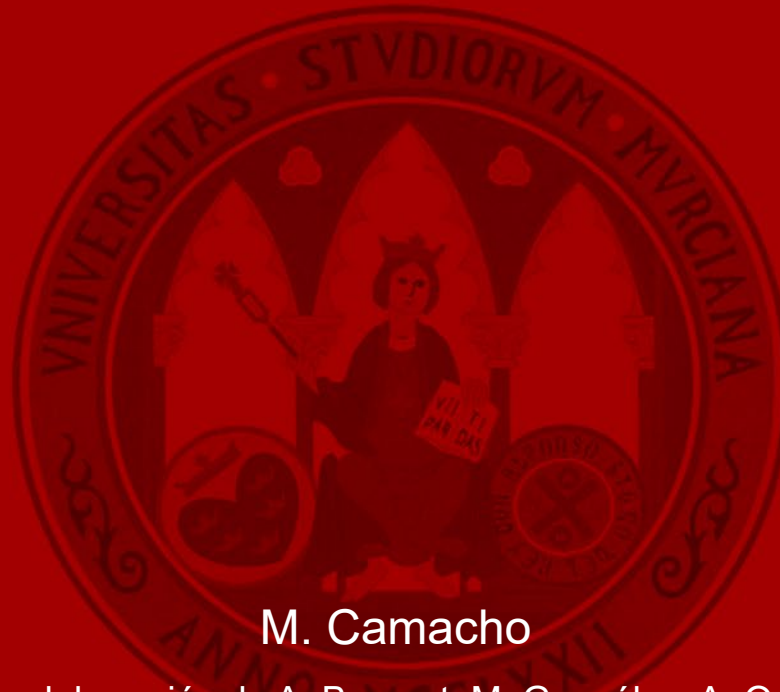


# TEMA 2: MODELOS VAR y MCE



M. Camacho

con la colaboración de A. Beyaert, M. González, A. Quesada  
Departamento de Métodos Cuantitativos para la Economía y la Empresa

UNIVERSIDAD DE  
**MURCIA**

## 1. Introducción

## 2. Modelos VAR

2.1 FE, FR y FF del VAR

2.2 Estimación y contrastes

2.3 Análisis de la estacionariedad

2.4 Elección del número máximo de retardos

2.5 Chequeo de los residuos del modelo

2.6 “Causalidad” de Granger

2.7 Predicción

2.8 Especificación estructural



## 3. Modelos de Corrección del Error

### 3.1 Introducción

### 3.2 Cointegración

### 3.3 Estimación de MCE

### *Bibliografía básica:*

Lutkepohl, H. 2005. Structural Vector Autoregressive Analysis for Cointegrated Variables.

Documento de Trabajo del European Economic Institute No. 2005/2

(<http://cadmus.eui.eu/bitstream/handle/1814/2817/ECO2005-2.pdf;jsessionid=DAD1000534EA3EE699EA879301EFB981?sequence=1>)

Lutkepohl, H. 2007. Econometric Analysis with Vector Autoregressive Models. Documento de Trabajo del European Economic Institute No. 2007/11

(<http://cadmus.eui.eu/bitstream/handle/1814/6918/ECO-2007-11.pdf?sequence=1>)

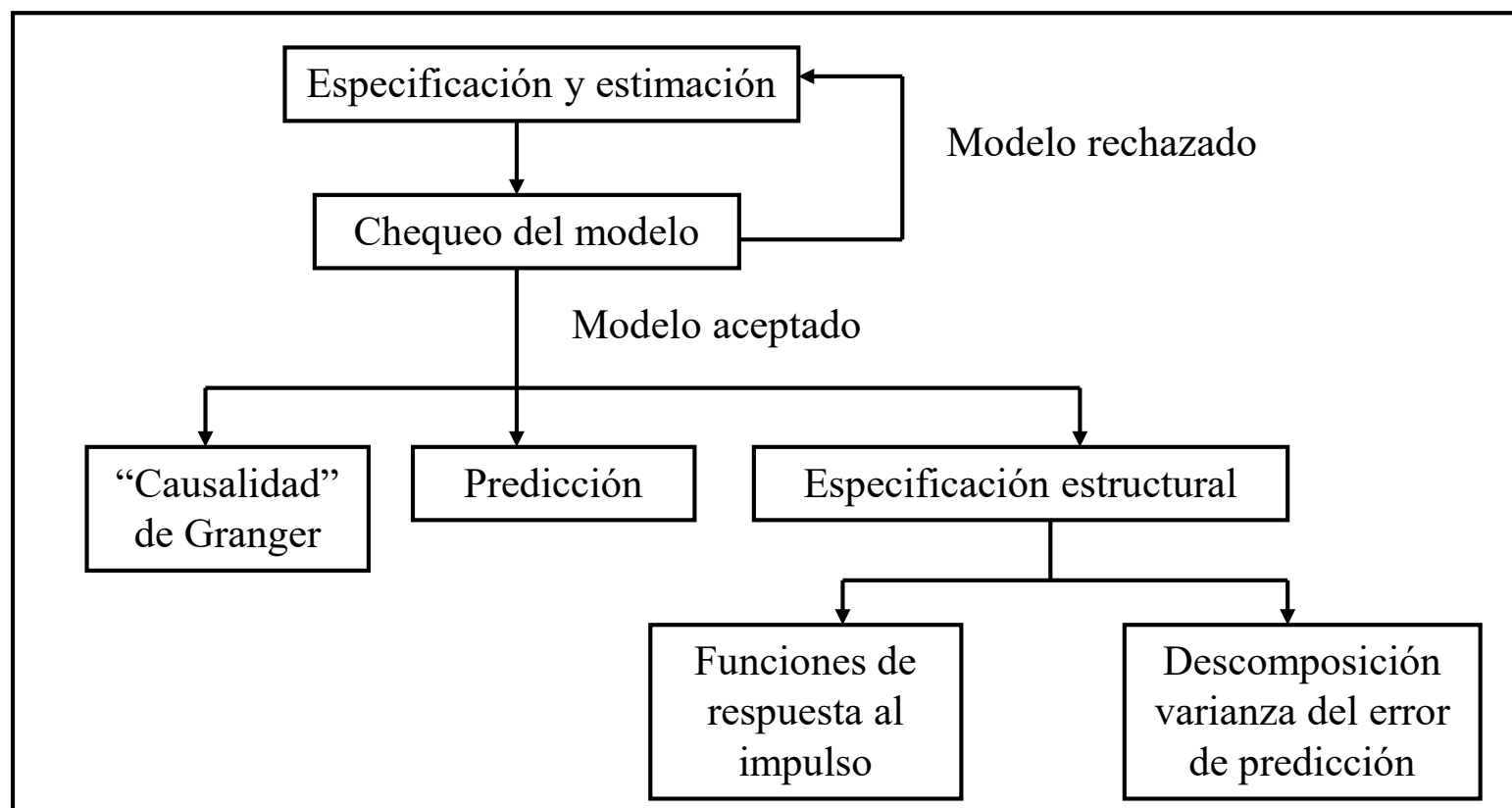
## ■ Críticas a los MES

- Complicado decidir qué variables son exógenas
- MES se proponen con muchas ecuaciones pero pocas observaciones
- Habían predicho muy mal los períodos de alto desempleo e inflación de 1970s
- Insuficientes retardos de las explicativas

## ■ Modelos VAR: caso particular de MES

- No contienen exógenas puras y se expresan generalmente en FR
- Usan pocas endógenas:  $M$  normalmente menor que 6
- Predecían mucho mejor que los MES más complejos
- Seleccionan cuidadosamente el número de retardos

- Resumen del análisis de los modelos VAR



## 2. Modelos VAR

### 2.1. FE, FR y FF del VAR

- Forma estructural: las endógenas del modelo función de otras endógenas del modelo y predeterminadas

$$A_0 Y_t = \tilde{\delta} + A_1 Y_{t-1} + \dots + A_p Y_{t-p} + \varepsilon_t \quad \varepsilon_t \sim \text{iid}(0, \Sigma)$$

- Forma reducida: las endógenas del modelo sólo función de predeterminadas

$$Y_t = A_0^{-1} \tilde{\delta} + A_0^{-1} A_1 Y_{t-1} + \dots + A_0^{-1} A_p Y_{t-p} + A_0^{-1} \varepsilon_t$$

$$Y_t = \delta + C_1 Y_{t-1} + \dots + C_p Y_{t-p} + U_t \quad E(U_t U_t') = \Omega = A_0^{-1} \Sigma (A_0^{-1})', \quad U_t = A_0^{-1} \varepsilon_t$$
$$C_i = A_0^{-1} A_i$$

- VMA (análogo a la Forma final del MES): las endógenas del modelo sólo función de ruidos y sus retardos. Partimos de la FR  $(I - C_1 L - \dots - C_p L^p) Y_t = \delta + U_t$

o bien  $C(L) Y_t = \delta + U_t$  y despejamos

$$Y_t = \mu + C(L)^{-1} U_t = \mu + U_t + \psi_1 U_{t-1} + \psi_2 U_{t-2} + \dots$$

## 2. Modelos VAR

### 2.1. FE, FR y FF del VAR

- Nos centramos en la FR (válida para predecir)  $Y_t = \delta + C_1 Y_{t-1} + \dots + C_p Y_{t-p} + U_t$

$$\begin{pmatrix} y_{1t} \\ \vdots \\ y_{Mt} \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_M \end{pmatrix} + \begin{pmatrix} c_{11}^1 & \dots & c_{1M}^1 \\ \vdots & \ddots & \vdots \\ c_{M1}^1 & \dots & c_{MM}^1 \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ \vdots \\ y_{Mt-1} \end{pmatrix} + \dots + \begin{pmatrix} c_{11}^p & \dots & c_{1M}^p \\ \vdots & \ddots & \vdots \\ c_{M1}^p & \dots & c_{MM}^p \end{pmatrix} \begin{pmatrix} y_{1t-p} \\ \vdots \\ y_{Mt-p} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ \vdots \\ u_{Mt} \end{pmatrix}$$

- FR para la variable endógena  $j$

➤ En  $t$ :

$$y_{jt} = \delta_j + c_{j1}^1 y_{1t-1} + \dots + c_{jM}^1 y_{Mt-1} + \dots + c_{j1}^p y_{1t-p} + \dots + c_{jM}^p y_{Mt-p} + u_{jt}$$

$$y_{jt} = X_t' \pi_j + u_{jt}, \quad u_{jt} \sim \text{iid}(0, \omega_j^2)$$

➤ Para todas las observaciones:

$$y_j = X \pi_j + u_j \quad u_j \sim (0, \omega_j^2 I_T)$$

## 2. Modelos VAR

### 2.1. FE, FR y FF del VAR

- FR para el modelo completo para todas las observaciones

$$\begin{bmatrix} y_1 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} X & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & X \end{bmatrix} \cdot \begin{bmatrix} \pi_1 \\ \vdots \\ \pi_M \end{bmatrix} + \begin{bmatrix} u_1 \\ \vdots \\ u_M \end{bmatrix}$$

o bien  $\tilde{y} = \tilde{X}\pi + \tilde{u}$

con  $\tilde{u} \sim (0, \tilde{\Omega})$  donde  $\tilde{\Omega} = \begin{bmatrix} \omega_1^2 I_T & \omega_{12} I_T & \cdots & \omega_{1M} I_T \\ \omega_{12} I_T & \omega_2^2 I_T & & \omega_{2M} I_T \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{1M} I_T & \omega_{2M} I_T & \cdots & \omega_M^2 I_T \end{bmatrix}$



## 2. Modelos VAR

### 2.1. FE, FR y FF del VAR

- Usaremos el ejemplo de Gali (1999)
- Modelo de tasa de crecimiento de la productividad ( $y_1=X$ ) y de las horas trabajadas ( $y_2=H$ ) para EE.UU. en la muestra 1947.1-2018.2
- Galí usa  $p=4$  pero nosotros en las especificaciones de **teoría** veremos  $p=1$

$$\begin{pmatrix} X_t \\ H_t \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} + \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} X_{t-1} \\ H_{t-1} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \quad \Omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{12} & \omega_{22} \end{pmatrix}$$

$$X_t = \delta_1 + c_{11}X_{t-1} + c_{12}H_{t-1} + u_{1t}$$

$$H_t = \delta_2 + c_{21}X_{t-1} + c_{22}H_{t-1} + u_{2t}$$

## 2. Modelos VAR

### 2.2. Estimación y contrastes

■ MCO ecuación por ecuación  $\hat{\pi}_{j,\text{MCO}} = (X'X)^{-1} X'y_j$  es

- Consistente: las explicativas son todas predeterminadas      llamemos  $\tilde{\Omega}^{-1} = \begin{pmatrix} w^{11}I & w^{12}I \\ w^{21}I & w^{22}I \end{pmatrix}$
- MCG no más eficiente: las explicativas son las mismas para todas las ecuaciones

$$\begin{aligned} \hat{\pi}_{\text{MCG}} &= (\tilde{X}'\tilde{\Omega}^{-1}\tilde{X})^{-1} \tilde{X}'\tilde{\Omega}^{-1}\tilde{y} = \left[ \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}' \begin{pmatrix} w^{11}I & w^{12}I \\ w^{21}I & w^{22}I \end{pmatrix} \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix} \right]^{-1} \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}' \begin{pmatrix} w^{11}I & w^{12}I \\ w^{21}I & w^{22}I \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \\ &= \left[ \begin{pmatrix} w^{11}I & w^{12}I \\ w^{21}I & w^{22}I \end{pmatrix} \begin{pmatrix} X'X & 0 \\ 0 & X'X \end{pmatrix} \right]^{-1} \begin{pmatrix} w^{11}I & w^{12}I \\ w^{21}I & w^{22}I \end{pmatrix} \begin{pmatrix} X'y_1 \\ X'y_2 \end{pmatrix} = \\ &= \begin{pmatrix} (X'X)^{-1} & 0 \\ 0 & (X'X)^{-1} \end{pmatrix} \begin{pmatrix} w^{11}I & w^{12}I \\ w^{21}I & w^{22}I \end{pmatrix}^{-1} \begin{pmatrix} w^{11}I & w^{12}I \\ w^{21}I & w^{22}I \end{pmatrix} \begin{pmatrix} X'y_1 \\ X'y_2 \end{pmatrix} = \begin{pmatrix} (X'X)^{-1} & 0 \\ 0 & (X'X)^{-1} \end{pmatrix} \begin{pmatrix} X'y_1 \\ X'y_2 \end{pmatrix} = \\ &= \begin{pmatrix} (X'X)^{-1} X'y_1 \\ (X'X)^{-1} X'y_2 \end{pmatrix} = \begin{pmatrix} \hat{\pi}_{1,\text{MCO}} \\ \hat{\pi}_{2,\text{MCO}} \end{pmatrix} \end{aligned}$$

## 2. Modelos VAR

### 2.2. Estimación y contrastes

- La estimación MCO es válida  $\Rightarrow$  aplicaremos la inferencia estudiada el curso anterior
- Contraste del ratio de la verosimilitud
  - Se usa mucho en el análisis de modelos VAR
  - Supongamos que queremos comparar dos modelos que estiman  $k_0$  y  $k_a$  parámetros respectivamente  $\Rightarrow$  Nos planteamos el contraste

$H_0$ : Modelo con  $k_0$  parámetros       $H_a$ : modelo con  $k_a$  parámetros

- Estimamos ambos modelos y calculamos el logaritmo de la función de verosimilitud (bajo normalidad es equivalente a usar las matrices de covarianzas)

$$2[\ln(L_a) - \ln(L_0)] = T[\ln|\hat{\Omega}_0| - \ln|\hat{\Omega}_a|] \sim \chi^2_{(k_a - k_0)}$$

## 2. Modelos VAR

### 2.3. Análisis de la estacionariedad

- Recuerda: un AR(1)  $y_t = c_0 + c_1 y_{t-1} + \varepsilon_t$  no tiene raíz unitaria cuando

$$|1 - c_1 L| = 0 \Rightarrow |L| = \left| \frac{1}{c_1} \right| > 1$$

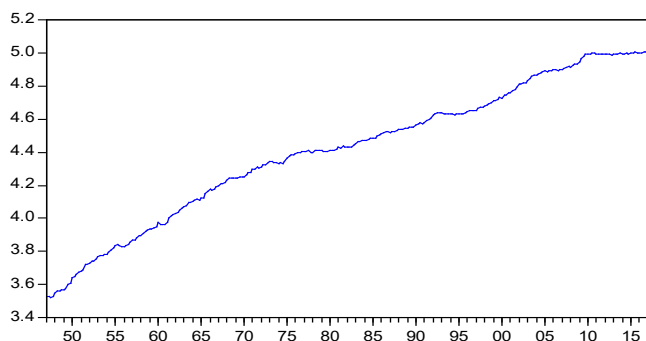
- Un AR (p)  $y_t = c_0 + c_1 y_{t-1} + \dots + c_p y_{t-p} + \varepsilon_t$  no tiene raíz unitaria si  $|L_i| > 1, \forall i$
- Un VAR  $Y_t = \delta + C_1 Y_{t-1} + \dots + C_p Y_{t-p} + U_t$  no tiene raíz unitaria si el determinante  $|I - C_1 L - \dots - C_p L^p| = 0$  tiene las raíces  $|L_i| > 1, \forall i$
- Si las series no son estacionarias por raíces unitarias
  - Generalmente se estacionarizan tramando las diferencias  $\Rightarrow$  modelo en diferencias
  - En caso de cointegración (al final del Tema) se usan modelos VECM

## 2. Modelos VAR

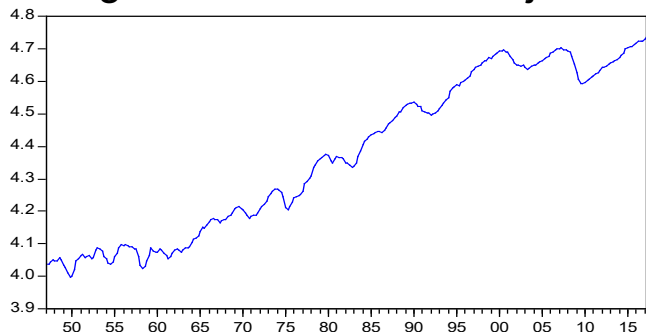
### 2.3. Análisis de la estacionariedad

- Es fácil de comprobar en Eviews: nos muestra la inversa de las raíces  $\Rightarrow$  todas deben ser  $< 1$  (“estar dentro del círculo unidad”)
- Ej: Modelo de Gali (1999): ¿Qué ocurre si no nos damos cuenta de que son  $I(1)$  usando ADF?

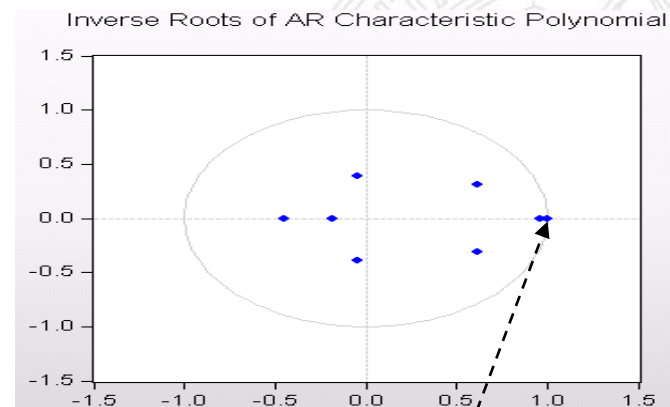
Logaritmo de productividad



Logaritmo de horas trabajadas



Si estimamos en niveles ( $p=4$ )



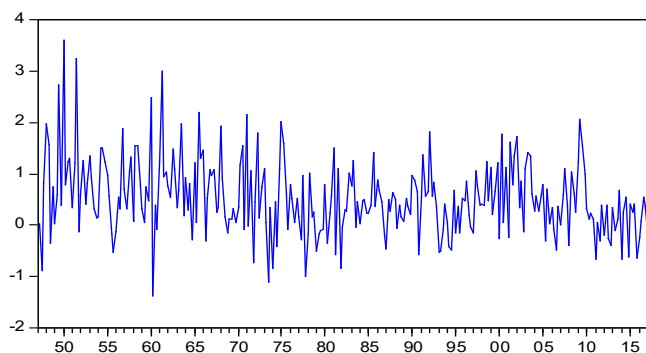
Hay una inversa de una raíz que está fuera del círculo unidad  $\Rightarrow$  no estacionario

## 2. Modelos VAR

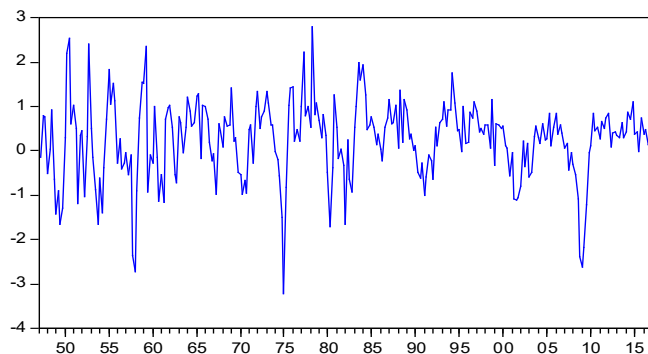
### 2.3. Análisis de la estacionariedad

- Ej: Modelo de Gali (1999): debemos usar ADF para comprobar que son  $I(1)$  y estacionarizarlas

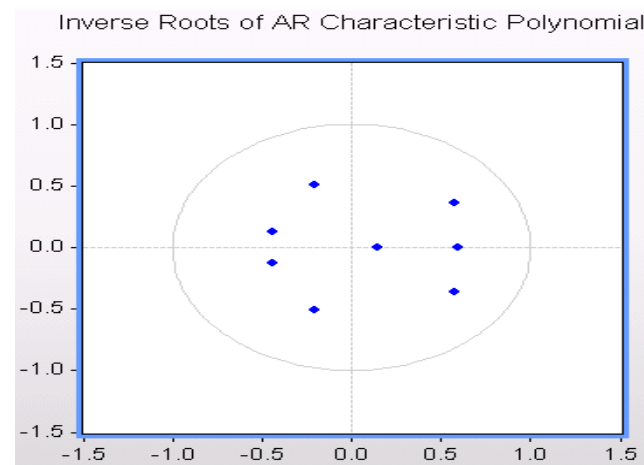
Crecimiento de productividad



Crecimiento de horas trabajadas



Si estimamos en diferencias ( $p=4$ )



Ahora todas las inversas de las raíces están dentro del círculo unidad  $\Rightarrow$  estacionario

## 2. Modelos VAR

### 2.4. Elección del número máximo de retardos

#### (1) Contrastes secuenciales

- Elegir  $p_{\max}$  grande y eliminar hasta el último significativo en los contrastes

$$H_0 : C_{p_{\max}} = 0 \quad H_0 : C_{p_{\max}-1} = 0 \quad \dots$$

- Problemas

- Si  $p_{\max}$  es demasiado alto podemos encontrar  $p$  demasiado grande porque siempre existe posibilidad de cometer errores de tipo 1 ( $RH_0$  cuando es cierta)
- Si  $p_{\max}$  es demasiado pequeño podemos no llegar al retardo óptimo porque el  $p$  verdadero fuese mayor que el  $p_{\max}$  elegido

## 2. Modelos VAR

### 2.4. Elección del número máximo de retardos

## (2) Criterios de selección Bayesianos

- Basados en elegir el modelo con **menor** valor del criterio

$$AIC(p) = \log|\Omega| + \frac{2}{T} pM^2$$

$$BIC(p) = \log|\Omega| + \frac{2\log(\log(T))}{T} pM^2$$

$$HQ(p) = \log|\Omega| + \frac{\log(T)}{T} pM^2$$

- Observaciones
  - Para un determinado p  $BIC(p) \geq HQ(p) \geq AIC(p)$
  - AIC tiende a sobreestimar p
  - HQ y BIC son consistentes  $\Rightarrow$  funcionan mejor en muestras grandes
  - En muestras pequeñas no está claro cuál funciona mejor



## 2. Modelos VAR

### 2.4. Elección del número máximo de retardos

### (3) Criterios de selección basados en capacidad predictiva

- FPE (Final Prediction Error)

$$\text{FPE}(p) = |\Omega| \left[ \frac{T + pM + 1}{T - pM - 1} \right]^M$$

- Se basa en la predicción de horizonte 1
- Elegimos el modelo con **menor** FPE
- En muestras grandes funciona peor que BIC y HQ
- En muestras pequeñas no está claro cuál funciona mejor



## 2. Modelos VAR

### 2.4. Elección del número máximo de retardos

■ El Eviews muestra un cuadro de comparación. En nuestro ejemplo

- Decidimos un máximo de 8 retardos (**SC=BIC**)

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-635.8757	NA	0.366898	4.673082	4.699525	4.683697
1	-573.0252	124.3196	0.238399	4.241943	4.321272	4.273787
2	-558.2186	29.07088	0.220254*	4.162773*	4.294988*	4.215847*
3	-556.3643	3.613448	0.223747	4.178493	4.363594	4.252796
4	-554.1548	4.273228	0.226707	4.191611	4.429598	4.287143
5	-549.2622	9.390985	0.225238	4.185071	4.475944	4.301833
6	-545.8836	6.435526	0.226279	4.189623	4.533382	4.327615
7	-545.6728	0.398354	0.232666	4.217383	4.614028	4.376604
8	-535.6294	18.83605*	0.222613	4.173109	4.622640	4.353560

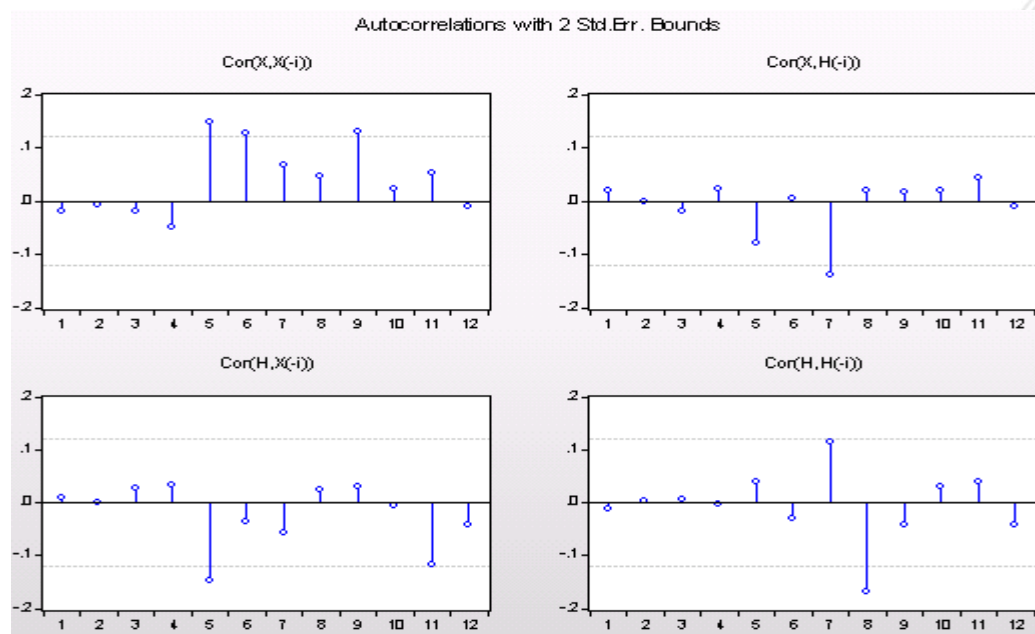
- Conclusión:
  - Según FPE, AIC, SC y HQ  $\Rightarrow p=2$
  - Según contrastes secuenciales  $\Rightarrow p=8$

## 2. Modelos VAR

### 2.5. Chequeo de los residuos del modelo

#### (1) Funciones de correlación

- Son un gráfico de  $\text{corr}(e_{h,t}, e_{j,t-i})$  para todo  $h, j=1, \dots, M$  con  $i=1, \dots, i_{\max}$
- Se presentan con las bandas de confianza  $\Rightarrow$  Si la correlación muestral está dentro de las bandas no podemos rechazar la nula de que la poblacional sea cero
- En nuestro ejemplo ( $i_{\max}=12$ ) están dentro de las bandas (no riesgo de correlación)



## 2. Modelos VAR

### 2.5. Chequeo de los residuos del modelo

#### (2) Contraste de Portmanteau $Y_t = \delta + C_1 Y_{t-1} + \dots + C_p Y_{t-p} + U_t$

- Sean las covarianzas (matrices  $M \times M$ ) de los ruidos  $V_j = E(U_t U_{t-j}')$
- Sean las estimaciones  $\hat{V}_j = \frac{1}{T} \sum_{t=j+1}^T \hat{U}_t \hat{U}_{t-j}'$
- El contraste se basa en que las autorrelaciones hasta orden  $h$  sean cero

$$H_0: V_1 = \dots = V_h = 0 \quad RH_0 \Leftrightarrow Q_h = T \sum_{j=1}^h \text{tr}(\hat{V}_j' \hat{V}_0^{-1} \hat{V}_j \hat{V}_0^{-1}) \underset{H_0}{\sim} \chi^2_{M^2(h-p)}$$
$$H_a: \text{no } H_0$$

- Para muestras pequeñas  $Q_h = T^2 \sum_{j=1}^h \frac{1}{T-j} \text{tr}(\hat{V}_j' \hat{V}_0^{-1} \hat{V}_j \hat{V}_0^{-1})$
- Sólo tiene sentido para valores de  $h > p$ : es una deficiencia grave porque los retardos  $h$  pequeños son los más importantes

## 2. Modelos VAR

### 2.5. Chequeo de los residuos del modelo

#### (2) Contraste de Portmanteau

- En nuestro ejemplo para un  $h$  de hasta 12 y  $p=4$

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	0.181661	NA*	0.182319	NA*	NA*
2	0.198888	NA*	0.199672	NA*	NA*
3	0.647143	NA*	0.652835	NA*	NA*
4	1.507743	NA*	1.526044	NA*	NA*
5	11.59370	0.0206	11.79740	0.0189	4
6	16.64514	0.0340	16.96069	0.0305	8
7	24.16179	0.0193	24.67221	0.0165	12
8	33.02089	0.0073	33.79478	0.0058	16
9	39.71788	0.0054	40.71667	0.0041	20
10	40.45663	0.0191	41.48309	0.0147	24
11	45.52865	0.0195	46.76485	0.0145	28
12	47.17693	0.0409	48.48777	0.0310	32

- Conclusión: Al 1% no parece que hay correlación serial (salvo  $h=8$  y  $h=9$ ). Al 5% habría problemas  $\Rightarrow$  comparar con el resultado de otros contrastes

## 2. Modelos VAR

### 2.5. Chequeo de los residuos del modelo

#### (3) Contraste de Breuch-Godfrey

- Sean  $\hat{U}_t$  los residuos y  $\hat{\Omega}$  la VARCOV estimada de  $Y_t = \delta + C_1 Y_{t-1} + \dots + C_p Y_{t-p} + U_t$
- Queremos contrastar  $H_0: U_t$  no autocorrelacionados hasta orden  $h$   
 $H_a: \text{no } H_0$
- Para un  $h$  realizamos la siguiente regresión auxiliar y obtenemos VARCOV estimada

$$\hat{U}_t = C_1 Y_{t-1} + \dots + C_p Y_{t-p} + B_1 \hat{U}_{t-1} + \dots + B_h \hat{U}_{t-h} + \varepsilon_t \Rightarrow \hat{\Sigma}_\varepsilon$$

- El contraste se puede plantear  $H_0: B_1 = B_2 = \dots = B_h = 0$   
 $H_a: \text{no } H_0$
- El estadístico y su distribución son  $Q_{LM} = T \left[ M - \text{tr} \left( \hat{\Omega}^{-1} \hat{\Sigma}_\varepsilon \right) \right]_{H_0} \sim \chi^2_{hM^2}$
- Tiene sentido para valores  $h \leq p$

## 2. Modelos VAR

### 2.5. Chequeo de los residuos del modelo

### (3) Contraste de Breuch-Godfrey

- En nuestro ejemplo para un h de hasta 12 y  $p=4$

Lags	LM-Stat	Prob
1	6.246841	0.1815
2	0.675979	0.9543
3	5.290329	0.2588
4	5.745759	0.2190
5	10.54728	0.0322
6	5.443480	0.2447
7	7.969604	0.0927
8	9.157498	0.0573
9	6.792705	0.1473
10	0.753506	0.9446
11	5.232831	0.2642
12	1.665734	0.7969

- Conclusión: Al 1% no hay correlación serial. Al 5% habría problemas sólo en  $h=5$

## 2. Modelos VAR

### 2.6. "Causalidad" de Granger

- Una variable  $y_i$  no causa a otra variable  $y_j$  en el sentido de Granger si los retardos de la primera no ayudan a predecir la segunda
- Si una variable causa en sentido de Granger a otra esto no quiere decir que la segunda sea el resultado de la primera
  - Ejemplo: observar que la gente abre los paraguas ayuda a predecir una tormenta, pero no la causa en ningún otro sentido
- Cómo contrastar la causalidad de Granger en un VAR

$$y_{jt} = \delta_j + \sum_{h=1}^p c_{jl}^h y_{lt-h} + \dots + \sum_{h=1}^p c_{ji}^h y_{it-h} + \dots + \sum_{h=1}^p c_{jM}^h y_{Mt-h} + u_{jt}$$

$$H_0 : y_i \not\rightarrow^G y_j \Leftrightarrow c_{ji}^h = 0 \quad \forall h = 1, \dots, p$$

$$H_a : y_i \xrightarrow{G} y_j \Leftrightarrow \text{algun } c_{ji}^h \neq 0$$

$\Rightarrow$  es un contraste de significatividad conjunta de  $c_{ji}^h \Rightarrow$  se hace de manera habitual



## 2. Modelos VAR

### 2.6. "Causalidad" de Granger

■ En nuestro ejemplo (p=4)

$$X_t = \delta_1 + c_{11}^1 X_{t-1} + c_{11}^2 X_{t-2} + c_{11}^3 X_{t-3} + c_{11}^4 X_{t-4} + c_{12}^1 H_{t-1} + c_{12}^2 H_{t-2} + c_{12}^3 H_{t-3} + c_{12}^4 H_{t-4} + u_{1t}$$

$$H_t = \delta_2 + c_{21}^1 X_{t-1} + c_{21}^2 X_{t-2} + c_{21}^3 X_{t-3} + c_{21}^4 X_{t-4} + c_{22}^1 H_{t-1} + c_{22}^2 H_{t-2} + c_{22}^3 H_{t-3} + c_{22}^4 H_{t-4} + u_{2t}$$

Dependent variable: X			
Excluded	Chi-sq	df	Prob.
H	28.57453	4	0.0000
All	28.57453	4	0.0000

Dependent variable: H			
Excluded	Chi-sq	df	Prob.
X	14.44715	4	0.0060
All	14.44715	4	0.0060

$$H_0 : H \not\rightarrow^G X$$

$$c_{12}^1 = c_{12}^2 = c_{12}^3 = c_{12}^4 = 0$$

$$H_a : H \rightarrow^G X$$

$$H_0 : X \not\rightarrow^G H$$

$$c_{21}^1 = c_{21}^2 = c_{21}^3 = c_{21}^4 = 0$$

$$H_a : X \rightarrow^G H$$

#### ■ Para predecir usamos siempre la FR

- Para simplificar supongamos que no hay constantes. Estimamos  $\hat{C}_1, \dots, \hat{C}_p$  en

$$Y_t = C_1 Y_{t-1} + \dots + C_p Y_{t-p} + U_t$$

- Queremos predecir las variables  $h$  períodos hasta  $T+h$

$$Y_{T+h} = C_1 Y_{T+h-1} + \dots + C_p Y_{T+h-p} + U_{T+h}$$

- Se puede demostrar que la predicción óptima (ECM mínimo) estimada es

$$\hat{Y}_{T+h/T} = \hat{E}(Y_{T+h} / Y_{T+h-1}, Y_{T+h-2}, \dots) = \hat{C}_1 Y_{T+h-1/T} + \dots + \hat{C}_p Y_{T+h-p/T}$$

- Se obtiene con predicción recursiva: usamos

$$Y_{T+h-j/T} = Y_{T+h-j} \text{ si } h \leq j \text{ y } Y_{T+h-j/T} = \hat{Y}_{T+h-j/T} \text{ si } h > j$$

#### ■ Podemos definir error de predicción $Y_{T+h} - Y_{T+h/T}$

## 2. Modelos VAR

### 2.7. Especificación estructural

#### 2.7.1. Introducción

- ¿Para qué queremos la especificación estructural?
  - Como hemos visto, si sólo queremos predecir, podemos usar FR
  - Pero necesitamos la FE si queremos dotar de intuición económica a los resultados
- Herramientas para el análisis económico
  - **Funciones de Respuesta al Impulso:** miden la reacción de las variables endógenas ante un shock (impulso) **inesperado** en alguna de las endógenas (en uno de los ruidos **estructurales**), manteniendo el resto de ruidos estructurales constantes
  - **Descomposición de la Varianza del Error de Predicción:** miden la proporción de la varianza del error de predicción que se explica por un impulso **inesperado** en alguna de las endógenas (en los ruidos **estructurales**), manteniendo el resto de ruidos estructurales constates

#### ■ ¿Cómo sería la respuesta a un impulso en un ruido de la VMA?

- La FR es  $(I - C_1L - \dots - C_pL^p)Y_t = C(L)Y_t = U_t$
- La VMA es  $Y_t = U_t + \psi_1 U_{t-1} + \psi_2 U_{t-2} + \dots$  donde  $\psi_i = f(C_1, \dots, C_p)$
- Supongamos que en  $t$  aumentamos 1 cada ruido ( $\tilde{1}$  es un vector de  $M$  unos)

➤ La respuesta de impacto es

$$\left. \begin{aligned} Y_t^* &= \psi_0 (U_t + \tilde{1}) + \psi_1 U_{t-1} + \psi_2 U_{t-2} + \dots \\ Y_t &= \psi_0 U_t + \psi_1 U_{t-1} + \psi_2 U_{t-2} + \dots \end{aligned} \right\} \Rightarrow E(Y_t^* - Y_t) = \psi_0$$

➤ Transcurrido 1 periodo, la respuesta es

$$\left. \begin{aligned} Y_{t+1}^* &= \psi_0 U_{t+1} + \psi_1 (U_t + \tilde{1}) + \psi_2 U_{t-1} + \dots \\ Y_{t+1} &= \psi_0 U_{t+1} + \psi_1 U_t + \psi_2 U_{t-1} + \dots \end{aligned} \right\} \Rightarrow E(Y_{t+1}^* - Y_{t+1}) = \psi_1$$

➤ Así, podemos calcular la respuesta tras  $h$  periodos:  $E(Y_{t+h}^* - Y_{t+h}) = \psi_h$

## 2. Modelos VAR

### 2.7. Especificación estructural

#### 2.7.2. Funciones de Respuesta a un Impulso

#### ■ ¿Cómo sería la respuesta a un impulso en un ruido de la VMA?

- Los  $\psi_j$  **se pueden calcular fácilmente** de dos maneras
  - Encontrando la relación que existe entre ellos y  $C_1, \dots, C_p$ .
  - Por simulación como cuando vimos las FRI en MES siguiendo los pasos

(1) Calculamos la evolución de la endógena  $i$  tras un shock inesperado en  $j$

$$Y_{i,t}^*(j), Y_{i,t+1}^*(j), \dots, Y_{i,t+h}^*(j), \dots$$

(2) Calculamos la evolución de la endógena  $i$  sin el shock

$$Y_{i,t}, Y_{i,t+1}, \dots, Y_{i,t+h}, \dots$$

(3) La diferencia es la respuesta de  $i$  a un shock inesperado en  $j$

$$\psi_h(i, j) = Y_{i,t+h}^*(j) - Y_{i,t+h}$$

(4) Repetimos (1) a (3) para  $i, j = 1, 2, \dots, M$

#### ■ Problema de identificación en modelos VAR:

(1) Los  $\psi_h$  NO son FRI: son respuestas a shocks de la FR ( $U_t$ ) y no de la FE ( $\varepsilon_t$ )

- Recordemos la relación entre shocks de las FE y FR

$$\text{FE: } A_0 Y_t = A_1 Y_{t-1} + \dots + A_p Y_{t-p} + \varepsilon_t \quad E(\varepsilon_t \varepsilon_t') = \Sigma$$

$$\text{FR: } Y_t = C_1 Y_{t-1} + \dots + C_p Y_{t-p} + U_t \quad U_t = A_0^{-1} \varepsilon_t \quad E(U_t U_t') = \Omega = A_0^{-1} \Sigma (A_0^{-1})'$$

$$\text{VMA: } Y_t = U_t + \psi_1 U_{t-1} + \psi_2 U_{t-2} + \dots$$

- Solución: usamos la relación  $U_t = A_0^{-1} \varepsilon_t$

$$Y_t = A_0^{-1} \varepsilon_t + \psi_1 A_0^{-1} \varepsilon_{t-1} + \psi_2 A_0^{-1} \varepsilon_{t-2} + \dots = R_0 \varepsilon_t + R_1 \varepsilon_{t-1} + R_2 \varepsilon_{t-2} + \dots$$

- Dado  $A_0$  podemos encontrar  $R_h$  a partir de  $R_h = \psi_h A_0^{-1}$

(2) Los  $R_h$  no representan FRI que buscamos como economistas: no podemos generar un shock en  $\varepsilon_{jt}$  manteniendo constantes los demás  $\varepsilon_{it}$  ya que  $\sigma_{ij} \neq 0$

#### ■ ¿Cómo resolvemos el problema de la identificación?

- Suponemos  $\sigma_{ij}=0 \quad \forall \quad i \neq j \Rightarrow$  Ahora  $R_h$  si representan FRI porque aíslan los shocks
- Buscamos  $A_0$  a partir de los parámetros de la FR  $\Rightarrow R_h = \psi_h A_0^{-1}$ 
  - ¿Podemos hacerlo?
    - FR hay  $M^2 p$  parámetros en  $C_j$  y  $\frac{1}{2}M(M+1)$  en  $\Omega$
    - Hay  $M^2$  parámetros en  $A_0$ ,  $M^2 p$  en  $A_i$ ,  $i=1, \dots, p$  y  $M$  parámetros en  $\Sigma$
    - Necesitamos imponer  $\frac{1}{2}M(M+1)$  restricciones adicionales
  - ¿Cómo imponemos las restricciones adicionales? Usando teoría económica
    - Restricciones de corto plazo
    - Restricciones de largo plazo

#### ■ Restricciones de corto plazo

- Se pueden imponer  $M$  restricciones  $\sigma_i^2=1 \ \forall i \Rightarrow$  las otras  $\frac{1}{2}M(M-1)$  restricciones se imponen en las **respuestas de impacto**  $R_0 = \psi_0 A_0^{-1} = A_0^{-1}$
- Las más conocidas son las restricciones de Cholesky
  - $A_0^{-1}$  se supone triangular inferior

$$A_0^{-1} = \begin{pmatrix} \alpha_{11} & 0 & 0 & \dots & 0 \\ \alpha_{21} & \alpha_{22} & 0 & \dots & 0 \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{M1} & \alpha_{M2} & \alpha_{M3} & \dots & \alpha_{MM} \end{pmatrix}$$

$\frac{1}{2}M(M-1)$  restricciones

$$\Sigma = I = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

$M$  restricciones

- $A_0$  también será triangular inferior  $\Rightarrow$  modelo recursivo
- Implican que las variables ordenadas más arriba en el VAR no responden de impacto a shocks no esperados de las variables más abajo



### ■ Restricciones de corto plazo

➤ ¿En qué consiste la descomposición de Cholesky?

- Una vez estimado  $\hat{\Omega}$  sabemos que es simétrica e invertible
- La descomposición de Cholesky asegura que siempre existe una matriz triangular inferior, que llamaremos  $A_0^{-1}$ , que cumple  $\hat{\Omega} = A_0^{-1} (A_0^{-1})'$

➤ En nuestro ejemplo

➤ Tenemos

$$\hat{\Omega} = A_0^{-1} (A_0^{-1})' \Rightarrow \begin{pmatrix} \hat{\omega}_{11} & \hat{\omega}_{12} \\ \hat{\omega}_{12} & \hat{\omega}_{22} \end{pmatrix} = \begin{pmatrix} \alpha_{11} & 0 \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} \alpha_{11} & \alpha_{21} \\ 0 & \alpha_{22} \end{pmatrix}$$

➤ Podemos resolver el sistema de 3 ecuaciones y 3 incógnitas

$$\alpha_{11}^2 = \hat{\omega}_{11}$$

$$\alpha_{21}\alpha_{11} = \hat{\omega}_{12}$$

$$\alpha_{21}^2 + \alpha_{22}^2 = \hat{\omega}_{22}$$

#### ■ Restricciones de largo plazo

- Se pueden imponer  $M$  restricciones  $\sigma_i^2=1 \forall i \Rightarrow$  las otras  $\frac{1}{2}M(M-1)$  restricciones se imponen en las **respuestas de largo plazo**
- Si las variables del VAR son estacionarias  $\Rightarrow R_\infty=0$
- Por eso las restricciones se imponen sobre  $\tilde{R}_\infty = \sum_{i=1}^{\infty} R_i$
- Demostremos que si las variables del VAR son las diferencias de variables no estacionarias  $\Rightarrow \tilde{R}_h = \sum_{i=1}^h R_i$  son las respuestas en los niveles
  - Sean las variables no estacionarias  $X_t$  y sus diferencias  $Y_t = \Delta X_t = X_t - X_{t-1}$
  - Las respuestas de impacto de los niveles a un shock en  $t$ 

$$\tilde{R}_0 = X_t^* - X_t = (X_t^* - X_{t-1}) - (X_t - X_{t-1}) = \Delta X_t^* - \Delta X_t = Y_t^* - Y_t = R_0$$
  - Las respuestas pasado un periodo
 
$$\tilde{R}_1 = X_{t+1}^* - X_{t+1} = (X_{t+1}^* - X_t^* + X_t^*) + (-X_{t+1} - X_t + X_t) = (X_{t+1}^* - X_t^*) - (X_{t+1} - X_t) + (X_t^* - X_t) = R_1 + R_0$$
  - Y pasados  $h$  periodos  $\tilde{R}_h = \sum_{i=1}^h R_i$

#### ■ Restricciones de largo plazo

- Por tanto, las  $\frac{1}{2}M(M-1)$  restricciones se imponen en las respuestas de largo plazo del nivel de las variables no estacionarias
- En nuestro ejemplo: la restricción permite identificar los shocks es  $\sum_{j=0}^{\infty} R_{12}^j = 0$

$$\tilde{R}_{\infty} = \begin{pmatrix} \tilde{R}_{11}^{\infty} & \tilde{R}_{12}^{\infty} \\ \tilde{R}_{21}^{\infty} & \tilde{R}_{22}^{\infty} \end{pmatrix} = \begin{pmatrix} \sum R_{11}^j & \sum R_{12}^j \\ \sum R_{21}^j & \sum R_{22}^j \end{pmatrix} = \begin{pmatrix} \sum R_{11}^j & 0 \\ \sum R_{21}^j & \sum R_{22}^j \end{pmatrix}$$

- $\varepsilon_1$  son shocks tecnológicos: son los únicos que afectan al nivel de la productividad en el largo plazo
- $\varepsilon_2$  son shocks de demanda: no pueden afectar al nivel de productividad en el largo plazo

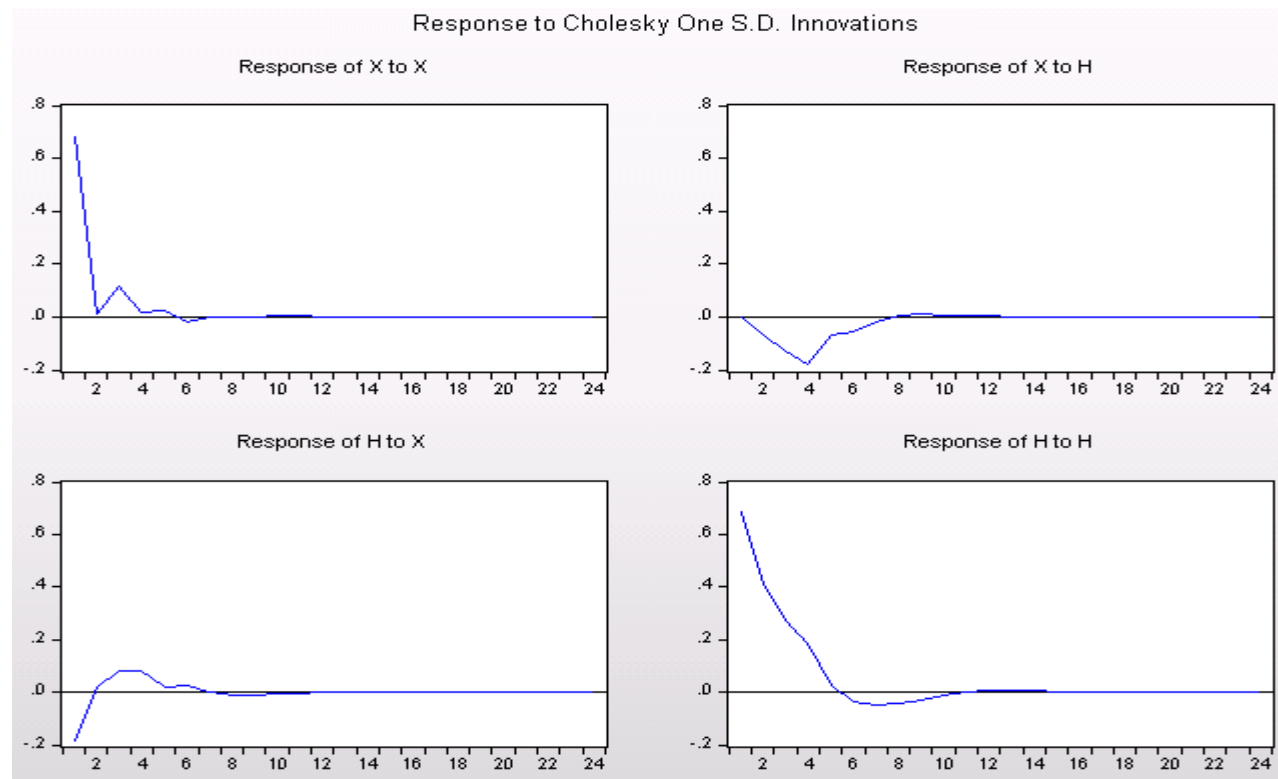
## 2. Modelos VAR

### 2.7. Especificación estructural

#### 2.7.2. Funciones de Respuesta a un Impulso

#### ■ En nuestro ejemplo

- Restricciones de **corto plazo**: respuestas de las **tasas de crecimiento**



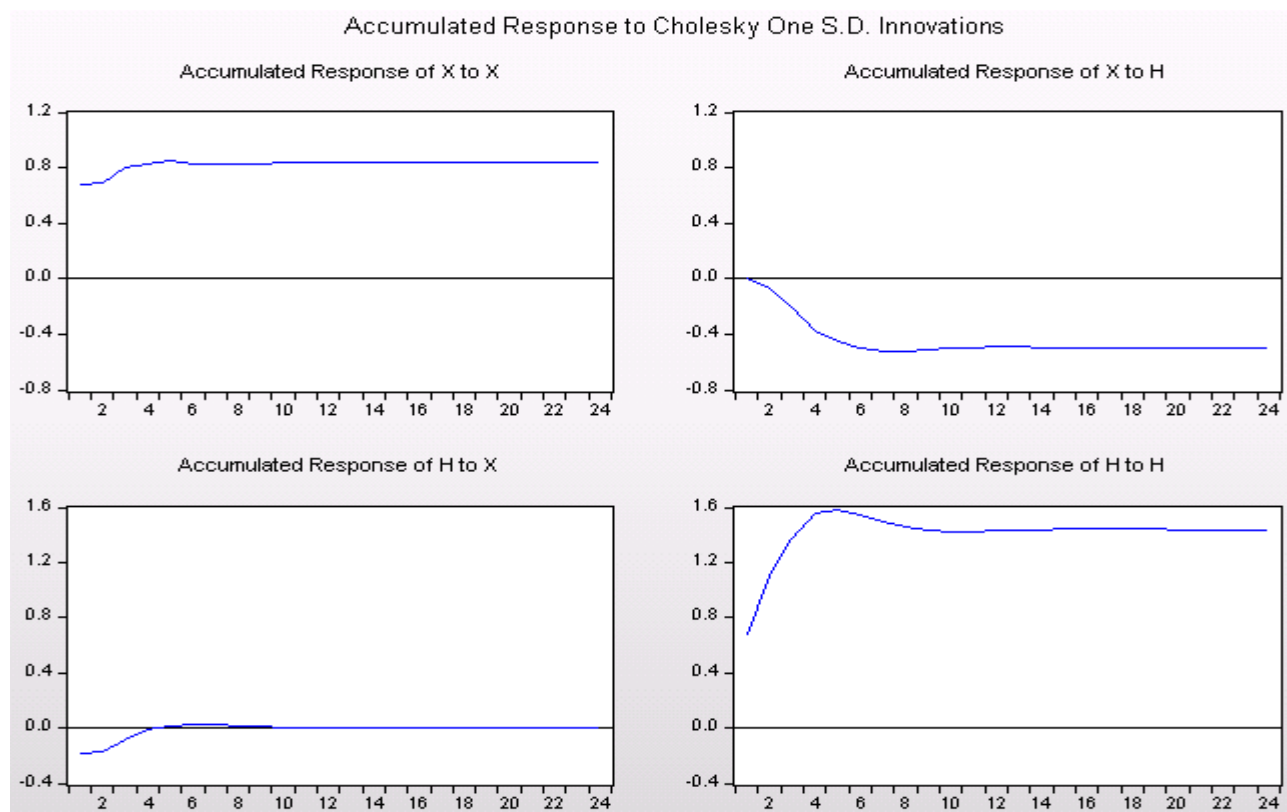
## 2. Modelos VAR

### 2.7. Especificación estructural

#### 2.7.2. Funciones de Respuesta a un Impulso

#### ■ En nuestro ejemplo

- Restricciones de **corto plazo**: respuestas de los **niveles**



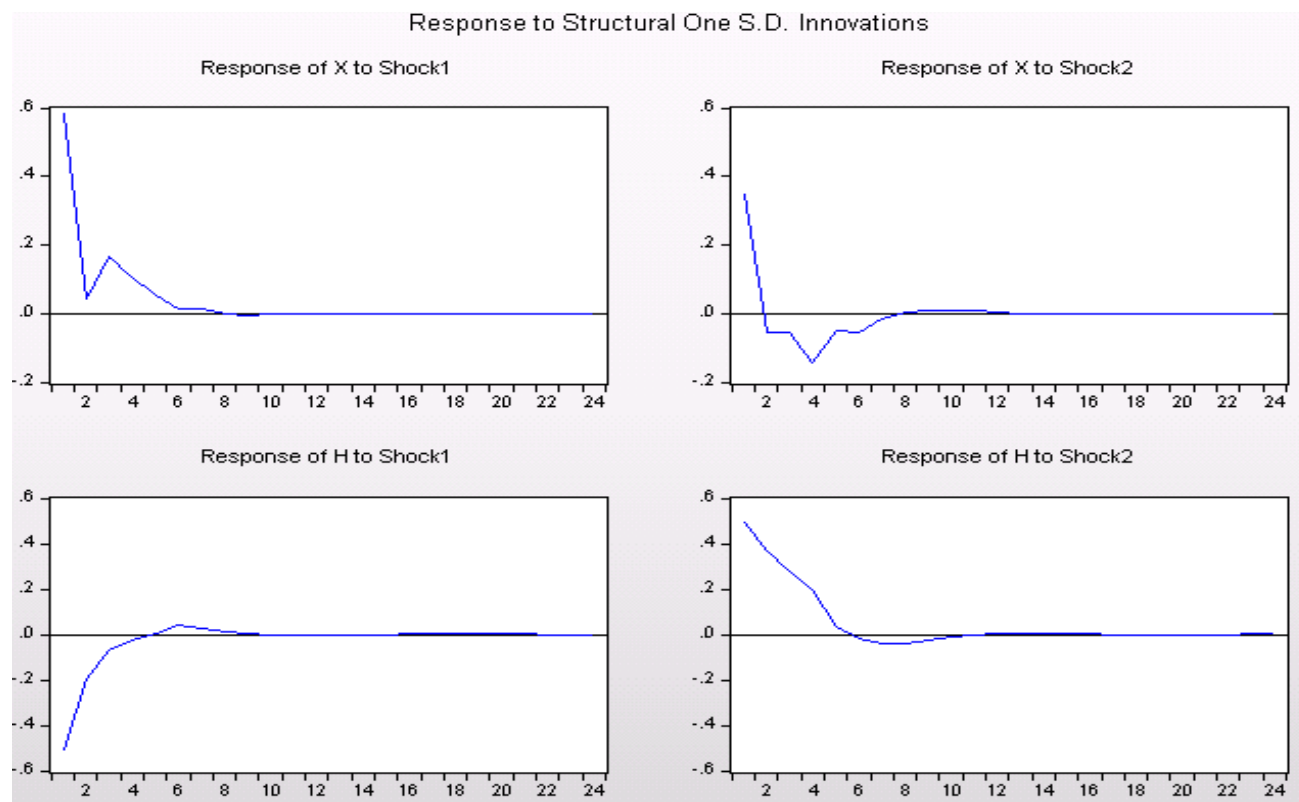
## 2. Modelos VAR

### 2.7. Especificación estructural

#### 2.7.2. Funciones de Respuesta a un Impulso

#### ■ En nuestro ejemplo

- Restricciones de **largo plazo**: respuestas de las **tasas de crecimiento**



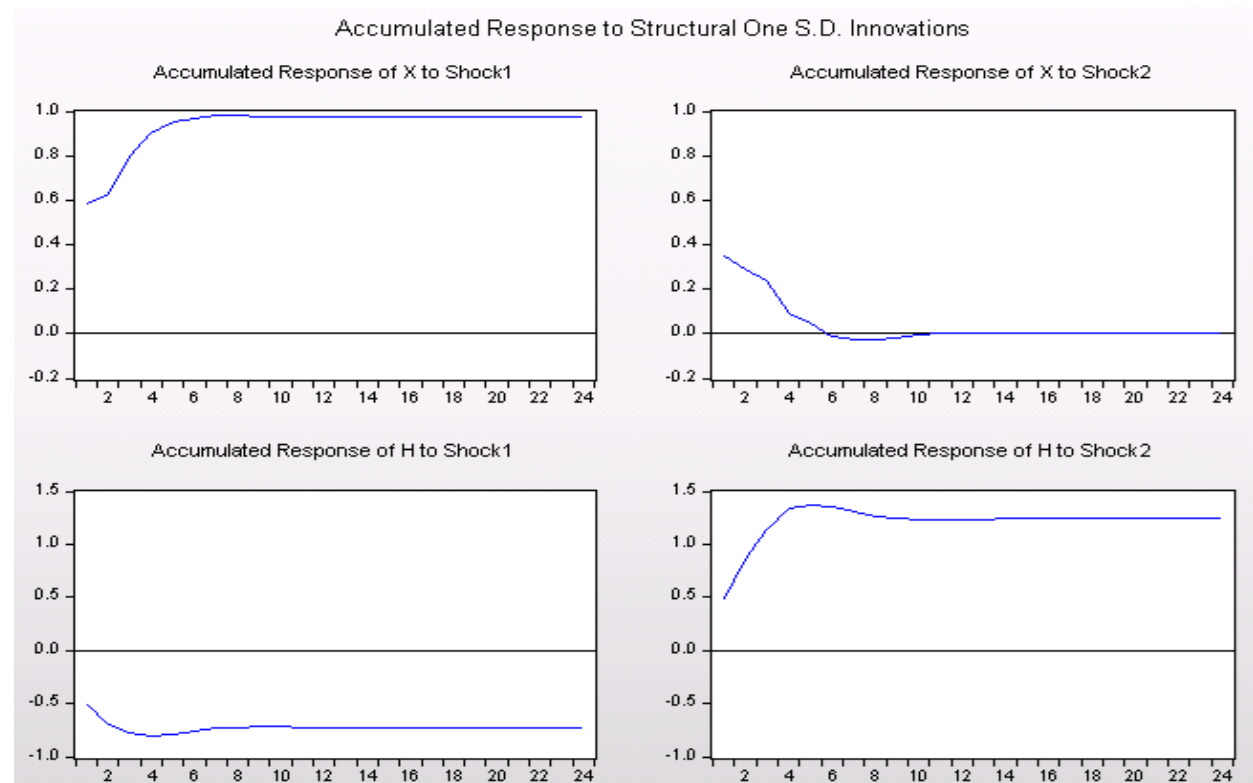
## 2. Modelos VAR

### 2.7. Especificación estructural

#### 2.7.2. Funciones de Respuesta a un Impulso

#### ■ En nuestro ejemplo

- Restricciones de **largo plazo**: respuestas de los **niveles**



La respuesta contracíclica del empleo a un shock en productividad llevó a Galí  $\Rightarrow$  a cuestionar los modelos de ciclos económicos reales

#### ■ Error de predicción y su varianza

- Recordemos el error de predicción de h periodos

$$Y_t = R_0 \varepsilon_t + R_1 \varepsilon_{t-1} + R_2 \varepsilon_{t-2} + \dots$$

$$Y_{t+h} = R_0 \varepsilon_{t+h} + R_1 \varepsilon_{t+h-1} + R_2 \varepsilon_{t+h-2} + \dots + R_{h-1} \varepsilon_{t+1} + R_h \varepsilon_t + R_{h+1} \varepsilon_{t-1} + R_{h+2} \varepsilon_{t-2} + \dots$$

$$Y_{t+h/t} = R_h \varepsilon_t + R_{h+1} \varepsilon_{t-1} + R_{h+2} \varepsilon_{t-2} + \dots$$

$$\xi_h = Y_{t+h} - Y_{t+h/t} = R_0 \varepsilon_{t+h} + R_1 \varepsilon_{t+h-1} + \dots + R_{h-1} \varepsilon_{t+1}$$

Error de predicción que se atribuye a  $\varepsilon$  futuros

- La varianza del error de predicción de h periodos (suponiendo  $\Sigma=I$ )

$$\text{var}(\xi_h) = E(\xi_h \xi_h') = R_0 R_0' + R_1 R_1' + \dots + R_{h-1} R_{h-1}'$$

$$\text{var}(\xi_{ih}) = \underbrace{\sum_{j=0}^{h-1} R_{i1}^j{}^2}_{\text{Contribución de } Y_1} + \dots + \underbrace{\sum_{j=0}^{h-1} R_{iM}^j{}^2}_{\text{Contribución de } Y_M}$$

**Contribución de  $Y_1$**    **Contribución de  $Y_M$**



## 2.8. Descomposición de la varianza del error de predicción

- En el ejemplo de Gali  $\xi_h = Y_{t+h} - Y_{t+h/t} = R_0 \varepsilon_{t+h} + R_1 \varepsilon_{t+h-1} + \dots + R_{h-1} \varepsilon_{t+1}$

$$\xi_h = \begin{pmatrix} X_{t+h} - X_{t+h/t} \\ H_{t+h} - H_{t+h/t} \end{pmatrix}$$

$$\text{var}(\xi_h) = \begin{pmatrix} R_{11}^0 & R_{12}^0 \\ R_{21}^0 & R_{22}^0 \end{pmatrix} \begin{pmatrix} R_{11}^0 & R_{12}^0 \\ R_{21}^0 & R_{22}^0 \end{pmatrix}' + \begin{pmatrix} R_{11}^1 & R_{12}^1 \\ R_{21}^1 & R_{22}^1 \end{pmatrix} \begin{pmatrix} R_{11}^1 & R_{12}^1 \\ R_{21}^1 & R_{22}^1 \end{pmatrix}' + \dots + \begin{pmatrix} R_{11}^{h-1} & R_{12}^{h-1} \\ R_{21}^{h-1} & R_{22}^{h-1} \end{pmatrix} \begin{pmatrix} R_{11}^{h-1} & R_{12}^{h-1} \\ R_{21}^{h-1} & R_{22}^{h-1} \end{pmatrix}'$$

$$\text{var}(X_{t+h} - X_{t+h/t}) = (R_{11}^{0^2} + R_{12}^{0^2}) + (R_{11}^{1^2} + R_{12}^{1^2}) + \dots + (R_{11}^{h-1^2} + R_{12}^{h-1^2}) = \underbrace{\sum_{j=0}^{h-1} R_{11}^{j^2}}_{\text{Contribución de X}} + \underbrace{\sum_{j=0}^{h-1} R_{12}^{j^2}}_{\text{Contribución de H}}$$

$$\text{var}(H_{t+h} - H_{t+h/t}) = (R_{22}^{0^2} + R_{21}^{0^2}) + (R_{22}^{1^2} + R_{21}^{1^2}) + \dots + (R_{22}^{h-1^2} + R_{21}^{h-1^2}) = \underbrace{\sum_{j=0}^{h-1} R_{22}^{j^2}}_{\text{Contribución de H}} + \underbrace{\sum_{j=0}^{h-1} R_{21}^{j^2}}_{\text{Contribución de X}}$$

## 2. Modelos VAR

### 2.8. Descomposición de la varianza del error de predicción

- En el ejemplo de Gali: la descomposición del error de predicción de  $x$

#### Identificación por Cholesky

Period	S.E.	X	H
1	0.679072	100.0000	0.000000
2	0.682473	99.02723	0.972770
3	0.704751	95.57878	4.421219
4	0.727246	89.82397	10.17603
5	0.730734	89.05099	10.94901
6	0.733126	88.52856	11.47144
7	0.733376	88.46904	11.53096
8	0.733381	88.46779	11.53221
9	0.733449	88.45171	11.54829
10	0.733503	88.44100	11.55900
11	0.733536	88.43519	11.56481
12	0.733543	88.43419	11.56581
13	0.733544	88.43422	11.56578
14	0.733545	88.43400	11.56600
15	0.733547	88.43369	11.56631
16	0.733548	88.43349	11.56651
17	0.733548	88.43342	11.56658
18	0.733548	88.43342	11.56658
19	0.733548	88.43342	11.56658
20	0.733548	88.43342	11.56658
21	0.733548	88.43341	11.56659
22	0.733548	88.43341	11.56659
23	0.733548	88.43341	11.56659
24	0.733548	88.43341	11.56659

Cholesky Ordering: X H

#### Identificación de largo plazo

Period	S.E.	Shock1	Shock2
1	0.679072	74.18085	25.81915
2	0.682473	73.83716	26.16284
3	0.704751	74.86221	25.13779
4	0.727246	72.46174	27.53826
5	0.730734	72.29208	27.70792
6	0.733126	71.85528	28.14472
7	0.733376	71.83098	28.16902
8	0.733381	71.83026	28.16974
9	0.733449	71.82014	28.17986
10	0.733503	71.80981	28.19019
11	0.733536	71.80341	28.19659
12	0.733543	71.80198	28.19802
13	0.733544	71.80202	28.19798
14	0.733545	71.80202	28.19798
15	0.733547	71.80186	28.19814
16	0.733548	71.80173	28.19827
17	0.733548	71.80168	28.19832
18	0.733548	71.80168	28.19832
19	0.733548	71.80168	28.19832
20	0.733548	71.80168	28.19832
21	0.733548	71.80167	28.19833
22	0.733548	71.80167	28.19833
23	0.733548	71.80167	28.19833
24	0.733548	71.80167	28.19833

Factorization: Structural

## 2. Modelos VAR

### 2.8. Descomposición de la varianza del error de predicción

- Otra utilidad: medición del **grado de conectividad**
- Llamemos  $d_{ij}^h$  a la proporción del error de la varianza de predicción de la variable  $i$  transcurridos  $h$  periodos desde que se produjo un shock en la variable  $j$ .
- Definamos la matriz de conectividad

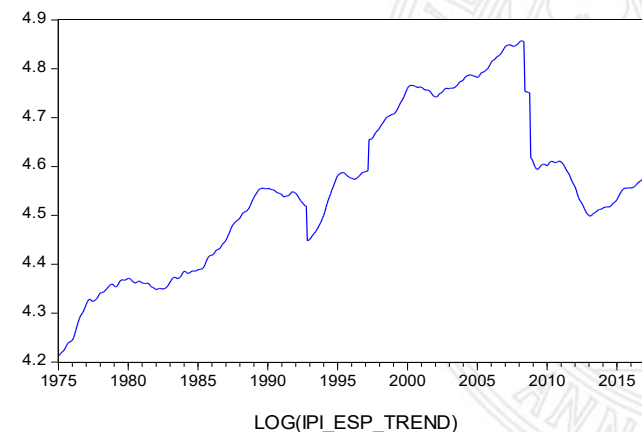
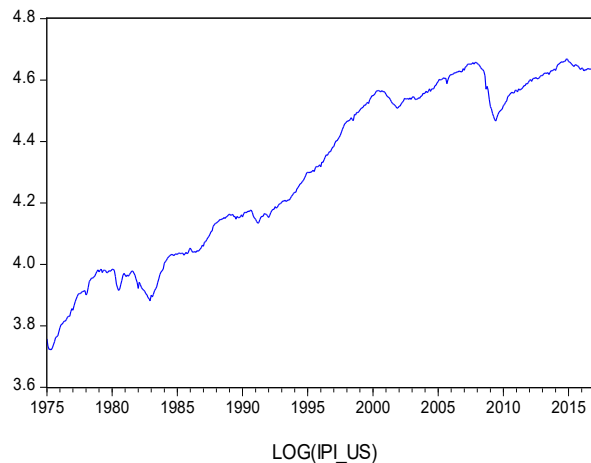
	$y_1$	$y_2$	...	$y_M$	De otros
$y_1$	$d_{11}^h$	$d_{12}^h$	...	$d_{1M}^h$	$\sum_{j=1}^M d_{1j}^h, j \neq 1$
$y_2$	$d_{21}^h$	$d_{22}^h$	...	$d_{2M}^h$	$\sum_{j=1}^M d_{2j}^h, j \neq 2$
.	.	.	.	.	.
.	.	.	.	.	.
$y_M$	$d_{M1}^h$	$d_{M2}^h$	...	$d_{MM}^h$	$\sum_{j=1}^M d_{Mj}^h, j \neq M$
Hacia otros	$\sum_{i=1}^M d_{i1}^h$ $i \neq 1$	$\sum_{i=1}^M d_{i2}^h$ $i \neq 2$		$\sum_{i=1}^M d_{iM}^h$ $i \neq M$	$\sum_{i,j=1}^M d_{ij}^h$ $i \neq j$

# 3. Modelos de Corrección del Error

## 3.1. Introducción

### ■ Regresión espuria

- Regresar variables no estacionarias en niveles  $\Rightarrow$  riesgo de regresión espuria
  - $R^2$  sistemáticamente alto
  - Estimaciones estadísticamente significativas  $\Rightarrow$  Granger y Newbold (1974)
  - Aunque no exista relación de causalidad entre las variables
  - Ejemplo: IPI de EE.UU. Y de España



$$\log(\hat{IPI}_{EE.UU_t}) = \underset{(0.19)}{-2.43} + \underset{(0.04)}{1.48} \log(IPI_{ESP_t}) \Rightarrow R^2 = 0.71$$

## 3. Modelos de Corrección del Error

### 3.1. Introducción

#### ■ ¿Por qué parece buena la regresión de los niveles?

- $R^2 = 1 - \frac{\hat{\sigma}_e^2}{\hat{\sigma}_Y^2}$  alto porque la varianza de Y es alta
- Cuando hay tendencia lineal el coeficiente es significativo porque la endógena tiene tendencia propia y la tendencia de la explicativa actúa de “variable proxy”
- Incluso cuando no hay tendencia lineal los estadísticos no tienen la distribución habitual  $\Rightarrow$  los contrastes de significatividad habituales no son válidos

#### ■ ¿Cómo hemos estimado la relación entre dos series no estacionarias?

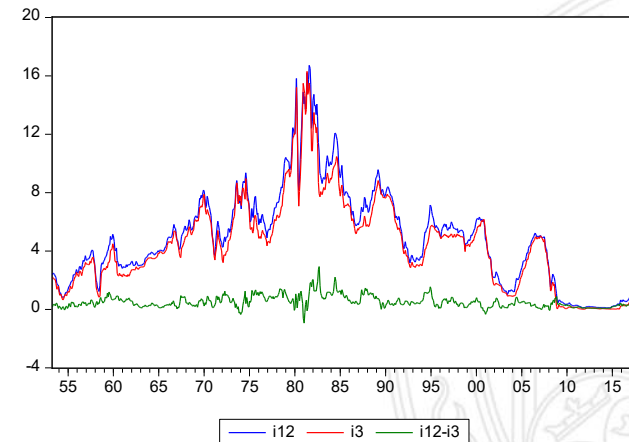
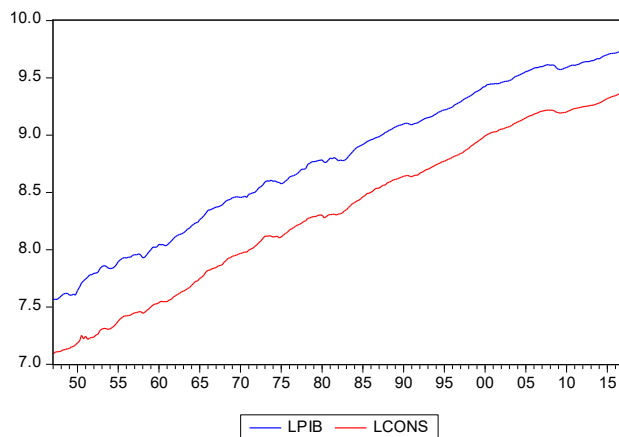
- Hasta ahora transformamos las series en estacionarias diferenciándolas
- Evitamos riesgos de regresión espuria
- Pero perdemos la información del nivel de las series: eliminamos la tendencia de las series y nos centramos en analizar cómo se relacionan “a corto plazo”

### 3. Modelos de Corrección del Error

#### 3.2. Cointegración

■ Pero existen variables que pueden tener relación no espuria a largo plazo por:

- Medidas de política económica: ejemplo de producción y consumo
- Fuerzas del mercado: estructura temporal de los tipos de interés



■ Decimos que esas variables cointegran

- No se separan mucho entre ellas en su dinámica en niveles  $\Rightarrow$  se dice que tienen una tendencia estocástica común
- Si lo hacen actúan las fuerzas que reducen los errores de equilibrio

### 3. Modelos de Corrección del Error

#### 3.2. Cointegración

#### ■ Definición de cointegración (adaptada de Engle-Granger)

- M variables  $I(d)$  están cointegradas si existe una combinación lineal no nula de todas ellas que sea  $I(d-1)$
- En economía es de especial interés series  $I(1)$  cuya combinación es  $I(0)$

#### ■ Contraste de cointegración de Engle y Granger

- Para simplificar, supongamos que tenemos dos variables  $y_{1t}$  y  $y_{2t}$  que son  $I(1)$
- Podría existir una (o varias) combinación lineal estacionaria  $\beta_1 y_{1t} + \beta_2 y_{2t} = \varepsilon_t^* \sim I(0)$  y si normalizamos  $y_{1t} + \tilde{\beta}_2 y_{2t} = \varepsilon_t \sim I(0)$
- ¿Como saber si existe tal relación lineal?
  - (1) A veces la **teoría económica** nos dice lo que valen  $\beta_1, \beta_2$  como en el caso de los tipos de interés  $\beta_1 = 1, \beta_2 = -1 \Rightarrow y_{1t} - y_{2t} = \varepsilon_t$
  - (2) Otras veces tendremos que **estimarla** regresando las series en niveles y obtenemos los residuos (**regresión de cointegración**)  $y_{1t} = a + b y_{2t} + \varepsilon_t \Rightarrow e_t$

### 3. Modelos de Corrección del Error

#### 3.2. Cointegración

- Ahora debemos comprobar si la combinación lineal es estacionaria
  - (1) Si los parámetros los da la teoría económica  $\Rightarrow$  **aplicamos ADF a  $\varepsilon_t$**
  - (2) A veces se estima porque queremos comprobar la teoría económica y otras veces porque no hay una teoría económica que nos indique cómo es la combinación lineal
    - **Aplicamos ADF a  $e_t$**  pero sin cte ni tendencia si hay cte en regresión de cointegración
    - Los valores críticos que aparecen en Eviews no son válidos porque ADF se realiza sobre una serie que proviene de una estimación (el residuo de la regresión de cointegración)
    - Los valores críticos debemos buscarlos en las tablas de Phillips y Ouliaris



### 3. Modelos de Corrección del Error

#### 3.2. Cointegración

- Tablas de Phillips y Ouliaris

Number of right-hand variables in regression, excluding trend or constant (n - 1)	Sample size (T)	Probability that $(\hat{\beta} - 1)/\hat{\sigma}_{\hat{\beta}}$ is less than entry						
		0.010	0.025	0.050	0.075	0.100	0.125	0.150
Case 1								
1	500	-3.39	-3.05	-2.76	-2.58	-2.45	-2.35	-2.26
2	500	-3.84	-3.55	-3.27	-3.11	-2.99	-2.88	-2.79
3	500	-4.30	-3.99	-3.74	-3.57	-3.44	-3.35	-3.26
4	500	-4.67	-4.38	-4.13	-3.95	-3.81	-3.71	-3.61
5	500	-4.99	-4.67	-4.40	-4.25	-4.14	-4.04	-3.94
Case 2								
1	500	-3.96	-3.64	-3.37	-3.20	-3.07	-2.96	-2.86
2	500	-4.31	-4.02	-3.77	-3.58	-3.45	-3.35	-3.26
3	500	-4.73	-4.37	-4.11	-3.96	-3.83	-3.73	-3.65
4	500	-5.07	-4.71	-4.45	-4.29	-4.16	-4.05	-3.96
5	500	-5.28	-4.98	-4.71	-4.56	-4.43	-4.33	-4.24
Case 3								
1	500	-3.98	-3.68	-3.42	—	-3.13	—	—
2	500	-4.36	-4.07	-3.80	-3.65	-3.52	-3.42	-3.33
3	500	-4.65	-4.39	-4.16	-3.98	-3.84	-3.74	-3.66
4	500	-5.04	-4.77	-4.49	-4.32	-4.20	-4.08	-4.00
5	500	-5.36	-5.02	-4.74	-4.58	-4.46	-4.36	-4.28

The probability shown at the head of the column is the area in the left-hand tail.

$$y_t = a + bx_t + \varepsilon_t$$

↓  
Constante en    Tendencia  
Cointegración    en series

No

No

Si

No

Si

Si

### 3. Modelos de Corrección del Error

#### 3.3. Estimación de MCE

#### ■ Modelos de corrección del error de Engle y Granger

- Supongamos que hay  $r$  relaciones de cointegración tales que

$$\begin{aligned} \beta_{11} Y_{1t} + \dots + \beta_{1M} Y_{Mt} &= \varepsilon_{1t} \sim I(0) \\ &\vdots \\ \beta_{r1} Y_{1t} + \dots + \beta_{rM} Y_{Mt} &= \varepsilon_{rt} \sim I(0) \end{aligned} \quad \beta' = \begin{pmatrix} \beta_{11} & \dots & \beta_{1M} \\ \vdots & & \vdots \\ \beta_{r1} & \dots & \beta_{rM} \end{pmatrix} \quad \varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \vdots \\ \varepsilon_{rt} \end{pmatrix}$$

- Engle y Granger demostraron que el MCE para variables que cointegran es

$$\Delta Y_t = \delta + \underbrace{\alpha \beta' Y_{t-1}}_{\text{Mecanismo de corrección del error}} + \underbrace{C(L) \Delta Y_{t-1}}_{\text{Modeliza el CP}} + U_t$$

NOTA:  $\beta$  por la teoría o estimado

## 3. Modelos de Corrección del Error

### 3.3. Estimación de MCE

#### ■ Ejemplo tipos a corto y a largo plazo

- Supongamos que nos fijamos de que  $\beta_1 = 1, \beta_2 = -1$  y queremos estimar la relación: realizamos la regresión de cointegración

$$y_t = a + bx_t + \varepsilon_t \Rightarrow \hat{y}_t = 0.26 + 1.06x_t \quad e_t = \hat{y}_t - (0.26 + 1.06x_t)$$

- Estimamos corto y largo plazo de las relaciones entre las series

$$\Delta y_t = a_0 + a(L)\Delta y_{t-1} + b(L)\Delta x_{t-1} + \alpha_1 e_{t-1} + u_{1t}$$

$$\Delta x_t = b_0 + c(L)\Delta y_{t-1} + d(L)\Delta x_{t-1} + \alpha_2 e_{t-1} + u_{2t}$$

- La dinámica de corto plazo se obtiene de  $a(L)$ ,  $b(L)$ ,  $c(L)$ ,  $d(L)$
- ¿Cómo afecta el ajuste a largo plazo en la dinámica del corto?
  - Errores de equilibrio en  $t-1$  se ajustan en  $t$
  - Se producirá un ajuste al equilibrio que actúa como atractor de las series

### 3. Modelos de Corrección del Error

#### 3.3. Estimación de MCE

#### ■ Ejemplo tipos a corto y a largo plazo $\hat{i}12_t = 0.26 + 1.06i3_t$

- Estimamos corto y largo plazo de la serie

$$\Delta \hat{i}12_t = 0.01 + 0.59\Delta i12_{t-1} - 0.24\Delta i3_{t-1} - 0.04(i12_{t-1} - 1.06i3_{t-1} - 0.26)$$

$$\Delta \hat{i}3_t = 0.01 - 0.05\Delta i3_{t-1} + 0.44\Delta i12_{t-1} + 0.12(i12_{t-1} - 1.06i3_{t-1} - 0.26)$$

- ¿Cómo afecta el ajuste a largo plazo en la dinámica del corto?
  - Imaginemos que en  $t$  se produce un error de equilibrio positivo:  $i12_t >> i3_t$
  - $i12$  caerá el periodo siguiente (coeficiente -0.04)
  - $i3$  subirá el periodo siguiente (coeficiente +0.12)
  - Se producirá un ajuste al equilibrio que actúa como atractor de las series
- Especificación y diagnóstico como en los VAR

## Lo que hemos aprendido:

- Los modelos VAR
  - Se usan principalmente para la predicción de variables económicas
  - Se suelen estimar en su FR (por MCO en cada ecuación)
  - No están tan directamente basados en la teoría económica como los MES
- Para estimar un VAR debemos tener en cuenta
  - Análisis de la estacionariedad: usamos variables estacionarias
  - Elección del número máximo de retardos
  - Chequear que los residuos del modelo son blancos
- Podemos usar el VAR para analizar la causalidad en el sentido de Granger, que hace referencia a la capacidad de los retardos de una variable para predecir otra.
- Las Funciones de Respuesta al Impulso y la Descomposición de la Varianza del Error de Predicción son las herramientas fundamentales para analizar económicamente las relaciones entre las variables del VAR
- Cuando un conjunto de variables no estacionarias están cointegradas
  - Existen relaciones lineales entre ellas que son estacionarias y podemos encontrar “atractores de sus dinámicas”
  - Los modelos más adecuados para estas variables son los MCE que capturan tanto la dinámica del corto plazo como la del largo plazo