



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

“DETERMINANTES DEL CRECIMIENTO
ECONÓMICO EN MÉXICO:
UN ANÁLISIS DE VECTORES AUTORREGRESIVOS”

REPORTE DE
SEMINARIO DE
TITULACIÓN

QUE PARA OBTENER EL TÍTULO DE:
ACTUARIO

P R E S E N T A:

LEO ERNESTO
SANDOVAL HERNÁNDEZ

TUTOR:
M.C. MICHEL ROJAS ROMERO
2009





Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Hoja de datos del Jurado

1. Datos del alumno

Sandoval

Hernández

Leo Ernesto

56 98 58 81

Universidad Nacional Autónoma de México

Facultad de Ciencias

Actuaría

91309158

2. Datos del tutor

M en C

Rojas

Romero

Michel

3. Datos del sinodal 1

Mtra

María del Pilar

Alonso

Reyes

4. Datos del sinodal 2

Act

Jaime

Vázquez

Alamilla

5. Datos del sinodal 3

Dra

Guillermina

Eslava

Gómez

6. Datos del sinodal 4

Mtro

José Antonio

Flores

Díaz

7. Datos del trabajo escrito

Determinantes del crecimiento económico en México:
un análisis de vectores autorregresivos.

172 p

2009

Deseo agradecer a:

Dios

Mi Padre, mi Hermano, mi Amigo.

M. en C. Michel Rojas Romero

Por ser mi profesor, por dirigir este proyecto de Seminario de titulación, por toda su dedicación y tiempo para la realización de este trabajo.

Mtra. María del Pilar Alonso Reyes

Act. Jaime Vázquez Alamilla

Dra. Guillermina Eslava Gómez

Mtro. José Antonio Flores Díaz

Sinodales del Seminario, por su tiempo e interés en revisar este reporte, por todas sus observaciones, correcciones y sugerencias.

“Vivimos en el mundo cuando amamos.

Sólo una vida vivida para los demás

merece la pena ser vivida”.

Albert Einstein

Dedico este trabajo a:

Mi madre, ejemplo de amor, fe, esperanza, responsabilidad y lucha cotidiana.

Mis hermanos, mis tíos y mis primos, orgullosamente universitarios. A toda mi familia.

Miguel Ángel Saloma Robles, amigo cuyas palabras han sido correctas y a tiempo, sobre todo para que este anhelo se vea cumplido.

Gustavo, Carlos y Malena, mis amigos y soportes incondicionales.

María Leticia Zebada, mi primera maestra y motivadora.

Act. José Antonio Perera y Mat. Felipa Martínez, profesores de la ENP 4 que con su guía y apoyo inclinaron la balanza a favor de la actuaría.

M. en C. Michel Rojas Romero, maestro y amigo.

Los sinodales de este reporte, que contribuyeron con sus invaluable conocimiento y disposición.

Los luchadores sociales, fuente de inspiración permanente.

Todos mis amigos que con su respaldo aportaron para la culminación de este proyecto.

ÍNDICE

Introducción	7
--------------------	---

PRIMERA PARTE

MARCO TEÓRICO VAR

CAPÍTULO 1

Modelos multiecuacionales de series de tiempo

1.1 Análisis VAR.....	9
1.1.1 Definición formal de un VAR en forma estándar.....	14
1.2 Estimación	17
1.2.1 La función de verosimilitud condicional.....	19
1.2.2 EMV de Φ	22
1.2.3 EMV de Ω	26
1.3 La función impulso–respuesta	28
1.3.1 Innovaciones ortogonales	29
1.4 Descomposición de varianza	31
1.5 Pruebas de hipótesis.....	35
1.5.1 Pruebas para el orden del VAR	35
1.5.2 Prueba para la causalidad de Granger	38

CAPÍTULO 2

Cointegración y modelos de corrección de error

2.1 Introducción.....	42
2.2 Combinaciones lineales de variables integradas	43
2.3 Integración, cointegración y corrección de error.....	44
2.4 Cointegración y tendencias comunes.....	49
2.5 Cointegración y corrección de error	52

2.5.1 El caso de n variables	63
2.6 Metodología de Johansen para probar cointegración	66
2.6.1 Raíces características, rango y cointegración	66
2.6.2 Pruebas de hipótesis.....	73
SEGUNDA PARTE	
APLICACIÓN DEL MODELO VAR	
CAPÍTULO 3	
Determinantes del crecimiento económico en México, de 1980 a 2007	
3.1 Introducción.....	80
3.2 Modelo teórico y metodología econométrica	81
3.3 Análisis empírico basado en el modelo reducido	90
3.4 Análisis empírico basado en el modelo ampliado	103
Conclusiones	117
APÉNDICES	
A Pruebas de hipótesis	121
A.1 Pruebas de raíz unitaria.....	121
A.1.1 Prueba de Dickey–Fuller	121
A.1.2 Prueba de Dickey Fuller Aumentada	126
A.2 Pruebas congruencia estadística para el VAR	127
A.2.1 Prueba de autocorrelación LM (multiplicador de Lagrange) de Breusch–Godfrey	127
A.2.2 Prueba de normalidad de Jarque–Bera	128
A.2.3 Prueba de heteroscedasticidad de White	131
B Cuadros de resultados – Modelo aplicado.....	134
C Series temporales macroeconómicas.....	158
Bibliografía y referencias	170

INTRODUCCIÓN

El objetivo de este trabajo es realizar una investigación aplicada a partir de teorías y metodologías econométricas establecidas respecto a los determinantes del crecimiento económico en México, durante el periodo que va del primer trimestre de 1980 al último trimestre de 2007. Para ello, se utiliza la metodología de vectores autorregresivos (VAR) aplicada a dos modelos de tipo post Keynesiano. El primero de ellos se especifica, estima y reparametriza utilizando las variables elasticidades ingreso de exportaciones e importaciones y dinámica del resto del mundo y el segundo, agregando las variables elasticidades precio de las importaciones y exportaciones, los términos de intercambio y los flujos netos de capital.

La primera parte del trabajo presenta un marco teórico dividido en dos capítulos, en los que se exponen algunos elementos de la teoría VAR, sin y con cointegración. En la segunda parte, se aplica la metodología VAR para hacer una exploración de los determinantes del crecimiento económico en México. Para cerrar se presentan un resumen y algunas conclusiones. De forma adicional se incluye un apéndice al respecto de las pruebas de raíz unitaria, así como los cuadros completos con los resultados de las pruebas econométricas utilizadas, y un anexo estadístico que describe la información que se usó en el estudio, así como la bibliografía.

En el análisis econométrico se utiliza información del producto interno bruto de México y de Estados Unidos, exportaciones e importaciones reales totales de México, flujos netos de capital y términos de intercambio.

PRIMERA PARTE

MARCO TEÓRICO VAR

CAPÍTULO I

Modelos multiecuacionales de series de tiempo

En este primer capítulo es presentada la teoría correspondiente a los modelos multiecuacionales de series de tiempo¹. Inicialmente se expone el concepto de Vector Autorregresivo (VAR), así como algunas de las herramientas utilizadas en el análisis VAR, tales como la causalidad de Granger, el análisis impulso–respuesta y la descomposición de la varianza. En los casos en los que estas herramientas son aplicadas, ayudan a entender la interrelación entre variables y a diseñar modelos más estructurados.

I.1 ANÁLISIS VAR

Respecto a la identificación dentro un sistema de ecuaciones, cuando no se sabe si en realidad alguna variable es exógena, cada una se trata de manera simétrica. En el caso bivariado, esto significa que la trayectoria temporal de $\{x_t\}$ es afectada

¹ En términos generales los modelos multiecuacionales se caracterizan por presentar un sistema interconectado de variables y ecuaciones, es decir, un sistema en el que la simultaneidad entre variables endógenas (o dependientes en un modelo de regresión) aparece en mayor o menor medida. Precisamente esa mayor o menor simultaneidad en las relaciones entre variables endógenas es un factor decisivo para determinar las propiedades de los distintos métodos de estimación. Esto no significa que sea la única variable a considerar (afectará también la identificación del modelo o el deseo de una estimación asintóticamente eficiente), pero sí resulta el primero de los factores analíticamente claves para una primera aproximación al método de estimación correcto. Y son precisamente las dificultades que representan tanto la identificación (el conocer la naturaleza de una variable como exógena o endógena, explicativa o explicada dentro de un modelo multiecuacional de series de tiempo) como su adecuada estimación lo que motivó el surgimiento de la teoría VAR.

por las realizaciones actuales y pasadas de la sucesión $\{z_t\}$ y viceversa. Esta simetría se puede representar en el sistema siguiente:

$$x_t = b_{10} - b_{12}z_t + \gamma_{11}x_{t-1} + \gamma_{12}z_{t-1} + e_{x_t} \quad (1.1)$$

$$z_t = b_{20} - b_{21}x_t + \gamma_{21}x_{t-1} + \gamma_{22}z_{t-1} + e_{z_t} \quad (1.2)$$

Se asume que las series x_t y z_t son estacionarias², e_{x_t} y e_{z_t} son procesos estocásticos de ruido blanco gaussiano, es decir, son independientes y cada uno sigue una distribución normal con media cero y desviación estándar σ_x y σ_z respectivamente ($e_{y_t} \sim i.i.d. N(0, \sigma_y^2)$), y las sucesiones de ruidos blancos gaussianos $\{x_t\}$ y $\{z_t\}$ no están correlacionadas.

Las ecuaciones (1.1) y (1.2) conforman un VAR de primer orden porque el rezago mayor es uno. La estructura de este VAR de dos variables incorpora efectos de retroalimentación entre ambas. De esta manera, $-b_{12}$ es el efecto contemporáneo de un cambio unitario de z_t sobre x_t , y γ_{12} es el efecto de un cambio unitario de z_{t-1} sobre la misma x_t . Los términos e_{x_t} y e_{z_t} son innovaciones o *shocks*³ en x_t y z_t , respectivamente. Es importante notar que si $b_{12} \neq 0$, entonces e_{z_t} tiene un efecto indirecto contemporáneo sobre x_t ; y del mismo modo, si $b_{21} \neq 0$, para e_{x_t} sobre z_t .

² “Se dice que un proceso estocástico es estacionario de segundo orden, estacionario en covarianza, o proceso estocástico débilmente estacionario, si su media y su varianza son constantes en el tiempo y si el valor de la covarianza entre dos periodos depende solamente de la distancia o rezago entre éstos y no del tiempo en el cual se ha calculado la varianza.” Gujarati (2004), p. 772.

En lo que se refiere a la estacionariedad de una serie macroeconómica, los conceptos vinculados a esta condición serán presentados más adelante tanto en el capítulo 2 de esta primera parte, como en el apéndice A.1, pues revisten una gran importancia en el tratamiento de los modelos VAR; ya que tratándose de series estacionarias o no estacionarias las involucradas en los modelos, será que se deba trabajar un VAR sin restricciones o un VAR cointegrado. Pero como se plantea en este punto, se arranca con el supuesto de que todas las series disponibles son estacionarias, y así será para todo este primer capítulo.

³ También llamados perturbaciones estocásticas, errores, impulsos o residuos muestrales, dado el caso.

El sistema de ecuaciones (1.1) y (1.2) no está en forma reducida porque x_t tiene un efecto indirecto contemporáneo sobre z_t y viceversa. Reescribiendo el sistema:

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} x_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} e_{x_t} \\ e_{z_t} \end{bmatrix}$$

o bien,

$$By_t = \Gamma_0 + \Gamma_1 y_{t-1} + e_t$$

donde

$$B = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}, y_t = \begin{bmatrix} x_t \\ z_t \end{bmatrix}, \Gamma_0 = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix}, \Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}, e_t = \begin{bmatrix} e_{x_t} \\ e_{z_t} \end{bmatrix}$$

El modelo VAR en forma estándar se obtiene como sigue:

$$B^{-1}By_t = B^{-1}\Gamma_0 + B^{-1}\Gamma_1 y_{t-1} + B^{-1}e_t$$

o

$$y_t = \Phi_0 + \Phi_1 y_{t-1} + \varepsilon_t$$

(1.3)

donde

$$B^{-1} = \frac{1}{1-b_{12}b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}, \Phi_0 = B^{-1}\Gamma_0 = \frac{1}{1-b_{12}b_{21}} \begin{bmatrix} b_{10} - b_{12}b_{20} \\ b_{20} - b_{10}b_{21} \end{bmatrix},$$

$$\Phi_1 = B^{-1}\Gamma_1 = \frac{1}{1-b_{12}b_{21}} \begin{bmatrix} \gamma_{11} - b_{12}\gamma_{21} & \gamma_{12} - b_{12}\gamma_{22} \\ \gamma_{21} - b_{21}\gamma_{11} & \gamma_{22} - b_{21}\gamma_{12} \end{bmatrix},$$

$$\varepsilon_t = B^{-1}e_t = \frac{1}{1-b_{12}b_{21}} \begin{bmatrix} e_{x_t} - b_{12}e_{z_t} \\ e_{z_t} - b_{21}e_{x_t} \end{bmatrix}.$$

Definiendo los elementos ϕ_{j0} de Φ_0 , ϕ_{lj} de Φ_1 y ε_{jt} de ε_t , con $j, l = 1, 2$; como

$$\Phi_0 = \begin{bmatrix} \phi_{10} \\ \phi_{20} \end{bmatrix}, \Phi_1 = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}, \varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

(1.3) se puede reescribir en la forma:

$$x_t = \phi_{10} + \phi_{11}x_{t-1} + \phi_{12}z_{t-1} + \varepsilon_{1t} \tag{1.4}$$

$$z_t = \phi_{20} + \phi_{21}x_{t-1} + \phi_{22}z_{t-1} + \varepsilon_{2t} \tag{1.5}$$

El sistema de ecuaciones formado por (1.1) y (1.2) se llama VAR estructural, y el que está compuesto por (1.4) y (1.5) es conocido como VAR en forma estándar. El modelo desarrollado para el análisis en este estudio tiene la forma del VAR estándar.

Precisando sobre la correlación entre las innovaciones ε_{jt} : ya que $\varepsilon_t = B^{-1}e_t$, tanto ε_{1t} como ε_{2t} están compuestos por los dos *shocks* e_{x_t} y e_{z_t} .

En efecto,

$$\varepsilon_{1t} = \frac{e_{x_t} - b_{12}e_{z_t}}{1 - b_{12}b_{21}} \tag{1.6}$$

$$\varepsilon_{2t} = \frac{e_{z_t} - b_{21}e_{x_t}}{1 - b_{12}b_{21}} \tag{1.7}$$

Ya que e_{x_t} y e_{z_t} son procesos de ruido blanco gaussiano, entonces ε_{1t} y ε_{2t} tienen media cero, varianzas constantes y no están autocorrelacionados. Se verifica entonces, para $\{\varepsilon_{1t}\}$:

$$E[\varepsilon_{1t}] = E \left[\frac{e_{x_t} - b_{12}e_{z_t}}{1 - b_{12}b_{21}} \right] = 0$$

y

$$\begin{aligned} Var[\varepsilon_{1t}] = E[\varepsilon_{1t}^2] &= E \left[\frac{e_{x_t} - b_{12}e_{z_t}}{1 - b_{12}b_{21}} \right]^2 = E \left[\frac{e_{x_t}^2 - 2b_{12}e_{x_t}e_{z_t} + b_{12}^2e_{z_t}^2}{(1 - b_{12}b_{21})^2} \right] \\ &= \frac{\sigma_x^2 + b_{12}^2\sigma_z^2}{(1 - b_{12}b_{21})^2} \end{aligned} \quad (1.8)$$

y

$$\begin{aligned} E[\varepsilon_{1t} \varepsilon_{1t-i}] &= E \left[\frac{(e_{x_t} - b_{12}e_{z_t})(e_{x_{t-i}} - b_{12}e_{z_{t-i}})}{(1 - b_{12}b_{21})^2} \right] \\ &= E \left[\frac{e_{x_t}e_{x_{t-i}} - b_{12}(e_{x_{t-i}}e_{z_t} + e_{x_t}e_{z_{t-i}}) + b_{12}^2e_{z_t}e_{z_{t-i}}}{(1 - b_{12}b_{21})^2} \right] = 0, \text{ para } i \neq 0 \end{aligned}$$

De manera similar se puede mostrar que ε_{2t} tiene media cero, varianza constante y autocovarianza cero. Es importante notar que ε_{1t} y ε_{2t} están correlacionados ya que:

$$\begin{aligned} E[\varepsilon_{1t} \varepsilon_{2t}] &= E \left[\frac{(e_{x_t} - b_{12}e_{z_t})(e_{z_t} - b_{21}e_{x_t})}{(1 - b_{12}b_{21})^2} \right] \\ &= E \left[\frac{e_{x_t}e_{z_t} - b_{21}e_{x_t}^2 - b_{12}e_{z_t}^2 + b_{12}b_{21}e_{x_t}e_{z_t}}{(1 - b_{12}b_{21})^2} \right] = \frac{-b_{21}\sigma_x^2 - b_{12}\sigma_z^2}{(1 - b_{12}b_{21})^2} \end{aligned} \quad (1.9)$$

En general, (1.9) no es cero y entonces las dos perturbaciones estarán correlacionadas. No será así en el caso especial de que no existan efectos contemporáneos de x_t sobre z_t y viceversa, es decir, cuando $b_{12} = b_{21} = 0$. La matriz de varianzas y covarianzas de los *shocks* (Ω) es entonces:

$$\Omega = \begin{bmatrix} Var(\varepsilon_{1t}) & Cov(\varepsilon_{1t}, \varepsilon_{2t}) \\ Cov(\varepsilon_{2t}, \varepsilon_{1t}) & Var(\varepsilon_{2t}) \end{bmatrix}$$

Y ya que los elementos de Ω son independientes del tiempo, esta matriz se puede escribir en la forma

$$\Omega = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} \tag{1.10}$$

donde $Var[\varepsilon_{it}] = \sigma_i^2$ y $Cov[\varepsilon_{1t}, \varepsilon_{2t}] = \sigma_{12} = \sigma_{21} = Cov[\varepsilon_{2t}, \varepsilon_{1t}]$.

1.1.1 DEFINICIÓN FORMAL DE UN VAR EN FORMA ESTÁNDAR

Las condiciones para la generalización formal de un VAR fueron propuestas por Sims (1980) y seguidas por diversidad de autores.

Partiendo de los modelos autorregresivos univariados, donde una variable escalar es modelada en términos de sus propios valores pasados, el proceso autorregresivo de orden p o $AR(p)$, es de la forma:

$$y_t = k + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

donde

- y_t es la variable escalar en el tiempo t modelada por sus valores pasados y_{t-i} , con $i = 1, \dots, p$.
- El número de rezagos p es el que indica el orden del proceso.
- Los valores ϕ_i son los coeficientes escalares de las variables rezagadas, correspondiendo ϕ_i a cada y_{t-i} .
- k es una constante escalar a la que se denomina el desplazamiento (o intercepto) del modelo.

Un modelo como el anterior sería determinístico de no ser por el sumando ε_t , como residuo muestral para el mismo; que teniendo un comportamiento de perturbación estocástica es debido a todos aquellos factores que afectan al modelo y que no es posible incluirlos en forma determinística; se supondrá entonces que esa perturbación tiene un comportamiento como el de un proceso de ruido blanco gaussiano, es decir, $\varepsilon_t \sim i. i. d. N(0, \sigma_\varepsilon^2)$.

Ahora, extendiendo el proceso anterior a n variables distintas en el tiempo t , se define un vector columna $\mathbf{y}_t = (y_{1t}, y_{2t}, \dots, y_{nt})'$, y se plantea un modelo para éste en términos de los valores pasados de las variables dentro del vector. El resultado es un vector autorregresivo o VAR.

El proceso de un vector autorregresivo de orden p o VAR(p) es:

$$\mathbf{y}_t = \mathbf{k} + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \dots + \Phi_p \mathbf{y}_{t-p} + \varepsilon_t \quad (1.11)$$

donde:

- \mathbf{k} es un vector desplazamiento formado por constantes escalares ($n \times 1$).
- las Φ_i son matrices de coeficientes ($n \times n$) para su correspondiente \mathbf{y}_{t-i} con $i = 1, \dots, p$.

- \mathbf{y}_{t-i} es el vector ($n \times 1$) que está formado por los valores que en el tiempo ($t - i$) ha tomado \mathbf{y}_t .
- Como en el caso univariado, se observa que el vector autorregresivo sería completamente determinístico si no fuera por el vector de perturbaciones estocásticas $\boldsymbol{\varepsilon}_t$, que es un vector ($n \times 1$) de procesos de ruido blanco gaussiano, con las propiedades

$$E[\boldsymbol{\varepsilon}_t] = \mathbf{0} \quad \text{para toda } t, \quad \text{y} \quad E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s'] = \begin{cases} \boldsymbol{\Omega} & s = t \\ \mathbf{0} & s \neq t \end{cases} \quad (1.12)$$

donde $\boldsymbol{\Omega}$ es la matriz de varianzas y covarianzas ($n \times n$) de los residuos muestrales.

Entonces para el caso multivariado, $\boldsymbol{\Omega}$ puede escribirse explícitamente como:

$$\boldsymbol{\Omega} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_n^2 \end{bmatrix}$$

donde $\sigma_j^2 = Var[\varepsilon_{jt}]$ y $\sigma_{jl} = \sigma_{lj} = Cov[\varepsilon_{jt}, \varepsilon_{lt}]$, con $j, l = 1, \dots, n$ y $j \neq l$; ya que los elementos de $\boldsymbol{\Omega}$ son independientes del tiempo. Esta matriz $\boldsymbol{\Omega}$ es asumida como simétrica definida positiva.

De este modo, los vectores de perturbaciones estocásticas $\boldsymbol{\varepsilon}_t$ no están correlacionados serialmente aunque sí podrían estarlo contemporáneamente.

A continuación se revisa el sustento estadístico para un problema que desde que esta modelación econométrica fue propuesta ha sido necesario fundamentar: el de la estimación tanto de los parámetros del VAR, como de su matriz de varianzas y covarianzas. En este importante punto se considerará a Hamilton (1994), el cual

aborda con detenimiento una detallada presentación de la obtención de los estimadores.

1.2 ESTIMACIÓN

Nuevamente se puede asumir que la discusión alrededor del planteamiento de modelos VAR tiene su origen cuando Sims (1980) critica las “*increíbles restricciones de identificación*” inherentes a los modelos econométricos estructurales de ecuaciones simultáneas y propone una estrategia alternativa de estimación.

Considerando la generalización multivariada de un proceso autorregresivo de orden p ya definido en (1.11), en la metodología VAR de Sims (1980), las variables incluidas se eligen de acuerdo al modelo económico relevante. Es aplicando pruebas para la longitud del rezago que se determina un valor óptimo para p .

Se tiene entonces que el vector \mathbf{k} contiene n parámetros y cada matriz Φ_i contiene n^2 parámetros, pues para \mathbf{y}_t , cada y_{jt} que lo integra con $j = 1, \dots, n$, y_{jt} es una combinación lineal tanto de los valores rezagados de sí misma como de los de las otras variables en el sistema, como puede verse en la siguiente forma explícita del VAR:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{nt} \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} + \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} & \dots & \phi_{1n}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} & \dots & \phi_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{n1}^{(1)} & \phi_{n2}^{(1)} & \dots & \phi_{nn}^{(1)} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{n,t-1} \end{bmatrix} +$$

$$\begin{aligned}
 & + \begin{bmatrix} \phi_{11}^{(2)} & \phi_{12}^{(2)} & \dots & \phi_{1n}^{(2)} \\ \phi_{21}^{(2)} & \phi_{22}^{(2)} & \dots & \phi_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{n1}^{(2)} & \phi_{n2}^{(2)} & \dots & \phi_{nn}^{(2)} \end{bmatrix} \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \\ \vdots \\ y_{n,t-2} \end{bmatrix} + \dots \\
 & + \begin{bmatrix} \phi_{11}^{(p)} & \phi_{12}^{(p)} & \dots & \phi_{1n}^{(p)} \\ \phi_{21}^{(p)} & \phi_{22}^{(p)} & \dots & \phi_{2n}^{(p)} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{n1}^{(p)} & \phi_{n2}^{(p)} & \dots & \phi_{nn}^{(p)} \end{bmatrix} \begin{bmatrix} y_{1,t-p} \\ y_{2,t-p} \\ \vdots \\ y_{n,t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{nt} \end{bmatrix}
 \end{aligned}
 \tag{1.13}$$

En consecuencia, es necesario estimar $n + pn^2$ coeficientes. Sin duda el VAR estará sobreparametrizado porque muchos coeficientes estimados no serán significativos. No obstante, el principal propósito de esta modelación es hallar interrelaciones importantes entre las variables y no hacer pronósticos de corto plazo. Por lo tanto, el imponer de manera impropia restricciones iniciales puede hacer perder información importante. Además, es probable que los regresores sean altamente colineales de modo que las pruebas t sobre coeficientes individuales no resulten confiables para reducir el modelo.

Se observa que el lado derecho de la expresión (1.13) por definición contiene sólo variables predeterminadas y que se asume que los términos de error no están correlacionados con varianza constante. En consecuencia, cada ecuación del sistema se puede estimar por el método de mínimos cuadrados ordinarios (MCO).

1.2.1 LA FUNCIÓN DE VEROSIMILITUD CONDICIONAL

Para la obtención de estimadores para los procesos autorregresivos $AR(p)$ se ha verificado la utilidad del método de máxima verosimilitud⁴. A través de éste mismo aplicado al vector autorregresivo es posible hallar sus estimadores máximo verosímiles, debido a su particular conveniencia.

Primero se calculará la función de verosimilitud condicional, y a continuación, se encontrarán los valores de θ (es decir, los verdaderos valores de los coeficientes del vector autorregresivo) para maximizar dicha función, a partir de las observaciones en el tiempo del vector \mathbf{y}_t .

Supóngase que se cuenta con las observaciones de cada una de las n variables del lado izquierdo de (1.11) para $(T + p)$ periodos de tiempo. En la formulación de la función de verosimilitud condicional, se toma en cuenta la estimación en las últimas T observaciones $(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T)$ condicionándolas sobre las primeras p observaciones $(\mathbf{y}_{-p+1}, \mathbf{y}_{-p+2}, \dots, \mathbf{y}_0)$:

$$f_{\mathbf{y}_T, \mathbf{y}_{T-1}, \dots, \mathbf{y}_1 | \mathbf{y}_0, \mathbf{y}_{-1}, \dots, \mathbf{y}_{-p+1}}(\mathbf{y}_T, \mathbf{y}_{T-1}, \dots, \mathbf{y}_1 | \mathbf{y}_0, \mathbf{y}_{-1}, \dots, \mathbf{y}_{-p+1}; \theta) \quad (1.14)$$

y maximizarla respecto a un vector θ , que contiene los elementos de $\mathbf{k}, \Phi_1, \Phi_2, \dots, \Phi_p$ y Ω . Será (1.14) entonces ésta, la función de verosimilitud, y el valor de θ que la maximiza el estimador máximo verosímil (EMV).

⁴ Este resultado univariado se puede consultar en **Hamilton** (1994), en la Sección 5.3, pp. 123-127. Este autor desarrolla también el método de máxima verosimilitud tanto para los estimadores de procesos de medias móviles MA(q), como para procesos autorregresivos de medias móviles ARMA (p,q).

Para obtener la función de verosimilitud sobre los valores de \mathbf{y} observados hasta el tiempo $t - 1$, se tiene que \mathbf{y} para el tiempo t será igual al vector de constantes

$$\mathbf{k} + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \dots + \Phi_p \mathbf{y}_{t-p} \quad (1.15)$$

más $\boldsymbol{\varepsilon}_t$, un vector de variables aleatorias con distribución $N(\mathbf{0}, \boldsymbol{\Omega})$. De modo que

$$\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots, \mathbf{y}_{-p+1} \sim N((\mathbf{k} + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \dots + \Phi_p \mathbf{y}_{t-p}), \boldsymbol{\Omega}) \quad (1.16)$$

Para simplificar el cálculo en la obtención de la función de verosimilitud y de los estimadores máximo verosímiles, sea \mathbf{x}_t un vector con un término constante y los p rezagos de \mathbf{y} :

$$\mathbf{x}_t \equiv \begin{bmatrix} 1 \\ \mathbf{y}_{t-1} \\ \mathbf{y}_{t-2} \\ \vdots \\ \mathbf{y}_{t-p} \end{bmatrix} \quad (1.17)$$

Así, \mathbf{x}_t es un vector de $[(np + 1) \times 1]$.

Y sea Φ' la matriz de $[n \times (np + 1)]$:

$$\Phi' \equiv [\mathbf{k} \quad \Phi_1 \quad \Phi_2 \quad \dots \quad \Phi_p] \quad (1.18)$$

entonces la media condicional en (1.16) es igual a $\Phi' \mathbf{x}_t$ de $(n \times 1)$. El j -ésimo renglón de Φ' contiene los parámetros de la j -ésima ecuación del VAR. Con esta notación abreviada, (1.16) puede reescribirse más compactamente:

$$\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots, \mathbf{y}_{-p+1} \sim N(\Phi' \mathbf{x}_t, \Omega) \quad (1.19)$$

Con la información conocida sobre la media y la varianza, resulta que la función de densidad condicional para la observación t -ésima es:

$$\begin{aligned} f_{\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots, \mathbf{y}_{-p+1}}(\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots, \mathbf{y}_{-p+1}; \boldsymbol{\theta}) \\ = (2\pi)^{-\frac{n}{2}} |\Omega^{-1}|^{\frac{1}{2}} \exp\left[(-1/2) (\mathbf{y}_t - \Phi' \mathbf{x}_t)' \Omega^{-1} (\mathbf{y}_t - \Phi' \mathbf{x}_t)\right] \end{aligned} \quad (1.20)$$

La densidad conjunta de las observaciones 1 a la t condicionadas sobre $\mathbf{y}_0, \mathbf{y}_{-1}, \dots, \mathbf{y}_{-p+1}$ satisface

$$\begin{aligned} f_{\mathbf{y}_t, \mathbf{y}_{t-1}, \dots, \mathbf{y}_1 | \mathbf{y}_0, \mathbf{y}_{-1}, \dots, \mathbf{y}_{-p+1}}(\mathbf{y}_t, \mathbf{y}_{t-1}, \dots, \mathbf{y}_1 | \mathbf{y}_0, \mathbf{y}_{-1}, \dots, \mathbf{y}_{-p+1}; \boldsymbol{\theta}) \\ = f_{\mathbf{y}_{t-1}, \dots, \mathbf{y}_1 | \mathbf{y}_0, \mathbf{y}_{-1}, \dots, \mathbf{y}_{-p+1}}(\mathbf{y}_{t-1}, \dots, \mathbf{y}_1 | \mathbf{y}_0, \mathbf{y}_{-1}, \dots, \mathbf{y}_{-p+1}; \boldsymbol{\theta}) \\ \times f_{\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots, \mathbf{y}_{-p+1}}(\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots, \mathbf{y}_{-p+1}; \boldsymbol{\theta}) \end{aligned}$$

Recursivamente, la función de verosimilitud para la muestra completa $\mathbf{y}_T, \mathbf{y}_{T-1}, \dots, \mathbf{y}_1$ condicionada sobre $\mathbf{y}_0, \mathbf{y}_{-1}, \dots, \mathbf{y}_{-p+1}$ es el producto de las funciones de densidad condicional individuales:

$$\begin{aligned} f_{\mathbf{y}_T, \mathbf{y}_{T-1}, \dots, \mathbf{y}_1 | \mathbf{y}_0, \mathbf{y}_{-1}, \dots, \mathbf{y}_{-p+1}}(\mathbf{y}_T, \mathbf{y}_{T-1}, \dots, \mathbf{y}_1 | \mathbf{y}_0, \mathbf{y}_{-1}, \dots, \mathbf{y}_{-p+1}; \boldsymbol{\theta}) \\ = \prod_{t=1}^T f_{\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots, \mathbf{y}_{-p+1}}(\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots, \mathbf{y}_{-p+1}; \boldsymbol{\theta}) \end{aligned} \quad (1.21)$$

La función de log-verosimilitud de la muestra es encontrada sustituyendo (1.20) en (1.21) y tomando logaritmos:

$$\begin{aligned}
 L(\boldsymbol{\theta}) &= \ln \left\{ \prod_{t=1}^T f_{\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots, \mathbf{y}_{-p+1}}(\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots, \mathbf{y}_{-p+1}; \boldsymbol{\theta}) \right\} \\
 &= \sum_{t=1}^T \ln \left[f_{\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots, \mathbf{y}_{-p+1}}(\mathbf{y}_t | \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots, \mathbf{y}_{-p+1}; \boldsymbol{\theta}) \right] \\
 &= \sum_{t=1}^T \ln \left\{ (2\pi)^{-\frac{n}{2}} |\boldsymbol{\Omega}^{-1}|^{\frac{1}{2}} \exp \left[\left(-\frac{1}{2} \right) (\mathbf{y}_t - \boldsymbol{\Phi}' \mathbf{x}_t)' \boldsymbol{\Omega}^{-1} (\mathbf{y}_t - \boldsymbol{\Phi}' \mathbf{x}_t) \right] \right\} \\
 &= -\left(\frac{Tn}{2} \right) \ln(2\pi) + \left(\frac{T}{2} \right) \ln |\boldsymbol{\Omega}^{-1}| - \left(\frac{1}{2} \right) \sum_{t=1}^T [(\mathbf{y}_t - \boldsymbol{\Phi}' \mathbf{x}_t)' \boldsymbol{\Omega}^{-1} (\mathbf{y}_t - \boldsymbol{\Phi}' \mathbf{x}_t)]
 \end{aligned} \tag{1.22}$$

Para encontrar el valor de $\boldsymbol{\theta}$ que maximiza (1.22) a partir de las observaciones en el tiempo del vector \mathbf{y}_t , primeramente se estudia la matriz de coeficientes $\boldsymbol{\Phi}$, y a continuación la matriz de varianzas y covarianzas $\boldsymbol{\Omega}$.

1.2.2 EMV DE $\boldsymbol{\Phi}$

Se muestra primero el EMV de $\boldsymbol{\Phi}$, para validarlo inmediatamente después. Sea $\boldsymbol{\Phi}'$ tal como fue definido en (1.18). El EMV de $\boldsymbol{\Phi}'$ estará dado por $\hat{\boldsymbol{\Phi}}'$, una matriz de $[n \times (np + 1)]$:

$$\hat{\Phi}' = \left[\sum_{t=1}^T (\mathbf{y}_t \mathbf{x}_t') \right] \left[\sum_{t=1}^T (\mathbf{x}_t \mathbf{x}_t') \right]^{-1} \quad (1.23)$$

Entonces el renglón j -ésimo de dimensiones $1 \times (np + 1)$ de $\hat{\Phi}'$ será:

$$\hat{\Phi}'_j = \left[\sum_{t=1}^T (y_{jt} \mathbf{x}_t') \right] \left[\sum_{t=1}^T (\mathbf{x}_t \mathbf{x}_t') \right]^{-1} \quad (1.24)$$

el cual es justamente el vector de coeficientes estimado por una regresión por mínimos cuadrados ordinarios de y_{jt} sobre \mathbf{x}_t .⁵

Así, los EMV de los coeficientes para la j -ésima ecuación de un VAR estarán dados por una regresión por MCO de y_{jt} sobre un término constante y p rezagos de todas las variables del sistema.

Para verificar el resultado propuesto en (1.23), es tomada la suma que aparece en el último término de (1.22) y se obtiene una expresión que resulta particularmente útil:

$$\begin{aligned} & \sum_{t=1}^T [(\mathbf{y}_t - \Phi' \mathbf{x}_t)' \Omega^{-1} (\mathbf{y}_t - \Phi' \mathbf{x}_t)] \\ &= \sum_{t=1}^T [(\mathbf{y}_t - \hat{\Phi}' \mathbf{x}_t + \hat{\Phi}' \mathbf{x}_t - \Phi' \mathbf{x}_t)' \Omega^{-1} (\mathbf{y}_t - \hat{\Phi}' \mathbf{x}_t + \hat{\Phi}' \mathbf{x}_t - \Phi' \mathbf{x}_t)] \end{aligned}$$

⁵ El que $\hat{\Phi}'_j$ sea efectivamente el estimador de Φ'_j obtenido por MCO, es discutido y demostrado en Hamilton (1994), en el apéndice 4.A, pp. 113-114.

$$= \sum_{t=1}^T [(\hat{\boldsymbol{\varepsilon}}_t + (\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi})' \mathbf{x}_t)' \boldsymbol{\Omega}^{-1} (\hat{\boldsymbol{\varepsilon}}_t + (\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi})' \mathbf{x}_t)] \quad (1.25)$$

donde el j -ésimo elemento del vector $\hat{\boldsymbol{\varepsilon}}_t$ ($n \times 1$), son los residuos muestrales para la observación t obtenidos por una regresión por MCO para y_{jt} sobre \mathbf{x}_t :

$$\hat{\boldsymbol{\varepsilon}}_t \equiv \mathbf{y}_t - \hat{\boldsymbol{\Phi}}' \mathbf{x}_t \quad (1.26)$$

Expandiendo la expresión (1.25):

$$\begin{aligned} & \sum_{t=1}^T [(\mathbf{y}_t - \boldsymbol{\Phi}' \mathbf{x}_t)' \boldsymbol{\Omega}^{-1} (\mathbf{y}_t - \boldsymbol{\Phi}' \mathbf{x}_t)] \\ &= \sum_{t=1}^T \hat{\boldsymbol{\varepsilon}}_t' \boldsymbol{\Omega}^{-1} \hat{\boldsymbol{\varepsilon}}_t + 2 \sum_{t=1}^T \hat{\boldsymbol{\varepsilon}}_t' \boldsymbol{\Omega}^{-1} (\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi})' \mathbf{x}_t + \sum_{t=1}^T \mathbf{x}_t' (\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi}) \boldsymbol{\Omega}^{-1} (\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi})' \mathbf{x}_t \end{aligned} \quad (1.27)$$

En este momento se considera sólo el término intermedio en (1.27). Puesto que es un escalar, no se verá modificado por la aplicación del operador Traza:

$$\begin{aligned} \sum_{t=1}^T \hat{\boldsymbol{\varepsilon}}_t' \boldsymbol{\Omega}^{-1} (\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi})' \mathbf{x}_t &= Tr \left[\sum_{t=1}^T \hat{\boldsymbol{\varepsilon}}_t' \boldsymbol{\Omega}^{-1} (\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi})' \mathbf{x}_t \right] \\ &= Tr \left[\sum_{t=1}^T \boldsymbol{\Omega}^{-1} (\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi})' \mathbf{x}_t \hat{\boldsymbol{\varepsilon}}_t' \right] \\ &= Tr \left[\boldsymbol{\Omega}^{-1} (\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi})' \sum_{t=1}^T \mathbf{x}_t \hat{\boldsymbol{\varepsilon}}_t' \right] \end{aligned}$$

(1.28)

Pero los residuos muestrales obtenidos por MCO son por su propia construcción ortogonales a las variables explicativas, dando entonces que $\sum_{t=1}^T \mathbf{x}_t \hat{\varepsilon}_{jt} = \mathbf{0}$ para todo j y así $\sum_{t=1}^T \mathbf{x}_t \hat{\varepsilon}_t' = \mathbf{0}$. Por lo tanto, (1.28) es idénticamente cero, y (1.27) queda como:

$$\begin{aligned} & \sum_{t=1}^T [(\mathbf{y}_t - \Phi' \mathbf{x}_t)' \Omega^{-1} (\mathbf{y}_t - \Phi' \mathbf{x}_t)] \\ &= \sum_{t=1}^T \hat{\varepsilon}_t' \Omega^{-1} \hat{\varepsilon}_t + \sum_{t=1}^T \mathbf{x}_t' (\hat{\Phi} - \Phi) \Omega^{-1} (\hat{\Phi} - \Phi)' \mathbf{x}_t \end{aligned} \quad (1.29)$$

Dado que Ω es una matriz definida positiva, entonces Ω^{-1} también lo es⁶. Así, definiendo el vector \mathbf{x}_t^* de dimensiones $n \times 1$ como:

$$\mathbf{x}_t^* \equiv (\hat{\Phi} - \Phi)' \mathbf{x}_t,$$

el último término de (1.29) toma la forma

$$\sum_{t=1}^T \mathbf{x}_t' (\hat{\Phi} - \Phi) \Omega^{-1} (\hat{\Phi} - \Phi)' \mathbf{x}_t = \sum_{t=1}^T (\mathbf{x}_t^*)' \Omega^{-1} \mathbf{x}_t^*$$

Esta última expresión es positiva para cualquier sucesión $\{\mathbf{x}_t^*\}_{t=1}^T$ distinta a $\mathbf{x}_t^* = \mathbf{0}$ para todo t . Así, el valor más pequeño de (1.29) puede ser alcanzado cuando $\mathbf{x}_t^* = \mathbf{0}$, es decir, cuando $\Phi = \hat{\Phi}$. Desde que (1.29) es minimizada mediante el establecimiento de $\Phi = \hat{\Phi}$, se sigue que (1.22) es maximizada por esta misma

⁶ Esto se obtiene del hecho de que Ω^{-1} puede ser escrita como $\mathbf{L}'\mathbf{L}$, con \mathbf{L} una matriz no singular. Este resultado es obtenido para el método de mínimos cuadrados generalizados (MCG), véase **Hamilton** (1994), Sección 8.3.1.

condición. Entonces la regresión por MCO provee los EMV de los coeficientes de un vector autorregresivo, que es lo que se quería probar.

1.2.3 EMV DE Ω

Es turno del EMV de Ω . Al evaluar la función de log-verosimilitud (1.22) en el EMV de $\hat{\Phi}$ acabado de obtener resulta:

$$L(\Omega, \hat{\Phi}) = -\left(\frac{Tn}{2}\right) \ln(2\pi) + \left(\frac{T}{2}\right) \ln|\Omega^{-1}| - \left(\frac{1}{2}\right) \sum_{t=1}^T \hat{\epsilon}_t' \Omega^{-1} \hat{\epsilon}_t \quad (1.30)$$

El objetivo es encontrar una matriz simétrica definida positiva Ω para la cual (1.30) tiene el valor más grande posible. Primero está el maximizar (1.30) escogiendo a Ω como una matriz cualquiera ($n \times n$) sin restricciones. Para este propósito se hace la derivación⁷ de (1.30) respecto a los elementos de Ω^{-1} :

$$\begin{aligned} \frac{\partial L(\Omega, \hat{\Phi})}{\partial \Omega^{-1}} &= \left(\frac{T}{2}\right) \frac{\partial \ln|\Omega^{-1}|}{\partial \Omega^{-1}} - \left(\frac{1}{2}\right) \sum_{t=1}^T \frac{\partial \hat{\epsilon}_t' \Omega^{-1} \hat{\epsilon}_t}{\partial \Omega^{-1}} \\ &= \left(\frac{T}{2}\right) \Omega' - \left(\frac{1}{2}\right) \sum_{t=1}^T \hat{\epsilon}_t \hat{\epsilon}_t' \end{aligned} \quad (1.31)$$

⁷ Considerando que para la derivación de matrices $\frac{\partial \mathbf{x}' \mathbf{A} \mathbf{x}}{\partial \mathbf{A}} = \mathbf{x} \mathbf{x}'$ y $\frac{\partial \ln |\mathbf{A}|}{\partial \mathbf{A}} = (\mathbf{A}')^{-1}$.

La función de verosimilitud es maximizada o minimizada cuando su derivada es cero, es decir:

$$\left(\frac{T}{2}\right)\mathbf{\Omega}' - \left(\frac{1}{2}\right)\sum_{t=1}^T \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t' = \mathbf{0}$$

Despejando $\mathbf{\Omega}'$:

$$\mathbf{\Omega}' = \left(\frac{1}{T}\right)\sum_{t=1}^T \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t' \tag{1.32}$$

La matriz $\mathbf{\Omega}$ que satisface (1.32) maximiza la función de verosimilitud entre la clase de todas las matrices ($n \times n$) sin restricciones. Se aprecia que de cualquier manera, el valor sin restricciones óptimo para $\mathbf{\Omega}$ que es especificado por (1.32) resulta ser el de una matriz simétrica y definida positiva. El EMV, es decir, el valor de $\mathbf{\Omega}$ que maximiza la función de log-verosimilitud entre la clase de todas las matrices simétricas definidas positivas, estará así también dada por (1.32):

$$\hat{\mathbf{\Omega}} = \left(\frac{1}{T}\right)\sum_{t=1}^T \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t' \tag{1.33}$$

El elemento en el renglón i y en la columna i de $\hat{\mathbf{\Omega}}$ es:

$$\hat{\sigma}_i^2 = \left(\frac{1}{T}\right)\sum_{t=1}^T \hat{\varepsilon}_{it}^2 \tag{1.34}$$

que es exactamente el promedio de los residuos muestrales de una regresión de la i -ésima variable en el VAR con un término constante y p rezagos de todas las variables. El elemento en el renglón i y la columna j de $\hat{\Omega}$ es:

$$\hat{\sigma}_{ij}^2 = \left(\frac{1}{T}\right) \sum_{t=1}^T \hat{\varepsilon}_{it}^2 \hat{\varepsilon}_{jt}^2 \quad (1.35)$$

el cual es el producto promedio de los residuos muestrales obtenidos por MCO tanto para la variable i , como para la variable j .

Por lo tanto, la estimación de Φ y Ω por máxima verosimilitud conduce a los estimadores que se obtienen por MCO para cada variable dentro del VAR. Una consecuencia muy importante de esto es que los estimadores son consistentes y asintóticamente eficientes⁸. De esta manera, aunado a la sencillez en su obtención, se tiene uno de los principales argumentos a favor de la utilización de la metodología VAR por parte de una cantidad importante de investigadores, y que le ha constituido como herramienta fundamental de la econometría moderna, a pesar de las críticas que también ha recibido por su carácter atóxico⁹.

1.3 LA FUNCIÓN IMPULSO—RESPUESTA

Una cuestión relevante en el planteamiento de la teoría VAR se refiere a lo que pasa cuando se presenta una perturbación en una innovación del VAR mismo y

⁸ La verificación de estas afirmaciones son presentados también por **Hamilton** (1994), Sección 11, pp. 298-302.

⁹ Como en el caso de **Cooley y LeRoy** (1985).

la clase de efecto que puede producir. Para ilustrar esto se puede considerar el siguiente sistema VAR(1) con dos variables, sin desplazamiento:

$$y_{1t} = a_{11}y_{1,t-1} + a_{12}y_{2,t-1} + \varepsilon_{1t}$$

$$y_{2t} = a_{21}y_{1,t-1} + a_{22}y_{2,t-1} + \varepsilon_{2t}$$

Una perturbación en ε_{1t} tiene un efecto uno-a-uno inmediato sobre y_{1t} , pero ninguno sobre y_{2t} . En el periodo $t + 1$, la perturbación en $y_{1,t}$ afecta a $y_{1,t+1}$ a través de la primera ecuación y asimismo afecta a $y_{2,t+1}$ a través de la segunda ecuación. Estos efectos son aplicados para el periodo $t + 2$, y así sucesivamente. De esta forma, una perturbación en una innovación en el VAR engendrará una reacción en cadena sobre el tiempo en todas las variables en el VAR. Son las Funciones Impulso–Respuesta las que calculan estas reacciones en cadena.

La discusión teórica de estas funciones revisten un complicado tratamiento revisado por Hamilton (1994), en el que considera el VAR reescribiéndolo como un vector de procesos de medias móviles MA(∞), y definiendo los impulsos a través de observar el comportamiento de las derivadas parciales de valores futuros de \mathbf{y} respecto a las innovaciones del modelo. Aquí, directamente a través del trabajo con innovaciones ortogonales, son revisadas las respuestas a las perturbaciones en las innovaciones originales del modelo.

1.3.1 Innovaciones ortogonales

Una objeción al procedimiento para el cómputo de funciones impulso–respuesta reside en que las innovaciones, en general, no son

contemporáneamente independientes de alguna otra innovación. El proponer que una de las innovaciones reciba una perturbación y la otra no es un tanto inverosímil. Una solución al problema ampliamente utilizada es transformar las innovaciones ε para producir un nuevo conjunto de innovaciones ortogonales. Al ser tomadas de dos en dos, éstas deberán estar no correlacionadas y tener varianzas unitarias. Sean las innovaciones ortogonales denotadas como el vector \mathbf{u} , y nuevamente sea tomado en cuenta el caso del VAR bivariado de primer orden, sin intercepto.

Sea $u_1 = b_{11}\varepsilon_1$, el requisito de una varianza muestral unitaria arroja que $b_{11} = 1/s_1$, donde s_1 es la desviación estándar muestral de ε_1 . A continuación se hace la regresión por mínimos cuadrados ordinarios de ε_2 sobre ε_1 para obtener el residuo $u_2^* = \varepsilon_2 - b_{21}\varepsilon_1$. Por su construcción este residuo no está correlacionado con ε_1 y por tanto tampoco lo está con u_1 . Sea denotado el error estándar de esta regresión por $s_{(21)}$, se sigue que $u_2 = u_2^*/s_{(21)}$ tendrá varianza unitaria y no estará correlacionada con u_1 . Las transformaciones pueden ser resumidas como

$$u_t = \mathbf{P}\varepsilon_t \quad \text{o} \quad \varepsilon_t = \mathbf{P}^{-1}u_t \quad (1.36)$$

donde $\mathbf{u}_t = [u_{1,t} \quad u_{2,t}]'$, $\varepsilon_t = [\varepsilon_{1,t} \quad \varepsilon_{2,t}]'$, y

$$\mathbf{P} = \begin{bmatrix} \frac{1}{s_1} & 0 \\ -\frac{b_{21}}{s_{(21)}} & \frac{1}{s_{(21)}} \end{bmatrix} \quad \text{dando} \quad \mathbf{P}^{-1} = \begin{bmatrix} s_1 & 0 \\ b_{21}s_1 & s_{(21)} \end{bmatrix} \quad (1.37)$$

La matriz de covarianzas muestral para las u es $\sum \frac{u_t u_t'}{n}$. De (1.36)

$$\frac{1}{n} \sum u_t u_t' = \mathbf{P} \left(\sum \varepsilon_t \varepsilon_t' \right) \mathbf{P}' = \mathbf{P} \hat{\Omega} \mathbf{P}'$$

pero la matriz de covarianzas muestral para las u es \mathbf{I} por construcción, y así

$$\hat{\mathbf{\Omega}} = \mathbf{P}^{-1}(\mathbf{P}^{-1})' \tag{1.38}$$

La ecuación (1.38) ilustra la factorización de Choleski de la matriz definida positiva $\hat{\mathbf{\Omega}}$. Se muestra como el producto de una matriz triangular inferior \mathbf{P}^{-1} y su transpuesta, la cual es triangular superior.

Las innovaciones ortogonalizadas fueron desarrolladas para tratar el problema de correlaciones no cero entre las innovaciones originales. Sin embargo, aparece un nuevo problema. El orden en el cual las variables ε son ortogonalizadas, puede tener efectos dramáticos sobre los resultados numéricos. La interpretación de funciones impulso–respuesta es de esta forma una operación un tanto arriesgada, y ha dado lugar a un intenso debate sobre su posible significancia económica. Sin embargo, su implementación en el análisis económico ha tenido un auge importante.

1.4 DESCOMPOSICIÓN DE VARIANZA

Los sistemas VAR tienen un importante uso en la producción de pronósticos para el corto plazo, aunque considerando que los VAR sin restricciones están sobreparametrizados, no son particularmente útiles para este propósito. No obstante, comprender las propiedades de los errores de pronóstico es muy útil para describir las interrelaciones entre las variables del sistema. En este aspecto el sistema VAR descansa en la proposición general de que las variables económicas tienden a desplazarse juntas en el tiempo y a ser autocorrelacionadas, lo que ofrece

una metodología predictiva con perspectivas atóricas desde el punto de vista económico.

Siguiendo a Johnston y Dinardo (1997) con el VAR(1) sin desplazamiento (nuevamente a modo de ilustración), se obtiene tanto a $\hat{\mathbf{y}}_{n+s}$ el vector de pronóstico óptimo de \mathbf{y} (con el mínimo error cuadrático medio) para el tiempo $n + s$ (es decir, s periodos hacia adelante), como a $\Sigma(s)$ la matriz de varianzas y covarianzas para los errores de predicción.

Siendo \mathbf{y}_t el VAR(1) mencionado arriba, entonces:

$$\mathbf{y}_t = \Phi \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \quad (1.39)$$

resulta que el vector de pronóstico óptimo de \mathbf{y}_{n+1} es la esperanza condicional de \mathbf{y}_{n+1} :

$$\hat{\mathbf{y}}_{n+1} = E[\mathbf{y}_{n+1} | \mathbf{y}_n, \dots, \mathbf{y}_1] = \Phi \mathbf{y}_n$$

Por un proceso reiterativo se tiene que:

$$\hat{\mathbf{y}}_{n+s} = \Phi^s \mathbf{y}_n \quad (1.40)$$

y la matriz $\Sigma(s)$ es:

$$\Sigma(s) = \boldsymbol{\Omega} + \Phi \boldsymbol{\Omega} \Phi' + \Phi^2 \boldsymbol{\Omega} (\Phi')^2 + \dots + \Phi^{s-1} \boldsymbol{\Omega} (\Phi')^{s-1} \quad (1.41)$$

Las fórmulas (1.40) y (1.41) sólo aplican a procesos de primer orden, $p = 1$ (aunque desde el principio no se restringió el número de variables en el vector \mathbf{y} , $2 \leq n$). Fórmulas similares pueden ser desarrolladas para procesos VAR de orden

superior. Es importante observar que las fórmulas están escritas en términos de las matrices verdaderas, así que ellas sólo toman en cuenta el error de las perturbaciones y no permiten la incertidumbre para los coeficientes.

En la práctica, generalmente las predicciones puntuales de \mathbf{y}_{n+s} y $\Sigma(s)$ son obtenidas a partir de las sustituciones de Φ por $\hat{\Phi}$, y Ω por $\hat{\Omega}$ en (1.40) o (1.41).

Entonces, para predicciones un periodo adelante, la matriz relevante es simplemente $Var[\boldsymbol{\varepsilon}] = \Omega$. Así $Var[\hat{\mathbf{y}}_{11}]$ está dada por el elemento superior izquierdo de Ω , y $Var[\hat{\mathbf{y}}_{21}]$ está dado por su elemento inferior derecho. Lo que se busca es expresar estas varianzas predictivas en términos de las varianzas de las innovaciones ortogonales. De la ecuación (1.36) se sigue que

$$\Omega = \mathbf{P}^{-1}Var[\mathbf{u}](\mathbf{P}^{-1})' = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} v_1 & 0 \\ 0 & v_2 \end{bmatrix} \begin{bmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{bmatrix} \quad (1.42)$$

donde c denota a los elementos de \mathbf{P}^{-1} y $v_i = Var[u_i]$ para $i = 1, 2$. Por construcción, cada \mathbf{u} tiene una varianza unitaria. Resolviendo la multiplicación en (1.42) resulta

$$Var[\hat{\mathbf{y}}_{11}] = c_{11}^2 v_1 + c_{12}^2 v_2 \quad \text{y} \quad Var[\hat{\mathbf{y}}_{21}] = c_{21}^2 v_1 + c_{22}^2 v_2$$

De (1.37), $c_{12} = 0$; y así, toda la varianza de $\hat{\mathbf{y}}_{11}$ se atribuye a la primera innovación ortogonal y es igual a c_{11}^2 . La varianza de $\hat{\mathbf{y}}_{21}$ es la suma de dos componentes, una proporción, $c_{21}^2/(c_{21}^2 + c_{22}^2)$, atribuida a la primera innovación ortogonal y la proporción, $c_{22}^2/(c_{21}^2 + c_{22}^2)$, asociada a la segunda innovación. Este resultado es la descomposición de la varianza del pronóstico.

Para pronósticos de dos ó más periodos en adelante se regresa a la matriz de varianzas y covarianzas de los errores de predicción dada en (1.41). Sustituyendo (1.38) en (1.41) es posible reescribir

$$\sum (s) = \mathbf{P}^{-1}(\mathbf{P}^{-1})' + (\mathbf{A}\mathbf{P}^{-1})(\mathbf{A}\mathbf{P}^{-1})' + \dots + (\mathbf{A}^{s-1}\mathbf{P}^{-1})(\mathbf{A}^{s-1}\mathbf{P}^{-1})' \quad (1.43)$$

Cálculos similares a los ya descritos son hechos para el número relevante de productos matriciales en (1.43), y las descomposiciones de la varianza son obtenidas.

La descomposición de la varianza del error de pronóstico muestra entonces la proporción de los movimientos en una sucesión debidos a sus propios *shocks* contra *shocks* en la otra variable. Si los *shocks* ε_{2t} no explican la varianza del error de pronóstico de $\{y_{1t}\}$ en todos los horizontes de pronóstico, se dice que la sucesión $\{y_{1t}\}$ es exógena. En esta circunstancia, $\{y_{1t}\}$ evoluciona independientemente de los *shocks* ε_{2t} y de la sucesión $\{y_{2t}\}$. En el otro extremo, los *shocks* ε_{2t} explican toda la varianza del error de pronóstico de $\{y_{1t}\}$ en todos los horizontes de pronóstico, y se dice que la sucesión $\{y_{1t}\}$ es endógena. En investigación aplicada es típico que una variable explique casi toda su varianza de error de pronóstico en horizontes cortos y pequeñas proporciones en horizontes largos. Este patrón se podría esperar si los *shocks* ε_{2t} tienen un efecto contemporáneo pequeño sobre y_{1t} pero actúan para afectar a la sucesión $\{y_{1t}\}$ con un rezago.

Como en el caso de las funciones de impulso–respuesta, los valores numéricos de las descomposiciones de la varianza son a menudo muy sensibles al orden para el cual las innovaciones originales fueron ortogonalizadas. Así, las precauciones ya indicadas para las funciones de impulso–respuesta aplican con el mismo rigor a las descomposiciones de varianza.

1.5 PRUEBAS DE HIPÓTESIS

1.5.1 Pruebas para el orden del VAR

En principio es posible incorporar un número grande de variables en un VAR, tal como fue definido en (1.11), aunque hay que tomar en cuenta que, en investigación aplicada, los grados de libertad se pierden rápidamente cuanto más variables se incluyen.

Además de las variables a incluir en el VAR, se tiene que la elección del orden del VAR, es decir, la cantidad de rezagos con el que cuenta, se realiza frecuentemente de manera arbitraria, con la recomendación general de colocar suficientes rezagos como para asegurarse de que los residuos se comporten como ruido blanco gaussiano. Sin embargo, si se elige una cantidad excesiva de rezagos, las estimaciones se tornan imprecisas. Es por ello que tradicionalmente se utiliza la denominada prueba de razón de verosimilitud para encontrar la longitud apropiada.¹⁰

En la búsqueda del estadístico de prueba de razón de verosimilitud LR (*likelihood ratio*), es necesario calcular el máximo valor alcanzado por (1.30)¹¹. Se toma entonces la función de log-verosimilitud evaluada en $\hat{\Omega}$ y $\hat{\Phi}$:

¹⁰ La función de verosimilitud permite comparar modelos, como cuando uno de ellos incluye una variable adicional con respecto al otro. Las diferencias en la función de verosimilitud se alteran arbitrariamente con la escala medida, por lo que la forma de compararlas adecuadamente es mediante un cociente. De ahí que cuando se comparan modelos que han sido estimados mediante este procedimiento se hable de la razón de verosimilitud.

¹¹ Hamilton (1994).

$$L(\hat{\Omega}, \hat{\Phi}) = -\left(\frac{Tn}{2}\right) \ln(2\pi) + \left(\frac{T}{2}\right) \ln|\hat{\Omega}^{-1}| - \left(\frac{1}{2}\right) \sum_{t=1}^T \hat{\varepsilon}_t' \hat{\Omega}^{-1} \hat{\varepsilon}_t \quad (1.44)$$

para una $\hat{\Omega}$ dada por (1.33). Ahora, el último término en (1.44) se reduce como sigue:

$$\begin{aligned} \left(\frac{1}{2}\right) \sum_{t=1}^T \hat{\varepsilon}_t' \hat{\Omega}^{-1} \hat{\varepsilon}_t &= \left(\frac{1}{2}\right) Tr \left[\sum_{t=1}^T \hat{\varepsilon}_t' \hat{\Omega}^{-1} \hat{\varepsilon}_t \right] = \left(\frac{1}{2}\right) Tr \left[\sum_{t=1}^T \hat{\Omega}^{-1} \hat{\varepsilon}_t \hat{\varepsilon}_t' \right] = \\ &= \left(\frac{1}{2}\right) Tr \left[\hat{\Omega}^{-1} (T\hat{\Omega}) \right] = \left(\frac{1}{2}\right) Tr[(T\mathbf{I}_n)] = Tn/2 \end{aligned}$$

Sustituyendo lo anterior en (1.44) resulta:

$$L(\hat{\Omega}, \hat{\Phi}) = -\left(\frac{Tn}{2}\right) \ln(2\pi) + \left(\frac{T}{2}\right) \ln|\hat{\Omega}^{-1}| - \left(\frac{Tn}{2}\right) \quad (1.45)$$

Esto hace que las pruebas para la razón de verosimilitud sean particularmente sencillas al momento de su aplicación.

Supóngase que se quiere probar la hipótesis nula de que un conjunto de variables fue generado a partir de un VAR gaussiano con p_0 rezagos contra la hipótesis alternativa de que fue con p_1 rezagos ($p_1 > p_0$). Es decir, con $1, 2, \dots, p_0, p_0 + 1, \dots, p_1$;

$$H_0: p_0 = p_1 \quad \text{vs.}$$

$$H_a: p_0 < p_1$$

Para estimar el sistema bajo H_0 , se realiza un conjunto de n regresiones por MCO de cada variable en el sistema sobre un término constante y sobre p_0 rezagos

de todas las variables en el sistema. Sea $\widehat{\Omega}_0 = \left(\frac{1}{T}\right) \sum_{t=1}^T \widehat{\varepsilon}_t(p_0)[\widehat{\varepsilon}_t(p_0)]'$ la matriz de varianzas y covarianzas de los residuos para estas regresiones.

El valor máximo para la log-verosimilitud bajo H_0 es entonces

$$L_0^* = -\left(\frac{Tn}{2}\right) \ln(2\pi) + \left(\frac{T}{2}\right) \ln |\widehat{\Omega}_0^{-1}| - \left(\frac{Tn}{2}\right)$$

Del mismo modo, el sistema es estimado bajo H_a por regresiones de MCO que incluyan p_1 rezagos de todas las variables. La log-verosimilitud maximizada bajo H_a resulta

$$L_1^* = -\left(\frac{Tn}{2}\right) \ln(2\pi) + \left(\frac{T}{2}\right) \ln |\widehat{\Omega}_1^{-1}| - \left(\frac{Tn}{2}\right)$$

donde $\widehat{\Omega}_1$ es la matriz de varianzas y covarianzas de los residuos para este segundo conjunto de regresiones. De duplicar la razón de log-verosimilitud se obtiene el estadístico de prueba LR :

$$\begin{aligned} LR &= 2(L_1^* - L_0^*) = 2 \left\{ \left(\frac{T}{2}\right) \ln |\widehat{\Omega}_1^{-1}| - \left(\frac{T}{2}\right) \ln |\widehat{\Omega}_0^{-1}| \right\} \\ &= T \ln(1/|\widehat{\Omega}_1|) - T \ln(1/|\widehat{\Omega}_0|) \\ &= -T \ln(|\widehat{\Omega}_1|) + T \ln(|\widehat{\Omega}_0|) \\ &= T \{ \ln(|\widehat{\Omega}_0|) - \ln(|\widehat{\Omega}_1|) \} \end{aligned} \tag{1.46}$$

Bajo la hipótesis nula, se tiene que LR se distribuye asintóticamente como una χ^2 con un número de grados de libertad igual al número de restricciones impuestas bajo H_0 . Cada ecuación en la especificación restringida por H_0 tiene

$(p_1 - p_0)$ menos rezagos en cada una de las n variables comparada con H_a ; así, H_0 impone $n(p_1 - p_0)$ restricciones sobre cada ecuación. Como existen n de tales ecuaciones, H_0 impone $n^2(p_1 - p_0)$ restricciones. De modo que el estadístico calculado en (1.46) se distribuye asintóticamente como una χ^2 con $n^2(p_1 - p_0)$ grados de libertad.

En la práctica, la regla de decisión para hallar el número de rezagos óptimo se observa al comparar el estadístico LR al nivel de confianza deseado (que generalmente es al 5%) comenzando por el valor de rezago más alto, y disminuyendo de uno en uno el orden hasta que la hipótesis nula sea rechazada, al obtener un valor de la probabilidad asociada al estadístico LR menor al valor crítico. Así se habrá llegado al valor óptimo de p (es decir p_0), propuesto en la hipótesis alternativa correspondiente.

Sims (1980) sugirió una modificación para la prueba de razón de verosimilitud tomando en cuenta muestras menores para los rezagos. Recomendó reemplazar en (1.46) el valor de T por el de $(T - k)$, donde $k = 1 + np_1$ es el número de parámetros estimados por ecuación.

La prueba así ajustada tiene asintóticamente la misma distribución que (1.46), pero el ajuste de Sims es menos proclive al rechazo de H_0 para muestras pequeñas.

1.5.2 Prueba para la causalidad de Granger

Una de las preguntas claves que acompañan al VAR es cuán útiles son algunas variables para pronosticar otras. A continuación se hace la presentación

particular de la relación predictiva entre dos variables propuesta por Granger (1969) y popularizada por Sims (1972).

A partir de la formulación general del VAR, los valores rezagados de cada variable aparecen en cada ecuación del mismo. Una prueba de causalidad es para saber si los rezagos de una variable entrarán en la ecuación para otra variable.

Considerando un modelo de dos variables con p rezagos, $\{y_{1t}\}$ no causa en el sentido de Granger a $\{y_{2t}\}$ si y sólo si todos los coeficientes de las entradas ϕ_{21} en las matrices Φ_i son todas cero para $i = 1, \dots, p$. Por lo tanto, si $\{y_{1t}\}$ no mejora el comportamiento del pronóstico de $\{y_{2t}\}$, entonces $\{y_{1t}\}$ no causa en el sentido de Granger a $\{y_{2t}\}$. Si todas las variables en el VAR son estacionarias, la forma directa de probar causalidad de Granger es usar una prueba F estándar de la hipótesis nula

$$H_0: \phi_{21}^{(1)} = \phi_{21}^{(2)} = \dots = \phi_{21}^{(p)} = 0 \quad \text{vs.}$$

$$H_a: \phi_{21}^{(i)} \neq 0 \text{ para alguna } i = 1, \dots, p.$$

donde H_0 establece que $\{y_{2t}\}$ no causa en el sentido de Granger a $\{y_{1t}\}$.

De manera directa se puede generalizar esta noción al caso de n variables de (1.11). Si todos los coeficientes $\phi_{jl}^{(i)}$ en el renglón j y en la columna l de las correspondientes matrices Φ_i se pueden igualar a cero, entonces al momento de observar la correspondiente ecuación del VAR para la variable j , se verá que no existe el polinomio para la variable l , por lo que la variable l no causa a la variable j en el sentido de Granger.

Se nota que la causalidad de Granger es muy distinta de una prueba de exogeneidad, vinculada al error de pronóstico. Para que $\{y_{2t}\}$ sea exógena, se requeriría que no sea afectada por el valor contemporáneo de $\{y_{1t}\}$. Sin embargo, la

causalidad de Granger se refiere sólo al efecto de valores pasados de $\{y_{1t}\}$ sobre el valor actual de $\{y_{2t}\}$. Por lo tanto, la causalidad de Granger, mide si los valores pasados y actuales de $\{y_{1t}\}$ ayudan a pronosticar los valores futuros de $\{y_{2t}\}$.

Sea entonces que tomando el primer renglón del sistema, el valor predictivo óptimo para el próximo periodo de y_{1t} depende solamente de sus propios valores retardados y no de los de y_{2t} (entonces, la restricción es tal que los coeficientes $\phi_{12}^{(i)} = 0$ para todo i).

$$\begin{aligned} & \hat{E}(y_{1t,t+1} | y_{1t,t}, y_{1t,t-1}, \dots, y_{2t,t}, y_{2t,t-1}, \dots) \\ &= k_1 + \phi_{11}^{(1)} y_{1t,t} + \phi_{11}^{(2)} y_{1t,t-1} + \dots + \phi_{11}^{(p)} y_{1t,t-p+1} \end{aligned} \tag{1.47}$$

La manera de implementar la prueba F para probar causalidad es calcular la suma de los residuos cuadrados (SRC) de (1.47),

$$SRC_1 = \sum_{t=1}^T \hat{\varepsilon}_{1t}^2$$

y comparar esto (a través de un cociente) con la suma de los residuos cuadrados de una regresión univariada para y_{1t} ,

$$SRC_0 = \sum_{t=1}^T \hat{e}_t^2$$

donde

$$y_{1t} = c_0 + \gamma_1 y_{1t,t-1} + \gamma_2 y_{1t,t-2} + \dots + \gamma_p y_{1t,t-p} + e_t \tag{1.48}$$

es estimado también por MCO. Se tiene así el estadístico de prueba:

$$S_1 \equiv \frac{(SRC_0 - SRC_1)/p}{SRC_1/(T - 2p - 1)} \quad (1.49)$$

Si S_1 es mayor al valor crítico para el 5% (el más común) para una distribución $F(p, T - 2p - 1)$, entonces es rechazada la hipótesis nula de que y_{2t} no causa en el sentido de Granger a y_{1t} ; esto es, si S_1 es suficientemente grande, se concluye que y_{2t} causa a y_{1t} en el sentido de Granger.

El estadístico de prueba dado en (1.49) podría tener una distribución exacta F para una regresión con regresores ajustados y perturbaciones gaussianas. Tanto con variables dependientes rezagadas como en las regresiones para la causalidad de Granger, la prueba es válida sólo asintóticamente.

Un contraste asintóticamente equivalente para probar si y_{2t} causa en el sentido de Granger a y_{1t} , con las mismas hipótesis de prueba

$$H_0: \phi_{21}^{(1)} = \phi_{21}^{(2)} = \dots = \phi_{21}^{(p)} = 0 \quad \text{vs.}$$

$$H_a: \phi_{21}^{(i)} \neq 0 \text{ para alguna } i = 1, \dots, p.$$

tendrá por estadístico de prueba:

$$S_2 \equiv \frac{T(SRC_0 - SRC_1)}{SRC_1} \quad (1.50)$$

con el cual se rechazaría la hipótesis nula de que y_{2t} no causa en el sentido de Granger a y_{1t} si S_2 es mayor al valor crítico al 5% para una variable χ^2 con p grados de libertad.

CAPÍTULO 2

Cointegración y modelos de corrección de error

2.1 INTRODUCCIÓN

Fue Granger (1981) quien propuso la relación entre cointegración y corrección de error, y posteriormente Engle y Granger (1987), extendieron y usaron esta relación para desarrollar procedimientos de estimación, pruebas y ejemplos empíricos. Cuando cada elemento y_{jt} de un vector de series de tiempo \mathbf{y}_t no es estacionario pero su combinación lineal es estacionaria, entonces las series de tiempo $\boldsymbol{\beta}'\mathbf{y}_t$ están cointegradas con vector de cointegración $\boldsymbol{\beta}' = (\beta_1, \beta_2, \dots, \beta_n)$. Debido a que pueden existir varios vectores de cointegración, $\boldsymbol{\beta}'$ en realidad puede ser una matriz.

Interpretando $\boldsymbol{\beta}'\mathbf{y}_t = 0$ como un equilibrio de largo plazo, la cointegración implica que las desviaciones de este equilibrio son estacionarias con varianza finita aún cuando las series mismas no sean estacionarias y tengan varianza infinita.

A lo largo de este capítulo han de presentarse los fundamentos de la cointegración, a partir de la definición de integración y a partir de ella la de la cointegración misma, pasando por los modelos de corrección de error tanto para el caso bivariado como el de n variables. Y para cerrar se presenta la metodología para

probar cointegración, propuesta por Johansen inicialmente en 1988, y aplicada sobre los modelos propuestos en la segunda parte de este estudio.

2.2 COMBINACIONES LINEALES DE VARIABLES INTEGRADAS

Una variable económica considerada como una serie de tiempo puede deambular ampliamente y, no obstante, se puede esperar que algunos pares de la serie se muevan de modo tal que no se alejen mucho. Típicamente, la teoría económica propone fuerzas que tienden a mantener a las series juntas. Ejemplos pueden ser las tasas de interés de corto y de largo plazo, ingreso familiar y gasto, precios del mismo bien en distintos mercados. Una idea parecida surge de la consideración de relaciones de equilibrio de largo plazo, donde el equilibrio es un punto estacionario caracterizado por fuerzas que tienden a regresar a la economía a su nivel de equilibrio siempre que ella se aleje de él. Si \mathbf{y}_t es un vector de variables económicas, se dice que están en equilibrio cuando se cumple con la restricción lineal específica

$$\boldsymbol{\beta}'\mathbf{y}_t = \mathbf{0}$$

En la mayoría de los periodos de tiempo, \mathbf{y}_t no está en equilibrio y la cantidad univariada

$$\mathbf{z}_t = \boldsymbol{\beta}'\mathbf{y}_t$$

se puede llamar error de equilibrio. El concepto de equilibrio¹² tiene relevancia en la especificación de modelos econométricos, cuando la economía tiene pequeños valores de \mathbf{z}_t más que valores grandes, es decir, cuando \mathbf{z}_t es estacionario.

Ahora, se da una base a estas ideas y se muestra que una clase de modelos, llamada de corrección de error, permite que los componentes de largo plazo obedezcan restricciones de corto plazo, mientras que los componentes de corto plazo tienen una especificación dinámica flexible. Un condición para que esto sea cierto, llamada cointegración, fue propuesta por Granger (1981) y por Granger y Weiss (1983) y se define en la próxima sección.

2.3 INTEGRACIÓN, COINTEGRACIÓN Y CORRECCIÓN DE ERROR

El teorema de descomposición de Wold muestra que una serie de tiempo estacionaria con componentes no deterministas tiene una representación $MA(\infty)$ que es generalmente aproximada por un proceso $ARMA(p, q)$ ¹³. Comúnmente, las series económicas deben diferenciarse antes de que la hipótesis de estacionariedad se pueda asumir. De acuerdo a lo anterior, Engle y Granger (1987) dan la siguiente definición de integración:

¹² Conviene decir que el término *desequilibrio* se usa de manera distinta por teóricos de la economía y econométricos. Mientras los primeros usan el término para referirse a una igualdad entre transacciones deseadas y reales, los segundos lo usan para referirse a cualquier relación de largo plazo entre variables estacionarias. La cointegración no requiere que la relación de largo plazo sea generada por fuerzas de mercado o por las reglas de comportamiento de los individuos. En Engle y Granger, la relación de equilibrio puede ser causal, de comportamiento o una relación en forma reducida entre variables con tendencia similar.

¹³ A este respecto puede consultarse **Box y Jenkins** (1976) o **Granger y Newbold** (1977).

Definición. Una serie sin componente determinista que tiene una representación invertible y estacionaria ARMA después de diferenciarla d veces, se dice que es integrada de orden d y se representa por $y_t \sim I(d)$.

En gran parte de lo que se presenta en esta segunda parte del marco teórico sólo serán considerados los casos $d = 0$ y $d = 1$. Para $d = 0$, y_t será estacionaria (en niveles); y para $d = 1$, el cambio de y_t será estacionario (en primeras diferencias). Hay diferencias importantes en apariencia entre series $I(0)$ y series $I(1)$:

1. Si $y_t \sim I(0)$, con media cero, entonces

- (i) la varianza de y_t es finita,
- (ii) una innovación tiene sólo un efecto temporal en el valor de y_t ,
- (iii) el espectro de y_t , $f(\omega)$ tiene la propiedad $0 < f(0) < \infty$,
- (iv) la cantidad esperada de veces que y_t cruza la línea $y = 0$ es finita
y
- (v) las autocorrelaciones $\rho(k)$ decrecen constantemente en magnitud para k suficientemente grandes, de manera que su suma es finita,

2. Si $y_t \sim I(1)$, con $y_0 = 0$, entonces,

- (i) la varianza de y_t tiende a infinito cuando t tiende a infinito,
- (ii) una innovación tiene un efecto permanente en el valor de y_t , ya que y_t es la suma de todos los cambios previos,

- (iii) el espectro de y_t , tiene aproximadamente la forma $f(\omega) \sim A\omega^{-2d}$ para ω pequeños de modo que en particular $f(0) = \infty$ tiene la propiedad $0 < f(0) < \infty$,
- (iv) la cantidad esperada de veces que y_t cruza la línea $y = 0$ es infinita y
- (v) las autocorrelaciones $\rho(k)$ tienden a uno para todo k cuando $t \rightarrow \infty$.

La varianza infinita para una serie $I(1)$ proviene completamente de la contribución de las bajas frecuencias, o parte de largo plazo de la serie. Así, una serie $I(1)$ es más suavizada teniendo largas ondas dominantes, en comparación con una serie $I(0)$. Debido a los tamaños relativos de las varianzas, siempre es cierto que la suma de una serie $I(0)$ con una serie $I(1)$ será $I(1)$. Además, si a y b son constantes, con $b \neq 0$ y si $y_t \sim I(d)$, entonces $a + by_t \sim I(d)$

Si dos series temporales x_t y y_t son $I(d)$, entonces es generalmente cierto que la combinación lineal $z_t = x_t - ay_t$ también es $I(d)$. Sin embargo, también es posible que $z_t \sim I(d - b)$, $b > 0$. Cuando esto sucede, opera una restricción especial para los componentes de largo plazo de la serie. Considerar el caso $d = b = 1$, de modo que x_t y y_t son $I(1)$, con componentes dominantes de largo plazo, pero z_t es $I(0)$ sin especialmente fuertes bajas frecuencias. La constante a es entonces tal que la mayor parte de los componentes de largo plazo de x_t y y_t se anulan. Para $a = 1$, la vaga idea de que x_t y y_t no pueden separarse se ha traducido en la más precisa proposición de que “su diferencia será $I(0)$ ”. El uso de la constante a únicamente sugiere que alguna ampliación necesita usarse antes de que la diferencia $I(0)$ se pueda lograr. Notar que generalmente no es cierto que existe una a que hace que $z_t \sim I(0)$.

Un caso parecido, considerando una distinta frecuencia importante, es cuando x_t y y_t son un par de series cada una con un componente estacional importante, y aún existe una a de modo que la serie derivada z_t no tiene estacionariedad. Aunque esto puede ocurrir se puede considerar una posibilidad.

Para formalizar estas ideas, en Engle y Granger (1987) se tiene la siguiente:

Definición: los componentes del vector \mathbf{y}_t son cointegrados de orden d, b , representado por $CI(d, b)$ si:

- (i) todos los componentes de \mathbf{y}_t son integrados de orden d , y
- (ii) existe un vector $\boldsymbol{\beta}$ ($\neq 0$) tal que la combinación lineal $\mathbf{z}_t = \boldsymbol{\beta}'\mathbf{y}_t \sim I(d - b)$, $b > 0$. El vector $\boldsymbol{\beta}$ se llama vector de cointegración.

En el caso particular de que $d = 1$ y $b = 1$ la cointegración significa que si todos los componentes de \mathbf{y}_t son $I(1)$, entonces el error de equilibrio \mathbf{z}_t será $I(0)$, raramente se alejará de cero si tiene media cero y cruzará con frecuencia la línea cero. Dicho esto de otra manera, la cointegración significa que el equilibrio ocasionalmente ocurrirá, al menos en una aproximación cercana, mientras si \mathbf{y}_t no está cointegrada, entonces \mathbf{z}_t puede divagar ampliamente y los cruces de la línea cero, serán muy raros, sugiriendo en estos casos, que el concepto de equilibrio no tiene implicaciones prácticas.

La reducción del orden de integración implica un tipo especial de relación con consecuencias interpretables y verificables. Si, no obstante, todos los elementos de \mathbf{y}_t son estacionarios, de modo que ellos son $I(0)$, entonces el error de equilibrio \mathbf{z}_t no tiene una propiedad distintiva si éste es $I(0)$. Puede ser que $\mathbf{z}_t \sim I(1)$, de modo que su espectro es cero en frecuencia cero, pero si cualquiera de las variables error de medición, esta propiedad generalmente no se observa y así este caso es en la realidad de poco interés. Al interpretar el concepto de cointegración, en el caso de

que $n = 2$, $d = b = 1$, Granger y Weiss (1983), muestran que una condición necesaria y suficiente para la cointegración es que la coherencia entre las dos series está en frecuencia cero.

Si \mathbf{y}_t tiene n componentes, entonces existe más de un vector de cointegración $\boldsymbol{\beta}$. Es claramente posible para algunas relaciones de equilibrio gobernar el comportamiento conjunto de las variables. En lo que sigue, se asume que existen exactamente r vectores de cointegración linealmente independientes, con $r \leq n - 1$ que se agrupan en el arreglo $\boldsymbol{\beta}$. Por construcción el rango de $\boldsymbol{\beta}$ será r que se llama “rango de cointegración” de \mathbf{y}_t .

En resumen, se tienen los siguiente resultados:

1. La cointegración se refiere a una combinación lineal de variables no estacionarias.
2. El vector de cointegración no es único. Si $\boldsymbol{\beta}$ es un vector de cointegración, entonces para todo valor de λ , $\lambda\boldsymbol{\beta}$ es también un vector de cointegración. Usualmente, se usa una de las variables para normalizar $\boldsymbol{\beta}$ haciendo sus coeficientes iguales a uno. Para normalizar $\boldsymbol{\beta}$ con respecto a y_{1t} , se elige $\lambda = 1/\beta_1$.
3. La definición de cointegración se refiere a variables que son integradas del mismo orden. Desde luego, esto no implica que todas las variables integradas estén cointegradas; generalmente, un conjunto de variables $I(d)$ no están cointegradas. Tal falta de cointegración implica que no existe equilibrio de largo plazo entre las variables de modo que éstas pueden vagar arbitrariamente alejándose unas de otras. Si dos variables son integradas de distintos órdenes, ellas no pueden estar cointegradas. Pero es posible encontrar relaciones de equilibrio entre grupos de variables que son integradas de distintos órdenes. Por ejemplo, si y_{1t} y y_{2t} son $I(2)$ y otra

variable y_{3t} es $I(1)$, no puede haber cointegración entre y_{1t} (o y_{2t}) y y_{3t} . Pero si y_{1t} y y_{2t} son $CI(2,1)$, existe una combinación lineal $\beta_1 y_{1t} + \beta_2 y_{2t}$ que es $I(1)$. Es posible que esta combinación lineal sea cointegrada con la variable $I(1)$. Se llama a este caso multicointegración, y es tratado por Granger y Lee (1990).

4. Si \mathbf{y}_t tiene n componentes no estacionarios, pueden existir $n - 1$ vectores de cointegración linealmente independientes. Por ejemplo, si \mathbf{y}_t es de dos variables, habrá al menos un vector de cointegración linealmente independiente. El número de vectores de cointegración se llama rango de cointegración de \mathbf{y}_t .
5. Mucha de la literatura sobre cointegración se enfoca en el caso donde cada variable contiene una sola raíz unitaria. La razón es que el análisis tradicional de regresión o series de tiempo se aplica cuando las variables son $I(0)$ y pocas variables económicas son integradas de un orden mayor a uno. Cuando esto es ambiguo, muchos autores usan el término cointegración para referir el caso en donde las variables son $CI(1,1)$. En el resto de este capítulo se usa esta convención. Desde luego, pueden surgir otras posibilidades. Por ejemplo, un conjunto de variables $I(2)$ pueden ser cointegradas $CI(2,1)$, de modo que existirá una combinación lineal que es $I(1)$.

2.4 COINTEGRACIÓN Y TENDENCIAS COMUNES

Una manera de entender las relaciones de cointegración es mediante la observación de Stock y Watson (1988) de que las variables cointegradas comparten tendencias estocásticas. Por ejemplo, sea \mathbf{y} un vector formado por dos variables $\mathbf{y}_t = (x_t, z_t)'$. Ignorando términos cíclicos y estacionales, cada variable de \mathbf{y}_t se

puede representar como una caminata aleatoria más un componente irregular (no necesariamente ruido blanco) en la forma:

$$x_t = \mu_{xt} + \varepsilon_{xt} \tag{2.1}$$

$$z_t = \mu_{zt} + \varepsilon_{zt} \tag{2.2}$$

donde μ_{xt} y μ_{zt} son procesos de caminata aleatoria que representan la tendencia estocástica de las variables x_t y z_t , respectivamente y ε_{xt} , ε_{zt} son componentes irregulares (estacionarios) de las variables x_t y z_t , respectivamente.

Las sucesiones $\{x_t\}$ y $\{z_t\}$ son $CI(1,1)$, si existen valores β_1 y β_2 distintos de cero para los cuales la combinación lineal $\beta_1 x_t + \beta_2 z_t$ es estacionaria. En efecto, sustituyendo en esta expresión x_t y z_t , se tiene

$$\begin{aligned} \beta_1 x_t + \beta_2 z_t &= \beta_1 (\mu_{xt} + \varepsilon_{xt}) + \beta_2 (\mu_{zt} + \varepsilon_{zt}) \\ \beta_1 x_t + \beta_2 z_t &= (\beta_1 \mu_{xt} + \beta_2 \mu_{zt}) + (\beta_1 \varepsilon_{xt} + \beta_2 \varepsilon_{zt}) \end{aligned} \tag{2.3}$$

Dado que el término $(\beta_1 \varepsilon_{xt} + \beta_2 \varepsilon_{zt})$ es estacionario, la condición necesaria y suficiente para que $\{x_t\}$ y $\{z_t\}$ sean $CI(1,1)$ es

$$\beta_1 \mu_{xt} + \beta_2 \mu_{zt} = 0 \tag{2.4}$$

es decir, que desaparezcan las tendencias estocásticas μ_{xt} y μ_{zt} , cuyos valores observados cambian continuamente en el tiempo. Dado que β_1 y β_2 no pueden ser simultáneamente iguales a cero, entonces (2.4) se cumple para todo t si y sólo si

$$\mu_{x_t} = \frac{-\beta_2 \mu_{z_t}}{\beta_1}$$

Para $\beta_1 \neq 0$ y $\beta_2 \neq 0$, la igualdad se cumple siempre y cuando las tendencias estocásticas sean idénticas a un escalar. Para el escalar $-\beta_2/\beta_1$, los procesos estocásticos $I(1)$, $\{x_t\}$ y $\{z_t\}$, tienen la misma tendencia estocástica si y sólo si son $CI(1,1)$.

En resumen, dos variables son cointegradas siempre que la tendencia de una variable sea una combinación lineal de las tendencias en la otra(s) variable(s). En este caso, es posible hallar un vector β tal que la combinación lineal $\beta_1 x_t + \beta_2 z_t + \beta_3 w_t$ no tenga tendencia.

Este resultado se puede generalizar al caso de n variables

$$\mathbf{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t \tag{2.5}$$

donde $\mathbf{y}_t = (y_{1t}, y_{2t}, \dots, y_{nt})'$, $\boldsymbol{\mu}_t$ es el vector de tendencias estocásticas $(\mu_{1t}, \mu_{2t}, \dots, \mu_{nt})'$ y $\boldsymbol{\varepsilon}_t$ es un vector $\boldsymbol{\varepsilon}_t (n \times 1)$ de componentes estacionarios.

Si una tendencia se puede expresar como una combinación lineal de las otras tendencias en el sistema, significa que existe un vector $\beta = \beta_1 \mu_{1t} + \beta_2 \mu_{2t} + \dots + \beta_n \mu_{nt}$ tal que $\beta = \mathbf{0}$.

Premultiplicando (2.5) por β se tiene

$$\beta \mathbf{y}_t = \beta \boldsymbol{\mu}_t + \beta \boldsymbol{\varepsilon}_t$$

Ya que $\beta\mu_t = \mathbf{0}$, entonces $\beta\mathbf{y}_t = \beta\boldsymbol{\varepsilon}_t$ y la combinación lineal $\beta\mathbf{y}_t$ es estacionaria. Este argumento se puede generalizar al caso de múltiples vectores de cointegración.

2.5 COINTEGRACIÓN Y CORRECCIÓN DE ERROR

El acercamiento inicial a la corrección de error será en forma heurística, haciendo referencias a tasas de interés.

Una característica importante de las variables cointegradas es que sus trayectorias son influidas por la magnitud de la desviación del equilibrio de largo plazo. Después de todo, si después de algún *shock* aleatorio que aleje al sistema de su equilibrio de largo plazo, éste retorna a este equilibrio, significa que los movimientos de al menos una de las variables del sistema responde a la magnitud del desequilibrio. Un ejemplo. Son las teorías de la estructura de los términos de la tasa de interés, las cuales establecen una relación de largo plazo entre las tasas de interés de largo y de corto plazo. Si la desviación de la tasa de corto plazo es grande respecto a la de largo plazo, la primera aumentará en relación a la segunda. Tal desviación puede reducirse mediante:

1. un incremento de la tasa de corto plazo y/o una disminución en la tasa de largo plazo,
2. un aumento en la tasa de largo plazo, pero un aumento proporcionalmente mayor en la tasa de corto plazo o
3. una disminución en la tasa de largo plazo pero una menor disminución en la tasa de corto plazo.

Sin disponer de una especificación dinámica del modelo, no es posible determinar cuál será la forma de reducir la desviación. Se puede decir, no obstante, que la dinámica de la tasa de interés de corto plazo será influida por su desviación de la tasa de interés de largo plazo. Este mecanismo se llama mecanismo de corrección de error, en el cual la dinámica de corto plazo de la tasa de interés depende de la desviación de la tasa de equilibrio de largo plazo. Asumiendo que las tasas de interés de corto y largo plazo son $I(1)$, un modelo de corrección de error (MCE) para la estructura del término de la tasa de interés puede tener la siguiente formulación:

$$\Delta r_{st} = \alpha_s(r_{Lt-1} - \beta r_{st-1}) + \varepsilon_{st} \quad \alpha_s > 0 \quad (2.6)$$

$$\Delta r_{Lt} = -\alpha_L(r_{Lt-1} - \beta r_{st-1}) + \varepsilon_{Lt} \quad \alpha_L > 0 \quad (2.7)$$

donde ε_{st} y ε_{Lt} son procesos de ruido blanco gaussiano que pueden estar correlacionados, r_{Lt} y r_{st} son las tasas de interés de largo y corto plazo, respectivamente y α_s , α_L y β son parámetros. En (2.6) y (2.7), las tasas de interés de corto y largo plazo cambian con *shocks* estocásticos (ε_{st} y ε_{Lt}) y cuando en periodos anteriores hay desviaciones del equilibrio de largo plazo ($r_{Lt-1} - \beta r_{st-1} > 0$). Si $(r_{Lt-1} - \beta r_{st-1}) > 0$, la tasa de interés de corto plazo aumentará y la de largo plazo disminuirá. En equilibrio de largo plazo, $r_{Lt} = \beta r_{st}$.

En este ejemplo, se puede percibir la relación entre MCE y variables cointegradas. En efecto, ya que por definición Δr_{st} es estacionaria, $\Delta r_{st} \sim I(0)$. Y Δr_{st} es $I(0)$ si el término derecho de (2.6) es $I(0)$. Como ε_{st} es estacionario, el término derecho en (2.6) es $I(0)$ si la combinación lineal $r_{Lt-1} - \beta r_{st-1}$ es

estacionaria (lo es si las tasas de ahorro son *CI*) Por lo tanto, la tasas de interés de corto y largo plazo deben ser cointegradas con el vector de cointegración $(1, -\beta)$.

Un argumento idéntico se aplica a la expresión (2.7). El punto esencial es que la representación de la corrección de error requiere que las dos variables sean *CI* $(1, 1)$. Este resultado se puede generalizar al modelo más general que incorpora al modelo los cambios rezagados de cada tasa de interés en las ecuaciones (2.6) y (2.7):

$$\Delta r_{st} = a_{10} + \alpha_s(r_{Lt-1} - \beta r_{st-1}) + \sum a_{11}^{(i)} \Delta r_{st-i} + \sum a_{12}^{(i)} \Delta r_{Lt-i} + \varepsilon_{st} \quad (2.8)$$

$$\Delta r_{Lt} = a_{20} + \alpha_L(r_{Lt-1} - \beta r_{st-1}) + \sum a_{21}^{(i)} \Delta r_{st-i} + \sum a_{22}^{(i)} \Delta r_{Lt-i} + \varepsilon_{Lt} \quad (2.9)$$

donde ε_{st} , ε_{Lt} , Δr_{st-i} y Δr_{Lt-i} son términos estacionarios. Siguiendo el mismo argumento anterior, se concluye que la representación de la corrección de error requiere que las dos variables sean cointegradas.

El MCE dado por (2.8) y (2.9) se puede considerar un VAR de dos variables en primeras diferencias extendido con términos de corrección de error $\alpha_s(r_{Lt-1} - \beta r_{st-1})$ y $-\alpha_L(r_{Lt-1} - \beta r_{st-1})$ donde los coeficientes α_s y α_L son conocidos como velocidades de ajuste de los parámetros. Cuanto mayor sea α_s , mayor será la respuesta de la desviación del equilibrio de largo plazo del periodo anterior; valores pequeños de α_s indicarán que la tasa de interés de corto plazo no responde al error de equilibrio del último periodo. La sucesión $\{\Delta r_{st}\}$ no es afectada por la sucesión $\{\Delta r_{Lt}\}$ si $\alpha_s = 0$ y todos los coeficientes $a_{12}^{(i)} = 0$. Si α_s y α_L son cero, no existe relación de equilibrio de largo plazo y el modelo no es de corrección de error o

cointegración. Para que el modelo sea MCE, al menos uno de los términos de velocidad de ajuste, α_s y α_L , debe ser distinto de cero.

El resultado dado para el MCE (2.8) y (2.9), se puede generalizar a un modelo de n variables de la siguiente manera. El vector $\mathbf{y}_t = (y_{1t}, y_{2t}, \dots, y_{nt})'$ de dimensión $(n \times 1)$, tiene una representación de corrección de error si se puede expresar como

$$\Delta \mathbf{y}_t = \boldsymbol{\pi}_0 + \boldsymbol{\pi} \mathbf{y}_{t-1} + \boldsymbol{\pi}_1 \Delta \mathbf{y}_{t-1} + \boldsymbol{\pi}_2 \Delta \mathbf{y}_{t-2} + \dots + \boldsymbol{\pi}_p \Delta \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t \quad (2.10)$$

donde $\boldsymbol{\pi}_0$ es un vector de dimensión $(n \times 1)$ de términos de ordenada al origen con elementos π_{i0} , $\boldsymbol{\pi}$ es una matriz con elementos π_{jl} tal que uno o más $\pi_{jl} \neq 0$, $\boldsymbol{\pi}_i$ son matrices de coeficientes de dimensión $(n \times n)$ con elementos $\pi_{jl}^{(i)}$ y $\boldsymbol{\varepsilon}_t$ es un vector $(n \times 1)$ de perturbaciones con elementos ε_{jt} . Los elementos ε_{jt} pueden estar correlacionados con ε_{lt} . Ahora bien, asumir que las variables en \mathbf{y}_t son $I(1)$.

Para las variables en \mathbf{y}_t , existe una representación de corrección de error, como (2.10), si existe una combinación lineal de las variables $I(1)$ que sea estacionaria.

Resolviendo (2.10) para $\boldsymbol{\pi} \mathbf{y}_{t-1}$ se obtiene

$$\boldsymbol{\pi} \mathbf{y}_{t-1} = \Delta \mathbf{y}_t - \boldsymbol{\pi}_0 - \sum_{i=1}^p \boldsymbol{\pi}_i \Delta \mathbf{y}_{t-i} - \boldsymbol{\varepsilon}_t$$

Ya que cada término del lado derecho de la igualdad es estacionario, $\boldsymbol{\pi} \mathbf{y}_{t-1}$ es estacionario. Como $\boldsymbol{\pi}$ contiene sólo constantes, cada renglón de $\boldsymbol{\pi}$ es un vector de cointegración de $\boldsymbol{\pi}_t$. Por ejemplo, para el primer renglón de $\boldsymbol{\pi}$,

$$\pi_{11} y_{1,t-1} + \pi_{12} y_{2,t-1} + \dots + \pi_{1n} y_{n,t-1}$$

cada serie es $I(1)$. Por lo tanto, el renglón $(\pi_{11}, \pi_{12}, \dots, \pi_{1n})$ es un vector de cointegración para \mathbf{y}_t .

La característica importante de la expresión (2.10) es la matriz $\boldsymbol{\pi}$. Dos aspectos son importantes de resaltar:

1. Si todos los elementos de $\boldsymbol{\pi}$ son cero, (2.10) es un VAR en primeras diferencias y no existe representación de corrección de error porque $\Delta \mathbf{y}_t$ no responde a la desviación de periodos previos del equilibrio de largo plazo.
2. Si uno o más de los π_{jl} son distintos de cero, $\Delta \mathbf{y}_t$ responde a la desviación de periodos previos del equilibrio de largo plazo. En este caso, no es apropiado estimar \mathbf{y}_t como un VAR en primeras diferencias si \mathbf{y}_t tiene una representación de corrección de error. La omisión del término $\boldsymbol{\pi} \mathbf{y}_{t-1}$ implica un error de especificación incorrecta si \mathbf{y}_t tiene una representación de corrección de error como en (2.10).

Para analizar la relación entre cointegración y corrección de error, sean consideradas las propiedades del VAR simple:

$$x_t = \phi_{11}x_{t-1} + \phi_{12}z_{t-1} + \varepsilon_{x_t} \tag{2.11}$$

$$z_t = \phi_{21}x_{t-1} + \phi_{22}z_{t-1} + \varepsilon_{z_t} \tag{2.12}$$

donde ε_{x_t} y ε_{z_t} son procesos de ruido blanco gaussiano que pueden estar correlacionados. Por medio del *operador rezago* L^{14} es posible reescribir las expresiones (2.11) y (2.12) como:

¹⁴ Recordar que $Ly_t = L(y_t) = y_{t-1}$.

$$(1 - \phi_{11}L)x_t - \phi_{12}Lz_t = \varepsilon_{x_t}$$

$$-\phi_{21}Lx_t + (1 - \phi_{22}L)z_t = \varepsilon_{z_t}$$

o matricialmente:

$$\begin{bmatrix} (1 - \phi_{11}L) & -\phi_{12}L \\ -\phi_{21}L & (1 - \phi_{22}L) \end{bmatrix} \begin{bmatrix} x_t \\ z_t \end{bmatrix} = \begin{bmatrix} \varepsilon_{x_t} \\ \varepsilon_{z_t} \end{bmatrix}$$

Resolviendo para x_t y z_t :

$$x_t = \frac{(1 - \phi_{22}L)\varepsilon_{x_t} + \phi_{12}L\varepsilon_{z_t}}{(1 - \phi_{11}L)(1 - \phi_{22}L) - \phi_{12}\phi_{21}L^2} \quad (2.13)$$

$$z_t = \frac{\phi_{21}L\varepsilon_{x_t} + (1 - \phi_{11}L)\varepsilon_{z_t}}{(1 - \phi_{11}L)(1 - \phi_{22}L) - \phi_{12}\phi_{21}L^2} \quad (2.14)$$

Se ha transformado el VAR bivariado en dos ecuaciones en diferencia de segundo orden con la misma ecuación característica inversa $(1 - \phi_{11}L)(1 - \phi_{22}L) - \phi_{12}\phi_{21}L^2 = 0$. Haciendo que $(1 - \phi_{11}L)(1 - \phi_{22}L) - \phi_{12}\phi_{21}L^2 = 0$ y resolviendo para L se obtienen las dos raíces de la ecuación característica inversa. Para trabajar con las raíces características (en oposición a las raíces características inversas), se define $\lambda = 1/L$ y se escribe la ecuación característica como

$$\lambda^2 - (\phi_{11} + \phi_{22})\lambda + (\phi_{11}\phi_{22} - \phi_{12}\phi_{21}) = 0 \quad (2.15)$$

Ya que las dos variables tienen la misma ecuación característica, las raíces características de (2.15) determinan la trayectoria temporal de las dos variables.

Enseguida se resumen las trayectorias temporales de $\{x_t\}$ y $\{z_t\}$:

1. Si ambas raíces características (λ_1, λ_2) están dentro del círculo unitario, (1.13) y (1.14) dan soluciones estables para $\{x_t\}$ y $\{z_t\}$. Si t es suficientemente grande o si las condiciones iniciales son tales que la solución homogénea es cero, la condición de estabilidad garantiza que las variables son estacionarias. Las variables no pueden ser $CI(1, 1)$ ya que cada una es estacionaria.
2. Si las raíces están fuera del círculo unitario, las soluciones son explosivas. Ninguna variable es estacionaria en diferencia de modo que ellas no pueden ser $CI(1, 1)$. De la misma manera, si ambas raíces características son unitarias, la segunda diferencia de cada variable será estacionaria. Ya que cada variable es $I(2)$, no pueden ser $CI(1, 1)$.
3. Como se puede ver en (1.11) y (1.12), si $\phi_{12} = \phi_{21} = 0$, la solución es trivial. Para que $\{x_t\}$ y $\{z_t\}$ sean procesos de raíz unitaria, es necesario que $\phi_{11} + \phi_{22} = 1$. De esto se sigue que $\lambda_1 = \lambda_2 = 1$ y que las dos variables evolucionan sin ninguna relación de largo plazo y, por lo tanto, las variables no pueden estar cointegradas.
4. Para que $\{x_t\}$ y $\{z_t\}$ sean $CI(1, 1)$, es necesario que una raíz característica sea uno y que la otra sea menor a uno en valor absoluto. En este caso cada variable tendrá la misma tendencia estocástica y la primera diferencia de cada variable será estacionaria. Por ejemplo, si $\lambda_1 = 1$, (2.13) tendrá la forma

$$x_t = \frac{(1 - \phi_{22}L)\varepsilon_{x_t} + \phi_{12}L\varepsilon_{z_t}}{(1 - L)(1 - \lambda_2L)}$$

o multiplicando por $(1 - L)$ se tiene

$$(1 - L)x_t = \Delta x_t = \frac{(1 - \phi_{22}L)\varepsilon_{x_t} + \phi_{12}L\varepsilon_{z_t}}{(1 - \lambda_2L)}$$

la cual es estacionaria si $|\lambda_2| < 1$.

En consecuencia, para asegurarse que las variables sean $CI(1,1)$ se debe establecer una de las raíces características igual a uno y la otra a un valor menor a uno en valor absoluto. Para que la más grande de las dos raíces sea igual a uno debe ocurrir que

$$\frac{1}{2}(\phi_{11} + \phi_{22}) + \frac{1}{2}\sqrt{(\phi_{11}^2 + \phi_{22}^2) - 2\phi_{11}\phi_{22} + 4\phi_{12}\phi_{21}} = 1$$

Simplificando, los coeficientes deben satisfacer

$$\phi_{11} = \frac{(1 - \phi_{22}) - \phi_{12}\phi_{21}}{1 - \phi_{22}} \tag{2.16}$$

Ahora para la segunda raíz característica. Ya que ϕ_{12} y/o ϕ_{21} deben ser distintas de cero si las variables están cointegradas, la condición $|\lambda_2| < 1$ requiere que

$$\phi_{22} > -1 \tag{2.17}$$

y

$$\phi_{12}\phi_{21} + (\phi_{22}^2) < 1 \tag{2.18}$$

Las ecuaciones (2.16), (2.17) y (2.18) son restricciones que se deben imponer sobre los coeficientes de (2.11) y (2.12) para asegurar que las variables sean $CI(1, 1)$. Para ver cómo estas restricciones sobre coeficientes se relacionan con la naturaleza de las soluciones, se reescriben (2.11) y (2.12) de la siguiente forma:

$$\begin{bmatrix} \Delta x_t \\ \Delta z_t \end{bmatrix} = \begin{bmatrix} \phi_{11} - 1 & \phi_{12} \\ \phi_{21} & \phi_{22} - 1 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x_t} \\ \varepsilon_{z_t} \end{bmatrix} \quad (2.19)$$

Ahora (2.16) implica que $\phi_{11} - 1 = \frac{-\phi_{12}\phi_{21}}{1-\phi_{22}}$. Manipulando, (2.19) se puede escribir como

$$\Delta x_t = -\left(\frac{-\phi_{12}\phi_{21}}{1-\phi_{22}}\right)x_{t-1} + \phi_{12}z_{t-1} + \varepsilon_{x_t} \quad (2.20)$$

$$\Delta z_t = \phi_{21}x_{t-1} - (1 - \phi_{22})z_{t-1} + \varepsilon_{z_t} \quad (2.21)$$

Las ecuaciones (2.20) y (2.21) forman un modelo de corrección de error. Si ϕ_{12} y ϕ_{21} son distintos de cero, se puede normalizar el vector de cointegración con respecto a las dos variables. Normalizando respecto a x_t se tiene:

$$\Delta x_t = \alpha_x(x_{t-1} - \beta z_{t-1}) + \varepsilon_{x_t}$$

$$\Delta z_t = \alpha_z(x_{t-1} - \beta z_{t-1}) + \varepsilon_{z_t}$$

donde, $\alpha_x = \frac{-\phi_{12}\phi_{21}}{1-\phi_{22}}$, $\beta = \frac{1-\phi_{22}}{\phi_{21}}$ y $\alpha_z = \phi_{21}$.

Se puede ver que x_t y z_t cambian en respuesta a la desviación del periodo previo del equilibrio de largo plazo ($x_{t-1} - \beta z_{t-1}$). Si $x_{t-1} = \beta z_{t-1}$, x_t y z_t cambian sólo en respuesta a los shocks ε_{x_t} y ε_{z_t} . Además, si $\alpha_x < 0$ y $\alpha_z > 0$, x_t decrece y z_t crece en respuesta a una desviación positiva del equilibrio de largo plazo.

Se observa que las condiciones (2.17) y (2.18) aseguran que $\beta \neq 0$ y que al menos uno de los parámetros de la velocidad de ajuste (α_x o α_z) es distinto de cero. Este modelo tiene la misma forma que el modelo de tasas de interés dado por las expresiones (2.6) y (2.7).

Aunque ϕ_{12} y ϕ_{21} no pueden ser ambos cero, un caso de especial interés surge si uno de los coeficientes es cero. Por ejemplo, si se hace $\phi_{12} = 0$, el coeficiente de velocidad de ajuste $\alpha_x = 0$. En este caso, x_t cambia sólo en respuesta a ε_{x_t} como $\Delta x_t = \varepsilon_{x_t}$. La sucesión $\{z_t\}$ hace toda la corrección para eliminar cualquier desviación del equilibrio de largo plazo.

Resaltan algunas implicaciones importantes de este modelo simple:

1. Las restricciones necesarias que aseguran que las variables son $CI(1,1)$ son garantía de que existe un MCE. En el ejemplo, $\{x_t\}$ y $\{z_t\}$ son procesos de raíz unitaria pero la combinación lineal $(x_t - \beta z_t)$ es estacionaria; el vector de cointegración es $\left[1, -\left(\frac{1-\phi_{22}}{\phi_{21}}\right)\right]$. Las variables tienen una representación de corrección de error con coeficientes de velocidad de ajuste $\alpha_x = \frac{-\phi_{12}\phi_{21}}{1-\phi_{22}}$ y $\alpha_z = \phi_{21}$. También se muestra que un MCE para variables $I(1)$ implica necesariamente cointegración. Este hallazgo ilustra el teorema de representación de Granger según el cual para todo conjunto de variables $I(1)$, corrección de error y cointegración son representaciones equivalentes.
2. Una cointegración necesita restricciones sobre coeficientes en un modelo VAR. Es importante darse cuenta de que un sistema cointegrado se puede ver como una forma restringida de un VAR general. Sea $\mathbf{y}_t = (x_t, z_t)'$ y $\varepsilon_t = (\varepsilon_{x_t}, \varepsilon_{z_t})'$ de modo que (2.19) puede escribirse en la forma

$$\Delta \mathbf{y}_t = \boldsymbol{\pi} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$$

(2.22)

No es apropiado estimar un VAR de variables cointegradas usando sólo primeras diferencias. Estimar (2.22) sin el término πy_{t-1} eliminaría la porción de corrección de error del modelo. También es importante notar que los renglones de π no son linealmente independientes si las variables son cointegradas.

Multiplicando cada elemento del renglón 1 por $-\left(\frac{1-\phi_{22}}{\phi_{21}}\right)$ se obtiene el correspondiente elemento en el renglón 2. Así, el determinante de π es cero y x_t y z_t tienen la representación de corrección de error dada por (2.20) y (2.21).

Este ejemplo de dos variables ilustra la importante idea de Johansen (1988) y de Stock y Watson (1988) de que se puede usar el rango de π para determinar si dos variables $\{x_t\}$ y $\{z_t\}$ son o no cointegradas. Comparar el determinante de π de la ecuación característica dada por (2.15). Si la mayor raíz característica es igual 1 ($\lambda_1 = 1$), entonces el determinante de π es cero y π tiene rango igual a uno. Si π ha de tener rango cero sería necesario que $\phi_{11} = 1$, $\phi_{22} = 1$ y $\phi_{12}\phi_{21} = 0$. El VAR representado por (2.11) y (2.12) no sería nada más que $\Delta x_t = \varepsilon_{x_t}$ y $\Delta z_t = \varepsilon_{z_t}$. En este caso las sucesiones $\{x_t\}$ y $\{z_t\}$ son procesos de raíz unitaria sin ningún vector de cointegración. Por último, si el rango de π es completo, entonces ninguna raíz característica puede ser uno, de modo que las sucesiones $\{x_t\}$ y $\{z_t\}$ son conjuntamente estacionarias.

3. En general, las dos variables en un sistema cointegrado responderán a una desviación del equilibrio de largo plazo. No obstante, es posible que uno (pero no ambos) de los parámetros de velocidad de ajuste sea cero. Por

ejemplo, si $\alpha_x = 0$, $\{x_t\}$ no responde a la discrepancia del equilibrio de largo plazo y $\{z_t\}$ hace todo el ajuste. En esta circunstancia se dice que $\{x_t\}$ es débilmente exógena. Como tal, es posible estimar un modelo econométrico para $\{z_t\}$ y probar hipótesis sin referencia a un modelo específico para $\{x_t\}$.

También es necesario reinterpretar la causalidad de Granger en un sistema cointegrado. En tal sistema, $\{x_t\}$ no causa a $\{z_t\}$ en el sentido de Granger si valores rezagados Δx_{t-i} no entran en la ecuación Δz_t y si z_t no responde a la desviación del equilibrio de largo plazo. Por lo tanto, $\{z_t\}$ debe ser débilmente exógena. Si $\phi_{21} = 0$ en (2.21), $\{z_t\}$ es débilmente exógena y no causa en el sentido de Granger a $\{x_t\}$. Similarmente, en el sistema cointegrado (2.8) y (2.9), $\{r_{Lt}\}$ es débilmente exógena y no causa en el sentido de Granger a $\{r_{st}\}$ si todos los $\phi_{12}^{(i)} = 0$ y si $\alpha_0 = 0$.

2.5.1 El caso de n variables

En el caso de n variables poco cambia. La relación entre cointegración, corrección de error y rango de la matriz $\boldsymbol{\pi}$ no cambia al agregar variables adicionales al sistema. La característica interesante en el caso de n variables es la posibilidad de vectores de cointegración múltiples. Considerar ahora una versión más general de (2.11) y (2.12) extendida a n variables:

$$\mathbf{y}_t = \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{2.23}$$

donde \mathbf{y}_t es el vector $(y_{1t}, y_{2t}, \dots, y_{nt})'$ de orden $n \times 1$, Φ_1 es una matriz cuadrada de orden n y $\boldsymbol{\varepsilon}_t$ un vector $(\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{nt})'$ de orden $n \times 1$. Restando \mathbf{y}_{t-1} de (2.23) se obtiene:

$$\begin{aligned}\Delta \mathbf{y}_t &= -(\mathbf{I} - \Phi_1)\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \\ &= \boldsymbol{\pi}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t\end{aligned}\tag{2.24}$$

como en (2.22) pero ahora $\boldsymbol{\pi}$ es la matriz cuadrada de orden n , $-(\mathbf{I} - \Phi_1)$ y π_{ij} representa el elemento en el renglón i y la columna j de $\boldsymbol{\pi}$. Como se ve, (2.24) es un caso especial de (2.10) tal que todos los $\boldsymbol{\pi}_i = 0$.

Nuevamente, el aspecto fundamental de la cointegración es el rango de la matriz $\boldsymbol{\pi}$ cuadrada de orden n . Si el rango de la matriz $\boldsymbol{\pi}$ es cero, cada elemento de $\boldsymbol{\pi}$ debe ser igual a cero. En este caso, (2.24) es equivalente a un VAR de n variables en primeras diferencias

$$\Delta \mathbf{y}_t = \boldsymbol{\varepsilon}_t$$

En (2.25) cada $\Delta y_{jt} = \varepsilon_{jt}$ de modo que la primera diferencia de cada variable en el vector \mathbf{y}_t es $I(0)$. Ya que cada $y_{jt} = y_{jt-1} + \varepsilon_{jt}$, todas las sucesiones $\{y_{jt}\}$ son procesos de raíz unitaria y no existe combinación lineal de las variables que sea estacionaria.

En el otro extremo, asumir que $\boldsymbol{\pi}$ tiene rango completo. La solución de largo plazo de (2.24) es dada por las n ecuaciones independientes

$$\pi_{11}y_{1t} + \pi_{12}y_{2t} + \dots + \pi_{1n}y_{nt} = 0$$

$$\pi_{21}y_{1t} + \pi_{22}y_{2t} + \dots + \pi_{2n}y_{nt} = 0$$

⋮

$$\pi_{n1}y_{1t} + \pi_{n2}y_{2t} + \cdots + \pi_{nn}y_{nt} = 0 \quad (2.25)$$

Cada una de estas n ecuaciones es una restricción independiente de la solución de largo plazo de las variables; las n variables en el sistema encaran n restricciones de largo plazo. En este caso, cada una de las n variables contenidas en el vector \mathbf{y}_t debe ser estacionaria con valores de largo plazo dados por (2.25). Las variables no pueden ser $CI(1, 1)$ ya que son estacionarias.

En casos intermedios, en los cuales el rango de $\boldsymbol{\pi}$ es igual a $r < n$, existen r vectores de cointegración. Con r ecuaciones independientes y n variables, hay $n - r$ tendencias estocásticas en el sistema. Si $r = 1$, hay un sólo vector de cointegración dado por un único renglón de la matriz $\boldsymbol{\pi}$. Cada sucesión $\{\Delta y_{jt}\}$ se puede escribir en forma de corrección de error. Por ejemplo, se puede escribir Δy_{jt} como

$$\Delta y_{jt} = \pi_{11}y_{1,t-1} + \pi_{12}y_{2,t-1} + \cdots + \pi_{1n}y_{n,t-1} + \varepsilon_{1t}$$

o, normalizando con respecto a $y_{1,t-1}$, se puede establecer $\alpha_1 = \pi_{11}$ y $\beta_{1l} = \pi_{1l}/\pi_{11}$ para obtener

$$\Delta y_{jt} = \alpha_1(y_{1,t-1} + \beta_{12}y_{2,t-1} + \cdots + \beta_{1n}y_{n,t-1}) + \varepsilon_{1t} \quad (2.26)$$

En el largo plazo, la sucesión $\{\Delta y_{jt}\}$ satisfacerá la relación

$$y_{1t} + \beta_{12}y_{2t} + \cdots + \beta_{1n}y_{nt} = 0$$

En consecuencia, el vector de cointegración normalizado es $(1, \beta_{12}, \beta_{13}, \dots, \beta_{1n})$, el parámetro de velocidad de ajuste es α_1 . De manera parecida, con dos vectores de cointegración los valores de largo plazo de las variables satisfacen las dos relaciones:

$$\pi_{11}y_{1t} + \pi_{12}y_{2t} + \dots + \pi_{1n}y_{nt} = 0$$

$$\pi_{21}y_{1t} + \pi_{22}y_{2t} + \dots + \pi_{2n}y_{nt} = 0$$

las cuales pueden normalizarse.

El punto principal aquí es que existen dos formas importantes de probar cointegración. La metodología de Engle y Granger (1987) busca determinar si los residuos de la relación de equilibrio son estacionarios. Las metodologías de Johansen (1988) y de Stock y Watson (1988) determinan el rango de π . Enseguida se expone la metodología de Johansen que es la que se utiliza en la investigación aplicada de este trabajo.

2.6 METODOLOGÍA DE JOHANSEN PARA PROBAR COINTEGRACIÓN

2.6.1 Raíces características, rango y cointegración

Los procedimientos de Johansen (1988) y de Stock y Watson (1988) dependen en gran medida de la relación entre el rango de una matriz y sus raíces características.

En el caso univariado, es posible considerar la estacionariedad de $\{y_t\}$ como dependiente de la magnitud $(\phi_1 - 1)$, esto es,

$$y_t = \phi_1 y_{t-1} + \varepsilon_t$$

o

$$\Delta y_t = (\phi_1 - 1)y_{t-1} + \varepsilon_t$$

Si $(\phi_1 - 1) = 0$, el proceso $\{y_t\}$ tiene una raíz unitaria. Descartando el caso en el cual $\{y_t\}$ es explosiva, si $(\phi_1 - 1) \neq 0$, se puede concluir que la sucesión $\{y_t\}$ es estacionaria. Las tablas DF (Dickey–Fuller) dan los estadísticos apropiados para probar formalmente la hipótesis nula $(\phi_1 - 1) = 0$ ¹⁵.

Considerar la generalización simple a n variables. Como en (2.23) sea

$$\mathbf{y}_t = \Phi_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$$

de modo que

$$\begin{aligned} \Delta \mathbf{y}_t &= \Phi_1 \mathbf{y}_{t-1} - \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \\ &= (\Phi_1 - \mathbf{I}) \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \\ &= \boldsymbol{\pi} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \end{aligned} \tag{2.27}$$

donde \mathbf{y}_t y $\boldsymbol{\varepsilon}_t$ son vectores de dimensión $(n \times 1)$, Φ_1 una matriz cuadrada de orden n , \mathbf{I} una matriz identidad cuadrada de orden n y $\boldsymbol{\pi} = (\Phi_1 - \mathbf{I})$.

El rango de $(\Phi_1 - \mathbf{I})$ es igual al número de vectores de cointegración. En analogía al caso univariado, si $(\Phi_1 - \mathbf{I})$ está compuesta solamente por ceros, entonces por el rango de $\boldsymbol{\pi}$ todas las sucesiones $\{y_{jt}\}$ son procesos de raíz unitaria.

¹⁵ Ver apéndice A.1 que trata las pruebas de raíz unitaria de Dickey-Fuller.

Ya que no existe combinación lineal de los procesos $\{y_{jt}\}$ que sea estacionaria, las variables no son cointegradas. Si se excluyen las raíces características que sean mayores que uno y si $\text{rango}(\boldsymbol{\pi}) = n$, (2.27) representa un sistema convergente de ecuaciones en diferencia, de modo que todas las variables son estacionarias.

Hay algunas formas de generalizar (2.27). La ecuación se puede modificar para permitir la presencia de término de desplazamiento; simplemente hay que hacer que

$$\Delta \mathbf{y}_t = \mathbf{k} + \boldsymbol{\pi} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \quad (2.28)$$

donde \mathbf{k} es el vector $(n \times 1)$ de constantes $(k_1, k_2, \dots, k_n)'$.

El efecto de la inclusión de diversos k_i es permitir la posibilidad de una tendencia temporal lineal en el mecanismo que genera los datos (MGD). Tal vez se quiera incluir el término desplazamiento al exhibir las variables una tendencia a crecer o decrecer. Aquí, el rango de $\boldsymbol{\pi}$ se puede ver como el número de vectores de cointegración que existen en los datos “sin tendencia”. En el largo plazo, $\boldsymbol{\pi} \mathbf{y}_{t-1} = \mathbf{0}$, de modo que cada sucesión $\{\Delta y_{jt}\}$ tiene un valor esperado de k_j . La agregación de todos estos cambios a lo largo de t proporciona la expresión determinista k_{jt} .

Una forma de incluir una constante en la relación de cointegración es restringir los valores de los diversos k_j . Por ejemplo, si el $\text{rango}(\boldsymbol{\pi}) = 1$, los renglones de $\boldsymbol{\pi}$ sólo pueden diferir por un escalar, de modo que es posible escribir cada sucesión $\{\Delta y_{jt}\}$ en (2.28) como

$$\Delta y_{1t} = \pi_{11} y_{1,t-1} + \pi_{12} y_{2,t-1} + \dots + \pi_{1n} y_{n,t-1} + k_1 + \varepsilon_{1t}$$

$$\Delta y_{2t} = s_2 (\pi_{11} y_{1,t-1} + \pi_{12} y_{2,t-1} + \dots + \pi_{1n} y_{n,t-1}) + k_2 + \varepsilon_{2t}$$

⋮

$$\Delta y_{nt} = s_n(\pi_{11}y_{1,t-1} + \pi_{12}y_{2,t-1} + \dots + \pi_{1n}y_{n,t-1}) + k_n + \varepsilon_{nt}$$

donde s_j son escalares tales que $s_j\pi_{1l} = \pi_{jl}$.

Si los k_j se pueden restringir de manera que $k_j = s_j k_j$, se sigue que todas las sucesiones $\{\Delta y_{jt}\}$ se pueden escribir con la constante incluida en el vector de cointegración:

$$\Delta y_{1t} = (\pi_{11}y_{1,t-1} + \pi_{12}y_{2,t-1} + \dots + \pi_{1n}y_{n,t-1} + k_1) + \varepsilon_{1t}$$

$$\Delta y_{2t} = s_2(\pi_{11}y_{1,t-1} + \pi_{12}y_{2,t-1} + \dots + \pi_{1n}y_{n,t-1} + k_2) + \varepsilon_{2t}$$

⋮

$$\Delta y_{nt} = s_n(\pi_{11}y_{1,t-1} + \pi_{12}y_{2,t-1} + \dots + \pi_{1n}y_{n,t-1} + k_n) + \varepsilon_{nt}$$

ó

$$\Delta \mathbf{y}_t = \mathbf{k} + \boldsymbol{\pi}^* \mathbf{y}_{t-1}^* + \boldsymbol{\varepsilon}_t \tag{2.29}$$

donde $\mathbf{y}_t = (y_{1t}, y_{2t}, \dots, y_{nt})'$, $\mathbf{y}_{t-1}^* = (y_{1,t-1}, y_{2,t-1}, \dots, y_{n,t-1}, 1)'$ y

$$\boldsymbol{\pi}^* = \begin{bmatrix} \pi_{11} & \pi_{12} & \dots & \pi_{1n} & k_1 \\ \pi_{21} & \pi_{22} & \dots & \pi_{2n} & k_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \pi_{n1} & \pi_{n2} & \dots & \pi_{nn} & k_n \end{bmatrix}$$

La característica interesante de (2.29) es que la tendencia lineal se purga del sistema. En esencia, los diversos k_j se han alterado de forma tal que la solución general para cada $\{y_{jt}\}$ no contiene una tendencia temporal. La solución al conjunto

de ecuaciones en diferencia (2.29) es tal que se espera que todos los Δy_{jt} sean iguales a cero cuando $\pi_{11}y_{1,t-1} + \pi_{12}y_{2,t-1} + \dots + \pi_{1n}y_{n,t-1} + k_j = 0$.

El modelo multivariado se puede generalizar para permitir un proceso autoregresivo de un orden alto. Sea

$$\mathbf{y}_t = \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \dots + \Phi_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t \quad (2.30)$$

donde \mathbf{y}_t es un proceso VAR como en (1.11) pero sin vector intersección \mathbf{k} .

La ecuación (2.30) se puede escribir en una forma más utilizable sumando y restando $\Phi_p \mathbf{y}_{t-p}$ al lado derecho para obtener

$$\mathbf{y}_t = \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \dots + \Phi_{p-2} \mathbf{y}_{t-p+2} + (\Phi_{p-1} + \Phi_p) \mathbf{y}_{t-p+2} - \Phi_p \Delta \mathbf{y}_{t-p+1} + \boldsymbol{\varepsilon}_t$$

Ahora, sumando y restando $(\Phi_{p-1} + \Phi_p) \mathbf{y}_{t-p+2}$ para obtener

$$\mathbf{y}_t = \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \dots - (\Phi_{p-1} + \Phi_p) \Delta \mathbf{y}_{t-p+2} - \Phi_p \Delta \mathbf{y}_{t-p+1} + \boldsymbol{\varepsilon}_t$$

Se puede continuar recursivamente de esta manera para obtener

$$\Delta \mathbf{y}_t = \boldsymbol{\pi} \mathbf{y}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\pi}_i \Delta \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t \quad (2.31)$$

donde $\boldsymbol{\pi} = (\mathbf{I} - \sum_{i=1}^{p-1} \Phi_i)$ y $\boldsymbol{\pi}_i = \sum_{j=i+1}^p \Phi_j$.

La característica clave a notar en (2.31) es el rango de la matriz $\boldsymbol{\pi}$. Si $\text{rango}(\boldsymbol{\pi}) = 0$, la matriz es nula y (2.31) es un modelo VAR en primeras diferencias. En cambio, si $\text{rango}(\boldsymbol{\pi}) = \mathbf{n}$, el proceso de vectores es estacionario. En casos intermedios, si el $\text{rango}(\boldsymbol{\pi}) = 1$, existe un solo vector de cointegración y la

expresión $\boldsymbol{\pi}\mathbf{y}_{t-1}$ es el término corrección de error. Para otros casos en los cuales $1 < \text{rango}(\boldsymbol{\pi}) < n$, existen vectores de cointegración múltiples.

El número de vectores de cointegración se puede obtener comprobando la significancia de las raíces características de $\boldsymbol{\pi}$. Se sabe que el rango de una matriz es igual al número de raíces características que son distintas de cero. Asumiendo que se ha obtenido la matriz $\boldsymbol{\pi}$ y ordenando las n raíces características tales que $\lambda_1 > \lambda_2 > \dots > \lambda_n$. Si las variables en \mathbf{y}_t no están cointegradas, el rango de $\boldsymbol{\pi}$ es cero y todas las raíces características serán igual a cero. Ya que $\ln(1) = 0$, la expresión $\ln(1 - \lambda_j)$ será cero si las variables no están cointegradas. Similarmente, si $\text{rango}(\boldsymbol{\pi}) = 1$, $0 < \lambda_1 < 1$ de modo que la primera expresión $\ln(1 - \lambda_1)$ será negativo y todas las otras $\lambda_j = 0$ de modo que $\ln(1 - \lambda_2) = \ln(1 - \lambda_3) = \dots = \ln(1 - \lambda_n) = 0$.

En la práctica, solamente se pueden obtener estimaciones de $\boldsymbol{\pi}$ y sus raíces características. La prueba para el número de raíces características que son insignificanamente distintas de uno se puede hacer utilizando los dos estadísticos de prueba siguientes:

$$\lambda_{traza}(r) = -T \sum_{j=r+1}^n \ln(1 - \hat{\lambda}_j) \tag{2.32}$$

$$\lambda_{max}(r, r + 1) = -T \ln(1 - \hat{\lambda}_{r+1}) \tag{2.33}$$

donde $\hat{\lambda}_j$ son los valores estimados de las raíces características que se obtienen de la matriz estimada $\boldsymbol{\pi}$, T es el número de observaciones utilizables.

Cuando los correspondientes valores de r son obtenidos, estos estadísticos son denominados λ_{traza} y λ_{max} .

Sea r_0 el número de vectores de cointegración distintos. Para el caso de λ_{traza} se tienen las hipótesis de prueba,

$$H_0: r_0 \leq r \quad \text{vs.}$$

$$H_a: r_0 > r$$

Debe notarse que λ_{traza} es igual a cero cuando todas las $\lambda_j = 0$. Cuanto mayores de cero sean las raíces características, la más negativa es $\ln(1 - \hat{\lambda}_j)$, y la más grande $\hat{\lambda}_j$ es el estadístico λ_{traza} .

Para λ_{max} , las hipótesis de prueba son las siguientes

$$H_0: r_0 = r \quad \text{vs.}$$

$$H_a: r_0 = r + 1$$

Nuevamente, si el valor estimado raíz característica es cercano a cero, λ_{max} será pequeño.

Para ambos estadísticos, sus valores críticos se obtienen usando la aproximación Monte Carlo. La distribución de los estadísticos depende de dos cosas:

1. El número de componentes no estacionarios bajo la hipótesis nula (es decir, $n - r$).
2. La forma del vector \mathbf{k} . Se usan distintas partes de tablas dependiendo si se incluye o no intercepto o desplazamiento en el vector de cointegración.

2.6.2 Pruebas de hipótesis

Como en las pruebas de raíz unitaria de Dickey–Fuller, en el procedimiento de Johansen es importante conocer la forma correcta de los regresores deterministas: los valores críticos de los estadísticos λ_{traza} y λ_{max} tienden a ser pequeños sin ningún regresor determinista y grandes con un intercepto incluido en el vector de cointegración. En lugar de postular la forma de \mathbf{k} , es posible restringir las formas del vector.

Uno de los aspectos más interesantes del procedimiento de Johansen es que permite probar formas restringidas del vector o vectores de cointegración. La clave para conocer todas esas pruebas de hipótesis es que *si existen r vectores de cointegración, sólo r combinaciones lineales de las variables son estacionarias.*

Todas las demás combinaciones lineales no son estacionarias. Así, supóngase que se reestima el modelo restringiendo los parámetros de $\boldsymbol{\pi}$. Si las restricciones no son vinculantes, se debe encontrar que el número de vectores de cointegración no disminuye.

Para probar la presencia de una intersección en el vector de cointegración en oposición a un desplazamiento en las restricciones \mathbf{k} , Se estiman las dos formas del modelo. Representando las raíces características ordenadas de la matriz no restringida $\boldsymbol{\pi}$ por $\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_n$ y las raíces características del modelo con intercepto(s) en el (los) vector(es) de cointegración por $\hat{\lambda}_1^*, \hat{\lambda}_2^*, \dots, \hat{\lambda}_n^*$. Suponiendo que la forma no restringida del modelo tiene r raíces características distintas de cero. El estadístico

$$-T \sum_{j=r+1}^n [\ln(1 - \hat{\lambda}_1^*) - \ln(1 - \hat{\lambda}_j)] \quad (2.34)$$

es asintóticamente distribuido como una χ^2 con $(n - r)$ grados de libertad.

La idea detrás de la prueba es que todos los valores de $\ln(1 - \hat{\lambda}_1^*)$ y de $\ln(1 - \hat{\lambda}_j)$ deberán ser equivalentes si la restricción no es vinculante. Por lo tanto, valores pequeños del estadístico de prueba implican que está permitido incluir la intersección en el vector de cointegración. Sin embargo, la posibilidad de hallar una combinación lineal estacionaria de las n variables es mayor con el intercepto en el vector de cointegración que si el intercepto no está en el vector de cointegración. Así, un valor grande de $\hat{\lambda}_{t+1}^*$ (y un correspondiente valor grande de $-T\ln(1 - \hat{\lambda}_{t+1}^*)$) implica que la restricción “infla” artificialmente el número de vectores de cointegración. Así, como probó Johansen (1991), si el estadístico es suficientemente grande, es posible rechazar la hipótesis nula de una intersección en el (los) vector(es) de cointegración y concluir que existe una tendencia lineal en las variables.

Para probar otras restricciones sobre el vector de cointegración, Johansen define las dos matrices α y β de dimensión $n \times r$, donde r es el rango de π .

Las propiedades de α y β son tales que $\pi = \alpha\beta'$. Es posible notar que β es la matriz de parámetros de cointegración y α es la matriz de peso con el que cada vector de cointegración entra en las n ecuaciones del VAR. En cierto sentido, α se puede ver como la matriz de la velocidad de los parámetros de ajuste. Debido a las restricciones entre ecuaciones, no es posible estimar α y β por MCO. Sin embargo, mediante estimación de máxima verosimilitud es posible:

- (a) estimar (2.31) como modelo de corrección de error,
- (b) determinar el rango de $\boldsymbol{\pi}$,
- (c) usar los r vectores de cointegración más significativos para formar $\boldsymbol{\beta}'$ y
- (d) elegir $\boldsymbol{\alpha}$ tal que $\boldsymbol{\pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$.

A continuación se observa el proceso en el caso de un vector de cointegración. Dado que el $\text{rango}(\boldsymbol{\pi}) = 1$, todos los renglones de $\boldsymbol{\pi}$ son múltiplos lineales unos de otros. Por lo tanto, las ecuaciones en (2.31) tienen la forma:

$$\Delta y_{1t} = \pi_{11}y_{1,t-1} + \pi_{12}y_{2,t-1} + \cdots + \pi_{1n}y_{n,t-1} + \cdots + \varepsilon_{1t}$$

$$\Delta y_{2t} = s_2(\pi_{11}y_{1,t-1} + \pi_{12}y_{2,t-1} + \cdots + \pi_{1n}y_{n,t-1}) + \cdots + \varepsilon_{2t}$$

⋮

$$\Delta y_{nt} = s_n(\pi_{11}y_{1,t-1} + \pi_{12}y_{2,t-1} + \cdots + \pi_{1n}y_{n,t-1}) + \cdots + \varepsilon_{nt}$$

donde s_j son escalares y, para simplificar notación, las matrices $\boldsymbol{\pi}_i\Delta\mathbf{y}_{t-i}$ no se han escrito.

Ahora definiendo $\alpha_j = s_j\pi_{1j}$ y $\beta_j = \pi_{1j}/\pi_{11}$, de modo que cada ecuación se puede escribir en la forma

$$\Delta y_{jt} = \alpha_j(\pi_{1,t-1} + \beta_2y_{2,t-1} + \cdots + \beta_ny_{n,t-1}) + \cdots + \varepsilon_{1t}; \quad j = 1, \dots, n$$

o

$$\Delta\mathbf{y}_t = \sum_{i=1}^{p-1} \boldsymbol{\pi}_i\Delta\mathbf{y}_{t-i} + \boldsymbol{\alpha}\boldsymbol{\beta}' + \boldsymbol{\varepsilon}_t$$

donde el único vector de cointegración es $\boldsymbol{\beta} = (1, \beta_2, \beta_3, \dots, \beta_n)'$ y la velocidad de ajuste de parámetros se dan por $\boldsymbol{\alpha} = (1, \alpha_2, \alpha_3, \dots, \alpha_n)'$. Una vez que $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ se determinan el ensayo de diversas restricciones sobre $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ es sencilla si se recuerda que el punto fundamental de que si hay r vectores de cointegración, sólo esas r combinaciones lineales de las variables son estacionarias. Así, los estadísticos de prueba λ_{traza} y λ_{max} descritos en el apartado anterior, implican la comparación del número de vectores de cointegración bajo sus hipótesis nula y alternativa.

Otra vez, sean $\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_n$ y $\hat{\lambda}_1^*, \hat{\lambda}_2^*, \dots, \hat{\lambda}_n^*$ las raíces características ordenadas de los modelos sin restricción y restringido, respectivamente. Para probar las restricciones sobre $\boldsymbol{\beta}$, a partir de las hipótesis nula y alternativa:

$$H_0: \beta_{r+1} = \dots = \beta_n = 0 \quad \text{vs.}$$

$$H_a: \beta_i \neq 0, \text{ para alguna } i = r + 1, \dots, n.$$

se forma el estadístico de prueba

$$T \sum_{j=1}^r [\ln(1 - \hat{\lambda}_1^*) - \ln(1 - \hat{\lambda}_j)] \tag{2.35}$$

que se distribuye asintóticamente como una χ^2 con un número de grados de libertad igual al número de restricciones sobre $\boldsymbol{\beta}$.

Valores pequeños de $\hat{\lambda}_j^*$ en relación a $(1 - \hat{\lambda}_j)$ (para $j \leq r$), implican un número reducido de vectores de cointegración. Por lo tanto, las restricciones incorporadas en la hipótesis nula de interés son vinculantes si el valor calculado del estadístico de prueba es mayor que en una tabla χ^2 .

Las restricciones sobre α se pueden probar de manera parecida. El procedimiento es restringir α y comparar las r raíces características más significativas de los modelos no restringido y restringido usando (2.35). Si el valor calculado de (2.35) es mayor que una χ^2 con grados de libertad igual al número de restricciones en α , las restricciones se pueden rechazar.

Si se puede determinar que existe un sólo vector de cointegración, es frecuente basarse en la estimación de un modelo de corrección de error para probar las restricciones sobre α . Si $r = 1$, y se prueba un sólo valor de α , el estadístico usual t es asintóticamente equivalente a la prueba de Johansen.

El programa econométrico Eviews ofrece la posibilidad de probar distintas especificaciones tanto de las ecuaciones del VAR como de las ecuaciones de cointegración (EC) de acuerdo a la inclusión de desplazamientos y tendencias temporales en ellas, así como de incorporar restricciones específicas sobre valores de α y β .

Las opciones de especificación dentro de la prueba para probar cointegración de Johansen con y sin restricciones, que dan por resultado tanto las pruebas para λ_{traza} y λ_{max} para el orden de cointegración, el cálculo de los coeficientes de cointegración, así como los coeficientes de cointegración normalizados, son las siguientes:

Si la prueba asume que no existe tendencia determinista en los datos:

1. No existe intercepto o tendencia en la EC o en el VAR.
2. Existe desplazamiento pero no tendencia en la EC. Sin intercepto en el VAR.

Si la prueba admite tendencia determinista lineal en los datos:

3. Existe desplazamiento y no tendencia en la EC y en el VAR.
4. Existen tendencia y desplazamiento en la EC. Sin tendencia en el VAR.

Si la prueba admite tendencia determinista cuadrática en los datos:

5. Intercepto y tendencia en la EC. Tendencia lineal en el VAR.

Y existe una sexta opción que permite:

6. Resumir el conjunto de las 5 suposiciones anteriores.

Para los casos 1 y 2 el VAR es estimado después de obtener las primeras diferencias. Para los casos 3, 4 y 5 la regresión es aplicada en niveles para la EC asumida con tendencia.

Entonces, para correr una prueba para el rango de cointegración, se necesita escoger a partir de las cinco posibles especificaciones aquella que parezca más plausible para la información a la mano, y se debe así también especificar el número de rezagos a incluir en el VAR.

SEGUNDA PARTE

APLICACIÓN DEL MODELO VAR

CAPÍTULO 3

Determinantes del crecimiento económico en México, de 1980 a 2007

3.1 INTRODUCCIÓN

El objeto de esta segunda parte será un estudio del crecimiento económico en México, en el periodo que va del primer trimestre de 1980 al último trimestre de 2007, identificando para ello factores macroeconómicos determinantes al considerar dos modelos de inspiración post Keynesiana, aplicando sobre cada modelo los métodos econométricos vistos en la primera parte.

En McCombie y Thirlwall (1994), se tiene que para la teoría Keynesiana es el comportamiento de la demanda el que determina el crecimiento económico y no la oferta. En el caso de una economía desarrollada o semi-industrializada, operante a través de un esquema abierto o cerrado, la principal restricción al crecimiento económico de un país estará dada por la sostenibilidad de la cuenta corriente de la balanza de pagos.

En tanto que para Davidson (1990–1991), para entender el comportamiento del sector externo de una economía nacional a largo plazo y, por lo tanto, para comprender las dificultades para expandir sostenidamente a la economía, es necesario conocer y estudiar dado el ingreso del resto del mundo, las elasticidades

ingreso de las exportaciones e importaciones. Y para Thirlwall y Hussain (1982), son variables determinantes del crecimiento económico los términos de intercambio y los flujos netos de capital.

El modelo de crecimiento económico restringido por la balanza de pagos debido a Prebisch (1949), en su versión analítica simplificada desarrollada por Thirlwall (1979) y Thirlwall y Husain (1982) es el que será utilizado para estimar los determinantes de crecimiento para el caso de la economía mexicana en el periodo propuesto.

En la sección siguiente es presentado en forma general el modelo teórico para una economía nacional desarrollada o semi-industrializada como es el caso de México, y cómo será aplicada la metodología econométrica al primer modelo reducido. Será principalmente a través de las pruebas de cointegración de Johansen que se tratará de establecer una relación de equilibrio de largo plazo para las variables involucradas.

En el segundo modelo de estudio, se aplicará la teoría económica considerando en parte a Atesoglu (1997), con una versión extendida del modelo de crecimiento económico restringido por la balanza de pagos, para observar el impacto sobre la estimación del crecimiento económico de México, al incluir en el modelo ampliado junto al ingreso del resto del mundo, tanto los flujos netos de capital como los términos de intercambio.

3.2 MODELO TEÓRICO Y METODOLOGÍA ECONOMÉTRICA

El modelo de Thirlwall, consiste en un sistema de ecuaciones que representa una economía de dos bienes, pequeña y abierta¹⁶.

¹⁶ Moreno-Brid y Pérez (1999).

Se tiene la identidad básica de la balanza de pagos:

$$P_d X + EK_f = P_f EM \quad (3.1)$$

donde:

- P_d = precio de los bienes y servicios domésticos exportados en pesos¹⁷
- X = cantidad de exportaciones
- E = tipo de cambio en pesos por dólar
- K_f = flujos netos de capital en dólares
- P_f = precio de las importaciones en dólares
- M = cantidad de importaciones

Se define la participación de las exportaciones en los ingresos totales como:

$$\theta = \frac{P_d X}{P_d X + EK_f} \quad (3.2)$$

De la misma forma, se tiene la participación de los flujos netos de capital en los ingresos totales:

$$(1 - \theta) = \frac{EK_f}{P_d X + EK_f} \quad (3.3)$$

¹⁷ El modelo de crecimiento económico presentado lo es para cualquier economía nacional. En este punto se contextualiza a pesos mexicanos porque al momento de realizar la aplicación del modelo a la economía mexicana se utilizará esta información. Mientras tanto, se seguirá presentando el modelo en forma general.

Tanto (3.2) como (3.3) son la parte proporcional de la cuenta que resulta de las importaciones financiadas por las exportaciones así como de los flujos netos de capital. Puesto que teóricamente θ puede tomar un valor menor, igual o mayor que uno, es posible que el valor de $(1 - \theta)$ sea positivo, cero, o negativo.

Al tomar (3.1) y aplicar el logaritmo natural a cada sumando con sus respectivas participaciones (θ , $(1 - \theta)$ y 1) como exponentes, de la siguiente manera

$$L(P_d X)^\theta + L(K_f E)^{1-\theta} = L(P_f EM)$$

$$\theta[L(P_d) + L(X)] + (1 - \theta)[L(K_f) + L(E)] = L(P_f) + L(E) + L(M)$$

resulta la ecuación (3.1) en forma dinámica:

$$\theta(p_d + x) + (1 - \theta)(k_f + e) = p_f + e + m \tag{3.4}$$

donde las letras minúsculas representan las tasas de crecimiento de cada una de las variables (identificadas con su logaritmo natural). Por otra parte, se tiene la función convencional de demanda de las exportaciones:

$$X = \left(\frac{P_d}{P_f E} \right)^\eta (W^\pi) \tag{3.5}$$

donde:

- η = elasticidad precio de las exportaciones ($\eta < 0$)
- W = producto del resto del mundo
- π = elasticidad ingreso de las exportaciones ($\pi > 0$)

así como la función convencional de la demanda de las importaciones:

$$M = \left(\frac{P_f E}{P_d} \right)^\phi (Y^\xi) \quad (3.6)$$

donde:

- ϕ = elasticidad precio de las importaciones ($\phi < 0$)
- Y = ingreso doméstico
- ξ = elasticidad ingreso de las importaciones ($\xi > 0$)

Tomando (3.5) y (3.6) y aplicando logaritmo natural de ambos lados de las ecuaciones resulta

$$L(X) = L \left[\left(\frac{P_d}{P_f E} \right)^\eta (W^\pi) \right] = \eta [L(P_d) - L(P_f) - L(E)] + \pi L(W)$$

$$L(M) = L \left[\left(\frac{P_f E}{P_d} \right)^\phi (Y^\xi) \right] = \phi [L(P_f) + L(E) - L(P_d)] + \xi L(Y)$$

Por lo tanto, las ecuaciones dinámicas de demanda de las exportaciones e importaciones son:

$$x = \eta(p_d - p_f - e) + \pi w \quad (3.7)$$

$$m = \phi(p_f + e - p_d) + \xi y \quad (3.8)$$

Es necesario resolver el sistema de ecuaciones para la tasa de crecimiento económico y . Despejando y de (3.8):

$$y = [m - \phi(p_f + e - p_d)]/\xi \quad (3.9)$$

y despejando m de (3.4):

$$m = \theta(p_d + x) + (1 - \theta)(k_f + e) - p_f - e \quad (3.10)$$

Ahora, sustituyendo (3.7) en (3.10) y (3.10) en (3.9) se tiene:

$$y = \frac{[\theta(p_d + \{\eta(p_d - p_f - e) + \pi w\}) + (1 - \theta)(k_f + e) - p_f - e] - \phi(p_f + e - p_d)}{\xi}$$

Resolviendo los paréntesis:

$$y = \frac{\theta p_d + \theta \eta p_d - \theta \eta p_f - \theta \eta e + \theta \pi w + k_f + e - \theta k_f - \theta e - p_f - e - \phi p_f - \phi e + \phi p_d}{\xi}$$

Reordenando y sumando $p_d - p_d$:

$$y = \frac{\theta \pi w + k_f + e - p_d - \theta k_f - \theta e + \theta p_d + \theta \eta p_d - \theta \eta p_f - \theta \eta e + \phi p_d - \phi p_f - \phi e + p_d - p_f - e}{\xi}$$

Factorizando resulta:

$$y = \frac{\theta \pi w + (1 - \theta)(k_f + e - p_d) + (\theta \eta + \phi + 1)(p_d - p_f - e)}{\xi} \quad (3.11)$$

La ecuación (3.11) es la que representa a la tasa de crecimiento económico, siendo ésta restringida por su balanza de pagos. Se observa como una función lineal de:

- la tasa de crecimiento del producto del resto del mundo, w ,
- la tasa de crecimiento de los flujos netos de capital en pesos constantes $(k_f + e - p_d)$,
- la evolución de los términos de intercambio $(p_d - p_f - e)$ y,
- la elasticidad ingreso de las importaciones, ξ .

Las ponderaciones en (3.11) estarán dadas por:

- la participación de las exportaciones y los flujos netos de capital en los ingresos totales, θ y $(1 - \theta)$,
- las elasticidades ingreso y precio de las exportaciones, π y η , y,
- la elasticidad precio de las importaciones, ϕ .

Para la teoría de estos modelos no existe distinción entre los precios internos y los precios de las exportaciones por lo que el recíproco de los términos de intercambio es equiparable al tipo de cambio real.

En el caso de las elasticidades precio de las exportaciones e importaciones (η y ϕ , respectivamente), el signo de ambas será negativo pues el modelo general modifica convenientemente el orden de los precios internos y externos, y del tipo de cambio nominal.

Ahora, es posible reducir la expresión (3.11) al suponer que los términos de intercambio permanecen constantes (es decir, $p_d - p_f - e = 0$):

$$y_b = \frac{\theta\pi w + (1 - \theta)(k_f + e - p_d)}{\xi}$$

$$(3.12)$$

Y si además, se asume el supuesto de una cuenta corriente equilibrada (es decir, $\theta = 1$), se tiene que la ecuación (3.12) queda reducida de esta forma:

$$y_b = \frac{\pi w}{\xi} \quad (3.13)$$

En la expresión (3.13), la tasa de crecimiento de equilibrio es el resultado de la interacción de tres variables: las elasticidades ingreso de las exportaciones e importaciones, y la tasa de crecimiento del resto del mundo. En esta última ecuación se observa que dadas las elasticidades ingreso del sector externo, un mayor crecimiento de la economía mundial permitiría alcanzar una superior tasa de crecimiento del producto interno sin poner en riesgo el equilibrio de la cuenta corriente; o bien, dado el crecimiento económico del resto del mundo, un incremento de la relación de las elasticidades ingreso de las exportaciones e importaciones permitiría alcanzar una mayor tasa de crecimiento del producto de largo plazo.

Lo anterior, en materia de política económica implica el tratamiento adecuado de las elasticidades ingreso de las exportaciones y las importaciones: *“Una política económica exitosa es aquella que incrementa el valor de π y/o reduce el valor de ξ , relajando así la restricción ligada a la balanza de pagos para, ulteriormente, aumentar la tasa de crecimiento económico”*.¹⁸

Para la aplicación de la metodología econométrica, la ecuación (3.13) es replanteada en forma estocástica:

$$y_{b,t} = \alpha w_t + u_t$$

¹⁸ Bairam y Dempster (1991) p. 1720.

(3.14)

donde α representa la relación de las elasticidades ingreso de las exportaciones e importaciones ($\alpha = \pi/\xi$), y u_t son perturbaciones estocásticas que se distribuyen como ruido blanco gaussiano. El valor esperado de α debe ser positivo ya que, por definición, tanto π como ξ son positivas.

Si α fuera menor que uno (y por lo tanto $y_b < w$) el desempeño de la economía en estudio estaría por debajo de la evolución económica del resto del mundo. Y en el caso de que $\alpha = 1$, esto implicaría que ambas tasas de crecimiento de largo plazo serían similares ($y_b = w$).

Entonces, lo deseable para cualquier economía es que α esté lo más posible por arriba del uno. Queda claro, por lo tanto, que la variable que determina la proporción entre el crecimiento económico del país en estudio (y_b) y del resto del mundo (w), es la relación de las elasticidades ingreso de las exportaciones e importaciones (α).

Para el caso de este estudio, resulta adecuado utilizar la tasa de crecimiento del producto interno bruto de Estados Unidos (*Gross Domestic Product*, GDP) como variable proxy de la tasa de crecimiento del producto del resto del mundo (w).

La especificación general del modelo VAR a partir de la ecuación (3.14) es la que incluye de forma bivariada tanto al logaritmo natural del producto interno bruto de México como el de Estados Unidos. Si

$$\mathbf{y} = \begin{bmatrix} L(Y_{MEX}) \\ L(Y_{EEUU}) \end{bmatrix}$$

entonces de forma explícita:

$$\begin{aligned} \begin{bmatrix} L(Y_{MEX,t}) \\ L(Y_{EEUU,t}) \end{bmatrix} &= \begin{bmatrix} \chi_0 \\ \gamma_0 \end{bmatrix} + \begin{bmatrix} \chi_{1,1} & \chi_{2,1} \\ \gamma_{1,1} & \gamma_{2,1} \end{bmatrix} \begin{bmatrix} L(Y_{EEUU,t-1}) \\ L(Y_{MEX,t-1}) \end{bmatrix} + \begin{bmatrix} \chi_{1,2} & \chi_{2,2} \\ \gamma_{1,2} & \gamma_{2,2} \end{bmatrix} \begin{bmatrix} L(Y_{EEUU,t-2}) \\ L(Y_{MEX,t-2}) \end{bmatrix} \\ &+ \dots + \begin{bmatrix} \chi_{1,p} & \chi_{2,p} \\ \gamma_{1,p} & \gamma_{2,p} \end{bmatrix} \begin{bmatrix} L(Y_{EEUU,t-p}) \\ L(Y_{MEX,t-p}) \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} \end{aligned}$$

o equivalentemente:

$$L(Y_{MEX,t}) = \chi_0 + \sum_{i=1}^p \chi_{1,i} L(Y_{EEUU,t-i}) + \sum_{i=1}^p \chi_{2,i} L(Y_{MEX,t-i}) + u_{1,t} \quad (3.15)$$

$$L(Y_{EEUU,t}) = \gamma_0 + \sum_{i=1}^p \gamma_{1,i} L(Y_{EEUU,t-i}) + \sum_{i=1}^p \gamma_{2,i} L(Y_{MEX,t-i}) + u_{2,t} \quad (3.16)$$

donde Y representa las variables producto interno bruto en niveles y L a su logaritmo natural, y la longitud de las sumas (p) depende los criterios de información para la longitud del rezago (vistos en la sección 1.5.1).

Esta forma funcional doble logarítmica para el modelo VAR resulta válida tomando en cuenta que los coeficientes representan las elasticidades de las variables¹⁹, y tanto χ como γ simbolizan la relación de las elasticidades ingreso de las exportaciones e importaciones²⁰.

En la siguiente sección se hará el contraste empírico del modelo de crecimiento económico restringido por la balanza de pagos empleando la metodología para estimar modelos VAR con cointegración.

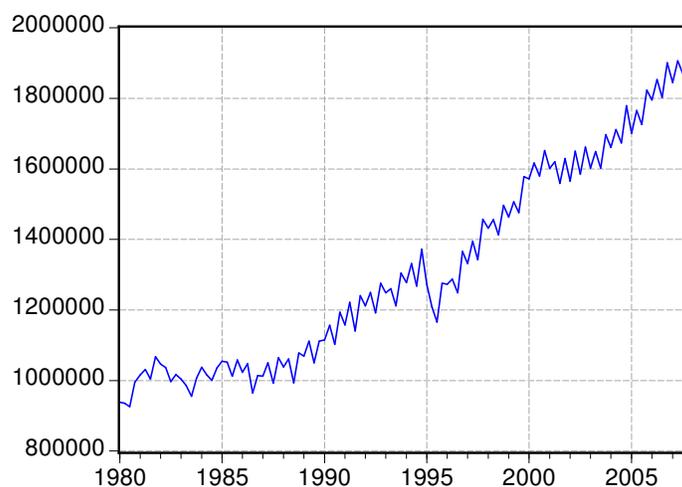
¹⁹ Maddala (2001).

²⁰ Guerrero (2004).

3.3 ANÁLISIS EMPÍRICO BASADO EN EL MODELO REDUCIDO

El análisis empírico comienza con la exploración gráfica del comportamiento del producto interno bruto de México (Y_{MEX} , en millones de pesos a precios de 1993, gráfica 1) y de los Estados Unidos (Y_{EEUU} , en miles de millones de dólares encadenados del 2000, gráfica 2)²¹, de sus tasas de crecimiento (gráficas 3 y 4) y los datos cruzados de ambas variables macroeconómicas (gráfica 5). En todo el estudio, tanto para la obtención de gráficos como de las tablas de resultados de las pruebas se hace uso del programa econométrico E-views 5.1.

GRÁFICA I

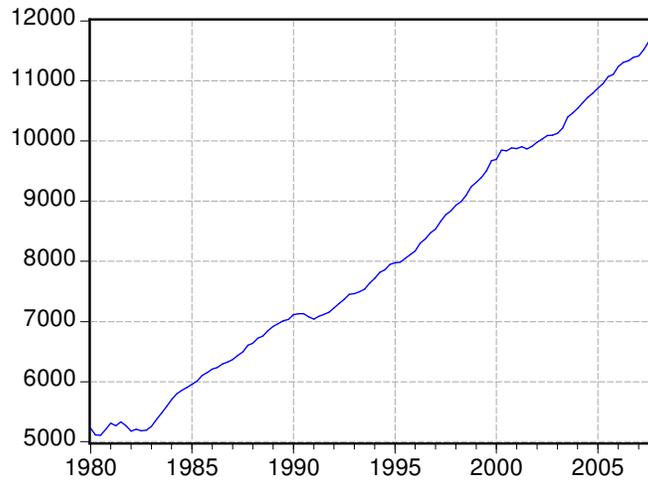


PIB de México, de 1980 a 2007.²²

²¹ Respecto al por qué se utilizan estas bases de datos con estos años base, consultar las notas a pie de página donde aparecen estas series macroeconómicas en el apéndice C.

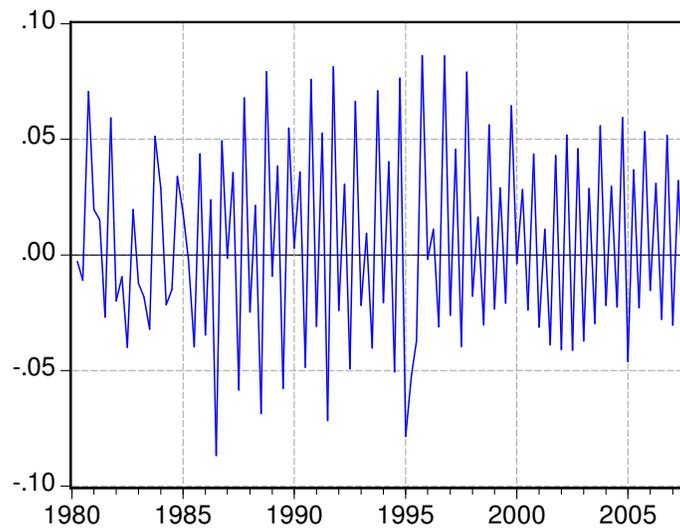
²² En el Apéndice C se encuentran las tablas con los valores de las variables macroeconómicas construidas y utilizadas a lo largo de este trabajo, así como sus fuentes.

GRÁFICA 2



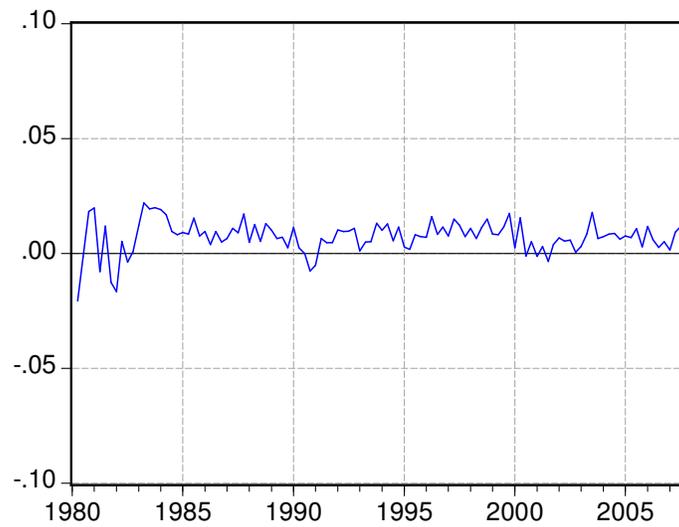
PIB de EEUU, de 1980 a 2007.

GRÁFICA 3



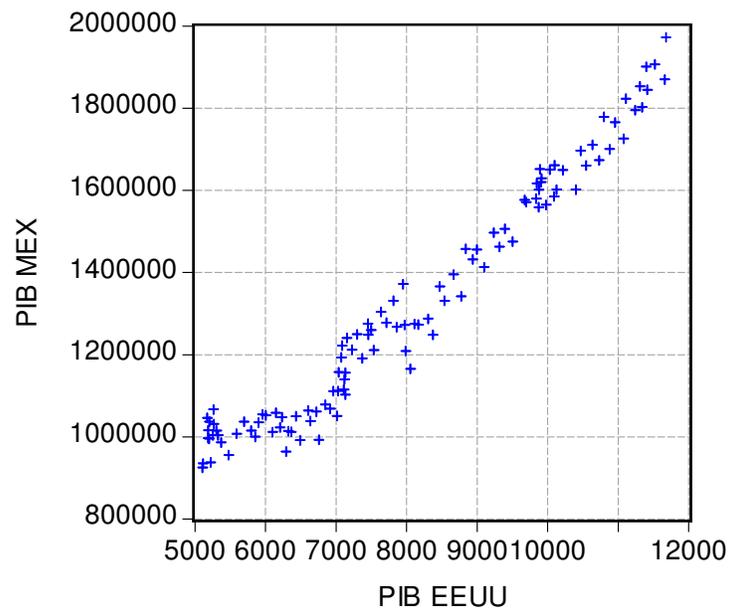
Tasas de crecimiento del PIB de México, de 1980 a 2007.

GRÁFICA 4



Tasas de crecimiento del PIB de EEUU, de 1980 a 2007.

GRÁFICA 5



PIB de México vs. PIB de EEUU, de 1980 a 2007.

De las gráficas 1 a 5 es posible observar:

- Las tendencias ascendentes del PIB de ambos países. Se tiene que entre 1980 y 2007 las tasas de crecimiento medias anuales de Estados Unidos y México fueron de 3.055% y 2.628% respectivamente²³, lo que concuerda en el parecido de las pendientes de ambos productos en las dos primeras gráficas. Ahora, si bien a lo largo del periodo estudiado se advierte una mayor tasa de crecimiento para el PIB de EU, resulta interesante que para intervalos de tiempo más largos (y en su caso anteriores al de este estudio), estas tasas de crecimiento presentan entre sí un comportamiento inverso²⁴.
- Las desaceleraciones económicas de la economía estadounidense a principios de las décadas de los ochenta, noventa y del año 2000, así como la volatilidad macroeconómica mexicana durante las décadas de los ochenta y noventa (particularmente entre 1994 y 1995), a partir de las gráficas 3 y 4.
- La significativa asociación positiva entre ambos productos. La gráfica 5 sugiere una relación lineal y relativamente estable entre los productos internos de ambos países.

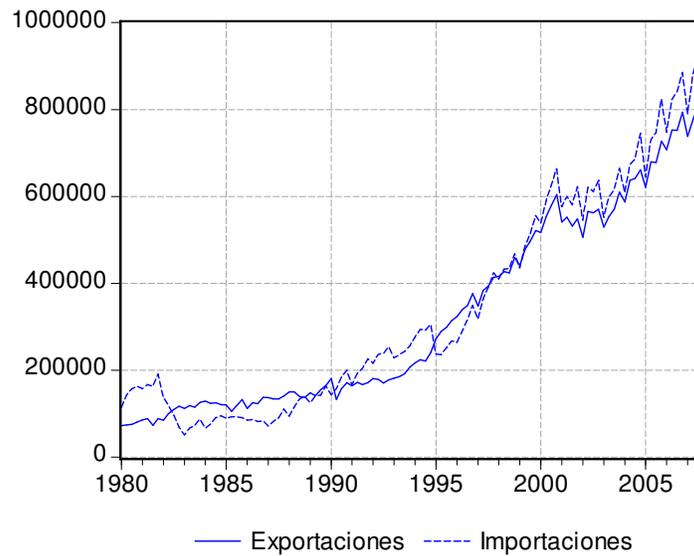
Ahora se presenta el comportamiento gráfico de las exportaciones e importaciones totales de México en millones de pesos a precios de 1993 (XT y MT , respectivamente), a lo largo del periodo en estudio (gráfica 6), así como los datos cruzados de ambas variables (gráfica 7), la participación de ambas en el producto interno bruto de México en tantos por ciento ($100 \times (\frac{XT}{Y_{MEX}})$ y $100 \times (\frac{MT}{Y_{MEX}})$),

²³ Fuente: elaboración propia con base en la información del Apéndice C.

²⁴ Guerrero (2004).

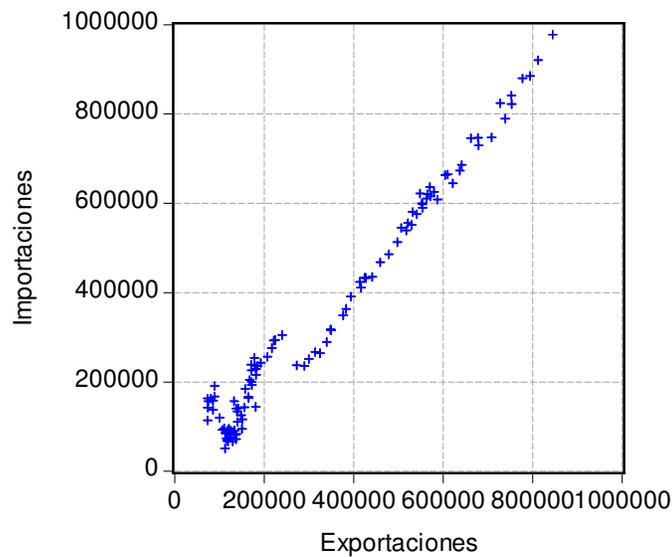
respectivamente, gráfica 8), y el cociente entre estas dos participaciones en tantos por uno (de las exportaciones respecto a las importaciones, gráfica 9).

GRÁFICA 6



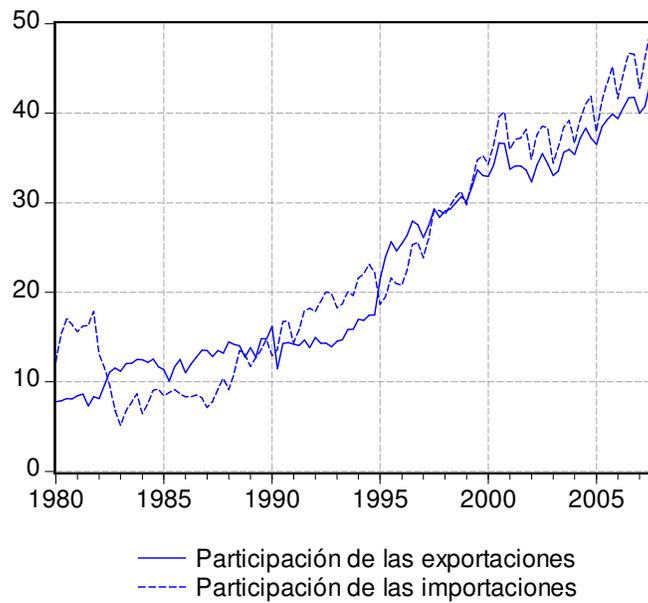
Exportaciones e importaciones totales de México, de 1980 a 2007.

GRÁFICA 7



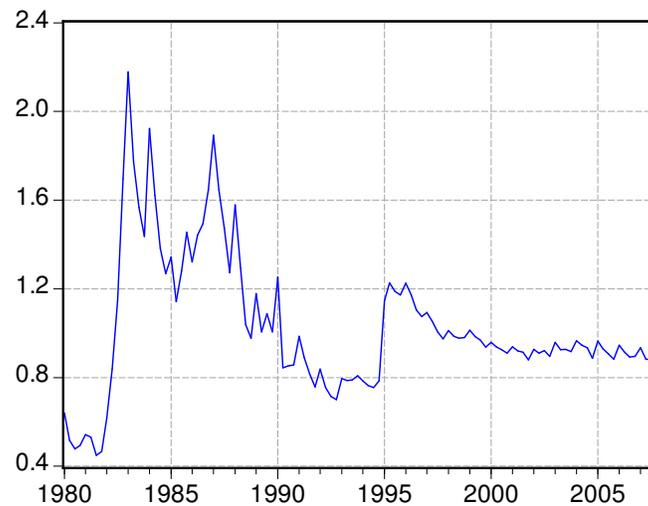
Exportaciones totales vs. importaciones totales de México, de 1980 a 2007.

GRÁFICA 8



Participación de las exportaciones e importaciones totales de México en su PIB en tantos por ciento, de 1980 a 2007.

GRÁFICA 9



Cociente de la participación de las exportaciones totales sobre las importaciones totales de México en su PIB en tantos por uno, de 1980 a 2007.

En el conjunto de las gráficas anteriores sobresale la siguiente información:

- La explosión y contracción de las importaciones totales durante la primera parte de la década de los ochenta.
- El acelerado crecimiento de las exportaciones e importaciones totales en el periodo estudiado –tanto en niveles como respecto al producto interno– y la retroalimentación entre ambas variables.
- En forma especial en la gráfica 9, la evolución del cociente de las exportaciones e importaciones. Si bien fue relativamente estable de los años cincuenta a finales de los setenta, se observa un gran salto hasta 2.18 en el primer trimestre del año 1983, y su posterior disminución hasta ubicarse por debajo de uno a partir del segundo trimestre de 1999 hasta el final del periodo.

Con el análisis gráfico inicial hecho, y previo a la estimación de los modelos VAR con cointegración es probado el orden de integración de las variables logaritmo del producto interno bruto de las economías estadounidense y mexicana. Fue aplicada la prueba Dickey–Fuller aumentada (DFA), para verificar el posible rechazo de la hipótesis nula de que cada serie tiene una raíz unitaria.

Cuadro 1
Pruebas de raíces unitarias²⁵
Producto interno bruto de EEUU y México 1980–2007

País	Variable	ADF
EEUU	L(Yeeuu)	-0.90791
	$\Delta(L(Yeeuu))$	-5.743323*
México	L(Ymex)	1.174323
	$\Delta(L(Ymex))$	-3.316293**

Hipótesis nula: la variable sometida a prueba tiene una raíz unitaria

(*) Rechazo de la hipótesis nula con un nivel de significancia del 1%

(**) Rechazo de la hipótesis nula con un nivel de significancia del 5%

Nota: En la prueba ADF (Dickey–Fuller Aumentada) el número de rezagos fue según el criterio de información de Schwartz, incluyendo un intercepto.

El cuadro 1 indica que el orden de integración respecto a las variables es en ambos casos uno (es decir, ambas series son $I(1)$), lo que desde un punto de vista macroeconómico tiene sentido, y refuerza (al tratarse de un modelo VAR balanceado, donde el primer sumando de (2.31) es $I(0)$) los resultados de las pruebas de Johansen que fueron revisadas en el apartado 2.6.2 de este estudio.

Para el estudio empírico de una relación de equilibrio de largo plazo entre las economías mexicana y norteamericana fueron estimados con la información disponible algunos modelos VAR con cointegración. Siguiendo los requisitos de la

²⁵ Una descripción de la prueba construida por Dickey y Fuller (1979) para probar si una variable tiene una raíz unitaria y utilizada aquí es presentada en el apéndice A.1. Los resultados completos obtenidos con Eviews 5.1 se presentan de los cuadros B1 al B4 en el Apéndice B.

metodología econométrica moderna²⁶ para las variables incluidas, primeramente se verificó su congruencia estadística mediante las pruebas de autocorrelación serial LM, de normalidad de Jarque–Bera, y las dos pruebas de heteroscedasticidad White (con y sin términos cruzados)²⁷. Los resultados fueron los siguientes:

- En el caso de la prueba de autocorrelación serial LM, se verifica que el número de rezagos conveniente para poder asumir que se trabaja con un modelo sin la presencia de correlación es 5.
- Al aplicar la prueba de normalidad de Jarque–Bera el alto valor del estadístico conjunto indica el no rechazo de la hipótesis nula de que los residuos muestrales se comportan normales multivariados.
- Y con respecto a la heteroscedasticidad, la prueba de White que no toma en consideración los términos de productos cruzados no rechaza la hipótesis nula de no heteroscedasticidad, es decir, existe evidencia empírica de homoscedasticidad en los residuos muestrales. Si bien la otra prueba sí rechaza la hipótesis nula.²⁸

A continuación se pueden observar los resultados sobre el modelo VAR estimado por el método de Johansen para probar cointegración, a partir de (3.14).

Como los resultados de la metodología de Johansen son sensibles al número de rezagos a incluir en el VAR, se realizó la prueba del estadístico *LR* de razón de verosimilitud para determinar el número de rezagos óptimo²⁹, que fue 5.

²⁶ Spanos (1986).

²⁷ La descripción de estos contrastes es presentada en el apéndice A.2, y los cuadros con los resultados pueden consultarse en el apéndice B (cuadros B5 al B8).

²⁸ El rechazo de la hipótesis nula para la prueba de White (con términos cruzados) puede atribuirse a la estimación hecha con un alto número de grados de libertad.

²⁹ Ver el cuadro B.9 en el apéndice B. Otros criterios (no incluidos en las descripciones de contrastes de este trabajo) para la elección del número de rezagos óptimo son computados al mismo tiempo por Eviews. En este caso todos coinciden en el mismo valor.

Cuadro 2

Prueba de Johansen del modelo reducido, vectores de cointegración normalizados
L(Ymex) y L(Yeeuu) 1980–2007

Número de rezagos	$\alpha = \pi / \xi$
2	1.3305
3	1.8140
4	2.1608
5	1.1111

Notas: Las pruebas de la traza y del valor propio máximo indicaron la existencia de cointegración entre las variables en todos los casos observados.

Es necesario recordar que en la obtención de las estimaciones de cada parámetro α en las ecuaciones de cointegración, se invierten sus signos al final.

Siendo cinco el número de rezagos óptimo según el criterio de información *LR*, ya se tenía que para este número se encuentra un modelo VAR congruente estadísticamente hablando (particularmente al realizar la prueba de autocorrelación serial *LM*). Adicionalmente se ha probado el modelo para un número de rezagos menor a cinco³⁰, y el hecho de que exista evidencia de cointegración aún en estos casos, hace que sea coherente la conclusión respecto al movimiento conjunto de ambas variables en el largo plazo.

Respecto a la información contenida en el cuadro 2 es necesario comentar que en la ecuación de cointegración no se incluyó una intersección. A este respecto, según la prueba de cointegración de Johansen, esto no fue necesario, al tomar en consideración la ecuación del modelo reducido (3.14). Además, que los valores de la

³⁰ Los resultados completos de estas pruebas se localizan en el apéndice B. Cuadros del B.10 al B.13.

relación de las elasticidades ingreso de las exportaciones e importaciones de la economía mexicana $-\alpha$ – sirven para calcular las tasas de crecimiento económico compatibles con un saldo cero en la cuenta corriente, como se verá a continuación.

La expresión resultante para la elaboración de un resumen de la actividad económica de México para el periodo 1980–2007 es la obtenida de (3.14):

$$\widehat{y}_{b,t} = \widehat{\alpha}w_t$$

Resulta que el crecimiento medio anual del producto observado³¹ ($y = 2.6281\%$) fue ligeramente menor que la cifra obtenida del modelo reducido de crecimiento económico restringido por la balanza de pagos ($y_b = 3.3949\%$), es decir,

$$\widehat{y}_{b,t} = (1.1111)(0.030554) = 0.033949$$

de la cantidad obtenida por la multiplicación de la tasa media anual de crecimiento económico de Estados Unidos³² ($w = 3.0554\%$) por la estimación de la relación entre las elasticidades ingreso de las exportaciones e importaciones de la economía mexicana ($\widehat{\alpha} = 1.1111$). Es así que el resultado del modelo estimado realiza una aproximación adecuada de la tasa de crecimiento observada.

Por otra parte, Guerrero (2004) en su estudio para las economías de México y Estados Unidos durante el periodo 1929–2003, realizó otra prueba estadística elaborada por Johansen en 1992, y obtuvo que la economía estadounidense no resultó exógenamente débil³³ respecto a la mexicana, lo cual fue inesperado. En este sentido, para el modelo obtenido, que deja de lado a otras muchas variables asociadas a ambas economías, la prueba estadística implica el tratamiento de los PIB

³¹ Elaboración propia con base en la información disponible del PIB de México en el apéndice C.

³² Elaboración propia con base en la información disponible del PIB de Estados Unidos en el apéndice C.

³³ Concepto revisado en la Sección 2.5 de este trabajo.

de ambos países como variables endógenas, es decir, como parte de un sistema. Así que se tiene una retroalimentación –o causalidad de ida y vuelta– entre las variables.³⁴

Respecto a una relación de cointegración a largo plazo entre estas variables el mismo autor realiza una partición de su periodo de estudio en cinco subperiodos para la obtención de la razón entre las elasticidades ingreso de las importaciones y las exportaciones, y encuentra que los valores estimados revelan su relativa estabilidad. Y subraya que *“por el valor unitario del parámetro estimado, parece que una característica bajo el nuevo modelo de desarrollo es que nuestro ritmo de crecimiento económico se aproxime mucho más al de Estados Unidos”*³⁵.

Y si bien el modelo de crecimiento económico restringido por la balanza de pagos es un planteamiento diseñado originalmente bajo una perspectiva de largo plazo, las estimaciones del parámetro α sirven para la aproximación de la evolución económica del periodo estudiado aquí.

Al sustraer la tasa de crecimiento anual observada con la obtenida mediante el modelo restringido resulta que la diferencia es negativa (-0.7215%), y se tiene una confirmación de cierta regularidad empírica a largo plazo. Lo anterior significa: *“la economía mexicana ha crecido, en general y en promedio, ligeramente por debajo de la tasa asociada al saldo cero en cuenta corriente”*.³⁶

³⁴ Un resultado interesante relacionado con esta idea es el que se puede observar en el cuadro B.14 del apéndice B, obtenido de la prueba de causalidad de Granger. Hecho el contraste para las variables $L(Ymex)$ y $L(Yeeuu)$ se tiene el rechazo de la hipótesis nula de no causalidad de $L(Ymex)$ respecto a $L(Yeeuu)$, es decir, $L(Ymex)$ causa en el sentido de Granger a $L(Yeeuu)$, y no en sentido contrario.

³⁵ Guerrero (2004).

³⁶ Ídem.

Igualmente con la estimación de un modelo VAR con cointegración, Moreno–Brid (1999) aplicó la “Ley de Thirlwall” a la economía mexicana para la segunda mitad del siglo XX³⁷, empleando la ecuación estocástica

$$L(Y_t) = \beta_0 + \beta_1 L(XT_t) + u_t \tag{3.17}$$

donde se tienen los logaritmos naturales tanto del producto interno bruto como de las exportaciones totales de México, y asimismo β_0 es el intercepto y $\beta_1 = 1/\xi$, es decir, β_1 es el recíproco de la elasticidad ingreso de las importaciones, y u_t es una perturbación aleatoria de ruido blanco gaussiano. A partir de la información disponible, y haciendo la regresión propuesta por (3.17) para obtener β_1 y por lo tanto ξ , se puede a su vez lograr una aproximación (por el valor de $\hat{\alpha}$) de las elasticidades ingreso de las exportaciones (π).³⁸

Es así que

$$\widehat{L(Y_t)} = \widehat{\beta}_0 + \widehat{\beta}_1 L(XT_t) = (10.53267) + (0.284223)L(XT_t)$$

por lo que $\hat{\xi} = \frac{1}{\hat{\beta}_1} = 3.5184$.

Con $\pi = \alpha\xi$, entonces la elasticidad ingreso de las exportaciones aproximada ($\hat{\pi}$) es igual a $3.9093 = (1.1111)(3.5184)$.

Al realizar un comparativo entre estas estimaciones de ξ y π con las que resultaron en los estudios citados de Guerrero y Brid–Pérez para distintos periodos a lo largo del siglo pasado, parece correcto deducir que los resultados de los más recientes modelos de desarrollo implantados en México son negativos en términos

³⁷ El primero en realizar este procedimiento fue Atesoglu (1997).

³⁸ Los resultados completos de esta regresión se encuentran en el cuadro B.15, del apéndice B.

del levantamiento de la restricción externa al crecimiento, pues se aprecia en el tiempo una mayor evolución en los valores de ξ respecto a los de π , y la relación entre éstas en periodos como el aquí estudiado (o más inmediatos) arrojan conservadores valores de α (en ocasiones por debajo de uno), que es justamente lo contrario a lo que debe buscarse en materia de política económica.

Como afirma Guerrero (2004) en su análisis de estas relaciones entre las elasticidades ingreso de las importaciones y las exportaciones de México país para los últimos 70 años del siglo pasado: *“históricamente la tasa de crecimiento de la economía mexicana ha perdido dinamismo, y actualmente presenta un alineamiento respecto al comportamiento de la economía norteamericana.”*

3.4 ANÁLISIS EMPÍRICO BASADO EN EL MODELO AMPLIADO

En esta sección se extenderá el estudio a un modelo que considere todo el conjunto de las variables incluidas en la ecuación para la determinación del crecimiento económico (3.11):

$$y = \frac{\theta\pi w + (1 - \theta)(k_f + e - p_d) + (\theta\eta + \phi + 1)(p_d - p_f - e)}{\xi}$$

En este punto se hace uso parcial del trabajo de Atesoglu (1997), que en su trabajo de forma incorrecta igualó en y a πw con las exportaciones totales³⁹. Se tiene la siguiente formulación econométrica para (3.11):

$$y_t = \beta_1 w_t + \beta_2 (k_{f,t} + e_t - p_{d,t}) + \beta_3 (p_{d,t} - p_{f,t} - e_t) + u_t$$

³⁹ Guerrero (2004).

$$(3.18)$$

donde

$$\beta_1 = \frac{\theta\pi}{\xi}$$

$$(3.19)$$

$$\beta_2 = \frac{(1 - \theta)}{\xi}$$

$$(3.20)$$

$$\beta_3 = \frac{\theta\eta + \phi + 1}{\xi}$$

$$(3.21)$$

Una diferencia clave entre las ecuaciones (3.14) y (3.18) es que no se incorpora ningún supuesto para reducir (3.18), y se opera sobre la definición contable de la balanza de pagos, y de las funciones de demanda de exportaciones e importaciones. Es en (3.18) que el crecimiento económico depende entonces de la evolución de la actividad productiva del resto del mundo w , así como de los flujos netos de capital en pesos constantes y de los términos de intercambio.

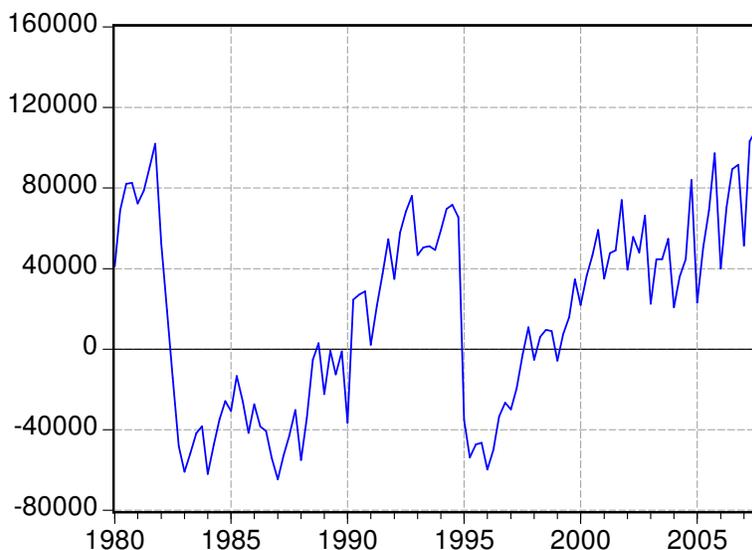
Lo que se puede esperar de los signos para los β_i en (3.18) es lo siguiente:

- β_1 será positivo por definición,
- β_2 al depender de $1 - \theta$, no es posible adelantar su signo. Si el país es exportador neto de capitales, entonces $1 - \theta < 0$. El sentido de la desigualdad será inverso si el país en cuestión es importador.

- Suponiendo que $\theta > 0$, entonces el signo de β_3 depende del valor que tomen las elasticidades precio de las exportaciones e importaciones. Por definición según (3.5) y (3.6), $\eta < 0$ y $\phi < 0$. Si su suma en valor absoluto es mayor que uno⁴⁰, entonces $\beta_3 < 0$. Si la suma da un valor en sentido inverso entonces $\beta_3 > 0$. Por lo que el efecto de los términos de intercambio sobre el comportamiento del producto es teóricamente ambiguo.

Por la inclusión de las corrientes netas de capital en el modelo, primeramente se muestra de forma gráfica su evolución en el periodo estudiado en millones de pesos a precios de 1993 (K , gráfica 10), su participación en términos de porcentaje en el PIB ($100 \times \left(\frac{K}{Y_{MEX}}\right)$, gráfica 11), así como la gráfica cruzada de estas dos variables macroeconómicas (gráfica 12).

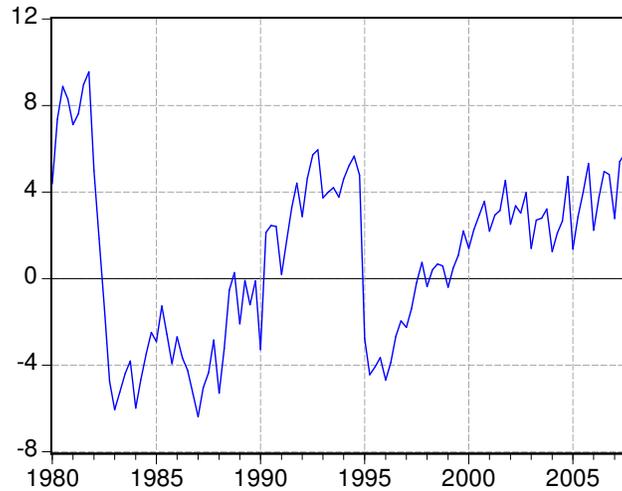
GRÁFICA 10



Flujos netos de capital, de 1980 a 2007.

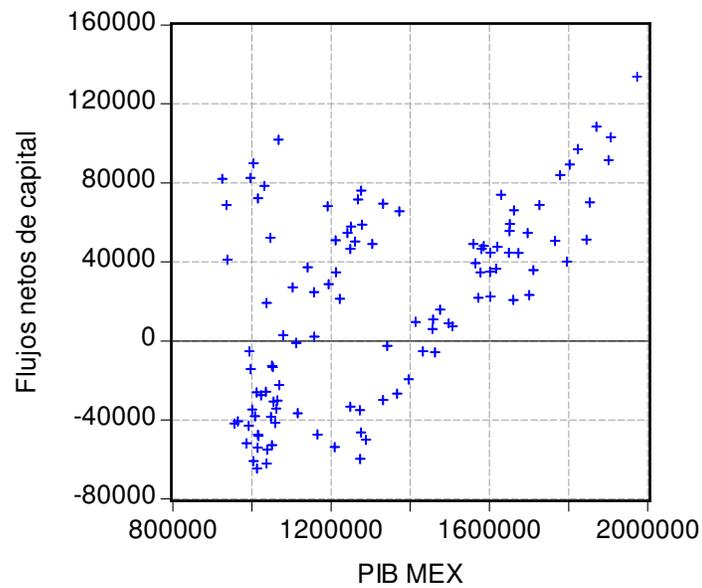
⁴⁰ El **teorema o condición de Marshall-Lerner** demuestra que para llevar a cabo la devaluación de una divisa y que ésta tenga un impacto positivo en la balanza comercial, la suma de las elasticidades precio de las importaciones y las exportaciones ha de ser, en valor absoluto, superior a 1.

GRÁFICA II



Participación de los flujos netos de capital sobre el PIB de México en tantos por ciento, de 1980 a 2007.

GRÁFICA I2



Flujos netos de capital vs. PIB de México, de 1980 a 2007.

Para entender el comportamiento de los flujos netos de capital es necesario recordar que hasta después del quiebre del sistema Bretton Woods aumentó significativamente la movilidad del capital a nivel internacional⁴¹, así como el hecho de que se ha dado la instauración de la liberalización de los mercados financieros, tanto en países desarrollados como semi-industrializados. Internamente, en materia de políticas de inversión extranjera en el periodo estudiado, si bien para los años ochenta, México tenía uno de los regímenes más restrictivos del mundo⁴², se dió la flexibilización de la legislación vigente por parte de la Comisión Nacional de Inversiones Extranjeras en 1984. Y la aparición en diciembre de 1993 de la Ley de Inversión Extranjera fue clave en el nuevo modelo de desarrollo.⁴³

En el periodo de estudio, México ha sido tanto un país de origen como de destino de capital, y puede ser considerado como un país exportador neto de capitales, con todo y la cifra récord para el último trimestre de 1981 de 9.55% para la participación de las corrientes netas de capital sobre el PIB, y los repuntes para 1992 y 2007, destacando lapsos de gran volatilidad así como de pronunciadas caídas. Como consecuencia de esto, en la gráfica 12 la dispersión de los puntos no sugiere ninguna asociación entre los flujos netos de capital y el producto interno bruto.

En el cuadro siguiente se da el resultado de la prueba Dickey-Fuller Aumentada sobre los flujos netos de capital y resalta (congruentemente con la gráfica 10) el que la variable K es estacionaria, $I(0)$, lo cual tiene sentido económico, pues solamente su acumulación generaría un comportamiento no estacionario. Es necesario señalar que esta variable macroeconómica será

⁴¹ Puyana (2003).

⁴² Blanco (1994).

⁴³ Gurría (1994)

considerada en niveles pues toma valores negativos a lo largo de varios trimestres en el periodo estudiado, e imposibilita considerar su logaritmo natural.

Cuadro 3
Prueba de raíz unitaria⁴⁴
Flujos netos de capital 1980–2007

	ADF
K	-2.727501*

Hipótesis nula: K tiene una raíz unitaria

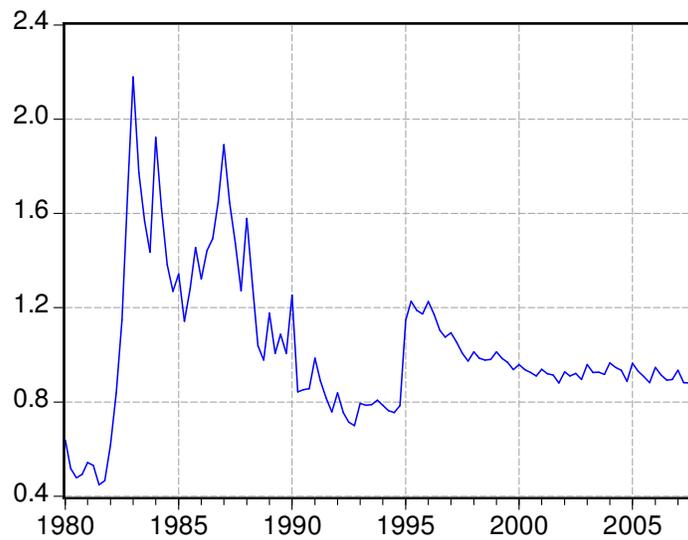
(*) Rechazo de la hipótesis nula con un nivel de significancia del 10%

Nota: En la prueba ADF el número de rezagos fue según el criterio de información de Schwartz, incluyendo un intercepto.

En las gráficas 13 y 14 se observa que la participación de las exportaciones totales en los ingresos totales de la balanza de pagos ($\theta = \frac{XT}{MT}$) permanece positivo, no así el valor de la participación de los flujos netos de capital en los ingresos totales de la balanza de pagos ($1 - \theta = \frac{K}{MT}$), que es negativo en algunos años y positivo en otros. Entre los años 1980 y 2007 sus valores promedio fueron 1.0305 y 0.0305 respectivamente. Se observa que en la gráfica 14 para $(1 - \theta)$ existe una ligera tendencia ascendente.

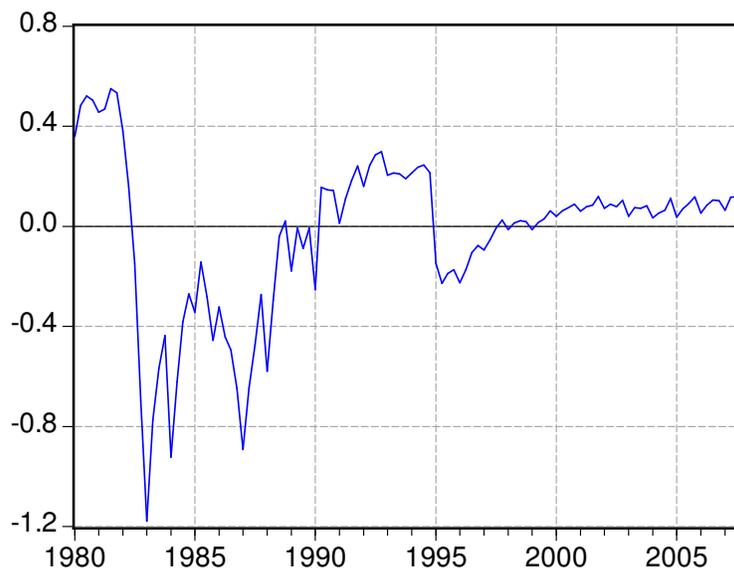
⁴⁴ En el cuadro B.16 del apéndice B se encuentra la salida original del programa Eviews.

GRÁFICA 13



Participación de las exportaciones totales en los ingresos totales de la balanza de pagos (θ), de 1980 a 2007.

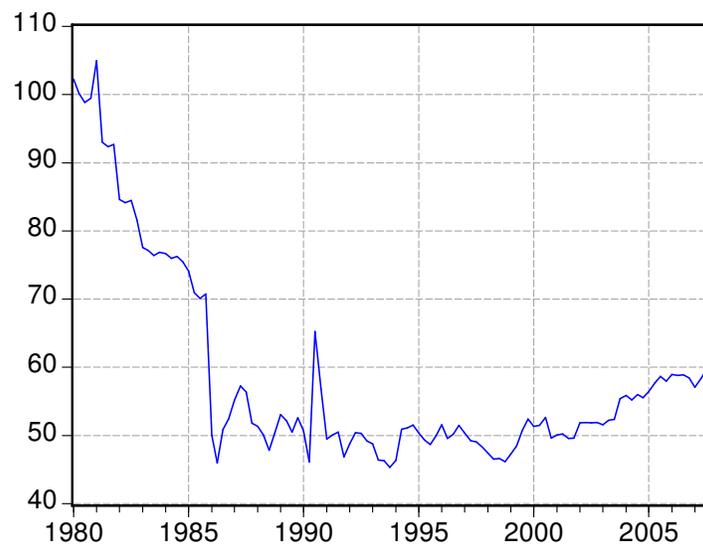
GRÁFICA 14



Participación de los flujos netos de capital en los ingresos totales de la balanza de pagos ($1 - \theta$), de 1980 a 2007.

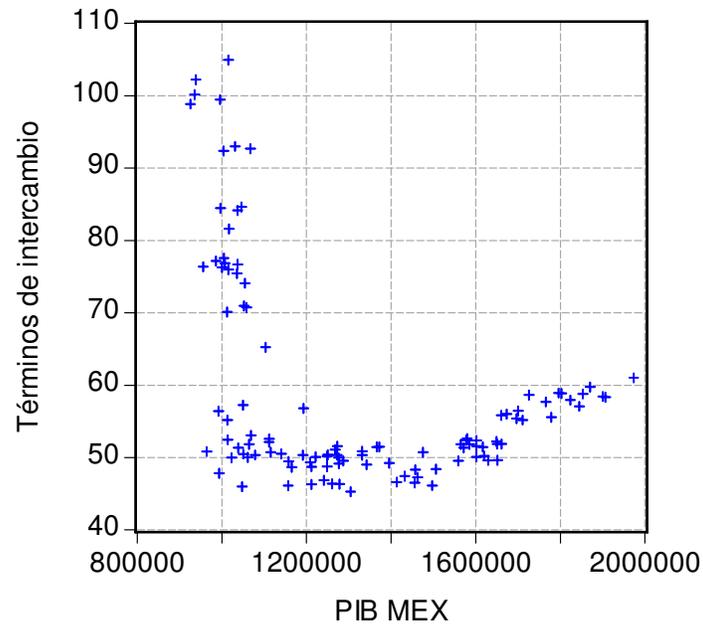
De las variables incluidas en el modelo ampliado, sólo resta englobar los términos de intercambio (*TI*). En la gráfica 15 se presenta su evolución para el periodo 1980–2007. Y en la gráfica 16 se alcanza a observar una cierta asociación negativa entre esta variable y el producto interno bruto de México.

GRÁFICA 15



Términos de intercambio, de 1980 a 2007.

GRÁFICA 16



Términos de intercambio vs. PIB de México, de 1980 a 2007.

El cuadro 4 muestra que el logaritmo natural de los términos de intercambio como podría esperarse a partir de su evolución en el periodo estudiado presentan evidencia de ser una serie estacionaria, $I(0)$.

Cuadro 4
Prueba de raíz unitaria⁴⁵
Términos de intercambio 1980–2007

	ADF
L(TI)	-2.758707*

Hipótesis nula: L(TI) tiene una raíz unitaria

(*) Rechazo de la hipótesis nula con un nivel de significancia del 10%

Nota: En la prueba ADF el número de rezagos es según el criterio de información de Schwartz, incluyendo un intercepto.

Para el caso del modelo econométrico ampliado que incluye a $L(Y_{MEX})$, $L(Y_{EEUU})$, K y $L(TI)$, las pruebas de congruencia estadística⁴⁶ (de correlación serial, de normalidad y de heteroscedasticidad de los residuos) no son tan satisfactorias como lo fueron las del modelo reducido, particularmente en la prueba de White donde se tiene el rechazo de la hipótesis nula y por lo tanto la presencia de heteroscedasticidad en los residuos, y esto dificulta el encontrar un movimiento conjunto de las variables a largo plazo como supone la cointegración, aunque al ser aplicada la prueba de Johansen a este modelo VAR ampliado resulta que las variables sí son cointegradas. De esta manera se tienen estimaciones de los coeficientes β_1 , β_2 y β_3 .

⁴⁵ Cuadro B.17 del apéndice B.

⁴⁶ Cuadros B.18 al B.20 del apéndice B. Y los criterios de aceptación o rechazo de las hipótesis de cada prueba en el apéndice A.

Cuadro 5

Prueba de Johansen del modelo ampliado, vectores de cointegración normalizados⁴⁷
L(Ymex), L(Yeeuu), K y L(TI) 1980–2007

Año	Periodicidad	Observaciones	β_1	β_2	β_3
1980–2007	Trimestral	106 después de ajustes	1.6736	6.87E-05	-3.7903

Notas: Las pruebas de la traza y del valor propio máximo indicaron la existencia de cointegración entre todas las variables.

Es necesario recordar que en la obtención de las estimaciones de los parámetros β_i en las ecuaciones de cointegración, se invierten sus signos al final.

Del cuadro 5 resulta que si bien el parámetro de los flujos de capital (β_2) es casi cero como resultado del valor reducido de $(1 - \theta)$, el hecho de que su signo sea positivo es revelador, pues su pequeña magnitud desde un punto de vista macroeconómico tiene sentido, ya que por un lado los inlfujos de capital externo representan una fuente alternativa de financiamiento de inversión, pero por el otro, existe evidencia acerca de una relación negativa con las importaciones.⁴⁸ En el caso del estudio de Guerrero (2004) para el intervalo de 1929–2003, se tiene que los flujos netos de capital presentan este alineamiento al resto de las variables hasta el último subperiodo, que va de 1996 a 2003, y que no es posible hablar de relaciones de cointegración para subperiodos anteriores.

⁴⁷ El criterio de razón de verosimilitud LR para la elección del número de rezagos óptimo (que fue cinco) para este VAR se puede consultar en el cuadro B.21 del apéndice B. La información completa de la prueba del cuadro 5 se encuentra en el cuadro B.22.

⁴⁸ López (1996).

El valor estimado de β_1 se encuentra en la cercanía del valor estimado de α , lo cual nos da una buena aproximación al modelo reducido. Sin embargo, el elevado valor negativo de β_3 se aleja (y por mucho) de los valores obtenidos para este ponderador en otros estudios⁴⁹ (acotados entre -1 y 0) en periodos relativamente cercanos al expuesto aquí, lo cual dificulta la valoración de su utilidad en todo el contexto.

Ahora bien, excluyendo las corrientes netas de capital del modelo ampliado, se obtienen mejores resultados en lo que respecta al movimiento conjunto de las variables que figuran en (3.18).

Volviendo a verificar la congruencia estadística del modelo⁵⁰, esta vez resulta satisfactoria, pues aunado a la no correlación serial y a la normalidad de los residuos, es posible no rechazar la hipótesis nula de la prueba de White, y por lo tanto se puede hablar de no heteroscedasticidad en este modelo ampliado donde se omiten los flujos netos de capital.

Cuadro 6
Prueba de Johansen, vectores de cointegración normalizados
Log(Ymex), Log(Yeeuu) y Log(TI) 1980-2007

Año	Periodicidad	Observaciones	β_1	β_3
1980-2007	Trimestral	106 después de ajustes	1.2812	0.4404

Notas: Las pruebas de la traza y del valor propio máximo indicaron la existencia de cointegración entre todas las variables.

Es necesario recordar que en la obtención de las estimaciones de los parámetros β_i en las ecuaciones de cointegración, se invierten sus signos al final.

⁴⁹ Particularmente **Guerrero** (2004).

⁵⁰ Cuadros del B.23 al B.25.

Igualmente como en el caso anterior se observa que las variables incluidas presentan cointegración.

De la aproximación observada de la participación de las exportaciones totales en los ingresos totales de la balanza de pagos al supuesto de una cuenta corriente equilibrada ($\theta \cong 1$) en la última parte del periodo en estudio, es posible observar que si bien los valores obtenidos para α y β_1 referidos a la tasa de crecimiento de la economía mexicana explicada por la tasa de crecimiento de la economía del resto del mundo, sus valores cercanos sirven para confirmar la validez de α (pues $\beta_1 = \theta\left(\frac{\pi}{\xi}\right) = \theta\alpha$).

La dificultad encontrada es en lo tocante a β_3 . Se espera de este coeficiente que sea negativo y la suma de η y ϕ (las elasticidades precio de las exportaciones y las importaciones) no mayor a uno en valor absoluto, pues es mostrado en otros casos, como en Guerrero (2004), para el caso de la economía mexicana que se cumple la condición de Marshall–Lerner. Por lo que aquí se tiene una ambigüedad en el resultado de la estimación de β_3 . Pues de un signo negativo asociado a los términos de intercambio resulta, en contexto de macroeconomía abierta moderna, la conveniencia de promover un tipo de cambio real competitivo, que implica el impulso de políticas que promuevan la productividad del trabajo y la competitividad de las empresas nacionales, no así el sostener la sobrevaluación del peso, sino más bien al revés para incrementar las exportaciones frente a las importaciones como ocurre en últimos tiempos. Lo cual tiene sentido, en contra del signo positivo del valor obtenido.

Aún así, el modelo ampliado que no incluye las corrientes de capital da una mejor explicación del movimiento conjunto de las variables involucradas que el modelo ampliado que sí las incluye, contando con una mayor congruencia

estadística. La ambigüedad presentada en el coeficiente de los términos intercambio afecta modelo en su conjunto, y no permite hacer más inferencias respecto a las relaciones entre las variables, sobre todo para explicar el crecimiento de la economía mexicana.

CONCLUSIONES

Es posible identificar los determinantes que permiten comprender y medir el crecimiento económico de un país, bajo una perspectiva post Keynesiana. Éstos son: la dinámica productiva del resto del mundo, el comportamiento de los flujos netos de capital y la evolución de los términos de intercambio. De un modo importante en la ecuación descrita para la observación del crecimiento económico restringido por la balanza de pagos destacan las elasticidades ingreso de las exportaciones y de las importaciones.

Para el estudio de las variables como determinantes del crecimiento económico en México en el periodo de 1980 a 2007, se plantearon dos modelos, el primero reducido y el segundo ampliado.

El modelo reducido plantea directamente, que el crecimiento económico del país depende de la dinámica productiva del resto del mundo (que en el caso en estudio es como variable proxy la dinámica productiva de la economía de los Estados Unidos), a partir de la relación existente entre las elasticidades ingreso de las exportaciones e importaciones $-\alpha-$. En materia de política económica, el objetivo de cualquier país en una situación como ésta, será incrementar lo más posible el valor de α , es decir, el implementar las políticas que lleven a un incremento de la elasticidad ingreso de las exportaciones, así como una disminución de la elasticidad ingreso de las importaciones.

En el modelo ampliado desde el punto de vista econométrico los efectos por la inclusión de las corrientes netas de capital y de los términos de intercambio son teóricamente ambiguos. Los flujos netos de capital relajan la restricción externa al

crecimiento, pero la significancia estadística de su parámetro β_2 depende de su magnitud respecto a los ingresos totales de la balanza de pagos $(1 - \theta)$. En la macroeconomía abierta contemporánea el recíproco de los términos de intercambio se conoce como tipo de cambio real. De aquí que es conveniente la implementación de políticas que impulsen a mejorar las exportaciones netas, y por consiguiente, promover el crecimiento económico. Mas el signo del coeficiente que relaciona el crecimiento económico con los términos de intercambio depende de la suma de las elasticidades precio de las exportaciones e importaciones. En el cumplimiento de la condición de Marshall–Lerner se espera que el signo de este coeficiente sea negativo.

Para el estudio de los modelos VAR propuestos en (3.14) y (3.18) fueron utilizadas las siguientes técnicas econométricas (presentadas a lo largo de la primera parte y del apéndice A de este trabajo):

- Pruebas de raíz unitaria Dickey–Fuller Aumentada a cada variable para conocer su orden de integración.
- Pruebas de razón de verosimilitud LR con el fin de obtener el número de rezagos óptimo a incluir en cada VAR.
- Contrastes para la congruencia estadística de los residuos, como las pruebas de correlación serial LM ,
- así como pruebas de normalidad de Jarque–Bera,
- y también pruebas de heteroscedasticidad de White.
- Pruebas para contrastar causalidad en el sentido de Granger uno a uno entre cada una de las variables.
- Y la metodología de Johansen para probar cointegración en cada modelo VAR.

Las variables en estudio, a partir de las pruebas de Johansen resultaron ser cointegradas en todos los casos, dando las estimaciones de los coeficientes de cointegración normalizados, y por tanto, de los parámetros de interés α y β_i de los modelos reducido y ampliado, con base en las estadísticas disponibles y construidas para este trabajo para el periodo 1980–2007 de forma trimestral.

Con los resultados obtenidos en este trabajo se observa en el intervalo en cuestión la confirmación de que el crecimiento económico de México está ligado básicamente a la dinámica productiva de los Estados Unidos⁵¹. De los valores específicos de los parámetros estimados en el modelo reducido ($\alpha = 1.1111$) y ampliado que excluye las corrientes de capital ($\beta_1 = 1.2812$), se puede observar que se mantienen las relaciones de movimiento conjunto entre las variables incluidas (PIB de México y PIB de EEUU).

En contexto histórico hay que apuntar que la relación entre las elasticidades ingreso de las exportaciones y las importaciones (α) ha ido disminuyendo, de modo que la tasa de crecimiento económico de nuestro país experimenta una significativa reducción, por lo que se ha tenido cada vez un más evidente alineamiento respecto a la dinámica productiva de los Estados Unidos.

En lo referente a las corrientes netas de capital, a pesar de su positivo efecto macroeconómico, en este estudio no parecen constituirse como solución a los problemas de la balanza de pagos (esto a partir de la falta de significación estadística del modelo que estima β_2). Y para los términos de intercambio, la evidencia econométrica obtenida aquí ($\beta_3 = 0.4404$) no da elementos para apoyar lo que se ha observado históricamente de las elasticidades precio de las exportaciones y las importaciones (y por tanto de la condición de Marshall–Lerner), que impone la revisión de las actuales políticas económicas para el incremento de la productividad y la competitividad en el sector externo.

⁵¹ En Guerrero (2004), es incluida la evolución de los términos de intercambio.

APÉNDICES

APÉNDICE A

PRUEBAS DE HIPÓTESIS

Habiendo mostrado en el cuerpo del marco teórico VAR los elementos más relevantes para el estudio tanto de modelos sin restricciones como de modelos con cointegración, en este apéndice se muestran primeramente las pruebas de estacionariedad conocidas como prueba de raíz unitaria elaboradas por Dickey y Fuller (1979 y 1981), que resulta necesario aplicar desde un principio en un análisis VAR para reconocer la condición de estacionariedad de cada serie temporal involucrada en el problema a estudiar. Y a continuación, una breve presentación de los contrastes realizados sobre cada modelo VAR propuesto (respecto a los modelos reducido y ampliado para los determinantes del crecimiento económico de la segunda parte) para probar su congruencia estadística a partir de sus residuos muestrales.

A.1 PRUEBAS DE RAÍZ UNITARIA

A.1.1 Prueba Dickey–Fuller

Para las pruebas de estacionariedad incluidas en el problema aplicado, se hizo uso de la prueba de raíz unitaria Dickey–Fuller Aumentada. Primeramente es

presentado su antecedente inmediato, la prueba de Dickey–Fuller publicada en 1979.

Considerando el proceso estocástico autorregresivo de orden uno, $AR(1)$,

$$y_t = \rho y_{t-1} + \varepsilon_t \quad t = 1, 2, \dots,$$

donde $y_0 = 0$, $\rho \in \mathbb{R}$, $\varepsilon_t \sim i. i. d. N(0, \sigma^2)$, se tiene una raíz unitaria cuando $\rho = 1$, y y_t se convierte en una caminata aleatoria sin intercepto, del cual se sabe que es un proceso estocástico no estacionario. En vez de hacer la regresión de y_t sobre su valor rezagado y_{t-1} para averiguar si la ρ estimada es estadísticamente igual a 1, se modifica la ecuación tomando la primera diferencia respecto a y_t :

$$y_t - y_{t-1} = \rho y_{t-1} - y_{t-1} + \varepsilon_t = (\rho - 1)y_{t-1} + \varepsilon_t$$

que puede ser expresada como

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

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

Se tienen entonces las siguientes hipótesis de prueba

$$H_0: \delta = 0 \quad \text{vs.}$$

$$H_a: \delta < 0$$

Si $\delta = 0$, entonces $\rho = 1$, es decir, se tiene una raíz unitaria, lo cual significa que la serie de tiempo es no estacionaria. También, si $\delta = 0$, se tiene que $\Delta y_t = \varepsilon_t$, y puesto que u_t es un término de error con ruido blanco gaussiano, entonces es estacionario, lo cual significa que las primeras diferencias de una serie de tiempo de caminata aleatoria son estacionarias.

En el intento de probar la hipótesis nula $\delta = 0$ (contra la hipótesis alternativa de que δ es negativa, y por lo tanto la serie estacionaria), se observa que el valor t del coeficiente estimado de y_{t-1} no sigue la distribución t incluso en muestras grandes; es decir, no tiene una distribución normal asintótica. Y por lo cual Dickey y Fuller (1979) probaron para este coeficiente estimado de y_{t-1} que su valor estimado sigue el del estadístico τ , al que se le conoce como la prueba Dickey–Fuller (DF).

Considerando que se puede probar asimismo la hipótesis nula

$$H_0: \rho = 1 \quad \text{vs.}$$

$$H_a: \rho < 1$$

Ésta es de interés en distintas aplicaciones dado que corresponde a la hipótesis de que es apropiado transformar series de tiempo tomando diferencias.

En 1970, Box y Jenkins (1976) sugirieron su estadístico de prueba

$$Q_K = n \sum_{k=1}^K r_k^2$$

donde

$$r_k = \frac{\sum_{t=k+1}^n \hat{e}_t \hat{e}_{t-k}}{\sum_{t=1}^n \hat{e}_t^2}$$

y los \hat{e}_t son los residuos del modelo ajustado. Bajo la hipótesis nula, el estadístico Q_K se distribuye aproximadamente como una variable aleatoria χ^2 con $K - p$ grados de libertad, donde p es el número de parámetros estimados. Si $\{y_t\}$ satisface la condición inicial de ser $AR(1)$ entonces bajo la hipótesis nula $p = 0$, y $\hat{e}_t = y_t - y_{t-1}$.

La prueba de razón de verosimilitud de la hipótesis nula $\rho = 1$ propuesta por Dickey y Fuller es una función de

$$\hat{\tau} = \frac{\hat{\rho} - 1}{S_e} \sqrt{\sum_{t=2}^n Y_{t-1}^2}$$

donde

$$S_e^2 = \frac{1}{n-2} \sum_{t=2}^n (Y_t - \hat{\rho}Y_{t-1})^2$$

Dickey y Fuller (1979) obtuvieron representaciones para las distribuciones límite de $\hat{\rho}$ y $\hat{\tau}$, dado que $|\rho| = 1$. Dichas representaciones les permitieron la construcción de tablas de puntos porcentuales para los estadísticos. Tanto $\hat{\tau}$ como $\hat{\rho}$ fueron generalizados a modelos que contienen tanto intercepto como tendencia temporal. Y fue con el método de Monte Carlo que compararon la potencia de los estadísticos $\hat{\rho}$ y $\hat{\tau}$ contra Q_K .

Los modelos de estudio bajo $H_0: \rho = 1$ vs. $H_a: \rho < 1$ son:

- a) $y_t = \rho y_{t-1} + \varepsilon_t$,
- b) $y_t = \mu + \rho y_{t-1} + \varepsilon_t$,
- c) $y_t = \mu + \beta t + \rho y_{t-1} + \varepsilon_t$,

donde $t = 1, 2, \dots$, $y_0 = 0$, $\rho \in \mathbb{R}$, $\varepsilon_t \sim i.i.d. N(0, \sigma^2)$, en cada caso.

Supóngase que se dispone de n observaciones y_1, y_2, \dots, y_n para su análisis y se definen los siguientes vectores de dimensión $1 \times (n - 1)$:

$$\mathbf{1}' = (1, 1, \dots, 1)$$

$$\mathbf{t}' = \left(1 - \binom{n}{2}, 2 - \binom{n}{2}, \dots, m n - 1 - \binom{n}{2} \right),$$

$$\mathbf{y}'_t = (y_2, y_3, \dots, y_n)$$

$$\mathbf{y}'_{t-1} = (y_1, y_2, \dots, y_{n-1})$$

Sean $\mathbf{u}_1 = \mathbf{y}_{t-1}$, $\mathbf{u}_2 = (\mathbf{1}, \mathbf{y}_{t-1})$, y $\mathbf{u}_3 = (\mathbf{1}, \mathbf{t}, \mathbf{y}_{t-1})$. Se define $\hat{\rho}_\mu$ como la última entrada en el vector $(\mathbf{u}'_2 \mathbf{u}_2)^{-1} \mathbf{u}'_2 \mathbf{y}_t$, y así también se define $\hat{\rho}_\tau$ como la última entrada en el vector $(\mathbf{u}'_3 \mathbf{u}_3)^{-1} \mathbf{u}'_3 \mathbf{y}_t$.

Los estadísticos análogos al estadístico t de la regresión para probar la hipótesis nula son:

$$\hat{t} = \frac{(\hat{\rho}-1)}{\sqrt{S_{e1}^2 c_1}},$$

$$\hat{t}_\mu = \frac{(\hat{\rho}_\mu-1)}{\sqrt{S_{e2}^2 c_2}},$$

$$\hat{t}_\tau = \frac{(\hat{\rho}_\tau-1)}{\sqrt{S_{e3}^2 c_3}},$$

donde S_{ek}^2 es el correspondiente error cuadrático medio de la regresión

$$S_{ek}^2 = \frac{1}{n - k - 1} [\mathbf{y}'_t (\mathbf{I} - \mathbf{u}_k (\mathbf{u}'_k \mathbf{u}_k)^{-1} \mathbf{u}'_k) \mathbf{y}_t]$$

y c_k es el elemento inferior derecho de $(\mathbf{u}'_k \mathbf{u}_k)^{-1}$.

De ser rechazada la hipótesis nula, significa que y_t es una serie de tiempo estacionaria con media cero en el primer caso, que y_t es estacionaria con una media

distinta de cero en el segundo caso, y que y_t es estacionaria alrededor de una tendencia determinista en el tercer caso.

Cada una de las tres especificaciones anteriores de la prueba DF tiene diferentes valores críticos. Como no hay una forma de saber cual especificación es la correcta para la serie en cuestión, resulta inevitable hacer pruebas de ensayo y error, a pesar de que se haga una búsqueda exhaustiva de datos. La posible existencia de correlación serial en la perturbación estocástica llevó a una expansión de esta prueba, presentada a continuación.

A.1.2 Prueba Dickey–Fuller Aumentada

Al llevar a cabo la prueba DF, se supone que el término de error ε_t no estaba correlacionado. Pero Dickey y Fuller (1981) desarrollaron una prueba para cuando dicho término si está correlacionado, lo cual se conoce como la prueba Dickey–Fuller Aumentada (DFA).

Para llevar a cabo esta prueba se aumenta a cualquiera de las tres ecuaciones vistas en el apartado anterior (bajo la hipótesis nula $H_0: \delta = 0$ vs. $H_a: \delta < 0$) los valores rezagados de la variable dependiente Δy_t .

Considerando, por ejemplo, el primer caso de la prueba DF:

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

Resulta que para ésta, la prueba DFA consiste en estimar la siguiente regresión:

$$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^m \alpha_i \Delta y_{t-i} + \varepsilon_t$$

donde ε_t es un proceso de ruido blanco gaussiano.

El número de términos de diferencia rezagados que se debe incluir, con frecuencia se determina de manera empírica, siendo la idea incluir los términos suficientes para que el término estocástico en la última ecuación no esté serialmente correlacionado⁵².

Resulta que los estadísticos para la prueba DFA siguen las mismas distribuciones asintóticas que los estadísticos para la prueba DF de la sección anterior, por lo que se pueden utilizar los mismos valores críticos para rechazar o no rechazar la hipótesis nula, es decir, si y_t es estacionaria o no, respectivamente.

A.2 PRUEBAS DE CONGRUENCIA ESTADÍSTICA PARA EL VAR

A.2.1 Prueba de Autocorrelación LM (*multiplicador de Lagrange*) de Breusch–Godfrey

Esta prueba fue desarrollada por los estadísticos Breusch y Godfrey en 1978 para evitar algunos de los problemas de la prueba d de Durbin–Watson para detectar autocorrelación, y es utilizada en un sentido más general pues permite regresoras no estocásticas como los valores rezagados de la variable regresada, esquemas autorregresivos de orden mayor y promedios móviles de términos de error con ruido blanco gaussiano de orden superior.

El método de la prueba de Breusch–Godfrey para un VAR sin restricciones como (1.11) implica primeramente estimar la ecuación de interés mediante MCO y obtener $\hat{\epsilon}_t$. A continuación se hace una nueva regresión auxiliar de $\hat{\epsilon}_t$ sobre \mathbf{y}_t , que incluye del lado derecho de la ecuación tanto los valores rezagados de \mathbf{y}_t como de $\hat{\epsilon}_t$

⁵² Al utilizar Eviews 5.1, en la aplicación de la prueba DFA a las series macroeconómicas en estudio, el número de rezagos se obtiene según el criterio de información de Schwartz, incluyendo un intercepto.

hasta el rezago h para ambos, y se calcula el coeficiente de determinación R^2 de esta regresión auxiliar.

Considerando a ρ_i para $i = 1, 2, \dots, h$ como los coeficientes de los respectivos $\hat{\varepsilon}_i$ dentro de la regresión auxiliar, se tiene que si todos estos ρ_i son significativamente iguales a cero entonces no existe correlación serial hasta el orden h para el modelo. Así se tiene la hipótesis nula

$$H_0: \rho_1 = \rho_2 = \dots = \rho_h = \mathbf{0} \quad \text{vs.}$$

$$H_a: \rho_i \neq \mathbf{0} \quad \text{para alguna } i = 1, 2, \dots, h$$

Si el tamaño de la muestra es grande, Breusch y Godfrey demostraron que el estadístico LM (por basarse en el principio multiplicador de Lagrange)

$$LM = T \times R^2$$

donde T es el número de observaciones y R^2 es el coeficiente de determinación de la regresión auxiliar, se distribuye asintóticamente como una χ^2 con n^2 grados de libertad, siendo n el número de variables de la regresión, dada $\mathbf{y}'_t = (y_1, y_2, \dots, y_n)$.

El objetivo será contar en el modelo, con una cantidad de rezagos h que permitan no rechazar la hipótesis nula con un adecuado nivel de confianza, es decir, para la no existencia de correlación serial, deberá presentarse un valor de la probabilidad asociada a LM relativamente alto, respecto a un nivel de significancia deseado.

A.2.2 Prueba de normalidad de Jarque–Bera

La prueba de normalidad utilizada sobre las variables de interés es una extensión multivariada de la prueba de Jarque Bera sobre la normalidad de los

residuos obtenidos por MCO, la cual compara los momentos tercero y cuarto de los residuos muestrales con los que corresponden a la distribución normal.

Para la prueba multivariada, se debe escoger una factorización de los k residuos (o innovaciones) que son ortogonales entre sí⁵³.

Sea \mathbf{P} una factorización matricial $k \times k$ tal que

$$\mathbf{u}_t = \mathbf{P}\boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \mathbf{I}_k)$$

donde $\boldsymbol{\varepsilon}_t$ es el vector de residuos reducido. Definiendo los vectores de momentos tercero y cuarto

$$\mathbf{m}_3 = \sum_t \frac{u_t^3}{T} \quad \text{y} \quad \mathbf{m}_4 = \sum_t \frac{u_t^4}{T}$$

entonces

$$\sqrt{T} \begin{bmatrix} \mathbf{m}_3 \\ \mathbf{m}_4 - 3 \end{bmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} 6\mathbf{I}_k & 0 \\ 0 & 24\mathbf{I}_k \end{bmatrix}\right)$$

Para una variable normalmente distribuida, el coeficiente de asimetría $\mathbf{m}_3 = 0$ y el de kurtosis $\mathbf{m}_4 = 3$. Por lo tanto, la prueba de normalidad constituye una prueba con la hipótesis nula conjunta

$$H_0: \mathbf{m}_3 = 0 \quad \text{y} \quad \mathbf{m}_4 = 3 \quad \text{vs.}$$

$$H_a: \mathbf{m}_3 \neq 0 \quad \text{ó} \quad \mathbf{m}_4 \neq 3.$$

Como cada componente es independiente del otro, es posible construir un estadístico χ^2 por la adición de los cuadrados de los componentes de estos momentos tercero y cuarto.

⁵³ Este concepto puede revisarse en el apartado 1.3.1 de este trabajo correspondiente a las innovaciones ortogonales.

En la aplicación de esta prueba en el programa Eviews se eligió la ortogonalización a partir de la factorización de Choleski⁵⁴, donde \mathbf{P} es la inversa de la matriz triangular inferior en la factorización de la matriz de varianzas y covarianzas de los residuos. El estadístico de prueba resultante depende del orden de las variables en el VAR.

En la tabla de resultados de esta prueba se observan los estadísticos de prueba para cada componente ortogonal y para la prueba conjunta. Para los componentes individuales, el coeficiente de asimetría estimado \mathbf{m}_3 y el de kurtosis \mathbf{m}_4 son presentados juntos en las primeras dos columnas con los p –values a partir de la distribución χ^2 con un grado de libertad. La columna con el estadístico Jarque–Bera

$$JB = T \left[\frac{\mathbf{m}_3^2}{6} + \frac{(\mathbf{m}_4 - 3)^2}{24} \right]$$

donde T es el número de observaciones, presenta JB contrastada con p –values tomados de una distribución χ^2 con 2 grados de libertad, esperando de acuerdo a la hipótesis nula que $JB = 0$.

Para la prueba conjunta, se tiene que

$$\lambda_3 = \frac{T\mathbf{m}_3'\mathbf{m}_3}{6} \sim \chi_{(k)}^2, \quad \lambda_4 = \frac{T(\mathbf{m}_4 - 3)'(\mathbf{m}_4 - 3)}{24} \sim \chi_{(k)}^2 \quad \text{y} \quad \lambda = \lambda_3 + \lambda_4 \sim \chi_{(2k)}^2.$$

Si el p –value calculado del estadístico JB es suficientemente bajo en una aplicación, lo cual sucederá si el valor del estadístico difiere en gran medida de cero, se puede rechazar la hipótesis de que los residuos están normalmente distribuidos. Por el contrario, si el p –value es razonablemente alto, lo cual pasa cuando el valor del estadístico está cercano a cero, no se rechaza la hipótesis nula de normalidad.

⁵⁴ ídem.

A.2.3 Prueba de heteroscedasticidad de White

Los contrastes aplicados en este trabajo para probar heteroscedasticidad (es decir, la varianza de los residuos muestrales no es constante) son la extensión para sistemas de ecuaciones de las pruebas uni–ecuacionales propuestas inicialmente por White (1980).

Para el caso uni–ecuacional de n variables del lado derecho de la ecuación, la prueba consiste en que a partir del modelo original se realice una nueva regresión auxiliar de los residuos muestrales al cuadrado sobre los regresores originales sumando a éstos del lado derecho del modelo auxiliar los cuadrados de los mismos, y en determinado caso pueden sumarse también o no, los productos cruzados de los regresores, dando así dos pruebas de White, una prueba con términos cruzados y otra que no los lleve; tomando en cuenta en cada caso el que hay que sumar también un término de intercepto.

De esta manera la hipótesis nula de no existencia de heteroscedasticidad es planteada en términos de los coeficientes β_i de los g regresores (excluyendo el intercepto) de la regresión auxiliar

$$H_0: \beta_1 = \beta_2 = \dots = \beta_g = 0 \quad \text{vs.}$$

$$H_a: \beta_i \neq 0 \quad \text{para alguna } i = 1, 2, \dots, g$$

La anterior hipótesis nula implica que la varianza de los residuos muestrales originales es igual al término de intercepto, o sea una constante, y por lo tanto se tiene homoscedasticidad.

Al obtener R^2 de esta regresión auxiliar, multiplicándolo por el número de observaciones T , se tiene que

$$LM = T \times R^2$$

se distribuye asintóticamente como una distribución χ^2 con g grados de libertad.

El criterio de decisión para no rechazar la hipótesis nula consiste en que los valores obtenidos para el estadístico LM no excedan el p -value para el nivel de significancia elegido.

Para el caso multivariado la regresión auxiliar se obtiene haciendo la regresión de cada producto cruzado de los residuos muestrales sobre los productos cruzados de los regresores originales y probando la significancia conjunta de esta nueva regresión. Como se anotó en el caso univariado, para la prueba realizada sin términos cruzados se tiene que la regresión incluye solamente los regresores originales en niveles y cuadrados, mientras que para la prueba con términos cruzados son incluidos todos los productos cruzados que no sean redundantes. Igualmente siempre es incluido un intercepto en la regresión auxiliar.

Y asimismo se tienen las hipótesis nula de no heteroscedasticidad y la alternativa

$$H_0: \beta_1 = \beta_2 = \dots = \beta_g = \mathbf{0} \quad \text{vs.}$$

$$H_a: \beta_i \neq 0 \quad \text{para alguna } i = 1, 2, \dots, g$$

La primera parte de los resultados de la aplicación de esta prueba en Eviews, presenta la significancia conjunta de los regresores excluyendo el intercepto de cada regresión de prueba. Como en el caso univariado se puede pensar cada regresión auxiliar como una prueba de su igualdad con una constante (y por lo tanto la igualdad a cero de todos sus coeficientes) de cada elemento en la matriz de

varianzas y covarianzas de los residuos. Bajo la hipótesis nula la ausencia de intercepto no debería ser conjuntamente significativa.

La última parte de los resultados muestra el valor χ^2 del estadístico LM para la significancia conjunta de todos los regresores en el sistema de las ecuaciones de prueba⁵⁵. El estadístico $LM = T \times R^2$ del sistema se distribuye asintóticamente como una χ^2 con mn grados de libertad, donde $m = \frac{k(k+1)}{2}$ es el número de todos los productos cruzados de los residuos en el sistema y n es el número del conjunto común de variables que aparecen del lado derecho en la regresión auxiliar.

Igualmente para el caso multivariado, el criterio de decisión para no rechazar la hipótesis nula consiste en valores para el estadístico LM que no excedan el valor χ^2 crítico para el nivel de significancia elegido.

En el caso de un modelo que tiene muchos regresores y por lo tanto las regresiones auxiliares implican las sumas tanto de los niveles como de los cuadrados y todos los productos cruzados de los regresores, esto afecta al consumir grados de libertad rápidamente. En los casos en los cuales el estadístico de prueba LM para White es significativo estadísticamente, la heteroscedasticidad puede no ser necesariamente la causa, sino los errores de especificación. Así, la prueba de White puede ser una prueba de heteroscedasticidad (pura) o de error de especificación o de ambos.

Se ha argumentado que si no están presentes términos con productos cruzados en el procedimiento entonces se tiene una prueba de heteroscedasticidad pura, e incluyendo dicho términos de productos cruzados, entonces es una prueba de heteroscedasticidad y de sesgo de especificación.

⁵⁵ Consultar Doornik (1995), para más detalles.

APÉNDICE B

CUADROS DE RESULTADOS – MODELO APLICADO

Cuadro B.1	Prueba de raíz unitaria DFA para L(Yeeuu)
Cuadro B.2	Prueba de raíz unitaria DFA para D(L(Yeeuu))
Cuadro B.3	Prueba de raíz unitaria DFA para L(Ymex)
Cuadro B.4	Prueba de raíz unitaria DFA para D(L(Ymex))
Cuadro B.5	Prueba de correlación serial de los residuos LM para el VAR reducido (p = 5)
Cuadro B.6	Prueba de normalidad de los residuos Jarque–Bera para el VAR reducido (p = 5)
Cuadro B.7	Prueba de heteroscedasticidad de White (sin términos cruzados) para el VAR reducido (p = 5)
Cuadro B.8	Prueba de heteroscedasticidad de White (con términos cruzados) para el VAR reducido (p = 5)
Cuadro B.9	Prueba de selección del orden del rezago LR para el VAR reducido
Cuadro B.10	Prueba de Johansen del modelo VAR reducido (número de rezagos: 2)
Cuadro B.11	Prueba de Johansen del modelo VAR reducido (número de rezagos: 3)
Cuadro B.12	Prueba de Johansen del modelo VAR reducido (número de rezagos: 4)
Cuadro B.13	Prueba de Johansen del modelo VAR reducido (número de rezagos: 5)
Cuadro B.14	Prueba de causalidad de Granger para L(Ymex) y L(Yeeuu)
Cuadro B.15	Estimación de la ecuación $L(Y_t) = \beta_0 + \beta_1 L(XT_t) + u_t$
Cuadro B.16	Prueba de raíz unitaria DFA para K
Cuadro B.17	Prueba de raíz unitaria DFA para L(TI)

- Cuadro B.18 Prueba de correlación serial de los residuos LM para el modelo VAR ampliado (incluyendo K) (p=5)
- Cuadro B.19 Prueba de normalidad de los residuos Jarque–Bera para el modelo VAR ampliado (incluyendo K) (p=5)
- Cuadro B.20 Prueba de heteroscedasticidad de White (sin términos cruzados) para el modelo VAR ampliado (incluyendo K) (p=5)
- Cuadro B.21 Prueba de selección del orden del rezago LR para el VAR ampliado (incluyendo K)
- Cuadro B.22 Prueba de Johansen del modelo VAR ampliado (incluyendo K, número de rezagos: 5)
- Cuadro B.23 Prueba de correlación serial de los residuos LM para el modelo VAR ampliado (excluyendo K) (p=5)
- Cuadro B.24 Prueba de normalidad de los residuos Jarque–Bera para el modelo VAR ampliado (excluyendo K) (p=5)
- Cuadro B.25 Prueba de heteroscedasticidad de White (sin términos cruzados) para el modelo VAR ampliado (excluyendo K) (p=5)
- Cuadro B.26 Prueba de selección del orden del rezago LR para el VAR ampliado (excluyendo K)
- Cuadro B.27 Prueba de Johansen del modelo VAR ampliado (excluyendo K, número de rezagos: 5)
- Cuadro B.28 Prueba de causalidad de Granger uno a uno para L(Ymex), L(Yeeuu), K, L(TI)

Cuadro B.1
Prueba de raíz unitaria DFA para L(Yeeuu)

Null Hypothesis: LOGYEEUU has a unit root

Exogenous: Constant

Lag Length: 2 (Automatic based on SIC, MAXLAG=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.907910	0.7825
Test critical values: 1% level	-3.491345	
5% level	-2.888157	
10% level	-2.581041	

*MacKinnon (1996) one-sided p-values.

Cuadro B.2
Prueba de raíz unitaria DFA para D(L(Yeeuu))

Null Hypothesis: DIFLOGYEEUU has a unit root

Exogenous: Constant

Lag Length: 1 (Automatic based on SIC, MAXLAG=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.743323	0.0000
Test critical values: 1% level	-3.491345	
5% level	-2.888157	
10% level	-2.581041	

*MacKinnon (1996) one-sided p-values.

Cuadro B.3
Prueba de raíz unitaria DFA para L(Ymex)

Null Hypothesis: LOGYMEX has a unit root

Exogenous: Constant

Lag Length: 9 (Automatic based on SIC, MAXLAG=12)

	t-Statistic	Prob.*
--	-------------	--------

Augmented Dickey-Fuller test statistic		1.174323	0.9979
Test critical values:	1% level	-3.495677	
	5% level	-2.890037	
	10% level	-2.582041	

*MacKinnon (1996) one-sided p-values.

Cuadro B.4
Prueba de raíz unitaria DFA para D(L(Ymex))

Null Hypothesis: DIFLOGYMEX has a unit root
Exogenous: Constant
Lag Length: 7 (Automatic based on SIC, MAXLAG=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.316293	0.0166
Test critical values:		
	1% level	-3.495021
	5% level	-2.889753
	10% level	-2.581890

*MacKinnon (1996) one-sided p-values.

Cuadro B.5
Prueba de correlación serial de los residuos LM para el VAR reducido (p= 5)

VAR Residual Serial Correlation LM Tests
H0: no serial correlation at lag order h
Sample: 1980Q1 2007Q4
Included observations: 107

Lags	LM-Stat	Prob
1	14.20849	0.0067
2	9.805823	0.0438
3	10.37077	0.0346
4	13.18187	0.0104
5	2.869434	0.5799
6	7.323657	0.1197
7	3.887773	0.4214
8	4.075337	0.3959
9	2.843537	0.5843
10	1.632081	0.8030
11	5.428186	0.2461

12 10.99308 0.0266

Probs from chi-square with 4 df.

Cuadro B.6

Prueba de normalidad de los residuos Jarque-Bera para el VAR reducido ($p=5$)

VAR Residual Normality Tests

Orthogonalization: Cholesky (Lutkepohl)

H0: residuals are multivariate normal

Sample: 1980Q1 2007Q4

Included observations: 107

Component	Skewness	Chi-sq	df	Prob.
1	-1.340446	32.04286	1	0.0000
2	-0.326675	1.903110	1	0.1677
Joint		33.94597	2	0.0000

Component	Kurtosis	Chi-sq	df	Prob.
1	7.410576	86.72877	1	0.0000
2	3.013169	0.000773	1	0.9778
Joint		86.72954	2	0.0000

Component	Jarque-Bera	df	Prob.
1	118.7716	2	0.0000
2	1.903883	2	0.3860
Joint	120.6755	4	0.0000

Cuadro B.7

Prueba de heteroscedasticidad de White (sin términos cruzados) para el VAR reducido ($p=5$)

VAR Residual Heteroskedasticity Tests: No Cross Terms (only levels and squares)

Sample: 1980Q1 2007Q4

Included observations: 107

Joint test:

Chi-sq	df	Prob.
73.59031	60	0.1117

Individual components:

Dependent	R-squared	F(20,86)	Prob.	Chi-sq(20)	Prob.
res1*res1	0.149806	0.757671	0.7549	16.02928	0.7148
res2*res2	0.330441	2.122140	0.0091	35.35721	0.0183
res2*res1	0.205201	1.110176	0.3549	21.95655	0.3429

Cuadro B.8

Prueba de heteroscedasticidad de White (con términos cruzados) para el VAR reducido (p = 5)

VAR Residual Heteroskedasticity Tests: Includes Cross Terms

Sample: 1980Q1 2007Q4

Included observations: 107

Joint test:

Chi-sq	df	Prob.
193.1869	141	0.0023

Individual components:

Dependent	R-squared	F(47,59)	Prob.	Chi-sq(47)	Prob.
res1*res1	0.486540	1.189506	0.2621	52.05981	0.2835
res2*res2	0.751508	3.796435	0.0000	80.41139	0.0017
res2*res1	0.583898	1.761538	0.0199	62.47712	0.0648

Cuadro B.9

Prueba de selección del orden del rezago LR para el VAR reducido

VAR Lag Order Selection Criteria

Endogenous variables: LOGYMEX LOGYEEUU

Exogenous variables: C

Sample: 1980Q1 2007Q4

Included observations: 104

Lag	LogL	LR	FPE	AIC	SC	HQ
0	165.0151	NA	0.000149	-3.134907	-3.084053	-3.114304
1	577.9458	802.0383	5.73e-08	-10.99896	-10.84640	-10.93715
2	616.9028	74.16829	2.93e-08	-11.67121	-11.41694	-11.56820
3	621.1769	7.972733	2.91e-08	-11.67648	-11.32050	-11.53226
4	645.4794	44.39887	1.97e-08	-12.06691	-11.60923	-11.88149
5	674.8923	52.60387*	1.21e-08*	-12.55562*	-11.99623*	-12.32900*
6	677.0577	3.789318	1.26e-08	-12.52034	-11.85924	-12.25251
7	681.0581	6.846925	1.26e-08	-12.52035	-11.75754	-12.21131
8	682.2980	2.074461	1.33e-08	-12.46727	-11.60276	-12.11703

* 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

Cuadro B.10
Prueba de Johansen del modelo VAR reducido (número de rezagos: 2)

Sample (adjusted): 1980Q4 2007Q4

Included observations: 109 after adjustments

Trend assumption: No deterministic trend

Series: LOGYMEX LOGYEEUU

Lags interval (in first differences): 1 to 2

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.181575	23.71300	12.32090	0.0004
At most 1	0.017031	1.872328	4.129906	0.2015

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 *	0.181575	21.84067	11.22480	0.0005
At most 1	0.017031	1.872328	4.129906	0.2015

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=l):

LOGYMEX	LOGYEEUU
-0.771525	1.026493
5.264883	-8.295750

Unrestricted Adjustment Coefficients (alpha):

D(LOGYMEX)	-0.002546	-0.004029
D(LOGYEEUU)	-0.002881	8.40E-05

1 Cointegrating Equation(s): Log likelihood 623.3094

Normalized cointegrating coefficients (standard error in parentheses)

LOGYMEX	LOGYEEUU
1.000000	1.330473
	(0.05051)

Adjustment coefficients (standard error in parentheses)

D(LOGYMEX)	0.001964
	(0.00237)
D(LOGYEEUU)	0.002223
	(0.00047)

Cuadro B.II

Prueba de Johansen del modelo VAR reducido (número de rezagos: 3)

Sample (adjusted): 1981Q1 2007Q4

Included observations: 108 after adjustments

Trend assumption: No deterministic trend

Series: LOGYMEX LOGYEEUU

Lags interval (in first differences): 1 to 3

Unrestricted Cointegration Rank Test (Trace)

Hypothesized		Trace	0.05	
No. of CE(s)	Eigenvalue	Statistic	Critical Value	Prob.**

None *	0.147000	21.20615	12.32090	0.0013
At most 1	0.036669	4.034632	4.129906	0.0529

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 *	0.147000	17.17152	11.22480	0.0041
At most 1	0.036669	4.034632	4.129906	0.0529

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=l):

LOGYMEX	LOGYEEUU
0.832876	-1.510854
5.318848	-8.317471

Unrestricted Adjustment Coefficients (alpha):

D(LOGYMEX)	-0.004260	-0.004368
D(LOGYEEUU)	-0.002147	0.000556

1 Cointegrating Equation(s): Log likelihood 646.0442

Normalized cointegrating coefficients (standard error in parentheses)

LOGYMEX	LOGYEEUU
1.000000	-1.814020
	(0.05926)

Adjustment coefficients (standard error in parentheses)

D(LOGYMEX)	-0.003548
	(0.00207)
D(LOGYEEUU)	-0.001788
	(0.00049)

Cuadro B.12
Prueba de Johansen del modelo VAR reducido (número de rezagos: 4)

Sample (adjusted): 1981Q2 2007Q4
 Included observations: 107 after adjustments
 Trend assumption: No deterministic trend
 Series: LOGYMEX LOGYEEUU
 Lags interval (in first differences): 1 to 4

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.164705	24.54528	12.32090	0.0003
At most 1 *	0.048223	5.288456	4.129906	0.0255

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 *	0.164705	19.25683	11.22480	0.0016
At most 1 *	0.048223	5.288456	4.129906	0.0255

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):

LOGYMEX	LOGYEEUU
0.376204	-0.812911
5.497553	-8.621228

Unrestricted Adjustment Coefficients (alpha):

D(LOGYMEX)	0.000146	-0.004275
D(LOGYEEUU)	-0.002399	-0.000130

1 Cointegrating Equation(s): Log likelihood 676.2556

Normalized cointegrating coefficients (standard error in parentheses)

LOGYMEX	LOGYEEUU
1.000000	-2.160825
	(0.13450)

Adjustment coefficients (standard error in parentheses)

D(LOGYMEX)	5.48E-05
	(0.00074)
D(LOGYEEUU)	-0.000902
	(0.00021)

Cuadro B.13

Prueba de Johansen del modelo VAR reducido (número de rezagos: 5)

Sample (adjusted): 1981Q3 2007Q4
 Included observations: 106 after adjustments
 Trend assumption: No deterministic trend
 Series: LOGYMEX LOGYEEUU
 Lags interval (in first differences): 1 to 5

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.220549	31.98936	12.32090	0.0000
At most 1 *	0.051260	5.577804	4.129906	0.0216

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 *	0.220549	26.41156	11.22480	0.0001
At most 1 *	0.051260	5.577804	4.129906	0.0216

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):

LOGYMEX	LOGYEEUU
-0.525055	0.583365
5.641267	-8.889241

Unrestricted Adjustment Coefficients (alpha):

D(LOGYMEX)	-4.28E-05	-0.004398
D(LOGYEEUU)	-0.002626	-9.84E-05

1 Cointegrating Equation(s): Log likelihood 679.9498

Normalized cointegrating coefficients (standard error in parentheses)

LOGYMEX	LOGYEEUU
1.000000	-1.111053
	(0.08924)

Adjustment coefficients (standard error in parentheses)

D(LOGYMEX)	2.25E-05
	(0.00105)
D(LOGYEEUU)	0.001379
	(0.00027)

Cuadro B.14
Prueba de causalidad de Granger para L(Ymex) y L(Yeeuu)

Pairwise Granger Causality Tests
Sample: 1980Q1 2007Q4
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Probability
LOGYEEUU does not Granger Cause LOGYMEX	110	3.53214	0.03276
LOGYMEX does not Granger Cause LOGYEEUU		0.89243	0.41274

Cuadro B.15
Estimación de la ecuación $L(Y_t) = \beta_0 + \beta_1 L(XT_t) + u_t$

Dependent Variable: LOGYMEX
Method: Least Squares

Date: 09/10/08 Time: 18:35
 Sample: 1980Q1 2007Q4
 Included observations: 112
 LOGYMEX=C(1)+C(2)*LOGXT

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	10.53267	0.084734	124.3029	0.0000
C(2)	0.284223	0.006801	41.78949	0.0000
R-squared	0.940744	Mean dependent var		14.06752
Adjusted R-squared	0.940206	S.D. dependent var		0.215675
S.E. of regression	0.052739	Akaike info criterion		-3.029234
Sum squared resid	0.305952	Schwarz criterion		-2.980689
Log likelihood	171.6371	Durbin-Watson stat		0.624406

Cuadro B.16
 Prueba de raíz unitaria DFA para K

Null Hypothesis: K has a unit root
 Exogenous: Constant
 Lag Length: 4 (Automatic based on SIC, MAXLAG=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.727501	0.0727
Test critical values:		
1% level	-3.492523	
5% level	-2.888669	
10% level	-2.581313	

*MacKinnon (1996) one-sided p-values.

Cuadro B.17
 Prueba de raíz unitaria DFA para L(TI)

Null Hypothesis: LOGTI has a unit root
 Exogenous: Constant
 Lag Length: 0 (Automatic based on SIC, MAXLAG=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.758707	0.0676
Test critical values:		
1% level	-3.490210	
5% level	-2.887665	

10% level

-2.580778

*MacKinnon (1996) one-sided p-values.

Cuadro B.18
Prueba de correlación serial de los residuos LM para el modelo VAR ampliado
(incluyendo K) (p=5)

VAR Residual Serial Correlation LM Tests

H0: no serial correlation at lag order h

Sample: 1980Q1 2007Q4

Included observations: 107

Lags	LM-Stat	Prob
1	22.75837	0.1203
2	22.15251	0.1383
3	11.75584	0.7606
4	21.79713	0.1499
5	18.40348	0.3008
6	13.14508	0.6621
7	17.40161	0.3601
8	17.87822	0.3311
9	9.582160	0.8875
10	10.64880	0.8306
11	12.25770	0.7261
12	12.36715	0.7183

Probs from chi-square with 16 df.

Cuadro B.19
Prueba de normalidad de los residuos Jarque-Bera para el modelo VAR
ampliado (incluyendo K) (p=5)

VAR Residual Normality Tests

Orthogonalization: Cholesky (Lutkepohl)

H0: residuals are multivariate normal

Sample: 1980Q1 2007Q4

Included observations: 107

Component	Skewness	Chi-sq	df	Prob.
1	-0.560788	5.608276	1	0.0179

2	-0.045088	0.036254	1	0.8490
3	-0.560271	5.597946	1	0.0180
4	-0.064163	0.073417	1	0.7864
Joint		11.31589	4	0.0232
Component	Kurtosis	Chi-sq	df	Prob.
1	2.564511	0.845526	1	0.3578
2	1.859109	5.803115	1	0.0160
3	3.337005	0.506342	1	0.4767
4	9.945513	215.0707	1	0.0000
Joint		222.2256	4	0.0000
Component	Jarque-Bera	df	Prob.	
1	6.453802	2	0.0397	
2	5.839370	2	0.0540	
3	6.104288	2	0.0473	
4	215.1441	2	0.0000	
Joint	233.5415	8	0.0000	

Cuadro B.20

Prueba de heteroscedasticidad de White (sin términos cruzados) para el modelo ampliado (incluyendo K) (p=5)

VAR Residual Heteroskedasticity Tests: No Cross Terms (only levels and squares)

Sample: 1980Q1 2007Q4

Included observations: 107

Joint test:

Chi-sq	df	Prob.
470.6083	400	0.0085

Individual components:

Dependent	R-squared	F(40,66)	Prob.	Chi-sq(40)	Prob.
res1*res1	0.327918	0.805059	0.7676	35.08726	0.6908

res2*res2	0.537855	1.920304	0.0092	57.55044	0.0356
res3*res3	0.338474	0.844234	0.7146	36.21676	0.6413
res4*res4	0.388614	1.048787	0.4243	41.58173	0.4017
res2*res1	0.468212	1.452740	0.0883	50.09867	0.1315
res3*res1	0.312979	0.751673	0.8330	33.48874	0.7568
res3*res2	0.572768	2.212071	0.0021	61.28618	0.0168
res4*res1	0.416868	1.179547	0.2720	44.60485	0.2843
res4*res2	0.524394	1.819256	0.0153	56.11013	0.0468
res4*res3	0.452690	1.364747	0.1298	48.43787	0.1691

Cuadro B.2I
Prueba de selección del orden del rezago LR para el VAR ampliado
(incluyendo K)

VAR Lag Order Selection Criteria

Endogenous variables: LOGYMEX LOGYEEUU K LOGTI

Exogenous variables: C

Sample: 1980Q1 2007Q4

Included observations: 104

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-971.9418	NA	1663.378	18.76811	18.86982	18.80932
1	-420.2637	1050.310	0.055877	8.466609	8.975145	8.672632
2	-376.9997	79.03991	0.033125	7.942302	8.857668*	8.313144
3	-348.0508	50.66058	0.025914	7.693285	9.015480	8.228945
4	-304.4074	73.01871	0.015329	7.161682	8.890706	7.862160
5	-272.5787	50.80354*	0.011428*	6.857283*	8.993137	7.722580*
6	-258.1787	21.87700	0.011973	6.888051	9.430735	7.918167
7	-249.4902	12.53142	0.014092	7.028658	9.978171	8.223592
8	-240.9551	11.65375	0.016763	7.172213	10.52856	8.531966

* 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

Cuadro B.22
Prueba de Johansen del modelo VAR ampliado
(incluyendo K, número de rezagos: 5)

Sample (adjusted): 1981Q3 2007Q4
 Included observations: 106 after adjustments
 Trend assumption: No deterministic trend
 Series: LOGYMEX LOGYEEUU K LOGTI
 Lags interval (in first differences): 1 to 5

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.212658	57.57014	40.17493	0.0004
At most 1 *	0.168628	32.22637	24.27596	0.0041
At most 2 *	0.107779	12.65052	12.32090	0.0440
At most 3	0.005290	0.562176	4.129906	0.5154

Trace test indicates 3 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 *	0.212658	25.34378	24.15921	0.0344
At most 1 *	0.168628	19.57585	17.79730	0.0267
At most 2 *	0.107779	12.08834	11.22480	0.0352
At most 3	0.005290	0.562176	4.129906	0.5154

Max-eigenvalue test indicates 3 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=l):

LOGYMEX	LOGYEEUU	K	LOGTI
0.230256	-0.385359	-1.58E-05	0.872750
4.480938	-7.825813	5.92E-06	1.710248
9.210016	-11.98180	2.71E-05	-5.549083

11.20799	-14.07969	-1.10E-05	-7.842295	
Unrestricted Adjustment Coefficients (alpha):				
D(LOGYMEX)	0.002747	-0.002006	-0.001110	-0.000882
D(LOGYEEUU)	0.001922	-0.000821	0.000291	0.000118
D(K)	-338.4441	-4039.566	-4549.243	117.3070
D(LOGTI)	-0.010745	-0.015600	0.008808	-0.000522
1 Cointegrating Equation(s):	Log likelihood	-294.5357		
Normalized cointegrating coefficients (standard error in parentheses)				
LOGYMEX	LOGYEEUU	K	LOGTI	
1.000000	-1.673613	-6.87E-05	3.790348	
	(1.85658)	(2.9E-05)	(4.19259)	
Adjustment coefficients (standard error in parentheses)				
D(LOGYMEX)	0.000632			
	(0.00036)			
D(LOGYEEUU)	0.000443			
	(0.00011)			
D(K)	-77.92879			
	(426.634)			
D(LOGTI)	-0.002474			
	(0.00128)			
2 Cointegrating Equation(s):	Log likelihood	-284.7478		
Normalized cointegrating coefficients (standard error in parentheses)				
LOGYMEX	LOGYEEUU	K	LOGTI	
1.000000	0.000000	-0.001678	82.09388	
		(0.00069)	(17.8734)	
0.000000	1.000000	-0.000962	46.78713	
		(0.00039)	(10.2355)	
Adjustment coefficients (standard error in parentheses)				
D(LOGYMEX)	-0.008357	0.014641		
	(0.00699)	(0.01220)		
D(LOGYEEUU)	-0.003235	0.005682		
	(0.00220)	(0.00384)		
D(K)	-18178.97	31743.31		
	(8077.75)	(14106.0)		
D(LOGTI)	-0.072378	0.126226		
	(0.02383)	(0.04162)		
3 Cointegrating Equation(s):	Log likelihood	-278.7036		

Normalized cointegrating coefficients (standard error in parentheses)

LOGYMEX	LOGYEEUU	K	LOGTI
1.000000	0.000000	0.000000	-3.085713 (0.16171)
0.000000	1.000000	0.000000	-2.023736 (0.10618)
0.000000	0.000000	1.000000	-50759.09 (9019.85)

Adjustment coefficients (standard error in parentheses)

D(LOGYMEX)	-0.018584 (0.01591)	0.027946 (0.02223)	-8.55E-08 (5.0E-08)
D(LOGYEEUU)	-0.000558 (0.00501)	0.002199 (0.00700)	-2.74E-08 (1.6E-08)
D(K)	-60077.57 (17737.6)	86251.45 (24786.8)	-0.142007 (0.05535)
D(LOGTI)	0.008745 (0.05353)	0.020689 (0.07481)	3.17E-07 (1.7E-07)

Cuadro B.23
Prueba de correlación serial de los residuos LM para el modelo VAR ampliado
(excluyendo K) (p=5)

VAR Residual Serial Correlation LM Tests

H0: no serial correlation at lag order h

Sample: 1980Q1 2007Q4

Included observations: 107

Lags	LM-Stat	Prob
1	15.39135	0.0807
2	9.517949	0.3909
3	9.917044	0.3572
4	17.85105	0.0369
5	8.951011	0.4418
6	5.897445	0.7501
7	7.994518	0.5347
8	11.31694	0.2546
9	6.402150	0.6991
10	5.572192	0.7819
11	9.215727	0.4176
12	12.61337	0.1809

Probs from chi-square with 9 df.

Cuadro B.24
 Prueba de normalidad de los residuos Jarque-Bera para el modelo VAR
 ampliado (excluyendo K) (p=5)

VAR Residual Normality Tests
 Orthogonalization: Cholesky (Lutkepohl)
 H0: residuals are multivariate normal
 Sample: 1980Q1 2007Q4
 Included observations: 107

Component	Skewness	Chi-sq	df	Prob.
1	-0.989793	17.47112	1	0.0000
2	-0.198138	0.700114	1	0.4027
3	0.229938	0.942875	1	0.3315
Joint		19.11411	3	0.0003
Component	Kurtosis	Chi-sq	df	Prob.
1	5.489007	27.62008	1	0.0000
2	2.436261	1.416867	1	0.2339
3	13.67118	507.6882	1	0.0000
Joint		536.7251	3	0.0000
Component	Jarque-Bera	df	Prob.	
1	45.09120	2	0.0000	
2	2.116981	2	0.3470	
3	508.6311	2	0.0000	
Joint	555.8392	6	0.0000	

Cuadro B.25
 Prueba de heteroscedasticidad de White (sin términos cruzados) para el
 modelo VAR ampliado (excluyendo K) (p=5)

VAR Residual Heteroskedasticity Tests: No Cross Terms (only levels and squares)

Sample: 1980Q1 2007Q4
Included observations: 107

Joint test:		
Chi-sq	df	Prob.
214.7671	180	0.0391

Individual components:					
Dependent	R-squared	F(30,76)	Prob.	Chi-sq(30)	Prob.
res1*res1	0.332986	1.264686	0.2053	35.62947	0.2205
res2*res2	0.418389	1.822387	0.0189	44.76766	0.0406
res3*res3	0.151782	0.453319	0.9913	16.24065	0.9806
res2*res1	0.293205	1.050921	0.4181	31.37293	0.3972
res3*res1	0.244519	0.819939	0.7237	26.16355	0.6668
res3*res2	0.237486	0.789009	0.7628	25.41097	0.7048

Cuadro B.26
Prueba de selección del orden del rezago LR para el VAR ampliado
(excluyendo K)

VAR Lag Order Selection Criteria
Endogenous variables: LOGYMEX LOGYEEUU LOGTI
Exogenous variables: C
Sample: 1980Q1 2007Q4
Included observations: 104

Lag	LogL	LR	FPE	AIC	SC	HQ
0	243.3339	NA	1.97e-06	-4.621806	-4.545526	-4.590903
1	737.2407	949.8208	1.76e-10	-13.94694	-13.64181	-13.82332
2	772.5192	65.80790	1.06e-10	-14.45229	-13.91833	-14.23597
3	781.5245	16.27893	1.06e-10	-14.45240	-13.68959	-14.14336
4	809.4152	48.80871	7.41e-11	-14.81568	-13.82403	-14.41393
5	849.5736	67.96039*	4.08e-11*	-15.41488*	-14.19439*	-14.92042*
6	855.8246	10.21793	4.33e-11	-15.36201	-13.91268	-14.77485
7	860.9819	8.132582	4.69e-11	-15.28811	-13.60994	-14.60824
8	865.2193	6.437587	5.19e-11	-15.19652	-13.28951	-14.42394

* 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

Cuadro B.27
Prueba de Johansen del modelo VAR ampliado
(excluyendo K, número de rezagos: 5)

Sample (adjusted): 1981Q3 2007Q4
 Included observations: 106 after adjustments
 Trend assumption: No deterministic trend
 Series: LOGYMEX LOGYEEUU LOGTI
 Lags interval (in first differences): 1 to 5

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.212878	42.10065	24.27596	0.0001
At most 1 *	0.131999	16.72723	12.32090	0.0086
At most 2	0.016111	1.721631	4.129906	0.2227

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 *	0.212878	25.37342	17.79730	0.0030
At most 1 *	0.131999	15.00560	11.22480	0.0104
At most 2	0.016111	1.721631	4.129906	0.2227

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=l):

LOGYMEX	LOGYEEUU	LOGTI
2.779578	-3.561123	-1.224183
2.815795	-5.511453	2.380850
-13.82964	17.65858	9.069237

Unrestricted Adjustment Coefficients (alpha):

D(LOGYMEX)	-9.95E-05	-0.002372	0.002100
D(LOGYEEUU)	0.002421	-0.000419	1.90E-05
D(LOGTI)	-0.006668	-0.018893	-0.001297

1 Cointegrating Equation(s): Log likelihood 852.5912

Normalized cointegrating coefficients (standard error in parentheses)

LOGYMEX	LOGYEEUU	LOGTI
1.000000	-1.281174	-0.440421
	(0.13699)	(0.30365)

Adjustment coefficients (standard error in parentheses)

D(LOGYMEX)	-0.000277
	(0.00521)
D(LOGYEEUU)	0.006729
	(0.00141)
D(LOGTI)	-0.018533
	(0.01598)

2 Cointegrating Equation(s): Log likelihood 860.0940

Normalized cointegrating coefficients (standard error in parentheses)

LOGYMEX	LOGYEEUU	LOGTI
1.000000	0.000000	-2.877018
		(0.15635)
0.000000	1.000000	-1.901847
		(0.10004)

Adjustment coefficients (standard error in parentheses)

D(LOGYMEX)	-0.006955	0.013426
	(0.00735)	(0.01219)
D(LOGYEEUU)	0.005548	-0.006310
	(0.00199)	(0.00331)
D(LOGTI)	-0.071731	0.127870
	(0.02133)	(0.03538)

Cuadro B.28

Prueba de causalidad de Granger uno a uno para L(Ymex), L(Yeeuu), K, L(TI)

Pairwise Granger Causality Tests
 Sample: 1980Q1 2007Q4
 Lags: 2

Null Hypothesis:	Obs	F-Statistic	Probability
LOGYEEUU does not Granger Cause LOGYMEX	110	3.53214	0.03276
LOGYMEX does not Granger Cause LOGYEEUU		0.89243	0.41274
K does not Granger Cause LOGYMEX	110	2.21020	0.11476
LOGYMEX does not Granger Cause K		16.4675	6.0E-07
LOGTI does not Granger Cause LOGYMEX	110	2.26100	0.10930
LOGYMEX does not Granger Cause LOGTI		0.14654	0.86387
K does not Granger Cause LOGYEEUU	110	5.70415	0.00445
LOGYEEUU does not Granger Cause K		3.29332	0.04100
LOGTI does not Granger Cause LOGYEEUU	110	1.04748	0.35446
LOGYEEUU does not Granger Cause LOGTI		0.34336	0.71017
LOGTI does not Granger Cause K	110	0.80082	0.45168
K does not Granger Cause LOGTI		2.65522	0.07500

APÉNDICE C

SERIES TEMPORALES MACROECONÓMICAS

Toda la información para las series que son utilizadas en la aplicación del modelo econométrico propuesto en este estudio corresponden al periodo 1980–2007, de forma trimestral. Posteriormente a la definición de los indicadores se presenta la información dividida en dos tablas para su consulta.

SERIE I YMEX PRODUCTO INTERNO BRUTO DE MÉXICO (PIB), en millones de pesos a precios de 1993⁵⁶. Fuente: INEGI. Sistema de Cuentas Nacionales de México. (www.inegi.gob.mx – Indicadores económicos de coyuntura › Oferta y demanda final de bienes y servicios › A precios constantes › A precios de 1993 › Oferta › Valores absolutos)

⁵⁶ A finales de 1993, el INEGI inició un ambicioso programa de investigación económico-estadística tendiente a modernizar al Sistema de Cuentas Nacionales de México (SCNM) desde el punto de vista conceptual y cuantitativo, acorde con los últimos lineamientos metodológicos de los organismos internacionales (Naciones Unidas, Fondo Monetario Internacional, Banco Mundial, entre otros). Para ello, las Cuentas Nacionales de México disponían ya de importantes acervos informativos de origen censal, múltiples encuestas sectoriales y de la explotación y aprovechamiento de información financiera y registros administrativos de entes privados y públicos. Para ese entonces había ya importantes investigaciones y elaboraciones estadísticas que pudieron ser incorporadas sobre la base del nuevo Sistema de Cuentas Nacionales de 1993 y permitieron un avance sustantivo en los cálculos.

A partir de entonces y sobre esa referencia metodológica se han venido ampliando y actualizando los diversos productos del Sistema de Cuentas Nacionales de México que tienen como base el año de 1993.

Con información de inegi.gob.mx, SCNM.

- SERIE 2 YEEUU PRODUCTO INTERNO BRUTO DE ESTADOS UNIDOS (REAL GROSS DOMESTIC PRODUCT), en miles de millones de dólares encadenados del 2000⁵⁷. Serie ajustada estacionalmente en las tasas anuales. Fuente: U.S. Bureau of Economic Analysis (BEA) (www.bea.gov – National Income and Product Accounts Table)
- SERIE 3 XT EXPORTACIONES TOTALES DE MÉXICO, en millones de pesos a precios de 1993. Fuente: INEGI. Sistema de Cuentas Nacionales de México. Cifras revisadas debido a mayor disponibilidad de información. (www.inegi.gob.mx – Indicadores económicos de coyuntura › Oferta y demanda final de bienes y servicios › A precios constantes › A precios de 1993 › Demanda › Valores absolutos)
- SERIE 4 MT IMPORTACIONES TOTALES DE MÉXICO, en millones de pesos a precios de 1993. Fuente: INEGI. Sistema de Cuentas Nacionales de México. Cifras revisadas debido a mayor disponibilidad de información. (www.inegi.gob.mx – Indicadores económicos de coyuntura › Oferta y demanda final de bienes y servicios › A precios constantes › A precios de 1993 › Oferta › Valores absolutos)
- SERIE 5 K FLUJOS NETOS DE CAPITAL, definidos contablemente como MT menos XT.

⁵⁷ Hasta los años noventa los cálculos del GDP (*Gross Domestic Product*, Producto interno bruto de los Estados Unidos) real que publicaba el Bureau of Economic Analysis se hacían considerando precios fijos sobre un año base elegido por la institución, que naturalmente presentan diferencias con respecto a otros años base al considerar en el tiempo, la producción y el precio de determinados bienes. Si el año base elegido era particularmente de los años ochenta, para los noventa cuando la producción de bienes tecnológicos se dispara y los precios retroceden, esto afectaba de un modo directo y muy importante al resto de la economía y por supuesto, el cálculo del PIB real.

Como consecuencia, los economistas oficiales que diseñan la contabilidad nacional estadounidense han adoptado un método que denominan “Chain-linking” (encadenamiento) que divide la diferencia entre usar uno de los primeros años como base o uno de los últimos. En la práctica, al momento de escribir esto, la contabilidad nacional de los estados Unidos presenta los datos del PIB real en miles de millones de dólares encadenados de 2000, como año base. Ver **Krugman y Wells** (2007).

- SERIE 6 **TI** TÉRMINOS DE INTERCAMBIO, índice base 1980 = 100.
Fuente: BANCO DE MÉXICO. (<http://www.banxico.org.mx> – Estadísticas > Índices de Precios Productor y de Comercio Exterior > Índices de Precios de Comercio Exterior > Sin unidad > Índices > Base 1980)
- SERIE 7 **XT/YMEX** PARTICIPACIÓN PORCENTUAL DE LAS EXPORTACIONES TOTALES EN EL PIB, calculada como XT entre YMEX por 100%.
- SERIE 8 **MT/YMEX** PARTICIPACIÓN PORCENTUAL DE LAS IMPORTACIONES TOTALES EN EL PIB, calculada como MT entre YMEX por 100%.
- SERIE 9 **K/YMEX** PARTICIPACIÓN PORCENTUAL DE LAS CORRIENTES NETAS DE CAPITAL EN EL PIB, calculada como K entre YMEX por 100%.
- SERIE 10 **θ (THETA)** PARTICIPACIÓN DE LAS EXPORTACIONES TOTALES EN LOS INGRESOS TOTALES DE LA BALANZA DE PAGOS, calculada como Exportaciones Totales de México (XT) entre Importaciones Totales de México (MT), a partir de la identidad básica de la balanza de pagos: $XT + K = MT$. **$\theta = XT/MT$.**
- SERIE 11 **$1 - \theta$** PARTICIPACIÓN DE LOS FLUJOS NETOS DE CAPITAL EN LOS INGRESOS TOTALES DE LA BALANZA DE PAGOS, calculada como 1 menos θ . **$1 - \theta = K/MT$.**

TABLA I. DE LA SERIE 1 A LA 6.

	YMEX	YEEUU	XT	MT	K	TI
1980/01	938,135	5,221	73,172	114,291	41,118	102.2380
1980/02	935,461	5,116	73,854	142,765	68,911	100.1350
1980/03	925,245	5,107	75,286	157,398	82,112	98.8540
1980/04	995,587	5,202	80,683	163,206	82,523	99.4530
1981/01	1,015,503	5,308	86,005	158,235	72,230	104.9610
1981/02	1,031,141	5,266	88,944	167,465	78,521	93.0040
1981/03	1,004,063	5,330	73,426	163,328	89,902	92.3480
1981/04	1,067,221	5,263	89,067	190,989	101,922	92.6860
1982/01	1,046,417	5,177	85,294	137,553	52,259	84.6280
1982/02	1,036,685	5,205	100,465	119,683	19,218	84.1400
1982/03	996,733	5,185	110,387	96,165	-14,222	84.4620
1982/04	1,016,646	5,190	117,401	69,213	-48,188	81.6010
1983/01	1,004,290	5,254	112,432	51,615	-60,817	77.5720
1983/02	986,440	5,372	118,730	66,949	-51,781	77.1560
1983/03	955,682	5,478	115,461	73,661	-41,800	76.3860
1983/04	1,007,248	5,591	125,764	87,567	-38,197	76.8670
1984/01	1,037,162	5,700	129,258	67,221	-62,036	76.7050
1984/02	1,015,362	5,798	123,715	76,272	-47,444	75.9790
1984/03	1,000,452	5,854	125,537	90,793	-34,744	76.2780
1984/04	1,035,536	5,902	121,038	95,347	-25,692	75.4420
1985/01	1,054,820	5,957	120,072	89,399	-30,673	74.1200
1985/02	1,052,454	6,008	106,067	92,818	-13,249	70.9470
1985/03	1,012,227	6,102	118,580	92,578	-26,002	70.1310
1985/04	1,058,455	6,149	132,552	91,046	-41,506	70.7460
1986/01	1,023,030	6,207	112,765	85,314	-27,451	50.0190

	YMEX	YEEUU	XT	MT	K	TI
1986/02	1,047,878	6,232	125,552	87,088	-38,464	45.9790
1986/03	964,237	6,292	123,019	82,336	-40,684	50.8770
1986/04	1,014,174	6,323	137,370	83,357	-54,013	52.4760
1987/01	1,012,635	6,365	136,908	72,341	-64,567	55.1680
1987/02	1,050,061	6,435	134,550	81,716	-52,834	57.2590
1987/03	992,042	6,493	133,850	90,872	-42,978	56.4010
1987/04	1,064,328	6,607	140,730	110,565	-30,164	51.8260
1988/01	1,038,644	6,639	150,036	95,030	-55,006	51.3550
1988/02	1,061,388	6,724	150,518	116,252	-34,266	50.0520
1988/03	993,274	6,759	139,427	134,190	-5,237	47.8400
1988/04	1,078,618	6,849	137,523	140,540	3,018	50.3830
1989/01	1,068,783	6,918	147,795	125,452	-22,343	53.0720
1989/02	1,111,605	6,964	141,885	141,015	-870	52.1540
1989/03	1,050,907	7,013	155,622	143,072	-12,550	50.5190
1989/04	1,111,908	7,031	164,897	163,796	-1,101	52.5850
1990/01	1,115,170	7,112	180,745	144,194	-36,551	50.7240
1990/02	1,156,562	7,130	132,731	157,368	24,637	46.1280
1990/03	1,102,849	7,131	157,587	184,716	27,129	65.2370
1990/04	1,193,417	7,077	171,508	200,258	28,749	56.7940
1991/01	1,157,545	7,041	164,625	166,819	2,194	49.4640
1991/02	1,221,764	7,087	171,950	193,288	21,338	50.1040
1991/03	1,140,122	7,121	167,171	204,550	37,379	50.5280
1991/04	1,241,096	7,154	171,405	226,078	54,672	46.8740
1992/01	1,211,845	7,228	181,278	216,090	34,812	48.7490
1992/02	1,249,936	7,298	179,110	237,008	57,898	50.4020
1992/03	1,191,296	7,370	170,501	238,760	68,258	50.3290
1992/04	1,276,025	7,451	177,914	253,990	76,076	49.1750

	YMEX	YEEUU	XT	MT	K	TI
1993/01	1,248,725	7,460	181,603	228,315	46,712	48.7750
1993/02	1,260,352	7,498	185,555	235,932	50,378	46.3850
1993/03	1,211,580	7,536	191,839	242,870	51,031	46.3030
1993/04	1,304,127	7,637	207,163	256,319	49,155	45.3080
1994/01	1,277,838	7,715	217,047	275,996	58,948	46.3560
1994/02	1,331,435	7,816	224,473	293,978	69,505	50.9050
1994/03	1,267,386	7,860	221,082	292,774	71,692	51.1120
1994/04	1,372,142	7,952	239,922	305,446	65,524	51.5160
1995/01	1,272,242	7,974	272,565	237,478	-35,087	50.3610
1995/02	1,209,053	7,988	289,427	235,755	-53,673	49.3320
1995/03	1,165,580	8,053	299,087	251,716	-47,371	48.6800
1995/04	1,275,557	8,112	313,951	267,533	-46,418	49.9710
1996/01	1,273,078	8,169	324,143	264,402	-59,741	51.5700
1996/02	1,287,401	8,303	339,646	289,648	-49,999	49.5630
1996/03	1,248,665	8,373	349,397	316,141	-33,256	50.1900
1996/04	1,366,292	8,471	376,063	349,412	-26,651	51.4620
1997/01	1,331,527	8,536	347,718	317,851	-29,867	50.3440
1997/02	1,395,247	8,666	383,331	363,830	-19,502	49.2450
1997/03	1,342,048	8,774	393,513	390,929	-2,584	49.0560
1997/04	1,457,278	8,838	413,559	424,403	10,843	48.3170
1998/01	1,431,862	8,936	415,907	410,656	-5,251	47.4450
1998/02	1,455,594	8,995	426,633	432,734	6,101	46.5320
1998/03	1,412,882	9,099	423,910	433,513	9,604	46.6330
1998/04	1,496,902	9,237	459,075	468,008	8,934	46.1610
1999/01	1,462,740	9,316	441,172	435,353	-5,820	47.2880
1999/02	1,506,308	9,393	478,699	486,117	7,418	48.4290
1999/03	1,475,502	9,502	497,228	513,138	15,910	50.7400

	YMEX	YEEUU	XT	MT	K	TI
1999/04	1,577,232	9,671	521,147	555,846	34,699	52.4140
2000/01	1,571,296	9,696	517,279	539,179	21,900	51.3510
2000/02	1,617,057	9,848	553,289	589,823	36,534	51.4650
2000/03	1,579,483	9,837	579,183	625,737	46,553	52.6450
2000/04	1,651,503	9,888	604,071	663,196	59,125	49.6480
2001/01	1,601,652	9,876	540,904	575,894	34,990	50.1120
2001/02	1,619,638	9,906	552,316	599,957	47,641	50.2370
2001/03	1,558,906	9,871	531,459	580,524	49,065	49.5650
2001/04	1,629,065	9,910	548,095	622,105	74,010	49.6370
2002/01	1,564,986	9,977	505,972	545,327	39,354	51.8470
2002/02	1,650,489	10,032	565,027	620,676	55,648	51.8970
2002/03	1,585,256	10,091	562,760	610,747	47,987	51.8510
2002/04	1,661,516	10,096	570,351	636,525	66,174	51.8880
2003/01	1,601,886	10,126	529,308	551,793	22,485	51.5830
2003/02	1,649,348	10,213	553,190	597,832	44,643	52.2450
2003/03	1,601,803	10,399	571,075	615,738	44,662	52.3790
2003/04	1,696,549	10,467	609,908	664,628	54,720	55.4070
2004/01	1,660,214	10,544	587,465	608,291	20,826	55.8670
2004/02	1,710,906	10,634	636,890	672,810	35,919	55.2130
2004/03	1,673,248	10,729	641,093	685,661	44,568	56.0280
2004/04	1,778,826	10,796	661,455	745,519	84,064	55.5560
2005/01	1,700,330	10,878	621,278	644,468	23,189	56.4390
2005/02	1,765,160	10,954	679,088	729,722	50,635	57.7000
2005/03	1,725,789	11,074	677,598	746,416	68,819	58.6510
2005/04	1,823,101	11,107	727,101	824,218	97,117	57.9680
2006/01	1,795,545	11,239	707,525	747,594	40,069	58.9410
2006/02	1,852,802	11,307	752,491	822,661	70,171	58.8030

	YMEX	YEEUU	XT	MT	K	TI
2006/03	1,802,588	11,337	751,895	841,235	89,340	58.8960
2006/04	1,900,767	11,396	793,540	885,080	91,540	58.4540
2007/01	1,844,784	11,413	738,144	789,504	51,360	57.0870
2007/02	1,906,175	11,520	776,704	879,899	103,195	58.3573
2007/03	1,869,887	11,659	811,588	920,031	108,443	59.7590
2007/04	1,972,745	11,676	844,059	977,749	133,689	61.0270

TABLA 2. DE LA SERIE 7 A LA 11.

	XT/YMEX	MT/YMEX	K/YMEX	THETA	1 - THETA
1980/01	7.799777%	12.182737%	4.382959%	0.64023195	0.35976805
1980/02	7.894937%	15.261473%	7.366536%	0.51731162	0.48268838
1980/03	8.136915%	17.011510%	8.874594%	0.47831823	0.52168177
1980/04	8.104052%	16.392977%	8.288925%	0.49436121	0.50563879
1981/01	8.469248%	15.581951%	7.112703%	0.54352935	0.45647065
1981/02	8.625793%	16.240745%	7.614952%	0.53112054	0.46887946
1981/03	7.312864%	16.266722%	8.953859%	0.44955975	0.55044025
1981/04	8.345709%	17.895962%	9.550253%	0.46634593	0.53365407
1982/01	8.151084%	13.145182%	4.994097%	0.62008154	0.37991846
1982/02	9.690984%	11.544766%	1.853782%	0.83942663	0.16057337
1982/03	11.074850%	9.647994%	-1.426857%	1.14789153	-0.14789153
1982/04	11.547903%	6.807959%	-4.739944%	1.69623570	-0.69623570
1983/01	11.195218%	5.139493%	-6.055725%	2.17827276	-1.17827276
1983/02	12.036192%	6.786892%	-5.249300%	1.77344691	-0.77344691
1983/03	12.081517%	7.707652%	-4.373864%	1.56747040	-0.56747040
1983/04	12.485888%	8.693685%	-3.792203%	1.43620206	-0.43620206

	XT/YMEX	MT/YMEX	K/YMEX	THETA	I - THETA
1984/01	12.462621%	6.481271%	-5.981349%	1.92286668	-0.92286668
1984/02	12.184350%	7.511777%	-4.672574%	1.62203310	-0.62203310
1984/03	12.548014%	9.075187%	-3.472827%	1.38267273	-0.38267273
1984/04	11.688480%	9.207452%	-2.481028%	1.26945872	-0.26945872
1985/01	11.383199%	8.475277%	-2.907922%	1.34310638	-0.34310638
1985/02	10.078076%	8.819208%	-1.258868%	1.14274160	-0.14274160
1985/03	11.714728%	9.145975%	-2.568753%	1.28086156	-0.28086156
1985/04	12.523112%	8.601770%	-3.921343%	1.45587627	-0.45587627
1986/01	11.022690%	8.339341%	-2.683349%	1.32176986	-0.32176986
1986/02	11.981569%	8.310932%	-3.670637%	1.44166373	-0.44166373
1986/03	12.758224%	8.538938%	-4.219286%	1.49412306	-0.49412306
1986/04	13.545027%	8.219198%	-5.325829%	1.64797425	-0.64797425
1987/01	13.519966%	7.143850%	-6.376116%	1.89253221	-0.89253221
1987/02	12.813537%	7.782051%	-5.031486%	1.64655012	-0.64655012
1987/03	13.492384%	9.160077%	-4.332307%	1.47295532	-0.47295532
1987/04	13.222388%	10.388264%	-2.834125%	1.27281988	-0.27281988
1988/01	14.445356%	9.149379%	-5.295977%	1.57883460	-0.57883460
1988/02	14.181206%	10.952779%	-3.228427%	1.29475870	-0.29475870
1988/03	14.037133%	13.509873%	-0.527260%	1.03902778	-0.03902778
1988/04	12.749906%	13.029663%	0.279757%	0.97852920	0.02147080
1989/01	13.828333%	11.737808%	-2.090525%	1.17810182	-0.17810182
1989/02	12.763968%	12.685739%	-0.078229%	1.00616668	-0.00616668
1989/03	14.808380%	13.614126%	-1.194253%	1.08772164	-0.08772164
1989/04	14.830122%	14.731067%	-0.099055%	1.00672425	-0.00672425
1990/01	16.207814%	12.930202%	-3.277613%	1.25348504	-0.25348504
1990/02	11.476375%	13.606546%	2.130171%	0.84344515	0.15655485

	XT/YMEX	MT/YMEX	K/YMEX	THETA	I - THETA
1990/03	14.289104%	16.748969%	2.459864%	0.85313339	0.14686661
1990/04	14.371203%	16.780186%	2.408983%	0.85643886	0.14356114
1991/01	14.221941%	14.411456%	0.189515%	0.98684972	0.01315028
1991/02	14.073929%	15.820421%	1.746492%	0.88960523	0.11039477
1991/03	14.662542%	17.941080%	3.278538%	0.81726084	0.18273916
1991/04	13.810788%	18.215950%	4.405162%	0.75817006	0.24182994
1992/01	14.958841%	17.831478%	2.872637%	0.83890080	0.16109920
1992/02	14.329512%	18.961592%	4.632080%	0.75571252	0.24428748
1992/03	14.312257%	20.042026%	5.729770%	0.71411226	0.28588774
1992/04	13.942859%	19.904801%	5.961942%	0.70047720	0.29952280
1993/01	14.543069%	18.283866%	3.740797%	0.79540449	0.20459551
1993/02	14.722436%	18.719556%	3.997120%	0.78647358	0.21352642
1993/03	15.833772%	20.045728%	4.211956%	0.78988263	0.21011737
1993/04	15.885217%	19.654429%	3.769212%	0.80822580	0.19177420
1994/01	16.985498%	21.598641%	4.613143%	0.78641513	0.21358487
1994/02	16.859479%	22.079768%	5.220289%	0.76357138	0.23642862
1994/03	17.443925%	23.100586%	5.656660%	0.75512914	0.24487086
1994/04	17.485195%	22.260531%	4.775336%	0.78547969	0.21452031
1995/01	21.424028%	18.666107%	-2.757920%	1.14775017	-0.14775017
1995/02	23.938353%	19.499118%	-4.439234%	1.22766334	-0.22766334
1995/03	25.659916%	21.595745%	-4.064171%	1.18819311	-0.18819311
1995/04	24.612845%	20.973804%	-3.639042%	1.17350413	-0.17350413
1996/01	25.461368%	20.768709%	-4.692659%	1.22594850	-0.22594850
1996/02	26.382333%	22.498628%	-3.883705%	1.17261966	-0.17261966
1996/03	27.981655%	25.318336%	-2.663319%	1.10519327	-0.10519327
1996/04	27.524327%	25.573747%	-1.950580%	1.07627273	-0.07627273
1997/01	26.114198%	23.871149%	-2.243049%	1.09396485	-0.09396485

	XT/YMEX	MT/YMEX	K/YMEX	THETA	I - THETA
1997/02	27.474064%	26.076352%	-1.397712%	1.05360076	-0.05360076
1997/03	29.321793%	29.129268%	-0.192525%	1.00660933	-0.00660933
1997/04	28.378893%	29.122960%	0.744067%	0.97445083	0.02554917
1998/01	29.046558%	28.679852%	-0.366706%	1.01278619	-0.01278619
1998/02	29.309895%	29.729024%	0.419129%	0.98590168	0.01409832
1998/03	30.003194%	30.682905%	0.679711%	0.97784725	0.02215275
1998/04	30.668315%	31.265116%	0.596801%	0.98091160	0.01908840
1999/01	30.160672%	29.762816%	-0.397856%	1.01336755	-0.01336755
1999/02	31.779631%	32.272126%	0.492495%	0.98473929	0.01526071
1999/03	33.698872%	34.777160%	1.078288%	0.96899437	0.03100563
1999/04	33.041871%	35.241860%	2.199989%	0.93757456	0.06242544
2000/01	32.920544%	34.314268%	1.393724%	0.95938355	0.04061645
2000/02	34.215766%	36.475066%	2.259300%	0.93805905	0.06194095
2000/03	36.669184%	39.616550%	2.947366%	0.92560267	0.07439733
2000/04	36.577051%	40.157108%	3.580057%	0.91084873	0.08915127
2001/01	33.771653%	35.956263%	2.184609%	0.93924259	0.06075741
2001/02	34.101204%	37.042671%	2.941467%	0.92059248	0.07940752
2001/03	34.091787%	37.239201%	3.147414%	0.91548117	0.08451883
2001/04	33.644757%	38.187850%	4.543093%	0.88103303	0.11896697
2002/01	32.330795%	34.845461%	2.514667%	0.92783373	0.07216627
2002/02	34.233923%	37.605556%	3.371633%	0.91034216	0.08965784
2002/03	35.499630%	38.526718%	3.027088%	0.92142887	0.07857113
2002/04	34.327162%	38.309886%	3.982724%	0.89603926	0.10396074
2003/01	33.042831%	34.446461%	1.403630%	0.95925185	0.04074815
2003/02	33.539902%	36.246594%	2.706692%	0.92532563	0.07467437
2003/03	35.652030%	38.440290%	2.788260%	0.92746517	0.07253483
2003/04	35.949901%	39.175290%	3.225389%	0.91766776	0.08233224

	XT/YMEX	MT/YMEX	K/YMEX	THETA	1 – THETA
2004/01	35.384915%	36.639321%	1.254406%	0.96576340	0.03423660
2004/02	37.225336%	39.324775%	2.099439%	0.94661283	0.05338717
2004/03	38.314269%	40.977849%	2.663580%	0.93499952	0.06500048
2004/04	37.184943%	41.910760%	4.725818%	0.88724095	0.11275905
2005/01	36.538688%	37.902511%	1.363824%	0.96401759	0.03598241
2005/02	38.471741%	41.340291%	2.868550%	0.93061128	0.06938872
2005/03	39.263076%	43.250732%	3.987656%	0.90780141	0.09219859
2005/04	39.882646%	45.209662%	5.327016%	0.88217084	0.11782916
2006/01	39.404477%	41.636049%	2.231572%	0.94640289	0.05359711
2006/02	40.613645%	44.400933%	3.787288%	0.91470252	0.08529748
2006/03	41.711951%	46.668182%	4.956232%	0.89379849	0.10620151
2006/04	41.748404%	46.564343%	4.815939%	0.89657454	0.10342546
2007/01	40.012493%	42.796539%	2.784046%	0.93494693	0.06505307
2007/02	40.746724%	46.160424%	5.413700%	0.88271989	0.11728011
2007/03	43.403060%	49.202481%	5.799420%	0.88213155	0.11786845
2007/04	42.786039%	49.562844%	6.776805%	0.86326843	0.13673157

BIBLIOGRAFÍA Y REFERENCIAS

- ATESOGLU, H. S. (1997) “Balance-of-payments-constrained growth model and its implications for the United States”, *Journal of Post Keynesian Economics*, 3: 327-35.
- BAIRAM, E. I. AND DEMPSTER, G. J. (1991) “The Harrod foreign trade multiplier and economic growth in Asian countries”, *Applied Economics*, 23.
- BLANCO, H. (1994), *Las Negociaciones Comerciales de México con el Mundo*, Fondo de Cultura Económica.
- BOX, GEORGE P. E. AND JENKINS, GWILYM M. (1976), *Time Series Analysis, Forecasting and Control*, San Francisco: Holden Day.
- COOLEY, T. F. AND LEROY, S. F. (1985) “Atheoretical Macroeconometrics: A critique”, *Journal of Monetary Economics*, 16: 283-308.
- DAVIDSON, PAUL. (1990-1991) “A post Keynesian positive contribution to “theory””, *Journal of Post Keynesian Economics*, 2: 298-303.
- DICKEY, D. A. AND FULLER, W. A. (1979) “Distribution of the Estimators for Autoregressive Time Series with a Unit Root”, *Journal of the American Statistical Association*, 74: 427-31.
- DICKEY, D. A. AND FULLER, W. A. (1981) “Likelihood ratio statistics for autorregresive time series with a unit root”, *Econometrica*, Vol. 49: 1057-72.
- ENDERS, WALTER. (2004), *Applied Econometrics Time Series*, Second Edition, J. Wiley Series in Probability and Statistics.
- ENGLE, R. AND GRANGER, C. W. J. (1987) “Cointegration and error correction: representation, estimation and testing”, *Econometrica*, 55: 251-76.
- GRANGER, C. W. J. AND WEISS, A. A. (1983) “Time Series Analysis of Error-Correcting Models”, *Studies in Econometrics, Time Series, and Multivariate Statistics*. New York: Academic Press: 255-78.

- GRANGER, C. W. J. (1969) “Investigating Causal Relations by Econometric Models and Cross-Spectral Methods,” *Econometrica*, 37: 424–38.
- GRANGER, C. W. J. (1981) “Some Properties of Time Series Data and Their Use in Econometric Model Specification,” *Journal of Econometrics*: 121-30.
- GRANGER, C. W. J. AND LEE, T. H. (1990) “Multicointegration,” G. F. Rhodes, Jr and T. B. Fomby (eds), *Advances in Econometrics: Cointegration, Spurious Regressions and Unit Roots*, JAI Press, New York: 17-84.
- GRANGER, C. W. J. AND NEWBOLD, P. (1977), *Forecasting Economic Time Series*. New York: Academic Press.
- GREENE, WILLIAM H. (1999), *Análisis econométrico*, Tercera edición, Madrid, Pearson Educación.
- GUERRERO, C. (2004) “Determinantes del Crecimiento Económico en México 1929-2003. Una perspectiva postkeynesiana”, *Cátedra “Estudios de México” ITESM*, Campus Ciudad de México.
- GUJARATI, DAMODAR N. (2004), *Econometría*, McGraw-Hill, Cuarta edición, Interamericana.
- GURRÍA, J. A. (1994) “Flujos de capital: el caso de México”, *Serie Financiamiento del Desarrollo*, CEPAL, 27.
- HAMILTON, JAMES D. (1994), *Time Series Analysis*, Princeton University Press.
- JOHANSEN, S. (1992) “Testing weak exogeneity and the order of cointegration in UK money demand data”, *Journal of Policy Modelling*, vol.14, 3: 313-34.
- JOHANSEN, S. (1995) *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press.
- JOHANSEN, S. (1988) “Statistical Analysis of Cointegration Vectors,” *Journal of Economic Dynamics and Control*, 12: 231-54.
- JOHANSEN, S. (1991) “Estimation and Hypotesis Testing of Cointegration Vectors in Gaussian Vector Autorregresive Models,” *Econometrica*, 59: 1551-80.
- JOHNSTON, JACK AND DINARDO, JOHN (1997), *Econometric Methods*, Fourth Edition, McGraw-Hill.
- KRUGMAN, PAUL Y WELLS, ROBIN (2007), *Introducción a la economía. Macroeconomía*, Reverté.
- LÓPEZ, J. (1996) “El costo de la fuga de capitales en México”, *Investigación económica*, Vol. LVI, 218: 51-71.

- MADDALA, G. S. (2001), *Introduction to Econometrics*, John Wiley & Sons.
- MCCOMBIE, J. S. L. AND THIRLWALL, A. P. (1994), *Economic Growth and the Balance-of-Payments Constraint*, St. Martin's Press.
- MORENO-BRID, J. C. (1999) “México's economic growth and the balance of payments constraint: a cointegration analysis”, *International Review of Applied Economics*, 2: 478-486.
- MORENO-BRID, J. C. Y PÉREZ, E. (1999) “Balance-of-payments-constrained growth in Central America: 1950-96”, *Journal of Post Keynesian Economics*, 1.
- PREBISCH, RAUL. (1949) “Desarrollo Económico de América Latina y sus principales problemas”, CEPAL.
- PUYANA, J. (2003) “Liberalización mundial de capitales, desequilibrios macroeconómicos y recesión mundial”, *Financiamiento del Desarrollo con Mercados de Dinero y Capital Globalizados*, G. Mantey y N. Levy coordinadoras, Porrúa.
- SCHMIDT, STEPHEN J. (2005), *Econometría*, McGraw-Hill.
- SIMS, CHRISTOPHER. (1972) “Money, Income and Causality,” *American Economic Review*, 62: 540-42.
- SIMS, CHRISTOPHER. (1980) “Macroeconomics and Reality,” *Econometrica*, Vol. 48, 1.
- SPANOS, A. (1986), *Statistical Foundations of Econometric Modelling*, Cambridge University Press.
- STOCK, J. S. AND WATSON, M. (1988) “Variable Trends in Economic Time Series,” *Journal of Economic Perspectives*, Vol 2, No. 3.
- THIRLWALL, A. P. (1979) “The balance of payments constraint as an explanation of international growth rate differences”, *Banca Nazionale del Lavoro Quarterly Review*, 128: 45-53.
- THIRLWALL, A. P. AND HUSSAIN, M. N. (1982) “The balance of payments constraint, capital flow and growth rate differences between countries”, *Oxford Economic Papers*, 3: 498-510.
- WHITE, H. (1980) “Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity” *Econometrica*, 48: 817-838.