

# Ejemplo de análisis de cointegración con Gretl

Marta Regúlez Castillo  
Departamento de Economía Aplicada III  
(Econometría y Estadística) (UPV-EHU)

Noviembre 2006

# Índice general

1.1. Datos . . . . .	1
1.2. Contrastes de raíces unitarias a nivel univariante . . . . .	2
1.3. Contrastes de cointegración . . . . .	5
1.4. Modelo de corrección del error (VECM) . . . . .	9

## 1.1. Datos

Este ejemplo utiliza los datos del ejemplo 8.6.5 del libro de Ruey S. Tsay, *Multivariate Time Series Analysis and its applications*. Se consideran dos series de tipos de interés: Weekly U.S. interest rates (3-month Treasury Bill y 6-month Treasury Bill) que se encuentran en las dos primeras columnas del fichero **w-tb3n6ms.txt**. Son datos semanales desde el 12 de diciembre de 1958 al 6 de agosto de 2004, con un total de 2383 observaciones.

Leemos los datos en Gretl

**Archivo** → **Abrir datos** → **Importar** → **ASCII**

Seleccionamos v3,v4,v5 que no nos interesan con la tecla de mayúsculas y el botón izquierdo. Con el botón derecho aparece una ventana en la que tenemos la opción de borrar. Informamos a Gretl del tipo de datos

**Datos** → **estructura del conjunto de datos** → **serie temporal**

→ **semanal** → **viernes** → **1958/12/12**

Y le damos nombre a las variables

**Variable** → **Editar atributos**

utilizando las etiquetas TB3 y TB6 respectivamente además de dar una pequeña descripción que aparecerá al lado del nombre. Guardamos los datos en formato gretl (**Archivo** → **guardar datos como**). Podemos obtener

**Variable** → **Estadísticos principales**

Estadísticos principales, usando las observaciones 58/12/12 - 04/08/06 para la variable TB3 (2383 observaciones válidas)

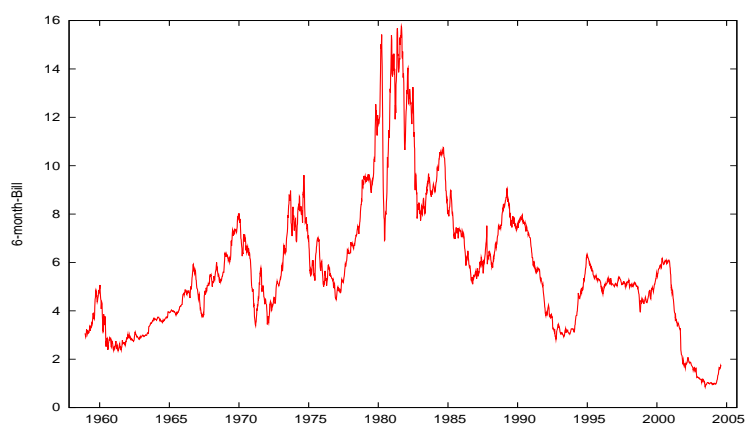
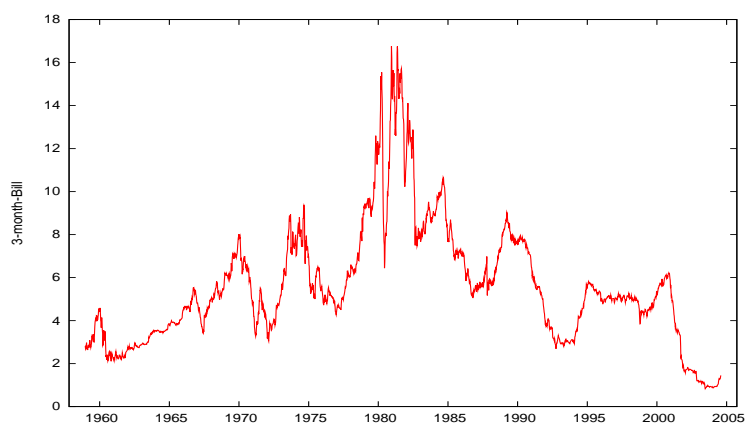
Media	Mediana	Mínimo	Máximo
5,59568	5,16000	0,840000	16,7600
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
2,76677	0,494447	1,10276	1,84815

Estadísticos principales, usando las observaciones 58/12/12 - 04/08/06 para la variable TB6 (2383 observaciones válidas)

Media	Mediana	Mínimo	Máximo
5,74950	5,31000	0,850000	15,7600
Desv. Típ.	C.V.	Asimetría	Exc. de curtosis
2,72353	0,473699	0,989282	1,44616

## 1.2. Contrastes de raíces unitarias a nivel univariante

Mirando a los gráficos de las series parecen no presentar una tendencia determinista clara pero sí una tendencia estocástica aunque posiblemente sin deriva. Dado que la media parece ser distinta de cero consideramos en el contraste de ADF incorporar una constante en la regresión aunque la hipótesis nula es la de raíz unitaria pero sin deriva (Caso 2 del Hamilton).



## 1.2. CONTRASTES DE RAÍCES UNITARIAS A NIVEL UNIVARIANTE3

Consideramos la regresión

$$\Delta TB3_t = \nu + \gamma TB3_{t-1} + \sum_{j=1}^{4-1} \gamma_j \Delta TB3_{t-j} + u_t$$

$$H_0 : \gamma = 0, \text{ versus } H_1 : \gamma < 0$$

Una vez hemos definido las primeras diferencias de TB3 y TB6 con

**Añadir** → **Primeras diferencias de las variables seleccionadas**

y sus retardos con

**Variable** → **definir nueva variable** →  $d - TB3 - j = d - TB3(-j)$

siendo  $j$  el retardo correspondiente, estimamos la regresión anterior con

**Modelo** → **Mínimos Cuadrados Ordinarios**

Estimaciones MCO utilizando las 2379 observaciones 59/01/09–04/08/06

Variable dependiente: d\_TB3

Variable	Coefficiente	Desv. típica	Estadístico $t$	valor p
const	0,0204353	0,00950288	2,1504	0,0316
TB3_1	-0,00371135	0,00152221	-2,4381	0,0148
d_TB3_1	0,271457	0,0204924	13,2468	0,0000
d_TB3_2	-0,0148460	0,0212326	-0,6992	0,4845
d_TB3_3	0,0381931	0,0205139	1,8618	0,0628

El **valor crítico para el estadístico**  $t_{\hat{\gamma}}$  bajo la hipótesis nula de raíz unitaria ( $\gamma = 0$ ) en el caso 2 (modelo bajo  $H_0$  sin deriva) es  $-2,86$  al  $0,05$  de significación y el tamaño muestral considerado (ver Tabla B.6, Case 2, columna  $0,05$ ). Dado que el valor muestral del estadístico  $-2,4381$  es mayor que el valor crítico  $-2,86$ , no cae en la zona de rechazo, por lo que no se rechaza la hipótesis de raíz unitaria en la serie TB3.

Lo mismo hacemos para la serie TB6. Consideramos la regresión

$$\Delta TB6_t = \nu + \gamma TB6_{t-1} + \sum_{j=1}^{4-1} \gamma_j \Delta TB6_{t-j} + u_t$$

Estimaciones MCO utilizando las 2379 observaciones 59/01/09–04/08/06

Variable dependiente: d\_TB6

Variable	Coefficiente	Desv. típica	Estadístico $t$	valor p
const	0,0188423	0,00868102	2,1705	0,0301
TB6_1	-0,00332840	0,00136431	-2,4396	0,0148
d_TB6_1	0,273770	0,0204870	13,3631	0,0000
d_TB6_2	0,0535491	0,0212198	2,5235	0,0117
d_TB6_3	0,0408834	0,0205125	1,9931	0,0464

Dado que en este caso el valor muestral del estadístico  $-2,4396$  es mayor que el valor crítico  $-2,86$ , no cae en la zona de rechazo, por lo que no se rechaza la hipótesis de raíz unitaria en la serie TB6.

Estos contrastes los podemos hacer también en el menú utilizando

**Variable → Contraste aumentado de Dickey Fuller**

y eligiendo en la ventana, el orden del retardo  $p - 1$ , distintas opciones para los términos deterministas en la regresión, etc. Por ejemplo para TB3 obtenemos.

Contrastes aumentados de Dickey-Fuller, orden 3, para TB3: tamaño muestral 2379 hipótesis nula de raíz unitaria:  $a = 1$

```

contraste sin constante
modelo: (1 - L)y = (a-1)*y(-1) + ... + e
Coef. de autocorrelación de primer orden de e: -0,003
valor estimado de (a - 1): -0,000773973
Estadístico de contraste: tau_nc(1) = -1,15121
valor p asintótico 0,2282

```

```

contraste con constante
modelo: (1 - L)y = b0 + (a-1)*y(-1) + ... + e
Coef. de autocorrelación de primer orden de e: -0,003
valor estimado de (a - 1): -0,00371135
Estadístico de contraste: tau_c(1) = -2,43813
valor p asintótico 0,1312

```

Valores p basados en el artículo de MacKinnon (JAE, 1996)

El segundo caso sería el que hemos considerado anteriormente y se obtiene el mismo valor del estadístico  $t$  cuyo valor p tiene en cuenta que no sigue una distribución estandar. No se rechaza la hipótesis nula de raíz unitaria en TB3 al 0,05.

También se tiene la opción de realizar el contraste KPSS. En este caso la hipótesis nula es la de estacionariedad de la serie. En el caso de TB3 eligiendo el orden de truncamiento en el retardo 3 y sin tendencia se obtiene

Contraste KPSS para TB3 (sin tendencia)

Parámetro de truncamiento de los retardos = 3

Estadístico de contraste = 8,99282

	10%	5%	2,5%	1%
Valores críticos:	0,347	0,463	0,574	0,739

Se rechaza la hipótesis nula que la serie sea estacionaria. Para TB6, ocurre lo mismo:

Contraste KPSS para TB6 (sin tendencia)

Parámetro de truncamiento de los retardos = 3

Estadístico de contraste = 9,29618

	10%	5%	2,5%	1%
Valores críticos:	0,347	0,463	0,574	0,739

### 1.3. Contrastes de cointegración

Vamos a utilizar primeramente el contraste de Engle y Granger

**Modelo → Series temporales → Contraste de cointegración**

→ **Engle-Granger**

Se eligen las series que se quieren contrastar si están o no cointegradas. Gretl utiliza la primera como la dependiente en la regresión cointegrante. Se elige si la regresión cointegrante tiene o no constante u otras variables deterministas como tendencia, y el orden del retardo para hacer el contraste ADF en las series individuales y en la del residuo de la regresión cointegrante.

Aunque se puede omitir que muestre los resultados de los contrastes univariantes de ADF en las series, también los muestra.

Etapa 1: contrastando la existencia de una raíz unitaria en TB3

Contraste aumentado de Dickey-Fuller, orden 3, para TB3 tamaño muestral 2379 hipótesis nula de raíz unitaria:  $a = 1$

```

contraste con constante
valor estimado de (a - 1): -0,00371135
Estadístico de contraste: tau_c(1) = -2,43813
valor p asintótico 0,1312

```

Etapa 2: contrastando la existencia de una raíz unitaria en TB6

Contraste aumentado de Dickey-Fuller, orden 3, para TB6 tamaño muestral 2379 hipótesis nula de raíz unitaria:  $a = 1$

contraste con constante  
 valor estimado de  $(a - 1)$ : -0,0033284  
 Estadístico de contraste:  $\tau_c(1) = -2,43963$   
 valor p asintótico 0,1308

Etapa 3: regresión cointegrante

Regresión cointegrante - estimaciones MCO utilizando las 2383 observaciones 58/12/12-04/08/06 Variable dependiente: TB3

VARIABLE	COEFICIENTE	DESV.TÍP.	ESTAD T	VALOR P
const	-0,227230	0,0103472	-21,961	<0,00001 ***
TB6	1,01277	0,00162648	622,674	<0,00001 ***

R-cuadrado = 0,993896  
 R-cuadrado corregido = 0,993894  
 Estadístico de Durbin-Watson = 0,164916  
 Coef. de autocorr. de primer orden. = 0,917536  
 Criterio de información de Akaike (AIC) = -534,739  
 Criterio de información Bayesiano de Schwarz (BIC) = -523,187  
 Criterio de Hannan-Quinn (HQC) = -530,535

Etapa 4: Contraste de Dickey-Fuller sobre los residuos

Orden de retardos 3 tamaño muestral 2379 hipótesis nula de raíz unitaria:  $a = 1$

valor estimado de  $(a - 1)$ : -0,0714629  
 Estadístico de contraste:  $\tau_c(2) = -8,40176$   
 valor p asintótico 3,55e-013

Valores p basados en el artículo de MacKinnon (JAE, 1996)

Hay evidencia de una relación cointegrante si: (a) La hipótesis de existencia de raíz unitaria no se rechaza para las variables individuales. (b) La hipótesis de existencia de raíz unitaria se rechaza para los residuos (uhat) de la regresión cointegrante.



Dados los resultados se confirman (a) y (b) por lo que hay evidencia de cointegración entre TB3 y TB6. El vector de cointegración estimado por la regresión de cointegración, normalizando con respecto a TB3, es  $\hat{\beta}'_{EG} = (1, -1, 0.12)$  tal que el residuo de cointegración  $\hat{w}_t = \hat{\beta}'_{EG}y_t + 0,227230$  es estacionario.

Consideremos ahora realizar los contrastes de la traza y del menor valor propio de Johansen:

### Modelo $\rightarrow$ Series temporales $\rightarrow$ Contraste de cointegración

#### $\rightarrow$ Johansen

Seleccionamos en la ventana el orden del retardo (del VECM asociado), en este caso seguimos eligiendo 3 retardos ( $p-1=3$ , siendo  $p$  el orden del VAR para las variables en niveles). Las variables a contrastar TB3 y TB6. Además hay varias opciones dependiendo de **cómo tratar a los términos deterministas del VAR** (constante y tendencia  $t$ ), llamémoslos  $\nu_t$ .

$$\Delta y_t = \nu_t + \alpha\beta'y_{t-1} + D_1\Delta y_{t-1} + \dots + D_{p-1}\Delta y_{t-p+1} + u_t$$

- Sin constante  $\nu_t = 0$ : Si en el VAR o VECM no incluimos ni siquiera constante elegiríamos la primera opción, sin constante. Por ejemplo en el caso que tenemos claro que las series son I(1) sin deriva y las series  $w_t = \beta'y_t$  tienen media cero.
- Constante restringida  $\nu_t = \nu_0 = \alpha c_0$  donde  $c_0$  es un vector  $k$ -dimensional de constantes distintas de cero. Por tanto, el VECM bajo este caso de constante restringida es

$$\Delta y_t = \alpha(\beta'y_{t-1} + c_0) + D_1\Delta y_{t-1} + \dots + D_{p-1}\Delta y_{t-p+1} + u_t$$

tal que los componentes de  $y_t$  son I(1) sin deriva o drift y el vector de residuos de cointegración  $w_t = \beta'y_t + c_0$  tiene de vector de medias a un vector de constantes distinto de cero.

- Constante no restringida  $\nu_t = \nu_0$  que es distinta de cero. Aquí las componentes de  $y_t$  son I(1) con deriva  $\nu_0$  y  $w_t$  puede tener una media distinta de cero.
- Tendencia restringida  $\nu_t = \nu_0 + \alpha c_1 t$

Este caso implica que hemos incorporado tanto un término constante como una tendencia determinista  $t$  pero entrarían en el VECM como

$$\Delta y_t = \nu_0 + \alpha(\beta'y_{t-1} + c_1 t) + D_1\Delta y_{t-1} + \dots + D_{p-1}\Delta y_{t-p+1} + u_t$$

tal que los componentes de  $y_t$  son I(1) con deriva  $\nu_0$  y  $w_t$  es estacionario alrededor de una tendencia lineal  $c_1 t$ .

- Tendencia no restringida  $\nu_t = \nu_0 + \nu_1 t$  donde  $\nu_1 \neq 0$ . En este caso tanto el término constante como la tendencia están sin restringir. Los componentes de  $y_t$  son I(1) y presentan una tendencia cuadrática determinista y  $w_t$  tiene una tendencia lineal.

En nuestro caso elegimos un orden del retardo en el VECM igual a 3 (que implica en el VAR un orden de  $p=4$ ) y constante restringida, considerando que los tipos de interés sean I(1) pero sin deriva. Obtenemos los siguientes resultados

Contraste de Johansen:

Número de ecuaciones = 2 Orden del retardo = 3

Periodo de estimación: 59/01/02 - 04/08/06 (T = 2380)

Caso 2: Constante restringida

Rango	Valor propio	Estad. traza [valor p]	Estad. Lmáx [valor p]
0	0,032151	83,271 [0,0000]	77,778 [0,0000]
1	0,0023056	5,4936 [0,2421]	5,4936 [0,2416]

Se rechaza la hipótesis nula de rango de cointegración 0 tanto con el estadístico de la traza como el del mayor valor propio. No se rechaza que sea 1, por lo que hay evidencia de que las dos series están cointegradas. Por tanto, hay una relación de cointegración y una tendencia estocástica común.

valor propio      0,032151      0,0023056

beta (vectores cointegrantes)

TB3                      -4,8412      -0,014602

TB6                      4,9014      0,38257

const                   -1,0912      -2,0799

alfa (vectores de ajuste)

TB3                      0,019594      -0,0081349

TB6                      0,0043613      -0,0085909

beta renormalizado

TB3	1,0000	-0,038168
TB6	-1,0124	1,0000
const	0,22540	-5,4366

alfa renormalizado

TB3	-0,094860	-0,0031122
TB6	-0,021114	-0,0032867

matriz de largo plazo (alfa \* beta')

	TB3	TB6	const
TB3	-0,094741	0,092927	-0,0044615
TB6	-0,020988	0,018090	0,013109

Dado que solamente hay un vector de cointegración, elegimos la primera columna de beta renormalizado (aquí se ha normalizado por el coeficiente de TB3), y la primera columna de alfa renormalizado. En el VECM estimado tenemos, además de las matrices que acompañan a  $\Delta y_{t-j}$ ,  $j = 1, 2, 3$  se obtiene:

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} -0,09 \\ -0,02 \end{bmatrix} \begin{bmatrix} 1 & -1,01 & 0,23 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ 1 \end{bmatrix} + D_1 \Delta y_{t-1} + D_2 \Delta y_{t-2} + D_3 \Delta y_{t-3} + u_t$$

## 1.4. Modelo de Corrección del error (VECM)

Podemos obtener la estimación completa del VECM

**Modelo** → **series temporales** → **VECM**

Seleccionamos orden del retardo 3, rango de cointegración 1, Variables endógenas TB3 y TB6 y finalmente constante restringida, y obtenemos

Sistema VECM, orden del retardo 3  
 estimaciones Máxima Verosimilitud, observaciones 59/01/02–04/08/06  
 ( $T = 2380$ )  
 Rango de cointegración = 1  
 Caso 2: Constante restringida

Vectores cointegrantes (Desviaciones típicas entre paréntesis)

TB3 <sub>t-1</sub>	1,00000 (0,000000)
TB6 <sub>t-1</sub>	-1,01244 (0,00855065)
const	0,225399 (0,0544098)

Log-verosimilitud = 3212,68

Determinante de la matriz de covarianzas = 0,000230449

AIC = -2,6896

BIC = -2,6605

HQC = -2,6790

Ecuación 1:  $\Delta$ TB3

Variable	Coefficiente	Desv. típica	Estadístico $t$	valor p
$\Delta$ TB3 <sub>t-1</sub>	0,0465648	0,0479721	0,9707	0,3318
$\Delta$ TB3 <sub>t-2</sub>	-0,206708	0,0480383	-4,3030	0,0000
$\Delta$ TB6 <sub>t-1</sub>	0,265020	0,0537407	4,9314	0,0000
$\Delta$ TB6 <sub>t-2</sub>	0,254742	0,0542175	4,6985	0,0000
EC1 <sub>t-1</sub>	-0,0948597	0,0199116	-4,7641	0,0000

Ecuación 2:  $\Delta$ TB6

Variable	Coefficiente	Desv. típica	Estadístico $t$	valor p
$\Delta$ TB3 <sub>t-1</sub>	-0,0419023	0,0431569	-0,9709	0,3317
$\Delta$ TB3 <sub>t-2</sub>	-0,0346324	0,0432164	-0,8014	0,4230
$\Delta$ TB6 <sub>t-1</sub>	0,316447	0,0483464	6,5454	0,0000
$\Delta$ TB6 <sub>t-2</sub>	0,0993926	0,0487754	2,0378	0,0417
EC1 <sub>t-1</sub>	-0,0211139	0,0179129	-1,1787	0,2386

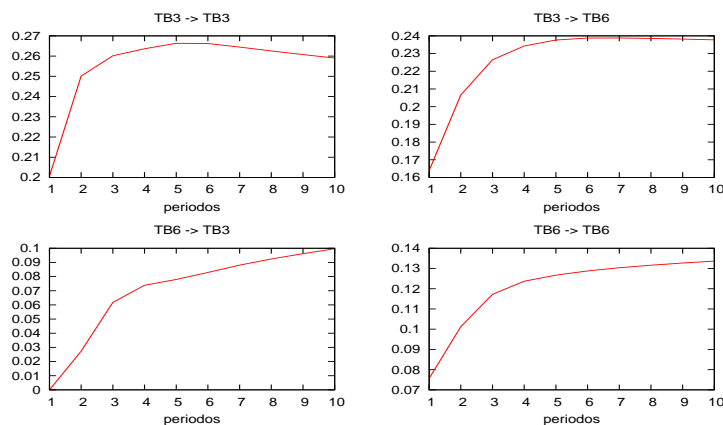
Matriz de covarianzas cruzadas entre ecuaciones

	$\Delta$ TB3	$\Delta$ TB6
$\Delta$ TB3	0,0402604	0,0328844
$\Delta$ TB6	0,0328844	0,0325837

determinante = 0,000230449

En la ventana asociada a los resultados de estimación del VECM también aparecen distintas opciones. En **Análisis** se encuentra la posibilidad de utilizar el VECM estimado para predecir y para el análisis impulso-respuesta.

### Análisis Impulso-Respuesta:



Respuestas a un shock de tamaño una desviación típica en TB3

periodo	TB3	TB6
1	0,200650	0,163889
2	0,250133	0,206611
3	0,260149	0,226532
4	0,263627	0,234299
5	0,266347	0,237686
6	0,266192	0,238753
7	0,264447	0,238823
8	0,262539	0,238551
9	0,260763	0,238169
10	0,259042	0,237747

Respuestas a un shock de tamaño una desviación típica en TB6

periodo	TB3	TB6
1	0,000000	0,0756568
2	0,0273166	0,101215
3	0,0617646	0,117265
4	0,0738896	0,123698
5	0,0779976	0,126712
6	0,0828904	0,128774
7	0,0880880	0,130382
8	0,0924361	0,131636
9	0,0961797	0,132692
10	0,0996748	0,133650

### Predicción

Para predecir considerando como origen de predicción el 6 de agosto de 2006, tenemos que ampliar el rango de observación con

#### Datos → Añadir observaciones

Por ejemplo añadimos 10 observaciones más, pudiendo así obtener las predicciones de  $h$ -periodos hacia adelante con  $h = 1, 2, \dots, 10$ .

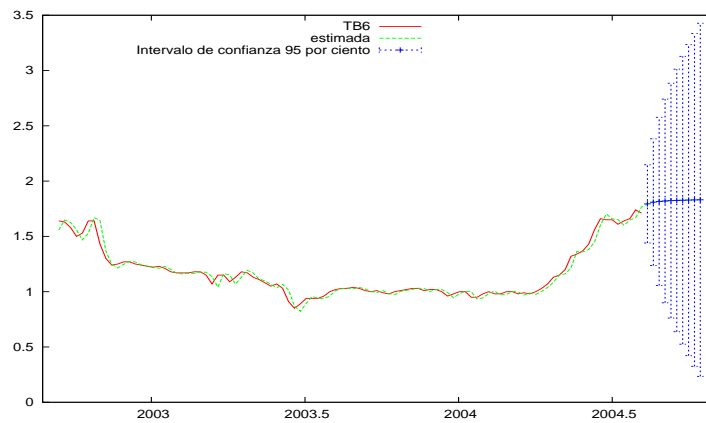
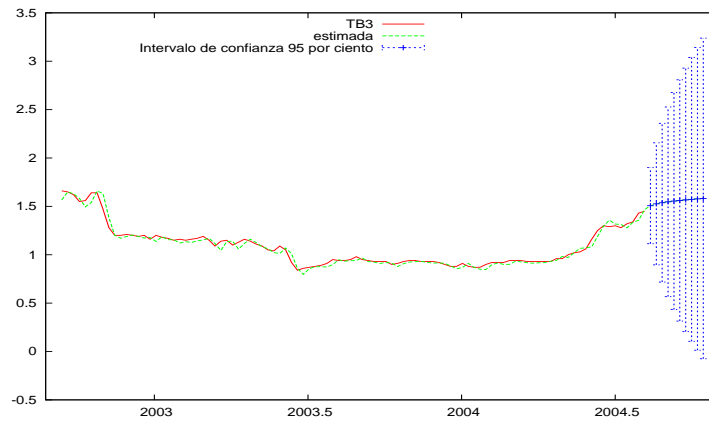
Volvemos a la pantalla de estimación del VECM y seleccionamos

#### Análisis → Predicciones → TB3

Observaciones	TB3	predicción	desv. típica	Intervalo de confianza 95 %	
				low	high
04/08/13		1,51	0,201	1,11	1,90
04/08/20		1,53	0,322	0,90	2,16
04/08/27		1,54	0,418	0,72	2,36
04/09/03		1,55	0,500	0,57	2,53
04/09/10		1,55	0,572	0,43	2,68
04/09/17		1,56	0,636	0,31	2,81
04/09/24		1,57	0,695	0,21	2,93
04/10/01		1,57	0,748	0,11	3,04
04/10/08		1,58	0,798	0,01	3,14
04/10/15		1,58	0,845	-0,07	3,24

#### Análisis → Predicciones → TB6

Observaciones	TB6	predicción	desv. típica	Intervalo de confianza 95 %	
				low	high
04/08/13		1,79	0,181	1,44	2,15
04/08/20		1,81	0,292	1,24	2,38
04/08/27		1,82	0,388	1,06	2,58
04/09/03		1,82	0,470	0,90	2,74
04/09/10		1,82	0,542	0,76	2,88
04/09/17		1,83	0,606	0,64	3,01
04/09/24		1,83	0,664	0,53	3,13
04/10/01		1,83	0,718	0,42	3,24
04/10/08		1,83	0,768	0,32	3,33
04/10/15		1,83	0,815	0,23	3,43



Debido a la no-estacionariedad por la raíz unitaria, los intervalos de predicción son muy amplios y no son informativos.