Denotemos la observación de  $y_t$  en el periodo anterior como  $y_{t-1}$ , de tal forma que tenemos las siguientes relaciones:

$$Ly_t = y_{t-1}$$

$$L^2 y_t = LLy_t = Ly_{t-1} = y_{t-2}$$

Y así sucesivamente, de tal forma que  $L^j y_t = y_{t-j}$ .

Finalmente, definimos el operador de diferencia  $\Delta$ .

$$\Delta y_t = y_t - y_{t-1} = (1-L)y_t$$

$$\Delta y_{t-1} = y_{t-1} - y_{t-2} = L(1-L)y_t.$$

Y así sucesivamente.

Haremos referencia a estos conceptos ampliamente a través del texto como piezas fundamentales en la construcción de modelos de series de tiempo.

## **1.2. Modelos Básicos**

La teoría econométrica establece que a partir de una muestra se puede encontrar el proceso estocástico generador (PEG) de la misma, lo cual implica que se debe seleccionar una serie y subsecuentemente modelarla. Existen varios modelos que pudieran representar el PEG de una serie de tiempo; aquí nos concentraremos en tres que son fundamentales en la práctica: un proceso autorregresivo (AR), de media móvil (MA), y autorregresivo de media móvil (ARMA).

## 1. Proceso Autorregresivo

Un proceso autorregresivo de primer orden AR(1), se define como:

$$y_t = \mu + \lambda y_{t-1} + \varepsilon_t \quad (1.1)$$

Donde  $\mu$  es una constante,  $\varepsilon_t$  es ruido blanco, e  $y_t$  sigue un proceso autorregresivo el cual supone que su esperanza condicional depende de su valor en el período pasado, es decir,  $E(y_t | y_{t-1}) = \mu + \lambda y_{t-1}$ . Así, un modelo autorregresivo de orden  $p$ , AR( $p$ ), se puede generalizar como:

$$y_t = \mu + \lambda_1 y_{t-1} + \lambda_2 y_{t-2} + \dots + \lambda_p y_{t-p} + \varepsilon_t \quad (1.2)$$

Donde su esperanza condicional y varianza están dadas por las expresiones

$$E(y_t | y_{t-1}, y_{t-2}, \dots, y_{t-p}) = \mu + \lambda_1 y_{t-1} + \lambda_2 y_{t-2} + \dots + \lambda_p y_{t-p}$$

$$\text{Var}(y_t | y_{t-1}, y_{t-2}, \dots, y_{t-p}) = \sigma_\varepsilon^2.$$

Ahora bien, haciendo uso del operador de rezago, la ecuación (1.1) puede expresarse como

$$y_t - \lambda y_{t-1} = \mu + \varepsilon_t = y_t - \lambda L y_t \quad \text{que a su vez es igual a:}$$

$$(1 - \lambda L) y_t = \mu + \varepsilon_t \quad (1.3)$$

De tal forma que el modelo AR( $p$ ) definido en (1.2), se puede replantear como

$y_t - \lambda_1 y_{t-1} - \lambda_2 y_{t-2} - \dots - \lambda_p y_{t-p} = \mu + \varepsilon_t$  que se simplifica como:

$$(1 - \lambda_1 L - \lambda_2 L^2 - \dots - \lambda_p L^p) y_t = \mu + \varepsilon_t \quad (1.4)$$

Donde  $(1 - \lambda_1 L - \lambda_2 L^2 - \dots - \lambda_p L^p)$  es un polinomio en el operador de rezago que permite plantear el modelo AR ( $p$ ) como  $\lambda(L) = \mu + \varepsilon_t$ . Aquí suponemos entonces que la variable de interés depende de la dinámica de sus observaciones pasadas. Como lo abordaremos adelante, identificar la magnitud de  $\lambda$  será fundamental para modelar apropiadamente el comportamiento de la serie. En particular, nótese que si  $\lambda = 1$  entonces el comportamiento contemporáneo de  $y_t$  está determinado por lo que se observó en el periodo  $t - 1$ . En tal caso, la serie no converge y se desplaza en el tiempo sin acotamiento.

## 2. Media Móvil

En este modelo  $y_t$  se expresa en términos de una constante y un proceso de media móvil.

Se dice que es de orden uno, MA(1), cuando se tiene una especificación como la siguiente:

$$y_t = \mu + \varepsilon_t - \theta \varepsilon_{t-1} \quad (1.5)$$

Donde  $\mu$  es la constante y  $\varepsilon_t - \theta \varepsilon_{t-1}$  el proceso de media móvil. Consecuentemente se tiene un proceso MA( $q$ ) cuando  $y_t$  se plantea de la siguiente forma general:

$$y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (1.6)$$

Si se toma la esperanza matemática en ambos miembros de la expresión (1.5) se tiene que

$$E(y_t) = E(\mu + \varepsilon_t - \theta \varepsilon_{t-1}) = \mu$$

y la varianza está dada por

$$Var(y_t) = \sigma_\varepsilon^2 (1 + \theta^2)$$

Haciendo uso del operador de diferencia, el MA(1) y MA( $q$ ) se pueden representar respectivamente como:

$$y_t = (1 - \theta L) \varepsilon_t \quad (1.7)$$

$$y_t = (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) \varepsilon_t = \theta(L) \varepsilon_t \quad (1.8)$$

Donde  $\theta(L) = (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q)$ . Como se mostrará adelante, los modelos MA son comúnmente empleados para realizar pronósticos o eliminar el patrón estacional de una serie.

### 3. Proceso Autorregresivo con Media Móvil

A partir de las consideraciones anteriores, un modelo mixto, autorregresivo de promedio (media) móvil, ARMA  $(p,q)$ , se representa como:

$$y_t = \alpha + \lambda_1 y_{t-1} + \lambda_2 y_{t-2} + \dots + \lambda_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (1.9)$$

En esta ecuación, como se puede ver, aparecen  $p$  términos autorregresivos que representan rezagos distribuidos de la variable dependiente y  $q$  rezagos de la variable de perturbación,  $\varepsilon_t$ , que corresponden a los promedios móviles y que se denominan “innovaciones”. Se le conoce así a  $\varepsilon_t$  por que recoge nueva información que se va conociendo cada periodo. En particular, nótese que el término de error que se define exhibe la siguiente forma:  $\varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$ .

Podemos especificar una representación formal de los 3 modelos mediante el operador de rezagos de la siguiente forma. En definitiva, nótese que los modelos ARMA $(p,q)$  involucran a los modelos AR $(p)$  y MA $(q)$  a través de una sola especificación. El modelo AR $(p)$  se corresponde con un modelo ARMA  $(p,0)$  y un modelo MA $(q)$  es idénticamente igual a un ARMA  $(0,q)$ .

### 1.3. Relación entre los Modelos AR y MA

Es posible mostrar que los modelos anteriormente introducidos presentan relaciones analíticas entre ellos.

#### 1. AR como MA

Tomando (1.1) y utilizando el operador de rezagos, se tiene  $(1 - \lambda L)y_t = \mu + \varepsilon_t$ .

Suponiendo que  $|\lambda| < 1$  se deriva que  $y_t = \frac{\mu}{1 - \lambda L} + \frac{\varepsilon_t}{1 - \lambda L}$  que a su vez es equivalente a la siguiente expresión redefinida,

$$y_t = \frac{\mu}{1 - \lambda} + (1 + \lambda L + \lambda^2 L^2 + \dots)\varepsilon_t$$

El término en paréntesis es una secuencia que se puede re-escribir de esta forma  $\sum_{i=0}^{\infty} \lambda^i \varepsilon_{t-i}$ ,

la cual representa un proceso de media móvil. Así la expresión anterior se escribe

$$y_t = \frac{\mu}{1 - \lambda} + \sum_{i=0}^{\infty} \lambda^i \varepsilon_{t-i} \quad (1.13)$$

Esta definición sugiere que un proceso AR(1) puede ser expresado como un proceso MA( $\infty$ ), lo que significa que  $y_t$  es una agregación de la historia de todas las innovaciones.

## 2. MA como AR

Se considera que un modelo MA(1) puede representarse como un modelo  $AR(\infty)$  debido al siguiente desarrollo:

Sea  $y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1}$ , y haciendo uso del operador de rezago, se tiene que  $y_t = \mu + \varepsilon_t - \theta L \varepsilon_t$  lo que es igual a  $y_t = \mu + (1 - \theta L) \varepsilon_t$ . Ahora bien, si cada término es dividido por  $(1 - \theta L)$ , se obtiene lo siguiente:

$$\frac{y_t}{1 - \theta L} = \frac{\mu}{1 - \theta} + \frac{1 - \theta L}{1 - \theta L} \varepsilon_t = \frac{\mu}{1 - \theta} + \varepsilon_t. \quad (1.14)$$

Considerando  $|\theta| < 1$ , se establece que  $\frac{1}{1 - \theta L} = (1 + \theta L + \theta^2 L^2 + \dots)$ , lo cual representa la

suma de una progresión geométrica decreciente de razón  $\theta L$ . En consecuencia,

$$\frac{\mu}{1 - \theta L} = (1 + \theta L + \theta^2 L^2 + \dots) \mu = \frac{\mu}{1 - \theta}^1, \text{ y a partir de ello, se puede plantear la siguiente}$$

expresión:

$$\frac{y_t}{1 - \theta L} = (1 + \theta L + \theta^2 L^2 + \dots) y_t = y_t + \sum_{i=1}^{\infty} \theta^i y_{t-i} \quad (1.15)$$

---

<sup>1</sup> . De acuerdo con Greene (1998), al ser  $\mu$  una constante, la expresión  $\frac{\mu}{1 - \theta L} = \mu = (\theta \mu + \theta^2 \mu + \dots)$  es igual a  $\frac{\mu}{1 - \theta}$ . Esto significa que el operador de retardos puede fijarse igual a 1 cuando opera con una constante.

La cual se puede definir de forma alternativa para afirmar que un modelo MA(1) puede representarse como un modelo  $AR(\infty)$ . La nueva especificación sería igual a:

$$y_t + \sum_{i=1}^{\infty} \theta^i y_{t-i} = \frac{\mu}{1-\theta} + \varepsilon_t . \quad (1.16)$$

donde

$$y_t = \frac{\mu}{1-\theta} - \sum_{i=1}^{\infty} \theta^i y_{t-i} + \varepsilon_t$$

A este punto el lector se preguntará ¿cuál es la utilidad de saber estas relaciones? Las mismas son fundamentales cuando se realizan derivaciones formales de modelos complejos, por ejemplo la caracterización de un sistema multivariado, ya que se puede simplificar la notación refiriéndose a ellas.<sup>2</sup>

#### 1.4. Funciones de Autocovarianza y Autocorrelación

Cada uno de los procesos que definimos anteriormente presenta relaciones de correlación y covarianza, las cuales en la práctica nos pueden ser útiles para identificar el proceso que una variable pudiera seguir. A continuación las presentamos.

---

<sup>2</sup> El lector que haya revisado algún artículo econométrico técnico, reconocerá que frecuentemente se enuncian en el texto supuestos sobre la caracterización de procesos, por ejemplo cuando se refieren a la posibilidad de invertir un proceso estocástico. Es en esas instancias cuando es útil conocer estas relaciones.

Definimos la autocovarianza de una serie  $y_t$  como

$$\gamma_j = \text{cov}(y_t, y_{t-j}) = E(y_t y_{t-j}) \quad (1.17)$$

Nótese que con  $j = 0$  tenemos  $\gamma_0 = E(y_t y_t) = E(y_t)^2 = \text{Var}(y_t)$

Para los modelos que hemos considerado en el texto podemos definir las funciones de autocovarianza de la siguiente forma:

a) Ruido blanco

$$\gamma_0 = \text{Var}(y_t)$$

b) AR(1)

$$\gamma_0 = E(y_t y_t) = E((\lambda y_t + \varepsilon_t)(y_t)) = \sigma^2$$

$$\gamma_1 = E(y_t y_{t-1}) = E((\lambda y_{t-1} + \varepsilon_t)(y_{t-1})) = \lambda \sigma^2$$

$$\gamma_2 = E(y_t y_{t-2}) = E((\lambda^2 y_{t-2} + \lambda \varepsilon_{t-1} + \varepsilon_t)(y_{t-2})) = \lambda^2 \sigma^2$$

.

.

$$\gamma_k = E(y_t y_{t-k}) = E((\lambda^k y_{t-k} + \dots + \varepsilon_t)(y_{t-k})) = \lambda^k \sigma^2$$

c) MA(1)

$$\gamma_0 = \text{Var}(y_t) = \text{Var}(\varepsilon_t + \theta\varepsilon_{t-1}) = \sigma^2 + \theta^2\sigma^2 = (1 + \theta^2)\sigma^2$$

$$\gamma_1 = E(y_t y_{t-1}) = E((\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-1} + \theta\varepsilon_{t-2})) = \theta\sigma^2$$

$$\gamma_2 = 0$$

.

.

$$\gamma_k = 0$$

Similarmente, la correlación entre  $y_t$  e  $y_{t-j}$  se define como la función de autocorrelación

$$\rho_j = \frac{E(y_t y_{t-j})}{\text{Var}(y_t)} = \frac{\gamma_j}{\text{Var}(y_t)} = \frac{\gamma_j}{\gamma_0} \quad (1.18)$$

Para los modelos que hemos considerado tenemos

a) Ruido blanco

$$\rho_0 = 1, \text{ debido a que } \frac{\gamma_0}{\gamma_0} = 1$$

b) AR(1)

$$\rho_0 = 1$$

$$\rho_1 = \frac{\lambda\sigma^2}{\sigma^2} = \lambda$$

$$\rho_2 = \frac{\lambda^2\sigma^2}{\sigma^2} = \lambda^2$$

.

.

$$\rho_k = \frac{\lambda^k\sigma^k}{\sigma^k} = \lambda^k$$

c) MA(1)

Considerando el modelo  $y_t = \varepsilon_t + \theta\varepsilon_{t-1}$  se tiene que

$$\rho_0 = 1$$

$$\rho_1 = \frac{\theta\sigma^2}{(1+\theta^2)\sigma^2} = \frac{\theta}{(1+\theta^2)}$$

$$\rho_2 = 0$$

.

.

$$\rho_k = 0$$

Como veremos en los ejercicios prácticos que se presentan más adelante, las funciones de autocovarianza y autocorrelación resultan fundamentales en la identificación del PEG que pudiera representar a una serie de tiempo.

### 1.5. Sugerencia de Procedimiento Práctico

Una vez especificados los modelos teóricos AR, MA y ARMA, presentamos ahora una descripción de los pasos que generalmente se siguen en la aplicación de los mismos, de acuerdo a lo sugerido por Box-Jenkins (1970). En capítulos posteriores ejemplificaremos la misma con ejercicios prácticos. Resumiendo tenemos las siguientes generalizaciones:

$$\text{AR}(p): y_t = \mu + \lambda_1 y_{t-1} + \lambda_2 y_{t-2} + \dots + \lambda_p y_{t-p} + \varepsilon_t$$

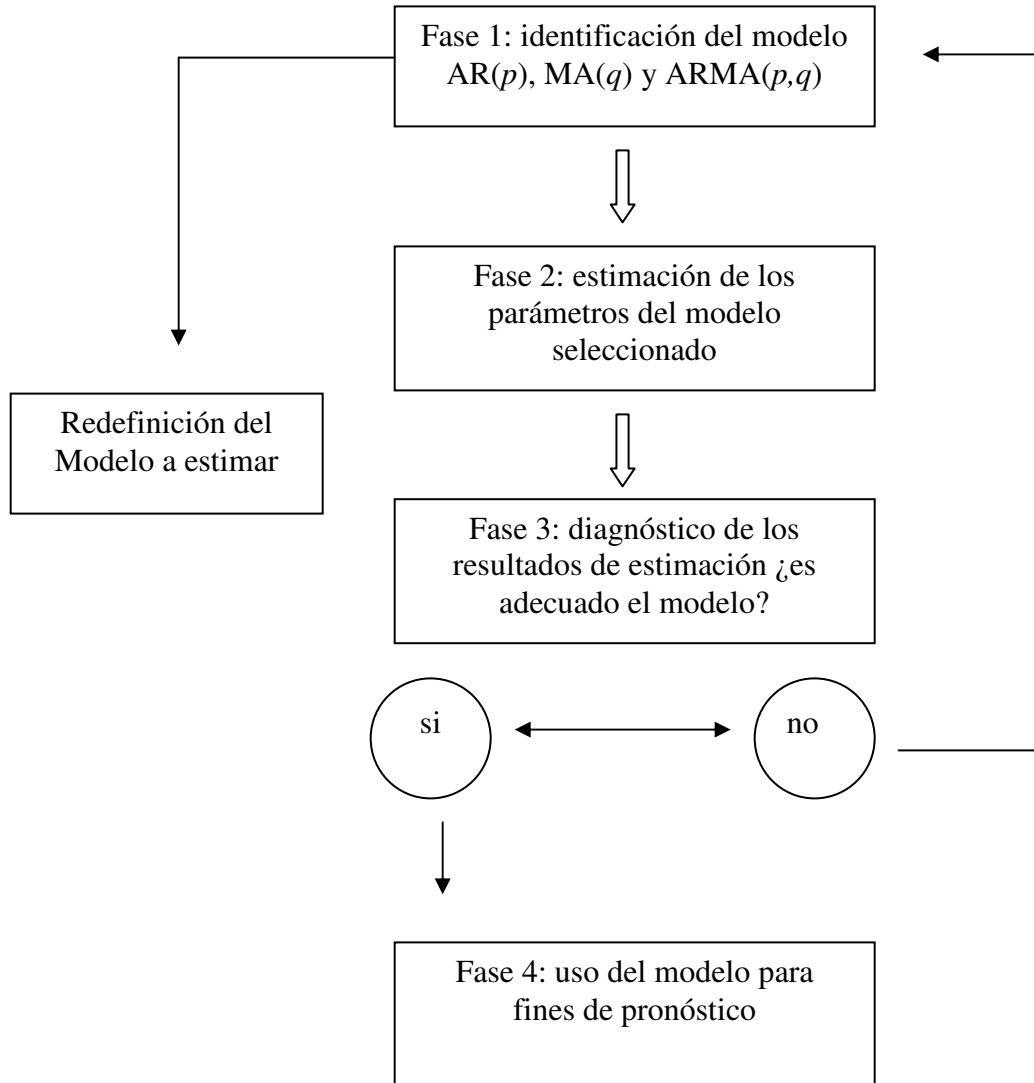
$$\text{MA}(q): y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

$$\text{ARMA}(p,q): y_t = \alpha + \lambda_1 y_{t-1} + \lambda_2 y_{t-2} + \dots + \lambda_p y_{t-p} + \varepsilon_t - \varphi_1 \varepsilon_{t-1} - \dots - \varphi_q \varepsilon_{t-q}$$

En el desarrollo de la metodología Box-Jenkins se plantean cuatro fases que son: identificación del modelo, estimación de parámetros, diagnóstico, y predicción. Cada una de éstas requiere de un trabajo cuidadoso, pues de ello depende que el modelo seleccionado cumpla satisfactoriamente con su objetivo y utilidad. En la Figura 1.1 se ilustran los pasos. Como se puede observar, la primera fase consiste en identificar los componentes  $p$  y  $q$  del modelo, lo cual puede lograrse apoyándose en el correlograma de las funciones de autocorrelación simple,  $\rho_k$ , y parcial,  $\rho_{kk}$ , de las series. Esta tarea se facilita en la medida

en que se toman como referentes los patrones teóricos señalados en el Cuadro 1.1, y que se derivan directamente de los modelos teóricos descritos en la sección anterior.

**Figura 1.1 Metodología Box – Jenkins**



**Cuadro 1.1. Patrones Teóricos**

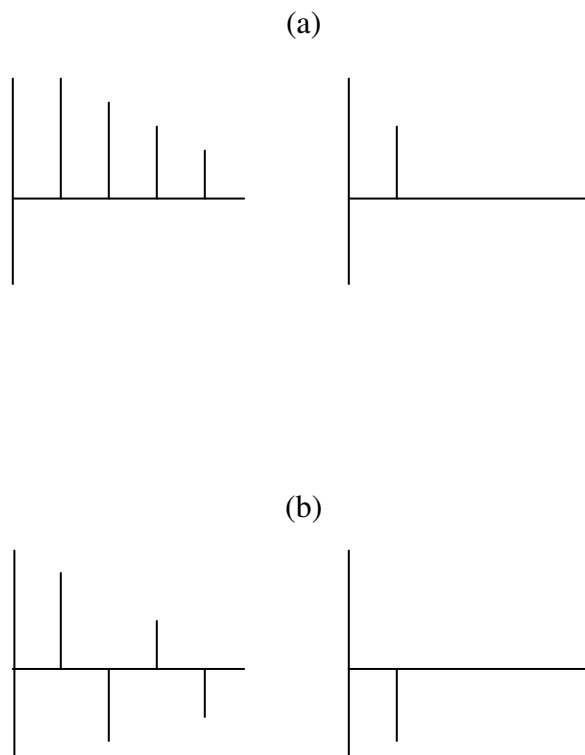
Modelo	Función de autocorrelación $\rho_k$	Función de autocorrelación Parcial $\rho_{kk}$
AR( $p$ )	Disminución exponencial combinando coeficientes positivos y negativos en forma regular o sinusoidal	Coefficientes significativos a lo largo de $p$ retardos
MA( $q$ )	Coefficientes significativos a lo largo de $q$ retardos	Decrecimiento exponencial
ARMA( $p,q$ )	Decrecimiento exponencial	Decrecimiento exponencial

Se plantea que la elección de un modelo AR( $p$ ) se efectúa cuando se observa que los coeficientes de autocorrelación simple decrecen exponencialmente y además las correlaciones parciales reflejan picos grandes a lo largo de  $p$  rezagos. Por su parte, se elige un modelo MA( $q$ ) cuando la función de autocorrelación parcial decrece exponencialmente o de forma sinusoidal o combinando valores positivos y negativos, conjugado con una función de autocorrelación simple que muestra coeficientes significativos en ciertos  $p$  retardos. Si ambas funciones decrecen exponencialmente, entonces se estima un modelo ARMA( $p,q$ ).

Como se indicó en la Sección 1.2, el modelo AR( $p$ ) es consistente con un ARMA( $p,0$ ), y se caracteriza por mostrar una función de autocorrelación muestral que disminuye aceleradamente hacia cero, o bien, de forma regular, sinusoidal o alternando valores positivos y negativos y que se conjuga con una función de autocorrelación parcial con valores distintos de ceros. Asimismo, el modelo MA( $q$ ) es congruente con un ARMA( $0,q$ ) que se distingue por manifestar una función de autocorrelación parcial que decrece rápidamente hacia cero, bien en forma regular, inusoidal o alternando los valores positivos y negativos, junto con una función de autocorrelación con tantos valores distintos de cero como el orden de la media

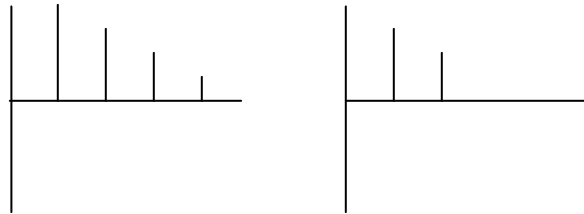
móvil. En la Figura 1.2 se ilustran algunos coeficientes de autocorrelación teóricos. Cada línea sobre el eje X está asociada a un coeficiente de autocorrelación simple y parcial y a un rezago temporal  $k$ . En realidad son representaciones teóricas que reflejan distintos patrones de comportamiento en dichas funciones, y cuyos valores se infiere que están representados sobre el eje Y.

**Figura 1.2. Correlogramas de Coeficientes de Autocorrelación Teóricos**  
**Proceso AR(1)=ARMA(1,0):  $y_t = \psi_0 + \psi_1 y_{t-1} + \varepsilon_t$**   
**Función de autocorrelación AC con  $k$  rezados**      **Función de autocorrelación parcial PAC con  $k$  rezagos**

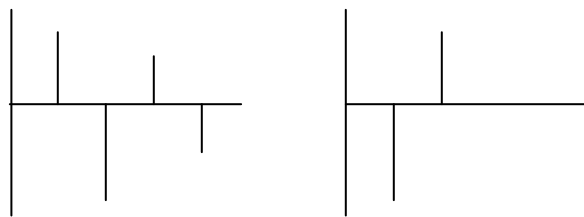


**Proceso AR(2) = ARMA(2,0):  $y_t = \psi_0 + \psi_1 y_{t-1} + \psi_2 y_{t-2} + \varepsilon_t$**   
**Función de autocorrelación AC con  $k$  rezagos**      **Función de autocorrelación parcial PAC con  $k$  rezagos**

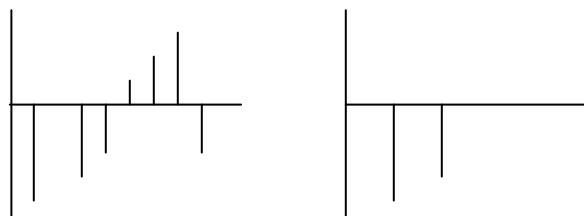
(c)



(d)



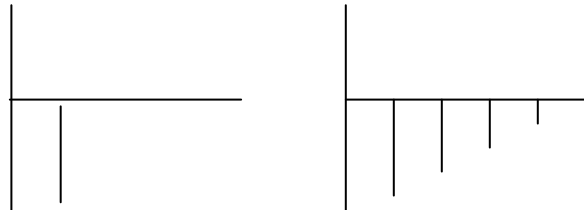
(e)



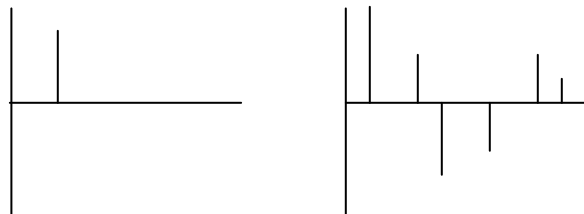
**Proceso: MA(1) = ARMA(0,1):  $y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1}$**

**Función de autocorrelación AC con  $k$  rezagos**                      **Función de autocorrelación parcial PAC con  $k$  rezagos**

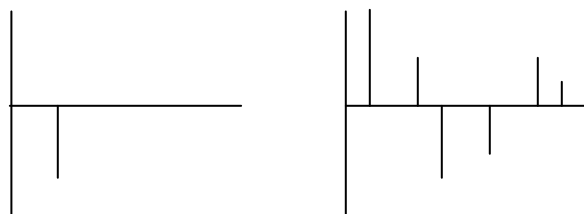
(f)



(g)



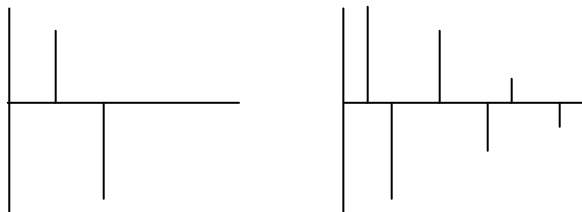
(h)



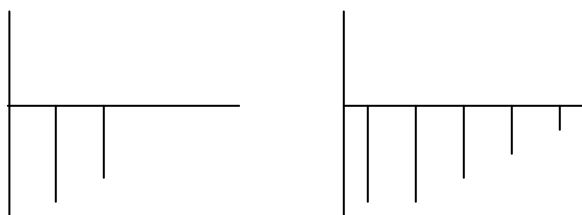
**Proceso MA(2)=ARMA(0,2):  $y_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2}$**

**Función de autocorrelación AC**                      **Función de autocorrelación parcial PAC**  
**con  $k$  rezagos**    **con  $k$  rezagos**

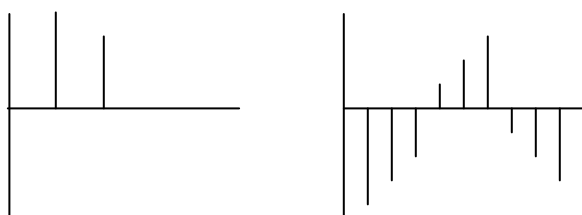
(i)



(j)



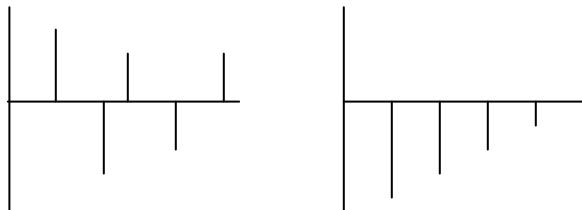
(k)



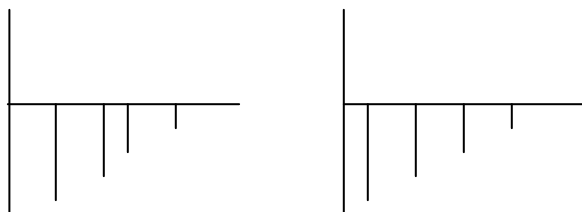
**Proceso ARMA(1,1):  $y_t = \psi_0 + \psi_1 y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$**

<b>Función de autocorrelación AC</b> <b>con <math>k</math> rezagos</b>	<b>Función de autocorrelación parcial PAC</b> <b>con <math>k</math> rezagos</b>
---	--

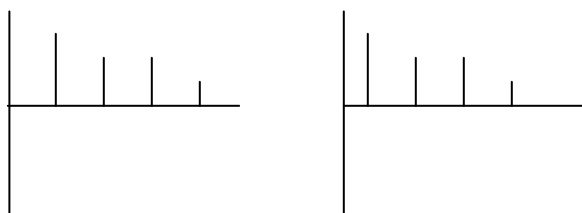
(l)



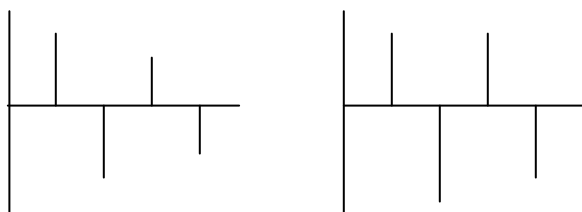
(m)



(n)



(o)



Por ejemplo, los diagramas (a) y (b) proyectan el caso de un modelo AR(1). En los dos diagramas la función de autocorrelación parcial manifiesta un pico en uno de los coeficientes teóricos, significando que es distinto de cero; lo que en términos muestrales significa que se ubica fuera del intervalo de confianza.

Un modelo AR(2) se caracteriza por presentar correlogramas teóricos como los ilustrados en el diagrama (c), (d) y (e). En el primer caso, la función de autocorrelación decrece en forma regular hacia cero, en el segundo también se observa un decrecimiento combinando valores positivos y negativos en los coeficientes, y en el tercero, resalta una disminución sinusoidal. En los tres diagramas el correlograma de la función de autocorrelación parcial se distingue por manifestar picos significativos en dos de sus coeficientes convirtiéndolos en valores distintos de cero.

Por su parte, un proceso MA(1), de acuerdo con los diagramas (f), (g) y (h), es aquel en donde se tiene que la función de autocorrelación parcial disminuye regularmente hasta llegar a cero, o bien alternando valores de sus coeficientes positivos y negativos o en forma sinusoidal, combinado esto con una función de autocorrelación cuyos coeficientes resaltan a lo largo de los  $q$  rezagos. Lo mismo se esperaría para un modelo MA(2), sólo que en este caso, tal y como se muestra en los diagramas, (i), (j) y (k), se tendrían 2 picos o coeficientes distintos de cero en el correlograma de la función de autocorrelación. En el caso del modelo ARMA, la regla de identificación consiste en ver si ambas funciones decrecen bajo distintas posibilidades. En los diagramas (l), (m), (n) y (o) se exponen algunos patrones consistentes con este tipo de modelo.

La segunda fase de la metodología Box – Jenkins, tiene que ver directamente con la estimación de los parámetros del modelo identificado. El objetivo fundamental es hacer una evaluación de la confiabilidad de los resultados de regresión a partir del análisis de diversos estadísticos de prueba. En principio, se tiene que contrastar la significancia estadística de los coeficientes estimados haciendo uso de la distribución normal. También se tienen que analizar las varianzas y compararlas en caso de que se tengan dos o más modelos alternativos con especificaciones distintas. De igual manera se tiene que valorar la relación de la bondad de ajuste medida a través del coeficiente de determinación múltiple normal o ajustado.

En la tercera fase, se realiza un diagnóstico para valorar la confiabilidad de los resultados de regresión. Esta tarea se desarrolla analizando el correlograma de los coeficientes de autocorrelación simple y parcial de los residuales del modelo identificado y estimado. El diagnóstico debe arrojar resultados que indique que los coeficientes de autocorrelación son iguales a cero, lo que significa que cada residuo debe ubicarse dentro del intervalo de confianza del correlograma. Esto se corrobora con las pruebas de Box-Pierce y Ljung-Box que, en general, permitirán concluir si los residuales del modelo estimado son ruido blanco. Bajo la hipótesis nula de que los residuos son ruido blanco, si cualquiera de ambos estadísticos es mayor al valor crítico correspondiente a una  $\chi^2$ , se rechaza dicha hipótesis y se determina que el modelo debe ser corregido o bien, que se debe retornar a la fase número uno, relacionada con la identificación del modelo.

Una vez concluida satisfactoriamente la tercera fase de la metodología descrita, se procede a desarrollar la cuarta y última etapa que consiste en utilizar el modelo para fines de

pronóstico. De no ser favorable la validación del modelo, entonces el econometrista está obligado a retornar a la primera fase con el interés de identificar otro tipo de modelo que arroje buenos estadístico  $t$ ,  $F$ , Durbin-Watson y coeficiente de determinación múltiple ajustado. En caso de que se hayan estimado dos o más modelos que resulten ser aceptables, se deberá elegir uno de ellos en función de los estadísticos ya señalados y de otros criterios que miden el margen de error de las estimaciones como el error cuadrático medio o el error absoluto medio.

Antes de mostrar un ejemplo concreto de la instrumentación de esta metodología, abordamos dos temas que son de particular importancia en el análisis de modelos univariados, el de la estacionariedad, y el de raíces unitarias.

## Capítulo 2: Estacionariedad

En este capítulo introducimos el concepto de estacionariedad, que fundamentalmente se refiere a condiciones sobre la distribución de una variable a través del tiempo. Asimismo, presentamos una primera aproximación a técnicas útiles para identificar estacionariedad en series de tiempo, éstas se refieren al análisis gráfico y de las funciones de correlación de las variables, que de hecho ejemplifican la instrumentación práctica de la metodología Box-Jenkins. En el capítulo 3 consideramos herramientas más formales que identifican puntualmente la naturaleza estocástica de una serie de tiempo.

### 2.1. Concepto

El concepto de estacionariedad puede ser entendido en sentido estricto y amplio. El primer caso se presenta cuando la función de distribución de probabilidad de una variable no varía con respecto a un desplazamiento en el tiempo. Esto significa que la distribución conjunta de  $m$  observaciones  $y_{t_1}, y_{t_2}, \dots, y_{t_m}$  es similar a la de  $y_{t_1+k}, y_{t_2+k}, \dots, y_{t_m+k}$  para todo  $m$  y  $k$  retardos. Como esta definición implica una restricción fuerte de los procesos en el sentido de que cada distribución de  $y(t)$  es independiente del tiempo, es que se define la estacionariedad débil, en sentido amplio o de segundo orden. Por tanto, una serie es débilmente estacionaria si su media es constante e independiente del tiempo, su varianza es finita y constante, y el valor de su covarianza entre dos periodos no depende del tiempo, sino de la distancia entre ellos. En consecuencia, una serie presenta estacionariedad débil si reúne las siguientes condiciones:

*Media* :  $E(y_t) = \mu$ , constante  $\forall_t$

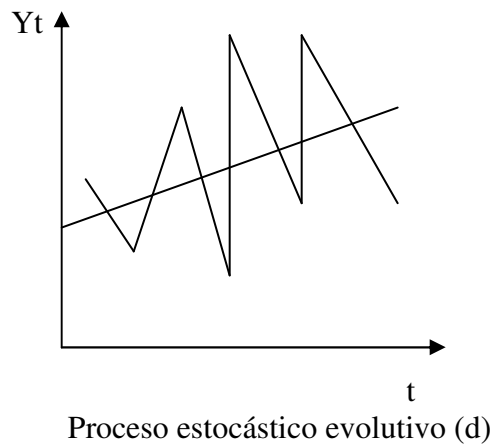
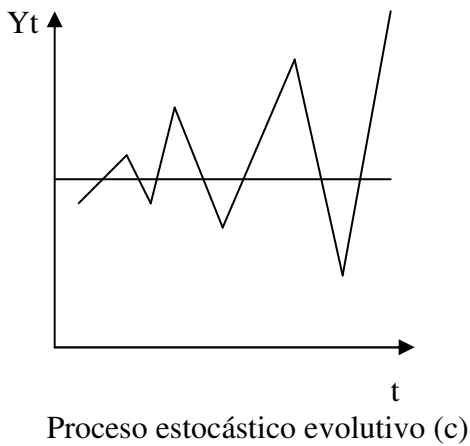
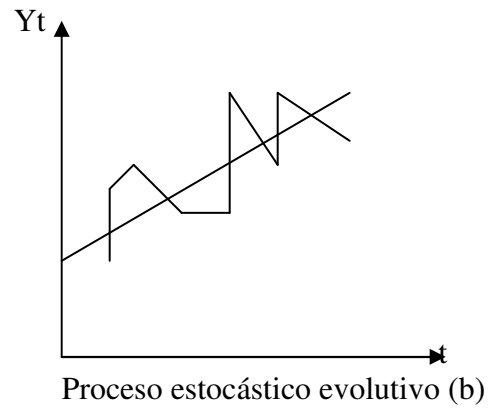
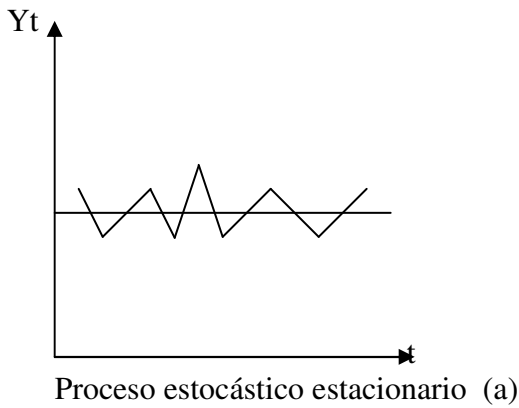
*Varianza* :  $Var(y_t) = E(y_t - \mu)^2 = \sigma^2 = \gamma_0$ , constante  $\forall_t$

*Co varianza* :  $Cov(y_t, y_{t+k}) = E[(y_t - \mu)(y_{t+k} - \mu)] = \gamma_k$  siendo  $t \neq t+k$

En la Figura 2.1, se ilustran distintos tipos de patrones que permiten apreciar el tipo de serie de tiempo que se tiene como realización de un proceso estocástico. Por ejemplo, la figura del panel (a) refleja el comportamiento típico de una serie estacionaria en un sentido amplio, en el panel (b) se ilustra el caso de una serie que no es estacionaria debido a la evolución de la media alrededor de la tendencia. En el panel (c), se constata la no estacionariedad producto de cambios en la varianza en el transcurso del tiempo y finalmente en el panel (d), no se percibe una serie con media y varianza constantes.

Un tipo de proceso que ha recibido marcada atención en el análisis econométrico es el llamado ruido blanco, cuya característica es que está formado por una sucesión de variables aleatorias con distribución normal, esperanza cero y varianza constante. Un ejemplo de esto sería la serie hipotética del panel (a), cuyos residuos reúne las características de ser ruido blanco, es decir,  $\varepsilon_t \sim N(0, \sigma_t^2)$  para todo  $t$ , tal que  $cov(\varepsilon_t, \varepsilon_{t+1}) = 0, \forall t \neq t+1$ .

**Figura 2.1. Tipos de Procesos Estocásticos**



Otro caso que ha recibido particular atención en los desarrollos de la teoría econométrica es el proceso estocástico ergódico, el cual se presenta cuando los valores de la serie alejados en el tiempo están poco correlacionados, es decir, cuando los coeficientes de autocorrelación decrecen al aumentar el retardo  $k$  lo que significa que  $\lim_{k \rightarrow \infty} \rho_k = 0$ . Si dicho proceso sigue una distribución normal, se puede decir que las observaciones de la serie son cada vez más independientes de las anteriores, lo que permite construir estimadores consistentes de los

parámetros. Al respecto, un estimador  $\hat{\delta}$  es consistente si se acerca al verdadero valor de  $\delta$  a medida que el tamaño de la muestra se incrementa, situación que puede ser representada como  $p \lim_{n \rightarrow \infty} \hat{\delta} = \delta$ , donde *plim* significa probabilidad en el límite.

Un proceso elemental que también es materia de análisis en la literatura especializada de series de tiempo, es el denominado caminata aleatoria (*random walk*), que se manifiesta de dos formas, una que contempla rumbo y otro que no. Este último caso se puede expresar en términos de la siguiente especificación:

$$y_t = y_{t-1} + \varepsilon_t \tag{2.1}$$

Donde  $\varepsilon_t$  reúne las características de ser ruido blanco y la variable  $y_t$  es una serie cuyo valor es igual al inmediatamente anterior más un choque aleatorio.

En esta lógica,  $y_{t-1} = y_{t-2} + \varepsilon_{t-1}$  y  $y_{t-2} = y_{t-3} + \varepsilon_{t-2}$ , por lo que sustituyendo progresivamente se tiene que

$$y_t = y_{t-1} + \varepsilon_t = \varepsilon_t + \varepsilon_{t-1} + y_{t-2} = \varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + y_{t-3} = \dots =$$

Por tanto, si el proceso no tiene pasado infinito e inició con el valor  $y_0$  como condición inicial, entonces  $y_t = \varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + y_0$  teniendo  $y_t = \sum \varepsilon_j + y_0$  para  $j = 1, 2, \dots, t$ .

En consecuencia, la media se expresa como

$$\mu = E(y_t) = \left( \sum \varepsilon_j + y_0 \right) = y_0 \sum E(\varepsilon_j) = y_0$$

y la varianza como

$$\text{var}(y_t) = E(y_t - y_0)^2 = E\left(\sum \varepsilon_j\right)^2 = \sum E(\varepsilon_j)^2 = t\sigma^2 .$$

Esto significa que la media de  $y_t$  es igual al valor inicial que es constante pero la varianza no lo es, ya que crece de manera indefinida en la medida que  $t$  aumenta, lo cual viola una condición de estacionariedad.

Una extensión del modelo anterior, es el proceso de caminata aleatoria con rumbo o constante que se define:

$$y_t = \beta + y_{t-1} + \varepsilon_t \tag{2.2}$$

Y que al igual que el caso anterior, es un proceso AR(1), con la diferencia que incluye un parámetro de variación  $\beta$ . Si (2.2) se reescribe como  $y_t - y_{t-1} = \Delta y_t = \beta + \varepsilon_t$ , entonces se deduce que la tendencia de  $y_t$  se incrementa si  $\beta > 0$  y disminuye en caso de que  $\beta < 0$ .

¿Por qué es importante conocer la estacionariedad de una serie de tiempo? Recuérdese que los supuestos de la regresión clásica suponen que los dos primeros momentos de un proceso estocástico son constantes, esto es, media constante (no necesariamente igual a 0) y varianza constante. Bajo esos supuestos los estimadores que se obtienen por medio de Mínimos Cuadrados Ordinarios (MCO's) son insesgados y eficientes. De tal forma que resulta fundamental determinar si los supuestos de media y varianza constantes se cumplen para proceder con una estimación; ya que si no lo fueran entonces los resultados de una regresión no serían confiables.

## **2.2 . Identificación de Estacionariedad: Primera Aproximación**

Aquí se describen distintas herramientas que ayudan a evaluar las condiciones de estacionariedad de una serie. Primero se explican las funciones de autocorrelación simple y autocorrelación parcial y subsecuentemente, se analizan los estadísticos de prueba propuestos por Box y Pierce (1970), y Ljung y Box (1978).

### **a) Funciones de autocorrelación simple y autocorrelación parcial**

Una primer prueba que se utiliza para diagnosticar si una serie es estacionaria, es la función de autocorrelación simple de la serie de tiempo  $y_t$ . Este instrumento permite determinar el grado de autocorrelación entre las observaciones individuales que conforman la muestra total de la serie  $y_1, y_2, \dots, y_T$ , considerando un desfase temporal igual a  $k$ . Su representación gráfica se ilustra a través de un correlograma que está conformado por todos

los coeficientes muestrales. En el Capítulo 1 se presentaron varias caracterizaciones para las funciones de correlación y covarianza de los modelos ARMA, aquí empleamos un lenguaje más general para ilustrar el ejercicio práctico. En particular, para un desfase de  $k$  periodos la autocorrelación se define a través de la siguiente expresión:

$$\rho_k = \frac{E[(y_t - \mu)(y_{t+k} - \mu)]}{\sqrt{E(y_t - \mu)^2 E(y_{t+k} - \mu)^2}} = \frac{Cov(y_t, y_{t+k})}{\sigma_{y_t} \sigma_{y_{t+k}}} \quad (2.3)$$

donde  $\rho_k$  es el coeficiente de autocorrelación,  $\mu$  la media de la serie analizada y el subíndice  $k$  representa el orden del rezago.

Se deduce que para una serie estacionaria son iguales las varianzas  $\sigma_{y_t}$  y  $\sigma_{y_{t+k}}$ ; por lo que la expresión (2.3), se redefiniría como:

$$\rho_k = \frac{E[(y_t - \mu_y)(y_{t+k} - \mu_y)]}{\sigma_y^2} = \frac{Cov(y_t, y_{t+k})}{\sigma_y^2} = \frac{\gamma_k}{\gamma_0} \quad (2.4)$$

Siendo  $\gamma_k$  la covarianza al rezago  $k$  y  $\gamma_0$  la varianza de  $y_t$ . En el caso particular de que la longitud del rezago fuera  $k = 1$ , se tendría:

$$\rho_0 = \frac{E(y_t - \mu_y)}{\sigma_y^2} = 1$$

Hay que considerar que para inferir sobre los coeficientes poblacionales, es necesario estimar los coeficientes muestrales, los cuales oscilan entre  $-1$  y  $+1$ . En este caso, la función de autocorrelación se expresa para el conjunto de la muestra como  $\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}$ , donde:

$$\hat{\gamma}_k = \frac{\sum (y_t - \bar{y})(y_{t+k} - \bar{y})}{T}$$

$$\hat{\gamma}_0 = \frac{\sum (y_t - \bar{y})^2}{T}$$

De tal forma que la función muestral estará dada por:

$$\hat{\rho}_k = \frac{\sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2}. \quad (2.5)$$

Donde  $\bar{y}$  es la media de la población

Para obtener el correlograma de la función de autocorrelación simple, es recomendable que se cuente con un tamaño de muestra lo suficientemente grande como para evaluar la autocorrelación de orden superior. Es importante destacar que no se podrían efectuar las estimaciones si  $k > T - 1$ . En la práctica es común que se estimen los coeficientes considerando  $k = n/4$ . Por ejemplo, si se dispone de una serie de 100 observaciones, puede calcularse  $\hat{\rho}_k$  para los 25 primeros coeficientes de la función de autocorrelación simple

(FAC). Así pues, la función de autocorrelación es un instrumento que ayuda a especificar el proceso estocástico. En particular la función de autocorrelación parcial de un proceso estocástico  $\{y_1, y_2, \dots, y_T\}$ , es similar al de FAC, con la diferencia que está corregido por los rezagos intermedios, ya que denota el efecto marginal que cada  $t-k$  tiene sobre  $t$ . Para encontrar la FAP de primer orden que es exactamente a la FAC del mismo orden, debido que no hay rezagos intermedios, se estima la regresión  $\tilde{y} = \delta_1 \tilde{y}_{t-1} + \varepsilon_{1t}$  donde  $\tilde{y}$  es la desviación de  $y_t$  respecto a su media. Para estimar la FAP de segundo orden, se requiere correr la regresión  $\tilde{y} = \lambda \tilde{y}_{t-1} + \delta_2 \tilde{y}_{t-2} + \varepsilon_{2t}$  donde  $\delta_2$  es justamente la FAP de orden(2), pero  $\lambda$  no representa la FAP de orden(1) ya así sucesivamente..

#### **b) Estadísticos de Box – Pierce y Ljung - Box**

Una vez que se ha calculado la función de autocorrelación muestral, la siguiente fase de análisis para evaluar si una serie temporal es estacionaria, consiste en verificar si un valor particular de  $\hat{\rho}_k$  es estadísticamente significativo. En este caso la hipótesis a contrastar es si  $\rho_k = 0$  para una  $k$  particular o bien si  $\rho_k = 0$  para  $k > 0$ , en cuyo caso, se utilizan otros estadísticos de prueba. Bartlett (1946) plantea que si una serie de tiempo se caracteriza por ser ruido blanco (serie puramente aleatoria), entonces los coeficientes de autocorrelación muestral para  $k > 0$  están distribuidos aproximadamente de acuerdo con una distribución normal con media 0 y desviación estándar igual a  $1/\sqrt{T}$ , donde  $T$  es el número de observaciones y  $\hat{\rho}_k \sim N(0, 1/T)$ .

Esta prueba de significancia individual está basada en un intervalo del 95% de confianza que sigue las propiedades de la distribución normal estándar y que se especifica como  $(\rho_k \pm 1.96\sqrt{1/T})$  La regla de decisión es que si este intervalo incluye el valor de cero, entonces no se rechaza la hipótesis de que  $\rho_k = 0$  y por tanto, se concluye que no hay autocorrelación de orden  $k$ . Alternativamente, se puede utilizar el estadístico de Box y Pierce para evaluar la hipótesis de que todos los coeficientes de autocorrelación son iguales a cero,  $H_0 : \rho_1 = \rho_2 = \dots, \rho_k = 0$  o bien de que  $\rho_k = 0$  para  $k > 0$ . El estadístico de prueba está definido por  $Q = T \sum_{k=1}^K \rho_k^2$ , donde  $T$  es el tamaño de la muestra y  $k$  la longitud del rezago. Éste es un estadístico que está distribuido como una  $\chi^2$  con  $k$  grados de libertad, por lo que  $Q(k) \sim \chi^2(k)$ . Si la  $Q$  calculada excede el valor teórico al nivel de significancia seleccionado, se puede rechazar la hipótesis de que todos los  $\rho_k$  son iguales a cero, por lo que al menos algunos de ellos deben ser diferentes de cero.

En el análisis de la estacionariedad también es frecuente utilizar el estadístico sugerido por Ljung y Box y que en realidad es una variante de la prueba anterior. La hipótesis que se contrasta es  $H_0 : \rho_1 = \rho_2 = \dots, \rho_k = 0$  y el estadístico está definido por la siguiente expresión:  $LB = T(T + 2) = \sum_{k=1}^K \left( \frac{\rho_k^2}{n - k} \right) \sim \chi^2(k)$ , que al igual que el estadístico  $Q$ , sigue una distribución  $\chi^2$  con  $k$  grados de libertad. De acuerdo con Johnston y DiNardo (1997), un aspecto singular de este estadístico es que se ha demostrado que para muestras pequeñas es más potente que el estadístico  $Q$ .

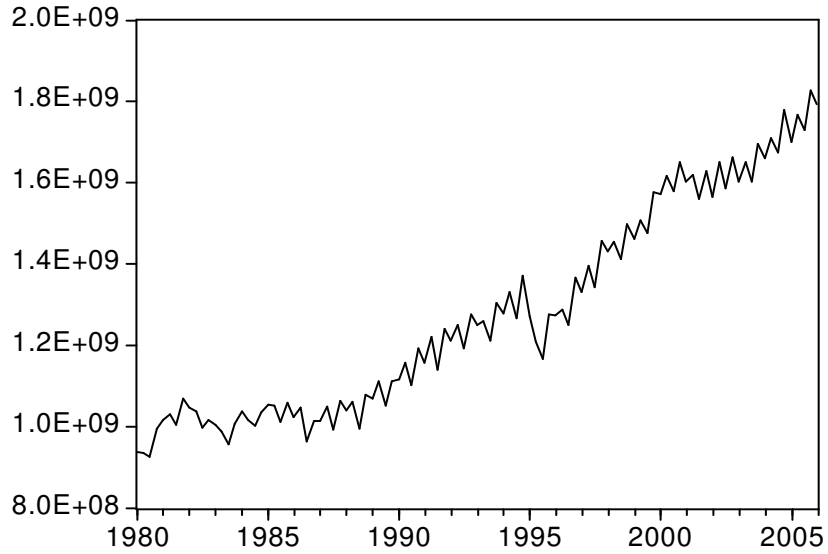
### **2.3. Ejemplo Numérico: el Producto Interno Bruto de México**

En este apartado se presenta un ejercicio práctico para ilustrar los conceptos teóricos que hasta este punto se han discutido. Se utiliza como serie objeto de análisis el Producto Interno Bruto de México (PIB). Los datos fueron obtenidos del Banco de México. De acuerdo a la Gráfica 2.1, una primera inspección de la serie indica que a lo largo del periodo 1980:01 – 2006:01 sigue una tendencia creciente con sus respectivas oscilaciones. Específicamente, llama la atención la caída registrada en el primer trimestre de 1995, la cual se explica por los efectos de la crisis económica que experimentó la economía nacional a fines de 1994. Para propósitos del desarrollo del ejercicio, se trabaja con la serie en logaritmos (LPIB), la cual se presenta en la Gráfica 2.2 mostrando fundamentalmente una dinámica similar a la serie en valores originales.<sup>3</sup>

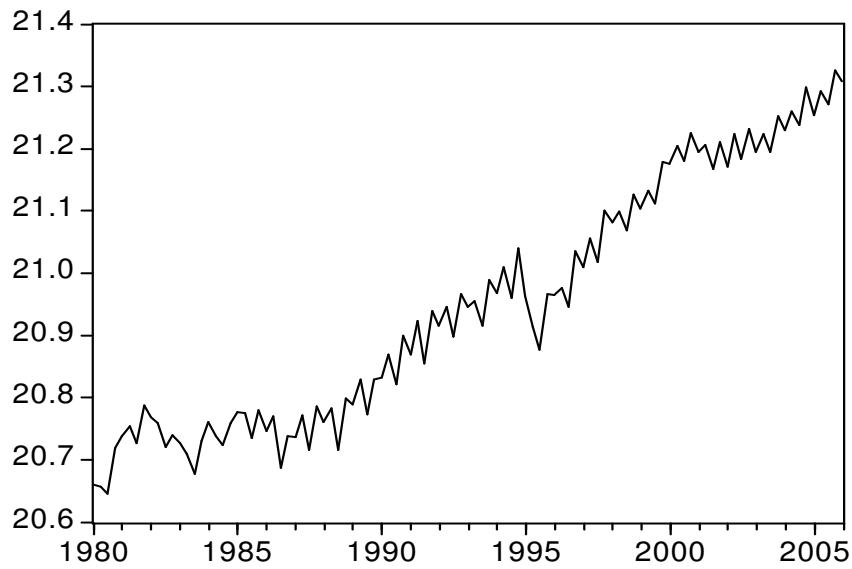
---

<sup>3</sup> A partir de este ejercicio y a través del resto del documento, las estimaciones se realizan utilizando el programa econométrico EViews, excepto para los ejercicios de ciclo común que se llevan a cabo en el Capítulo 9, para las cuales se empleó GAUSS.

**Gráfica 2.1. Producto Interno Bruto de México (1980:01 -2006:01), miles de pesos (1993=100)**

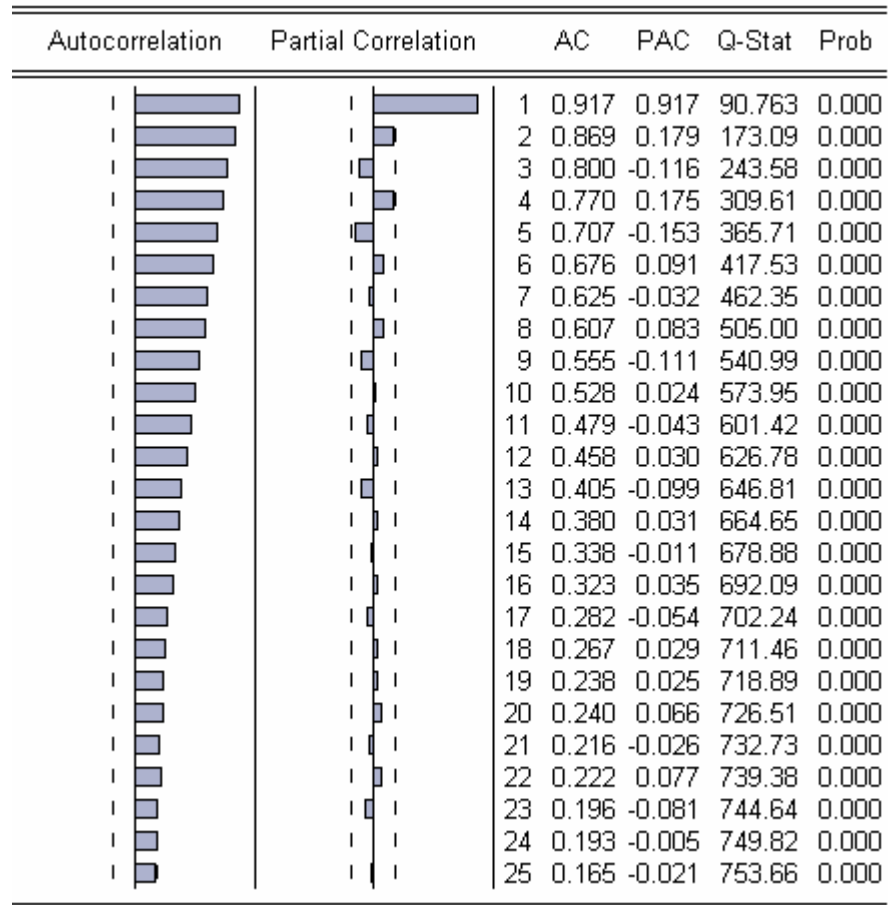


**Gráfica 2.2. Logaritmo del Producto Interno Bruto de México (1980:01 -2006:01), miles de pesos (1993=100)**



A partir de la inspección visual de la serie, notamos que la misma no parece evolucionar alrededor de una media, lo cual puede sugerir que la serie es no-estacionaria. Para continuar con el análisis, se estiman los coeficientes de autocorrelación simple y parcial representados en el correlograma de la Figura 2.2 Las barras que aparecen en las dos primeras columnas, corresponden a la gráfica de los coeficientes de autocorrelación simple y autocorrelación parcial respectivamente, que se refieren a los valores de la tercera y cuarta columna que tienen como encabezado AC (autocorrelaciones simples) y PAC (autocorrelaciones parciales). La quinta columna reporta los valores del estadístico  $Q$  de Ljung – Box y en la sexta y última columna, se reportan las probabilidades asociadas a dicho estadístico.

**Figura 2.2. Correlograma de LPIB(1993=100)**  
**T = 105 y k =25**



Al observar las barras de la primera columna, se puede constatar que con excepción del rezago número 25, todos los coeficientes  $\rho_k$  son positivos y se encuentran fuera del intervalo de confianza del 95%, determinado por las líneas punteadas ubicadas en los extremos de la línea continua. Aunque el ancho de las barras tiende a disminuir desde el primer rezago hasta el último, está claro que no hay evidencia suficiente para aceptar la hipótesis individual de que cada  $\rho_k = 0$ , ya que los intervalos no contienen el valor de cero, lo que a su vez sugiere, que la serie LPIB es, efectivamente, no-estacionaria.

Para evaluar la hipótesis conjunta de que todos los coeficientes de autocorrelación simple (AC) son simultáneamente iguales a cero, nos referimos a la prueba de Box – Pierce y Ljung –Box. En este caso la hipótesis a probar es que  $H_0 : \rho_1, \rho_2, \rho_3 \dots \rho_k = 0$ . Si el valor del estadístico de Ljung y Box, que es el que se muestra en el correlograma para  $k = 25$ , excede al valor de la  $\chi^2$  teórico a un nivel de significancia del 95% y con  $k$  grados de libertad, se rechaza la hipótesis de que todos los coeficientes son iguales a cero, lo que se corrobora con la probabilidad reportada en el correlograma de 0.000 asociada al retardo  $k = 25$ . En este sentido, se puede concluir que el PIB se comporta como una serie no-estacionaria en niveles.

Retomando el correlograma de la Figura 2.2, se agrega que cada valor del estadístico  $Q$  se va obteniendo sólo cambiando el valor correspondiente a la suma de los coeficientes de autocorrelación en el estadístico de Ljung y Box y de la diferencia  $n - k$ . Por ejemplo, para obtener la  $Q$  correspondiente al primer coeficiente de autocorrelación muestral,  $\rho_1^2 = 0.840889$  y  $n - k = 104$ . Así pues, la evolución de la serie presentada tanto en la Gráfica 2.1 como en la 2.2, arroja una clara tendencia que hace suponer que la serie no es estacionaria. El lento decrecimiento de la FAC correspondiente a la columna AC del correlograma corrobora esta inferencia. En el Cuadro 2.1 se reportan los intervalos de confianza.

**Cuadro 2.1. Rangos del Intervalo de Confianza**

<b>Prob (<math>\hat{\rho}_k - 0.191276407 \leq \rho_k \leq \hat{\rho}_k + 0.191276407</math>) = 0.95</b>		
$\hat{\rho}_1 = 0.917$	→	(0.725723593, 1.108276407)
$\hat{\rho}_2 = 0.869$	→	(0.677723593, 1.060276414)
$\hat{\rho}_3 = 0.800$	→	(0.608723586, 0.991276413)
$\hat{\rho}_4 = 0.770$	→	(0.578723586, 0.961276413)
$\hat{\rho}_5 = 0.707$	→	(0.515723586, 0.898276413)
.....		.....
.....		.....
.....		.....
$\hat{\rho}_{23} = 0.196$	→	(0.004723586, 0.387276413)
$\hat{\rho}_{24} = 0.193$	→	(0.0017235862, 0.38427641)
$\hat{\rho}_{25} = 0.165$	→	(-0.026276413, 0.356276413)

Así pues, la evidencia que se presenta hasta este punto señala que la serie del PIB es no-estacionaria. Sin embargo, a pesar de que parece contundente este resultado, es necesario realizar un análisis más formal para determinar puntualmente la naturaleza estocástica de la serie; dicho análisis se lleva a cabo en las secciones posteriores.

## Capítulo 3: Raíces Unitarias

En este capítulo profundizamos nuestro análisis sobre la propiedad de estacionariedad de una serie de tiempo. En particular, determinamos las condiciones que debe exhibir un proceso estocástico para identificar puntualmente la existencia de estacionariedad. Las mismas se evidencian por medio de la estimación de pruebas de hipótesis sobre los coeficientes de un PEG. Dichas pruebas se conocen comúnmente como pruebas de raíces unitarias. Aquí abordamos el sustento teórico de las mismas y presentamos una descripción de las más ampliamente utilizadas; complementando la discusión teórica con diversas estimaciones prácticas

### 3.1. Concepto

Para introducir el concepto de raíz unitaria presentamos un desarrollo formal. Considera la representación de un modelo AR(1)

$$y_t = \phi y_{t-1} + \varepsilon_t \quad (3.1)$$

Usando el operador de diferencia se re-escrive como

$$(1 - \phi L)y_t = \varepsilon_t \quad (3.2)$$

Invirtiendo la expresión se obtiene

$$y_t = (1 - \phi L)^{-1} \varepsilon_t \quad (3.3)$$

Representado como una expansión de Taylor se escribe

$$y_t = (1 + \phi L + \phi^2 L^2 + \dots) \varepsilon_t = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j} \quad (3.4)$$

Lo cual muestra, como ya se mostró anteriormente, que el proceso AR(1) se puede representar como un proceso  $MA(\infty)$ . El mismo converge si  $|\phi| < 1$ . Nótese entonces que ésta es una condición de convergencia. Si  $|\phi| = 1$  la serie no converge y si  $|\phi| > 1$  el proceso es explosivo.

Alternativamente, podemos re-escribir (3.3) como

$$\phi(L)y_t = \varepsilon_t \quad (3.5)$$

Donde  $\phi(L) = (1 - \phi L)$

Resolviendo por las raíces del polinomio, esto es  $\phi(L) = 0$ , se obtiene  $L = \frac{1}{\phi}$ .

Considerando la condición de convergencia anterior  $|\phi| < 1$ , tenemos que esto implica que  $L > 1$ .

Esto es, decimos coloquialmente que una serie presenta raíz unitaria si existen raíces del polinomio representativo dentro del círculo unitario, en tal caso,  $|\phi| = 1$ ,  $L = 1$ , y el sistema sería inestable, ya que la condición de convergencia no se cumpliría.

Recordará el lector que dentro de los pasos que se siguen para instrumentar la metodología Box-Jenkins, y que fueron indicados en el Capítulo 1, se plantea que previo a la identificación del modelo de predicción más apropiado se debe determinar si la serie en cuestión es estacionaria, en otras palabras, si no presenta raíz unitaria. Al respecto, la teoría econométrica señala que si bien el análisis del correlograma es un instrumento empírico de gran utilidad, es necesario que se apliquen otras pruebas más potentes, las cuales identifican formalmente la magnitud del parámetro  $\phi$ .

Retomemos la expresión (3.1)

$$y_t = \phi y_{t-1} + \varepsilon_t \quad (3.6)$$

Sustrayendo en ambos lados la variable  $y_{t-1}$  se obtiene

$y_t - y_{t-1} = \phi y_{t-1} - y_{t-1} + u_t$ , expresión que se puede re-escribir de la siguiente forma

$$\Delta y_t = \alpha y_{t-1} + u_t \quad (3.7)$$

Donde  $\Delta$  es el operador de la primera diferencia y  $\alpha = (\phi - 1)$ . Al realizar una prueba de hipótesis sobre el valor de  $\phi$  se valoraría en (3.6) si  $\phi = 1$  y en (3.7) si  $\alpha = 0$ . Estas pruebas de hipótesis constituyen de hecho el fundamento de las diversas pruebas de raíz unitaria que se encuentran en la literatura. Si se comprueba estadísticamente que  $\phi = 1$ , entonces se determina que la serie es no-estacionaria, de lo contrario se concluye que la serie no contiene raíz unitaria o es estacionaria.<sup>4</sup> Ahora bien, si se acepta la hipótesis nula de que el parámetro  $\alpha = 0$ , entonces la expresión (3.7) se reduce a  $\Delta y_t = (y_t - y_{t-1}) = u_t$ , lo que significa que la primera diferencia de la variable  $y_t$  es igual al error aleatorio  $u_t$ , y éste al ser ruido blanco implica que la serie en primera diferencia es estacionaria.

En este contexto, una serie que se diferencia en una observación y resulta ser estacionaria, se dice que es integrada de orden 1, o bien I(1). A su vez, si un proceso de caminata aleatoria se diferencia dos veces para ser estacionaria, se dice que es I(2), y si se tiene que diferenciar  $d$  veces será de orden I( $d$ ). En resumen, si una serie se caracteriza por ser I(1),.....,I( $d$ ), se determina que no es estacionaria en niveles y toma ( $d$ ) diferencias para que sea estacionaria.

Dentro de las pruebas formales que existen para determinar la presencia de raíz unitaria en un proceso estocástico, mencionamos las sugeridas por Dickey y Fuller (1979), la Dickey-Fuller Aumentada que se probó en Said y Dickey (1984), la prueba Phillips-Perron (1988), y la desarrollada por Kwiatkowski, Denis, Peter C. B. Phillips, Peter Schmidt & Yongcheol Shin (1992). A continuación presentamos cada una de éstas.

---

<sup>4</sup> Es posible, por supuesto, encontrar que  $\phi > 1$  lo cual indicaría que la serie es explosiva.

### 3.2. Pruebas Básicas de Raíz Unitaria

- Prueba Dickey-Fuller (DF)

La prueba básica DF se aplica generalmente a las siguientes regresiones:

$$\Delta y_t = \alpha y_{t-1} + u_t \quad (3.8)$$

$$\Delta y_t = \phi_1 + \alpha y_{t-1} + u_t \quad (3.9)$$

$$\Delta y_t = \phi_1 + \phi_2 t + \alpha y_{t-1} + u_t \quad (3.10)$$

Donde la primer ecuación representa una caminata aleatoria simple, la segunda una caminata aleatoria con variación,  $\phi_1$ , y la tercera añade un componente de tendencia,  $t$ . En cualquiera de las tres especificaciones si la hipótesis nula  $H_0 : \alpha = 0$  no se rechaza, se concluye que las serie objeto de estudio presenta una raíz unitaria.

Cabe destacar que uno de los supuestos fundamentales de la prueba DF es que el término de error es un proceso de ruido blanco. En este sentido es que se recomienda analizar el estadístico Durban- Watson para verificar si se cumple con el supuesto de no correlación serial. En caso de que las observaciones  $u_t$  estén autocorrelacionadas, se recomienda desarrollar la prueba Dickey-Fuller-Aumentada (DFA), que corrige por esta deficiencia.

- Prueba Dickey-Fuller Aumentada (DFA)

Esta prueba consiste en estimar la siguiente regresión:

$$\Delta y_t = \phi_1 + \phi_2 t + \alpha y_{t-1} + \sum_{i=1}^p \delta_i \Delta y_{t-i} + \varepsilon_t \quad (3.11)$$

La característica de esta técnica es que incorpora como regresores adicionales las diferencias rezagadas de la variable dependiente y se puede aplicar a cualquiera de los tres modelos anteriormente mencionados (ecuaciones 3.8, 3.9 y 3.10). El propósito de incluir los rezagos de las diferencias es precisamente controlar por el problema de autocorrelación en el término de error, del cual se alertó en la especificación DF. En este caso, se evalúa la hipótesis  $H_0 : \alpha = 0$  contra  $H_1 : \alpha < 1$ . Es importante señalar que se debe especificar el número de rezagos del término  $\Delta y_{t-i}$ , por lo que se sugiere que estos *lags*, se determinen empíricamente hasta que el término de error sea ruido blanco. En la práctica es común que se utilice criterios de información como el de Akaike o el de Schwarz para determinar el número óptimo de rezagos, lo cual implica elegir la estimación que arroje los valores de los criterios más bajos.

Una de las inquietudes que surgen al momento de estimar cualquiera de las tres regresiones relacionadas con la prueba de raíces unitarias, es si se debe o no incluir una constante o si además de la constante hay que añadir una variable de tendencia, o ninguna de las dos posibilidades. Es importante llevar a cabo una selección adecuada, pues de ello depende la

distribución asintótica del estadístico de prueba de la hipótesis nula. Se sugiere que si la serie original presenta tendencia, es recomendable incluir en la regresión el término independiente y de tendencia lineal, si la serie no la refleja y tiene un valor medio distinto de cero, es conveniente incluir el término constante. Si por el contrario la serie fluctúa en torno al valor medio cero, no es imprescindible que se incorpore ningún regresor adicional en la ecuación.

- Prueba Phillips-Perron (PP)

La prueba PP se basa en un método no-paramétrico para controlar por el problema de correlación serial de orden elevado en una serie, esto es, de un proceso  $AR(p)$  para  $p > 1$ . En particular, a diferencia de la prueba DFA que resuelve el problema por medio de incluir términos de rezago de la primera diferencia de la variable dependiente, la PP incluye un factor de corrección estimando un estadístico de prueba diferente al utilizado en la prueba DFA. Consideremos el proceso autorregresivo  $y_t = \mu + \rho y_{t-1} + u_t$ . Supongamos que  $u_t$  presenta autocorrelación serial. Una prueba en el espíritu de la DFA utiliza un estadístico de prueba como  $T(\hat{\rho}-1)$ , mientras que la prueba PP emplea el estadístico  $Z\hat{\rho}_\mu = T(\hat{\rho}-1) - CF$ , donde  $CF$  es un factor de corrección. De tal forma que el estadístico modificado incorpora el posible problema de autocorrelación. Al igual que en las dos pruebas anteriores, aquí se debe determinar la prudencia de incluir una constante, o tendencia determinística.

- Prueba KPSS

A diferencia de las pruebas DF, DFA, y PP que toman como hipótesis nula la no-estacionariedad de las series, esto es  $H_0 : \alpha = 0$ , la prueba KPSS considera como hipótesis nula la estacionariedad de la serie, así que la hipótesis nula se refiere a  $H_0 = \alpha < 1$ . Algunos autores argumentan que esta prueba pudiera ser más exigente, y por ende más potente, que las pruebas que plantean la no-estacionariedad como hipótesis nula. Sin embargo, no existe un consenso en cuanto a la potencia relativa de las diferentes pruebas que se encuentran disponibles.<sup>5</sup>

### **3. 3. Pruebas de Raíz Unitaria sobre Series que Presentan Corte Estructural**

En 1989 Perron reconoció que algunas series de tiempo presentan un comportamiento estocástico que se caracteriza por la presencia de cortes estructurales. En tal caso, Perron argumenta que las pruebas convencionales de raíz unitaria no reportan un poder de prueba satisfactorio, por lo cual es necesario instrumentar alguna prueba que tome en cuenta la existencia de cortes estructurales. Al respecto se encuentra en la literatura una amplia gama de pruebas, dentro de las que podemos mencionar aquellas que identifican el corte o los cortes estructurales exógenamente, y otros que lo, o los, determinan endógenamente. Perron (1989) presenta un modelo que se refiere al primer tipo, Zivot y Andrews (1992) al segundo. Aquí presentamos un modelo en el espíritu de Perron (1989).

---

<sup>5</sup> Le teoría puntual sobre ésta y otras pruebas estructurales de raíz unitaria escapan el alcance del texto. Para una discusión detallada sobre las mismas se recomienda Patterson (2000).

Comenzamos por ilustrar un ejemplo de serie de tiempo que presenta corte estructural. Refiriéndonos a la Gráfica 2.1, que ilustra el PIB de México, se puede apreciar un cambio en la tendencia de la serie en 1982 y otro en 1994, lo cual no debe ser sorprendente ya que corresponden a los años de crisis económicas más acentuadas en el país. En este caso se recomendaría instrumentar una prueba que incluye el corte. Una posibilidad, que sigue de cerca las propuestas en Perron (1989), es estimar una ecuación como ésta:

$$\Delta y_t = \mu + \alpha y_{t-1} + \beta t + \phi_1 DVTB_t + \phi_2 DVU_t + \phi_3 DVT_t + \sum_{j=1}^p \delta_j \Delta y_{t-j} + \varepsilon_t \quad (3.12)$$

Donde  $DVTB_t = 1$  si  $t = T_b + 1$  e igual a 0 si  $t \neq T_b + 1$ ,  $DVU_t = 1$  si  $t > T_b$  e igual a 0 si  $t \leq T_b$ ,  $DVT_t = t$  si  $t > T_b$  e igual a 0 si  $t \leq T_b$ ; siendo  $T_b$  la fecha del corte estructural. El término  $DVTB_t$  controla por un cambio en el nivel de la serie en el periodo  $T_b + 1$ ,  $DVU_t$  por un cambio en la pendiente de la serie en  $T_b + 1$ , y  $DVT_t$  un cambio en la tasa de crecimiento de la pendiente. La prueba de raíz unitaria en este caso se llevaría a cabo considerando la hipótesis nula  $H_0 : \alpha = 0$ . Nótese que en esta prueba el corte estructural se determina exógenamente, esto es, se presume que se conoce  $T_b$ , lo cual puede presentar algunos problemas de inferencia, de acuerdo a lo señalado por Zivot y Andrews (1992), quienes argumentan que la elección de la fecha del corte se encuentra correlacionada necesariamente con los datos. Sin embargo, se debe mencionar que en el ejercicio práctico de identificación de raíz unitaria, pruebas que consideran corte estructural se emplean fundamentalmente para corroborar los resultados que se encuentran mediante la aplicación de pruebas convencionales.

### 3. 4. Ejercicio Práctico de Raíces Unitarias

En este apartado presentamos un ejercicio práctico para determinar la existencia de raíz unitaria, aplicando las pruebas anteriormente mencionadas. En particular, se utiliza como variable de análisis el PIB de México. Para propósitos del desarrollo del ejercicio se trabaja con la serie en logaritmos (LPIB). El análisis se inicia aplicando la prueba de DF sin considerar intercepto o tendencia. Los resultados se reportan en el Cuadro 3.1. Se observa que, en términos absolutos, el estadístico de prueba (1.441681) es menor que los valores críticos de MacKinnon al 1%, 5% y 10%, por lo que no se rechaza la hipótesis nula, esto es, la variable presenta raíz unitaria.

**Cuadro 3.1.**  $\Delta Y_t = \alpha Y_{t-1} + u_t$

Null Hypothesis: LPIB has a unit root					
Exogenous: None					
Lag Length: 0 (Automatic based on SIC, MAXLAG=0)					
			t-Statistic	Prob.*	
Augmented Dickey-Fuller test statistic			1.441681	0.9623	
Test critical values:	1% level		-2.587387		
	5% level		-1.943943		
	10% level		-1.614694		
*MacKinnon (1996) one-sided p-values.					
Augmented Dickey-Fuller Test Equation					
Dependent Variable: D(LPIB)					
Method: Least Squares					
Date: 05/29/06 Time: 12:09					
simple (adjusted): 1980:2 2006:1					
Included observations: 104 after adjustments					
	Variable	Coefficient	Std. Error	t-Statistic	Prob.
	LPIB(-1)	0.000296	0.000205	1.441681	0.1524
R-squared		-0.000233	Mean dependent var		0.006232
Adjusted R-squared		-0.000233	S.D. dependent var		0.043825
S.E. of regresión		0.043830	Akaike info criterion		-3.407423
Sum squared resid		0.197871	Schwarz criterion		-3.381996
Log likelihood		178.1860	Durbin-Watson stat		3.332869

Los resultados del Cuadro 3.2, corresponden a una regresión en donde sólo se introduce el intercepto, e igualmente se obtiene un estadístico de prueba menor que los valores críticos.

**Cuadro 3.2.**  $\Delta Y_t = \phi_1 + \alpha Y_{t-1} + u_t$

Null Hypothesis: LPIB has a unit root					
Exogenous: Constant					
Lag Length: 0 (Automatic based on SIC, MAXLAG=0)					
			t-Statistic	Prob.*	
Augmented Dickey-Fuller test statistic			-0.882478	0.7903	
Test critical values:	1% level		-3.494378		
	5% level		-2.889474		
	10% level		-2.581741		
*MacKinnon (1996) one-sided p-values.					
Augmented Dickey-Fuller Test Equation					
Dependent Variable: D(LPIB)					
Method: Least Squares					
Date: 05/29/06 Time: 12:14					
simple (adjusted): 1980:2 2006:1					
Included observations: 104 after adjustments					
	Variable	Coefficient	Std. Error	t-Statistic	Prob.
	LPIB(-1)	-0.019443	0.022033	-0.882478	0.3796
	C	0.413502	0.461528	0.895942	0.3724
R-squared		0.007577	Mean dependent var		0.006232
Adjusted R-squared		-0.002153	S.D. dependent var		0.043825
S.E. of regresión		0.043872	Akaike info criterion		-3.396031
Sum squared resid		0.196326	Schwarz criterion		-3.345177
Log likelihood		178.5936	F-statistic		0.778768
Durbin-Watson stat		3.293234	Prob(F-statistic)		0.379593

Los resultados del Cuadro 3.3, incorporan una variable de tendencia y se aprecia que tanto ésta como el intercepto son estadísticamente significativos, e interesantemente el estadístico de prueba es mayor en términos absolutos a cualquiera de los valores críticos. Así, encontramos que de acuerdo a los resultados de las dos primeras especificaciones se concluye que existe una raíz unitaria en el PEG de la serie del PIB, sin embargo, los resultados de la tercera especificación apuntan hacia la no existencia de una raíz unitaria, ¿cómo resolvemos esta contradicción? Notamos que en las tres especificaciones se aprecia un problema de correlación en los residuales, el estadístico Durbin-Watson es muy elevado,

por lo cual no podríamos formular ninguna conclusión razonable en base a estos resultados, de tal suerte que para atender este problema procedemos a estimar la prueba DFA.

**Cuadro 3.3.**  $\Delta Y_t = \phi_1 + \phi_2 t + \alpha Y_{t-1} + u_t$

Null Hypothesis: LPIB has a unit root					
Exogenous: Constant, Linear Trend					
Lag Length: 0 (Automatic based on SIC, MAXLAG=0)					
			t-Statistic	Prob.*	
Augmented Dickey-Fuller test statistic			-5.038217	0.0004	
Test critical values:	1% level		-4.048682		
	5% level		-3.453601		
	10% level		-3.152400		
*MacKinnon (1996) one-sided p-values.					
Augmented Dickey-Fuller Test Equation					
Dependent Variable: D(LPIB)					
Method: Least Squares					
Date: 05/29/06 Time: 12:18					
simple (adjusted): 1980:2 2006:1					
Included observations: 104 after adjustments					
	Variable	Coefficient	Std. Error	t-Statistic	Prob.
	LPIB(-1)	-0.400492	0.079491	-5.038217	0.0000
	C	8.260828	1.638804	5.040766	0.0000
	@TREND(1980:1)	0.002560	0.000517	4.950655	0.0000
	R-squared	0.201374	Mean dependent var		0.006232
	Adjusted R-squared	0.185560	S.D. dependent var		0.043825
	S.E. of regresión	0.039550	Akaike info criterion		-3.594057
	Sum squared resid	0.157988	Schwarz criterion		-3.517776
	Log likelihood	189.8910	F-statistic		12.73362
	Durbin-Watson stat	<b>2.703309</b>	Prob(F-statistic)		0.000012

Los resultados de la prueba DFA que incluye una constante y una tendencia en su especificación se presentan en el Cuadro 3.4. En este caso se determinó que el número de rezagos óptimos en  $\Delta y_{t-i}$ , es igual a 1. Se nota que el estadístico de prueba (-2.244557) es inferior a los valores críticos en términos absolutos; por lo que se puede concluir que la serie del PIB no es estacionaria. Asimismo ya no se aprecia un problema importante de autocorrelación.

**Cuadro 3.4.**  $\Delta Y_t = \phi_1 + \phi_2 t + \alpha Y_{t-1} + \sum_{i=1}^P \delta_i \Delta Y_{t-i} + u_t$

Null Hypothesis: LPIB has a unit root					
Exogenous: Constant, Linear Trend					
Lag Length: 1 (Automatic based on SIC, MAXLAG=1)					
Augmented Dickey-Fuller test statistic		-2.244557		0.4599	
Test critical values:	1% level	-4.049586			
	5% level	-3.454032			
	10% level	-3.152652			
*MacKinnon (1996) one-sided p-values.					
Augmented Dickey-Fuller Test Equation					
Dependent Variable: D(LPIB)					
Method: Least Squares					
Date: 05/29/06 Time: 12:23					
Simple (adjusted): 1980:3 2006:1					
Included observations: 103 after adjustments					
	Variable	Coefficient	Std. Error	t-Statistic	Prob.
	LPIB(-1)	-0.163896	0.073019	-2.244557	0.0270
	D(LPIB(-1))	-0.589192	0.081627	-7.218117	0.0000
	C	3.385374	1.505140	2.249208	0.0267
	@TREND(1980:1)	0.001100	0.000472	2.331219	0.0218
R-squared	0.476732	Mean dependent var		0.006320	
Adjusted R-squared	0.460875	S.D. dependent var		0.044030	
S.E. of regresión	0.032329	Akaike info criterion		-3.987638	
Sum squared resid	0.103472	Schwarz criterion		-3.885318	
Log likelihood	209.3634	F-statistic		30.06521	
Durbin-Watson stat	1.776941	Prob(F-statistic)		0.000000	

Seguimos con las pruebas PP y KPSS considerando únicamente la especificación que incluye constante y tendencia (aquí sólo reportamos los estadísticos). En el Cuadro 3.5 presentamos el resumen de los resultados.

### Cuadro 3.5. Prueba PP y Prueba KPSS

Null Hypothesis: LOG(PIB) has a unit root		
Exogenous: Constant, Linear Trend		
Bandwidth: 9 (Newey-West using Bartlett kernel)		
	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-5.708493	0.0000
Test critical values:	1% level	-4.049586
	5% level	-3.454032
	10% level	-3.152652
*MacKinnon (1996) one-sided p-values.		

Null Hypothesis: LOG(PIB) is stationary		
Exogenous: Constant, Linear Trend		
Bandwidth: 8 (Newey-West using Bartlett kernel)		
		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.198339
Asymptotic critical values*:	1% level	0.216000
	5% level	0.146000
	10% level	0.119000
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)		

De acuerdo a la prueba PP la serie es estacionaria, pero de acuerdo a la prueba KPSS la serie es no-estacionaria al 5% (que es el nivel de significancia estándar). Este ejercicio ilustra puntualmente los dilemas a los que se enfrenta quien practica el análisis econométrico de series de tiempo. En este caso tenemos que de 3 pruebas robustas (las tres especificaciones DF presentan problemas de correlación residual), dos indican que la serie es no-estacionaria y una que es estacionaria. En principio podríamos decidir en base al resultado que presenta la mayoría de las pruebas, pero eso sería poco formal. Alternativamente, notando que la serie presenta un corte estructural, podríamos estimar una prueba similar a la sugerida en (3.12). Afortunadamente para nosotros, dicha prueba ya se

realizó, encontrándose que el PIB de México es en efecto una serie no-estacionaria. Castillo y Díaz-Bautista (2002) estiman especificaciones en las que se determina el corte estructural tanto exógenamente como endógenamente. Los autores muestran que al considerar los cortes estructurales que son evidentes en la serie del PIB, pruebas de raíz unitaria indican que la serie es de hecho no-estacionaria.

¿A qué se debe que se obtengan resultados contradictorios bajo diferentes especificaciones? Fundamentalmente al bajo poder de prueba que en general presentan las pruebas de raíz unitaria. Ésta ha sido una preocupación de los estudiosos del tema por años, de tal suerte que en la actualidad se encuentran varias pruebas que de alguna u otra manera tratan de mejorar su poder predictivo. En el programa econométrico EViews, por ejemplo, se ofrecen 6 pruebas. Por esto, corresponde al econometrista que realiza el análisis práctico determinar el resultado que sea más acorde con la teoría, ya que es poco útil identificar resultados que no sean intuitivamente válidos.<sup>6</sup>

Tomando el resultado de no-estacionariedad del PIB en niveles como válido, procedemos a ilustrar el concepto de integración que se discutió en la Sección 3.1. Así, aplicamos las pruebas de raíz unitaria a la primera diferencia del logaritmo del PIB, que denotamos como  $D(LPIB)$ . Los resultados de la estimación de la prueba DFA se muestran en el Cuadro 3.6. Notamos que en términos absolutos, el estadístico de prueba es superior a los valores críticos, observando también que el estadístico de Durbin – Watson es igual a 1.82. En

---

<sup>6</sup> Citamos un caso real, la serie del índice nacional de precios al consumidor de México resulta consistentemente no estacionaria, no sólo en niveles, pero también en primeras diferencias, por lo cual la serie de inflación sería una serie no-estacionaria. Esto, por supuesto, no se apega a la teoría económica estándar. Así, existe un consenso implícito que considera a la serie de inflación como estacionaria.

consecuencia, se concluye que la serie en primera diferencia es estacionaria; lo que permite considerar a la serie del PIB en niveles como integrada de orden 1, I(1), ya que se tomó una diferencia para que resultara estacionaria.

**Cuadro 3.6. Resultados en Primera Diferencia de la Prueba DFA**

Null Hypothesis: D(LPIB) has a unit root					
Exogenous: Constant, Linear Trend					
Lag Length: 0 (Automatic based on SIC, MAXLAG=0)					
			t-Statistic	Prob.*	
Augmented Dickey-Fuller test statistic			-22.50054	0.0000	
Test critical values:					
	1% level		-4.049586		
	5% level		-3.454032		
	10% level		-3.152652		
*MacKinnon (1996) one-sided p-values.					
Augmented Dickey-Fuller Test Equation					
Dependent Variable: D(LPIB,2)					
Method: Least Squares					
Date: 05/29/06 Time: 13:02					
Sample (adjusted): 1980:3 2006:1					
Included observations: 103 after adjustments					
	Variable	Coefficient	Std. Error	t-Statistic	Prob.
	D(LPIB(-1))	-1.671855	0.074303	-22.50054	0.0000
	C	0.007033	0.006647	1.058138	0.2925
	@TREND(1980:1)	6.86E-05	0.000109	0.627242	0.5319
R-squared	0.835067	Mean dependent var			-0.000152
Adjusted R-squared	0.831768	S.D. dependent var			0.080396
S.E. of regresión	0.032975	Akaike info criterion			-3.957419
Sum squared resid	0.108737	Schwarz criterion			-3.880679
Log likelihood	206.8071	F-statistic			253.1527
Durbin-Watson stat	1.824011	Prob(F-statistic)			0.000000

## Capítulo 4: Series Integradas y los Modelos ARMA

En este capítulo extendemos nuestra discusión de los modelos ARMA tomando la idea de estacionariedad e integración que recién discutimos en el Capítulo 3. Haremos referencia a un modelo que es particularmente interesante como técnica de predicción en el análisis económico y financiero; y que nos ayuda a vislumbrar escenarios prospectivos en el mediano y largo plazo. Nos referimos a los modelos autorregresivos integrados con media móvil o ARIMA. Cabe mencionar que aquí continuamos considerando un análisis para modelos univariados, las aplicaciones de los conceptos de estacionariedad e integración para modelos multivariados se abordan a partir del Capítulo 6.

### 4. 1. Definición

Como lo indicamos anteriormente, si una serie en niveles es no-estacionaria es recomendable diferenciarla para trabajar con la serie estacionaria y evitar posibles resultados espurios. En caso de que la variable  $y_t$  tenga que diferenciarse, entonces se deberá estimar un modelo ARIMA  $(p,d,q)$ , donde  $p$  denota el orden del rezago en  $y_t$ ,  $q$  el orden del rezago del promedio móvil asociado a  $\varepsilon_t$ , y  $d$  el orden de diferenciación de la serie.

En este mismo sentido, se puede decir que un ARMA  $(1,1)$  se corresponde con un ARIMA  $(1,0,1)$ . Un ARIMA  $(p,d,q)$ , cuando involucra a una serie integrada de orden  $d$ , se puede especificar también como:

$$\Delta^d y_t = \lambda_1 \Delta^d y_{t-1} + \dots + \lambda_p \Delta^d y_{t-p} + \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \dots + \varphi_p \varepsilon_{t-p} \quad (4.1)$$

Donde  $\Delta^d y_t$  indica que la serie original o en niveles,  $y_t$ , ha sido diferenciada  $d$  veces; en este mismo sentido, una serie integrada de orden dos, se expresaría como  $\Delta^2 y_t = \Delta(\Delta y_{t-1}) = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$ .

## 4.2. Tratamiento de la Estacionalidad en los Modelos ARIMA

Una serie de tiempo que incluye datos estacionales debe recibir un tratamiento adecuado en el proceso de modelación. Aquí hacemos referencia a un modelo ARIMA, que capta el comportamiento estacional en una forma muy parecida a como se realiza para el componente regular de la serie. Un proceso ARIMA estacional, como su nombre lo indica, es un proceso que muestra variaciones estacionales, de tal forma que el valor de  $y_t$  está vinculado con el valor de  $y_{t-s}$ . Por ejemplo, si los datos de la serie son trimestrales y la periodicidad de la variaciones es anual, se tendrá que  $s = 4$ , mientras que si los datos son mensuales y la periodicidad es anual, entonces,  $s = 12$ .

Como el tratamiento de la estacionalidad implica trabajar con variables retardadas, es que se retoma el término de operador de rezago, que como ya se mencionó en la teoría previa, se expresa como  $Ly_t = y_{t-1}$  y al ser aplicado en forma sucesiva para  $s$  periodos se tiene que  $L^s y_t = y_{t-s}$ . A partir de esto, un modelo ARMA  $(p, d, q)$  puede expresarse como:

$$(1 - \lambda_1 L - \dots - \lambda_p L^p)(1 - L)^d = (1 + \theta_1 L + \dots + \theta_q L^q) \varepsilon_t \quad (4.2)$$

Que resumidamente se expresa de la siguiente forma:

$$\lambda_p(L) \nabla^d Y_y = \theta_q(L) \varepsilon_t \quad (4.3)$$

Donde

$$\nabla = (1 - L)$$

$\lambda_p(L)$  = polinomio de orden  $p$  en  $L$  con parámetro  $\lambda_i$

$\theta_q(L)$  = polinomio de orden  $q$  en  $L$  con parámetro  $\theta_i$

Algunos modelos sencillos que son frecuentemente empleados para caracterizar datos estacionales son:

$$\text{AR}(1): y_t = \mu + \lambda_1 y_{t-1} + \varepsilon_t \quad (4.4)$$

$$\text{MA}(1): y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} \quad (4.5)$$

Los cuales, deben incorporar un regresor adicional según la frecuencia de la estacionalidad.

Para el caso de datos trimestrales los autorregresivos y media móvil con componente estacional se expresan como:

$$\text{SAR}(1): y_t = \mu + \lambda_1 y_{t-s} + \varepsilon_t \quad s = 4 \text{ datos trimestrales} \quad (4.6)$$

$$\text{SMA}(1): y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-s} \quad s = 4 \text{ datos trimestrales} \quad (4.7)$$

Por tanto, cuando una serie presenta un componente estacional, significa que se debe llevar a cabo una o varias diferencias de orden estacional. En el caso particular de que se tuviera una serie con datos mensuales, se requeriría efectuar el siguiente cálculo:

$$\nabla_{12} y_t = (1 - L^{12}) y_t = y_t - y_{t-12} \quad (4.8)$$

$$\nabla_{12}^2 y_t = (1 - L^{12})^2 y_t = (y_t - y_{t-12}) - (y_{t-12} - y_{t-24}) \quad (4.9)$$

En este sentido, un modelo SARMA (1,1), donde S hace referencia al componente estacional, se expresa como:

$$\lambda_1^{12}(L) y_t = \theta_1^{12}(L) \varepsilon_t \quad (4.10)$$

Que a su vez, se puede plantear como

$$(1 - \lambda_{12} L^{12}) y_t = (1 + \theta_{12} L^{12}) \varepsilon_t \quad (4.11)$$

De modo que despejando  $y_t$  se tiene:

$$y_t = \lambda_{12} y_{t-12} + \varepsilon_t + \theta_{12} \varepsilon_{t-12} \quad (4.12)$$

Con base en estas consideraciones, si se tiene una serie  $y_t$  que no es estacionaria en niveles, entonces se debe trabajar con una serie transformada en diferencias de logaritmos,  $dly_t$ , para asegurar que no tenga tendencia en media y varianza. Subsecuentemente se debe correr el correlograma de la serie con una cierta longitud de retardos y analizar los coeficientes de autocorrelación simple y parcial. Si por ejemplo, se tiene una serie mensual y se aprecian en la función de autocorrelación coeficientes significativos que se reproducen en ciertos retardos con frecuencia estacional, (12, 24, 36), es un indicio de que existe tendencia en la parte estacional que habrá que corregir adecuadamente. Lo anterior requiere que se genere la serie con una primera diferencia en la parte estacional a través de la operación  $ddly_t = dly_t - dly_{t-12}$ . Una vez que se tiene esta serie, se vuelve a obtener el correlograma para identificar el modelo más apropiado. Si por ejemplo, el modelo identificado fuese un ARMA (1,1)\* AR<sub>12</sub>(1), es decir, un modelo autorregresivo de orden 1 y una medida móvil de orden 1 en la parte regular y una autorregresivo de orden 1 en la parte estacional, se determina que el modelo a estimar es  $ddly_t$  AR(1) MA(1) SAR(12). A continuación presentamos un ejemplo para aclarar ideas. Primero presentamos un modelo ARIMA que no considera patrón estacional, en seguida se incluye un ejercicio para un ARIMA estacional.

#### **4.3. Ejemplo para Modelo ARIMA: Índice de Volumen Físico de la Producción Manufacturera**

En esta sección se desarrolla un ejercicio práctico para ilustrar la metodología de los modelos univariados ARIMA. La variable elegida es el Índice del Volumen Físico de la Producción Manufacturera (IVFPM) con base 1993=100 correspondiente al periodo

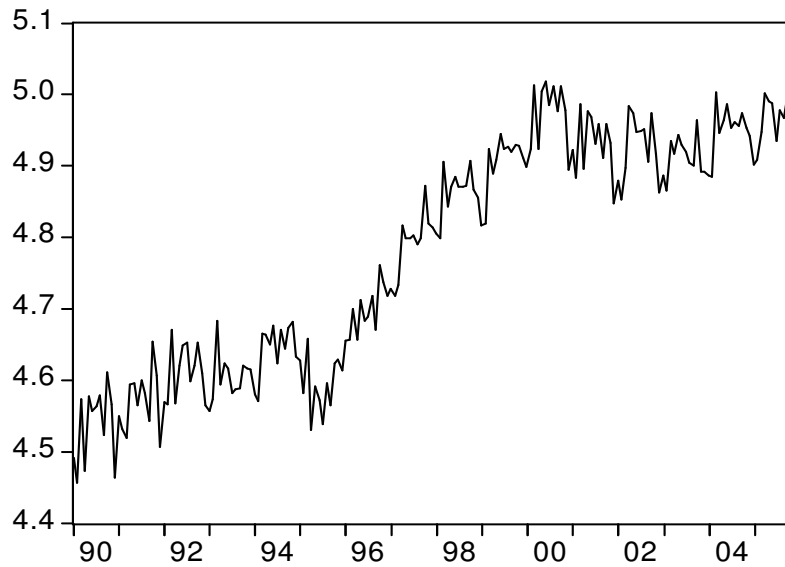
1990:01 – 2005:12. Los datos fueron obtenidos del Banco de Información Económica del Instituto Nacional de Estadística, Geografía e Informática (INEGI). Aquí incluimos una descripción detallada de los pasos a seguir en EViews.

Previo a la estimación y validación del modelo  $ARIMA(p,d,q)$ , primeramente transformamos la serie a logaritmos. Para realizar esta tarea, se utiliza el comando *Genr*, el cual conduce a un cuadro de diálogo en el que se debe especificar la instrucción  $LIVFPM=LOG(IVFPM)$ . Una vez realizado dicho procedimiento, se procede a graficar la serie para efectuar una primera exploración de su comportamiento y determinar su estacionariedad. En la Gráfica 4.1 se puede observar que la serie manifiesta un comportamiento diferenciado pero ascendente a lo largo del periodo de estudio. Si bien se nota un movimiento creciente, que desciende en 1995, también se aprecia que después del año 2000 se comporta relativamente estable hasta el año 2005. Por tanto, se puede decir que el comportamiento de LIVFPM se caracteriza por presentar tres tramos perfectamente diferenciados, pero que no obstante a ello, se identifica una clara tendencia positiva.

Después de realizar esta primera aproximación, analizamos su correlograma. En este ejercicio tenemos que el número de observaciones es 192, y se opta por especificar 48 rezagos. Al observar el correlograma de LIVFPM ilustrado en la Figura 4.1, distinguimos que la función de autocorrelación decrece exponencialmente y de forma lenta; mientras que la función de autocorrelación arroja un valor estadísticamente significativo en el rezago número uno, pues se aprecia claramente como la primer barra de la columna PAC supera los límites de confianza que permite interpretar al coeficiente de autocorrelación como no

significativo. En este sentido, parece haber elementos que permiten anticipar que la serie LIVFPM en niveles no es estacionaria.

**Gráfica 4.1. Logaritmo del Índice de Volumen Físico de la Producción Manufacturera, (1993=100)**



**Figura 4.1. Correlograma de LIVFPM**

Sample: 1990:01 2005:12  
 Included observations: 192

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.938	0.938	171.64	0.000
		2	0.914	0.285	335.52	0.000
		3	0.893	0.111	492.65	0.000
		4	0.859	-0.077	638.86	0.000
		5	0.851	0.165	783.12	0.000
		6	0.826	-0.043	919.73	0.000
		7	0.813	0.079	1052.9	0.000
		8	0.791	-0.079	1179.6	0.000
		9	0.770	0.019	1300.4	0.000
		10	0.756	0.003	1417.3	0.000
		11	0.745	0.106	1531.6	0.000
		12	0.744	0.087	1646.0	0.000
		13	0.707	-0.255	1750.1	0.000
		14	0.690	-0.004	1849.6	0.000
		15	0.660	-0.107	1941.2	0.000
		16	0.641	0.120	2028.2	0.000
		17	0.639	0.085	2115.0	0.000
		18	0.608	-0.126	2194.0	0.000
		19	0.605	0.079	2272.9	0.000
		20	0.585	-0.026	2347.1	0.000
		21	0.562	-0.017	2415.8	0.000
		22	0.551	-0.009	2482.4	0.000
		23	0.538	0.047	2546.2	0.000
		24	0.531	-0.021	2608.7	0.000
		25	0.500	-0.123	2664.5	0.000
		26	0.484	0.026	2717.2	0.000
		27	0.457	-0.061	2764.3	0.000
		28	0.437	0.006	2807.7	0.000
		29	0.428	0.011	2849.6	0.000
		30	0.400	-0.048	2886.4	0.000
		31	0.396	0.033	2922.6	0.000
		32	0.368	-0.065	2954.1	0.000
		33	0.341	-0.064	2981.3	0.000
		34	0.329	0.031	3006.8	0.000
		35	0.307	0.017	3029.3	0.000
		36	0.300	-0.008	3050.7	0.000
		37	0.270	-0.091	3068.2	0.000
		38	0.250	0.004	3083.3	0.000
		39	0.227	-0.020	3095.8	0.000
		40	0.211	0.070	3106.7	0.000
		41	0.199	-0.023	3116.5	0.000
		42	0.178	-0.018	3124.4	0.000
		43	0.171	-0.039	3131.7	0.000
		44	0.142	-0.028	3136.8	0.000
		45	0.123	-0.001	3140.6	0.000
		46	0.112	0.034	3143.9	0.000
		47	0.094	0.002	3146.1	0.000
		48	0.089	-0.032	3148.2	0.000

Para determinar categóricamente la naturaleza estocástica de la serie, se realizan algunas de las pruebas de raíz unitaria que se discutieron en el Capítulo 3. En el Cuadro 4.1, se reportan los resultados de las pruebas de Dickey - Fuller y Dickey - Fuller - Aumentada. Los resultados indican que la serie en niveles es no-estacionaria aún cuando la regresión 2 arroja un estadístico de prueba,  $\tau$ , superior a los valores críticos, pues el valor de Durban-Watson indica la posible existencia de autocorrelación. A partir de esto, se realiza la prueba de Dickey Fuller Aumentada con uno y dos rezagos, y en ambos casos se concluye que la serie no es estacionaria. Al realizar la prueba para la primera diferencia de la serie (D(LIVFPM)), considerando constante, tendencia y un rezago, se concluye que la serie sí es estacionaria. De tal forma que concluimos que la serie es integrada de orden 1.

**Cuadro 4.1. Resultados de las Pruebas de Estacionariedad**

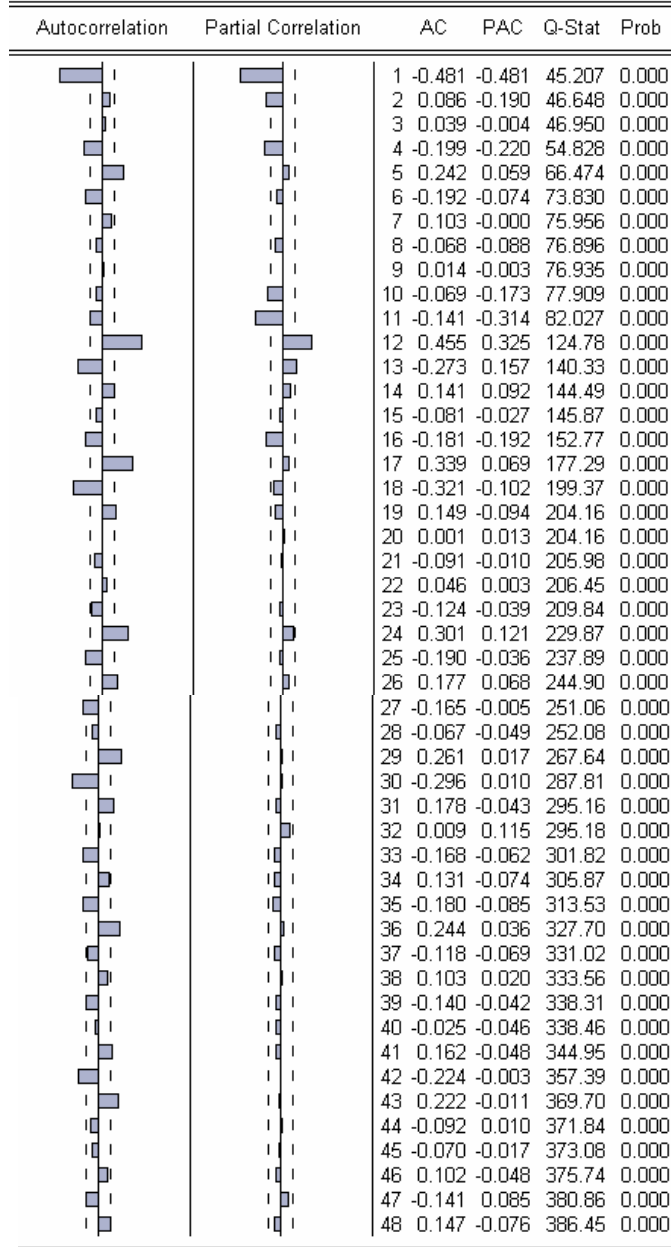
Regresión	Regresión	Estadísticos DF* y DFA**	Valores críticos de MacKinnon (1%, 5% y 10%) respectivamente	Ho: raíz unitaria
1	LIVFPM(C)	-2.292351	-3.464643 -2.876515 -2.574831	se acepta DW = 2.87
2	LIVFPM (C,T)	-5.499604	-4.006824 -3.433525 -3.140623	se rechaza DW =2.56
3	LIVFPM(C,T,R=1)	-3.319301	-4.007084 -3.433651 -3.140697	se acepta DW =2.0 Criterio de Akaike = -3.504695
4	LIVFPM(C,T,R=2)	-2.578015	-4.007347 -3.433778 -3.140772	Se acepta DW = 1.95 Criterio de Akaike = -3.545186
5	D(LIVFPM), (C,T)	-23.22104	-4.007084 -3.433651 -3.140697	se rechaza DW = 2.16
6	D(LIVFPM)(C,T,R=1)	-14.80325	-4.007347 -3.433778 -3.140772	se rechaza DW = 1.97

C= constante, T= tendencia, R = retardo (1,2,...,k), T = tendencia, 1 = primera diferencia

Una vez que se ha determinado que la serie DLIVFPM es estacionaria, se genera su correlograma para identificar el tipo de modelo que se tiene que estimar. Primero se estima el correlograma que se presenta en la Figura 4.2.

**Figura 4.2. Correlograma de DLIVFPM**

Sample: 1990:01 2005:12  
Included observations: 192



Aunque los coeficientes de autocorrelación simple y parcial indican que se pueden estimar distintos modelos  $ARIMA(p,d,q)$ ; se procede a estimar una ecuación autoregresiva bajo la siguiente instrucción en EViews, LS DLIVFPM AR(1) AR(12). Si se observa, la función de autocorrelación refleja un decrecimiento que combina coeficientes positivos y negativos en forma sinusoidal, acompañada de una función de autocorrelación parcial con picos en los rezagos  $k = 1$  y  $k = 12$ . Los resultados obtenidos de dicha estimación son los siguientes:

**Cuadro 4.2. Resultados de Regresión**

Dependent Variable: DLIVFPM				
Method: Least Squares				
Date: 08/04/06 Time: 19:26				
Sample (adjusted): 1991:01 2005:12				
Included observations: 180 after adjustments				
Convergence achieved after 2 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.383936	0.057394	-6.689462	0.0000
AR(12)	0.424132	0.054253	7.817733	0.0000
R-squared	0.410939	Mean dependent var		0.002758
Adjusted R-squared	0.407630	S.D. dependent var		0.046382
S.E. of regression	0.035698	Akaike info criterion		-3.816398
Sum squared resid	0.226833	Schwarz criterion		-3.780920
Log likelihood	345.4758	Durbin-Watson stat		2.336226
Inverted AR Roots	.90	.78+.46i	.78-.46i	.44-.80i
	.44+.80i	-.03-.93i	-.03+.93i	-.50-.80i
	-.50+.80i	-.84+.46i	-.84-.46i	-.97

Los coeficientes de ambos autorregresivos son estadísticamente significativos y reportan valores inferiores a la unidad que reflejan cambios en términos de elasticidades. Por su parte, el estadístico Durban-Watson igual a 2.3, es ligeramente superior al valor de referencia que es 2, lo que sugiere que no existe un problema de autocorrelación importante. Sin embargo, para verificar que dicho modelo es idóneo para realizar

predicciones a futuro, se tiene que hacer un diagnóstico del mismo analizando el correlograma de los coeficientes de los residuales del modelo. Este correlograma se presenta en la Figura 4.3.

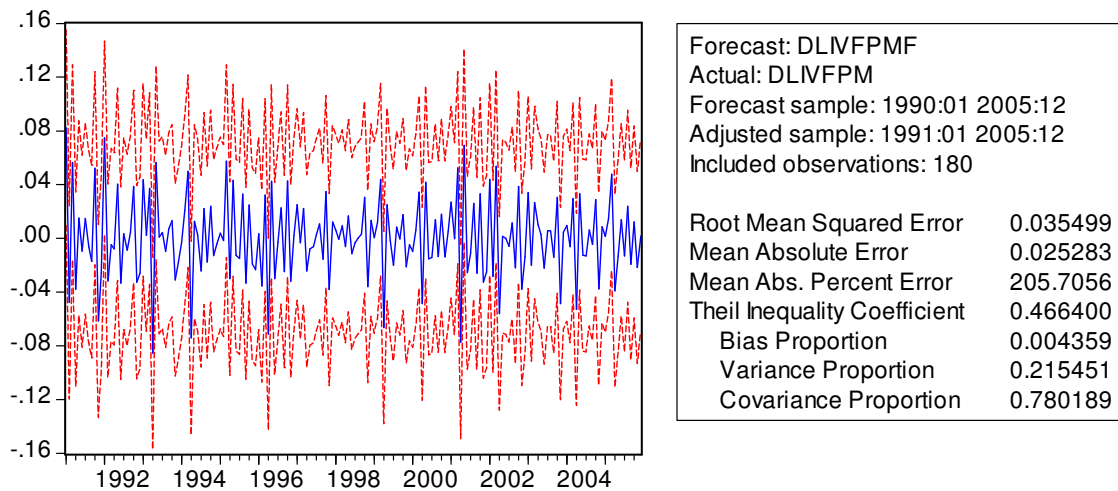
**Figura 4.3. Correlograma de los Residuales del Modelo Estimado**

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.174	-0.174	5.5729	0.018
		2	-0.128	-0.163	8.5829	0.014
		3	0.008	-0.049	8.5949	0.035
		4	-0.043	-0.076	8.9314	0.063
		5	0.071	0.044	9.8634	0.079
		6	-0.008	-0.001	9.8759	0.130
		7	-0.055	-0.043	10.451	0.164
		8	-0.007	-0.030	10.460	0.234
		9	-0.051	-0.073	10.965	0.279
		10	-0.146	-0.197	15.037	0.131
		11	0.096	-0.001	16.827	0.113
		12	0.021	-0.012	16.912	0.153
		13	0.110	0.131	19.272	0.115
		14	-0.031	0.018	19.467	0.148
		15	-0.097	-0.049	21.340	0.126
		16	-0.133	-0.211	24.879	0.072
		17	0.133	0.021	28.409	0.040
		18	-0.056	-0.120	29.033	0.048
		19	-0.010	-0.039	29.054	0.065
		20	-0.006	-0.053	29.063	0.087
		21	-0.055	-0.019	29.676	0.099
		22	-0.003	-0.058	29.678	0.126
		23	0.103	0.114	31.895	0.102
		24	0.031	0.016	32.100	0.125
		25	0.057	0.060	32.798	0.136
		26	0.042	0.021	33.166	0.157
		27	-0.092	-0.010	34.971	0.140
		28	-0.039	-0.106	35.304	0.161
		29	0.090	0.106	37.055	0.145

En razón de que las probabilidades asociadas al estadístico  $Q$  son mayores que 0.05, salvo en los primeros rezagos del correlograma, se concluye que los coeficientes de autocorrelación no son significativos, lo cual se puede corroborar con las barras ilustradas en el propio correlograma. Este procedimiento de diagnóstico constituye un método sencillo para evaluar si los residuales son o no puramente aleatorios. Después de observar los resultados presentados en la Figura 4.3, se determina que el modelo con dos autorregresivos [AR(1) y AR(12)], es una opción apropiada para efectuar pronósticos.

Aunado a esto, se evalúa la capacidad predictiva del modelo bajo otros criterios estadísticos utilizando en EViews el comando *Forecast* ubicado en el menú de la ventana de la ecuación de regresión. Los estadísticos se reportan en la Gráfica 4.2.

**Gráfica 4.2. Evaluación de la Capacidad Predictiva a través de un Método Estático**



Cabe señalar que estos resultados son producto de estimar el modelo a través de un método de pronóstico estático que se eligió en el cuadro de diálogo que produce la activación del comando *Forecast*. Es importante destacar que la predicción estática es una predicción paso a paso y la dinámica es en cadena. En el primer caso, se utilizan los valores verdaderos de las variables rezagadas para efectuar el pronóstico mientras que la segunda opción, se basa en los valores estimados. La característica es que las predicciones coincidirán en el primer periodo, pero a partir del segundo los valores serán diferentes. Conviene puntualizar que si DLIVFPM es la variable analizada, entonces al utilizar el comando *Forecast*, se genera la variable pronosticada denotada como DLIVFPMF, donde la última letra *F*, hace referencia justamente al pronóstico. En la medida que los estadísticos de la Gráfica 4.2 se aproximen a cero, significa que la capacidad predictiva del modelo es

satisfactoria. Los resultados sugieren que no hay sesgo, y la raíz de error cuadrático medio y el error absoluto medio son realmente bajos. Además el coeficiente de desigualdad de Theil que oscila entre cero y uno, no reporta un valor muy cercano a cero pero tampoco a uno. En este sentido, todo indica que el modelo estimado previamente es una opción aceptable para efectuar pronósticos fuera de la muestra.

Como se puede constatar, los resultados presentados en la Figura 4.3 y la Gráfica 4.2 son complementarios y ayudan a fortalecer la interpretación del modelo en términos de su capacidad de predicción fuera de la muestra. También es importante señalar, que los estadísticos resumidos al lado derecho de la Gráfica 4.2, ayudan a elegir entre dos modelos o más según el valor numérico que arrojen. En este caso, se ha optado por describir los mejores resultados de estimación con respecto a otros modelos alternativos que también se estimaron y cuyos resultados fueron menos alentadores. En cualquier caso, el objetivo de este apartado, ha sido en todo momento ilustrar la teoría con un ejercicio práctico.

Finalmente en el Cuadro 4.3 se resumen los pronósticos para la variable LIVFPM e IVFPM correspondientes al periodo 2006:01 – 2006:03 considerando también los últimos valores de la serie original para efecto de comparación. Para obtener directamente los valores pronosticados de LIVFPM, se estimo el mismo modelo pero especificando en EViews la instrucción  $LS\ D(LIVFPM)\ AR(1)\ AR(12)$ .

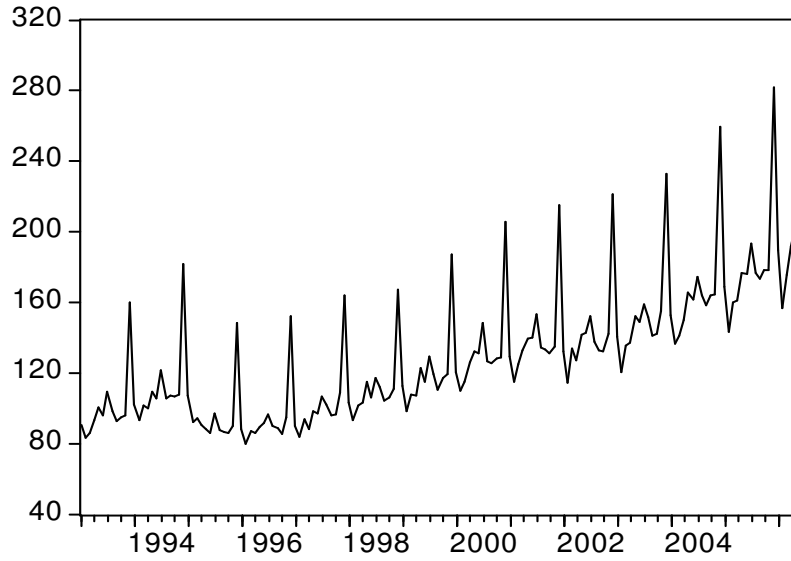
**Cuadro 4.3. Pronósticos**

<b>Periodo</b>	<b>LIVFPM</b>	<b>IVFPM</b>
2005:10	4.999237	148.3000
2005:11	4.982236	145.8000
2005:12	4.960044	142.6000
2006:01	4.951891	141.4421
2006:02	4.957523	142.2410
2006:03	4.972224	144.3475

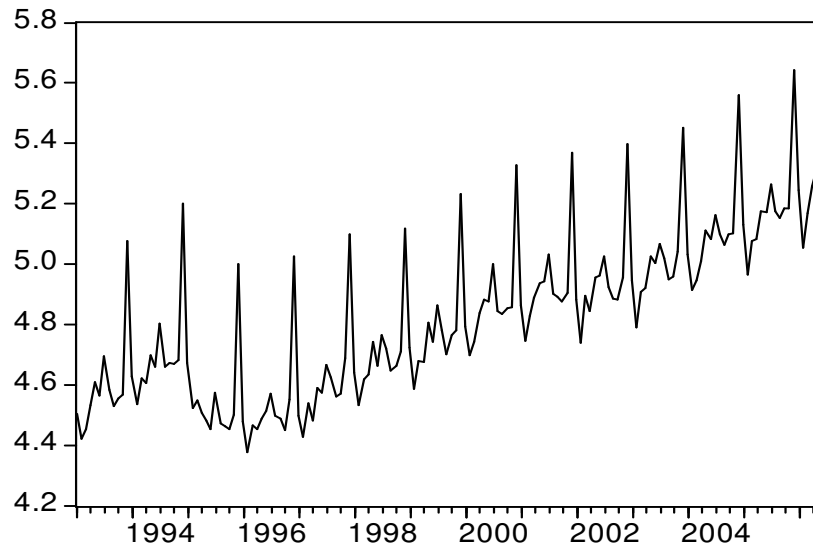
**Ejemplo de modelo ARIMA para serie con patrón estacional: Ventas de la ANTAD**

En este apartado se especifica y estima un modelo univariado para las ventas de la Asociación Nacional de Tiendas de Autoservicio y Departamentales, A.C (ANTAD) para el periodo 1993:01 – 2006:06. Los datos se obtuvieron del Sistema de Información Económica de Banco de México. La particularidad de esta serie de tiempo es que presenta un patrón estacional que requiere un tratamiento específico en el contexto de la metodología ARIMA. En la Gráfica 4.3 se describe la trayectoria de la serie y se puede ver cómo refleja una mayor dispersión en los últimos años así como un marcado patrón estacional y de tendencia. En la Gráfica 4.4, se describe la trayectoria de la serie transformada en logaritmos, la cual muestra una menor dispersión en el último tramo del periodo. Conviene advertir que aun así, la serie en logaritmos (LANTAD) no elimina la tendencia por lo que se procede a diferenciarla para efecto de que sea estacionaria y subsecuentemente se pueda modelar.

**Gráfica 4.3. Ventas de la ANTAD**



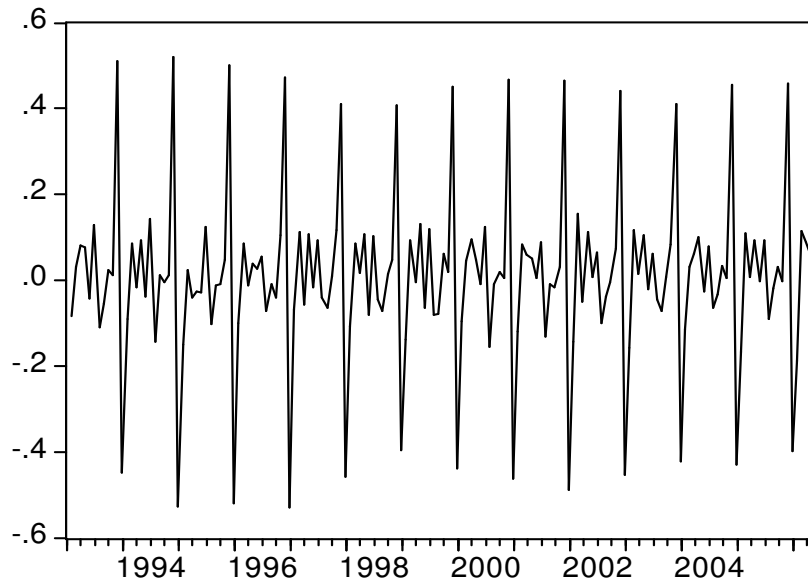
**Gráfica 4.4. Logaritmo de las ventas de la ANTAD**



Para obtener la primera diferencia de la serie en logaritmos, también se utiliza el cuadro de diálogo del comando *Genr*, en el cual se especifica  $DLVENTAS=LVENTAS-LVENTAS(-1)$ , donde D hace referencia a la primera diferencia y (-1) a un rezago de primer orden. La

Gráfica 4.5 muestra como la tendencia ha sido totalmente eliminada y los valores varían alrededor de cero.

**Gráfica 4.5. Primera diferencia del logaritmo de las ventas de la ANTAD**



Hasta aquí se puede plantear que la primera diferencia del logaritmo de las ventas de la ANTAD es una serie estacionaria, razón por la cual se elige para estudiarse en el marco de un modelo ARIMA  $(p,d,q)$  que considera un componente estacional. Al aplicar la prueba de estacionariedad DFA sobre la serie en diferencias, y cuyos resultados se reporta en el Cuadro 4.4, se determina que es estacionaria, pues en términos absolutos el estadístico de prueba  $\tau$  excede a los tres valores críticos correspondientes al 1%, 5% y 10% de significancia estadística.

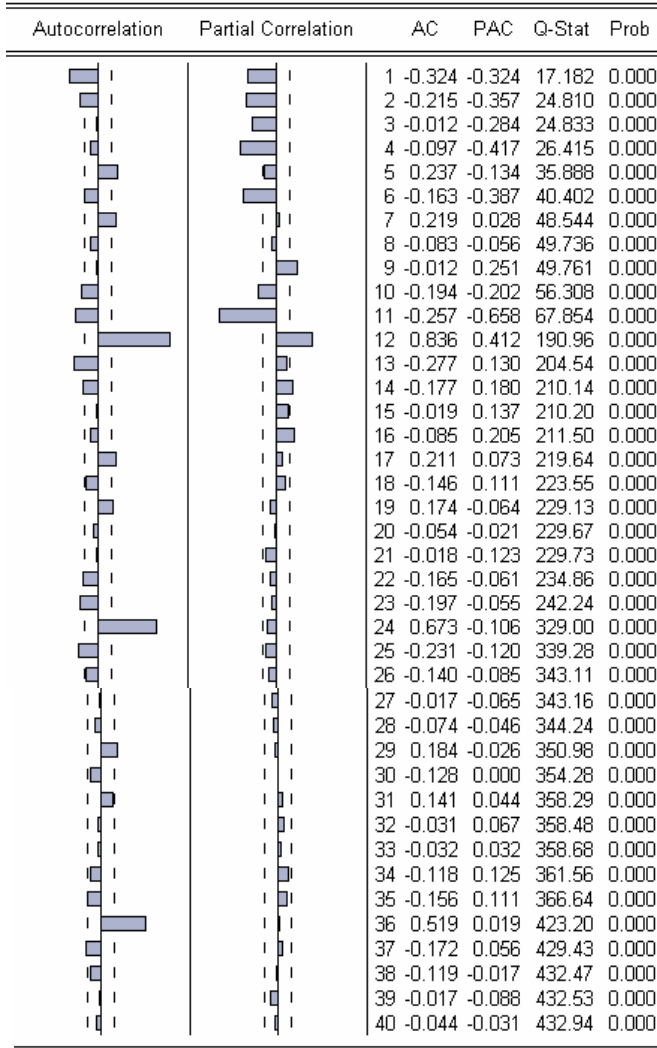
#### Cuadro 4.4. Prueba de Estacionariedad

Null Hypothesis: D(LANTAD) has a unit root					
Exogenous: Constant, Linear Trend					
Lag Length: 0 (Automatic based on SIC, MAXLAG=0)					
			t-Statistic	Prob.*	
Augmented Dickey-Fuller test statistic			-17.54148	0.0000	
Test critical values:	1% level		-4.016433		
	5% level		-3.438154		
	10% level		-3.143345		
*MacKinnon (1996) one-sided p-values.					
Augmented Dickey-Fuller Test Equation					
Dependent Variable: D(LANTAD,2)					
Method: Least Squares					
Date: 08/09/06 Time: 11:56					
Sample (adjusted): 1993:03 2006:06					
Included observations: 160 after adjustments					
	Variable	Coefficient	Std. Error	t-Statistic	Prob.
	D(LANTAD(-1))	-1.323735	0.075463	-17.54148	0.0000
	C	0.004562	0.030554	0.149324	0.8815
	@TREND(1993:01)	3.28E-05	0.000326	0.100437	0.9201
	R-squared	0.662153	Mean dependent var		0.000504
	Adjusted R-squared	0.657850	S.D. dependent var		0.325760
	S.E. of regresión	0.190549	Akaike info criterion		-0.459246
	Sum squared resid	5.700481	Schwarz criterion		-0.401587
	Log likelihood	39.73969	F-statistic		153.8540
	Durbin-Watson stat	2.231600	Prob(F-statistic)		0.000000

A partir de las consideraciones anteriores, a continuación se obtiene el correlograma de la serie DLVENTAS con  $k = 40$  tomando en cuenta que  $T = 161$ . Los resultados se ilustran en la Figura 4.4.

**Figura 4.4. Correlograma de DLVENTAS**

Sample: 1993:01 2006:09  
Included observations: 161

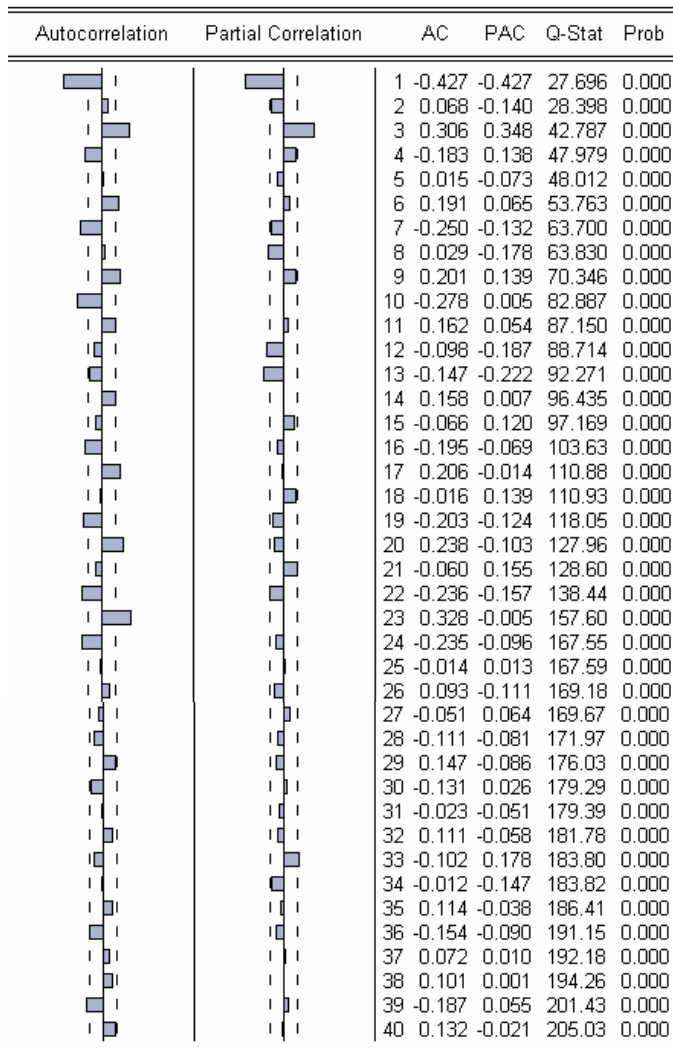


Al observar la función de autocorrelación se pueden identificar coeficientes sistemáticamente significativos asociados a los retardos (1, 12, 24 y 36), que sugieren que existe una tendencia en la parte estacional que es imprescindible corregir para luego especificar el modelo  $ARIMA(p,d,p)$  más apropiado. El tratamiento de la estacionalidad requiere calcular la serie con una primera diferencia en la parte estacional que como bien se aprecia en el correlograma anterior, y en la Gráfica de DLVENTAS, se manifiesta en diciembre de cada año. Para llevar a cabo esta transformación, se recurre de nuevo al

comando *Genr* y en el cuadro de diálogo se especifica la instrucción  $DDLVENTAS = DLVENTAS - DLVENTAS(-12)$ , donde DD hace referencia a la diferencia de una serie que ya está expresada en primera diferencia, y (-12) denota el orden del rezago que captura la estacionalidad de la serie. Otra forma alternativa de obtener lo anterior, es especificando en el propio cuadro de diálogo la instrucción  $DDLVENTAS = DLOG(VENTAS,1,12)$ . Concluida esta secuencia de pasos, se estima de nuevo la función de autocorrelación simple y parcial pero para DDLVENTAS, cuyo correlograma se ilustra en la Figura 4.5.

**Figura 4.5. Correlograma de DDLVENTAS**

Sample: 1993:01 2006:09  
Included observations: 149



Con base en este instrumento, se determina que un modelo idóneo a estimar es el siguiente:

DDLVENTAS AR(1) AR(3) SAR(12) MA(1), que contempla un autorregresivo de orden 1 y 3 y una media móvil de orden 1 en la parte regular, y un autorregresivo en la parte estacional SAR(12), donde S hace referencia al componente estacional. En otros términos, SAR(12), es la forma de señalar un proceso autorregresivo de orden 1 en la parte estacional considerando que la serie estudiada se trabaja con una frecuencia mensual.

Lo anterior se concluye a partir de observar las graficas de los coeficientes de autocorrelación simple y parcial, donde se puede apreciar que ambas funciones experimentan un decrecimiento conforme va aumentando el orden del rezago, lo que sugiere que se debe estimar un modelo tanto con un componente autorregresivo como de media móvil. Se eligió un autorregresivo de orden 1 en razón de que los coeficientes de ambas funciones de autocorrelación son significativos de acuerdo a su representación gráfica, las cuales exceden las líneas punteadas que se ubican en los extremos de la línea continua y que establecen los límites de confianza.

En este mismo sentido, se ha introducido el autorregresivo de orden 3, debido a que también es estadísticamente significativo. La idea de incorporar a la ecuación una media móvil de orden 1, responde al hecho de que la función de autocorrelación parcial decrece en forma regular con un coeficiente en el rezago 1 significativo, y además se conjuga con coeficiente significativo en el primer rezago de la función de autocorrelación simple. Hay que tomar en cuenta que este proceso de identificación se realiza tomando en cuenta los patrones teóricos de la función de autocorrelación simple y parcial explicados en el Capítulo 1. Esta especificación que consideramos apropiada se estima y se reportan los resultados en el Cuadro 4.5.

### Cuadro 4.5. Resultados de Estimación

Dependent Variable: DDLANTAD					
Method: Least Squares					
Sample (adjusted): 1995:05 2006:06					
Included observations: 134 after adjustments					
Convergence achieved after 10 iterations					
Backcast: 1995:04					
Variable	Coefficient	Std. Error	t-Statistic	Prob.	
AR(1)	-0.236968	0.106068	-2.234115	0.0272	
AR(3)	0.395972	0.074182	5.337821	0.0000	
SAR(12)	-0.289666	0.071703	-4.039793	0.0001	
MA(1)	-0.472638	0.109207	-4.327924	0.0000	
R-squared	0.455364	Mean dependent var		0.001799	
Adjusted R-squared	0.442796	S.D. dependent var		0.041875	
S.E. of regression	0.031258	Akaike info criterion		-4.063662	
Sum squared resid	0.127021	Schwarz criterion		-3.977160	
Log likelihood	276.2654	Durbin-Watson stat		1.913478	
Inverted AR Roots	.87-.23i	.87+.23i	.66	.64-.64i	
	.64+.64i	.23-.87i	.23+.87i	-.23+.87i	
	-.23-.87i	-.45-.63i	-.45+.63i	-.64-.64i	
	-.64-.64i	-.87-.23i	-.87+.23i		
Inverted MA Roots	.47				

Los estadísticos  $t$  indican que los cuatro regresores son significativos a un nivel de confianza del 95%. Aunado a esto, el estadístico de Durbin-Watson y el coeficiente de determinación múltiple son aceptables, por lo que se puede considerar como una especificación adecuada. Para juzgar la validez del modelo, también se puede utilizar la matriz de varianzas y covarianzas que se resume en el Cuadro 4.6. Este instrumento nos ayuda a tener un punto de referencia sobre el vínculo entre los parámetros del modelo. Por ejemplo, una covarianza elevada indica que los parámetros involucrados son redundantes y que por lo tanto habría que eliminar alguno sin afectar necesariamente la capacidad explicativa del modelo. Si por el contrario se tiene covarianzas pequeñas, significa que los parámetros se pueden mantener dentro de la especificación del modelo. En este sentido, los resultados derivados de la matriz de varianzas y covarianzas indican que no hay problemas con respecto a la redundancia de parámetros en el modelo.

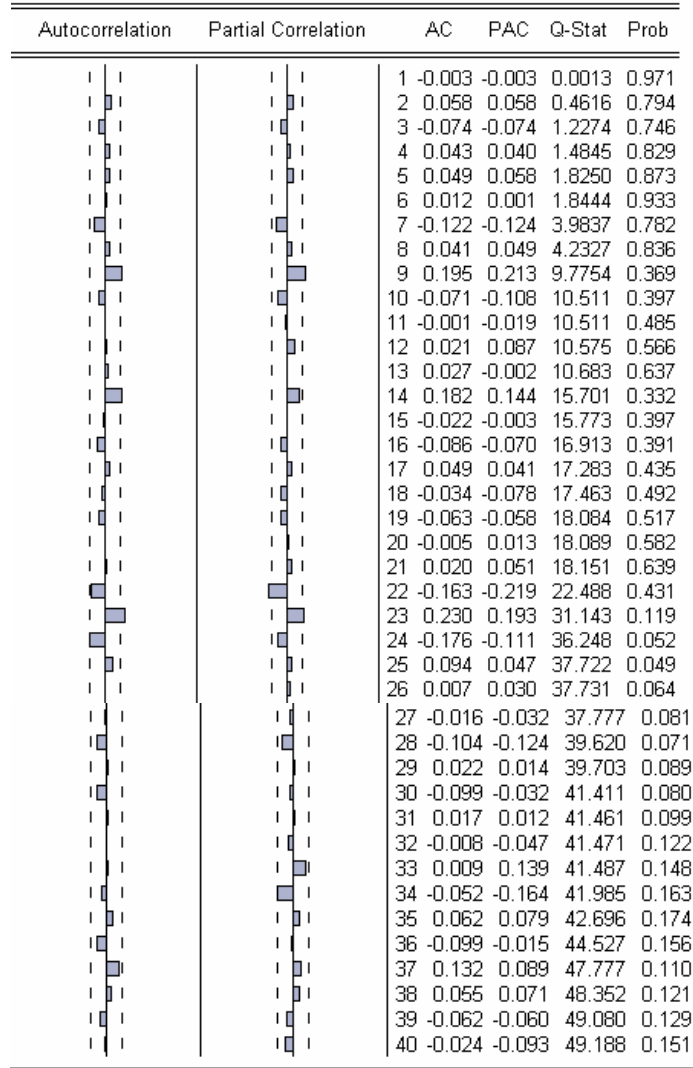
**Cuadro 4.6. Matriz de Varianzas y Covarianzas**

	AR(1)	AR(3)	SAR(12)	MA(1)
AR(1)	0.011250	0.000287	-0.000287	-0.008106
AR(3)	0.000287	0.005503	-0.000633	-0.001443
SAR(12)	-0.000287	-0.000633	0.005141	0.000796
MA(1)	-0.008106	-0.001443	0.000796	0.011926

Finalmente, para determinar si el modelo permite un ajuste razonable de los datos utilizados, se obtiene el correlograma de la función de autocorrelación simple y parcial de los residuales del modelo estimado previamente, y se analizan para ver si son puramente aleatorios (estacionarios). El mismo se presenta en la Figura 4.6. Las probabilidades de los estadísticos  $Q$  son superiores a 0.05, por lo que se pueden considerar como ruido blanco. Esta prueba convencional sugiere que no hay necesidad de buscar otro modelo alternativo. Al evaluar su capacidad de pronóstico, Cuadro 4.7, se obtiene una raíz del error cuadrático medio de 0.0410 y un error absoluto medio de 0.0329, corroborando su idoneidad para proyectar fuera de la muestra. Ambos criterios de capacidad de pronóstico se detallan en el siguiente capítulo.

**Figura 4.6. Correlograma de los Residuales del Modelo Estimado**

Sample: 1993:01 2006:09  
Included observations: 134



**Cuadro 4.7. Pronósticos**

Datos reales	Periodo	ddlantad	ddlantad
	2006:01	0.003600	0.003600
2006:02	-0.016765	-0.016765	
2006:03	-0.004734	-0.004734	
2006:04	0.017951	0.017951	
2006:05	-0.058581	-0.058581	
2006:06	0.000179	0.000179	
pronóstico	2006:07	0.019691	0.019691
	2006:08		-0.012355
	2006:09		0.003592

## **Capítulo 5: Temas Selectos de Modelos Univariados.**

En este capítulo presentamos una serie de métodos que son ampliamente utilizados en el análisis de series de tiempo, como los promedios móviles y de suavizamiento exponencial. Los primeros permiten la realización de pronósticos a partir del promedio de observaciones pasadas, mientras que los segundos ayudan a generar pronósticos utilizando promedios ponderados.

### **5. 1. Criterios para Evaluar la Capacidad Predictiva**

El análisis de series de tiempo se convierte en un tema fundamental cuando se procura llevar a cabo predicciones de una variable con el interés de tomar decisiones futuras. En la literatura existen diversos modelos que facilitan la realización de pronósticos, con distinta estructura matemática y capacidad predictiva. En términos generales se puede hablar de modelos MA, de suavizamiento exponencial, y ARMA. Aquí se efectúa una descripción general de cada uno de estos tipos de modelo, con el propósito de que el lector tenga una visión integral de las diferentes metodologías propuestas. Para tal efecto, primeramente se describen una serie de estadísticos que regularmente se utilizan para evaluar la capacidad predictiva de una modelo. Hay que considerar que la tarea de elegir entre una especificación u otra, se facilita cuando se dispone de herramientas muy claras como las que se describen en el Cuadro 5.1. Los estadísticos descritos sirven para comparar resultados obtenidos a partir de diferentes modelos, es decir, ayudan a determinar cuál modelo o técnica de pronósticos es más efectiva.

**Cuadro 5.1. Estadísticos para Evaluar la Capacidad Predictiva**

	<b>INDICADOR</b>	<b>NOTACIÓN</b>
1	Error cuadrático medio	$EM = \frac{\sum_{i=1}^n e^2_i}{n}$
2	Error absoluto medio	$EAM = \frac{\sum_{i=1}^n  e_i }{n}$
3	Raíz del error cuadrático medio	$RECM = \sqrt{\frac{\sum_{i=1}^n e^2_i}{n}}$
4	Error absoluto medio del porcentaje de error	$EAMPE = \frac{1}{n} \sum_{i=1}^n \left  \frac{e_i}{y_i} \right $
5	Coefficiente de desigualdad de Theil	$CDT = \frac{\sqrt{\sum_{i=1}^n (\hat{y}_i - y_i)^2 / n}}{\sqrt{\sum_{i=1}^n \hat{y}_i^2 / n + \sum_{i=1}^n y_i^2 / n}}$
6	Proporción de sesgo	$\frac{(\bar{\hat{y}} - \bar{y})^2}{\sum_{i=1}^n (\hat{y}_i - y_i)^2 / n}$
7	Proporción de varianza	$\frac{(S_{\hat{y}} - S_y)^2}{\sum_{i=1}^n (\hat{y}_i - y_i)^2 / n}$
8	Proporción de covarianza	$\frac{2(1-r)S_{\hat{y}}S_y}{\sum_{i=1}^n (\hat{y}_i - y_i)^2 / n}$

Para cualquiera de los casos,  $n$  es el número de observaciones,  $e_i$  el error de predicción que registra la diferencia entre cada valor observado de la serie  $y_i$  y su valor estimado. Específicamente, el error absoluto medio no toma en cuenta en su cálculo los signos de los valores que asumen los errores. La raíz del error cuadrático es un estadístico complementario al anterior. Por su parte, el error absoluto medio del porcentaje de error es una medida adicional que involucra en el cálculo los valores absolutos de la variable dependiente. El coeficiente de Theil es un estadístico más complejo y sus valores oscilan entre cero y uno. La principal característica de estos primeros cuatro estadísticos, es que en la medida en que su valor se aproxima a cero, indica que el modelo analizado tiene una adecuada capacidad predictiva; lo que significa que tiende a minimizar los errores de predicción. La descomposición del error cuadrático medio permite obtener varios componentes que igualmente se utilizan como criterios de evaluación de la capacidad predictiva, estos son: la proporción del sesgo, proporción de varianza y proporción de covarianza.

## **5. 2. Promedio Simple y Promedio Móvil**

Considerando que un número importante de series económicas presenta algún componente irregular que en la práctica dificulta la realización de un buen pronóstico, es que se recomienda suavizarlas mediante algún método estadístico. En este subapartado, se describe lo que es un promedio simple y móvil. El primero consiste simplemente en promediar todas las observaciones disponibles a través de la siguiente expresión:

$$\hat{y}_{t+1} = \frac{1}{n} \sum_{i=1}^n y_i \quad (5.1)$$

Donde  $y_i$  es una observación real o el nivel,  $n$  el número de observaciones, e  $\hat{y}_{t+1}$  el pronóstico generado para el periodo  $t+1$ . Una vez calculado el primer pronóstico, el siguiente se estima mediante la misma fórmula pero añadiendo la nueva observación  $\hat{y}_{t+1}$ . Conviene destacar que esta técnica es recomendable aplicarla cuando se tiene cierta certeza de que los datos que integran la serie no han estado influidos por un entorno cambiante. Adicionalmente, existen los promedios móviles que a diferencia de los simples, incluyen las últimas observaciones de la serie histórica mensual o trimestral.

Un promedio móvil de orden  $k$ , se obtiene a través de:

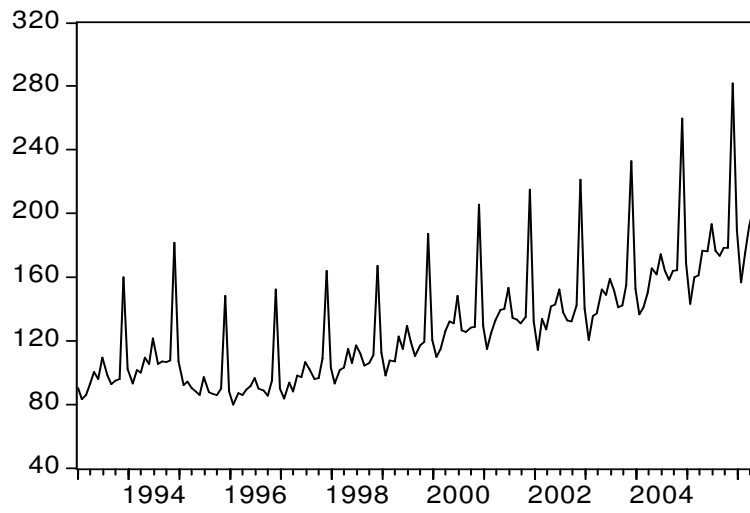
$$\hat{y}_{t+1} = \frac{y_t + y_{t-1} + y_{t-k+1}}{k} \quad (5.2)$$

Donde  $\hat{y}_{t+1}$  es el valor pronosticado,  $y_t$ , el valor real en el periodo  $t$  y  $k$  es el número de términos que integran el promedio móvil. Es importante anotar que este método se caracteriza por darle un peso igual a cada observación independientemente de la magnitud de la cifra. Un promedio como este, se puede definir como la media aritmética de las observaciones más recientes  $k$ . La idea de aplicar un procedimiento de esta naturaleza es siempre con el propósito de transformar la serie original y obtener una nueva que se distinga por un comportamiento más suave a lo largo de un intervalo de tiempo. En realidad es un método que consiste en ir recalculando el promedio incorporando la nueva

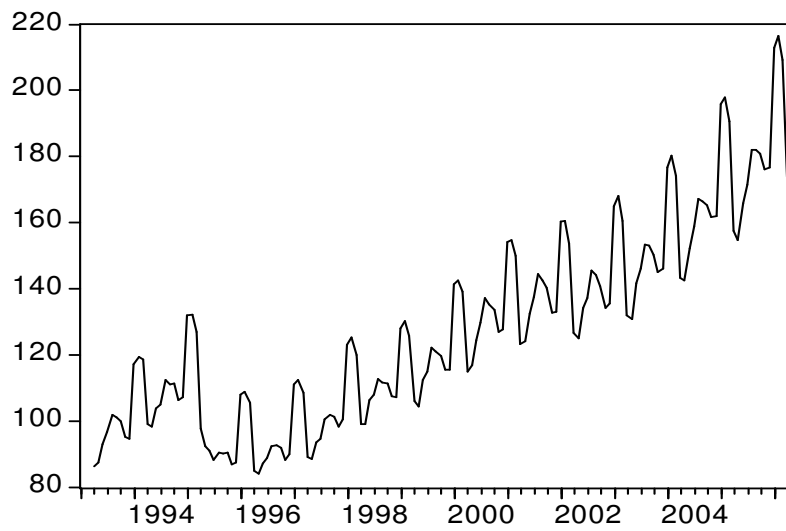
observación y eliminando aquella que ocurrió  $k$  periodos en el pasado. Para determinar si es mejor un promedio móvil trimestral o mensual, se deben calcular para ambos ejercicios el error medio, error absoluto medio o la raíz cuadrada del error cuadrático medio y con base en ello, se elige el modelo que reporta los valores más pequeños de los estadísticos.

A continuación se presenta un ejercicio de promedio móvil para tres y cinco meses de las ventas de la ANTAD y se calcula el error absoluto medio y el error cuadrático medio para elegir entre uno y otro promedio o serie suavizada. En la Gráfica 5.1 se describe la evolución de las ventas en valores originales, los cuales proyectan un componente de tendencia y estacionalidad al cierre de cada año. Con el propósito de disponer de una serie suavizada, se calculan promedios móviles trimestrales y semestrales. La Gráfica 5.2 ilustra el primer caso y la Gráfica 5.3 el segundo. Al observar las Gráficas, se puede notar que la serie que considera un promedio móvil de 6 meses es mucho más suavizada que la de tres meses. Al calcular el error absoluto medio (EAM) y el error cuadrático medio (ECM), se determina que el promedio móvil más apropiado para suavizar la serie analizada es el que contempla seis meses. Para este caso el EAM fue de 14.6 y el ECM de 612.3, mientras que para la serie suavizada con un promedio de tres meses el EAM fue de 18.6 y el ECM de 787.2. En este sentido, resulta obvio que la serie producto de aplicar un promedio móvil de 6 meses es la más apropiada para fines de pronóstico.

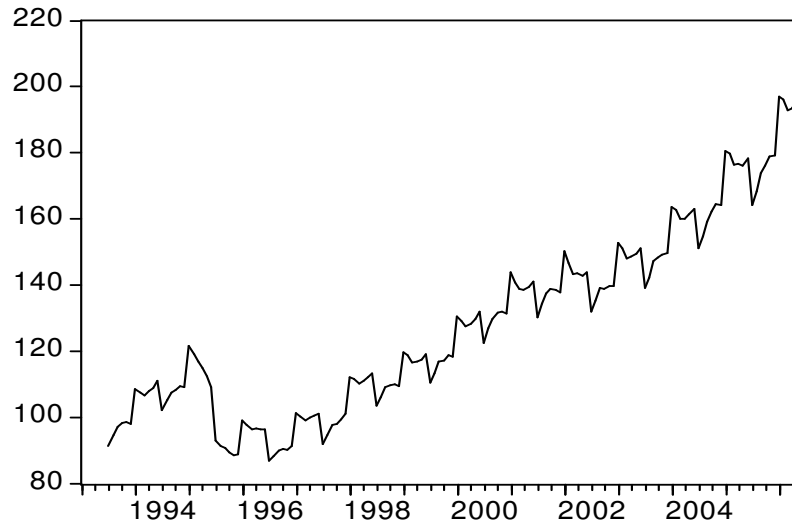
**Grafica 5.1. Comportamiento de las Ventas de la ANTAD**



**Gráfica 5.2. Promedio Móvil de Tres Meses de las Ventas de la ANTAD**



**Gráfico 5.3. Promedio Móvil de Seis Meses de las Ventas de la ANTAD**



### 5.3. Suavizamiento Exponencial Simple y Doble

Alternativamente a los métodos previamente señalados, se encuentran los que llevan a cabo un suavizamiento exponencial correspondiente a cada observación de la serie original. El esquema de suavizamiento se inicia aplicando la siguiente ecuación básica:

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha) \hat{y}_t \quad (5.3)$$

Donde  $\hat{y}_{t+1}$  es el nuevo valor suavizado o valor pronosticado para el periodo  $t+1$ ,  $\alpha$  es una constante de suavizamiento que puede asumir valores sujetos a la restricción  $(0 < \alpha < 1)$ ,

$y_t$  es la observación real correspondiente al periodo  $t$  e  $\hat{y}_t$  es el antiguo valor suavizado o pronóstico del periodo  $t$ . La ecuación (5.3) puede ser reformulada como:

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha)\hat{y}_t = \alpha y_t + \hat{y}_t - \alpha \hat{y}_t = \hat{y}_t + \alpha(y_t - \hat{y}_t) \quad (5.4)$$

El nuevo pronóstico es igual al anterior ajustado en  $\alpha$  veces el error del pronóstico antiguo. Ahora bien, replanteando la ecuación (5.4) para  $t$ , se tiene  $\hat{y}_t = \alpha y_{t-1} + (1 - \alpha)\hat{y}_{t-1}$  por tanto, al sustituir esta última expresión en (5.3) se obtiene

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha)[\alpha y_{t-1} + (1 - \alpha)\hat{y}_{t-1}] \quad (5.5)$$

Que a su vez es igual a:

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha)y_{t-1} + (1 - \alpha)^2 \hat{y}_{t-1} \quad (5.6)$$

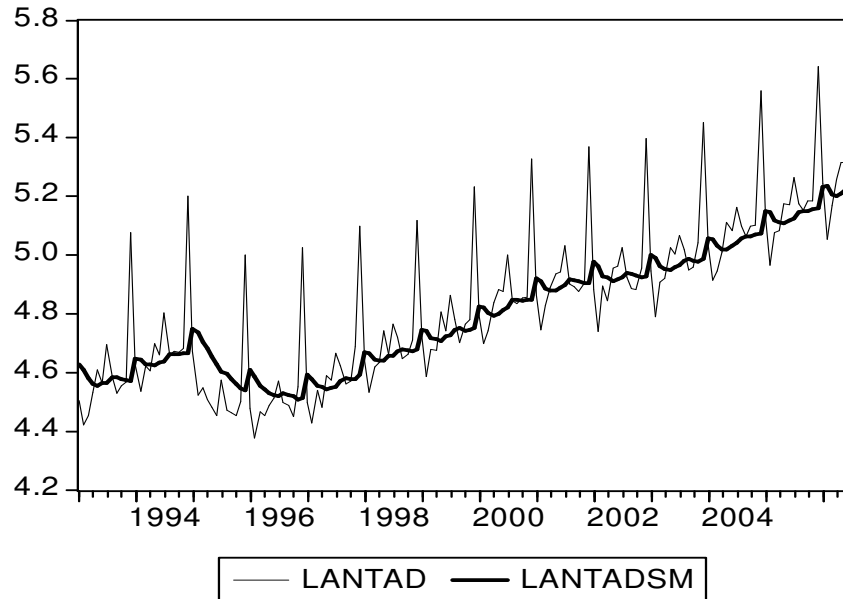
Donde  $\hat{y}_{t+1}$  sigue siendo un valor suavizado y la velocidad a la que pierden su impacto las observaciones pasadas depende de  $\alpha$ . El suavizamiento producido puede considerarse que es obtenido por una media móvil acompañada de ponderaciones decrecientes en forma de progresión geométrica. Si bien el suavizamiento exponencial está basado en promedios móviles como el método anterior, se diferencia por el hecho de que son valores ponderados. La ponderación exponencial es porque le concede mayor peso a las observaciones inmediatamente anteriores y menos a las de periodos más atrasados.

A partir de (5.6), se determina que:

$$\alpha \sum_{s=1}^{\infty} (1-\alpha)^s = \frac{\alpha}{1-(1-\alpha)} = 1 \quad (5.7)$$

El valor asignado a esta constante es fundamental, si se desea que las predicciones sean estables y que se suavicen las variaciones aleatorias, se sugiere utilizar una constante de valor pequeño. Por otra parte, si interesa una respuesta rápida a un cambio real en el patrón de las observaciones, es conveniente elegir un valor mayor para la constante de suavizamiento. Este método está considerado en el cuadro de diálogo del EViews denominado *Exponential Smoothing* identificado como método de suavizamiento exponencial simple con un parámetro (*Smoothing Method: Single*). Para ilustrar su aplicación tomamos la serie de ventas del ANTAD y le aplicamos el suavizamiento simple. La Gráfica 5.4 presenta el resultado y en el Cuadro 5.2 se muestran los estadísticos de evaluación. Notamos en la Gráfica que la serie suavizada LANTADSM ya no presenta un patrón estacional tan pronunciado como la serie original.

**Gráfica 5.4. Suavizamiento Simple de LANTAD**



**Cuadro 5.2. Estadísticos de Evaluación**

Date: 08/18/06	Time: 14:39
Sample: 1993:01	2006:06
Included observations:	162
Method:	Single Exponential
Original Series:	LANTAD
Forecast Series:	LANTADSM
Parameters: Alpha	0.1520
Sum of Squared Residuals	4.123723
Root Mean Squared Error	0.159546
End of Period Levels: Mean	5.239528

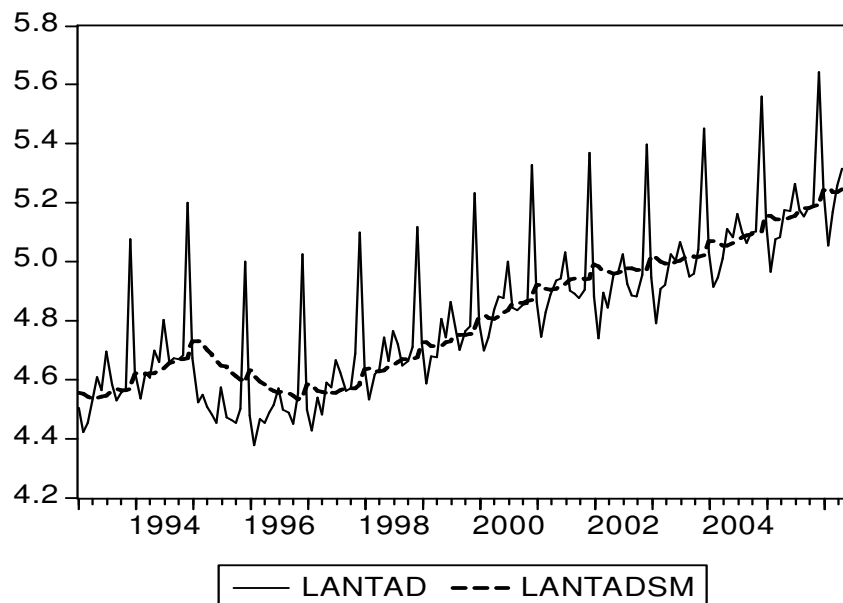
Una técnica de suavizamiento que se recomienda aplicar cuando la serie analizada ilustra una clara tendencia, es el suavizamiento doble con un parámetro, que también se conoce en la literatura como suavizamiento exponencial de segundo orden. Éste consiste en aplicar la técnica de suavizamiento simple dos veces usando el mismo parámetro. De acuerdo con Mendenhall y Reinmuth (1981), el pronóstico suavizado está determinado por la ecuación:

$$\hat{y}_{t+T} = [2 + \alpha T / (1 - \alpha) S_t - (1 + \alpha T / (1 - \alpha))] S_t(2) \quad (5.8)$$

Donde  $S_t(2) = \alpha S_t + (1 - \alpha) S_{t-1}(2)$

Siendo  $S_t(2)$  una estadística doblemente suavizada, es decir, un suavizamiento de los valores suavizados ( $S_t$ ). Se considera al igual que los demás métodos que la constante de suavizamiento  $\alpha$  se encuentra dentro del intervalo ( $0 < \alpha < 1$ ). Al igual que el método anterior, éste también está incorporado en la opción de *Exponential Smoothing* de EViews como (*Smoothing Method: Double*). Continuando con el ejemplo de la serie de ventas de ANTAD, presentamos los resultados de instrumentar esta técnica en la Gráfica 5.5 y en el Cuadro 5.3.

**Gráfica 5.5. Suavizamiento Doble de LANTAD**

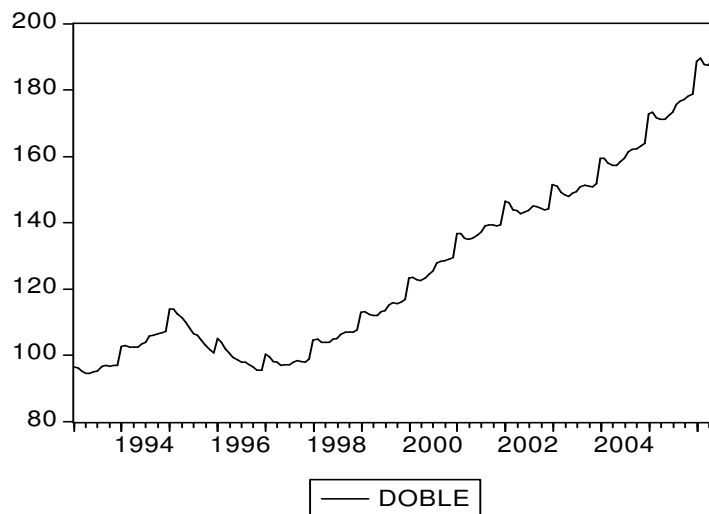
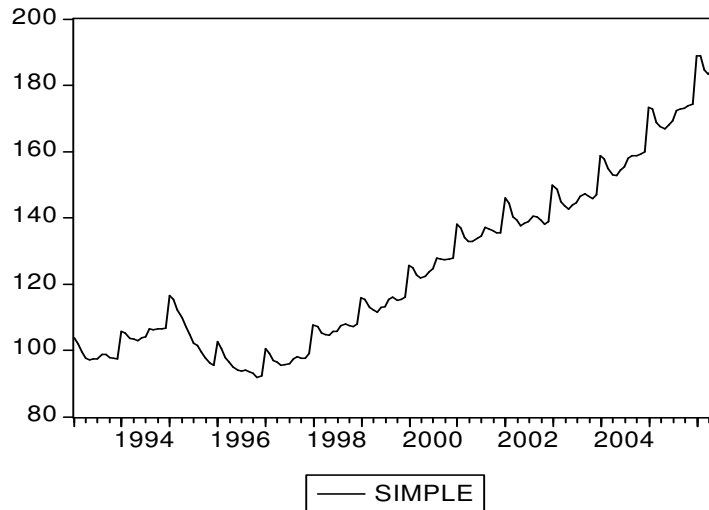


### Cuadro 5.3. Estadísticos de Evaluación

Date: 08/18/06	Time: 14:45
Simple: 1993:01	2006:06
Included observations:	162
Method:	Double Exponential
Original Series:	LANTAD
Forecast Series:	LANTADSM
Parameters: Alpha	0.0500
Sum of Squared Residuals	3.983832
Root Mean Squared Error	0.156817
End of Period Levels: Mean	5.262345
Trend	0.006260

Al comparar los estadísticos de los Cuadros 5.2 y 5.3, se constata que el método de suavizamiento doble arroja una suma de residuos al cuadrado y una raíz del error cuadrático medio más pequeño, lo que sugiere que el suavizamiento doble es más eficiente que el generado por el método de suavizamiento simple. Aunado a esto, hay que considerar que la constante de suavizamiento doble es menor que la del suavizamiento simple. A este respecto, conviene recordar que una constante más pequeña permite un mejor suavizamiento de las variaciones aleatorias. Por ello, no es fortuito que los estadísticos de prueba avalen como mejor método a este último. El lector podrá observar que las trayectorias de las series suavizadas con el método simple y doble ilustradas en la Gráfica 5.6, son diferentes, pues se aprecia con toda claridad como el método de suavizamiento doble produce una serie más estable, lo que lo convierte en más eficaz. Ahora bien, al observar la raíz del error cuadrático medio que arroja el suavizamiento simple y doble, se aprecia que éste último arroja un estadístico ligeramente menor, convirtiéndolo en un método más efectivo, pues es justamente el que más suaviza la serie analizada.

**Gráfica 5.6. Ventas de la ANTAD con Suavizamiento Simple y Doble**



#### **5.4. Suavizamiento Exponencial Lineal Ajustado a la Tendencia**

Mientras que el suavizamiento exponencial simple se aplica normalmente a series que no experimentan importantes cambios aleatorios, éste método se aplica cuando los datos de una serie muestran tendencia ascendente o descendente. Es importante mencionar que esta técnica está estrechamente relacionada con las aportaciones de Holt (1957) quien entre

otros autores desarrolló las bases teóricas de su construcción. También es usual que a este método se le llame técnica de suavizamiento exponencial doble con dos parámetros  $\alpha$  y  $\beta$ , cuya estructura matemática está definida por tres ecuaciones. De acuerdo con Makridakis et al. (1998) la primera ecuación que produce la serie suavizada se expresa como:

$$L_t = \alpha y_t + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (5.9)$$

Donde la tendencia se estima a través de:

$$T_t = \beta(L_t - T_{t-1}) + (1 - \beta)T_{t-1} \quad (5.10)$$

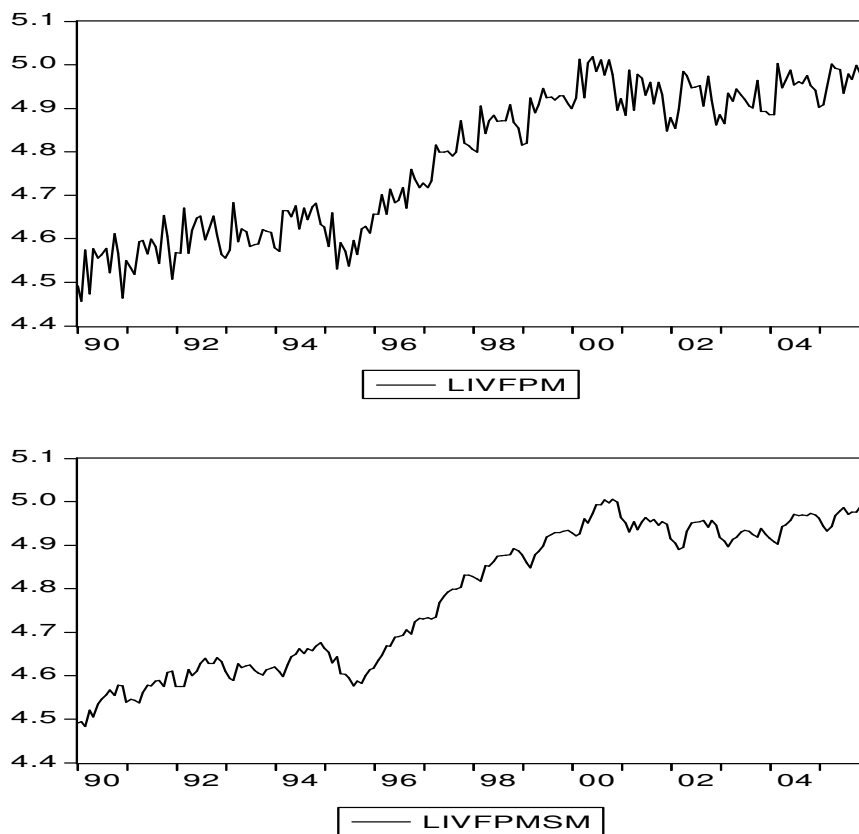
Y el pronóstico con base a:

$$\hat{y}_{t+p} = L_t + pT_t \quad (5.11)$$

El término  $L_t$  define el nuevo valor suavizado,  $\alpha$  es una constante de suavizamiento,  $y_t$  es el valor real de la serie en el periodo  $t$ ,  $\beta$  es una constante de suavizamiento para el estimado de tendencia,  $T_t$  es el estimado de la tendencia,  $p$  es periodo a pronosticar en el futuro y finalmente  $\hat{y}_{t+p}$  es el pronóstico para el periodo  $t+p$ . Este método se utiliza cuando se desea suavizar una serie que manifiesta tendencia pero no estacionalidad y es justamente el método propuesto por Holt y ampliado por Winters (1960), que se identifica

en el paquete mencionado como *Smoothing Method: Holt – Winter – No Seasonal (two parameters)*). Aplicamos esta técnica a la serie logaritmo del Índice de Volumen Físico de la Producción Manufacturera (LIVFPM), y obtenemos la ilustración de la Gráfica 5.7 y el Cuadro 5.4. Se ha optado por utilizar esta serie en virtud de que no refleja como en el caso de las ventas de la ANTAD, una estacionalidad muy marcada.

**Gráfica 5.7. Suavizamiento Holt-Winters de LIVFPM**



La primera gráfica corresponde a los valores originales de la serie LIVFPM y la segunda es la serie suavizada con el método previamente descrito. Como se puede ver, ésta última tiene un comportamiento más estable a lo largo de todo el periodo analizado. Por su parte, los resultados del Cuadro 5.4 arrojan una serie de estadísticos que ayudan a evaluar la

capacidad de predicción del método de suavizamiento. Como se puede observar, tanto la suma de residuales al cuadrado como la raíz del error cuadrático medio son realmente bajas, lo que sugiere que es un método de suavizamiento indicado para pronosticar la serie considerada.

**Cuadro 5.4. Estadísticos de Evaluación**

Date:	08/28/06	Time:	02:43
Sample:	1990:01	2005:12	
Included observations:	192		
Method:	Holt-Winters No Seasonal		
Original Series:	LIVFPM		
Forecast Series:	LIVFPMSM		
Parameters:	Alpha		0.3700
	Beta		0.0000
Sum of Squared Residuals			0.317972
Root Mean Squared Error			0.040695
End of Period Levels:	Mean		4.978432
	Trend		0.003279

**5.5. Suavizamiento Exponencial para Tendencia y Estacionalidad Multiplicativa**

Este es un método asociado a los trabajos de Winters (1960) que considera tres parámetros de suavizamiento exponencial. Su rasgo distintivo con respecto a la metodología anterior, es que incorpora una ecuación adicional para estimar la estacionalidad, pues su estructura es la apropiada para pronosticar series con tendencia y variación estacional multiplicativa.

Las ecuaciones involucradas son:

$$L_t = \alpha y_t / S_{t-s} + (1 - \alpha)(L_{t-1} + T_{t-1}) \tag{5.12}$$

Que estima la serie  $L_t$  a través de un suavizamiento exponencial. En tanto que la tendencia se estima con base a:

$$T_t = \beta(L_t + L_{t-1}) + (1 - \beta)T_{t-1} \quad (5.13)$$

Y la estacionalidad a través de la siguiente ecuación:

$$S_t = \gamma y_t / L_t + (1 - \gamma)S_{t-s} \quad (5.14)$$

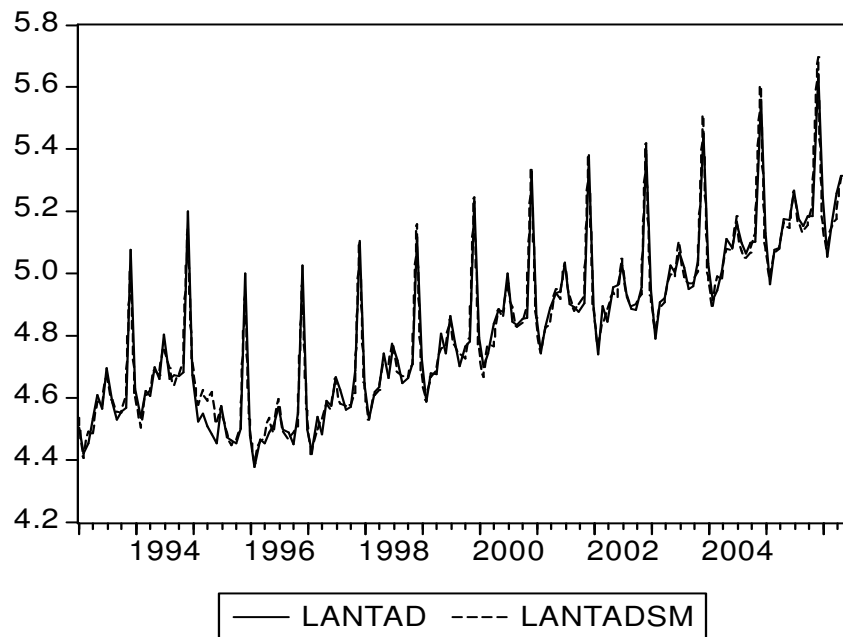
Y el pronóstico para el periodo  $t+p$  se determina mediante:

$$\hat{y}_{t+p} = (L_t + pT_t)S_{t-s+p} \quad (5.15)$$

El término  $L_t$  es el valor suavizado o nivel actual estimado,  $\alpha$  es una constante de suavizamiento para el nivel,  $y_t$  es la nueva observación o valor real en el periodo  $t$ ,  $\beta$  es una constante de suavizamiento para el estimado de la tendencia,  $T_t$  es el estimador de la tendencia,  $\gamma$  es una constante de suavizamiento para el estimado de la estacionalidad,  $S_t$  es el estimado de la estacionalidad,  $p$  el número de periodos futuros a pronosticar,  $s$  es la longitud de la estacionalidad, y finalmente  $\hat{y}_{t+p}$  es el pronóstico para el periodo  $p$  en el futuro. Este método es consistente con el registrado también en EViews como *Smoothing Method: Holt – Winter – Multiplicative (three parameters)*. En la Gráfica 5.8, se puede ver como las series LANTAD y LANTADSM tienden a ajustarse, con la característica de que

la serie estimada no presenta la estabilidad que sí arrojan las estimaciones previas en donde claramente se ve que LANTADSM no presenta las oscilaciones tan marcadas que reporta éste método de estimación. Los estadísticos presentados en el Cuadro 5.5, al igual que en los casos anteriores, sirven para evaluar la capacidad predictiva del método. Particularmente, el valor de la suma de residuales al cuadrado, y de la raíz del error cuadrático medio, permiten concluir que el método de suavizamiento utilizado es bueno.

**Gráfica 5.8. Suavizamiento con Tres Parámetros de LANTAD**



**Cuadro 5.5. Estadístico de Evaluación**

Date:	08/18/06	Time:	14:53
Simple:	1993:01	2006:06	
Included observations:	162		
Method:	Holt-Winters Multiplicative Seasonal		
Original Series:	LANTAD		
Forecast Series:	LANTADSM		
Parameters:	Alpha		0.5800
	Beta		0.0000
	Gamma		0.0000
Sum of Squared Residuals			0.190116
Root Mean Squared Error			0.034257
End of Period Levels:	Mean		5.327269
	Trend		0.004115
	Seasonals:	2005:07	1.015066
		2005:08	0.995283
		2005:09	0.987145
		2005:10	0.988266
		2005:11	0.996422
		2005:12	1.090962
		2006:01	0.993184
		2006:02	0.967823
		2006:03	0.983970
		2006:04	0.985455
		2006:05	1.001043
		2006:06	0.995383

**5.6. Filtro Hodrick – Prescott**

Aquí consideramos un método de suavizamiento que comúnmente se emplea en la descomposición de una serie de tiempo en su componente de tendencia y su componente cíclico. Al respecto, Hodrick y Prescott (1997) sugieren un filtro para estimar la tendencia estocástica de una serie. Si  $y_t$  es una serie tiempo  $t = 1, 2, 3, \dots, T$  y  $\tau_t$  su tendencia, entonces a partir de ello se tiene que  $C_t = y_t - \tau_t$  que representa una medida de las fluctuaciones cíclicas. Lo que subyace detrás de este planteamiento, es el interés por obtener una tendencia suavizada. Los autores plantean que el componente de tendencia de  $y_t$  es aquel que minimiza la ecuación:

$$\sum_{t=1}^T (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t)(\tau_t - \tau_{t-1})]^2 \quad (5.16)$$

Que también se puede expresar como:

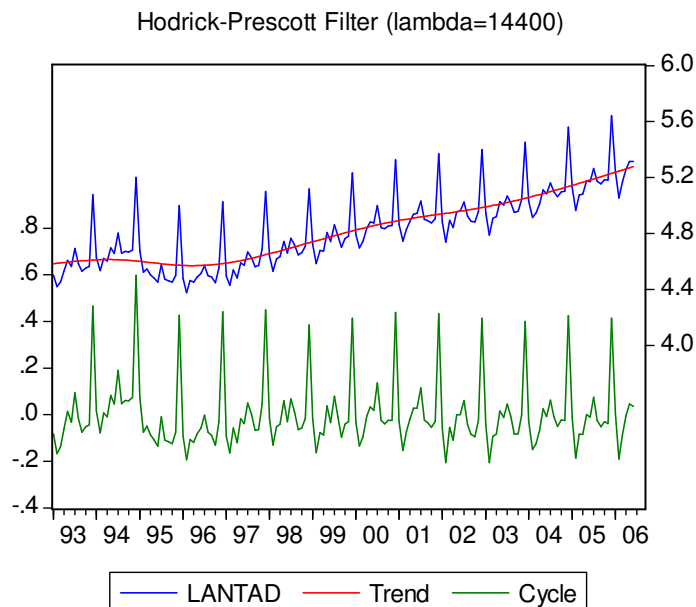
$$\sum_{t=1}^T C_t^2 + \lambda \sum_{t=3}^T (\Delta^2 \tau_t)^2 \quad (5.17)$$

Donde  $\Delta^2 = (1 - L)^2$  con  $L$  que denota el operador de rezago.

Al trabajar con  $y_t = \ln y_t$  y  $\tau_t = \ln \tau_t$  se pretende que  $C_t$  quede expresado directamente en términos de variaciones porcentuales de la serie respecto a la tendencia. El primer término de las ecuaciones (5.16) y (5.17), mide el grado de ajuste entre la tendencia y la serie original, el segundo término de ambas ecuaciones registra la sumatoria de cuadrados de la segunda diferencia del componente estacional y se interpreta como una medida del grado de suavidad. El parámetro  $\lambda$  penaliza la variabilidad de la tendencia medida como la segunda diferencia al cuadrado. En la medida en que este parámetro sea más pequeño, se considera que más suave es la tendencia. Si por ejemplo  $\lambda = 0$ ,  $\tau_t$  es igual a la serie de tiempo original en cuyo caso  $C_t = 0$ . Por el contrario, si  $\lambda = \infty$  los valores de la tendencia son expresados por la línea de mínimos cuadrados ordinarios y se les asigna la máxima ciclicidad posible. Se sugiere que cuando se dispone de información anual, se utiliza un parámetro  $\lambda = 100$ , para datos trimestrales  $\lambda = 1600$  y para información mensual  $\lambda = 14400$ .

Con estos valores se infiere que se aproxima a un filtro ideal, el cual se opera en EViews a través del cuadro de diálogo *Hodrick – Prescott Filter*, en el cual se debe especificar la serie objeto de análisis y el valor del parámetro en las opciones de *Smoothed Series* y *Smoothing Parameter* respectivamente. La Gráfica 5.9, reporta la descomposición de la serie LANTAD presentando por un lado su tendencia estimada y por el otro, el componente cíclico.

**Gráfica 5.9. Aplicación del Filtro Hodrick - Prescott**



El Filtro HP es ampliamente utilizado para establecer la existencia de características comunes entre tendencias y ciclos en un grupo de variables. Por ejemplo, es común estimar el producto interno bruto (PIB) potencial simplemente considerando la tendencia que se obtiene del Filtro como la series de tendencia de largo plazo. Una vez realizado este ejercicio para el PIB de diferentes países, se procede a estimar la correlación que existe entre las series; si se encuentra una correlación significativa, se concluye que las series

están sincronizadas. Dicho análisis presenta problemas serios, que se comentarán ampliamente en el Capítulo 9. Por ahora notamos que uno de ellos se refiere al hecho que la serie de tendencia que se obtiene estimando el Filtro cambia cada vez que se le agrega una observación. Así, si por ejemplo se estima el Filtro para  $t$ , y en  $t+1$  se desea estimar de nuevo, el ejercicio no podría sólo incluir la nueva observación y agregarla a la serie ya estimada, se tendría que instrumentar una vez más el Filtro obteniéndose una serie para  $t+1$  completamente diferente a la que se obtuvo para  $t$ ; por lo cual no es razonable argumentar en nuestro ejemplo que la serie de tendencia que se deriva del Filtro representa el PIB potencial puntualmente. Cabe sin embargo destacar que para ejercicio menos formales, como el suavizar una serie, el Filtro representa una herramienta importante.

### **5.7. La Tendencia y Estacionalidad en los Modelos Clásicos de Regresión**

Si bien existen modelos de pronóstico sencillos como los ya enunciados, también hay otros de gran utilidad que ayudan a modelar una serie cuando tiene tendencia o estacionalidad. Cuando una serie reporta una tendencia clara, se estima introduciendo en la ecuación de regresión simple la variable tiempo, asignando valores en orden cronológico a cada observación 1, 2, ..... $n$ . En este caso, se estima la siguiente ecuación:

$$\hat{y}_{t+1} = \delta_0 + \delta_1 * tiempo + e_i \quad (5.18)$$

Conviene advertir que en ciertas ocasiones la tendencia puede ser cuadrática, cúbica o exponencial, en cuyo caso cambiará la especificación de la ecuación (5.18). La selección de

un modelo u otro se realizará en función del  $\bar{R}^2$ , criterio de información de Akaike o criterio de información de Schwarz. El coeficiente de determinación ajustado por grados de libertad se define como:

$$\bar{R}^2 = 1 - \frac{SRC/(n-k)}{STC/(N-1)} = 1 - (1 - R^2) \frac{n-1}{n-k} \quad (5.19)$$

Donde SRC y STC son las sumas de residuos al cuadrado y la suma total de cuadrados respectivamente. Un aspecto importante de este estadístico es que es menor que el  $R^2$  normal, y además, penaliza cuando se agregan nuevas variables que no resultan ser estadísticamente significativas. Al contrario del  $R^2$  que es una función no decreciente del número de variables explicativas, el  $\bar{R}^2$  sí disminuye cuando se incorpora una variable irrelevante. Además, puede reportar un signo negativo aun cuando el coeficiente de determinación múltiple no ajustado no lo sea. En definitiva,  $\bar{R}^2$  se incrementará sólo en la medida en que el valor absoluto de  $t$  de la variable adicional sea mayor que 1.

Otro estadístico que ayuda a elegir el modelo más idóneo es el criterio de información de Akaike que se define como  $CIA = e^{2k/n} \sum ui^2 / n = e^{2k/n} SRC / n$ , donde  $k$  es el número de regresores (incluyendo el intercepto) y  $n$  el número de observaciones. Al utilizar las propiedades de los logaritmos<sup>7</sup>, se linealiza la expresión anterior y se obtiene la siguiente:

$$\ln CIA = (2k/n) + \ln(SRC/n) \quad (5.20)$$

---

<sup>7</sup> Las propiedades son:  $\ln(uv) = \ln u + \ln v$ ,  $\ln(u/v) = \ln u - \ln v$ ,  $\ln(1/v) = -\ln v$ ,  $\ln u^n = n \ln u$ ,  $\ln e = 1$  y  $\ln 1 = 0$

Donde  $\ln CIA$  es el logaritmo natural de  $CIA$ ,  $(2k/n)$  es el factor de penalización y  $SRC$  es la suma de los residuales al cuadrado. La regla de decisión establece que si se comparan dos modelos con base al  $CIA$ , se elegirá aquel que tenga el menor valor de  $CIA$ , y lo mismo aplica para el criterio de Schwarz que se define como  $CIS = n^{k/n} \sum ui^2 / n = e^{k/n} SRC / n$ , y que al llevar a cabo la transformación logarítmica respectiva se obtiene la ecuación final:

$$\ln CIA = k/n \ln(SRC / n) \quad (5.21)$$

Donde  $\ln CIA$  es el logaritmo natural del criterio,  $k/n$  es el factor de penalización y  $SRC$  es al igual que en el caso anterior la suma de residuos al cuadrado. Si por otra parte, se dispone de una serie con una frecuencia temporal trimestral, y se quiere modelar y pronosticar con base a la estacionalidad, se tiene que estimar una ecuación de la forma:

$$y_t = \sum_{i=1}^s \mathcal{D}_{it} + \ell_t \quad (5.22)$$

Donde se tendrán tantos parámetros  $\gamma$  como variables dicótomas que son parte de la ordenada al origen y que se identifican también como factores estacionales. Si como se señaló, la frecuencia temporal es trimestral, entonces se tendrá que  $s = 4$ . En consecuencia, la primera variable dicótoma  $D_{it}$ , registrará valores de (1) para el segundo trimestre y (0) para el resto, la segunda valores de (1) para el tercer trimestre y ceros para el resto, y así sucesivamente. Lo que sí es importante cuidar en este proceso de especificación, es no añadir la ordenada al origen y además  $s$  variables  $D_{it}$ ; en razón de que se genera un

problema de colinealidad perfecta o bien, se cae en lo que se llama la trampa de la variable dicótoma. Por tanto, la regla que debe prevalecer es que si una variable tiene  $s$  categorías, se deben introducir  $s-1$  variables dicótomas. Ahora bien, si de acuerdo a los resultados de regresión se determinara que ninguno de los parámetros  $\gamma$  fuese estadísticamente significativo, entonces sólo deberá quedar en la ecuación la ordenada al origen común o en su defecto la ordenada al origen y la variable indicadora que sólo sea significativa a un determinada nivel de confianza estadística.

Otra forma de llevar a cabo el pronóstico de la variable  $y_t$ , en caso de que se tuviera indicio en base a un análisis gráfico de que presenta tendencia y estacionalidad, es integrar a la ecuación (5.18) y (5.19) en una sola que se resumiría de la siguiente forma:

$$y_t = \beta t + \sum_{i=1}^s \gamma D_{it} + \ell_t \quad (5.23)$$

En este caso, se tendrían  $s$  regresores asociados con su respectivo parámetro  $\gamma$  más  $\beta$  que mediría el comportamiento de  $y_t$  en el transcurso del tiempo, que bien puede estar capturando una relación lineal positiva o negativa. Una versión ampliada de la ecuación (5.22), es cuando la ecuación además de incluir las  $s-1$  variables dicótomas, también incluye un variable explicativa  $x_t$ . Hay que tomar en cuenta que en el ámbito de la economía un número significativo de series basadas en información trimestral o mensual, regularmente presentan tendencia y estacionalidad pero además mantienen una relación de causalidad con otras variables explicativas. En este contexto, a veces resulta interesante

eliminar el componente estacional (desestacionalización de la serie) para concentrarse por ejemplo en la tendencia. Cuando este es el propósito y por ejemplo  $s = 4$ , se tiene que especificar una ecuación como la siguiente:

$$y_t = \gamma_1 + \gamma_2 D_{2t} + \gamma_3 D_{3t} + \gamma_4 D_{4t} + \delta x_t + \varepsilon \quad (5.24)$$

Donde los interceptos diferenciales  $\gamma_2, \gamma_3, \gamma_4$  están asociados a,

$D_2 = 1$  para segundo trimestre y 0 para el resto

$D_3 = 1$  para el tercer trimestre y 0 para el resto

$D_4 = 1$  para el cuatro trimestre y 0 para el resto

En caso de que sólo el coeficiente  $\gamma_3$  y  $\delta$  resultaran ser estadísticamente significativos, entonces la ecuación (5.24) se tendría que replantear como una versión restringida como la siguiente:

$$y_t = \gamma_1 + \gamma_3 D_{3t} + \delta x_t + \varepsilon \quad (5.25)$$

En definitiva, para determinar qué tipo de modelo de regresión es el más apropiado para pronosticar, se sugiere que primero se grafique la serie en cuestión y se lleve a cabo una exploración de su comportamiento para posteriormente pensar en el tipo de especificación más adecuada y a partir de distintas versiones, elegir el modelo más apropiado en función de los criterios previamente explicados.

## Parte II: Modelos Multivariados

### Capítulo 6. Introducción a los Sistemas Multivariados

En la Parte I del texto abordamos varios modelos que aproximan el PEG de una serie de tiempo. En adelante nos enfocaremos a caracterizar la interacción entre dos o más series. En este capítulo iniciamos definiendo lo que se entiende por un sistema multivariado, y presentamos los argumentos que dictan la necesidad de analizar series de tiempo por medio de la instrumentación de metodologías particulares; que constituyen fundamentalmente técnicas de cointegración.

#### 6.1. Fundamentos

Comenzamos por indicar que un sistema multivariado de series de tiempo consiste simplemente en un conjunto de series.<sup>8</sup> Así, podríamos ilustrar un conjunto de tres series de tiempo como sigue:

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix}$$

---

<sup>8</sup> Indicamos que un análisis teórico profundo de los temas que aquí trataremos se encuentra en Hamilton (1994).

La relación estocástica que existe entre estas variables puede ser estudiada de diversas formas, comúnmente se supone que existe una variable dependiente, o endógena, y una serie de variables explicativas o exógenas.<sup>9</sup> De tal forma que podríamos postular la siguiente relación:

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 z_t + \varepsilon_t$$

Esta ecuación puede ser estimada por medio de Mínimos Cuadrados Ordinarios (MCO's) obteniendo los estimadores que corresponden a las  $\beta$ 's. Sin embargo, es importante identificar la naturaleza estocástica de las series para determinar la eficiencia de dichos estimadores. En particular, si las series son estacionarias, los estimadores cuentan con las propiedades de eficiencia y no-sesgo propias de la estimación por MCO's. Alternativamente, si las series del sistema son no-estacionarias entonces puede, o no puede, ser que los resultados de la estimación sean válidos. Veamos por qué:

Consideremos la ecuación

$$y_t = x_t \beta + \varepsilon_t \tag{6.1}$$

Si suponemos que tanto  $y_t$  como  $x_t$  son iid, esto es, su media es constante y su varianza independiente del tiempo, entonces los estimadores  $\beta$  serían insesgados y presentarían

---

<sup>9</sup> Existen pruebas econométricas formales para determinar la exogeneidad de variables en un sistema; las cuales no abordamos en el presente texto. El lector interesado puede consultar por ejemplo Johnston y DiNardo (1997).

mínima varianza dentro del conjunto de estimadores lineales. Sin embargo, es poco probable que series de tiempo sean iid. Tomemos una de las series de la ecuación (6.1) y modelemos su PEG como un proceso autorregresivo:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \varepsilon_t \quad (6.2)$$

Si  $|\alpha_1| < 1$  entonces  $y_t$  es estacionaria, pero si  $|\alpha_1| = 1$  entonces  $y_t$  es no-estacionaria y tendríamos lo siguiente:

$$y_t = \alpha_0 + y_{t-1} + \varepsilon_t \quad (6.3)$$

Si modelamos la secuencia en el tiempo de  $y_t$  tendríamos

$$y_t = t\alpha_0 + y_0 + \sum_{j=1}^t \varepsilon_j \quad (6.4)$$

Donde  $y_0$  se asume sea un valor inicial de la secuencia de  $y_t$ . Suponiendo que  $\varepsilon_j$  es iid, la media y varianza de  $y_t$ , de acuerdo a (6.4), son  $t\alpha_0 + y_0$  y  $t\sigma_\varepsilon^2$  respectivamente. Nótese que ambos momentos se encuentran expresados como función del tiempo, esto es, ambos ya no son constantes, luego entonces, al estimar una ecuación como (6.1), donde tanto  $y_t$  como  $x_t$  sean no estacionarias, los estimadores por medio de MCO's no serían confiables y/o los estadísticos válidos.

¿Qué hacer entonces cuando las series de interés son no-estacionarias? Una alternativa consistiría en estimar la ecuación utilizando las primeras diferencias de las series.

Retomemos el PEG indicado en (6.3):

$$y_t = \alpha_0 + y_{t-1} + \varepsilon_t$$

Si substraemos  $y_{t-1}$  de ambos lados de la ecuación tenemos

$$y_t - y_{t-1} = \alpha_0 + \varepsilon_t \tag{6.5}$$

Que puede re-escribirse como

$$\Delta y_t = \alpha_0 + \varepsilon_t \tag{6.6}$$

Calculando la media y varianza de (6.6) se obtiene  $\alpha_0$  y  $\sigma_\varepsilon^2$  respectivamente, esto es, los dos primeros momentos son independientes del tiempo. En tal caso, si tanto  $y_t$  como  $x_t$  son no-estacionarias en niveles, pero estacionarias en primeras diferencias, entonces estimando la ecuación  $\Delta y_t = \Delta x_t \beta + \varepsilon_t$  por medio de MCO's se obtendrían estimadores para las  $\beta$ 's eficientes e insesgados, y los estadísticos de la estimación serían confiables.

¿Por qué entonces no se estima la relación entre cualquier conjunto de series de tiempo no-estacionarias utilizando sus primeras diferencias? Simplemente por que al transformar las

series en su primera diferencia, ya no se está estimando la relación entre las variables originales, sino entre una transformación lineal de las mismas; así, la interpretación de los coeficientes ya no es equivalente a la que se deriva cuando se estiman las relaciones en niveles.

Tomemos un ejemplo:

Supongamos que nos interesa estimar la relación entre el consumo ( $y_t$ ) y el ingreso ( $x_t$ ).

Previa transformación logarítmica de las variables, postulamos la siguiente ecuación:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \quad (6.7)$$

El coeficiente estimado,  $\beta_1$ , se interpretaría como la elasticidad del consumo con respecto al ingreso en el largo plazo. Sin embargo, como lo indicamos anteriormente, siendo que muy probablemente tanto el consumo como el ingreso son series no-estacionarias,  $\beta_1$  no tendría las propiedades de eficiencia que se desean. De tal suerte que lo apropiado sería estimar la relación considerando las primeras diferencias de las dos variables; así, el estimador sería confiable. Desafortunadamente, la interpretación del mismo no sería clara, esto es, ya no se podría interpretar como la elasticidad de largo plazo, ya que la variable dependiente sería la diferencia logarítmica del consumo (tasa trimestral de crecimiento del consumo) y la explicativa la diferencia logarítmica del ingreso (tasa trimestral de crecimiento del ingreso). Podríamos, en todo caso, interpretarlo como una elasticidad de corto plazo o la respuesta instantánea del consumo con respecto al ingreso.

Para ilustrar la idea estimamos una ecuación como (6.7) considerando el consumo privado y el PIB de México. La serie del PIB corresponde a la empleada en los ejemplos que se presentaron en la Parte I del texto. El consumo se refiere al consumo privado medido en idénticas unidades al PIB y su fuente es la base de datos de indicadores económicos y financieros de Banco de México. El periodo de estimación va del primer trimestre de 1993 al cuarto trimestre de 2005.

Antes de realizar la estimación, verificamos la estacionariedad de las series. Para el PIB se mostró en la Parte I que es de hecho una serie no-estacionaria, así que aquí sólo presentamos los resultados de la prueba KPSS para el consumo. Notamos en el Cuadro 6.1 que la hipótesis nula de estacionariedad para la serie en niveles se rechaza al 10% y al 5%, pero la misma no se rechaza para la serie en primeras diferencias a niveles convencionales. De tal forma que podemos concluir razonablemente que la serie de consumo privado está integrada de orden 1.

**Cuadro 6.1. Prueba de Raíz Unitaria KPSS**

Null Hypothesis: LOG(CONSUMO) is stationary		
Exogenous: Constant		
Bandwidth: 91.5 (Andrews using Parzen kernel)		
		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.566589
Asymptotic critical values*:	1% level	0.739000
	5% level	0.463000
	10% level	0.347000
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)		

Null Hypothesis: D(LOG(CONSUMO)) is stationary		
Exogenous: Constant		
Bandwidth: 4.31 (Andrews using Parzen kernel)		
		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.290040
Asymptotic critical values*:	1% level	0.739000
	5% level	0.463000
	10% level	0.347000
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)		

Enseguida estimamos la regresión entre el consumo y el ingreso y presentamos los resultados en el Cuadro 6.2. Nótese que el coeficiente  $\beta_1$  es aproximadamente 1, lo cual no es sorprendente en el caso de México, ya que la dependencia del consumo en el ingreso, a falta de acceso al mercado de crédito, es elevada. Los resultados, si embargo, pueden no ser confiables debido a las razones anteriormente indicadas.

**Cuadro 6.2. Resultados de  $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$** 

Dependent Variable: LOG(CONSUMO)				
Method: Least Squares				
Date: 08/10/06 Time: 13:23				
Sample (adjusted): 1993Q1 2005Q4				
Included observations: 52 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.252331	0.862475	-2.611472	0.0119
LOG(PIB)	1.089860	0.040848	26.68078	0.0000
R-squared	0.934371	Mean dependent var		20.75881
Adjusted R-squared	0.933059	S.D. dependent var		0.137516
S.E. of regresión	0.035579	Akaike info criterion		-3.796396
Sum squared resid	0.063295	Schwarz criterion		-3.721348
Log likelihood	100.7063	F-statistic		711.8638
Durbin-Watson stat	0.862964	Prob(F-statistic)		0.000000

Estimemos ahora la ecuación en primeras diferencias, así tendríamos la certeza que los coeficientes y estadísticos obtenidos son válidos. La ecuación que se estima es  $\Delta y_t = \beta_0 + \beta_1 \Delta x_t + \varepsilon_t$  y los resultados se presentan en el Cuadro 6.3. Evidentemente, el coeficiente 0.86 ya no se puede interpretar como la elasticidad de largo plazo, en todo caso, pudiera interpretarse como la respuesta instantánea del consumo al ingreso corriente, o en otras palabras, una elasticidad de corto plazo.

**Cuadro 6.3 Resultados de  $\Delta y_t = \beta_0 + \beta_1 \Delta x_t + \varepsilon_t$** 

Dependent Variable: DLOG(CONSUMO)				
Method: Least Squares				
Date: 08/10/06 Time: 13:58				
Sample (adjusted): 1993Q2 2005Q4				
Included observations: 51 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.001761	0.004513	0.390298	0.6980
DLOG(PIB)	0.862270	0.100839	8.550954	0.0000
R-squared	0.598751	Mean dependent var		0.008197
Adjusted R-squared	0.590563	S.D. dependent var		0.049661
S.E. of regresión	0.031776	Akaike info criterion		-4.021756
Sum squared resid	0.049477	Schwarz criterion		-3.945998
Log likelihood	104.5548	F-statistic		73.11881
Durbin-Watson stat	2.212700	Prob(F-statistic)		0.000000

¿Cuál es la alternativa entonces si se desea conocer la elasticidad de largo plazo obteniendo resultados confiables? La respuesta es una, y se refiere a instrumentar un análisis de cointegración. El concepto de cointegración, así como diferentes técnicas de estimación, se presenta detalladamente en el siguiente capítulo.

## Capítulo 7: Cointegración

### 7.1. Concepto

La implementación de alguna metodología de cointegración es común en la actualidad, ya que nos permite identificar relaciones estables de largo plazo entre variables de interés. Comencemos por definir el concepto. Considera dos procesos estocásticos integrados de orden 1

$$y_t \sim I(1)$$

$$x_t \sim I(1)$$

En general, una combinación lineal de estas variables se esperaría estuviera integrada de orden 1, esto es

$$y_t - \rho_1 x_t = \varepsilon_t \sim I(1) \tag{7.1}$$

Sin embargo, en algunas instancias las series  $y_t$  y  $x_t$  pueden compartir una tendencia estocástica común; de tal forma que la combinación lineal de las series podría estar integrada de orden 0, o ser estacionaria, formalmente tendríamos:

$$y_t - \rho_1 x_t = \varepsilon_t \sim I(0) \tag{7.2}$$

En dado caso, se dice que  $y_t$  y  $x_t$  están cointegradas. El lector familiarizado con el concepto se habrá percatado que la definición que presentamos de cointegración no es la que comúnmente se describe en libros de texto, que se refiere a la propiedad que exhiben series de tiempo de “moverse juntas”. Esta frase, que originalmente postularan Engle y Granger (1987), creemos ejemplifica apropiadamente el caso en el que las series se asocian positivamente, pero ¿podríamos afirmar que dos series que exhiben una relación negativa se mueven juntas? Pensemos en la relación entre el consumo y el tipo de cambio real. En la literatura se ha mostrado que las series presentan una fuerte asociación negativa y de hecho cointegran.<sup>10</sup> En este caso, es claro que las series no se “mueven juntas”, ya que un incremento en el nivel del tipo de cambio real (depreciación del peso) se asocia con una caída en el consumo; así que precisamos que el concepto de cointegración se refiere más puntualmente a la propiedad que exhiben series de tiempo de compartir tendencia común.

A partir de esta caracterización, es posible determinar los coeficientes de cointegración, que de hecho representan los ponderadores que hacen posible que la combinación lineal de  $y_t$  y  $x_t$  sea estacionaria. En nuestro ejemplo,  $y_t - \rho_1 x_t$ , ya que estamos normalizando con respecto a  $y_t$  (esto es, dividimos los ponderadores por el coeficiente de  $y_t$ ), los coeficientes serían 1 y  $-\rho_1$ . Nótese que los mismos no son únicos, de hecho cualquier múltiplo de ellos constituirían coeficientes de cointegración. Asimismo, es importante destacar que el signo negativo de los coeficientes *normalizados*, en efecto indica una asociación positiva entre las variables  $y_t$  y  $x_t$ , ¿por qué?, por que al escribir la relación  $y_t - \rho_1 x_t = \varepsilon_t$  en forma de ecuación se tiene  $y_t = \rho_1 x_t + \varepsilon_t$ , de tal forma que la variable

---

<sup>10</sup> Ver Castillo (2003) por ejemplo.

$x_t$  presenta una asociación positiva con  $y_t$ . Así, el vector de cointegración en este caso es  $(1, -\rho_1)$ .

La instrumentación del ejercicio de cointegración se puede llevar a cabo bajo diferentes metodologías. A continuación se describen algunas de las más comúnmente utilizadas.

## **7.2. Metodologías de Cointegración**

En la práctica, el análisis de cointegración puede al menos cumplir con dos propósitos: el primero se refiere a identificar la existencia de relaciones estables de largo plazo entre variables. El segundo va un paso más adelante y no sólo se concreta en identificar la existencia de una tendencia común entre las series, sino que pretende determinar el número de relaciones de cointegración que existe en un sistema. Este ejercicio es particularmente interesante, por ejemplo, en el análisis de descomposición de series de tiempo. A saber, una serie de tiempo consiste fundamentalmente de un componente tendencial y un componente cíclico. De tal forma que al realizar un ejercicio de cointegración, lo que se está haciendo es, visto desde una óptica diferente a la descrita en la sección anterior, identificar los componentes tendenciales de las series y determinar si existe un comportamiento común entre ellos, pero ¿por qué es importante conocer el número de relaciones de cointegración en dicho análisis? Por que en un sistema de  $n$  variables, la suma de los componentes

tendenciales comunes ( $r$ ) y componentes cíclicos comunes ( $s$ ), debe cumplir con la condición  $r + s \leq n$ .<sup>11</sup>

A partir de esta caracterización podemos mencionar otra instancia en la que puede ser interesante conocer el número de vectores de cointegración, y éste es el caso en el que el sistema consiste de más de dos series. Supón que se considera un sistema de tres variables, entonces es posible que existan hasta 2 vectores de cointegración, cada uno de los cuales sería único. En tal caso, podría ser interesante conocer los dos vectores y comparar la magnitud de los coeficientes de cointegración para poder determinar aquellos que pudieran tener más sentido de acuerdo a la teoría que se esté evaluando. Ésta, por supuesto, es sólo una de un sin número de instancias en las cuales sería deseable determinar el número de vectores de cointegración.

Para cumplir con el primer propósito antes mencionado, simplemente determinar la existencia de cointegración, podemos referirnos fundamentalmente a dos metodologías: a) El método de dos etapas sugerido por Engle y Granger (1987), y b) Modelos de corrección de error. Para el segundo propósito, determinar el número de vectores de cointegración, típicamente se instrumenta la metodología sugerida por Johansen (1991). Como se mostrará a continuación, existen diferencias importantes entre estas metodologías, sin embargo, queda a la discreción del investigador determinar aquella que satisfaga puntualmente los requerimientos de su análisis.

---

<sup>11</sup> El estudio de tendencias/ciclos comunes es muy amplio y escapa el alcance del presente texto, en el Capítulo 9 presentamos una breve descripción del mismo pero referimos al lector interesado a revisar artículos como Vahid y Engle (1991), y para el caso de México, Castillo y Herrera (2005), Herrera (2004), o Herrera y Castillo (2003).

- Engle- Granger

Esta metodología consiste en instrumentar una estimación bietápica que prueba la idea fundamental de cointegración; que una combinación lineal de series no-estacionarias es en sí misma estacionaria. Para introducir la metodología a continuación se ilustra un caso bivariado.

Considera dos series integradas de orden 1,  $y_t \sim I(1)$  y  $x_t \sim I(1)$ . Podemos representar una ecuación de regresión entre las mismas como sigue:<sup>12</sup>

$$y_t = \gamma x_t + \varepsilon_t \tag{7.3}$$

Donde  $\varepsilon_t$  representa el término de error que suponemos es ruido blanco.

La primera etapa de la metodología consiste en estimar la ecuación (7.3) por medio de MCO's y obtener los errores estimados,  $\hat{\varepsilon}_t$ . En la segunda etapa se realiza una prueba de estacionariedad para  $\hat{\varepsilon}_t$ ; si el término de error es estacionario, entonces se concluye que  $y_t$  y  $x_t$  se encuentran cointegradas, y la elasticidad de largo plazo de  $y_t$  con respecto a  $x_t$  es el coeficiente  $\gamma$ . Si por el contrario  $\hat{\varepsilon}_t$  resulta no-estacionario entonces se concluye que  $y_t$  y  $x_t$  no comparten una tendencia común, y por ende los resultados de la primera etapa no son confiables, en otras palabras, la relación entre  $y_t$  y  $x_t$  es espuria.

---

<sup>12</sup> En este ejemplo no incluimos constante para facilitar la exposición. Sin embargo notamos que los resultados que se obtiene aquí se aplican estrictamente al caso en el que la ecuación incluye una constante.

Nótese que la ecuación (7.3) se puede re-escribir como  $y_t - \alpha_t = \varepsilon_t$  de tal forma que  $\varepsilon_t$  representa una combinación lineal de  $y_t$  y  $x_t$ , así que al estimar  $\hat{\varepsilon}_t$  y realizar la prueba de estacionariedad sobre la serie, se está verificando el concepto mismo de cointegración, la existencia de una combinación lineal de  $y_t$  y  $x_t$  que es estacionaria. Cabe destacar que para realizar la prueba de estacionariedad sobre  $\hat{\varepsilon}_t$  los valores críticos no son los que reportan los programas econométricos comunes, sino los que reportan Phillips y Ouliaris (1990).

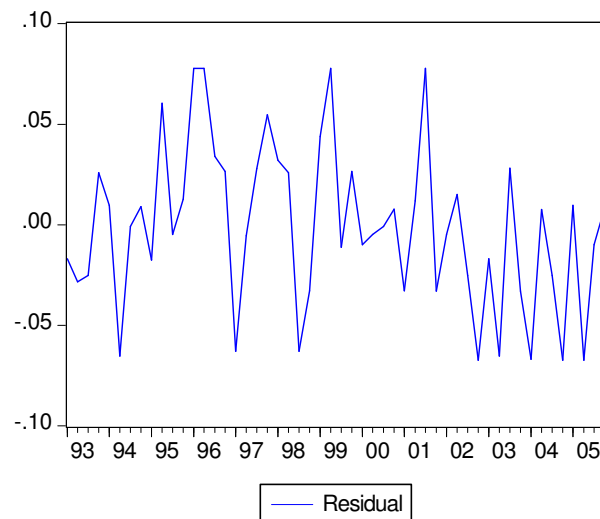
Para ilustrar la instrumentación de la metodología, aquí continuamos con nuestro ejemplo de consumo e ingreso. Siguiendo con el procedimiento anteriormente descrito, en la primera etapa estimamos la ecuación  $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$  por medio de MCO's obteniendo los resultados que se presentaron en el Cuadro 6.2; y que a aquí se reproducen:

**Cuadro 7.1 Resultados de  $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$**

Dependent Variable: LOG(CONSUMO)				
Method: Least Squares				
Date: 08/10/06 Time: 13:23				
Sample (adjusted): 1993Q1 2005Q4				
Included observations: 52 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.252331	0.862475	-2.611472	0.0119
LOG(PIB)	1.089860	0.040848	26.68078	0.0000
R-squared	0.934371	Mean dependent var		20.75881
Adjusted R-squared	0.933059	S.D. dependent var		0.137516
S.E. of regresión	0.035579	Akaike info criterion		-3.796396
Sum squared resid	0.063295	Schwarz criterion		-3.721348
Log likelihood	100.7063	F-statistic		711.8638
Durbin-Watson stat	0.862964	Prob(F-statistic)		0.000000

A continuación creamos la serie de residuales estimados,  $\hat{\varepsilon}_t$ , y realizamos la prueba de estacionariedad de acuerdo a lo sugerido en la segunda etapa. Para darnos una idea del comportamiento de  $\hat{\varepsilon}_t$ , preliminar a llevar a cabo una prueba formal de raíces unitarias, en la Gráfica 7.1 se presenta una ilustración de los mismos. Razonablemente, podríamos indicar que la serie de residuales es estacionaria, ya que fluctúa alrededor de una media constante, que en este caso es 0.

**Gráfica 7.1. Serie de Residuales**



Para verificar nuestra sospecha, llevamos a cabo la prueba DFA obteniendo el estadístico de prueba,  $t$ , y comparándolo con valores críticos de Phillips y Ouliaris (1990). Tomando un ejemplo, para 50 observaciones los valores críticos proporcionados por los autores son -4.085 al 1%, -3.438 al 5%, y -3.094 al 10%. Al realizar la prueba DFA sobre los residuales se obtiene un estadístico de prueba de -5.91, el cual es mayor a cualquiera de los valores críticos listados. Como se recordará, la hipótesis nula en la prueba DFA es la existencia de raíz unitaria, la cual se rechaza de acuerdo a los resultados obtenidos. Así, se puede

concluir que  $\hat{\varepsilon}_t$  es una serie estacionaria y por tanto las series de consumo e ingreso se encuentran cointegradas con coeficientes de cointegración (1, -1). En otras palabras, la elasticidad de largo plazo del consumo con respecto al ingreso es de 1.

Para el caso multivariado la idea es, en esencia, la anteriormente presentada. La ecuación a estimar en la primera etapa para un sistema de  $n$  variables sería:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_n x_{nt} + \varepsilon_t \quad (7.4)$$

Habiendo estimado la ecuación por medio de MCO's, se realizaría la prueba de estacionariedad sobre  $\hat{\varepsilon}_t$ . Si se encontrara que la serie es estacionaria, las elasticidades de largo plazo serían los coeficientes  $\beta_i$  para  $i = 1, \dots, n$  (suponiendo que tanto la variable dependiente como las independientes se expresen en logaritmos).

Ilustramos la estimación continuando con nuestro ejemplo del consumo, pero ahora incluimos adicionalmente al PIB el agregado monetario M1 en términos reales, esto es, M1 nominal ajustado por el Índice Nacional de Precios al Consumidor.<sup>13</sup> Ambas series provienen de la base de datos de indicadores económicos y financieros de Banco de México. Primero verificamos la naturaleza estocástica de M1 real, previa transformación logarítmica, presentando los resultados de la prueba KPSS en el Cuadro 7.2. La hipótesis nula de estacionariedad se rechaza para la serie en niveles al 5%, para la serie en primeras

---

<sup>13</sup> M1 real puede considerarse como una medida de riqueza o profundidad del mercado de capitales. Ver Castillo (2003) para una discusión detallada sobre los argumentos que podrían incluirse en una función de consumo.

diferencias la misma no se rechaza a ningún nivel de confianza, por lo cual concluimos que la serie de M1 real es I(1).

**Cuadro 7.2. Prueba de Raíz Unitaria KPSS**

Null Hypothesis: M1/INPCPROM is stationary		
Exogenous: Constant		
Bandwidth: 48.2 (Andrews using Bartlett kernel)		
		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.472917
Asymptotic critical values*:	1% level	0.739000
	5% level	0.463000
	10% level	0.347000
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)		

Null Hypothesis: D(M1/INPCPROM) is stationary		
Exogenous: Constant		
Bandwidth: 4.14 (Andrews using Bartlett kernel)		
		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.287373
Asymptotic critical values*:	1% level	0.739000
	5% level	0.463000
	10% level	0.347000
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)		

Enseguida estimamos una ecuación como (7.4) considerando al consumo como variable dependiente y al PIB y M1 real como explicativas. Los resultados se presentan en el Cuadro 7.3. El PIB continúa presentando una relación positiva con el consumo pero el coeficiente es ahora .75, que es menor al coeficiente de 1 que se obtiene al estimar la ecuación sin el agregado monetario. Con respecto a M1, se encuentra una relación positiva y significativa con una elasticidad de .19.

**Cuadro 7.3. Resultados de  $y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t$** 

Dependent Variable: LOG(CONSUMO)				
Method: Least Squares				
Date: 08/25/06 Time: 13:56				
Sample (adjusted): 1993Q1 2005Q4				
Included observations: 52 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.498250	0.603214	2.483779	0.0165
LOG(PIB)	0.752463	0.055668	13.51688	0.0000
LOG(MIREAL)	0.192878	0.027174	7.097978	0.0000
R-squared	0.967642	Mean dependent var		13.85105
Adjusted R-squared	0.966321	S.D. dependent var		0.137516
S.E. of regression	0.025237	Akaike info criterion		-4.465078
Sum squared resid	0.031208	Schwarz criterion		-4.352506
Log likelihood	119.0920	F-statistic		732.6504
Durbin-Watson stat	1.631473	Prob(F-statistic)		0.000000

A continuación obtenemos la serie de errores estimados y realizamos las pruebas de raíz unitaria sobre ella. Los resultados se presentan en el Cuadro 7.4. El estadístico de prueba de -3.55 indica el rechazo de la hipótesis nula de no-estacionariedad (no de acuerdo a los valores críticos que se presentan en el Cuadro, pero a los que corresponden a Phillips y Ouliaris (1990)), por lo que se puede concluir razonablemente que las series en el sistema cointegran, con un vector normalizado con respecto al consumo de (1, -.75, -.19).

**Cuadro 7.4. Prueba de Raíz Unitaria ADF para Residuales**

Null Hypothesis: RESID02 has a unit root		
Exogenous: None		
Lag Length: 5 (Automatic based on SIC, MAXLAG=10)		
		t-Statistic
		Prob.*
Augmented Dickey-Fuller test statistic		-3.556704
Test critical values:	1% level	-2.616203
	5% level	-1.948140
	10% level	-1.612320
*MacKinnon (1996) one-sided p-values.		

A pesar de que la instrumentación de la metodología de Engle y Granger es técnicamente amigable, en la actualidad es poco utilizada, principalmente por dos razones: la primera tiene que ver con un problema de normalización, esto es, nótese que las variables endógena y exógenas se determinan a priori, y el vector de cointegración se normaliza con respecto a la variable endógena. El problema es que el vector generalmente cambia cuando se modifica el orden de las variables; de tal forma que cuando se invierten las variables los coeficientes de cointegración no son necesariamente el recíproco de los coeficientes que se obtuvieron originalmente. Esto es, la elección de la variable sobre la cual se realiza la normalización es determinante en el resultado de los coeficientes de cointegración. Para ilustrar este hecho puntualmente, re-estimemos la ecuación que sólo considera al consumo e ingreso pero considerando al ingreso como variable endógena y al consumo como la exógena. Los resultados de la estimación se presentan en el Cuadro 7.5.

**Cuadro 7.5. Resultados de  $x_t = \phi_0 + \phi_1 y_t + \varepsilon_t$** 

Dependent Variable: LOG(PIB)				
Method: Least Squares				
Date: 08/10/06 Time: 15:17				
Sample (adjusted): 1993Q1 2005Q4				
Included observations: 52 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.316666	0.667056	4.972096	0.0000
LOG(CONSUMO)	0.857332	0.032133	26.68078	0.0000
R-squared	0.934371	Mean dependent var		21.11385
Adjusted R-squared	0.933059	S.D. dependent var		0.121967
S.E. of regresión	0.031556	Akaike info criterion		-4.036375
Sum squared resid	0.049790	Schwarz criterion		-3.961327
Log likelihood	106.9457	F-statistic		711.8638
Durbin-Watson stat	0.866188	Prob(F-statistic)		0.000000

Ahora los coeficientes de cointegración son (1,-.86), lo cual no es congruente con lo que se obtuvo anteriormente. Si entendemos a los coeficientes como “ponderadores” que permiten que una combinación de las variables sea estacionaria, entonces deberíamos esperar que al invertir el orden de las variables de consumo e ingreso, los coeficientes fueran una vez más 1 y -1.

La segunda razón se refiere al hecho de que por medio de esta estimación no es posible determinar la relación de corto plazo entre las variables, esto es, la ecuación que se estima está diseñada para determinar relaciones de largo plazo, pero no incluye términos que pudieran indicar la dinámica de corto plazo existente en el sistema.<sup>14</sup>

<sup>14</sup> Otro problema potencial con la estimación simple del procedimiento de dos etapas, es que la distribución de los parámetros del vector de cointegración pudieran no distribuirse normalmente, por lo cual sería necesario ajustar por esta variación. Engle y Yoo (1991) proveen una sugerencia para subsanar este posible

En la práctica no es trivial atender el problema de normalización, por lo que generalmente lo que procede es realizar pruebas de exogeneidad sobre las variables, y así determinar con mayor precisión el orden de las mismas en el sistema. Con respecto a la segunda deficiencia (que de hecho sólo es inconveniente cuando se contempla analizar relaciones de corto plazo) es posible instrumentar un Modelo de Corrección de Error (MCE), que permite modelar explícitamente la dinámica de corto plazo. En seguida se presenta su desarrollo.

- Modelo de Corrección de Error

El fundamento teórico de un MCE se basa en el argumento que es posible modelar la dinámica de un sistema en término de sus desviaciones de un estado estacionario o de equilibrio. Esto es, nos imaginamos que en el largo plazo el sistema de variables fluctúa alrededor de una tendencia común que describe su estado de equilibrio, y que sólo se desvía temporalmente debido a choques transitorios. Para formalizar la idea consideremos el siguiente modelo:

$$y_t - y_{t-1} = \phi_1 \Delta x_t + \phi_2 (y_{t-1} - y^*_{t-1}) + \varepsilon_t \quad (7.5)$$

Supongamos que existe un nivel deseado de  $y_t$  que denotamos  $y^*$ . En tal caso podemos interpretar el término  $y_{t-1} - y^*_{t-1}$  como un desequilibrio rezagado, de tal forma que el parámetro  $\phi_2$  sería un coeficiente de ajuste, o de corrección de error, de ahí el término del modelo. Para que el sistema converja, el coeficiente del término de corrección de error

---

problema, el cual básicamente consiste en incluir un tercer paso que aproxima la distribución de los parámetros a una normal.

tendría que exhibir ciertas características. En particular,  $\phi_2$  deberá ser menor a uno en términos absolutos (que es en realidad una condición de estacionariedad, como se discutió en el análisis de modelos univariados), negativo (de tal forma que si el sistema se encuentra por arriba (debajo) de su tendencia de largo plazo, el mismo regrese a su trayectoria de equilibrio disminuyendo (aumentando) su nivel), y estadísticamente significativo.

Existen varios modelos que hacen referencia a este concepto y que se estiman de diversas formas. A continuación se describen dos de los más frecuentemente utilizados.

a) MCE en dos etapas bajo el espíritu de Engle y Granger

Este modelo es similar al propuesto por Engle y Granger en el sentido que requiere de la implementación de dos etapas para probar la existencia de cointegración. Partimos de un modelo bivariado como sigue:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \quad (7.6)$$

Nótese que podemos seguir pensando en  $\varepsilon_t$  como una caracterización de la relación lineal entre  $y_t$  y  $x_t$ , esto es

$$\varepsilon_t = y_t - \beta_0 - \beta_1 x_t$$

Así, reparametrizando la ecuación en primeras diferencias se tiene

$$\Delta y_t = \alpha_1 \Delta x_t + \alpha_2 (y_{t-1} - \beta_0 - \beta_1 x_{t-1}) + \eta_t \quad (7.7)$$

Y el término en paréntesis representa lo que anteriormente llamamos término de desequilibrio o de corrección de error, siendo en este caso  $\alpha_2$  el coeficiente del mismo.

Esta expresión se simplifica como

$$\Delta y_t = \alpha_0 + \alpha_1 \Delta x_t + \alpha_2 \varepsilon_{t-1} + \eta_t \quad (7.8)$$

De tal forma que el parámetro  $\alpha_2$  es el coeficiente de ajuste. Otra interpretación que se le puede dar a  $\alpha_2$  es el de término de velocidad de ajuste, ya que representa la magnitud por la que el sistema regresa a su estado estacionario en cada periodo. Así, si por ejemplo  $\alpha_2$  presenta una magnitud de .75, esto indicaría que el 75% de la desviación de equilibrio se recupera en el primer periodo, 75% de la desviación restante se recupera en el segundo periodo, y así sucesivamente.

La estimación del modelo se realiza en dos etapas. En la primera se estima la ecuación (7.6) y se obtienen el error estimado,  $\varepsilon_t$ . En la segunda etapa se estima la ecuación en primeras diferencias  $\Delta y_t = \alpha_0 + \alpha_1 \Delta x_t + \alpha_2 \varepsilon_{t-1} + \eta_t$  y se evalúa la existencia de cointegración de acuerdo a las características estadísticas de  $\alpha_2$ . Nótese que en la ecuación para la segunda etapa las variables en el sistema son estacionarias. En la práctica es común

incluir rezagos de las variables endógenas y exógenas, de tal forma que una especificación elaborada sería como la que sigue<sup>15</sup>:

$$\Delta y_t = \gamma_0 + \gamma_1 \Delta x_t + \gamma_2 \Delta y_{t-1} + \gamma_3 \Delta x_{t-1} + \phi \varepsilon_{t-1} + \eta_t \quad (7.9)$$

De tal suerte que la prueba de cointegración se realizaría sobre el coeficiente  $\phi$ . Si el coeficiente resulta con magnitud menor a uno, negativo, y estadísticamente significativo, entonces se concluiría que  $y_t$  y  $x_t$  cointegran con coeficientes de cointegración normalizados  $(1, -\beta_1)$  y velocidad de ajuste de  $\hat{\phi}$ . Para ilustrar la instrumentación de esta metodología continuamos con el ejemplo de la relación entre el consumo y el ingreso.<sup>16</sup>

En la primera etapa estimamos  $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$  por MCO's obteniendo los resultados que se presentaron en el Cuadro 6.2 y que se replicaron en el Cuadro 7.1. De esta estimación construimos la serie de errores estimados y los incluimos en una ecuación como (7.9). Los resultados se presentan en el Cuadro 7.6; donde RESID(-1) se refiere a  $\hat{\varepsilon}_{t-1}$ .

---

<sup>15</sup> Formalmente el número de rezagos a incluir en primeras diferencias sería igual al número de rezagos óptimos que se identificaran en el Vector Autorregresivo en Niveles,  $p$ , menos 1.

<sup>16</sup> Un modelo similar se presenta en Conesa (1998) donde se analiza el nivel de traspaso del tipo de cambio y los salarios al nivel de inflación.

**Cuadro 7.6. Resultados de  $\Delta y_t = \beta_0 + \beta_1 \Delta x_t + \beta_2 \Delta x_{t-1} + \beta_3 y_{t-1} + \phi e_{t-1} + \eta_t$**

Dependent Variable: DLOG(CONSUMO)				
Method: Least Squares				
Date: 08/14/06 Time: 14:53				
Sample (adjusted): 1993Q3 2005Q4				
Included observations: 50 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.006365	0.003481	1.828230	0.0741
DLOG(PIB)	0.554145	0.100475	5.515242	0.0000
DLOG(PIB(-1))	-0.700525	0.151805	-4.614627	0.0000
DLOG(CONSUMO(-1))	0.198157	0.117716	1.683343	0.0992
RESIDEG(-1)	-0.528011	0.110710	-4.769306	0.0000
R-squared	0.810055	Mean dependent var		0.008135
Adjusted R-squared	0.793171	S.D. dependent var		0.050163
S.E. of regresión	0.022813	Akaike info criterion		-4.628310
Sum squared resid	0.023420	Schwarz criterion		-4.437108
Log likelihood	120.7078	F-statistic		47.97763
Durbin-Watson stat	1.556520	Prob(F-statistic)		0.000000

Nótese que el coeficiente del término de error es igual a -0.53, y cumple con las condiciones necesarias para identificar cointegración; menor a uno en términos absolutos, negativo, y estadísticamente significativo. Así, se prueba que el consumo y el ingreso se encuentran cointegrados con coeficientes (1,-1).

La velocidad de ajuste en este caso es aceptable, de tal forma que un 53% de la desviación de equilibrio se recupera en el primer periodo y así sucesivamente. Se dice que es aceptable por que si hiciéramos el cálculo del número de periodos que tomaría al sistema regresar a su estado de equilibrio, después de un choque que lo desviara, sería menor al número de

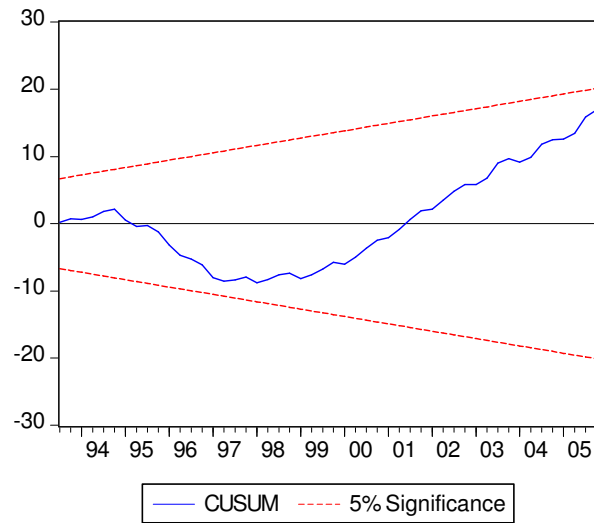
observaciones que se incluyen en la estimación.<sup>17</sup> De tal forma que es importante en la práctica asegurarse que la velocidad de ajuste estimada es lo suficientemente rápida para reestablecer equilibrio en el sistema dentro del periodo que se considera en el análisis. Si tuviéramos por ejemplo una velocidad de ajuste con magnitud 0.05, entonces necesitaríamos un número muy amplio de observaciones para que el sistema tuviera tiempo de regresar a su estado estacionario, ya que sólo el 5 por ciento de la desviación se recuperaría cada periodo.

Cabe mencionar que tanto para el modelo de Engle y Granger como para el MCE es recomendable realizar las pruebas pertinentes para los residuales; así como las pruebas de estabilidad de coeficientes. Esta última es de particular interés pues después de todo estamos estimando relaciones de largo plazo, ¿cómo podemos argumentar que los coeficientes de cointegración son de hecho constantes a través del tiempo? Alguien familiarizado con la dinámica del consumo y del ingreso en México, por ejemplo, podría señalar que es poco probable que el vector de cointegración sea estable debido al cambio estructural que presentan las series en el periodo 1994-1995, y tendría razón. Así que para verificar la estabilidad de la función estimada realizamos las pruebas de estabilidad recursiva CUSUM y CUSUM cuadrado. Las Gráficas 7.2 y 7.3 muestran los resultados.

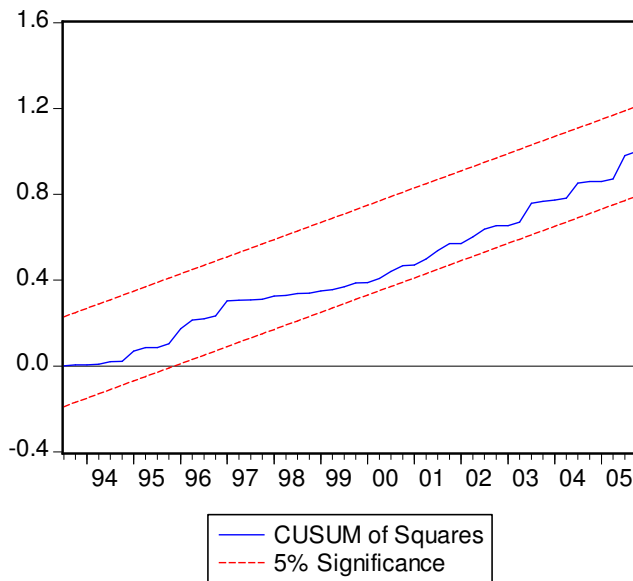
---

<sup>17</sup> Pregunta para el lector ávido de aprender ¿Cuántos periodos se demoraría el sistema en regresar a su estado estacionario siendo que la velocidad de ajuste es .53?

**Gráfica 7.2 Prueba CUSUM**



**Gráfica 7.3 Prueba CUSUM Cuadrado**



Como se puede observar, la línea de prueba se mantiene dentro de las bandas de tolerancia al 5% de significancia, de tal forma que podemos indicar que la función es estable. Claro

que en la práctica sería recomendable realizar otras pruebas de estabilidad disponibles en paquetes econométricos.<sup>18</sup>

b) MCE a partir de un modelo autorregresivo de rezagos distribuidos

Este modelo es una extensión del concepto de corrección de error. Aquí se parte de una representación del sistema como un proceso autorregresivo de rezagos distribuidos (ARD) y se realizan estimaciones recursivas yendo de lo general a lo particular hasta encontrar una especificación aceptable. Formalmente, suponemos que el ARD se representa de la siguiente forma:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t \quad (7.10)$$

Remplazando  $y_t$  con  $y_{t-1} + \Delta y_t$  y  $x_t$  con  $x_{t-1} + \Delta x_t$  en (7.10) obtenemos

$$y_{t-1} + \Delta y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_0 (x_{t-1} + \Delta x_t) + \beta_1 x_{t-1} + \varepsilon_t$$

Resolviendo por  $\Delta y_t$  se tiene

$$\Delta y_t = \alpha_0 + y_{t-1}(\alpha_1 - 1) + \beta_0 \Delta x_t + \beta_1 x_{t-1} + \varepsilon_t$$

---

<sup>18</sup> Existen varias pruebas formales y sofisticadas diseñadas para determinar la estabilidad de un vector de cointegración, referirse por ejemplo a Hansen y Johansen (1992, 1999).

Que a su vez se simplifica de ésta forma:

$$\Delta y_t = \alpha_0 + \beta_0 \Delta x_t - (1 - \alpha_1) y_{t-1} + (\beta_0 + \beta_1) x_{t-1} + \varepsilon_t \quad (7.11)$$

Multiplicando y dividiendo (7.11) por  $(1 - \alpha_1)$  la ecuación se puede re-escribir como

$$\Delta y_t = \beta_0 \Delta x_t - (1 - \alpha_1) \left[ y_{t-1} - \frac{\alpha_0}{1 - \alpha_1} - \left( \frac{\beta_0 + \beta_1}{1 - \alpha_1} \right) x_{t-1} \right] + \varepsilon_t \quad (7.12)$$

Nótese que en el término en corchetes las variables  $y_t$  y  $x_t$  se encuentran rezagadas un

periodo, de tal forma que podemos pensar en la expresión  $y_{t-1} - \frac{\alpha_0}{1 - \alpha_1} - \left( \frac{\beta_0 + \beta_1}{1 - \alpha_1} \right) x_{t-1}$

como el término de corrección de error. El modelo en (7.12) se puede estimar por medio de MCO's, especificando la ecuación tal y como aparece. Alternativamente, se puede estimar (7.10) y obtener directamente los estimadores. Pero ¿cómo se calculan las elasticidades de largo plazo a partir de las estimaciones? Para simplificar el proceso, podemos reparametrizar las expresiones anteriores y escribirlas en una forma a estimar generalizada que es aparentemente más elaborada, pero que en efecto es más amigable. En niveles sería una expresión como

$$y_t = \alpha_0 + \alpha_1^* y_{t-1} + \beta_1 x_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_n \Delta y_{t-n} + \varphi_0 \Delta x_t + \varphi_1 \Delta x_{t-1} + \dots + \varphi_k \Delta x_{t-k} + \psi D_t + \phi T + \varepsilon_t \quad (7.13)$$

Y en primera diferencias como

$$\begin{aligned} \Delta y_t = & \alpha_0 + \alpha_1 y_{t-1} + \beta_1 x_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_n \Delta y_{t-n} + \varphi_0 \Delta x_t + \varphi_1 \Delta x_{t-1} + \dots \\ & + \varphi_k \Delta x_{t-k} + \psi D_t + \phi T + \varepsilon_t \end{aligned} \quad (7.14)$$

En ambas expresiones se incluyen dos términos que “enriquecen” la ecuación,  $D_t$  que es un vector de variables determinísticas y  $T$  que es una tendencia. La estimación de estas ecuaciones se hace recursivamente, comenzando por incluir rezagos de las diferencias de  $y_t$  y  $x_t$ , tantas como sea “razonable”. Comenzamos entonces con un proceso de depuración eliminando las variables que no sean significativas. Este proceso continúa hasta encontrar una ecuación satisfactoria, no sólo en términos de la significancia de los coeficientes pero en términos de su eficiencia y estabilidad. Esto es, preferentemente debe ser una ecuación que apruebe las pruebas convencionales sobre los residuales, y que sea estable.

Una vez que se ha estimado una ecuación satisfactoria (o de hecho conjuntamente) se determina la existencia de cointegración, la cual se identifica al analizar el coeficiente de la variable dependiente en niveles rezagada un periodo,  $y_{t-1}$ , el cual es  $\alpha_1$  en nuestro ejemplo para la ecuación (7.14). Para el caso de la ecuación (7.13)  $\alpha_1$  se obtiene a partir de  $\alpha_1^*$  notando que es igual al negativo de  $1 - \alpha_1^*$ . Por ejemplo, si  $\alpha_1^*$  toma el valor de 0.52, entonces  $\alpha_1$  es igual a -0.42.

Al igual que en el caso del coeficiente de ajuste en el MCE de dos etapas, la existencia de cointegración se determina examinando  $\alpha_1$ , que sea menor a 1 en valor absoluto, negativo, y estadísticamente significativo.<sup>19</sup> Para el caso de la ecuación (7.13) la significancia de  $\alpha_1$  se infiere a partir de la significancia de  $\alpha_1^*$ . Aquí hay que puntualizar que Bewley (1979) mostró que la distribución de los parámetros en estas estimaciones no es normal, por lo que se debe realizar una transformación para determinar la significancia de los estimados. Aunque no presentamos el detalle de la misma, indicamos que para fines prácticos valores mayores a 3 en términos absolutos pueden considerarse aceptables.<sup>20</sup> El lector cauteloso debe preguntarse a este punto ¿por qué es posible calcular  $\alpha_1$  a partir de  $\alpha_1^*$ ? simplemente por que, como se indicó en la presentación formal del modelo, estas expresiones son simplemente una reparametrización la una de la otra, de tal forma que los coeficientes y elasticidades pueden calcularse en base a operaciones aritméticas simples.

La elasticidad de largo plazo de  $y_t$  con respecto a  $x_t$  se obtienen dividiendo el coeficiente de  $x_{t-1}$  por el valor absoluto de  $\alpha_1$ .<sup>21</sup> Los coeficientes  $\gamma$ 's representan las elasticidades de corto plazo de  $y_t$  con respecto a su historial, y los coeficientes  $\varphi$  las correspondientes a la historia de  $x_t$ .

---

<sup>19</sup> Nótese que a través de los diferentes modelos que hemos tratado en el texto, incluyendo los modelos univariados, esta condición de estabilidad es la misma, lo cual no es una coincidencia, parte del hecho que es una condición de convergencia, sin la cual los modelos considerados no tendrían sentido.

<sup>20</sup> Referimos al lector al artículo original.

<sup>21</sup> Seguimos suponiendo que las estimaciones se realizan empleando las transformaciones logarítmicas de las variables.

Para ejemplificar la estimación y convencer al lector de que las reparametrizaciones son, de hecho, transformaciones lineales la una de la otra, y por lo cual es posible calcular los coeficientes de ajuste aritméticamente, estimamos a continuación las ecuaciones para el caso del consumo y el ingreso.

Iniciamos con la estimación de la siguiente ecuación:

$$\Delta y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_1 x_{t-1} + \gamma_1 \Delta y_{t-1} + \gamma_2 \Delta y_{t-2} + \gamma_3 \Delta y_{t-3} + \gamma_4 \Delta y_{t-4} + \varphi_0 \Delta x_t + \varphi_1 \Delta x_{t-1} + \varphi_2 \Delta x_{t-2} + \varphi_3 \Delta x_{t-3} + \varphi_4 \Delta x_{t-4} + \varepsilon_t$$

Donde la variable  $y_t$  se refiere al consumo y  $x_t$  al PIB. En esta primera aproximación se incluyen hasta 4 rezagos de ambas variables en diferencias; lo cual es estándar en la práctica. Como se indicó en su oportunidad, si la periodicidad de las series es trimestral, se procura incluir al menos un año de historia, se fuera mensual, entonces hasta 12 periodos sería razonable. Los resultados de la estimación se presentan en el Cuadro 7.7.

### Cuadro 7.6. Resultados de Ecuación

Dependent Variable: DLOG(CONSUMO)				
Method: Least Squares				
Date: 08/16/06 Time: 16:21				
Sample (adjusted): 1994Q2 2005Q4				
Included observations: 47 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.965261	0.552230	-3.558772	0.0011
LOG(CONSUMO(-1))	-0.147276	0.086108	-1.710375	0.0960
LOG(PIB(-1))	0.237851	0.099033	2.401723	0.0218
DLOG(CONSUMO(-1))	-0.387529	0.166869	-2.322355	0.0262
DLOG(CONSUMO(-2))	-0.600943	0.165819	-3.624087	0.0009
DLOG(CONSUMO(-3))	-0.448619	0.157886	-2.841409	0.0074
DLOG(CONSUMO(-4))	0.272184	0.154411	1.762730	0.0867
DLOG(PIB)	1.041414	0.097241	10.70959	0.0000
DLOG(PIB(-1))	0.308814	0.203805	1.515243	0.1387
DLOG(PIB(-2))	0.635366	0.183469	3.463061	0.0014
DLOG(PIB(-3))	0.499436	0.181252	2.755469	0.0092
DLOG(PIB(-4))	-0.283572	0.183309	-1.546962	0.1309
R-squared	0.955882	Mean dependent var		0.008453
Adjusted R-squared	0.942016	S.D. dependent var		0.051290
S.E. of regression	0.012351	Akaike info criterion		-5.734385
Sum squared resid	0.005339	Schwarz criterion		-5.262007
Log likelihood	146.7581	F-statistic		68.93892
Durbin-Watson stat	2.188701	Prob(F-statistic)		0.000000

Observamos que varios coeficientes resultan no significativos, aun considerando los valores críticos estándar, tal es el caso del coeficiente de la variable de consumo en niveles rezagada un periodo,  $\alpha_1$ . Proseguimos entonces a eliminar variables que no resultan significativas y a identificar una ecuación que sea satisfactoria. Después de un arduo proceso de eliminación se encontró la siguiente ecuación:

$$\Delta y_{t-1} = \alpha_0 + \alpha_1 y_{t-1} + \beta_1 x_{t-1} + \gamma_2 \Delta y_{t-2} + \varphi_0 \Delta x_t + \varphi_1 \Delta x_{t-1} + \varepsilon_t \quad (7.15)$$

Los resultados de su estimación se presentan en el Cuadro 7.8.

**Cuadro 7.8. Resultados de  $\Delta y_{t-1} = \alpha_0 + \alpha_1 y_{t-1} + \beta_1 x_{t-1} + \gamma_2 \Delta y_{t-2} + \varphi_0 \Delta x_t + \varphi_1 \Delta x_{t-1} + \varepsilon_t$**

Dependent Variable: DLOG(CONSUMO)				
Method: Least Squares				
Date: 08/16/06 Time: 14:41				
Sample (adjusted): 1993Q4 2005Q4				
Included observations: 49 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.395092	0.631166	-3.794710	0.0005
LOG(CONSUMO(-1))	-0.478852	0.091156	-5.253125	0.0000
LOG(PIB(-1))	0.584506	0.105545	5.537966	0.0000
DLOG(CONSUMO(-2))	-0.289049	0.084529	-3.419541	0.0014
DLOG(PIB)	0.852203	0.117004	7.283538	0.0000
DLOG(PIB(-1))	-0.516007	0.094571	-5.456307	0.0000
R-squared	0.842549	Mean dependent var		0.008433
Adjusted R-squared	0.824241	S.D. dependent var		0.050638
S.E. of regresión	0.021229	Akaike info criterion		-4.752594
Sum squared resid	0.019379	Schwarz criterion		-4.520943
Log likelihood	122.4386	F-statistic		46.02021
Durbin-Watson stat	1.651245	Prob(F-statistic)		0.000000

Notamos primero que la totalidad de los coeficientes presentan estadísticos de prueba mayores a 3 en términos absolutos, lo cual nos da una aceptable idea de su significancia. El coeficiente de ajuste,  $\alpha_1$ , es significativo, negativo y con un valor menor a la unidad en valor absoluto; lo cual confirma la existencia de cointegración. Para calcular la elasticidad de  $y_t$  con respecto a  $x_t$  procedemos como se indicó anteriormente, dividimos el coeficiente de  $x_{t-1}$  por el valor absoluto de  $\alpha_1$ , esto es  $0.58/0.48$  teniendo como resultado 1.2. Con respecto a la dinámica de corto plazo, notamos un resultado interesante, que el efecto de la diferencia logarítmica del PIB rezagado un periodo sobre el consumo es negativo. Esto, en principio, puede parecer sospechoso, pues es difícil argumentar que el

efecto del PIB sobre el consumo sea negativo (este mismo efecto se observa para el coeficiente del logaritmo del consumo rezagado dos periodos), sin embargo es importante destacar que la interpretación de los términos de corto plazo no es convencional, ya que representan cambios transitorios. De tal forma que puede ser el caso que un choque transitorio al PIB con un rezago de dos periodos pudiera tener un efecto negativo en el consumo contemporáneo. El resto de los estadísticos de importancia como lo son la  $r^2$ , DW, y estadística  $F$  parecen aceptables.

Con el propósito de verificar la robustez de la ecuación a continuación se presenta una batería estándar de pruebas de residuales y estabilidad.

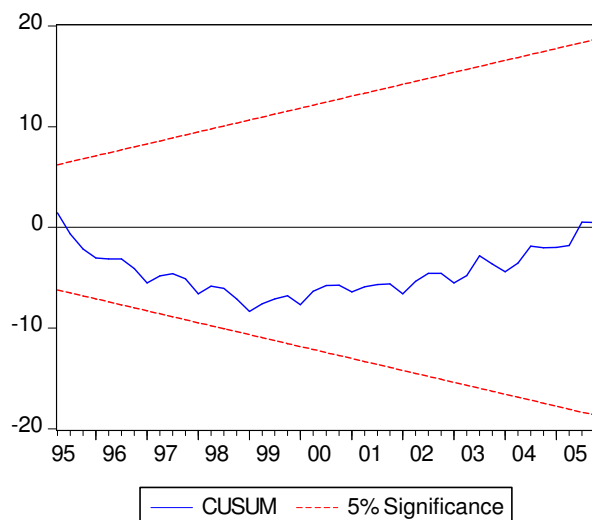
**Cuadro 7.9. Pruebas sobre Residuales**

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	2.705514	Prob. F(4,39)	0.044158
Obs*R-squared	10.64349	Prob. Chi-Square(4)	0.030877

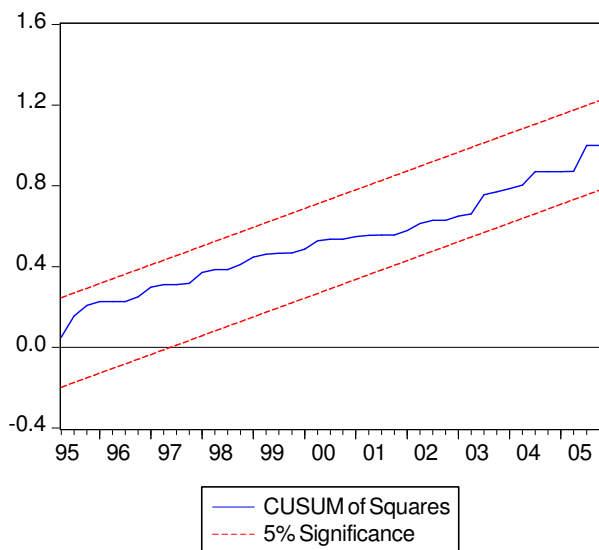
ARCH Test:			
F-statistic	2.188866	Prob. F(4,40)	0.087639
Obs*R-squared	8.081061	Prob. Chi-Square(4)	0.088654

White Heteroskedasticity Test:			
F-statistic	1.764903	Prob. F(9,39)	0.106851
Obs*R-squared	14.18119	Prob. Chi-Square(9)	0.116024

**Gráfica 7.4 Prueba CUSUM**



**Gráfica 7.5 Prueba CUSUM Cuadrado**



En general se obtienen resultados aceptables, con excepción de la presencia de cierta evidencia de un problema de correlación. ¿Podemos entonces estar satisfechos con esta ecuación? En realidad no, pues nótese que la elasticidad del consumo con respecto al PIB es mayor a 1, que no es acorde a lo que la teoría económica sugiere. En tal caso se podría proseguir en dos vertientes, una consistiría en continuar “buscando” una especificación

aceptable que produzca una elasticidad del consumo con respecto al PIB acorde con lo esperado, o alternativamente podríamos complementar la función de consumo con variables adicionales que pudieran ser parte de la misma. En el Cuadro 7.10 presentamos una estimación que incluye adicionalmente al PIB, el tipo de cambio real (tcr) y dos series binarias que controlan por estacionalidad en el primer y cuarto trimestre del año, @seas(1) y @seas(4) respectivamente.<sup>22</sup> Asimismo, ampliamos el periodo de estimación a 1986. En el Cuadro 7.11 se muestran las pruebas de residuales y en las Gráficas 7.6 y 7.7 las pruebas de estabilidad CUSUM y CUSUM cuadrado respectivamente.

Como puede apreciarse, la ecuación no presenta problemas de autocorrelación o ARCH para 2 rezagos; la prueba de heteroscedasticidad sugiere la necesidad de estimar la ecuación ajustando los errores estándar; y de hecho los resultados que aparecen en el Cuadro ya reflejan este ajuste. Las ilustraciones de estabilidad sugieren que la especificación es robusta, aunque para la prueba CUSUM cuadrado parece que la línea de prueba toca la banda de confianza alrededor del año 2001.

---

<sup>22</sup> Cabe señalar que en este ejercicio las variables que controlan por estacionalidad no son centradas, esto es, la suma de sus valores para cada año no es 0. En este caso simplemente tomamos las variables que asignan el valor de 1 para el trimestre de interés y 0 para cualquier otro trimestre. En la práctica, se recomienda emplear variables centradas. Para nuestro ejemplo, la variable que corresponde al primer trimestre del año tendría que tomar el valor de .75 en ese trimestre y de -.25 en cada uno de los tres trimestres restantes; de tal forma que la suma de esos valores para el año sería 0.

### Cuadro 7.10. Resultados de Regresión

Dependent Variable: DLOG(CONSUMO)				
Method: Least Squares				
Date: 11/12/07 Time: 19:36				
Sample: 1986Q1 2005Q4				
Included observations: 80				
White Heteroskedasticity-Consistent Standard Errors & Covariance				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.366203	0.279919	1.308246	0.1950
LOG(CONSUMO(-1))	-0.304615	0.066418	-4.586349	0.0000
LOG(PIB(-1))	0.285756	0.063182	4.522784	0.0000
LOG(TCR(-1))	-0.057758	0.018883	-3.058790	0.0031
DLOG(PIB)	0.620099	0.077767	7.973854	0.0000
DLOG(CONSUMO(-3))	-0.238446	0.071154	-3.351131	0.0013
@SEAS(1)	-0.064216	0.005687	-11.29238	0.0000
@SEAS(4)	-0.046154	0.008815	-5.236016	0.0000
R-squared	0.904384	Mean dependent var		0.007841
Adjusted R-squared	0.895088	S.D. dependent var		0.049862
S.E. of regression	0.016150	Akaike info criterion		-5.319131
Sum squared resid	0.018780	Schwarz criterion		-5.080929
Log likelihood	220.7653	F-statistic		97.28786
Durbin-Watson stat	1.924951	Prob(F-statistic)		0.000000

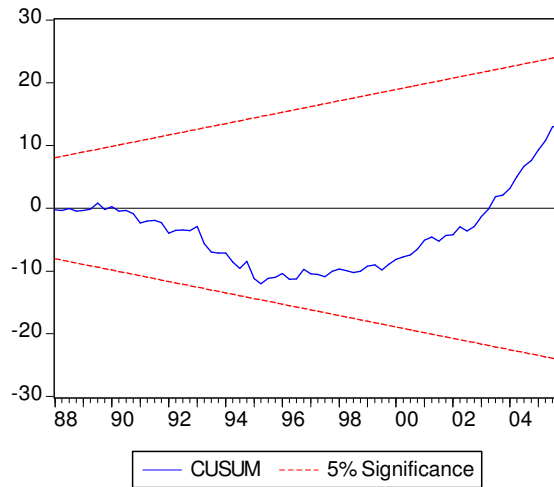
### Cuadro 7.11. Pruebas sobre Residuales

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	0.047632	Prob. F(2,70)	0.953515
Obs*R-squared	0.108725	Prob. Chi-Square(2)	0.947089

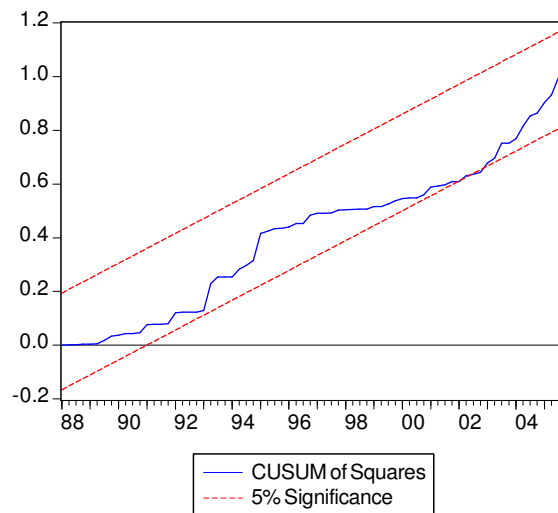
ARCH Test:			
F-statistic	0.882597	Prob. F(2,75)	0.417959
Obs*R-squared	1.793588	Prob. Chi-Square(2)	0.407875

White Heteroskedasticity Test:			
F-statistic	2.599270	Prob. F(12,67)	0.006701
Obs*R-squared	25.41265	Prob. Chi-Square(12)	0.012984

**Gráfica 7.6. Prueba CUSUM**



**Gráfica 7.7. Prueba CUSUM**



Ahora regresamos al objetivo de mostrar que las re-parametrizaciones de la ecuación de corrección de error son equivalentes. Con tal fin, se estima la ecuación (7.15) pero considerando la variable dependiente en niveles, en lugar de la variable en primera diferencia. La forma funcional específica es la siguiente:

$$y_{t-1} = \alpha_0 + \alpha_1 y_{t-1} + \beta_1 x_{t-1} + \gamma_2 \Delta y_{t-2} + \varphi_0 \Delta x_t + \varphi_1 \Delta x_{t-1} + \varepsilon_t \quad (7.16)$$

Los resultados se presentan en el Cuadro 7.12.

**Cuadro 7.12. Resultados de**  $y_{t-1} = \alpha_0 + \alpha_1 y_{t-1} + \beta_1 x_{t-1} + \gamma_2 \Delta y_{t-2} + \varphi_0 \Delta x_t + \varphi_1 \Delta x_{t-1} + \varepsilon_t$

Dependent Variable: LOG(CONSUMO)				
Method: Least Squares				
Date: 08/16/06 Time: 17:49				
Sample (adjusted): 1993Q4 2005Q4				
Included observations: 49 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.395092	0.631166	-3.794710	0.0005
LOG(CONSUMO(-1))	0.521148	0.091156	5.717123	0.0000
LOG(PIB(-1))	0.584506	0.105545	5.537966	0.0000
DLOG(CONSUMO(-2))	-0.289049	0.084529	-3.419541	0.0014
DLOG(PIB)	0.852203	0.117004	7.283538	0.0000
DLOG(PIB(-1))	-0.516007	0.094571	-5.456307	0.0000
R-squared	0.978351	Mean dependent var		20.76783
Adjusted R-squared	0.975834	S.D. dependent var		0.136563
S.E. of regresión	0.021229	Akaike info criterion		-4.752594
Sum squared resid	0.019379	Schwarz criterion		-4.520943
Log likelihood	122.4386	F-statistic		388.6552
Durbin-Watson stat	1.651245	Prob(F-statistic)		0.000000

Destacamos que la única diferencia con respecto a la estimación de la ecuación (7.15) es el coeficiente de  $y_{t-1}$ , que en el Cuadro 7.7 es -0.48 y aquí es 0.52, que es simplemente igual a  $1 - 0.48$ . En otras palabras, para obtener  $\alpha_1$  de esta estimación sólo tendríamos que calcular el negativo de  $1 - 0.52$ , que es precisamente -0.48. De tal forma que para estimar la elasticidad del consumo con respecto al ingreso en este caso, se procedería de igual forma que se hizo en la estimación anterior, se divide el coeficiente de  $x_{t-1}$  por el valor absoluto

de  $\alpha_1$ , y se obtiene 1.2. Las pruebas sobre los residuales y de estabilidad para esta estimación producen resultados idénticos a los que se presentaron anteriormente.

Comparando esta metodología con el modelo de corrección de error en dos etapas, argumentamos que puede considerarse más rico al permitir la libertad de incluir en una sola estimación las dinámicas de largo y de corto plazo, de tal forma que no es necesario instrumentar dos etapas para identificar la existencia de cointegración. Ésta, por supuesto, es una ventaja aun más clara con respecto a la metodología de Engle y Granger, que no contempla términos de corto plazo. Sin embargo, el problema de la normalización sigue presente, ya que al cambiar el orden de las variables el vector de cointegración se modifica.

Asimismo, al igual que las metodologías anteriores, el MCE derivado del proceso autorregresivo no permite identificar el número de vectores de cointegración, esto es, sabemos que las series comparten tendencia común, pero no sabemos cuántas. Aunque en este caso, dado que el sistema consiste de dos variables, sólo puede haber un vector,<sup>23</sup> en un sistema de tres o más variables resulta indispensable identificar el número de vectores de cointegración. En el capítulo siguiente se presenta la metodología de Johansen, que contempla la posibilidad de determinar el número de relaciones estables de largo plazo presentes en un sistema; y la libertad de normalizar la matriz de coeficientes de cointegración con respecto a cualquier variable.

---

<sup>23</sup> Como se mostrará en el siguiente capítulo, el número máximo de vectores de cointegración que se pueden identificar en un sistema de  $n$  variables es  $n-1$ .

## Capítulo 8: Introducción al Análisis de VAR's y la Metodología de Cointegración de Johansen.

En capítulos anteriores modelamos la interacción entre series de tiempo suponiendo que no existen efectos de retroalimentación entre las variables. En este capítulo ampliamos nuestro entendimiento de la relación entre series de tiempo permitiendo la posibilidad que exista retroalimentación a partir de choques idiosincráticos. Esta dinámica se captura por medio de un vector autorregresivo (VAR) que es fundamentalmente una generalización del análisis de procesos autorregresivos, en el cual, en lugar de considera una sola variable, se considera un vector de variables. A partir de la especificación VAR, podremos derivar la metodología planteada en Johansen (1991) para identificar la existencia de cointegración. En el capítulo siguiente describimos algunas aplicaciones del análisis VAR.

### 8. 1. Fundamentos

Un VAR es un sistema de dos o más series de tiempo que se modela considerando rezagos de las variables y la interacción dinámica que pudiera existir entre ellas. Consiste fundamentalmente de dos dimensiones, el número de variables ( $k$ ) y el número de rezagos ( $p$ ). Ilustremos un VAR de 3 variables y 1 rezago:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix} + \begin{pmatrix} \Pi_{11.1} & \Pi_{12.1} & \Pi_{13.1} \\ \Pi_{21.1} & \Pi_{22.1} & \Pi_{23.1} \\ \Pi_{31.1} & \Pi_{32.1} & \Pi_{33.1} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \epsilon_{3t} \end{pmatrix} \quad (8.1)$$

En general, un VAR de  $k$  variables y  $p$  rezagos se define de la siguiente manera:

$$y_t = \phi + \Pi_1 y_{t-1} + \Pi_2 y_{t-2} + \dots + \Pi_p y_{t-p} + \varepsilon_t \quad (8.2)$$

Donde

$$\phi' = (\phi_1, \phi_2, \dots, \phi_k)$$

$$y_t' = (y_{1t}, y_{2t}, \dots, y_{kt})$$

$$\varepsilon_t' = (\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{kt})$$

$$\Pi_p = \begin{bmatrix} \Pi_{11,p} & \cdot & \cdot & \Pi_{1k,p} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \Pi_{k1,p} & \cdot & \cdot & \Pi_{kk,p} \end{bmatrix}$$

Al igual que en el análisis de procesos autorregresivos que se llevó a cabo en el Capítulo 1, aquí es importante determinar las propiedades estocásticas del sistema, esto es, identificar la estacionariedad o estabilidad del VAR.

Para ilustrar el concepto de estacionariedad consideremos un VAR de 2 variables y 1 rezago que representamos de la siguiente forma:

$$y_t = \Pi_1 y_{t-1} + \varepsilon_t \quad (8.3)$$

O más explícitamente

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{bmatrix} \Pi_{11.1} & \Pi_{12.1} \\ \Pi_{21.1} & \Pi_{22.1} \end{bmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \quad (8.4)$$

La condición necesaria y suficiente para la estabilidad del VAR es que todos los valores característicos (eigenvalues) de  $\Pi_1$  tengan módulos menores a 1; los cuales de hecho son las raíces de la matriz  $\Pi_1$  en (8.3). Notamos que éstas se pueden obtener resolviendo la ecuación del polinomio característico

$$|\Pi_1 - \nu I| = 0 \quad (8.5)$$

o bien  $|I - \rho \Pi_1| = 0$

Que es el inverso del polinomio. Aquí hay que precisar que la condición de estabilidad que se refiere al valor de los eigenvalues considera la ecuación  $|\Pi_1 - \nu I| = 0$ , en el caso de la expresión  $|I - \rho \Pi_1| = 0$  la condición se traduce a que ninguno de los eigenvalues tenga módulo menor a 1. Presentemos un ejemplo con números para aclarar la idea.

Suponemos que la matriz  $\Pi_1$  está dada por

$$\Pi_1 = \begin{bmatrix} .313 & .250 \\ .125 & .313 \end{bmatrix}$$

Aplicando la expresión (8.5) tenemos

$$|\Pi_1 - vI| = \begin{vmatrix} .313 & .250 \\ .125 & .313 \end{vmatrix} - \begin{vmatrix} v & 0 \\ 0 & v \end{vmatrix} = \begin{vmatrix} .313v & .250 \\ .125 & .313v \end{vmatrix} = v^2 - .626v + .067 \quad (8.6)$$

Resolviendo por las raíces del polinomio se obtiene  $v_1 = .488$  y  $v_2 = .137$  las cuales son ambas menores a 1, lo cual indica que el VAR es estable, en otras palabras, todos sus eigenvalues son menores a 1. Podemos traducir este resultado de acuerdo al análisis que se realizó en el capítulo referente a raíces unitarias, en este caso, dado que ningún eigenvalue es igual a 1, entonces el VAR no presenta raíz unitaria y es estacionario (cuando los eigenvalues son mayores a 1 entonces el proceso es explosivo). Un VAR que presenta esta característica de estabilidad podría ser estimado en niveles, y los resultados serían confiables. Sin embargo, si se encontrara que no todos los eigenvalues son menores a 1 y en particular que fuera alguno igual a 1, entonces se deduciría que el VAR presenta raíz unitaria y por lo tanto es inestable o no-estacionario.

En el caso en el que se encuentre que el VAR es inestable, es posible proceder en dos sentidos; el primero consiste en estimar el VAR con series estacionarias; esto es, se tendría que diferenciar las series que son no-estacionarias de tal forma que se obtuvieran series estacionarias y estimar el VAR en primeras diferencias. Este escenario, como se indicó en el Capítulo 6, no es muy atractivo pues se perdería información importante sobre la dinámica de largo plazo.

La segunda posibilidad consiste en identificar la existencia de cointegración entre las series, tal y como se llevó a cabo en los modelos multivariados anteriormente presentados. En este caso la posibilidad de identificar relaciones de cointegración en el VAR sigue el argumento presentado en Engle y Granger (1987), que muestra que cuando una combinación lineal de variables  $I(1)$  puede representarse como una combinación lineal que es en sí misma  $I(0)$ , entonces ese conjunto de variables puede representarse por medio de un modelo de corrección de error. Este argumento, en términos generales, es conocido como el teorema de representación de Granger. Formalmente, la representación de corrección de error sigue la siguiente caracterización:

$$\Delta y_t = \alpha z_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{p-1} \Delta y_{t-(p-1)} + \Phi(L) \varepsilon_t \quad (8.7)$$

Donde  $z_{t-1} = \beta' y_{t-1}$  representa los  $r$  vectores de cointegración. Nótese aquí un detalle importante, el número de rezagos del VAR representado como corrección de error o VEC por sus siglas en inglés (Vector Error Correction Model) contiene  $p-1$  rezagos, que es un rezago menos que el número de rezagos óptimos para el VAR. Esta observación es particularmente importante en la aplicación práctica del VEC, ya que incluir  $p$  rezagos en lugar de  $p-1$  produciría resultados no confiables. A partir de esta representación VEC se deriva la prueba de cointegración de Johansen; que se describe a continuación.

## 8.2. Metodología de Johansen

La metodología de Johansen se basa fundamentalmente en el análisis de un VEC.

Tomemos el siguiente modelo

$$\Delta y_t = \Pi y_{t-1} + \varepsilon_t \quad (8.8)$$

Donde  $\Pi = \Pi_1 - I$

Destacamos que en la discusión de la sección anterior, el análisis se realizó sobre la matriz  $\Pi_1$ , que es la matriz de coeficientes del VAR; aquí la matriz relevante es  $\Pi$ . Comencemos por suponer que encontramos una raíz unitaria en el VAR de interés de tal forma que uno de los eigenvalues de  $\Pi_1$  es igual a 1. Esto implicaría que el rango de  $\Pi$  es reducido, esto es, en una matriz de dimensión  $k \times k$  existen sólo  $n$  renglones o columnas independientes, siendo  $n < k$ . Cuando el rango de la matriz  $\Pi$  es reducido, entonces sabemos que existe una combinación lineal de las variables en el sistema que es estacionaria, en otras palabras, existe al menos un vector de cointegración. Así, la identificación de la existencia de cointegración se reduce a determinar el rango de la matriz  $\Pi$ , y en particular si el rango es reducido. Dicha tarea se puede realizar empleando al menos 2 criterios: uno se refiere a identificar los eigenvalues de  $\Pi$ , ya que el número de eigenvalues distintos a 0 es en sí su rango. Alternativamente, se puede determinar que el rango de  $\Pi$  es reducido si su determinante es igual a 0. Ilustramos esta idea con un ejemplo:<sup>24</sup>

---

<sup>24</sup> Aquí se presentan sólo las generalidades de las propiedades de  $\Pi$  y  $\Pi_1$ , para una discusión detallada al respecto se puede consultar Patterson (2000) por ejemplo.

Supón que tenemos un VAR como sigue:

$$y_t = \Pi_1 y_{t-1} + \varepsilon_t$$

$$\text{Donde } \Pi_1 = \begin{bmatrix} .250 & .031 \\ .250 & .469 \end{bmatrix}$$

Calculando las raíces de  $|\Pi_1 - \nu I| = 0$  se obtiene que una de ellas es igual a 1, de tal forma que existe una raíz unitaria.<sup>25</sup> Escribamos ahora el sistema en forma de corrección de error:

$$\Delta y_t = \Pi y_{t-1} + \varepsilon_t$$

$$\text{Donde } \Pi = \begin{bmatrix} -.750 & .031 \\ .250 & -.531 \end{bmatrix} \text{ ya que } \Pi = \Pi_1 - I.$$

Calculando los eigenvalues de  $\Pi$  se obtiene que uno de ellos es igual a 0. De tal forma que el rango de  $\Pi$  es reducido y por lo tanto existe al menos una relación de cointegración, que de hecho en este caso debe ser sólo 1 vector de cointegración, ya que el sistema consiste de 2 variables.

Las opciones entonces con respecto a identificar cointegración en base al rango de  $\Pi$  se pueden resumir como sigue:

---

<sup>25</sup> Dejamos al lector la tarea de verificar este resultado.

Rango de cointegración      Interpretación

$r = n$	La matriz $\Pi$ tienen rango completo, el VAR es estacionario en niveles.
$1 \leq r \leq n - 1$	La matriz $\Pi$ tienen rango reducido, existen $r$ relaciones de cointegración.
$r = 0$	Implica que $\Pi = 0$ . No existen relaciones de cointegración y el VAR debe ser estimado en primeras diferencias.

Supongamos ahora que la matriz  $\Pi$  tiene rango reducido, la pregunta ahora es ¿cuántos vectores de cointegración existen? La respuesta a esta interrogante es clara, el número de vectores de cointegración está dado por el número de eigenvalues de la matriz  $\Pi$  diferentes a cero. De tal forma que si consideramos un sistema de  $n$  variables, el número máximo de vectores de cointegración puede ser  $n-1$  (ya que implícitamente el rango de  $\Pi$  es reducido). Así, si por ejemplo el sistema cuenta con dos variables,  $r$  ó el número de vectores de cointegración puede ser a lo más 1.

Johansen (1991) deriva estadísticos de prueba para determinar el número de vectores de cointegración. Los estadísticos que proporciona Johansen son los siguientes:

a) Estadístico de traza

$$\text{Traza}(r_0 | k) = -T \sum_{i=r_0+1}^n \ln(1 - \hat{\lambda}_i)$$

Donde  $\hat{\lambda}_i$  son los eigenvalues estimados, los cuales se ordenan como  $\lambda_1 > \lambda_2 > \dots > \lambda_n$ .  $r_0$  se encuentra en el rango  $[0, n-1]$ .

b) Estadístico del máximo eigenvalue

$$\lambda_{\max}(n-1) = -T \ln(1 - \hat{\lambda}_n)$$

En la instrumentación práctica de la prueba de Johansen, la prueba de hipótesis es secuencial, comenzando por probar la existencia de al menos un vector de cointegración hasta la posible existencia de  $n-1$  vectores.

Una vez que se determina que existe cointegración en las variables, entonces se procede a identificar los coeficientes de cointegración. Regresemos a la representación VEC de nuestro sistema:

$$\Delta y_t = \Pi y_{t-1} + \varepsilon_t$$

Partiendo de la especificación de corrección de error de la ecuación (8.7) notamos que

$\Pi = \alpha\beta'$  por lo cual el VEC se puede re-escribir como

$$\Delta y_t = \alpha\beta' y_{t-1} + \varepsilon_t \tag{8.8}$$

A partir de esta expresión podemos identificar dos componentes importantes, la matriz de coeficientes de ajuste,  $\alpha$ , y la matriz de coeficientes de equilibrio o cointegración,  $\beta$ . Aquí es importante destacar que cada una de las columnas de  $\beta$  potencialmente representa un vector de cointegración; por lo cual es posible determinar el vector de cointegración relevante tomando cualquiera de sus columnas y normalizando con respecto a cualquiera de los coeficientes.<sup>26</sup>

El lector atento habrá notado que en este último párrafo se menciona una de las ventajas de la metodología de Johansen sobre las metodologías estudiadas en el capítulo anterior, y se refiere a tener la libertad de normalizar el vector de cointegración con respecto a la variable que sea de interés, sin tener que preocuparse por la posibilidad de que los coeficientes cambien al modificar el ordenamiento de las variables. Así, concluimos la descripción de esta metodología señalando que queda a la discreción del lector emplear aquella que satisfaga más puntualmente sus requerimientos de investigación.

### **8.3. Ejercicio Práctico: Función de Consumo**

Ahora ilustramos la estimación de la metodología de Johansen; seguimos con nuestras variables de interés: el consumo y el ingreso. El primer paso consiste en determinar el número de rezagos óptimos para el VAR, al que anteriormente nos referimos como  $p$ . Esto lo hacemos estimando un VAR en niveles y refiriéndonos a los criterios de información

---

<sup>26</sup> Destacamos que en algunos paquetes econométricos, como en EViews, la hoja de resultados reporta  $\beta'$  por lo cual las filas, y no las columnas, se refieren a los vectores de cointegración.

estándar. Considerando el logaritmo del consumo y el del ingreso, en el VAR se incluyen hasta 8 rezagos y se obtienen los resultados que se presentan en el Cuadro 8.1.

**Cuadro 8.1 Prueba de Exclusión de Rezagos**

VAR Lag Order Selection Criteria						
Endogenous variables: LOG(CONSUMO) LOG(PIB)						
Exogenous variables: C						
Date: 08/15/06 Time: 18:43						
Sample: 1993Q1 2005Q4						
Included observations: 44						
Lag	LogL	LR	FPE	AIC	SC	HQ
0	124.2860	NA	1.32e-05	-5.558454	-5.477355	-5.528379
1	172.9184	90.63315	1.74e-06	-7.587201	-7.343902	-7.496974
2	201.3092	50.32918	5.75e-07	-8.695874	-8.290377	-8.545496
3	231.9574	51.54459	1.72e-07	-9.907153	-9.339456	-9.696624
4	258.5401	42.29067	6.19e-08	-10.93364	-10.20374*	-10.66296
5	259.9251	2.077578	7.04e-08	-10.81478	-9.922684	-10.48395
6	268.3018	11.80352	5.86e-08	-11.01372	-9.959425	-10.62274
7	271.9435	4.800429	6.08e-08	-10.99743	-9.780940	-10.54630
8	284.6019	15.53529*	4.21e-08*	-11.39100*	-10.01230	-10.87971*
* indicates lag order selected by the criterion						
LR: sequential modified LR test statistic (each test at 5% level)						
FPE: Final prediction error						
AIC: Akaike information criterion						
SC: Schwarz information criterion						
HQ: Hannan-Quinn information criterion						

De acuerdo a estos, el número de rezagos óptimos en niveles puede ser 4 u 8, de tal forma que para estimar el VEC, los números a considerar,  $p-1$ , serían 3 o 7, para nuestros propósitos escogemos 3 rezagos.<sup>27</sup>

Enseguida se realiza la prueba de cointegración, para la cual en el paquete econométrico EViews se puede escoger entre 5 opciones. Aquí instrumentamos la que se refiere a incluir una tendencia determinística; los resultados de la prueba se reportan en el Cuadro 8.2.<sup>28</sup> Nótese que bajo ambos estadísticos de prueba, el de la traza y el del máximo eigenvalue, se identifica la existencia de un vector de cointegración. En particular, para el caso de la prueba de la traza, la probabilidad sobre la hipótesis nula de la existencia de ninguna relación de cointegración es 0, lo cual indica el rechazo de la misma, esto es, existe al menos un vector de cointegración.<sup>29</sup> Enseguida se reporta para la hipótesis nula de “a lo más un vector de cointegración” una probabilidad de .20 que sugiere el no rechazo de la hipótesis; de tal forma que se concluye que existe un vector de cointegración.

---

<sup>27</sup> En la práctica seleccionar el número de rezagos con los que se inicia la prueba (en nuestro ejemplo 8), se realiza en base a la intuición del investigador. Es estándar, por ejemplo, comenzar con 12 rezagos cuando se trata de datos con frecuencia mensual, y con 4 rezagos cuando se trata de observaciones trimestrales.

<sup>28</sup> Recomendamos referirse a la explicación técnica en EViews para identificar claramente las diferencias que existen entre las opciones que se ofrecen para realizar la prueba de cointegración. En el fondo, cabe recordar que el ejercicio de cointegración, así como cualquier ejercicio que modele series de tiempo, tiene como fin último identificar la dinámica que generó la o las series con las que se trabaja, por lo cual la especificación que se escoja debe ser una aproximación razonable al comportamiento que exhiben las variables.

<sup>29</sup> Aquí la probabilidad se refiere al comúnmente conocido p-value que es estándar en pruebas de hipótesis.

### Cuadro 8.2. Resultados de Prueba de Johansen

Date: 08/15/06 Time: 18:47				
Sample (adjusted): 1994Q1 2005Q4				
Included observations: 48 after adjustments				
Trend assumption: Linear deterministic trend (restricted)				
Series: LOG(CONSUMO) LOG(PIB)				
Lags interval (in first differences): 1 to 3				
Unrestricted Cointegration Rank Test (Trace)				
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.586137	51.02668	25.87211	0.0000
At most 1	0.165428	8.680120	12.51798	0.2010
Trace test indicates 1 cointegrating eqn(s) at the 0.05 level				
* denotes rejection of the hypothesis at the 0.05 level				
**MacKinnon-Haug-Michelis (1999) p-values				
Unrestricted Cointegration Rank Test (Maximum Eigenvalue)				
Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.586137	42.34656	19.38704	0.0000
At most 1	0.165428	8.680120	12.51798	0.2010
Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.05 level				
* denotes rejection of the hypothesis at the 0.05 level				
**MacKinnon-Haug-Michelis (1999) p-values				
Unrestricted Cointegrating Coefficients (normalized by b'S11*b=I):				
LOG(CONSUMO)	LOG(PIB)	@TREND(80Q2)		
1.702062	-31.98589	0.174833		
36.28156	-25.80340	-0.136977		

Una vez identificada la existencia de cointegración, se procede a determinar el vector de coeficientes, lo cual se lleva a cabo a partir de la matriz  $\beta'$  (como se indicó anteriormente, el programa reporta  $\beta'$  y no  $\beta$ ). En este caso tenemos que  $\beta'$  es igual a

**Cuadro 8.3. Matriz  $\beta'$**

LOG(CONSUMO)	LOG(PIB)
1.702062	-31.98589
36.28156	-25.80340

Ya que la normalización se puede realizar con respecto a cualquier coeficiente para cada fila en la matriz, aquí escogemos los valores 36.28156 y -25.80340. Normalizando con respecto al consumo, esto es, dividiendo los coeficientes por 36.28156, obtenemos el vector de cointegración (1, -.71), que indica una relación positiva entre el consumo y el ingreso y con una elasticidad de largo plazo de .71 (como se indicó en el Capítulo 8, la relación cualitativa que existe entre las variables es la opuesta al signo del coeficiente normalizado). Podríamos haber construido el vector considerando la primera fila, sin embargo la elasticidad de largo plazo no sería congruente con la teoría. A continuación estimamos un sistema de tres variables, con el propósito de ilustrar un caso en el que puede encontrarse más de un vector de cointegración.

Nuestro sistema ahora incluye al agregado monetario M1 real como variable adicional en la función de consumo. Comenzamos por determinar el número de rezagos óptimos a partir del VAR en niveles. Los resultados del ejercicio de exclusión de rezagos se presentan en el Cuadro 8.4

### Cuadro 8.4. Prueba de Exclusión de Rezagos

VAR Lag Order Selection Criteria						
Endogenous variables: LOG(CONSUMO) LOG(PIB) LOG(MIREAL)						
Exogenous variables: C						
Date: 08/27/06 Time: 21:33						
Sample: 1993Q1 2005Q4						
Included observations: 44						
Lag	LogL	LR	FPE	AIC	SC	HQ
0	179.3545	NA	6.63e-08	-8.016112	-7.894462	-7.970998
1	254.5789	136.7717	3.27e-09	-11.02631	-10.53972	-10.84586
2	301.9868	79.73149	5.74e-10	-12.77213	-11.92058	-12.45633
3	325.3834	36.15838	3.03e-10	-13.42652	-12.21003	-12.97538
4	356.8365	44.32024	1.13e-10	-14.44711	-12.86567*	-13.86064
5	363.0471	7.904389	1.35e-10	-14.32032	-12.37393	-13.59851
6	380.7924	20.16508	9.77e-11	-14.71783	-12.40650	-13.86068
7	384.9762	4.183801	1.36e-10	-14.49892	-11.82263	-13.50642
8	414.6663	25.64148*	6.27e-11*	-15.43938*	-12.39814	-14.31154*

\* indicates lag order selected by the criterion

LR: sequential modified LR test statistic (each test at 5% level)

FPE: Final prediction error

AIC: Akaike information criterion

SC: Schwarz information criterion

HQ: Hannan-Quinn information criterion

De acuerdo a estos el número de rezagos óptimos para el VEC es de 3 o 7. Consideramos 3 para nuestro ejercicio y realizamos la prueba de cointegración. Reportamos los resultados en el Cuadro 8.5.

### Cuadro 8.5. Resultados de Prueba de Johansen

Date: 08/27/06 Time: 21:37				
Sample (adjusted): 1994Q1 2005Q4				
Included observations: 48 after adjustments				
Trend assumption: Linear deterministic trend (restricted)				
Series: LOG(CONSUMO) LOG(PIB) LOG(MIREAL)				
Lags interval (in first differences): 1 to 3				
Unrestricted Cointegration Rank Test (Trace)				
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.505035	65.05000	42.91525	0.0001
At most 1 *	0.336301	31.29310	25.87211	0.0096
At most 2	0.214954	11.61665	12.51798	0.0704
Trace test indicates 2 cointegrating eqn(s) at the 0.05 level				
* denotes rejection of the hypothesis at the 0.05 level				
**MacKinnon-Haug-Michelis (1999) p-values				
Unrestricted Cointegration Rank Test (Maximum Eigenvalue)				
Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.505035	33.75690	25.82321	0.0037
At most 1 *	0.336301	19.67645	19.38704	0.0454
At most 2	0.214954	11.61665	12.51798	0.0704
Max-eigenvalue test indicates 2 cointegrating eqn(s) at the 0.05 level				
* denotes rejection of the hypothesis at the 0.05 level				
**MacKinnon-Haug-Michelis (1999) p-values				
Unrestricted Cointegrating Coefficients (normalized by b'S11*b=I):				
LOG(CONSUMO)	LOG(PIB)	LOG(MIREAL)	@TREND(80Q2)	
-1.173952	38.60550	-0.291573	-0.221358	
100.1587	-79.59472	-22.14119	0.081246	
59.33312	-41.59916	-5.025612	-0.146434	

Tanto el estadístico de traza como el estadístico de máximo eigenvalue indican la existencia de dos vectores de cointegración. Para el caso del estadístico de traza, notamos que la hipótesis nula de “ningún vector de cointegración” se rechaza a niveles convencionales. Asimismo, rechazamos la hipótesis de “a lo más un vector de cointegración”, mientras que la hipótesis de “a lo más dos vectores de cointegración” no se rechaza a un nivel de confianza de 95%. De acuerdo a estos resultados concluimos entonces que existen dos vectores de cointegración; lo cual tiene sentido ya que el sistema consiste de tres variables.<sup>30</sup> Los posibles vectores de cointegración se aprecian en el Cuadro 8.6

**Cuadro 8.6. Matriz  $\beta'$**

LOG(CONSUMO)	LOG(PIB)	LOG(MIREAL)
-1.173952	38.60550	-0.291573
100.1587	-79.59472	-22.14119
59.33312	-41.59916	-5.025612

Normalizando con respecto a los coeficientes de consumo, podríamos tomar la segunda o la tercera fila para determinar el vector de cointegración, ya que éstas producen resultados en línea con la teoría. En particular, para la segunda fila obtenemos (1, -.79, -.22) y con respecto a la tercera fila (1, -.70, -.08); que si tuviéramos que escoger entre estos dos nos

<sup>30</sup> Aquellos lectores que hayan realizado este análisis en EViews, reconocerán que en algunas ocasiones la prueba señala que el número de vectores de cointegración es igual al número de variables en el sistema, esto es  $r=n$ , lo cual no tienen sentido pues la existencia de cointegración se acota a que la matriz  $\Pi$  exhiba rango reducido. Una posible causa de este resultado contradictorio se debe a que se incluyen variables exógenas en el sistema, por lo cual los estadísticos de prueba pueden ser no confiables. En todo caso, cuando la prueba indica que  $r=n$  se debe re-especificar el sistema pues el resultado carece de sentido.

inclinariamos por el primero, ya que una elasticidad del consumo con respecto al PIB de .79 es más cercana a lo que se encuentra en la literatura que la elasticidad de .70.<sup>31</sup>

Notamos finalmente que en el cuadro de salida de EViews se reporta en este caso, primero un vector de cointegración normalizado que generalmente se calcula en base a la primera fila, y enseguida un segundo vector de cointegración normalizado, que de hecho asigna el valor de 0 a alguno de los coeficientes. La asignación de 0 es necesaria para identificar los posibles vectores en el sistema, pero de hecho podría no hacerse esta simplificación y obtener directamente el vector deseado de la matriz  $\beta$ , como lo hicimos en nuestros ejercicios.<sup>32</sup>

Formalmente, la identificación del vector de cointegración concluye el análisis de Johansen, notando que aquí sólo incluimos la discusión de los resultados que reporta EViews hasta la matriz  $\beta'$ . El lector notará al realizar su prueba que también se reporta la matriz que corresponde a  $\alpha$ , que contiene los coeficientes de ajuste. En general, paquetes econométricos como GAUSS y RATS (con su complemento CATS) reportan similares resultados, aunque estos paquetes permiten realizar ejercicios relativamente más sofisticados a los disponibles en EViews.

El lector interesado se preguntará ¿por qué el vector de cointegración que se obtuvo al instrumentar las metodologías de Engle y Granger y de corrección de error no es similar al que se obtuvo en este ejercicio? La respuesta es simple, por que las especificaciones son

---

<sup>31</sup> Ver Castillo (2003) o Villagomez (1997), por ejemplo.

<sup>32</sup> Para una discusión detallada sobre la restricción de 0 en alguna de las variables del sistema en caso de que se encuentre más de un vector de cointegración se puede consultar Patterson (2000).

diferentes. En este caso la ecuación incluye una tendencia y el periodo de estimación está ajustado para incluir sólo 48 observaciones. Podríamos, sin embargo, tratar de aproximar los resultados estimando una ecuación en la primera etapa de los modelos de Engle y Granger, y corrección de error, similar a la que estimamos aquí. Con tal fin, consideramos ahora una ecuación que incluye una tendencia para periodos muestrales iguales. Los resultados de su estimación presentan en el Cuadro 8.7. El vector normalizado sería (1, -.89), que se asemeja un poco más a lo que se obtuvo con la metodología de Johansen pero no es idéntico.

**Cuadro 8.7. Resultados de  $y_t = \beta_0 + \beta_1 x_t + \beta_2 T + \varepsilon_t$**

Dependent Variable: LOG(CONSUMO)				
Method: Least Squares				
Date: 08/15/06 Time: 19:01				
Sample: 1994Q1 2005Q4				
Included observations: 48				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.592634	2.538987	0.627272	0.5337
LOG(PIB)	0.899846	0.123831	7.266720	0.0000
@TREND	0.002091	0.001033	2.024051	0.0489
R-squared	0.943923	Mean dependent var		20.77027
Adjusted R-squared	0.941431	S.D. dependent var		0.136929
S.E. of regresión	0.033138	Akaike info criterion		-3.915788
Sum squared resid	0.049417	Schwarz criterion		-3.798838
Log likelihood	96.97891	F-statistic		378.7338
Durbin-Watson stat	0.972309	Prob(F-statistic)		0.000000

Esta diferencia en los resultados que se obtienen a partir de la estimación de diferentes metodologías es común en la práctica, y es aquí cuando el investigador debe hacer uso de su criterio para determinar los resultados que van más de acuerdo con la teoría. Por ejemplo, en este caso es evidente que la ecuación anterior exhibe un problema de

autocorrelación, de tal forma que pudiera ser más eficiente estimar una especificación que atiende este problema. En el Cuadro 8.8 se presentan los resultados para una ecuación que incluye al consumo rezagado dos periodos y una variable estacional para el primer trimestre; pero que ya no considera la tendencia. Ahora la elasticidad de largo plazo del consumo con respecto al PIB es similar a la que se obtuvieron con la metodología de Johansen. Asimismo, ya no existe un problema tan grave de autocorrelación. Dados estos resultados, subrayamos que el ejercicio de cointegración requiere no sólo de conocer la mecánica para estimar modelos, sino que es fundamental que se encuentren resultados consistentes con la teoría que sustenta el ejercicio econométrico.

**Cuadro 8.7. Resultados de  $y_t = \beta_0 + \beta_1 x_t + \beta_2 T + \varepsilon_t$**

Dependent Variable: LOG(CONSUMO)				
Method: Least Squares				
Date: 11/12/07 Time: 20:23				
Sample: 1994Q1 2005Q4				
Included observations: 48				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.829693	0.458249	-3.992789	0.0002
LOG(PIB)	0.771578	0.085732	8.999912	0.0000
LOG(CONSUMO(-2))	0.341666	0.077082	4.432508	0.0001
@SEAS(1)	-0.040609	0.008507	-4.773425	0.0000
R-squared	0.967810	Mean dependent var		13.86251
Adjusted R-squared	0.965615	S.D. dependent var		0.136929
S.E. of regression	0.025391	Akaike info criterion		-4.429180
Sum squared resid	0.028367	Schwarz criterion		-4.273247
Log likelihood	110.3003	F-statistic		440.9560
Durbin-Watson stat	1.684704	Prob(F-statistic)		0.000000

Con este ejemplo concluimos nuestra presentación del análisis de cointegración, sugerimos al lector interesado revisar las referencias citadas a través de la discusión para profundizar en el aspecto teórico.

## **Capítulo 9: Temas Selectos en el Análisis de VAR's**

¿Se confina el análisis de un VAR a realizar pruebas de cointegración? La respuesta por supuesto es no. Existe una amplia gama de aplicaciones de un VAR dentro de las cuales aquí se tratan tres ampliamente utilizadas en la práctica. Comenzamos con la identificación de relaciones de causalidad según Granger, seguimos con la estimación de la función impulso-respuesta, y concluimos con la descomposición de la varianza. Asimismo, se incluye un análisis multivariado de la descomposición tendencia-ciclo de series de tiempo, el cual se basa en la dinámica estocástica de un VAR. Debemos destacar, por supuesto, que en la práctica los ejercicios que aquí se presentan generalmente se realizan para identificar el comportamiento de corto plazo de las variables en el sistema, ya que el análisis relevante para el largo plazo se refiere al análisis de cointegración.

### **9.1. Causalidad Según Granger**

En primera instancia podría parecer que la prueba se refiere a identificar si existe causalidad entre dos o más series de tiempo. Sin embargo, el propósito de este ejercicio es, en efecto, probar si valores rezagados de una variable influyen significativamente sobre el comportamiento de otra variable; de tal forma que no se determina causalidad en el sentido en el que comúnmente se entiende la palabra. Cabe señalar que este ejercicio generalmente se lleva a cabo cuando se conoce que no existe cointegración entre series no-estacionarias, y de hecho sería preferible estimar el VAR en primeras diferencias, llevando a cabo la prueba de causalidad de Granger también con series estacionarias. En caso de identificar

cointegración, entonces no se estimaría un VAR sino un VEC, como se indicó en el capítulo anterior.

Formalmente la prueba de causalidad se caracteriza como sigue:

Consideremos un VAR de dos variables y un rezago, esto es  $k = 2$  y  $p = 1$ .

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{bmatrix} \Pi_{11.1} & \Pi_{12.1} \\ \Pi_{21.1} & \Pi_{22.1} \end{bmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \quad (9.1)$$

En este caso, notamos que el valor rezagado de  $y_{2t}$  afecta el comportamiento de  $y_{1t}$ ; similarmente, el valor rezagado de  $y_{1t}$  influye sobre  $y_{2t}$ . De tal forma que si se estimaran los parámetros de la matriz  $\Pi_1$ ,  $\Pi_{12.1}$  y  $\Pi_{21.1}$ , y ambos resultaran estadísticamente significativos, se diría que  $y_{2t}$  Granger causa  $y_{1t}$  y  $y_{1t}$  Granger causa  $y_{2t}$ . Pero si por ejemplo  $\Pi_{12.1}$  fuera estadísticamente igual a 0, entonces tendríamos

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{bmatrix} \Pi_{11.1} & 0 \\ \Pi_{21.1} & \Pi_{22.1} \end{bmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \quad (9.2)$$

De tal forma que el valor rezagado de  $y_{2t}$  no influiría sobre  $y_{1t}$  y así se diría que  $y_{2t}$  no Granger causa  $y_{1t}$ .

En la práctica la instrumentación de la prueba contempla los siguientes pasos.

1. Especificar el VAR
2. Determinar el número de rezagos de acuerdo a algún criterio de información.
3. Estimar el VAR considerando el número de rezagos óptimos e identificar si valores rezagados de una variable influyen significativamente sobre otra u otras.

Cabe mencionar que varios paquetes econométricos incluyen pruebas de causalidad de Granger por lo cual el ejercicio se puede realizar directamente. Para ilustrar el mismo, a continuación presentamos un ejemplo considerando dos variables estacionarias: la tasa de desempleo de Estados Unidos y su tasa de inflación:<sup>33</sup>

Primero especificamos el VAR con la tasa de desempleo primero y enseguida la tasa de inflación. Segundo, realizamos la prueba de exclusión de rezagos para determinar el número óptimo de los mismos. Los resultados se presentan en el Cuadro 9.1.

---

<sup>33</sup> Nótese que a través del análisis de VAR's no se ha hecho hincapié en cuanto al ordenamiento de las variables en el sistema, esto es, acerca de cuál variable se incluye primero y cuál, o cuáles, enseguida. Generalmente se incluye primero la variable que se considera dependiente y enseguida las independientes. Aunque reconocemos que es posible por medio de pruebas de exogeneidad determinar qué variables deben considerarse endógenas y cuáles exógenas. Para propósitos prácticos suponemos que las variables incluidas en el VAR son endógenas, de tal forma que cada una requiere de una ecuación al modelar el VAR. Si hubiera alguna variable exógena, ésta no requeriría de una ecuación propia. El lector interesado puede referirse a la discusión metodológica en Juselius y Hansen (1995).

### Cuadro 9.1. Prueba de Exclusión de Rezagos

Lag	LogL	LR	FPE	AIC	SC	HQ
0	360.4309	NA	6.39e-05	-3.982566	-3.947089	-3.968181
1	914.3297	1089.334	1.42e-07	-10.09255	-9.986121*	-10.04940
2	923.2178	17.28226	1.34e-07	-10.14686	-9.969478	-10.07494
3	934.5615	21.80516	1.24e-07	-10.22846	-9.980120	-10.12777
4	941.9779	14.09127	1.19e-07*	-10.26642*	-9.947126	-10.13696*
5	942.6845	1.326833	1.24e-07	-10.22983	-9.839578	-10.07160
6	946.8001	7.636604	1.24e-07	-10.23111	-9.769907	-10.04411
7	951.2517	8.161267	1.23e-07	-10.23613	-9.703970	-10.02036
8	952.9598	3.093684	1.26e-07	-10.21066	-9.607551	-9.966128
9	953.6305	1.199755	1.31e-07	-10.17367	-9.499604	-9.900367
10	954.7573	1.990711	1.35e-07	-10.14175	-9.396725	-9.839673
11	967.5196	22.26300*	1.23e-07	-10.23911	-9.423128	-9.908262
12	970.9636	5.931361	1.24e-07	-10.23293	-9.345996	-9.873316

\* indicates lag order selected by the criterion  
 LR: sequential modified LR test statistic (each test at 5% level)  
 FPE: Final prediction error  
 AIC: Akaike information criterion  
 SC: Schwarz information criterion  
 HQ: Hannan-Quinn information criterion

Como se indicó en su oportunidad, cuando se trata con datos mensuales, es común que se incluyan 12 rezagos para iniciar el análisis; así lo hacemos aquí. A partir de los resultados del Cuadro, notamos que el número óptimo de rezagos puede ser 1, 4 u 11. Para nuestros propósitos seleccionamos 4.<sup>34</sup> Enseguida aplicamos la prueba de causalidad de Granger y obtenemos los resultados que se presentan en el Cuadro 9.2.

<sup>34</sup> A diferencia del ejercicio de cointegración donde el número de rezagos óptimo es  $p-1$ , aquí el número es  $p$ , y que se trabaja con series estacionarias.

### Cuadro 9.2. Prueba de Causalidad de Granger

Pairwise Granger Causality Tests			
Date: 08/23/06 Time: 14:51			
Sample: 1990M01 2005M12			
Lags: 4			
Null Hypothesis:	Obs	F-Statistic	Probability
INFLACION does not Granger Cause DESEMPLEO	188	2.07525	0.08591
DESEMPLEO does not Granger Cause INFLACIÓN		1.10981	0.35342

De acuerdo a los valores de probabilidad, no se rechaza la hipótesis nula de que la inflación no causa al desempleo al nivel de confianza del 10%, aunque sí se rechaza al 5%. En contraste, la hipótesis nula de que el desempleo no causa a la inflación no se rechaza a ningún nivel de significancia convencional. Así, en base a estos resultados, se puede indicar que choques rezagados del desempleo no tienen efecto sobre la dinámica de la inflación.

En la práctica es común encontrarse con series que no son estacionarias, sobre todo en el ámbito financiero. A continuación se consideran dos series frecuentemente analizadas en este sector: el Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores, (IPC) y el Índice Industrial Dow Jones (Dow). Debido a que ambas se encuentran integradas de orden 1, como se puede verificar en los Cuadros 9.3 y 9.4, se estima el VAR en primeras diferencias, ordenando al IPC en la primera posición y al Dow como segundo. Como primer paso, se determina el número de rezagos óptimos en el sistema; que de acuerdo a los resultados del Cuadro 9.5 sería igual a 0, 2, o 12; para nuestro ejercicio seleccionamos 2.

### Cuadro 9.3. Pruebas KPSS de Raíz Unitaria IPC

Null Hypothesis: LOG(IPC) is stationary		
Exogenous: Constant, Linear Trend		
Lag length: 0 (Spectral OLS AR based on AIC, MAXLAG=12)		
		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		4.315592
Asymptotic critical values*:	1% level	0.216000
	5% level	0.146000
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)		

Null Hypothesis: D(LOG(IPC)) is stationary		
Exogenous: Constant, Linear Trend		
Lag length: 0 (Spectral OLS AR based on AIC, MAXLAG=12)		
		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.052280
Asymptotic critical values*:	1% level	0.216000
	5% level	0.146000
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)		

### Cuadro 9.4. Pruebas KPSS de Raíz Unitaria Dow

Null Hypothesis: LOG(DOW) is stationary		
Exogenous: Constant, Linear Trend		
Lag length: 0 (Spectral OLS AR based on AIC, MAXLAG=12)		
		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		40.85470
Asymptotic critical values*:	1% level	0.216000
	5% level	0.146000
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)		

Null Hypothesis: D(LOG(DOW)) is stationary		
Exogenous: Constant, Linear Trend		
Lag length: 0 (Spectral OLS AR based on AIC, MAXLAG=12)		
		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.089046
Asymptotic critical values*:	1% level	0.216000
	5% level	0.146000
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)		

### Cuadro 9.5. Prueba de Exclusión de Rezagos

VAR Lag Order Selection Criteria						
Endogenous variables: DLOG(IPC) DLOG(DOW)						
Exogenous variables: C						
Date: 11/14/07 Time: 11:34						
Sample: 1995M01 2005M12						
Included observations: 132						
Lag	LogL	LR	FPE	AIC	SC	HQ
0	368.2235	NA	1.33e-05	-5.548841	-5.505162*	-5.531092
1	369.2823	2.069378	1.40e-05	-5.504277	-5.373240	-5.451030
2	385.9072	31.99041	1.15e-05*	-5.695564*	-5.477170	-5.606818*
3	388.6754	5.242846	1.17e-05	-5.676900	-5.371149	-5.552657
4	389.1998	0.977211	1.24e-05	-5.624239	-5.231130	-5.464498
5	390.9601	3.227166	1.28e-05	-5.590304	-5.109837	-5.395064
6	394.1698	5.787341	1.30e-05	-5.578331	-5.010506	-5.347593
7	395.1523	1.741614	1.36e-05	-5.532611	-4.877428	-5.266375
8	397.5340	4.149982	1.39e-05	-5.508091	-4.765551	-5.206357
9	400.0414	4.292993	1.43e-05	-5.485476	-4.655579	-5.148244
10	402.1377	3.525597	1.47e-05	-5.456632	-4.539377	-5.083902
11	403.9237	2.949469	1.52e-05	-5.423086	-4.418473	-5.014857
12	412.3536	13.66672*	1.43e-05	-5.490206	-4.398235	-5.046479

\* indicates lag order selected by the criterion

LR: sequential modified LR test statistic (each test at 5% level)

FPE: Final prediction error

AIC: Akaike information criterion

SC: Schwarz information criterion

HQ: Hannan-Quinn information criterion

En base a lo anterior se lleva a cabo la prueba de causalidad. De acuerdo a los estadísticos que se muestran en el Cuadro 9.6 no se rechaza la hipótesis que el IPC no causa al Dow; sin embargo se rechaza la causalidad en la dirección opuesta, esto es, los choques rezagados del Dow sí tienen un efecto significativo sobre el comportamiento de la tasa de crecimiento mensual del IPC.

### Cuadro 9.6. Prueba de Causalidad de Granger

Pairwise Granger Causality Tests			
Date: 11/14/07 Time: 11:12			
Sample: 1995M01 2005M12			
Lags: 2			
Null Hypothesis:	Obs	F-Statistic	Probability
DLOG(DOW) does not Granger Cause DLOG(IPC)	132	17.8178	1.5E-07
DLOG(IPC) does not Granger Cause DLOG(DOW)		0.90959	0.40530

Este segundo ejercicio de causalidad es particularmente intuitivo, ya que se esperaría que el comportamiento de la Bolsa Mexicana de Valores respondiera a los acontecimientos en los mercados accionarios de los Estados Unidos. Una ilustración más puntal sobre la respuesta precisa que tienen las variables a choques dentro del sistema se representa por medio de la función impulso-respuesta, que se discute a continuación.

### 9. 2. Función Impulso-Respuesta

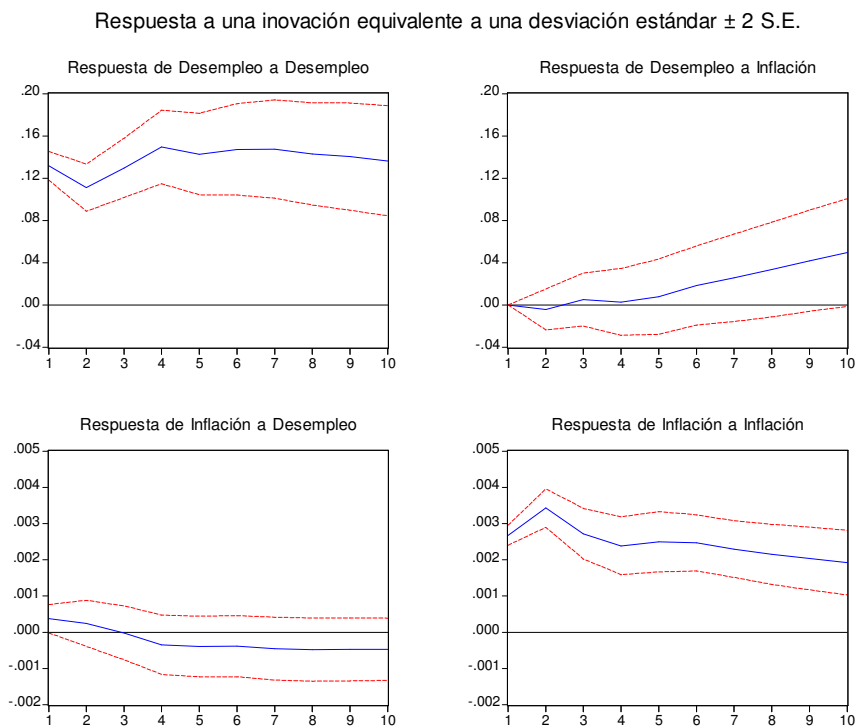
Una vez estimado el VAR, se puede continuar con la estimación de la función impulso-respuesta. La misma nos permite evaluar la respuesta de las variables a un choque o perturbación sobre un término de error. Consideremos el siguiente VAR:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{bmatrix} \Pi_{11.1} & \Pi_{12.1} \\ \Pi_{21.1} & \Pi_{22.1} \end{bmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \quad (9.3)$$

Supongamos que se desea conocer el efecto que tendría un cambio en  $\varepsilon_{1t}$  sobre las variables en el sistema. Evidentemente, un choque en  $\varepsilon_{1t}$  afectaría a  $y_{1t}$ , pero no a  $y_{2t}$ . Sin

embargo, en el periodo  $t+1$  el choque afectaría a  $y_{2t+1}$  a través del efecto en  $y_{1t}$ ; en el periodo  $t+2$  el efecto sería sobre  $y_{2t+2}$  y así sucesivamente. La función impulso-respuesta capta precisamente esta serie de respuestas o influencias sobre las variables del VAR. Al igual que en el caso de la prueba de causalidad según Granger, el ejercicio se puede realizar calculando las respuestas periodo por periodo, o simplemente instrumentando el ejercicio en algún programa econométrico. Aquí se continúa con el ejemplo del desempleo y la tasa de interés y se estima la función impulso-respuesta disponible en el paquete EViews. La Figura 9.1 presenta la misma.

**Figura 9.1. Función Impulso-Respuesta**

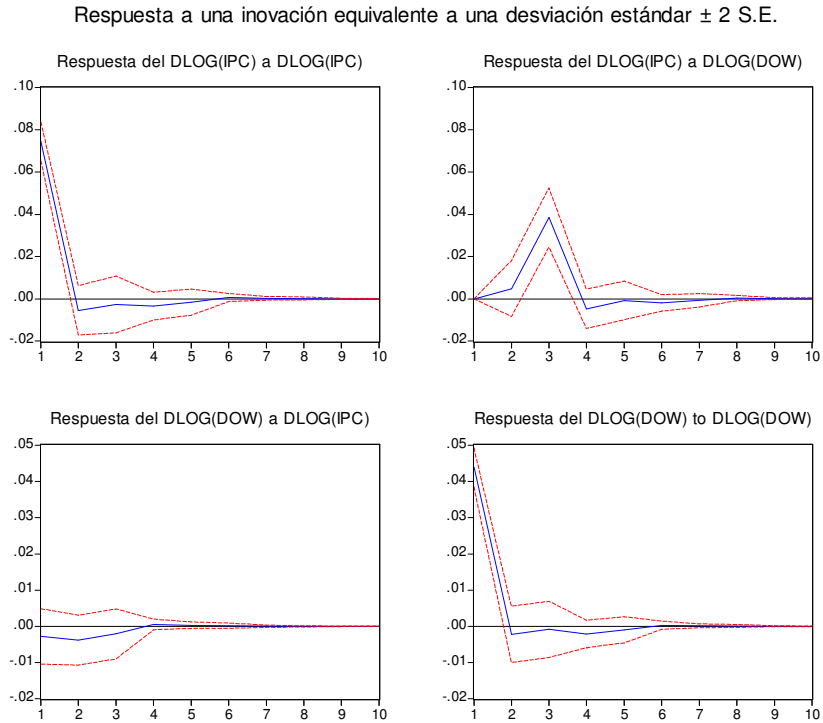


En principio se supone una innovación o choque equivalente a una desviación estándar sobre la inflación. Los cuadros de interés son particularmente el cuadro superior derecho y el inferior izquierdo, ya que ilustran la respuesta de una variable con respecto a la otra. En

el primero se nota que el desempleo muestra una respuesta a la inflación rezagada, para los primeros 4 periodos la respuesta es esencialmente 0. A partir del periodo 5 aparece cierta respuesta pero no es muy importante. En la gráfica inferior izquierda se aprecia una pequeña respuesta de la inflación con respecto al desempleo en el periodo 1, que se torna negativa a partir del periodo 5. Así, a partir del examen visual de la gráfica, nos podemos formar una idea de la respuesta de las variables en el sistema a choques o innovaciones en las mismas. En este caso, se aprecia una respuesta *vis-a-vis* poco importante.

En la Figura 9.2 se muestra la función para el caso del IPC y el Dow. Una vez más, las gráficas de interés se encuentran en la parte superior derecha y la inferior izquierda. Notamos que la primera diferencia del IPC responde significativamente al Dow, especialmente en el tercer periodo; esta respuesta desaparece para el cuarto periodo. En el caso de la respuesta del Dow al IPC, se encuentra que es tenue y para el tercer periodo se aproxima a 0. De esta forma se pueden sentar las bases para un estudio más detallado sobre la relación estocástica que guardan estas variables, sobre todo si se reconoce que ambas son no-estacionarias en niveles; por lo cual una vez determinada la relación que existe en sus primeras diferencias, se podría continuar con un análisis de cointegración como se describió en el capítulo anterior. En este sentido, destacamos que si las series se encuentran cointegradas, entonces se procedería a estimar un VEC y a partir de ello se identificaría la función impulso-respuesta. La interpretación de las gráficas es similar a la que se describió anteriormente. A pesar de que esta alternativa es atractiva, en la práctica la implementación de la metodología de cointegración prevalece sobre ejercicios impulso-respuesta que consideran un VEC.

## Figura 9.2. Función Impulso-Respuesta



Finalmente, indicamos que es posible en el paquete econométrico EViews modelar un vector autoregresivo estructural (SVAR), el cual impone restricciones sobre el efecto que el choques (o choques) del término de error tiene sobre las variables del sistema. Una aplicación de este tipo de especificaciones se refiere a la teoría de la neutralidad monetaria, donde se supone que variables nominales no presenta efecto alguno sobre variables reales. Aquí no profundizamos en estos temas pero referimos al lector interesado a consultar Johnston y DiNardo (1997) o Patterson (2000) para una discusión amplia al respecto.<sup>35</sup>

---

<sup>35</sup> La estimación de un VAR estructural es relativamente amigable en EViews. Lo que es verdaderamente importante no es, en todo caso, ejecutar las instrucciones del programa, sino formular el sistema apropiadamente de acuerdo a la teoría para sustentar la estimación.

### 9. 3. Descomposición de la varianza

Este ejercicio se puede llevar a cabo como una aplicación adicional de un VAR. El propósito del mismo es identificar la porción de la varianza de cada serie que se puede atribuir a innovaciones dentro del sistema. Consideremos el VAR que incluye el desempleo y la inflación; la descomposición de la varianza que arroja el paquete Eviews se presenta Cuadro 9.7.

**Cuadro 9.7. Descomposición de la Varianza**

Variance Decomposition of DESEMPLEO:			
Period	S.E.	DESEMPLEO	INFLACION
1	0.131754	100.0000	0.000000
2	0.172318	99.93809	0.061914
3	0.215799	99.90358	0.096416
4	0.262531	99.92309	0.076908
5	0.298930	99.87249	0.127515
6	0.333783	99.59362	0.406383
7	0.365823	99.16950	0.830499
8	0.394197	98.56263	1.437368
9	0.420559	97.74655	2.253445
10	0.444897	96.74661	3.253391
Variance Decomposition of INFLACION:			
Period	S.E.	DESEMPLEO	INFLACION
1	0.002704	1.905510	98.09449
2	0.004376	1.040907	98.95909
3	0.005152	0.752403	99.24760
4	0.005689	0.991990	99.00801
5	0.006226	1.230189	98.76981
6	0.006709	1.390599	98.60940
7	0.007105	1.644631	98.35537
8	0.007438	1.913993	98.08601
9	0.007726	2.151937	97.84806
10	0.007976	2.368062	97.63194
Cholesky Ordering: DESEMPLEO INFLACION			

La primera columna se refiere a los periodos durante los cuales se supone una respuesta de las variables a las perturbaciones. En la segunda se reporta el error estándar del pronóstico de las variables una vez que se consideran los choques. La tercera y cuarta columnas en este caso reportan la varianza que se atribuye al choque que recibe cada una de las variables. Así, por ejemplo, en la porción superior que considera la descomposición de la varianza del desempleo, notamos que la varianza en el pronóstico de esta variable para el primer periodo se atribuye en su totalidad al choque en la variable de desempleo, esto es, un choque en la variable de inflación no explica en lo absoluto la varianza en el desempleo. Interesantemente, en la parte inferior del Cuadro, que reporta la descomposición de la varianza para la inflación, el efecto de un choque en el empleo sobre la inflación en el primer periodo no es 0, aunque el mismo es pequeño.

En el Cuadro 9.8 se muestra la descomposición de la varianza para el ejemplo que considera al IPC y al Dow. En la parte superior del Cuadro se observa que hasta un 21 por ciento de la varianza del IPC es explicada por la variación en el Dow, lo cual es consistente con las pruebas de causalidad de Granger y funciones impulso-respuesta que se presentaron anteriormente. En la parte inferior del Cuadro los resultados apuntan en la misma dirección, ya que la participación del IPC en la varianza del Dow es prácticamente insignificante.

**Cuadro 9.8. Descomposición de la Varianza**

Variance Decomposition of DLOG(IPC):			
Period	S.E.	DLOG(IPC)	DLOG(DOW)
1	0.074528	100.0000	0.000000
2	0.074884	99.58555	0.414445
3	0.084224	78.82479	21.17521
4	0.084428	78.60968	21.39032
5	0.084446	78.61071	21.38929
6	0.084471	78.56996	21.43004
7	0.084474	78.56464	21.43536
8	0.084475	78.56351	21.43649
9	0.084475	78.56338	21.43662
10	0.084475	78.56332	21.43668
Variance Decomposition of DLOG(DOW):			
Period	S.E.	DLOG(IPC)	DLOG(DOW)
1	0.043970	0.392425	99.60757
2	0.044193	1.141095	98.85891
3	0.044251	1.364438	98.63556
4	0.044306	1.374087	98.62591
5	0.044318	1.377536	98.62246
6	0.044319	1.379154	98.62085
7	0.044319	1.379146	98.62085
8	0.044319	1.379210	98.62079
9	0.044319	1.379220	98.62078
10	0.044319	1.379221	98.62078
Cholesky Ordering: DLOG(IPC) DLOG(DOW)			

Concluimos esta sección comentando que el ejercicio de descomposición de varianza se puede igualmente instrumentar considerando una ortogonalización estructural de los impulsos, como se indicó anteriormente respecto al ejercicio de neutralidad monetaria. En tal caso, una vez estimado el VEC (VAR en la versión 5 de Eviews; ya que la

descomposición no se puede realizar sobre un VEC) se podría también estimar la descomposición de la varianza, aunque en tal caso se tendría que hacer referencia a la varianza *condicional*, ya que se estima restringiendo la respuesta de alguna variable en el sistema. Clarida y Gali (1994) presentan un ejercicio muy transparente sobre la estimación de un VAR estructural y la descomposición de la varianza. El mismo se refiere a estimar el efecto de choques nominales sobre el tipo de cambio real, de tal forma que la restricción que se impone en el sistema es sobre la respuesta del tipo de cambio real a choques en variables nominales, haciendo referencia a la teoría de la neutralidad nominal. El lector interesado encontrará en esta referencia un ejemplo claro sobre la aplicación de la descomposición de la varianza.

#### **9.4. Descomposición Tendencia-Ciclo**

Aquí se presenta una aplicación del análisis multivariado de series de tiempo que en recientes fechas ha tomado particular importancia; y se refiere a la identificación de ciclos comunes en un conjunto de variables. Nótese que a través del análisis de cointegración que se llevó a cabo en el Capítulo 8 se entendió el concepto como la existencia de tendencia(s) común(es) entre series de tiempo. Esto sugiere que si los componentes tendenciales de dos series, por ejemplo, se comportan de manera similar, entonces se podría construir una tendencia que les sea común a ambos. De similar manera, si dos series presentan un comportamiento cíclico similar, se podría entonces construir un ciclo común. En efecto, una serie de tiempo se puede entender como la suma de dos componentes, uno de tendencia y otro estacionario; que constituye su ciclo. El estudio sobre la descomposición de series de

tiempo en sus componentes es amplio, y escapa el alcance del texto, sin embargo aquí presentamos una discusión que puede motivar al lector a profundizar en el tema.<sup>36</sup>

Beveridge y Nelson (1981) sugieren que una serie de tiempo puede ser descompuesta en dos partes, una que se refiere a una tendencia que sigue un proceso de caminata aleatoria y otra que se representa como un proceso ARMA( $p, q$ ), y que se entiende como el ciclo de la serie. Formalmente, la serie  $y_t$  se descompone como:

$$y_t = \tau_t + c_t \tag{9.4}$$

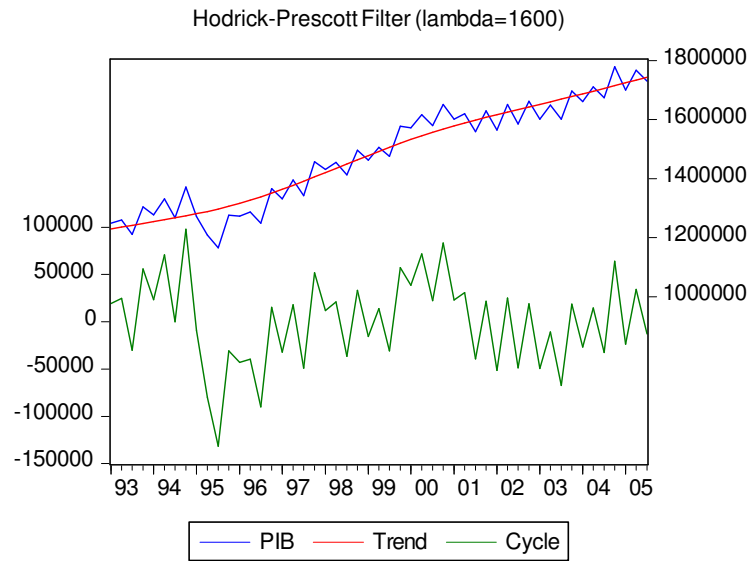
Donde  $\tau_t = \phi + \tau_{t-1} + \eta_t$  y  $A(L)c_t = B(L)\varepsilon_t$  con los polinomios  $A(L)$  y  $B(L)$  representando los componentes  $p$  y  $q$  del proceso ARMA. La tarea de descomposición se puede realizar por medio de diversas metodologías. Como se comentó en el Capítulo 5, comúnmente se empleo un Filtro para realizar el ejercicio.<sup>37</sup> Sin embargo, esta metodología presenta serias deficiencias; que ilustramos con el siguiente ejemplo que considera el PIB. Primero, aplicamos el filtro Hodrick-Prescott (HP) a la serie para el periodo 1993:1 a 2005:3. La serie filtrada se presenta en la Gráfica 9.1.

---

<sup>36</sup> La literatura relevante es extensa, aquí mencionamos algunos de los trabajos más representativos: Beveridge y Nelson (1981), Watson (1986), Stock y Watson (1988), Blanchard y Quah (1989), Engle y Kozicki (1993), y Vahid y Engle (1993).

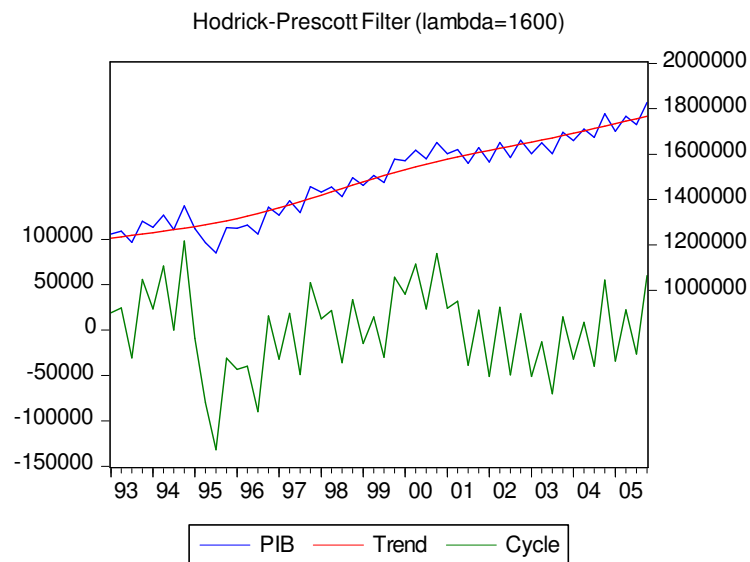
<sup>37</sup> Aunque en el texto sólo nos referimos al Filtro HP, indicamos que los comentarios se aplican igualmente a cualquier otro tipo de filtro que no incluya restricciones de cointegración o ciclo común.

**Gráfica 9.1. Filtro HP 1993:1-2005:3**



Enseguida, se calcula la tendencia considerando el periodo hasta el cuarto trimestre del 2005, esto es, se le suma una observación más a la serie. Los resultados se ilustran en la Gráfica 9.2.

**Gráfica 9.2 Filtro HP 1993:1-2005:4**



En principio, ya que sólo sumamos una observación, podríamos esperar que sólo el último valor de la tendencia cambiara; con respecto a la serie que termina en 2005:3. Si examinamos las gráficas, éste es el caso. Sin embargo, como lo ilustra el Cuadro 9.9, que muestra los valores de las series de tendencia para el año 2005, los valores entre las dos series para los primeros trimestres de 2005 son diferentes, esto es, la tendencia estimada al agregar una observación a la serie original cambia totalmente; de tal forma que no se puede considerar la tendencia que se obtiene con el Filtro HP como confiable.

**Cuadro 9.9. Comparación de Tendencias HP**

2005Q1	1724105.75	1734284.50
2005Q2	1733398.75	1745153.70
2005Q3	1742683.69	1756081.59
2005Q4		1767047.04

Alternativamente, es posible llevar a cabo la descomposición tendencia-ciclo considerando un sistema multivariado, dentro del cual es factible introducir restricciones que consideren posibles relaciones de cointegración.

En las siguientes líneas describimos brevemente la idea general de esta descomposición.

Supongamos que existe un vector de  $n$  variables que se encuentran integradas de orden 1,

$y_t$ , y cuya transformación en primeras diferencias es estacionaria y se representa como:

$$\Delta y_t = C(L)e_t = C(1)e_t + (1-L)C^*(L)e_t \quad (9.5)$$

Donde  $C(L)$  es una matriz polinomial en el operador de rezagos  $L$ . Integrando ambos lados, (9.5) se puede re-escribir como

$$y_t = C(1)\sum_{i=0}^{\infty} e_{t-i} + C^*(L)e_t \quad (9.6)$$

Que es precisamente una versión multivariada de la descomposición tendencia-ciclo de Beveridge y Nelson, donde el primer término del lado derecho se refiere a la tendencia y el segundo al ciclo.

En este sistema multivariado la tendencia para una serie en particular se puede estimar por medio de un VEC, por ejemplo, normalizando el vector de cointegración con respecto a la variable de interés. De esta forma se incluiría la restricción de cointegración en el cálculo de la tendencia. Esta metodología ya representa una mejora importante con respecto a la metodología del Filtro, que no impone ninguna restricción sobre la dinámica de las series. Sin embargo, como lo argumenta Vahid y Engle (1993), resulta importante considerar la posible existencia no sólo de relaciones estables de largo plazo entre las variables, sino también la existencia de posibles relaciones estables de corto plazo.

Esto es, retomando la expresión (9.6), notamos que la existencia de cointegración implica que  $\alpha' C(1) = 0$ , donde  $\alpha$  incluye los  $r$  vectores de cointegración. De igual forma, para el componente cíclico sería posible encontrar que  $\alpha' C^*(L) = 0$ ; lo cual indicaría que existen relaciones estables de corto plazo en el sistema, tal como la condición  $\alpha' C(1) = 0$  sugiere la

existencia de relaciones estables de largo plazo. Cuando se cumple la condición  $\alpha' C^*(L) = 0$ , se dice que existen  $s$  relaciones de ciclo común y  $\alpha$  agrupa los vectores de comovimiento, con la restricción  $r + s \leq n$ . En otras palabras, el concepto de ciclo común es el equivalente al de cointegración pero para el corto plazo. ¿Por qué es importante tomar en cuenta la posible existencia de comovimientos cíclicos? Por que tal y como se muestra en Issler y Vahid (2001) para variables de Estados Unidos, y Herrera (2004) para el caso del PIB de México, no incluir las restricciones que se refieren a los vectores de comovimiento resulta en la subestimación del efecto de choques transitorios, por ende, en un cálculo no eficiente de la tendencia común. Así, para un sistema en el que se identifica cointegración, la técnica que produce estimaciones de tendencias y ciclos comunes más eficientes se refiere a ésta sugerida por Vahid y Engle.<sup>38</sup>

En general, la metodología de Vahid y Engle considera en primera instancia el cálculo de las correlaciones canónicas entre las diferencias de las variables y su historia, la cual se define por el rezago del término de corrección de error y el rezago de las diferencias de los elementos de la matriz de variables explicativas,  $X_t$ . Enseguida, se estiman las correlaciones canónicas cuadradas ( $\lambda_i^2 \forall i = 1, 2, \dots, n$ ) que se obtienen resolviendo un problema de eigenvalues. Una vez realizado este ejercicio, se evalúa la posible rotación del sistema de tal forma que las variables que se encuentran en el lado derecho del mismo (aquellas que constituyen la historia) no presenten correlación entre ellas. Enseguida se procede a realizar una prueba de hipótesis sobre el número de combinaciones lineales de las variables

---

<sup>38</sup> En el caso en el que no exista cointegración entre las variables, se recomienda emplear la metodología sugerida por Engle y Kozicki (1993) para series estacionarias (de tal forma que si el sistema lo constituyen series I(1) se tendría que trabajar con primeras diferencias). La misma se fundamenta en la estimación del sistema en dos etapas, siendo los instrumentos los rezagos de las series estacionarias.

diferenciadas que cumplen con las siguientes características: no pueden ser pronosticadas, y eliminan los patrones de correlación serial en el sistema. En notación de Vahid y Engle,  $s$  representa este número de combinaciones lineales y de hecho es igual al número de correlaciones canónicas igual a cero, y constituyen vectores de comovimiento; que de hecho es posible mostrar que  $s \leq n - r$ . La estadística de prueba para la hipótesis nula de que las  $s$  correlaciones canónicas cuadradas son igual a cero ( $\lambda_j^2 = 0, \forall j = 1, 2, \dots, s$ ) se calcula con la siguiente expresión:<sup>39</sup>

$$C(p, s) = -(T - p - 1) \sum_{j=1}^s \log(1 - \lambda_j^2)$$

Bajo la hipótesis nula de que la dimensión del espacio de vectores de co-movimiento es al menos  $s$ , esta estadística tiene una distribución  $\chi^2$  con grados de libertad igual a  $s(np + r) - s(n - s)$ .<sup>40</sup>

Para ilustrar la aplicación de esta metodología, se presentan resultados obtenidos en Castillo y Ramírez (2006) para un sistema que considera el PIB de México, el de Estados Unidos y el de Canadá para el periodo 1980:1-2006:4.<sup>41</sup> La fuente de los datos para Estados Unidos es el Bureau of Economic Analysis, y Statistics Canada para Canadá. El Cuadro 9.10 muestra los resultados del análisis de cointegración y ciclo común normalizando los vectores con respecto al PIB de Estados Unidos (EU).

---

<sup>39</sup> Esta discusión sigue de cerca de Herrera (2004).

<sup>40</sup> Aquí incluimos sólo una reseña de la metodología de ciclos comunes, Issler y Vahid (2001) incluyen una descripción detallada de la misma.

<sup>41</sup> Las estimaciones se realizaron instrumentando el programa en GAUSS desarrollado por Issler y Vahid, modificado por Jorge Herrera de Banco de México, y ajustado por Castillo y Ramírez. Ver Goldberger (1994) para una ilustración de los fundamentos del programa GAUSS.

**Cuadro 9.10. Resultados de Prueba de Cointegración y de Ciclos Comunes**

<u>Cointegración</u>					
Hipótesis R	Estadístico de Traza	0.05 Valor Crítico	Vector Normalizado de Cointegración		
			EU	CAN	MEX
0	37.62	29.80			
1	11.03	15.49	1	-0.55	-0.13
2	0.96	3.84			
<u>Ciclo Común</u>					
Hipótesis S	Probabilidad		Vector Normalizado de Comovimiento		
	0.05		EU	CAN	MEX
>0	0.65		1	0	-0.72
>1	0.18		0	1	-0.48
>2	0.00				

El vector de cointegración indica una asociación positiva entre el PIB de Estados Unidos y los correspondientes a Canadá (CAN) y México (MEX). La magnitud del coeficiente de Canadá es mayor relativo al de México; lo cual es congruente con la idea sobre la elevada sincronización de las economías de Estados Unidos y Canadá. Los resultados del análisis de ciclo común muestran la existencia de dos relaciones estables de corto plazo, con un vector de comovimiento que sugiere una relación positiva entre los ciclos de Estados Unidos y Canadá con respecto a México.

En este caso se tiene que la suma del número de vectores de cointegración ( $r$ ) y el número de ciclos comunes ( $s$ ) es igual al número de variables ( $n$ ) por lo que se puede realizar una descomposición de las series en tendencia y ciclo común de en el espíritu de Beveridge y Nelson (1981). En particular, siguiendo con la notación anterior, el componente de tendencia y cíclico puede derivarse reconociendo que

$$y_t = \tilde{\alpha}(\tilde{\alpha}'\tilde{\alpha})^{-1}\tilde{\alpha}'y_t + \alpha(\alpha'\alpha)^{-1}\alpha'y_t$$

Donde el primer término de la expresión representa la tendencia y el segundo el ciclo. En el Cuadro 9.10 se presentan los componentes  $\tilde{\alpha}(\tilde{\alpha}'\tilde{\alpha})^{-1}\tilde{\alpha}'$  y  $\alpha(\alpha'\alpha)^{-1}\alpha'$ .

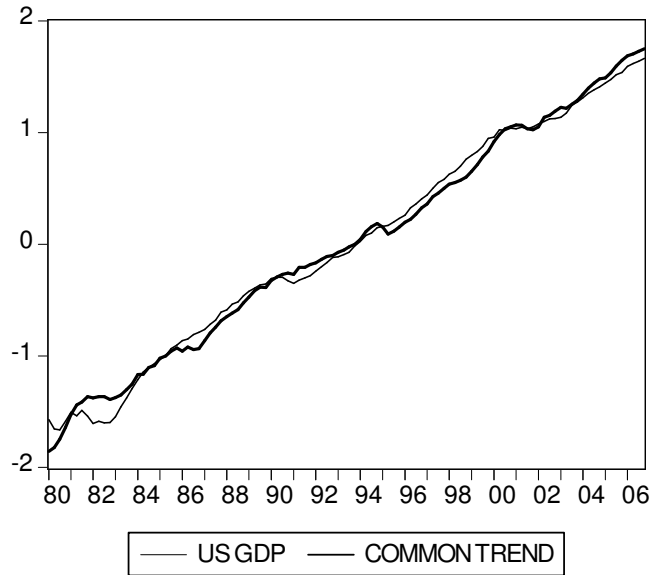
**Cuadro 9.10. Componentes Tendencial y Cíclico**

Matriz del Componente Tendencial		
EU	CAN	MEX
0.62	0.47	-0.83
0.05	0.94	0.1
-0.26	0.32	0.44
Matriz del Componente Cíclico		
EU	CAN	MEX
0.38	-0.47	0.83
-0.05	0.06	-0.1
0.26	-0.32	0.56

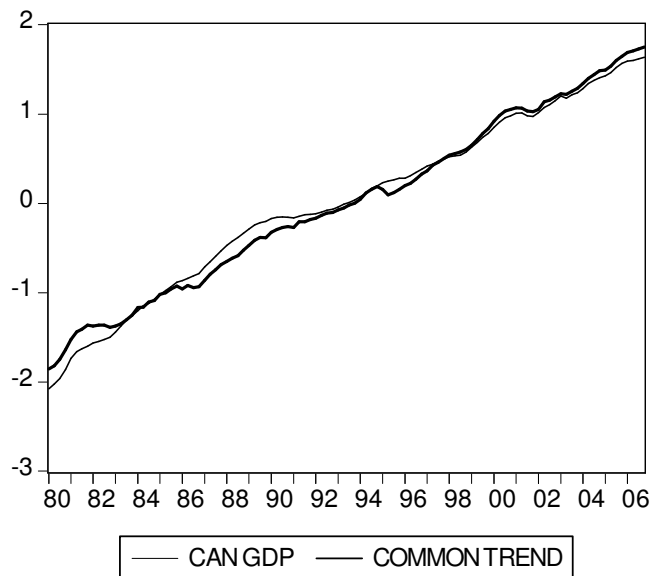
A partir de estos coeficientes es posible construir la tendencia común multiplicando los números de la diagonal principal de la parte superior del cuadro por el vector de series que se han considerado. Esto es, se multiplica la serie del PIB de Estados Unidos por 0.62, la de Canadá por 0.94 y la de México por 0.44. Las series que resultan se suman y se obtiene la tendencia común del sistema. La tendencia, conjuntamente con el PIB de cada país, se ilustra en las Gráficas 9.3, 9.4 y 9.5. En cada una de ellas se aprecian los diferentes periodos de auge y desaceleración que han experimentado las economías. Por ejemplo, en el caso de Estados Unidos la recesión de principios de los 80 es evidente. El periodo de

crecimiento de mediados de los 90 es también claro en la Gráfica de Canadá. Para México, la crisis de 1994 es captada puntualmente por la ilustración.<sup>42</sup>

**Gráfica 9.3. Tendencia Común y PIB de EU**

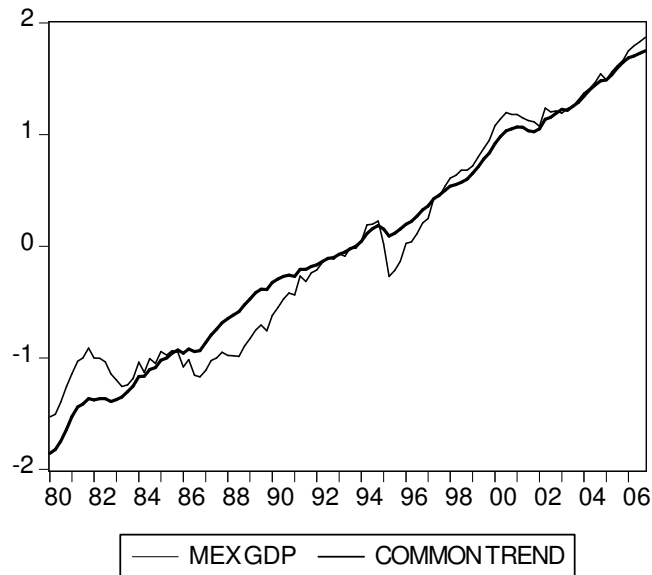


**Gráfica 9.4 Tendencia Común y PIB de Canadá**



<sup>42</sup> Aplicaciones prácticas de esta construcción para el caso de México se encuentran en Herrera y Castillo (2003) y Castillo y Herrera (2005).

**Gráfica 9.5. Tendencia Común y PIB de México**



Evidentemente, el análisis sobre la descomposición tendencia-ciclo de una serie de tiempo es muy amplio y profundo. Aquí sólo presentamos las ideas básicas y referimos al lector interesado a revisar las referencias que indicamos.

## Bibliografía

Bartlett M. S. (1946): "On the Theoretical Specification of Sampling Properties of Autocorrelated Time Series" *Journal of the Royal Statistical Society Series B*, 27, pp. 27-41.

Beveridge S. y C. R. Nelson (1981): "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of Business Cycle" *Journal of Monetary Economics*, 7, pp. 151-174.

Bewley (1979): "The Direct Estimation of the Equilibrium Response in a Linear Dynamic Model" *Economics Letters*, 3, pp. 357-361.

Blanchard O. y D. Quah (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances" *American Economic Review*, 79, pp. 655-673.

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

Box G. y D. Pierce (1970): "Distribution of Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Model" *Journal of the American Statistical Association*, 65, pp. 1509-1526.

Castillo R. (2003): "Restricciones de Liquidez, Canal de Crédito, y Consumo en México" *Economía Mexicana, Nueva Época*, 12, pp. 65-101.

Castillo R. y A. Díaz-Bautista (2002): "Testing for Unit Roots: Mexico's GDP" *Momento Económico*, 124, pp. 2-10.

Castillo R. y J. Herrera (2005): "El Efecto del Gasto Público sobre el Gasto Privado en México: Un Ejercicio Econométrico" *Estudios Económicos*, 20, pp. 173-196.

Castillo R. y R. Ramírez (2006): "How Integrated are the NAFTA Economies" *Documento de Trabajo*, Universidad Autónoma de Baja California.

Conesa A. (1998): "Pass-Through del Tipo de Cambio y del Salario: Teoría y Evidencia para la Industria Manufacturera en México" Documento de Trabajo 1998-03, Banco de México.

Clarida R. y J. Gali (1994): "Sources of Real Exchange Rate Fluctuations: How Important are Nominal Shocks?" *Carnegie-Rochester Conference Series on Public Policy*.

Carrascal U., Y. González, González R. B.(2001). *Análisis Econométrico con EViews*. Alfaomega Grupo Editor. México.

Dickey D. y W. Fuller. (1979): "Distribution of the Estimators for Autorregresive Time Series White a Unit Root" *Journal of the American Statistical Association*, 74, pp. 427-431.

Engle R y C. W. Granger (1987): Co-integration and Error Correction: Representation, Estimation, and Testing" *Econometrica*, 55, pp. 251-276.

Engle R. y S. Kozicki (1993): "Testing for Common Features" *Journal of Business and Economic Statistics*, 11, pp. 369-395.

Engle, R. y B. Yoo (1991): "Cointegrated Economic Time Series: An Overview with New Results" en: Engle, R. y C. W. Granger (Eds.), *Long-run Economic Relationships - Readings in Cointegration*, Oxford et al.

Goldberger, A. (1994): *A Course in Econometrics*. Harvard University Press.

Greene H. W. (1999): *Análisis Económico*. Tercera edición, Prentice Hall.

Hamilton, J. (1994): *Time Series Analysis*, Princeton University Press.

Hansen H. y S. Johansen (1999): "Some Tests for Parameter Consistency in Cointegrated Var Models" *Econometrics Journal*, 2, pp. 306-333.

Hansen H. y S. Johansen (1992): "Recursive Estimation in Cointegrated VAR Models" *Discussion Paper* (92-13) Institute of Economics University of Copenhagen.

Herrera J. (2004): "Business Cycles in Mexico and the United States: Do They Share Common Movements?" *Journal of Applied Economics*, 7, pp. 303-323.

Herrera J. y R. Castillo (2003): "Trends and Cycles: How Important are Long-run and Short-run Restrictions? The Case of Mexico" *Estudios Económicos*, 18, pp. 133-155.

Hodrick R., y E. C. Prescott (1997): "Postwar U.S. Business Cycles: an Empirical Investigation" *Journal of Money, Credit and Banking*, 29, pp. 1-16.

Holt C. C. (1957): "Forecasting Seasonal and Trend by Exponentially Weighted Moving Average" Office of Naval Research, Memorandum núm. 52.

Issler J. y F. Vahid (2001): "Common Cycles and the Importance of Transitory Shocks to Macroeconomics Aggregates" *Journal of Monetary Economics*, 47, pp. 449-475.

Johansen S. (1991): "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models" *Econometrica*, 59, pp. 1551-1580.

Johnston J. y J. DiNardo (1997): *Econometric Methods* 4a Edition, The McGraw-Hill Companies.

Juselius K. y H. Hansen (1995): *CATS in RATS. Manual to Cointegration Analysis of Time Series*. Distributed by ESTIMA, Evanstone, Illinois, USA.

Kwiatkowski, D., P. Phillips, P. Schmidt, y Y. Shin (1992): "Testing the Null Hypothesis of Stationary against the Alternative of a Unit Root" *Journal of Econometrics*, 54, pp. 159-178.

Ljung G.M. y G. Box (1978): "On a Measure of Lack of Fit in Time Series Models" *Biometrika*, 66, pp. 66-72.

Makridakis, S., S. C. Wheelwright y R. J. Hyndman (1998): *Forecasting Methods and Applications*, 3rd edition, New York, John Wiley - Sons.

Mendenhall, W., y E. J. Reinmuth (1981): *Estadística para Administración y Economía*, Grupo Editorial Iberoamericano.

Patterson K. (2000): *An Introduction to Applied Econometrics: a Time Series Approach* Palgrave. London.

Perron P. (1989): "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis" *Econometrica*, 57, pp.1361-1401.

Phillips, P. C. B. y S. Ouliaris (1990): "Asymptotic Properties of Residual Based Tests for Cointegration" *Econometrica*, 58, pp. 165-193.

Phillips, P.C.B. y P. Perron (1988): "Testing for a Unit Root in Time Series Regression," *Biometrika*, 75, pp. 335-346.

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

Stock J. y M. Watson (1988): "Testing for Common Trends" *Journal of the American Statistical Association*, 83, pp. 1097-1107

Vahid F. y Engle R. (1993): "Common Trends and Common Cycles" *Journal of Applied Econometrics*, 8, pp. 341-360.

Villagomez, A. (1997): "Private Saving, Interest Rates and Liquidity Constraints in LDC's: Recent Evidence." *Applied Economics*, 29, pp. 607-615.

Watson M. (1986): "Univariate Detrending Methods with Stochastic Trends" *Journal of Monetary Economics*, 18, pp. 49-75

Winters, P.R. (1960): "Forecasting Sales by Exponentially Weighted Moving Averages" *Management Science*, 6, pp. 324-342.

Zivot E. and D. Andrews (1982): "Further Evidence on the Great Crash, the Oil Price Shock and the Unit Root Hypothesis" *Journal of Business and Economic Statistics*, 10, pp. 251-270.