

Cambios estructurales y procesos no lineales

Gabriel V. Montes-Rojas

Motivación

- Muchas de las series temporales en macroeconomía y en finanzas son no lineales:
 - asimetría en la ocurrencia y persistencia de los shocks negativos y los positivos (business cycle);
 - irreversibilidad de los efectos temporales;
 - diferencias en los efectos en las colas de la distribución (asimetría y curtosis);
 - efectos asimétricos de la heteroscedasticidad (leverage and feedback effects).
- Las soluciones propuestas en la literatura:
 - usar otras distribuciones con colas pesadas o asimétricas (ej. t);
 - modelos estocásticos y determinísticos para acomodar la heteroscedasticidad;
 - considerar más de un régimen en la media y en la varianza condicional (regime switching).

Procesos lineales vs. no lineales

Una serie de tiempo $\{x_t : t = 0, 1, 2, \dots\}$ es **lineal** si

$$x_t = \mu + \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i},$$

donde $\psi_0 = 1$ y $\{\epsilon_t\}$ es ruido blanco. ϵ_t es la nueva información que se adquiere en t , también llamada innovación o shock. ψ son los “pesos” de las innovaciones del pasado en el presente.

Si esto no se satisface entonces tenemos un proceso no lineal. En general, podemos plantear una función $f(\cdot)$ tal que

$$x_t = f(\epsilon_t, \epsilon_{t-1} \dots)$$

Procesos lineales vs. no lineales

Supongamos que \mathcal{F}_{t-1} es el campo σ generado por toda la información hasta $t - 1$.
Entonces,

$$\mu_t = E(x_t | \mathcal{F}_{t-1}) = g(\mathcal{F}_{t-1}), \quad \sigma_t^2 = \text{Var}(x_t | \mathcal{F}_{t-1}) = h(\mathcal{F}_{t-1})$$

donde $g(\cdot)$ y $h(\cdot)$ son funciones bien definidas con $h(\cdot) > 0$.
Consideremos modelos de este tipo:

$$x_t = g(\mathcal{F}_{t-1}) + \sqrt{h(\mathcal{F}_{t-1})}u_t, \quad u_t = \varepsilon_t / \sigma_t$$

Entonces podemos tener, no linealidad en la media y/o en la varianza.

Cambio estructural

Cambio estructural: Consideremos el proceso $\{x_t : t = 0, 1, 2, \dots\}$. Si el comportamiento del proceso es diferente para $t \leq h$ y para $t > h$, entonces tenemos un **cambio estructural o corte** en h .

Por ejemplo:

- Cambio estructural en el intercepto: $x_t = \alpha_1 I[t \leq h] + \alpha_2 I[t > h] + \epsilon_t$.
- Cambio estructural en la tendencia: $x_t = \alpha_0 + \beta_1 t I[t \leq h] + \beta_2 t I[t > h] + \epsilon_t$.
- Cambio estructural en el parámetro autorregresivo:
 $x_t = \alpha_0 + \phi_1 x_{t-1} I[t \leq h] + \phi_2 x_{t-1} I[t > h] + \epsilon_t$.

Cambio estructural

Los contrastes para cambios estructurales dependen fundamentalmente de si se conoce el momento h en el cual ocurre el cambio:

- Si el periodo h en el cual ocurre un cambio estructural **se conoce**, entonces usamos el **contraste de Chow**.
- Si el periodo en el cual podría haber ocurrido el cambio estructural **no se conoce**, entonces usamos el contraste de Hansen (1992) o Elliott y Muller (2006):

Hansen, B. (1992), "Parameter Instability in Linear Models," Journal of Policy Modeling, 14, 517-533.

Elliott, G. and Muller, U.K., (2006), "Efficient Tests for General Persistent Time Variation in Regression Coefficients," Review of Economic Studies, 73, 907-940.

Cambio estructural

Contraste de Chow

Supongamos que se sospecha que en $t = h$ ocurrió un cambio estructural. Consideremos un modelo de regresión con variables $\{y_t, x_t\}$ (donde x_t puede contener lags de y , o sea y_{t-1} , etc., así como una constante y/o tendencia). Supongamos cambios en los parámetros de la regresión:

$$y_t = \beta_1 x_t I[t \leq h] + \beta_2 x_t I[t > h] + \epsilon_t, t = 1, 2, \dots, T \quad (1)$$

Aquí tenemos que el efecto de x en y es β_1 para $t \leq h$ pero β_2 para $t > h$. También podemos escribirlo de la siguiente manera:

$$y_t = \begin{cases} \beta_1 x_t + \epsilon_t & \text{si } t \leq h \\ \beta_2 x_t + \epsilon_t & \text{si } t > h \end{cases}, t = 1, 2, \dots, T \quad (2)$$

Es conveniente re-escribir el modelo anterior como

$$y_t = \beta x_t + \gamma_h x_t I[t > h] + \epsilon_t, t = 1, 2, \dots, T \quad (3)$$

[Ejercicio: Comprobar que estos dos modelos son equivalentes.]

Cambio estructural

Contraste de Chow

Un contraste de Chow para cambio estructural es:

$$H_0 : \beta_1 = \beta_2 \text{ (no hay cambio estructural)}$$

$$H_A : \beta_1 \neq \beta_2 \text{ (hay un cambio estructural en } h \text{)}$$

o también

$$H_0 : \gamma_h = 0 \text{ (no hay cambio estructural)}$$

$$H_A : \gamma_h \neq 0 \text{ (hay un cambio estructural en } h \text{)}$$

¿Cómo simular un cambio estructural en STATA?

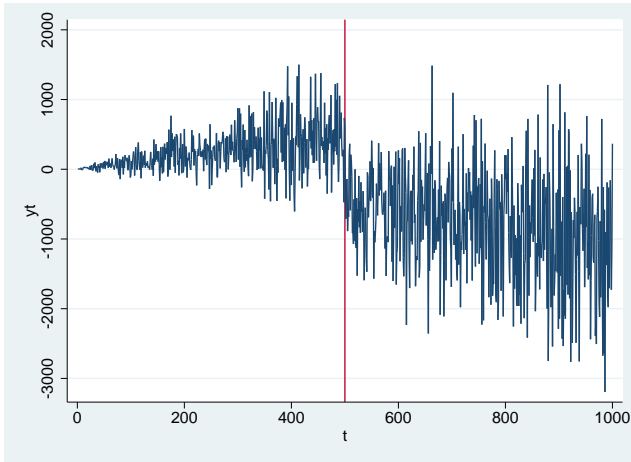
```
clear
global T=1000
set obs $T
gen t=_n
tsset t

gen et=rnormal(0,1)
gen yt=.
replace yt=et in 1
global alpha1=1
global alpha2=-1
global beta1=1
global beta2=-1
global h=500 /*el cambio ocurre en t=500*/

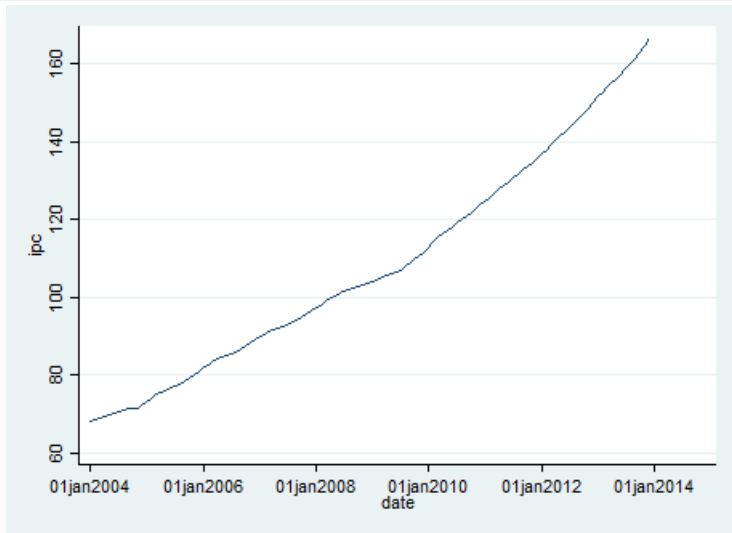
forvalues i=2(1)$T {
quietly replace yt=$alpha1 + $beta1*t+et['i'] +t*et if t<=$h in 'i'
quietly replace yt=$alpha2 + $beta2*t+et['i'] +t*et if t>$h in 'i'
}

line yt* t, xline($h)
```

Cambio estructural

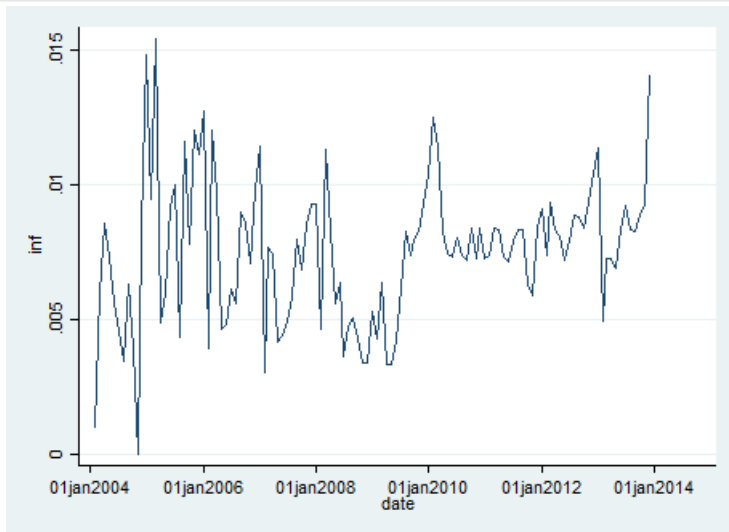


Indice de Precios al Consumidor GBA, base abril 2008=100



(Ver <http://www.indec.gov.ar/informacion-de-archivo.asp>)

Inflación mensual, en base a IPC abril 2008



(Ver <http://www.indec.gov.ar/informacion-de-archivo.asp>)

¿Cómo implementar el contraste de Chow en STATA?

- 1 Supongamos que estamos interesados en la regresión $reg\ y\ x$
- 2 Especificar el momento del cambio estructural: `global h=##` (donde `##` indica el tiempo en el cual ocurrió el cambio estructural)
- 3 Crear una variable *dummy* indicando antes y después:

```
gen after=0
replace after=1 if time>=$h
gen afterx=after*x
reg y x after afterx
test after afterx
```

[Repasar modelos de variables binarias y contrastes de hipótesis múltiples.]

- 4 Esto se puede implementar directamente en STATA después de una regresión usando el comando `estat sbknown`. Ver <https://www.stata.com/manuals14/tsestatsbknown.pdf>.

Cambio estructural

Contraste de Hansen

- Supongamos no se sabe cuando podría haber ocurrido un cambio estructural. h entonces podría haber sido cualquier valor $h \in [1, \dots, T]$.
- Para armar un test de cambio estructural en este caso debemos probar “todos” los posibles valores de h .
- Este contraste también se llama contraste de **estabilidad de los parámetros**, ya que contrasta por todos los posibles cambios en los parámetros de la regresión.
- El contraste se construye a partir de un contraste de Wald-F para cada h , $\{F_h\}_{h=1}^{T-1}$, y luego considerar el supremo (sup) o promedio (ave) de todos los h . Aunque es intuitivo la distribución de estos estadísticos no es estándar.
- En STATA:

```
ssc install ql1  
ql1 y x
```

- En STATA se puede implementar también usando el comando `estat sbsingle`. Ver <https://www.stata.com/manuals14/tsestatsbsingle.pdf>.

Modelos de umbrales (threshold models)

- Los modelos no lineales se pueden aproximar por funciones lineales por partes. En teoría una cantidad suficiente de partes (¿infinita?) puede aproximar cualquier función no lineal.
- Uno de los modelos más usados en series temporales son los **modelos autoregresivos de umbrales (threshold autoregressive models, TAR)** de Tong and Lim (1980), Tong (1983); y los SETAR (self-exciting threshold autoregressive) de Tong (1990), Tsay (1989) and Chan (1993).
- Los SETAR son modelos AR lineales por partes determinados en el espacio de los parámetros que determinan los umbrales.
- Los modelos de umbrales usan modelos lineales por partes para obtener una mejor aproximación de la media condicional.

SETAR(1,1)

Ejemplo SETAR(1,1):

$$y_t = \theta_{01}I(y_{t-1} \leq \gamma) + \theta_{11}y_{t-1}I(y_{t-1} \leq \gamma) + \theta_{02}I(y_{t-1} > \gamma) + \theta_{12}y_{t-1}I(y_{t-1} > \gamma) + \epsilon_t$$

- Hansen (1997): “(...) they are relatively simple to specify, estimate, and interpret, at least in comparison with many other nonlinear time-series models.”
- Los trabajos de Hansen (1996, 1997, 2000) desarrollan la teoría de estos modelos y proponen contrastes para no linealidades.
- Si los parámetros de umbrales γ se conocen *a priori* entonces el modelo se puede estimar con variables dummy ($I(y_{t-1} \leq \gamma)$) e interacciones. Los modelos son similares a los vistos para evaluar cambios estructurales. Por ejemplo, $\gamma = 0$ evalúa si el efecto de y_{t-1} es diferente siendo esta positiva o negativa.
- Si γ no se conoce hay que estimarlo y el modelo se vuelve más complejo.

¿Cómo estimar γ ?

$$(\hat{\theta}, \hat{\gamma}) = \arg \min_{(\theta, \gamma)} \sum_{t=1}^T (y_t - x_t(\gamma)' \theta)^2.$$

donde $\theta = (\theta_1, \theta_2)$ es un vector de 2×2 ,

$x(\gamma) = [1, y_{t-1}] \otimes (1\{y_{t-1} \leq \gamma\}, 1\{y_{t-1} > \gamma\})$, donde $1\{\cdot\}$ es una función indicador y \otimes es el producto de Kronecker.

Esto también se puede implementar en dos pasos:

- Consideremos una grilla de valores de $\gamma \in \Gamma$.
- Para cada γ se estiman los parámetros $\hat{\theta}_\gamma$ por MCO.
- Luego encontrar $\hat{\gamma}$,

$$\hat{\gamma} = \arg \min_{\gamma} \sum_{t=1}^T (y_t - x_t(\gamma)' \hat{\theta}_\gamma)^2.$$

Modelos SETAR (self-exciting threshold autoregressive)

Una serie de tiempo y_t sigue un modelo SETAR con k regimenes y variable de umbral y_{t-d} si

$$y_t = \phi_0^{(j)} + \phi_1^{(j)} y_{t-1} + \dots + \phi_p^{(j)} y_{t-p} + \epsilon_t \text{ si } \gamma_{j-1} \leq y_{t-d} < \gamma_j$$

- k y d son números enteros positivos, $j = 1, \dots, k$.
- el índice j se usa para determinar el regimen.
- γ_j son parámetros de umbrales $-\infty < \gamma_0 < \gamma_1 < \dots < \gamma_{k-1} < \gamma_k < \infty$.
- $\{\epsilon_t\}$ son ruido blanco con media 0 y varianza σ_2 .
- el parámetro d es el parámetro de retraso (delay parameter).

¿Cómo contrastar para efectos de umbrales no lineales?

Consideremos el modelo SETAR(1,1)

$$y_t = \theta_{01}I(y_{t-1} \leq \gamma) + \theta_{11}y_{t-1}I(y_{t-1} \leq \gamma) + \theta_{02}I(y_{t-1} > \gamma) + \theta_{12}y_{t-1}I(y_{t-1} > \gamma) + \epsilon_t$$

- Si γ se conoce se usa la hipótesis nula $H_0 : \theta_{01} = \theta_{02}, \theta_{11} = \theta_{12}$. Esto se implementa como un contraste F o Wald en modelos de regresión.
- Si γ no se conoce esto se vuelve problemático porque bajo la hipótesis nula $H_0 : \theta_{01} = \theta_{02}, \theta_{11} = \theta_{12}$, γ no se puede estimar. Se dice que γ no está identificado bajo H_0 .

Contrastes de Hansen

- Hansen (1996), siguiendo a Davies (1977, 1987) y Andrews and Ploberger (1994) proponen usar contrastes **sup Wald** y **ave Wald**. Estos son tests Kolmogorov-Smirnov

$$KS \sup \mathcal{W} = \sup_{\gamma \in \mathcal{G}} \mathcal{W}_T(\gamma),$$

y

$$KS \text{ ave} \mathcal{W} = \text{ave}_{\gamma \in \mathcal{G}} \mathcal{W}_T(\gamma),$$

donde

$$\mathcal{W}_T(\gamma) = T(R\hat{\theta}_\gamma)'[R\hat{V}_{\hat{\theta}_\gamma}^{-1}R']^{-1}(R\hat{\theta}_\gamma),$$

R es una matriz que construye las hipótesis de interés, $\hat{V}_{\hat{\theta}_\gamma}$ es la matriz varianza-covarianza de MCO de $\hat{\theta}_\gamma$, y \mathcal{G} es el espacio de los posibles valores de γ .

- Notar que para γ fijo, $\mathcal{W}_T(\gamma)$ es un contraste F de $H_0 : \theta_{01} = \theta_{02}, \theta_{11} = \theta_{12}$.

Contrastes de Hansen

Problema: La distribución asintótica de estos contrastes de Wald depende de la función de varianza-covarianza $\hat{V}_{\hat{\theta}_\gamma}$. Entonces es una distribución no standard.

Solución: Simular el p-valor.

Contrastes de Hansen

Simulación:

- Generar una grilla Γ de valores γ , $\Gamma = (\gamma_1, \dots, \gamma_g)$.
- Consideremos $j = 1, \dots, J$ simulaciones.
- Generar $\{v_{tj}\}_{t=1}^T$ iid $N(0, 1)$ y luego

$$\hat{S}_{T,j}^*(\gamma) = \frac{1}{\sqrt{T}} \sum_{t=1}^T x_t(\gamma) \hat{\epsilon}_t(\gamma) v_{tj},$$
 para cada $\gamma \in \Gamma$, donde $\hat{\epsilon}_t(\gamma)$ son los residuos MCO.
- Construir $\mathcal{W}_{T,j}^*(\gamma) =$

$$\left(R(X_T(\gamma)' X_T(\gamma))^{-1} S_{T,j}^*(\gamma) \right)' [R \hat{V}_{\hat{\theta}_\gamma}^{-1} R']^{-1} \left(R(X_T(\gamma)' X_T(\gamma))^{-1} S_{T,j}^*(\gamma) \right).$$
- Calcular $\sup_{\gamma \in \Gamma} \mathcal{W}_{T,j}^*(\gamma)$ y $\text{ave}_{\gamma \in \Gamma} \mathcal{W}_{T,j}^*(\gamma)$.

Esto produce una muestra de J simulaciones del proceso. Hansen (1996) demuestra que el p-valor se puede aproximar por la distribución de $\{\mathcal{W}_{T,j}^*\}_{j=1}^J$.

Modelos de Smooth Transition (STAR)

Chan and Tong (1986) y Terasvirta (1994).

- Una crítica de los modelos SETAR es que la media condicional no es continua, tiene un salto.
- Los umbrales $\{\gamma_j\}$ son los puntos de discontinuidad de la función de media condicional.
- Se dice que una serie temporal $\{x_t\}$ sigue un STAR(p) con 2 regímenes si

$$x_t = c_1 + \sum_{i=1}^p \phi_{1,i} x_{t-i} + F\left(\frac{x_{t-d} - \Delta}{s}\right) \left(c_2 + \sum_{i=1}^p \phi_{2,i} x_{t-i} \right) + \epsilon_t$$

- Δ y s son parámetros que representan la locación y la escala, respectivamente de la transición, y $F(\cdot)$ es una función de transición suave (logística, exponencial, o una función de distribución como de la normal)
- Una ventaja de estos modelos es que la función de media condicional es diferenciable...
- ... pero Δ y s son difíciles de estimar.

Markov switching models

- Hamilton (1989) enfatiza que la economía realiza transiciones no periódicas entre distintos estados. La idea de usar modelos de probabilidad para cambiar de régimen en series temporales no lineales es de Tong (1983).
- En los modelos Markov switching autoregressive (MSA) la transición se encuentra regulada por una cadena de Markov entre dos estados no observable. Markov chain
- Estos modelos son diferentes de los SETAR porque la transición entre regímenes esta dada por una variable no observable (el estado en la cadena de Markov) y no por un rezago. En consecuencia se dice que los SETAR usan un esquema **determinístico** para explicar la transición y los MSA un esquema **estocástico**. En la práctica los MSA implican que uno nunca sabe con certeza en que régimen se encuentra.

Markov switching models

- Una serie temporal $\{x_t\}$ sigue un modelo MSA(2) si

$$x_t = \begin{cases} c_1 + \sum_{i=1}^p \phi_{1,i} x_{t-i} + \epsilon_{1t} & \text{si } s_t = 1, \\ c_2 + \sum_{i=1}^p \phi_{2,i} x_{t-i} + \epsilon_{2t} & \text{si } s_t = 2, \end{cases}$$

donde s_t toma valores $\{1, 2\}$ y sigue una cadena de Markov con probabilidades de transición

$$P(s_t = 2 | s_{t-1} = 1) = w_1, \quad P(s_t = 1 | s_{t-1} = 2) = w_2$$

- $\{\epsilon_{1t}\}$ y $\{\epsilon_{2t}\}$ son ruido blanco mutuamente independientes;
- w_i pequeño significa que el modelo se queda más tiempo en el estado i .

Lecturas sugeridas

- Enders, W. "Applied Econometric Series", cap. 7.
- Johnston, J. y Dinardo, J. "Econometric Methods", cap. 4.
- Tsay, R. "Analysis of Financial Time Series", cap. 4.