

# CURSO INTERNACIONAL: CONSTRUCCIÓN DE ESCENARIOS ECONÓMICOS Y ECONOMETRÍA AVANZADA

Instructor: Horacio Catalán



# TEORÍA DE COINTEGRACIÓN

## **Co-Integration and Error Correction: Representation, Estimation, and Testing**

Robert F. Engle; C. W. J. Granger

*Econometrica*, Vol. 55, No. 2. (Mar., 1987), pp. 251-276.



# Efectos de las propiedades estocásticas de las series en un modelo de regresión

Sea el caso de dos variables que se definen como camino aleatorio y se especifican en un modelo de regresión

$$(1) \quad y_t = y_{t-1} + u_t \quad u_t \rightarrow IID(0, \sigma_u^2)$$

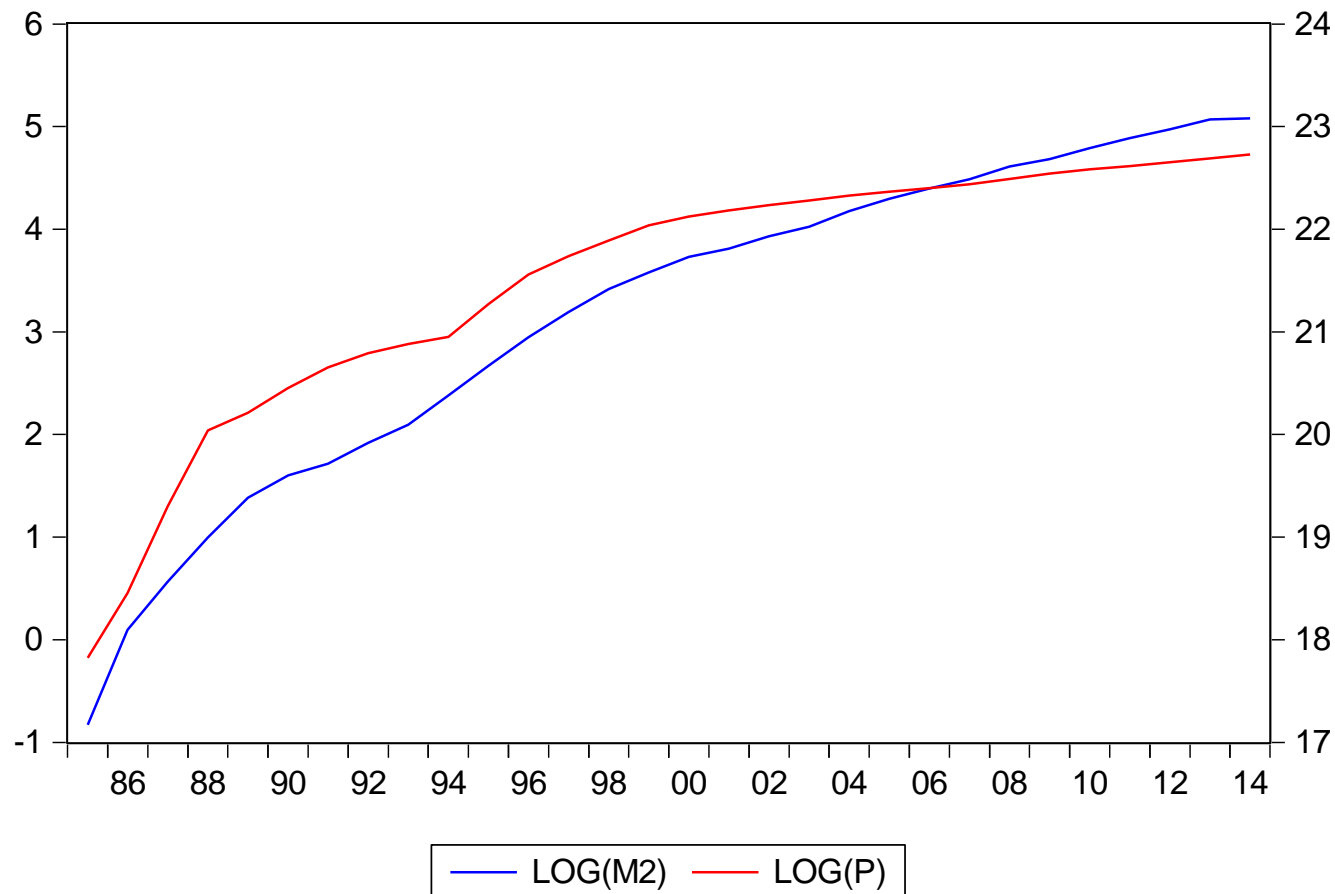
$$(2) \quad x_t = x_{t-1} + v_t \quad v_t \rightarrow IID(0, \sigma_v^2)$$

$$E(u_t v_t) = 0 \quad \forall t, s; \quad E(u_t u_{t-k}) = E(v_t v_{t-k}) = 0 \quad \forall k \neq 0$$

$$(3) \quad y_t = \beta_0 + \beta_1 x_t + z_t$$

# Ejemplo el nivel de precios en función del agregado monetario

$$\ln P_t = \beta_0 + \beta_1 \ln M2_t + u_t$$



Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-13.10433	0.602416	-21.75295	0.0000
LOG(M2)	0.784393	0.028395	27.62472	0.0000
R-squared	0.964607	Mean dependent var	3.490005	
Adjusted R-squared	0.963343	S.D. dependent var	1.297665	
S.E. of regression	0.248450	Akaike info criterion	0.117190	
Sum squared resid	1.728368	Schwarz criterion	0.210604	
Log likelihood	0.242143	Hannan-Quinn criter.	0.147074	
F-statistic	763.1254	Durbin-Watson stat	0.289218	
Prob(F-statistic)	0.000000			

## Problemas de la regresión espuria

- 1) Los estimadores son estadísticamente significativos, presentando estadísticos  $t$  y  $F$  elevados, que rechazan la hipótesis nula.
- 2) El valor de la  $R^2$  es muy cercano a 1, indicando que el modelo tiene un buen ajuste
- 3) El estadístico  $DW$  tiende a cero

Una regla para determinar si la regresión es falsa

$$DW < R^2$$

Cuando dos variables presentan camino aleatorio indica que la varianza de ambas series aumenta con el tiempo:

$$Y_t = Y_{t-1} + u_t \quad \text{Var}(Y_t) = T\sigma^2_Y$$

$$X_t = X_{t-1} + e_t \quad \text{Var}(Y_t) = T\sigma^2_X$$

La serie  $Y_t$  se aleja de su media por lo tanto se generan valores de  $R^2$  cercanos a uno, señalando que el ajuste del modelo es muy bueno. Sin embargo se debe a que las series se mueven juntas

Son series con memoria larga en consecuencia la función de autocorrelación no es convergente y se presenta un problema de Autocorrelación en los errores

## La estimación en primera diferencia muestra un cambio importante en los estadísticos

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.026380	0.037627	-0.701087	0.4892
DLOG(M2)	0.959897	0.140192	6.847000	0.0000
R-squared	0.634549	Mean dependent var	0.169146	
Adjusted R-squared	0.621014	S.D. dependent var	0.214327	
S.E. of regression	0.131944	Akaike info criterion	-1.146412	
Sum squared resid	0.470046	Schwarz criterion	-1.052116	
Log likelihood	18.62297	Hannan-Quinn criter.	-1.116880	
F-statistic	46.88140	Durbin-Watson stat	1.913021	
Prob(F-statistic)	0.000000			



## Econometrics—Alchemy or Science?

By DAVID F. HENDRY



### **Econometric Modelling of the Aggregate Time-Series Relationship Between Consumers' Expenditure and Income in the United Kingdom**

James E. H. Davidson; David F. Hendry; Frank Srba; Stephen Yeo

*The Economic Journal*, Vol. 88, No. 352 (Dec., 1978), 661-692.

## Ecuación del consumo

$$\begin{aligned} \Delta_4 \ln \hat{C}_t = & 0.47 \Delta_4 \ln Y_t - 0.21 \Delta_1 \Delta_4 \ln Y_t - 0.10 \ln (C/Y)_{t-4} \\ & (0.04) \quad (0.05) \quad (0.02) \\ & + 0.01 \Delta_4 D_t^0 - 0.13 \Delta_4 \ln P_t - 0.28 \Delta_1 \Delta_4 \ln P_t, \end{aligned} \quad (45)$$

(0.003)      (0.07)      (0.15)

$$R^2 = 0.77, \quad \hat{\sigma} = 0.0061, \quad dw = 1.8, \quad z_1(20) = 21.8, \quad z_2(12) = 19.$$

- Un modelo con series en primera diferencia y sus rezagos genera mejores resultados que modelos en niveles
- Las propiedades de estacionaridad de las variables son relevantes en la especificación del modelo econométrico

- Existe el problema de la regresión espuria. Un modelo que parece tener buenos resultados pero en realidad la relación entre las variables es falsa
- Los modelos con las variables en primeras diferencias generan mejores resultados
- El modelo de corrección de errores era superior a otras especificaciones teóricas
- Es importante las propiedades de estacionaridad de las series (pruebas de raíz unitaria)

# Prueba Dickey-Fuller

Dickey, D.A. y W.A. Fuller (1979), “Distribution of the Estimators for Autoregressive Time Series With a Unit Root”, *Journal of the American Statistical Association*, vol. 74, pp. 427-431

Dickey, D.A. y W.A. Fuller (1981), “Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root”, *Econometrica*, vol. 49, pp. 1057-1022



David Dickey



Wayne  
Fuller

De la ecuación AR(1)

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

Restando en ambos lados de la ecuación  $y_{t-1}$

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

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

La especificación de la prueba puede definirse como:

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

$$\alpha = (\phi - 1)$$

## Definición de la hipótesis nula

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

$$H_0 : \alpha = 0 \Rightarrow (\phi - 1) = 0 \Rightarrow \phi = 1$$

Si el parámetro alfa es igual a cero implica que la serie sigue un camino aleatorio, en consecuencia es no estacionaria

$$H_1 : \alpha \neq 0 \Rightarrow (\phi - 1) \neq 0 \Rightarrow \phi \neq 1$$

$$\alpha < 0 \Rightarrow (\phi - 1) < 0 \Rightarrow \phi < 1$$

Si el parámetro alfa es distinto de cero y es negativo, entonces la serie es estacionaria

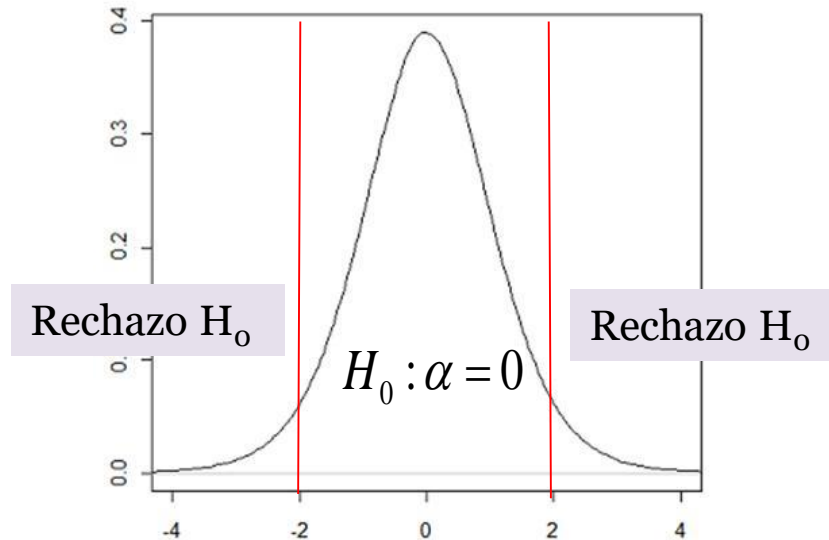
En el contexto de la prueba Dickey-Fuller hay dos condiciones para que la serie de tiempo sea estacionaria:

- Rechazar la hipótesis nula
- Que el estimador alfa sea negativo

Esto se puede probar mediante un estadístico t de Student

$$\hat{t}_{\hat{\alpha}} = \frac{\hat{\alpha}}{\text{SE}(\hat{\alpha})}$$

# Distribución t-Student



## Condiciones de estacionaridad

$$\hat{t}_{\hat{\alpha}} = \frac{\hat{\alpha}}{SE(\hat{\alpha})} < 0$$

Prob de rechazo  $< 0.05$



# Consideraciones en la prueba Dickey-Fuller

- 1) Distribución del estadístico bajo la hipótesis nula, tiende a valores negativos
- 2) Los resultados de la estimación del parámetro  $\alpha$  son afectados por la presencia de autocorrelación en los errores de la prueba
- 3) Los componentes de constante y tendencia deben ser incorporados en la especificación de la prueba

# Dickey-Fuller Aumentada (ADF)

Modelo A

$$\Delta y_t = \delta_0 + \delta_1 T + \alpha y_{t-1} + \sum_{i=1}^k \gamma_i \Delta y_{t-i} + \varepsilon_t$$

Modelo B

$$\Delta y_t = \delta_0 + \alpha y_{t-1} + \sum_{i=1}^k \gamma_i \Delta y_{t-i} + \varepsilon_t$$

Modelo C

$$\Delta y_t = \alpha y_{t-1} + \sum_{i=1}^k \gamma_i \Delta y_{t-i} + \varepsilon_t$$

# Cointegración



El análisis de cointegración es esencial cuando se tiene una combinación de variables que presenten una similitud en el orden de integración. Si se tiene una ecuación con las siguientes condiciones:

Sean las variables  $X_t \sim I(1)$      $Y_t \sim I(1)$

$$Y_t = \beta_0 + \beta_1 X_t + u_t$$

Una combinación lineal de estas variables que sea estacionaria. Entonces, se dice que las variables  $Y$ ,  $X$  están cointegradas

$$Y_t - \beta_0 - \beta_1 X_t = u_t$$

Puede ser  $I(0)$

El análisis de cointegración es en esencia un análisis multivariado

Un conjunto de variables se dice que son cointegradas si existe una combinación lineal que genere un proceso estocástico estacionario

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \cdots + \beta_k x_{kt} + u_t$$

$$y_t - \beta_1 x_{1t} - \beta_2 x_{2t} - \cdots - \beta_k x_{kt} = u_t$$

Vector normalizado

$$\begin{bmatrix} 1 - \beta_1 - \beta_2 \cdots - \beta_k \end{bmatrix} \begin{bmatrix} y_t \\ x_{1t} \\ x_{2t} \\ \vdots \\ x_{kt} \end{bmatrix} = \mathbf{u}_t$$

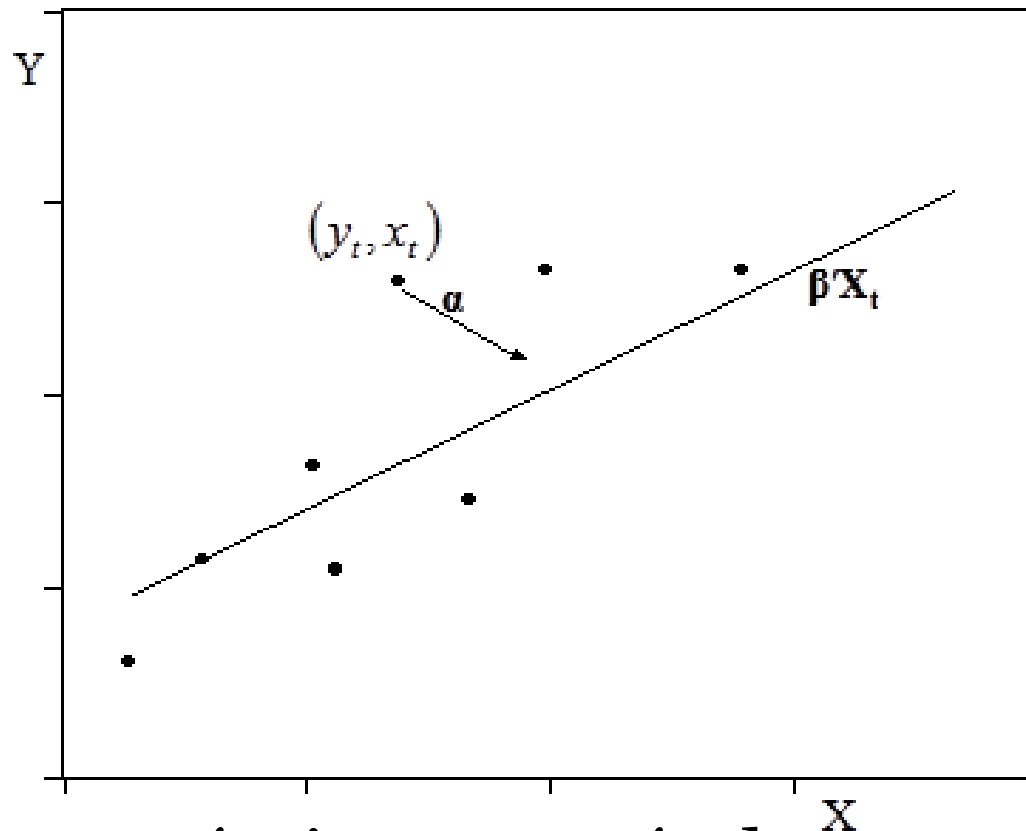
Así el análisis de cointegración se basa en estimar los valores del vector  $\beta$  que generan un proceso estocástico estacionario

$$\beta'X_t = u_t \longrightarrow I(0)$$

Se dice que las variables del modelo están cointegrados

Si las series cointegran la regresión entre las dos variables es significativa ( no es espuria) y no se pierde información valiosa de largo plazo lo cual sucedería si se estima la regresión en primeras diferencias.

El vector de cointegración es un atractor de la combinación de los puntos formados por el par de series  $(y_t, x_t)$



Los agentes económicos a través de sus acciones generan que las variables económicas se muevan alrededor de la relación estructural  $\beta'X_t$

## Prueba de cointegración: Residuales OLS

1) Estimar la ecuación de largo plazo por MCO

$$Y_t = b_0 + b_1 X_t + u_t$$

2) Guardar la serie de los residuales

$$\hat{u}_t = Y_t - \hat{b}_0 - \hat{b}_1 X_t$$

3) Aplicar una prueba Dickey-Fuller Simple

$$\Delta u_t = \alpha u_{t-1} + e_t$$

$$\hat{t}_\alpha = \frac{\hat{\alpha}}{\text{SE}(\hat{\alpha})} < 0$$



4) Comprobar que la serie de los errores sea estacionario

$$\hat{t}_{\alpha} = \frac{\hat{\alpha}}{SE(\hat{\alpha})} < 0 \quad \text{y la prob} < 0.05$$

Si se cumple esta condición se dice que las variables cointegran, se mantienen juntas en el tiempo

$$LGAS = -9.17 - 0.18 * LPRG + 1.22 * LY$$

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.275588	0.0244
Test critical values:		
1% level	-2.647120	
5% level	-1.952910	
10% level	-1.610011	

Observaciones importantes sobre la definición de cointegración.

1) La cointegración se refiere a una combinación lineal de variables no estacionarias.

- Pueden ser posibles relaciones no lineales.
- El vector de cointegración no es único.
- Se realiza una normalización del vector de cointegración.

2) Todas las variables deben ser del mismo orden de integración

❑ Aún si todas las variables son del mismo orden de integración no se asegura que cointegren.

❑ No existe claridad en el uso del término “relación de equilibrio”.

3) Si  $X_t$  tiene  $n$  componentes, debe haber  $n-1$  vectores de cointegración. El número de vectores se denomina rango de cointegración

# MODELO DE CORRECCIÓN DE ERRORES

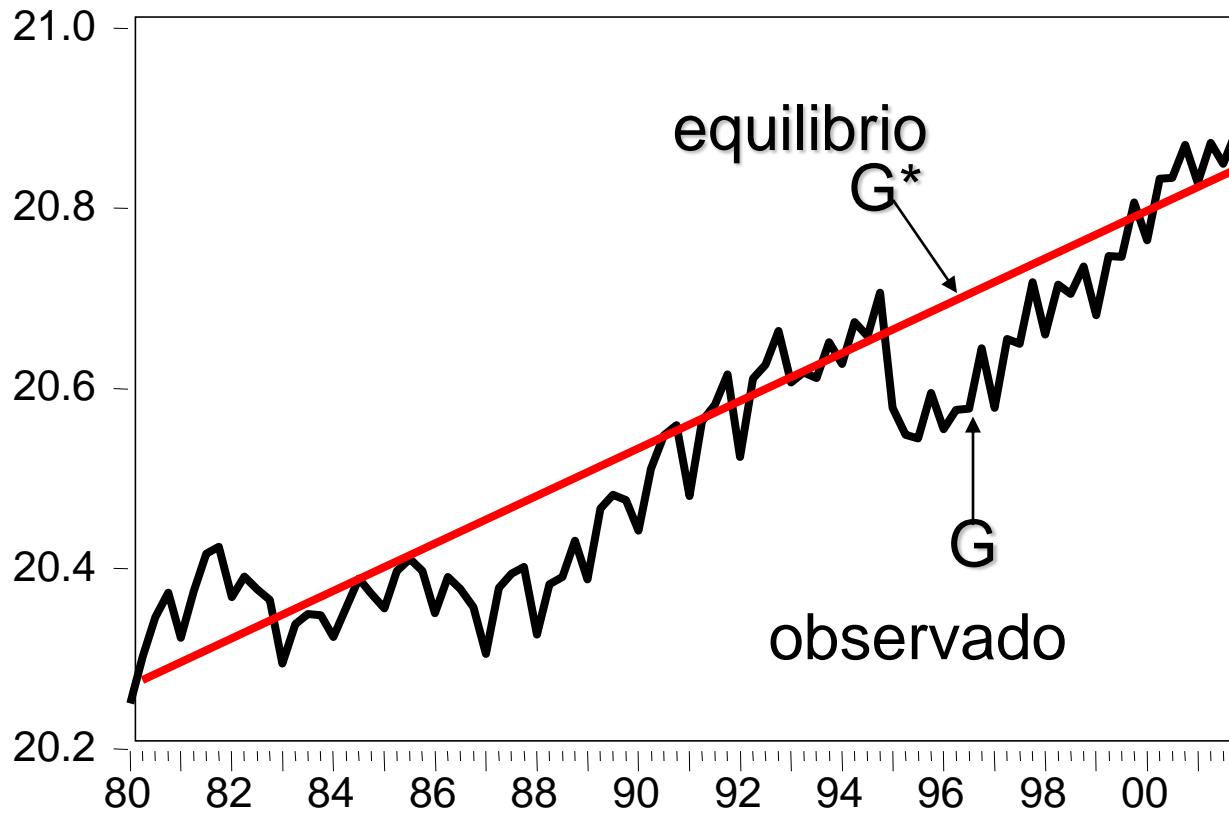
Relación de equilibrio

$$y_t = k_0 + k_1 x_t + u_t$$

Modelo de corrección de errores

$$\Delta y_t = \alpha \Delta x_t + \gamma [y_{t-1} - k_0 - k_1 x_{t-1}] + v_t$$

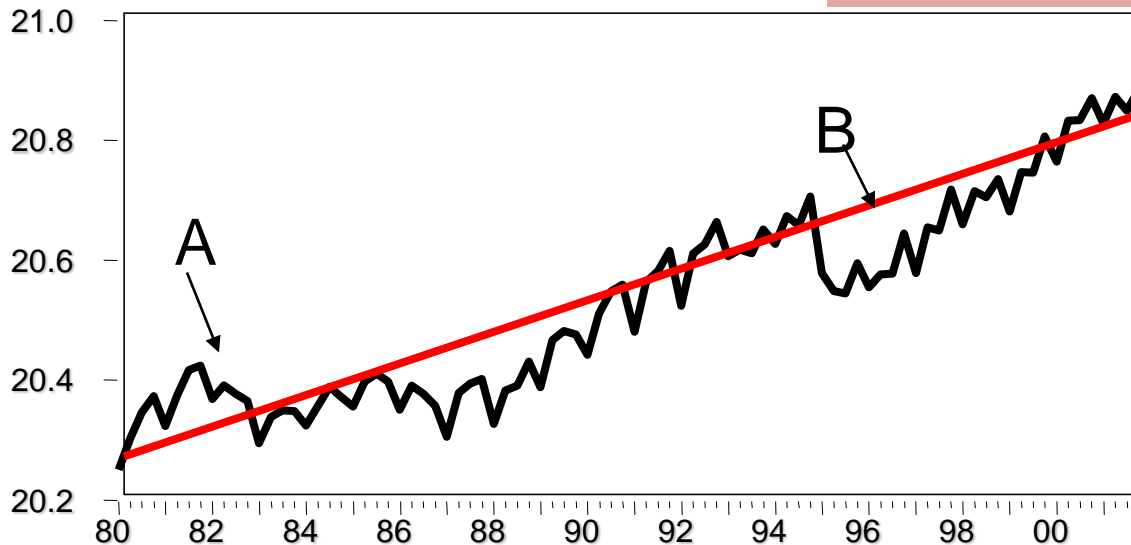
$\gamma$  es el coeficiente del mecanismo de corrección de errores toma valores entre  $-1$  y  $0$



Relación  
De  
Equilibrio

Cuando  $u > 0$  implica que  $G > G^*$

Cuando  $u < 0$  implica que  $G < G^*$



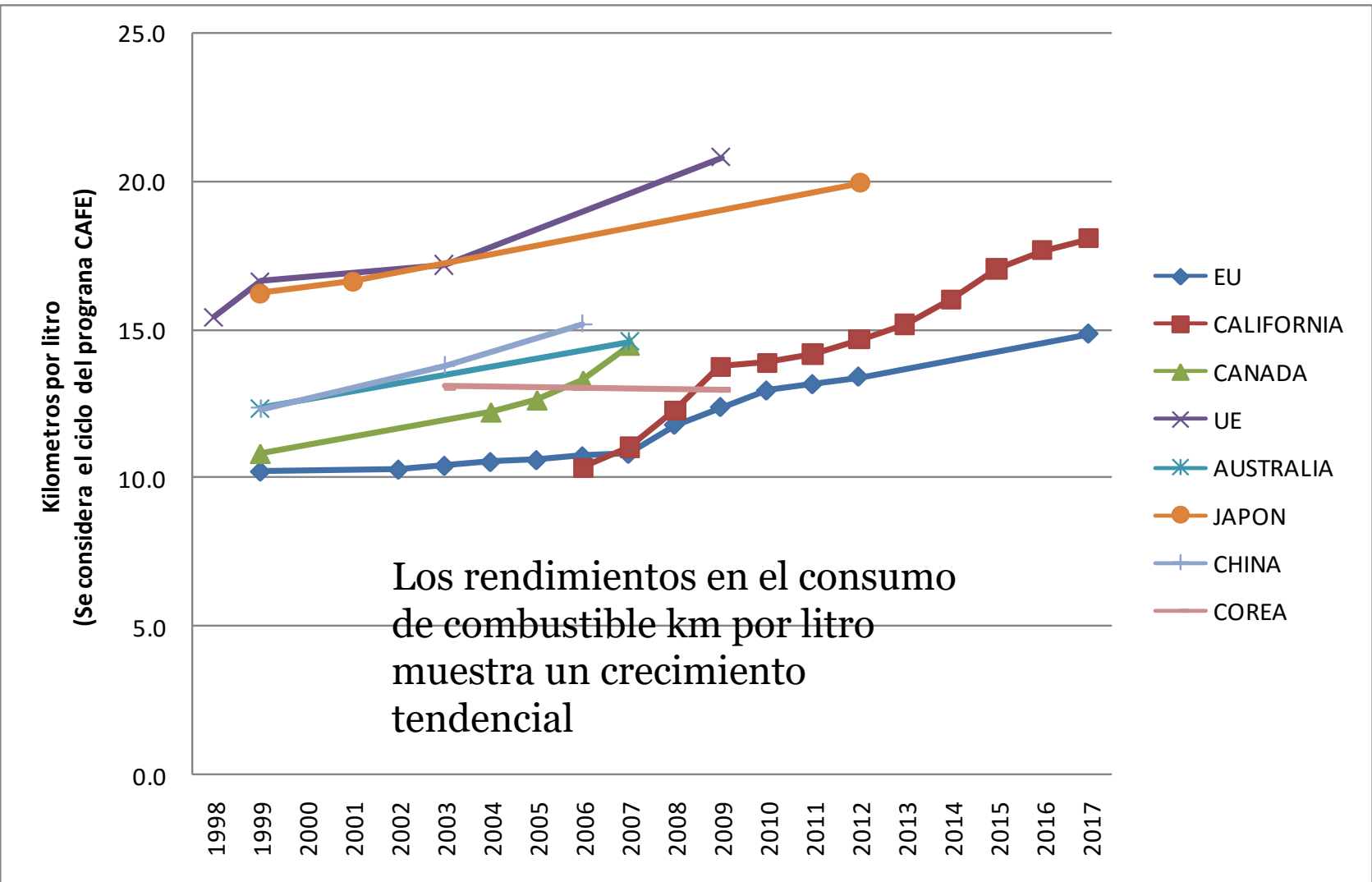
A)  $G > G^* \quad \longrightarrow \quad ECM = (G - G^*) > 0 \quad \text{Si} \quad \gamma < 0$

$$\Delta G_t = \beta_2 \Delta Y_t + \gamma [ECM_{t-1}] + U_t \quad \text{Efecto negativo}$$

B)  $G < G^* \quad \longrightarrow \quad ECM = (G - G^*) < 0 \quad \text{Si} \quad \gamma < 0$

$$\Delta G_t = \beta_2 \Delta Y_t + \gamma [ECM_{t-1}] + U_t \quad \text{Efecto positivo}$$

# Tendencia en la evolución de los rendimientos a nivel internacional



Con base en información de International Council on Clean Transportation

# Estudios de meta-análisis

Autor	Variable	Elasticidad ingreso		Elasticidad precio	
		CP	LP	CP	LP
Espey (1996)	CG		0.60 (n = 21)		-0.53 (n = 70)
Espey (1998)	CG	0.47 (n = 345)	0.88 (n = 345)	-0.26 (n = 277)	-0.58 (n = 363)
Hanly, Dargay y Goodwin (2002)	CG	0.39 (n = 45)	1.08 (n = 50)	-0.25 (n = 64)	-0.64 (n = 51)
	V-KM	0.30 (n = 7)	0.73 (n = 7)	-0.10 (n = 3)	-0.29 (n = 3)
	V-KM/V	-0.005 (n = 3)	0.17 (n = 4)	-0.10 (n = 3)	-0.30 (n = 3)
	SV	0.32 (n = 15)	0.81 (n = 15)	-0.08 (n = 8)	-0.25 (n = 8)
	CG/V	0.07 (n = 1)	0.93 (n = 1)	-0.08 (n = 1)	-1.10 (n = 1)
Graham y Glaister (2002b)	CG	0.47 (n = 333)	0.93 (n = 150)	-0.25 (n = 377)	-0.77 (n = 213)
	V-KM			-0.15 (n = 31)	-0.31 (n = 72)
	EG			0.10 (n = 31)	0.46 (n = 72)
	SV	0.28 (n = 5)	-0.74 (n = 5)		
Brons, Nijkamp, Rietveld (2008)	CG			-0.36 (n = 94) y -0.34 (n = 222)	-0.81 (n = 64) y -0.84 (n = 90)
	EG			0.09 (n = 11) y 0.14 (n = 222)	0.20 (n = 4) y 0.31 (n = 90)
	V-KM/V			-0.55 (n = 3) y -0.12 (n = 222)	-0.29 (n = 90)
	SV			-0.10 (n = 9) y -0.08 (n = 222)	-0.77 (n = 5) y -0.24 (n = 90)
	CG/V			-0.25 (n = 97) y -0.26 (n = 222)	-0.78 (n = 15) y -0.60 (n = 90)
	V-KM			0.03 (n = 8) y -0.20 (n = 222)	-0.32 (n = 2) y -0.53 (n = 90)



# Evidencia sobre elasticidad eficiencia

Autor	Variable dependiente	Método	Países y periodo de estimación	Elasticidad ingreso	Elasticidad precio	Elasticidad eficiencia
Bentzen (1994)	Gasolina	Cointegración y ECM	Dinamarca (1948-1991)	1.04	-0.41	-0.014
Johnston y Dinardo (1997)	Gasolinas	Cointegración y ECM	Estados Unidos (1959-1990)	0.99	-0.13	-0.51
Dargay (1997)	Gasolinas	ML	Países de la OCDE y Asia (1992)	1.06 (0.69 a 1.43)	-0.2* y -0.5**	-0.2* y -0.5**
Medlock III (2009)	Gasolinas	OLS	Estados Unidos (1980-2005)	0.16	-0.02	-0.45
Broadstock y Hunt (2010)	Gasolinas	STSM	Reino Unido (1960-2007)	0.53	-0.12	-0.32
	Gasolinas	STSM	Reino Unido (1964-2003)	0.57	-0.12	-0.27

Notas: Gasolinas: Consumo total de gasolina y diesel; ECM: Modelo de Corrección de Error; ML: Máxima Verosimilitud; STSM: Modelos Estructurales de Series de Tiempo. OLS: Mínimos Cuadrados Ordinarios; \*escenario bajo; y \*\* escenario alto.

## Diferentes relaciones de largo plazo

$$gas_t = \beta_0 + \beta_1 prg_t + \beta_2 Y_t + \beta_3 \left( \frac{km}{lt} \right) + u_t$$

$$gas_t = \beta_0 + \beta_1 prg_t + \beta_2 Y_t + \beta_3 \left( \frac{autos}{pob} \right) + u_t$$

$$gas_t = \beta_0 + \beta_1 prg_t + \beta_2 Y_t + \beta_3 \left( \frac{km}{lt} \right) * Venta + u_t$$

# CURSO INTERNACIONAL: CONSTRUCCIÓN DE ESCENARIOS ECONÓMICOS Y ECONOMETRÍA AVANZADA

Instructor: Horacio Catalán

