

UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA  
DIPARTIMENTO  
DI SCIENZE  
STATISTICHE

UNIVERSITA' DEGLI STUDI DI PADOVA

**Facoltà di scienze statistiche**

Corso di Laurea Triennale in Scienze Statistiche e  
Gestione delle Imprese

Tesi di laurea

**Processi a memoria lunga e break  
strutturali:  
un'applicazione al tasso d'inflazione**

Relatore: Prof.ssa Luisa Bisaglia

Laureanda: Lucrezia Bonollo

N. matricola 545332-GEI

ANNO ACCADEMICO 2007-2008



*Ai miei genitori,  
a mio fratello,  
alla professoressa Luisa Bisaglia  
e a Matteo*



# Sommario

- Introduzione\_\_\_\_\_ -7-

- **Capitolo 1**    **Processi stazionari a memoria lunga**\_\_\_\_\_ 9

1.1 Che cosa sono i processi a memoria lunga\_\_\_\_\_ 9

1.2 Modelli ARIMA frazionari\_\_\_\_\_ 11

1.3 Metodi per la stima dei parametri\_\_\_\_\_ 14

1.3.1 Stimatore di Whittle\_\_\_\_\_ -14-

1.3.2 Metodo del log-periodogramma\_\_\_\_\_ -16-

1.3.3 Local Whittle method\_\_\_\_\_ -18-

- **Capitolo 2**    **Break strutturali**\_\_\_\_\_ 19

2.1 Cosa sono i break strutturali\_\_\_\_\_ -19-

2.2 Stima dei break strutturali\_\_\_\_\_ -20-

2.2.1 Test:  $l$  versus  $l+1$  break\_\_\_\_\_ -21-

2.2.2 Stima nel caso di un numero sconosciuto di break\_\_\_\_\_ -22-

|                                    |                                                   |             |
|------------------------------------|---------------------------------------------------|-------------|
| <b>- Capitolo 3</b>                | <b>Presentazione dei dati: il tasso</b>           |             |
|                                    | <b>d'inflazione</b>                               | <b>23</b>   |
| 3.1                                | L'inflazione: evidenza di memoria lunga           | 23          |
| 3.2                                | Il tasso d'inflazione italiano e statunitense     | -24-        |
| 3.3                                | Presentazione dei dati                            | -25-        |
| <br>                               |                                                   |             |
| <b>- Capitolo 4</b>                | <b>Analisi</b>                                    | <b>27</b>   |
| 4.1                                | Strategia empirica                                | 27          |
| 4.2                                | Applicazione ai dati reali                        | -28-        |
| 4.3                                | Analisi di un sottocampione delle serie originali | -42-        |
| <br>                               |                                                   |             |
| <b>- Capitolo 5</b>                | <b>Conclusioni</b>                                | <b>-51-</b> |
| <br>                               |                                                   |             |
| <b>- Riferimenti bibliografici</b> |                                                   | <b>-53-</b> |

# Introduzione

Autorità monetarie, istituti finanziari, fondi pensioni e investitori privati domandano, oggi, sempre più una maggiore precisione dei modelli statistici per l'inflazione che permettano una corretta stima di variabili, quali la ricchezza, il reddito e i profitti nel lungo e nel breve periodo.

Recentemente è stato adottato un nuovo approccio per lo studio, appunto, dell'inflazione. Sembra, infatti, evidente che molte serie storiche economiche sperimentino l'esistenza di occasionali *break* strutturali. Allo stesso tempo, per alcune di queste variabili c'è un'evidenza di memoria lunga. In particolare, sembra che il tasso di inflazione abbia entrambe le caratteristiche. Una causa di ciò potrebbe essere che le due tipologie sono difficili da distinguere usando gli attuali strumenti econometrici. Infatti, recenti studi hanno mostrato che trascurare *break* occasionali potrebbe condurre a scoprire erroneamente proprietà di memoria lunga nelle serie storiche.

In particolare, se un processo a memoria breve è affetto da *break* occasionali, la sua realizzazione dà l'impressione di persistenza e la stima del parametro di memoria lunga differisce da zero.

L'importanza di distinguere tra modelli a memoria lunga e modelli con *break* strutturali occasionali è giustificata anche dalla prospettiva di previsione.

Lo scopo di questo lavoro è quello di analizzare se, serie storiche reali sono meglio descritte come realizzazioni di processi a memoria lunga o come processi con *break* strutturali. Tuttavia, la letteratura relativa a tale soggetto non è ancora ben sviluppata; sebbene l'interesse nella possibilità di distinguere processi a memoria lunga e processi con *break* occasionali sia notevolmente cresciuto.

Questo lavoro è strutturato come segue. Nel primo capitolo vengono affrontati i processi a memoria lunga e, in particolare, i diversi metodi per stimare il parametro di memoria lunga,  $d$ . Nel secondo capitolo vengono presentati i *break* strutturali ed esposti i metodi per stimarli. Nel capitolo 3 esaminiamo il tasso di inflazione e il suo comportamento nelle serie storiche

successivamente analizzate. Nel capitolo 4 viene descritta e, successivamente applicata, una procedura empirica per analizzare il tasso di inflazione quando i dati mostrano persistenza, al fine di rilevare effetti sospetti causati da una non corretta specificazione del modello. La strategia si basa su un differente comportamento dello stimatore di memoria lunga quando questo viene applicato nella serie originale e in quella “purificata” dai *break*. In particolare l’analisi viene effettuata su serie economiche relative all’indice dei prezzi al consumo di Italia e USA che coprono un periodo di tempo che va da gennaio 1961 ad aprile 2008.

# Capitolo 1

## Processi stazionari a memoria lunga

### 1.1 Che cosa sono i processi a memoria lunga

L'interesse per i processi a memoria lunga è nato negli anni '50 nel campo dell'ingegneria, in particolare nell'idrologia: il primo studio fu proposto da Hurst (1951) a proposito delle proprietà di riempimento dei bacini idrici. Solo a partire dagli studi di Mandelbrot, nei primi anni '60, il tema della memoria lunga si è rapidamente sviluppato, anche se è solo dall'inizio degli anni '80 che questo tipo di modellistica ha iniziato a destare vero interesse tra gli econometrici.

In questo capitolo vengono descritti alcuni processi stocastici che possono essere utilizzati per modellare dati con dipendenza di lungo periodo, ossia che presentano memoria lunga.

Le serie temporali generate da questo tipo di processi presentano caratteristiche ben precise. Innanzitutto le osservazioni tendono ad assumere valori al di sopra o al di sotto della media, per periodi di tempo relativamente lunghi: valori elevati tendono ad essere seguiti da valori elevati, mentre valori bassi sono tendenzialmente seguiti da valori bassi. Inoltre, nelle serie sono presenti cicli non periodici in successione casuale; risulta, infatti, evidente l'assenza di *trend* persistenti o di cicli. Le serie, poi, appaiono complessivamente stazionarie sia in media che in varianza, ma la varianza della media campionaria converge a zero più lentamente di  $n^{-1}$ .

Importante, inoltre, sottolineare che la funzione di autocorrelazione (autocovarianza) converge molto lentamente a zero; questo implica che la dipendenza tra osservazioni successive decade lentamente all'aumentare della distanza temporale tra le stesse.

Analogamente, la funzione di densità spettrale presenta un polo all'origine, cioè diverge all'origine come una potenza. Asintoticamente:

$$f(\lambda) \sim c_f(\lambda)|\lambda|^{-\beta} \text{ quando } \lambda \rightarrow 0,$$

dove la  $f(\lambda)$  è la funzione di densità spettrale,  $c_f(\lambda)$  è una costante,  $\lambda \in (-\pi, \pi)$  e  $\beta \in (0, 1)$ .

Se si assume che il processo osservato sia stazionario, è possibile definire un processo a memoria lunga o con dipendenza di lungo periodo o dipendente in senso forte nel seguente modo:

**Definizione 1.1.1** *Il processo  $X = (X_t, t \in Z)$  viene chiamato processo stazionario con memoria lunga o dipendenza di lungo periodo, se è stazionario e se esiste un numero reale  $H^1 \in (1/2, 1)$  e una funzione  $L_f(\lambda) > 0$  che varia lentamente a zero, tale che*

$$f(\lambda) \sim L_f(\lambda)|\lambda|^{1-2H},$$

*quando  $\lambda$  tende a 0 per  $\lambda \in (-\pi, \pi)$ . Dove  $f(\lambda)$  è la funzione di densità spettrale.*

Poiché la conoscenza della funzione di densità spettrale è equivalente alla conoscenza della funzione di autocorrelazione, infatti:

$$f(\lambda) = (1/2\pi) \sum_{k=-\infty}^{\infty} e^{-i\lambda k} \rho(k),$$

dove  $\rho(k)$  è la funzione di autocorrelazione, un processo a memoria lunga può essere definito nel seguente modo:

**Definizione 1.1.2** *Il processo  $X = (X_t, t \in Z)$  viene chiamato processo stazionario con memoria lunga o dipendenza di lungo periodo, se è stazionario e se esiste un numero reale  $H^1 \in (1/2, 1)$  e una funzione  $L_{\rho(k)}$  che varia lentamente all'infinito, tale che:*

$$\rho(k) \sim L_{\rho(k)} k^{2H-2},$$

*quando  $k$  tende a  $\infty$ .*

---

<sup>1</sup> Tra il parametro  $H$  e il parametro  $d$  esiste una semplice relazione lineare del tipo:  
 $H = d + 1/2$ .

Esiste una classe di modelli per cui sono soddisfatte le condizioni date nelle due definizioni precedenti. Si tratta dei modelli ARIMA frazionari che tratteremo nel prossimo paragrafo.

## 1.2 Modelli ARIMA frazionari

I modelli ARIMA (p,d,q) Frazionari, detti anche ARFIMA (p,d,q), sono stati introdotti da Hosking (1981) e Granger e Joyeux (1980) come una generalizzazione dei modelli ARMA(p,q) per descrivere la dinamica di fenomeni che presentano una struttura di dipendenza che persiste nel tempo. Questi modelli risultano utili per interpretare e modellare processi le cui realizzazioni evidenziano correlazioni seriali significative anche se misurate in tempi molto distanti. Appare quindi necessario, per adeguare la modellistica standard alla dinamica di molti fenomeni reali, generalizzare la classe dei processi lineari ARIMA (Box e Jenkins, 1970; Piccolo 1990) al fine di poter tener conto in modo esplicito di una componente di lunghissimo periodo, che spieghi i comportamenti, prima evidenziati, rilevabili nella funzione di autocorrelazione globale.

In tale prospettiva, la classe dei modelli ARFIMA (*AutoRegressive Fractional Integrated Moving Average*) offre uno strumento flessibile per descrivere tale tipo di dinamica.

I modelli ARMA (p,q) possono essere pensati come un ARIMA (p,0,q). I modelli ARFIMA (p,d,q), a loro volta, generalizzano i modelli ARIMA (p,d,q) in quanto il parametro d può assumere un qualsiasi valore reale.

A questo punto si può definire un processo ARFIMA (p,d,q) come segue:

**Definizione 1.2.1**  $X = (X_t, t \in Z)$  è un processo ARFIMA (p,d,q) con  $d \in (-1/2, 1/2)$ , se è stazionario<sup>2</sup> e se soddisfa l'equazione alle differenze

$$\Phi(B)\Delta^d X_t = \Theta(B)\epsilon_t$$

---

<sup>2</sup> Il processo è stazionario se le radici di  $\Phi(B) = 0$  cadono fuori del circolo unitario.

dove  $\epsilon_t \sim w.n.(0, \sigma^2)$  e  $\Phi(.)$  e  $\Theta(.)$  sono polinomi di grado  $p$  e  $q$  rispettivamente<sup>3</sup>.

Si osservi che, in caso di  $d \in \mathbf{R}$ , allora:

$$\Delta^d = (1 - B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-1)^k B^k = \sum_{k=0}^{\infty} \pi_k B^k,$$

dove:

$$\pi_k = \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)} = \prod_{0 < j \leq k} \frac{j-1-d}{j}, j = 0, 1, 2, \dots$$

dove  $\Gamma(.)$  è la funzione gamma.

Se  $p = q = 0$ , il processo  $X = (X_t, t \in \mathbf{Z})$  è spesso chiamato Rumore Integrato Frazionario (Fractionally Integrated Noise) e viene denotato come ARFIMA (0,d,0), e viene trattato in questo lavoro in quanto descrive la sola dinamica di memoria lunga. La letteratura fa riferimento a tali modelli, come modelli FI(d).

Tale processo si presenta così:

$$\Delta^d X_t = \epsilon_t^4$$

I processi FI(d) possono essere espressi in forma  $MA(\infty)$ , come nel seguente teorema:

**Teorema 1.2.1** Se  $d \in (-1/2, 1/2)$  allora esiste un'unica soluzione stazionaria  $X_t$  di  $\Delta^d X_t = \epsilon_t$  data da:

$$X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} = \Delta^{-d} \epsilon_t$$

dove  $\psi_0 = 1$  e

$$\psi_j = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} = \prod_{k=1}^j \frac{k-1+d}{k}, j = 1, 2, \dots$$

---

<sup>3</sup>  $\Phi(.)$  e  $\Theta(.)$  sono due polinomi di grado  $p$  e  $q$  rispettivamente, tali che:

$$\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$$

$$\Theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$$

e  $B$  è tale che  $B^j X_t = X_{t-j}$ ,  $j = 0, 1, \dots$

<sup>4</sup> Viene assunta per convenienza e senza perdita di generalità  $\sigma^2 = 1$  e  $\mu = 0$ .

Inoltre, la funzione di densità spettrale è:

$$f(\lambda) = \frac{\sigma^2}{2\pi} |2 \sin(\lambda/2)|^{-2d}, \quad -\pi \leq \lambda \leq \pi$$

la varianza del processo è:

$$\gamma(0) = \sigma^2 \frac{\Gamma(1-2d)}{\Gamma^2(1-d)}$$

e la funzione di autocorrelazione è:

$$\rho(k) = \frac{\Gamma(k+d)\Gamma(1-d)}{\Gamma(k-d+1)\Gamma(d)} = \prod_{j=1}^k \frac{j-1+d}{j-d}, \quad k = 1, 2, \dots$$

Quando  $j$  tende a  $\infty$  è possibile dimostrare che:

$$\psi_j \sim \frac{j^{d-1}}{\Gamma(d)}$$

e

$$\rho(k) \sim k^{2d-1} \frac{\Gamma(1-d)}{\Gamma(d)}.$$

Inoltre, quando  $\lambda$  tende a 0:

$$f(\lambda) = \frac{\sigma^2}{2\pi} \lambda^{-2d}.$$

Il parametro  $d$  misura l'intensità della dipendenza. Quando  $d \in (0, 1/2)$  il processo ARFIMA (p,d,q) è un processo stazionario con memoria lunga, le sue correlazioni sono tutte positive e decadono ad un tasso iperbolico. Quando  $d = 0$  si ottiene nuovamente l'usuale modello ARMA (p,q). Quando  $d \in (-1/2, 0)$  il processo è stazionario con memoria intermedia o dipendenza negativa o antipersistente. Il parametro  $d$  quindi misura l'intensità della dipendenza. Inoltre se  $d \in (0, 1/2)$  il processo è stazionario e invertibile. Quindi testare l'ipotesi nulla di dipendenza debole contro l'ipotesi alternativa di memoria lunga equivale a testare  $d = 0$  contro  $d \neq 0$ . Da qui in avanti viene fatto riferimento solo al caso  $d \in (0, 1/2)$ , cioè al caso di memoria lunga in senso stretto. In tal caso, il processo  $FI(d)$  possiede una funzione di autocorrelazione che tende molto lentamente a zero e una funzione di densità spettrale dominata dalle componenti a bassa frequenza.

## 1.3 Metodi per la stima dei parametri

In questo paragrafo descriveremo alcune delle principali metodologie per stimare il parametro di memoria lunga  $d$  (o  $H = d + 1/2$ ). È possibile classificare i vari metodi in due gruppi: metodi parametrici e semiparametrici. Per quanto riguarda il primo gruppo consideriamo il metodo di *Whittle* (Fox e Taqqu, 1986). Appartenenti al secondo gruppo, invece, consideriamo il metodo di *Geweke- Porter Hudak* (1983) (denominato da qui in avanti con **GPH**) e il metodo *local Whittle* (Kunsch, 1987, Robinson 1995).

### 1.3.1 Stimatore di Whittle

Questo metodo si basa sulla massimizzazione di un'approssimazione della funzione di verosimiglianza.

Whittle (1951) ha introdotto tale approssimazione nel caso di variabili casuali debolmente dipendenti e la letteratura riguardante tale argomento fa quasi interamente riferimento al suo lavoro.

Si consideri una realizzazione,  $X = (X_t, t = 1, 2, \dots, N)$ , di un processo gaussiano lineare, invertibile e causale. La funzione di verosimiglianza si presenta nel seguente modo:

$$L_N(X, \vartheta) = -\frac{N}{2} \log 2\pi - \frac{1}{2} \log |\Sigma_N(\vartheta)| - \frac{1}{2} X' \Sigma_N^{-1}(\vartheta) X$$

dove  $\vartheta = (\sigma^2, H, \vartheta_3, \dots, \vartheta_M)$  è il vettore di parametri sconosciuti e  $\Sigma_N(\vartheta)$  è la matrice di varianze e covarianze.

Per semplificare i calcoli ed evitare di incorrere in alcuni problemi, è possibile approssimare i due termini  $\log |\Sigma_N(\vartheta)|$  e  $\Sigma_N^{-1}(\vartheta)$  mediante le seguenti espressioni:

$$\log |\Sigma_N(\vartheta)| \sim \frac{N}{2\pi} \int_{-\pi}^{\pi} \log f(\lambda, \vartheta) d\lambda,$$

e

$$\Sigma_N^{-1}(\vartheta) \sim A(\vartheta) = [\alpha_{j-k}(\vartheta)]_{j,k=1,2,\dots,N},$$

dove

$$\alpha_{j-k}(\vartheta) = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \frac{e^{i(j-k)\lambda}}{f(\lambda, \vartheta)} d\lambda.$$

Minimizzando, rispetto a  $\vartheta$ , la funzione di log-verosimiglianza approssimata è possibile ottenere in tal modo un'approssimazione dello stimatore di massima verosimiglianza, del vettore di parametri  $\vartheta$ .

La funzione di log-verosimiglianza approssimata si presenta perciò nel seguente modo:

$$L_W(X, \vartheta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(\lambda, \vartheta) d\lambda + \frac{X' A(\vartheta) X}{N}$$

dove

$$\frac{X' A(\vartheta) X}{N} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I_N(\lambda)}{f(\lambda, \vartheta)} d\lambda$$

dove  $I_N(\lambda)$  è il periodogramma, cioè:

$$I_N(\lambda) = \frac{1}{2\pi N} \left| \sum_{j=1}^N e^{ij\lambda} (X_j - \bar{X}_N)^2 \right|.$$

La funzione di log-verosimiglianza diventa così:

$$L_W(X, \vartheta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left( \log f(\lambda, \vartheta) + \frac{I_N(\lambda)}{f(\lambda, \vartheta)} \right) d\lambda.$$

L'indice W denota lo stimatore di Whittle, mentre  $\hat{\vartheta}_W$  il valore di  $\vartheta$  che minimizza la funzione di cui sopra.

Lo stimatore di *Whittle* è computazionalmente semplice, tuttavia questo stimatore ha lo svantaggio di richiedere una specificazione a priori della forma parametrica di  $f(\lambda, \vartheta)$ .

### 1.3.2 Metodo del log-periodogramma

Il metodo del log- periodogramma, GPH, è particolarmente semplice da utilizzare, in quanto non richiede la conoscenza del modello sottostante i dati. Questo modello è stato introdotto originariamente da Geweke e Porter-Hudack (1983) e poi sviluppato da Robinson (1995) e si basa essenzialmente sulla proprietà della funzione di densità spettrale che in un intorno dello zero diventa infinita.

La funzione spettrale di un processo ARFIMA(p,d,q) può essere approssimata nel seguente modo:

$$f(\lambda) \sim (4 \sin^2(\frac{\lambda}{2}))^{-d}, \quad (1)$$

quando  $\lambda$  tende a 0. Inoltre, il periodogramma  $I(\lambda)$  è una stima asintoticamente non distorta della funzione di densità spettrale, cioè:

$$\lim_{\lambda \rightarrow 0} E[I(\lambda)] = f(\lambda)$$

Quindi è possibile stimare il parametro  $d$  applicando il metodo dei minimi quadrati alla seguente regressione:

$$\ln(I(\lambda_{j,N})) = \ln c_f + \beta \ln(4 \sin^2(\frac{\lambda_{j,N}}{2})) + \epsilon_j, \quad j = 1, \dots, g(N)$$

dove  $\beta = -d$ ,  $I(\lambda_{j,N})$  è il periodogramma calcolato alle frequenze di Fourier  $\lambda_{j,N} = (2\pi j/N)$ , con  $j = 1, 2, \dots, m$  e gli  $\epsilon_j, j = 1, 2, \dots, m$  sono termini di errore *i.i.d.*.

Importante notare che la (1) si caratterizza per il comportamento di  $f(\lambda)$  nell'intorno di 0, ma non specifica nulla riguardo al comportamento di breve periodo del processo. Perciò, stime robuste e test nel dominio delle frequenze possono essere effettuate usando il periodogramma (o alcune sue funzioni) valutate in un intorno degenerato della frequenza di zero. Per fare ciò è necessario scegliere il numero di frequenze  $m$ , in modo tale che tenda ad infinito al tendere ad infinito della grandezza del campione ( $N$ ), ma più lentamente di  $N$ , ossia in modo tale che  $m/N$  tenda a 0. Se il numero di ordinate  $m$  è scelto in modo tale che  $m = g(N)$ , dove  $g(N)$  è tale che  $\lim_{N \rightarrow \infty} g(N) = \infty$ ,  $\lim_{N \rightarrow \infty} \{g(N)/N\} = 0$ ,  $\lim_{N \rightarrow \infty} \{(\log(N))^2/g(N)\} = 0$ ,

allora lo stimatore dei minimi quadrati di  $d$  nell'equazione (1) convergerà ad una distribuzione non degenera, in particolare:

$$(\hat{d}_{GPH} - d) / \{\text{var}(\hat{d}_{GPH})\}^{1/2} \longrightarrow N(0, 1).$$

dove  $\text{var}(\hat{d}_{GPH})$  è l'usuale stimatore dei minimi quadrati.

Uno dei maggiori problemi riscontrati nell'applicazione dello stimatore GPH è stata appunto la scelta di  $m$ . Diebold e Rudebusch (1989) scelgono  $m = N^{1/2}$ , mentre Sowell (1992) ha argomentato che  $m$  si dovrebbe basare sul più corto ciclo associato al comportamento di lungo periodo. Per esempio, con una serie di 40 anni dove 2 anni sono il ciclo più corto,  $m$  dovrebbe essere scelto di numerosità  $40/2 = 20$  ordinate. Questa regola è deliberatamente indipendente dalla frequenza del campione dei dati in quanto, con dati mensili,  $m$  verrebbe scelta ugualmente come  $160/8 = 20$ . Un'altra possibilità è quella di scegliere  $m$  tale che la regressione della varianza residuale sia approssimativamente uguale a  $\pi^2/6$ . Tuttavia, bisogna essere consapevoli che, dalla scelta di  $g(N)$  dipenderà anche la stima di  $d$ . In questo lavoro verrà considerato  $m = N^{1/2}$

Robinson (1995), come accennato precedentemente, sviluppa ed estende i risultati ottenuti da Geweke e Porte-Hudack per tentare di dare risposta ad alcuni problemi da loro riscontrati.

Egli considera tutti i modelli per cui il comportamento della funzione di densità spettrale quando  $\lambda$  tende a 0 è del tipo:

$$f(\lambda) \sim C|\lambda|^{-2d}$$

Dove  $0 < C < \infty$  e  $d \in (-1/2, 1/2)$ .

Poiché i modelli ARFIMA(p,d,q), possiedono una funzione di densità spettrale del tipo sopra riportata quando  $\lambda$  tende a 0, è possibile considerare la semplice equazione di regressione:

$$\ln(I(\lambda_{j,N})) = \alpha + \beta \ln(|\lambda_{j,N}|) + \epsilon_j, j = 1, 2, \dots, m$$

dove  $\beta = -2d$  per ottenere, applicando il metodo dei minimi quadrati, uno stimatore consistente e asintoticamente normale di  $d$ .

### 1.3.3 Local Whittle method

Lo stimatore *local Whittle* fu suggerito da Kunsch (1987) e Robinson (1995b). Questo stimatore non necessita della condizione di Gaussianità né tantomeno richiede la specificazione della densità spettrale. Infatti, tutto ciò che importa è che nell'intorno di zero la densità spettrale di un processo a memoria lunga al tendere di  $\lambda$  a zero, sia del tipo:

$$f(\lambda, d) = C|\lambda|^{-2d}$$

dove  $d \in (-1/2, 1/2)$  e  $0 < C < \infty$ . In tal modo, sostituendo  $f(\lambda, d)$  con  $C|\lambda|^{-2d}$ , è possibile ottenere una versione locale di  $L_w(\vartheta)$  che dipende solo da  $L$  e  $d$ , e non dai coefficienti ARMA, come di seguito:

$$R(L, d) = \frac{1}{m} \sum_{j=1}^m \left( \frac{I(\lambda) \lambda^{2d}}{L} + \log L \lambda - 2d \right)$$

dove  $m$  rappresenta il numero di frequenze incluse nell'intorno della frequenza zero, un intero più piccolo di  $N/2$ , in modo tale che  $1/m + m/N \rightarrow 0$  al tendere di  $N$  a 0. Lo stimatore *local Whittle*  $\hat{d}_{mLW}$  è così ottenuto minimizzando l'equazione  $R(L, d)$ .

Quando non ci sono cambiamenti in media, Robinson (1995 b) ha provato, inoltre, che:

$$m^{1/2}(\hat{d}_{mLW} - d) \xrightarrow{d} N(0, 1/4).$$

# Capitolo 2

## Break strutturali

### 2.1 Cosa sono i break strutturali

In molte analisi di tipo statistico e/o econometrico, una delle assunzioni fondamentali, riguardante i dati (le serie storiche) a disposizione, è quella di stazionarietà. Se i dati sono non stazionari l'inferenza statistica potrebbe essere non attendibile. Nel caso di serie temporali, la non stazionarietà può essere determinata dalla presenza di trend nel livello della serie o, caso che interessa a noi, dalla presenza di "rotture" (*break*) strutturali o cambiamenti di regime. In particolare, nelle serie storiche di tipo economico ciò può avvenire per diverse ragioni: ad esempio a causa di cambiamenti nelle politiche economiche, nella struttura dell'economia o a causa dell'introduzione di nuove tecnologie che modificano uno specifico settore. Questi cambiamenti o "*break*" possono derivare da gradualità evoluzioni dei coefficienti di regressione della popolazione su un lungo periodo di tempo o da cambiamenti discreti nei coefficienti a date ben distinte. Se questi *break* vengono trascurati possono portare a fuorvianti inferenze e previsioni. Tali questioni, riguardanti cambiamenti strutturali delle serie storiche, sono stati studiati in econometria e statistica (Bai (1994), Zivot e Andrews (1992)) . Tuttavia, gran parte di questi studi riguardano l'analisi di cambiamenti singoli. Recentemente l'interesse è cresciuto e si sono sviluppati modelli più complessi che studiano cambiamenti strutturali multipli verificatisi in date sconosciute.

Il primo problema causato dai *break* consiste nel fatto che lo stimatore di regressione OLS (ordinary least square) stimerà nell'intero campione una relazione valida "in media", nel senso che la sua stima sarà frutto di una combinazione di differenti periodi, ma a seconda del periodo in cui ci si trova, questa funzione di regressione "media" può essere molto diversa dalla vera

funzione di regressione per quella particolare parte del campione. Come conseguenza, le previsioni ottenute dalla funzione di regressione “media” potrebbero essere pessime.

Inoltre, la presenza di *break* strutturali potrebbe, erroneamente, portarci a credere che la serie storica sia caratterizzata dalla presenza di memoria lunga, in quanto la funzione di autocorrelazione può risultare significativa anche per grandi ritardi (perché a causa dei *break* ci sono molte osservazioni sia sopra che sotto la media).

In questa sezione, vengono presentati alcuni test per stimare i *break* nelle serie storiche. Nei successivi capitoli questa procedura viene presa in considerazione per una stima empirica dei *break* nelle serie storiche relative al *Consumer Price Index* di Italia e USA, al fine di discernere tra effettiva memoria lunga o apparenza di memoria lunga causata da *break* trascurati.

Per quanto riguarda il problema di testare cambiamenti strutturali, contributi recenti sono quelli di Andrews (1993) e di Andrews e Ploberger (1994). Questioni riguardanti le proprietà di distribuzione delle stime dei parametri sono, invece, trattate da Bai (1994a,b).

## 2.2 Stima dei break strutturali

Uno dei metodi più usati per stimare e testare *break* multipli la cui data di realizzo è sconosciuta è la procedura di Bai-Perron (d'ora in avanti BP) (Bai e Perron, 1998, 2003). Esso stima cambiamenti multipli in modelli di regressione lineari, minimizzando la somma dei residui al quadrato. Si basa su un tipo di test in cui abbiamo l'ipotesi nulla di nessun cambiamento contro l'ipotesi alternativa di un numero pre specificato o arbitrario di cambiamenti; alternativamente un altro test può prevedere  $I$  cambiamenti sotto l'ipotesi nulla e  $I+1$  cambiamenti sotto l'ipotesi alternativa. Quest'ultimo test è particolarmente utile in quanto permette di determinare consistentemente un appropriato numero di cambiamenti nei dati.

Tale test può essere costruito tenendo conto di una differente correlazione seriale negli errori, di differenti distribuzioni nei dati e negli errori nei diversi segmenti o imponendo una struttura comune.

Qui di seguito viene presentato uno dei metodi sopra elencati che è verosimilmente il più utile e di particolare rilevanza in ambito pratico. Inoltre, questo metodo viene usato dal programma statistico R che abbiamo preso in considerazione per la nostra analisi nei capitoli seguenti.

Successivamente si fa riferimento al caso in cui il test venga costruito tenendo in conto di una differente correlazione seriale degli errori.

### 2.2.1 Test: $l$ versus $l+1$ break

Questo paragrafo considera il caso di un test la cui ipotesi nulla è di  $l$  break contro l'ipotesi alternativa di un addizionale *break* esistente. Idealmente, si vorrebbe far basare il test sulla differenza tra la somma dei residui al quadrato ottenuta con  $l$  e  $l+1$  break. La distribuzione di questo test statistico è comunque difficile da ottenere. Perciò di seguito viene proposta una strategia differente. Per il modello con  $l$  break, la stima degli stessi, indicata con  $\hat{N}_1, \dots, \hat{N}_l$ , è ottenuta minimizzando globalmente la somma dei residui al quadrato. Successivamente si testa ogni  $l+1$  segmento (ottenuto usando la partizione stimata  $\hat{N}_1, \dots, \hat{N}_l$ ) per la presenza di un *break* addizionale. La grandezza dei segmenti viene assunta fissa per comodità.

Il test equivale all'applicazione dei test  $l+1$  ad ogni segmento contenente le osservazioni da  $\hat{N}_{i-1} + 1$  a  $\hat{N}_i$  ( $i = 1, \dots, l+1$ ) usando per convenzione  $\hat{N}_0 = 0$  e  $\hat{N}_{l+1} = N$  con ipotesi nulla di nessun cambiamento strutturale contro l'ipotesi alternativa di un cambiamento singolo. Si rifiuta l'ipotesi nulla in favore di un modello con  $l+1$  break se il valore minimo complessivo della somma dei residui al quadrato (su tutti i segmenti dove un *break* addizionale è incluso) è sufficientemente più piccola della somma dei residui al quadrato del modello

con  $l$  *break*. La data del *break* così selezionato è quella associata con questo valore minimo. Più precisamente, il test è definito come:

$$F_N(l+1|l) = \left\{ S_N(\hat{N}_1, \dots, \hat{N}_l) - \min_{1 \leq i \leq l+1} \cdot \inf_{\tau \in \Lambda_{i,\eta}} S_N(\hat{N}_1, \dots, \hat{N}_{i-1}, \tau, \hat{N}_i, \dots, \hat{N}_l) \right\} / \hat{\sigma}^2,$$

dove

$$\Lambda_{i,\eta} = \left\{ \tau; \hat{N}_{i-1} + (\hat{N}_i - \hat{N}_{i-1})\eta \leq \tau \leq \hat{N}_i - (\hat{N}_i - \hat{N}_{i-1})\eta \right\}$$

$S_T(\cdot)$  è la somma dei residui al quadrato e  $\hat{\sigma}^2$  è una stima consistente di  $\sigma^2$  sotto l'ipotesi nulla.

### 2.2.2 Stima nel caso di un numero sconosciuto di break

Il metodo si sviluppa partendo con lo stimare un modello con un piccolo numero di *break* che sono pensati essere necessari (o partendo con nessun *break*). Successivamente si svolge un test di costanza dei parametri per ogni sottocampione (quelli ottenuti spezzando la serie ai punti dove si stimano i *break*), aggiungendo un *break* ad un sottocampione associato con un rifiuto del test  $F_T(l+1|l)$ . Questo processo è ripetuto aumentando  $l$  sequenzialmente fino al punto in cui il test sbaglia nel rifiutare l'ipotesi nulla di assenza di cambiamenti strutturali addizionali. Un chiaro vantaggio di questo modello è che esso permette facilmente di tenere in considerazione l'effetto di possibili correlazioni seriali degli errori.

## Capitolo 3

# Presentazione dei dati: il tasso d'inflazione

### 3.1 L'inflazione: evidenza di memoria lunga

Modelli a memoria lunga iniziano ad essere studiati in Econometria e Finanza da quando Mandelbrot (1969) introdusse specificazioni di memoria lunga nei cosiddetti *price processes*. Modelli a memoria lunga, poi, sono stati largamente usati negli anni '80 quando Geweke e Porter-Hudak (1983) svilupparono uno stimatore basato sulla regressione del log-periodogramma per l'ordine di integrazione del parametro  $d$  nel modello ARFIMA  $(p,d,q)$  di Granger e Joyeux (1980) e Hosking (1981). Essi applicarono il loro stimatore ai dati relativi all'indice dei prezzi al consumo del dopoguerra negli Stati Uniti d'America (USA) e scoprirono che l'inflazione poteva essere generata da un processo  $FI(d)$ , con  $d$  attorno allo 0.5, chiaramente differente dal valore di zero e uno, assunti nella precedente letteratura. Questa fu una caratterizzazione interessante e alternativa del comportamento di lungo periodo dell'inflazione con importanti implicazioni economiche.

In particolare una delle caratteristiche di questo tipo di serie con memoria lunga è che l'autocorrelazione decresce a zero molto lentamente come per serie non stazionarie.

## 3.2 Il tasso d'inflazione italiano e statunitense

In questo lavoro vengono esaminate le serie storiche relative al tasso di inflazione mensile di Italia e Stati Uniti d'America.

Il campione copre un periodo che va da gennaio 1961 ad aprile 2008 con un totale di 568 osservazioni per ogni Paese (l'anno base è il 2000).

In questo capitolo vengono analizzate brevemente dal punto di vista storico le serie relative ai due Paesi mettendo l'accento su quei cambiamenti che possono aver portato ai *break* di cui sopra.

Per quanto riguarda l'Italia, la seconda metà degli anni '70 è caratterizzata da forti fluttuazioni, conseguenza di due shock petroliferi e di frequenti deprezzamenti della moneta. Con l'inizio degli anni '80 il tasso di inflazione variò meno rispetto a prima e ciò fu dovuto ad una politica economica restrittiva da parte del Governo Italiano. Negli anni '90, quando l'Italia si stava lentamente preparando ad entrare a far parte dell'Unione Monetaria Europea, rispettando i parametri di Maastricht, questa politica fu addirittura maggiormente indispensabile e quindi implementata.

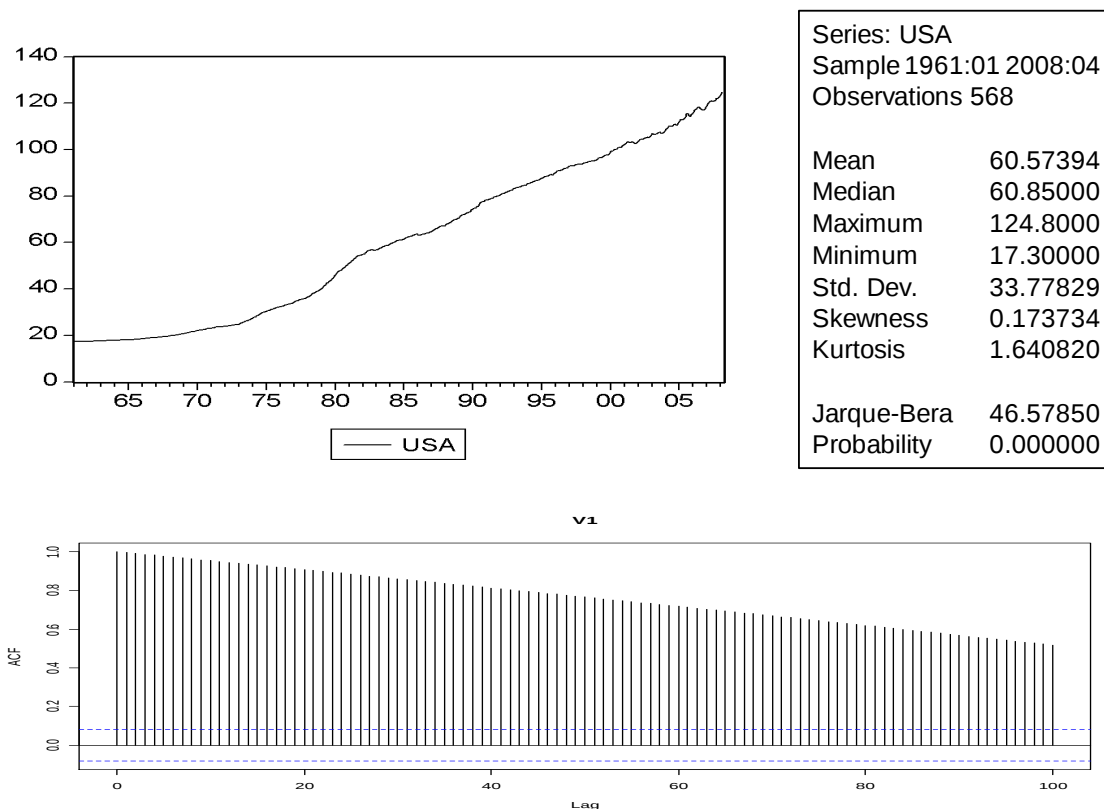
Per quanto riguarda invece il tasso di inflazione americano, è noto che questo subì un notevole numero di *break* strutturali nel periodo del dopoguerra. Ci sono diverse spiegazioni a supporto di tali cambiamenti. Una di queste, ad esempio, è relativa ai cambiamenti nelle politiche economiche (nel breve periodo con controlli di prezzo diretti tra il 1950 e il 1970, nel lungo periodo con controlli indiretti dell'inflazione iniziati negli anni '80) o ai cambiamenti nel processo di generazione di shock dei prezzi. Questi shock possono essere dovuti ai processi tecnologici, a cambiamenti nella disoccupazione, a fluttuazioni del reale costo del lavoro, a shock petroliferi, a cambiamenti nella distribuzione dei cambiamenti di prezzi, ai sindacati, ecc. Cecchetti, Hooper, Kasman, Schoenholtz, e Watson (2007) forniscono una lunga discussione sull'evidenza empirica di cambiamenti nelle proprietà di serie storiche riguardanti l'inflazione. È stato riconosciuto che i modelli empirici dovrebbero considerare tali cambiamenti e allo stesso tempo la caratteristica di persistenza dell'inflazione.

### 3.3 Presentazione dei dati

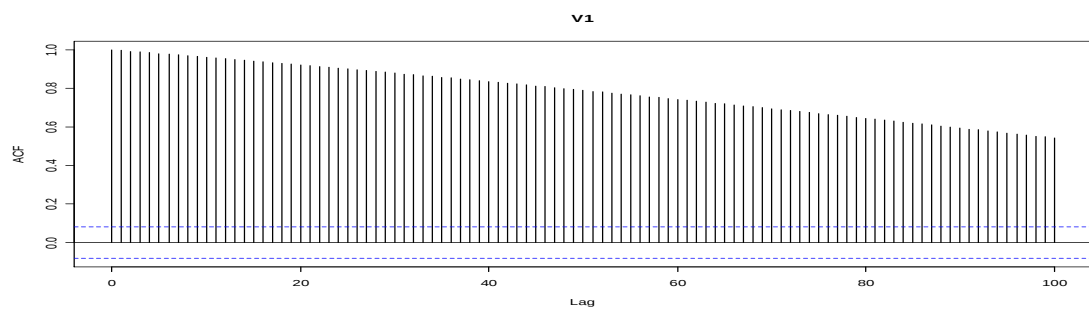
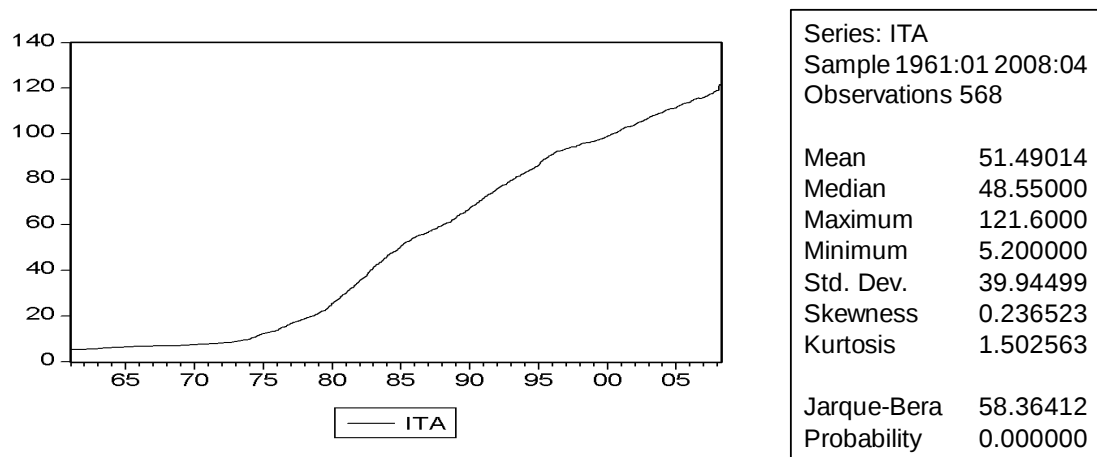
Le serie del tasso di inflazione sono costruite prendendo 100 volte la differenza prima del logaritmo della serie dell'indice dei prezzi al consumo. Le serie sono state poi destagionalizzate attraverso il metodo, additivo, delle medie mobili.

I dati sono stati reperiti su un database presente nel sito dell'OECD: [www.sourceoecd.org](http://www.sourceoecd.org).

Le figure 1 e 2 riportano il grafico delle serie originarie e la funzione di autocorrelazione globale per USA e Italia. Come si può vedere, entrambe le serie presentano un andamento costantemente crescente e le funzioni di autocorrelazione campionarie decrescono molto lentamente a zero, forse per la presenza di memoria lunga o, forse, per la presenza di *break* strutturali trascurati.



**Figura 1:** Usa, indice dei prezzi al consumo e funzione di autocorrelazione



**Figura 2:** Italia, indice dei prezzi al consumo e funzione di autocorrelazione

# Capitolo 4

## Analisi

### 4.1 Strategia empirica

Come discusso precedentemente, quando un processo a memoria corta è affetto da *break* occasionali, in alcuni casi la sua realizzazione potrebbe dare l'impressione di persistenza e la stima del parametro di lunga memoria  $d$  differire da zero. In questo paragrafo viene presentata la procedura empirica che useremo per verificare se l'impressione di memoria lunga è effettivamente dovuta a persistenza nei dati o se invece ciò è provocato da *break* strutturali trascurati.

Per fare questo seguiremo i seguenti passaggi:

- Stimare  $d$  nella serie  $x_t$  con i tre metodi, parametrico e semiparametrici descritti nei capitoli precedenti, ottenendo quindi  $\hat{d}$ ;
- Stimare le date dei *break* nella serie  $x_t$  con la procedura di Bai-Perron (BP);
- Ottenere la nuova serie priva dei *break* precedentemente trovati come  $x'_t = x_t - m_t$  dove  $m_t$  è la media del campione di ogni regime;
- Stimare  $d$  nella nuova serie  $x'_t$  ottenendo  $\hat{d}'$ .

Useremo i tre diversi metodi descritti nei capitoli precedenti per stimare il parametro di memoria lunga prima e dopo aver trovato i possibili *break* nella serie.

Se la stima di  $d$  differisce da zero questo significa che potremmo essere in presenza di un processo a memoria lunga. Se, quindi, stimato nuovamente il parametro  $d$  dopo aver depurato la serie degli eventuali *break*, questo

presenterà un valore attorno allo zero si può concludere che l'impressione di lunga memoria è dovuta al fatto di aver trascurato *break* occasionali.

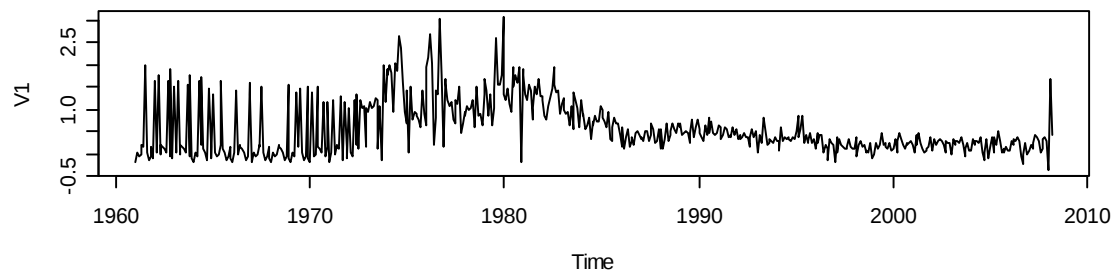
Successivamente l'analisi continuerà considerando le sottoserie ottenute dividendo le serie originali nei punti corrispondenti ai *break* trovati. Le nuove sottoserie verranno quindi considerate indipendentemente l'una dall'altra e per ciascuna verrà stimato, *ex novo*, il parametro  $d$  di memoria lunga per vedere con maggiore evidenza se queste presentano o no memoria lunga e quindi se la persistenza è reale o se frutto della mancata depurazione dei cambiamenti strutturali dalle serie.

## 4.2 Applicazione ai dati reali

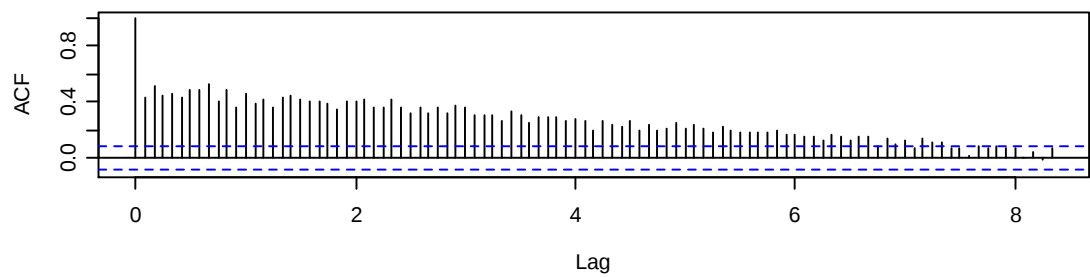
In questo paragrafo ci occupiamo di applicare la strategia empirica di cui sopra alle serie reali di dati relative all'indice dei prezzi al consumo di Italia e Stati Uniti.

Per quanto riguarda il software utilizzato, un'analisi preliminare è stata condotta con il programma Econometric Views (EViews), mentre la stima del parametro di memoria lunga e dei *break* presenti nelle serie sono stati calcolati tramite apposite routine compilate in linguaggio R, versione 2.1.1.

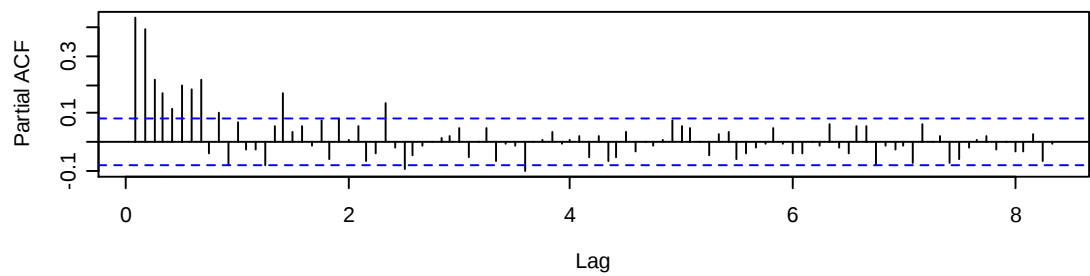
Le figure 3 e 4 riportano i grafici delle serie dell'inflazione statunitense e italiana rispettivamente, dove per serie dell'inflazione intendiamo  $100 * \log(C_t - C_{t-1})$ , con  $C_t$  la serie dell'Indice dei Prezzi al Consumo (Consumer Price Index).



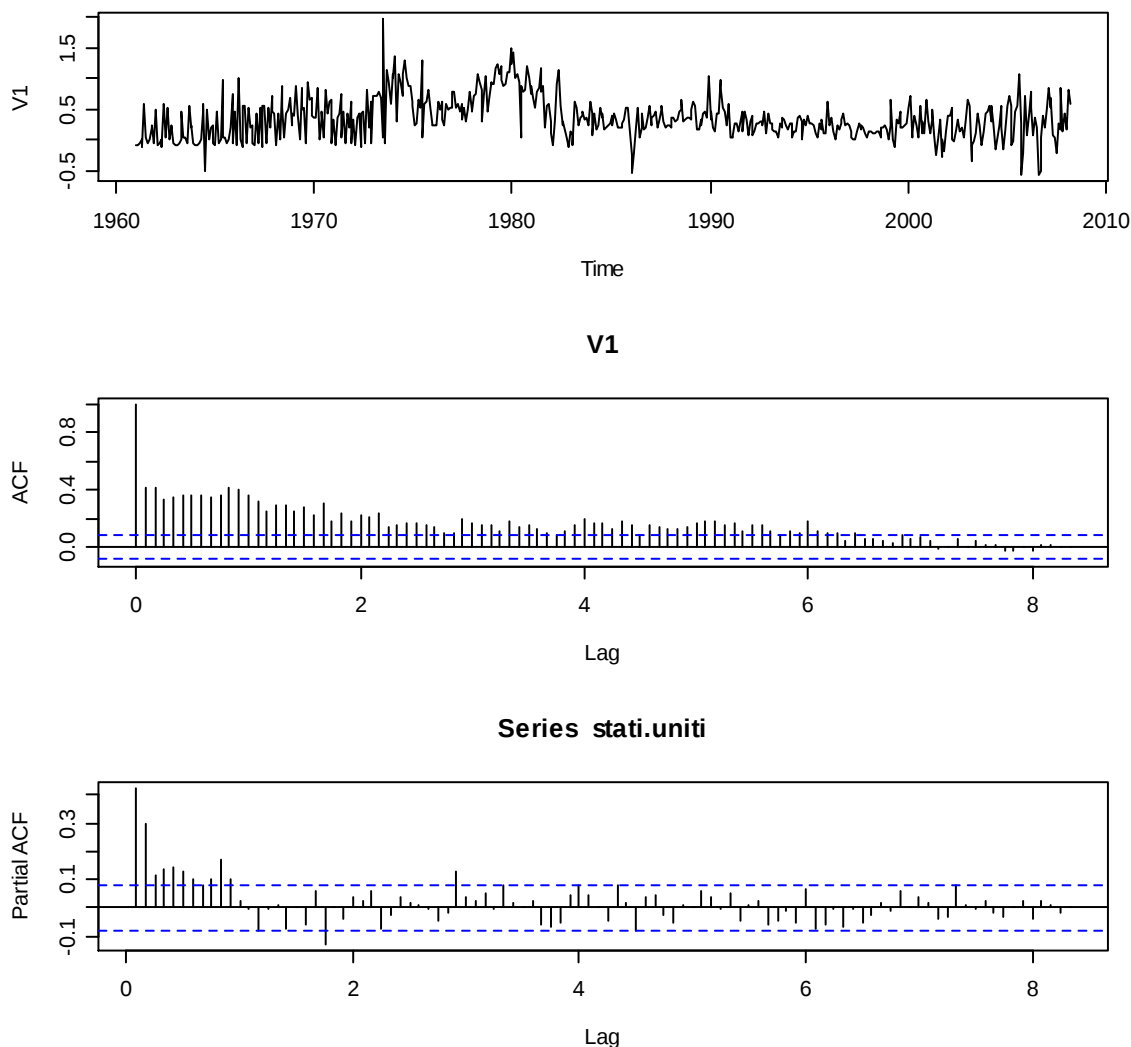
**V1**



**Series italy**



**Figura 3:** Italia, serie dell'inflazione



**Figura 4:** USA, serie dell'inflazione

I grafici evidenziano come, in entrambi i casi, la funzione di correlazione decresce molto lentamente a zero. Sembrano, inoltre, esserci delle variazioni nei livelli delle due serie.

Procediamo dapprima, con le stime del parametro di memoria lunga nelle serie dell'inflazione. Le tabelle 1 e 2 riportano i risultati delle stime rispettivamente di Italia e Stati Uniti.

Per lo stimatore GPH vengono riportate due stime; la prima derivata con il metodo GPH esposto nel capitolo 1; la seconda derivata, invece, con lo

stimatore GPH ripreso e sviluppato da Robinson, al quale si è accennato sempre nel capitolo 1.

| Stimatori: | Whittle | GPH                           | Local Whittle |
|------------|---------|-------------------------------|---------------|
| $\hat{d}$  | 0.2813  | 0.7250<br>0.2362 <sup>5</sup> | 0.7474        |

**Tabella 1:**

Italia, stima del parametro di lunga memoria  $\hat{d}$  sull'intero campione

| Stimatori: | Whittle | GPH                           | Local Whittle |
|------------|---------|-------------------------------|---------------|
| $\hat{d}$  | 0.2888  | 1.0643<br>1.5259 <sup>5</sup> | 0.8762        |

**Tabella 2:**

USA, stima del parametro di lunga memoria  $\hat{d}$  sull'intero campione

Come si può vedere chiaramente dalle due tabelle i valori stimati si discostano significativamente da zero. Specialmente lo stimatore *GPH* mostra un valore elevato di  $\hat{d}$ , nel caso USA anche leggermente maggiore di uno.

La strategia quindi procede stimando, attraverso il metodo di BP, gli eventuali *break* nelle serie.

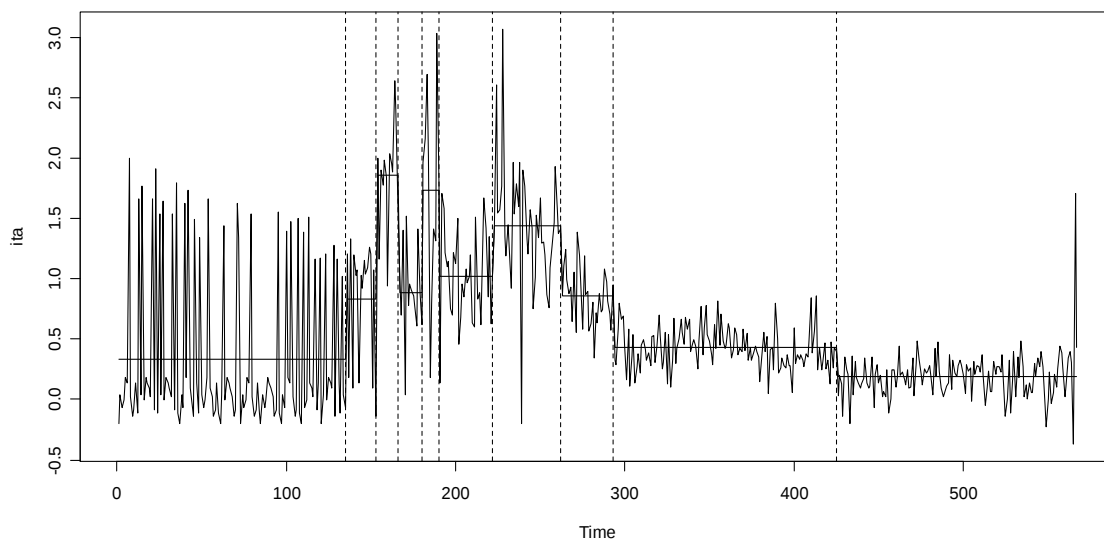
La lunghezza minima stabilita tra due *break* è di 12 osservazioni circa (un anno).

Analizziamo le due serie separatamente cominciando dapprima con quella italiana.

---

<sup>5</sup> Nell'analisi è stato considerato lo stimatore GPH ripreso e sviluppato da Robinson, al quale è stato accennato nel capitolo 1.

In figura 5 viene riportato il grafico della serie dell'inflazione con i *break* e le medie stimate per ogni periodo.



**Figura 5:** Italia, serie dell'inflazione e *break* stimati

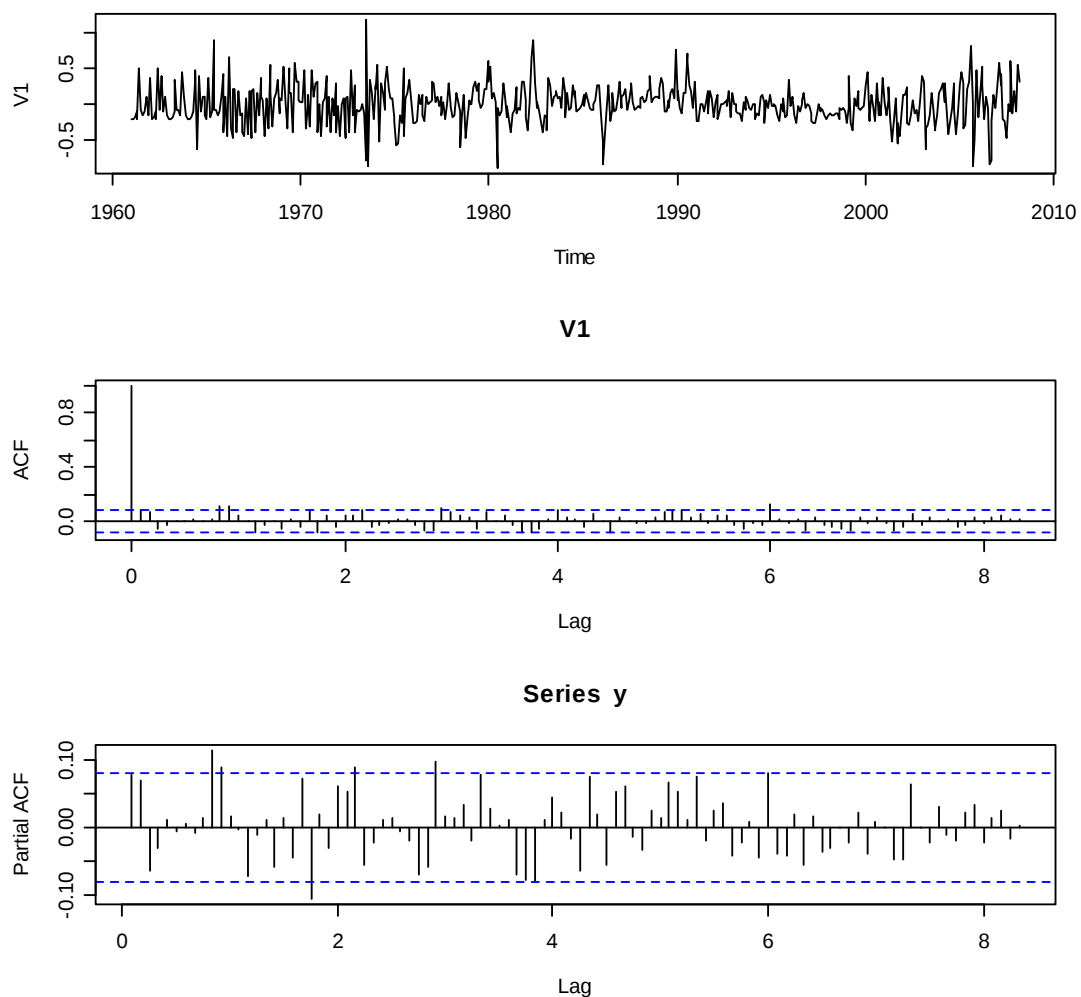
La serie presenta nove *break* rilevati sull'intero campione. La tabella 3 riporta le date in cui i *break* si verificano.

| Serie  | Numero di <i>break</i> | Date dei <i>break</i> |
|--------|------------------------|-----------------------|
| Italia | 9                      | 1972(03)              |
|        |                        | 1973(09)              |
|        |                        | 1974(10)              |
|        |                        | 1975(12)              |
|        |                        | 1976(10)              |
|        |                        | 1979(06)              |
|        |                        | 1982(10)              |
|        |                        | 1985(05)              |
|        |                        | 1996(05)              |

**Tabella 3:**  
Italia, numero di *break* e relative date

Come si nota, la maggior parte dei *break* si verificano nel corso degli anni '70, infatti, come accennato nel capitolo precedente, questo periodo è caratterizzato da forti fluttuazioni, conseguenza del forte aumento del prezzo del petrolio registrato in quel particolare decennio a fronte degli shock petroliferi verificatisi e conseguenza di frequenti deprezzamenti della moneta.

Otteniamo ora la nuova serie priva dei *break* precedentemente stimati come  $x'_t = x_t - m_t$  dove  $m_t$  è la media del campione per ogni regime, vi riportiamo il grafico della serie depurata dai *break*, e delle rispettive funzioni di autocorrelazione, in figura 6.



**Figura 6:** Italia, serie dell'inflazione depurata dai *break* strutturali

Il grafico presenta una media che è più costante nel tempo. Anche la funzione di autocorrelazione empirica ha un andamento differente, infatti, decresce velocemente a zero, come per un processo a memoria corta. Questo ci porta a pensare che, come sospettato, l'apparenza di memoria lunga sia causata dall'aver trascurato i *break*.

Andiamo a verificare ciò stimando nuovamente il parametro di memoria lunga nella serie attraverso i tre metodi descritti nei capitoli precedenti.

|            | Whittle | GPH                            | Local Whittle |
|------------|---------|--------------------------------|---------------|
| $\hat{d}$  | 0.2814  | 0.7250<br>0.2362 <sup>6</sup>  | 0.7474        |
| $\hat{d}'$ | 0.000   | -0.1821<br>0.0472 <sup>6</sup> | 0.000         |

**Tabella 4:**

Italia, tre stime di  $\hat{d}$  nella serie originale e in quella depurata dai *break*

La tabella 4 riporta un confronto con la stima di  $d$  effettuata sulla serie originale.

Le nuove stime sono molto vicine allo zero. Questo vuol dire che, dopo aver rimosso i punti di *break* dalla serie, il valore della stima del parametro di memoria lunga si muove verso lo zero, qualsiasi metodo di stima venga usato. In particolare con il metodo *Whittle* e il metodo *local Whittle*, il valore della stima è praticamente indistinguibile da zero. Per quanto riguarda il metodo GPH, il valore stimato di  $\hat{d}'$  risulta essere negativo, ciò caratterizza una memoria di tipo antipersistente. Questo ci suggerisce, come detto in precedenza, che probabilmente la persistenza forte dei dati è dovuta al fatto di aver trascurato *break* occasionali.

---

<sup>6</sup> Nell'analisi è stato considerato lo stimatore GPH ripreso e sviluppato da Robinson, al quale si è accennato nel capitolo 1.

La conclusione potrebbe essere che la serie dell'inflazione riguardante l'Italia è caratterizzata da *break* strutturali e non da memoria lunga. Ma per essere veramente certi di ciò andiamo ad analizzare le sottoserie ottenute dividendo la serie originale nei punti corrispondenti ai *break* stimati. Le sottoserie vengono quindi considerate indipendentemente. Per ciascuna di esse viene stimato, attraverso i tre metodi il parametro  $\hat{d}$ .

| Sottoserie     | Date              | Whittle<br>$\hat{d}''$ | Local Whittle<br>$\hat{d}''$ | GPH<br>$\hat{d}''$ |
|----------------|-------------------|------------------------|------------------------------|--------------------|
| <b>1:134</b>   | 1961:01 – 1972:02 | 0.000                  | 0.5344                       | 0.8619             |
| <b>135:152</b> | 1972:03 – 1973:08 | 0.000                  | 0.000                        | 0.0134             |
| <b>153:165</b> | 1973:09 – 1974:09 | 0.000                  | 0.4878                       | 1.6084             |
| <b>166:179</b> | 1974:10 – 1975:11 | 0.000                  | 0.000                        | -2.0338            |
| <b>180:189</b> | 1975:12 – 1976:09 | 0.000                  | 0.6440                       | 0.7453             |
| <b>190:221</b> | 1976:10 – 1979:05 | 0.000                  | 0.000                        | -0.6486            |
| <b>222:261</b> | 1979:06 – 1982:09 | 0.031                  | 0.1541                       | 0.3019             |
| <b>262:292</b> | 1982:10 – 1985:04 | 0.068                  | 0.2513                       | 0.2521             |
| <b>293:424</b> | 1985:05 – 1996:04 | 0.102                  | 0.2252                       | 0.2123             |
| <b>425:567</b> | 1996:05 – 2008:04 | 0.000                  | 0.2443                       | 0.0434             |

**Tabella 5:**  
Italia, stima di  $d$  nelle sottoserie

Nella tabella 5, tutte le stime, ricavate con il metodo Whittle, si avvicinano a zero, eccezion fatta per la stima relativa al periodo maggio 1985 – aprile 1996 che restituisce un valore di 0.102. E' necessario far presente che la letteratura, riguardante i metodi di stima del parametro di memoria lunga, dimostra che il metodo *Whittle* rivela una performance leggermente migliore rispetto ad altri stimatori (Bisaglia e Gerolimetto (2007a,b)) e che, in generale, la performance di tutti i metodi diviene peggiore al crescere del

numero di osservazioni incluse nel campione. Questo dipende probabilmente dal fatto che più lunga è la serie e più evidente diviene il comportamento non stazionario così che il livello di somiglianza tra processi con *break* occasionali e processi a memoria lunga cresce.

La colonna, in tabella 5, relativa allo stimatore *local Whittle*, invece, presenta delle stime che si discostano quasi sempre da zero, addirittura alcune di esse in modo significativo; come nel caso della stima relativa alla sottoserie che va da dicembre 1975 a settembre 1976. Il numero delle osservazioni appartenenti a questa sottoserie è, comunque, troppo piccolo affinché la stima possa essere statisticamente significativa.

Per quanto riguarda, infine, lo stimatore GPH (riportato nell'ultima colonna della tabella 5), nella tabella emerge un valore della stima anomalo, pari a -2.0338, relativamente al periodo 1974:10 – 1975:11. Questo non deve stupire in quanto la sottoserie relativa presenta solo 13 osservazioni, un numero evidentemente piccolo per ottenere risultati statisticamente affidabili. Lo stesso ragionamento viene fatto per quanto riguarda il valore 1.6084 restituito da una sottoserie (1973:09 – 1974:09) di sole 12 osservazioni.

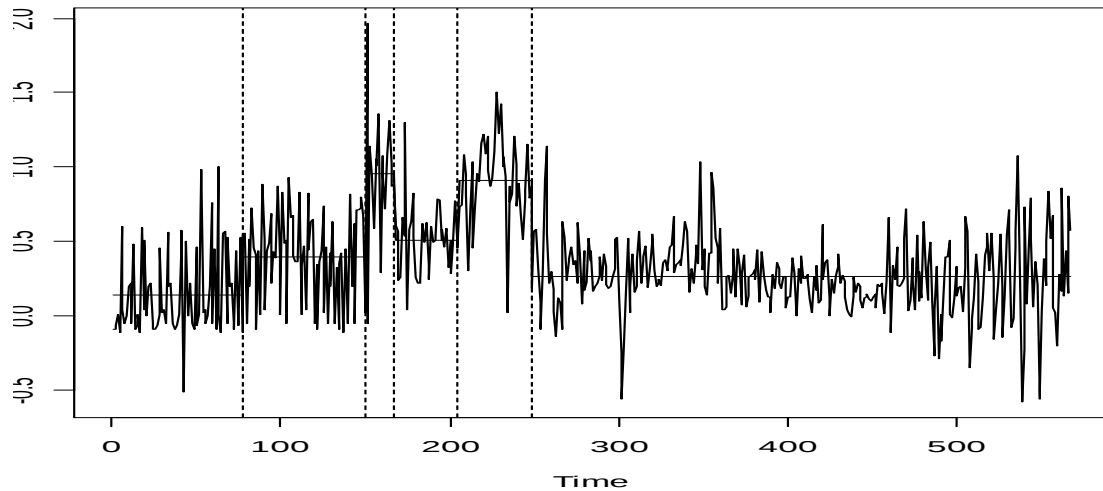
Il parametro  $d$ , ricordiamo, rivela l'intensità della dipendenza che, in caso di stime negative, come appaiono in tabella, è addirittura antipersistente.

A questo punto una conclusione può essere tratta a proposito dell'andamento dell'inflazione italiana.

Dipendenza di memoria lunga e cambiamenti strutturali sono fenomeni strettamente connessi e difficilmente distinguibili. Attraverso la procedura empirica presentata abbiamo distinto i due fenomeni e trovato, in particolare, che l'inflazione italiana sembra essere caratterizzata da *break* strutturali piuttosto che da memoria lunga.

Ripetiamo ora la medesima analisi per la serie dell'inflazione di Stati Uniti d'America.

Per prima cosa, in figura 7, riportiamo il grafico della serie dell'inflazione nel quale sono stati evidenziati i *break* trovati e le relative medie per ogni sottoperiodo.



**Figura 7:** USA, serie dell'inflazione con i rispettivi *break* stimati

Nel caso americano il numero dei *break* stimati è relativamente piccolo. Si tratta di 5 *break* per l'esattezza. La tabella 6 elenca le date in corrispondenza dei *break*.

| Serie | Numero dei <i>break</i> | Date dei <i>break</i>                                    |
|-------|-------------------------|----------------------------------------------------------|
| USA   | 5                       | 1967(05)<br>1973(06)<br>1974(11)<br>1978(01)<br>1981(08) |

**Tabella 6:**  
USA, numero di *break* e relative date

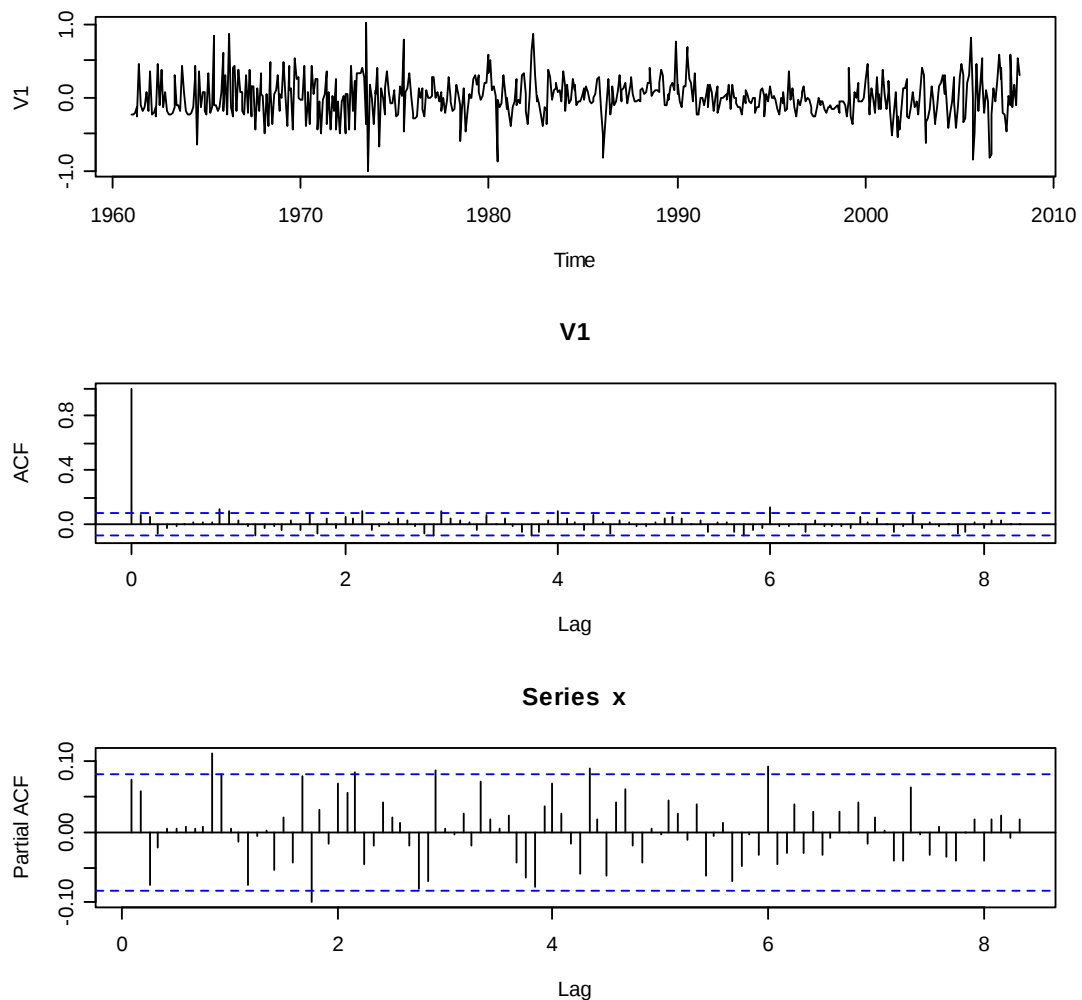
Andiamo a ricercare una spiegazione storica a supporto di questi *break*.

Il boom economico degli Anni '60 fu seguito da un periodo di instabilità politica ed economica. Nel 1971, gli Stati Uniti, in seguito allo stress economico causato dall'impegno in Vietnam, dichiararono la fine della 'convertibilità in oro del dollaro', pilastro dell'economia capitalista, dando inizio ad una lunga fase di instabilità e disordine monetario mondiale. Questo fatto, aggiunto alla decisione dei Paesi produttori di petrolio di quadruplicare il costo di questa materia prima per riprendersi dalla guerra arabo-israeliana, causò nel 1973 il cosiddetto 'shock petrolifero', che colpì pesantemente tutti i Paesi industrializzati, in particolare Italia e Giappone. La conseguente crisi delle industrie portò ad un brusco calo dei prezzi e quindi ad una forte inflazione.

Interessante notare che uno dei cinque *break*, corrispondente al 1973, è simile a quello trovato da Bos, Franses and Ooms (2001), per il quale essi fornirono una spiegazione solo dal punto di vista storico. La nostra analisi rileva poi un regime addizionale tra i due shock petroliferi del 1974 e del 1978.

Otteniamo ora la nuova serie priva dei *break* precedentemente stimati come  $x'_t = x_t - m_t$  dove  $m_t$  è la media del campione per ogni regime.

La figura 8 riporta il grafico della serie depurata dai *break* e delle rispettive funzioni di autocorrelazione.



**Figura 8:** USA, serie dell'inflazione depurato dei *break* strutturali.

Il grafico presenta una media che sembra essere costante nel tempo. La funzione di autocorrelazione empirica decresce velocemente a zero, nel modo tipico di un processo a memoria corta. Tutto ciò ci porta a pensare che la serie presenti *break* strutturali e non quindi memoria lunga.

Andiamo a verificare ciò analizzando la nuova serie e stimando quindi nuovamente il parametro di memoria lunga nella serie attraverso tre metodi: *Whittle*, *local Whittle* e *GPH*. Per il metodo *GPH*, come precedentemente nel caso italiano, vengono riportate due stime. La seconda è relativa allo

stimatore GPH ripreso e sviluppato da Robinson a cui si è accennato nel capitolo 1.

|            | Whittle | GPH                            | Local Whittle |
|------------|---------|--------------------------------|---------------|
| $\hat{d}$  | 0.2888  | 1.0643<br>1.5259 <sup>7</sup>  | 0.8762        |
| $\hat{d}'$ | 0.0516  | 0.0637<br>-0.1355 <sup>7</sup> | 0.0788        |

**Tabella 7:**

USA, tre stime di  $\hat{d}$  nella serie originale e in quella depurata dai *break*

La tabella 7 riporta, i valori di  $\hat{d}$  prima e dopo il calcolo dei *break*.

In tutti e tre i metodi, il valore della stima del parametro di memoria lunga si muove verso lo zero. Questo ci suggerisce, come detto in precedenza, che la persistenza forte dei dati potrebbe essere dovuta al fatto di aver trascurato dei *break* occasionali nella serie.

Nuovamente il metodo Whittle restituisce una stima di  $d$  più vicina allo zero rispetto agli altri due metodi. In particolare con il metodo GPH ripreso e sviluppato da Robinson risulta un  $\hat{d}$  negativo.

La conclusione a cui si giunge è, perciò, che la serie dell'inflazione degli Stati Uniti è caratterizzata da *break* strutturali e non da memoria lunga. Ma per essere veramente certi di ciò andiamo ad analizzare le sottoserie ottenute dividendo la serie originale nei punti corrispondenti ai *break* stimati. Le sottoserie vengono quindi considerate indipendentemente. Per ciascuna di esse viene stimato, attraverso il metodo *Whittle*, *local Whittle* e GPH il parametro  $\hat{d}$ .

---

<sup>7</sup> Nell'analisi è stato considerato lo stimatore GPH ripreso e sviluppato da Robinson, al quale si è accennato nel capitolo 1.

| Sottoserie     | Date              | Whittle<br>$\hat{d}''$ | Local Whittle<br>$\hat{d}''$ | GPH<br>$\hat{d}''$ |
|----------------|-------------------|------------------------|------------------------------|--------------------|
| <b>1:76</b>    | 1961:01 – 1967:04 | 0.000                  | 0.1074                       | 0.1887             |
| <b>77:149</b>  | 1967:05 – 1973:05 | 0.000                  | 0.3107                       | 0.3850             |
| <b>150:166</b> | 1973:06 – 1974:10 | 0.000                  | 0.000                        | -0.2738            |
| <b>167:204</b> | 1974:11 – 1977:12 | 0.000                  | 0.000                        | 0.3562             |
| <b>205:247</b> | 1978:01 – 1981:07 | 0.2959                 | 0.2251                       | 0.2236             |
| <b>248:567</b> | 1981:08 – 2008:04 | 0.2616                 | 0.2291                       | 0.3550             |

**Tabella 8:**  
USA, stima di  $d$  nelle sottoserie

Come nell'analisi precedente, anche in questo caso *Whittle* (terza colonna della tabella 8) fornisce, per quasi tutte le sottoserie, una stima di  $\hat{d}$  vicina allo zero. Le ultime due stime, invece, differiscono da zero, riportando a memoria lunga, ma non c'è da stupirsi, in quanto è possibile che i due fenomeni coesistano. In particolare, per quanto riguarda l'ultima sottoserie, che presenta un valore dalla stima di 0.2616 relativa al periodo 1981:08 – 2008:04, questa copre un periodo di circa 27 anni, pari a ben 319 osservazioni, ed è quindi possibile aspettarsi memoria lunga.

Guardando alla quarta colonna della tabella 8, lo stimatore local Whittle per due delle sottoserie restituisce una stima praticamente indistinguibile da zero, mentre per le altre è comunque non elevatissima. Come riscontrato prima, potrebbe questo essere il caso in cui i due fenomeni, memoria lunga e *break* strutturali, coesistano, cioè, in cui entrambi questi fenomeni si presentino nella stessa serie.

In fine, l'ultimo stimatore, il GPH (colonna 5, tabella 8) riporta stime che si discostano tutte da 0. In particolare, una di queste (periodo 1973:06 –

1974:10) presenta dipendenza antipersistente, ma è relativa ad una sottoserie di sole 16 osservazioni, statisticamente poco numerosa.

A questo punto può essere data una conclusione a proposito dell'andamento dell'inflazione degli Stati Uniti.

Dipendenza di memoria lunga e *break* strutturali, come già ripetuto in questo lavoro, sono fenomeni strettamente connessi e difficilmente distinguibili. Per la serie dell'inflazione statunitense, dopo aver eseguito la procedura empirica, descritta all'inizio di questo capitolo, si giunge a delle conclusioni diverse per quanto concerne la serie dell'inflazione italiana. La serie americana sembra, infatti, essere caratterizzata da entrambi i fenomeni, memoria lunga e *break* strutturali.

Nel prossimo paragrafo viene esposta, in breve, l'analisi che è stata ricondotta, nel medesimo modo, su un sottocampione delle serie originali.

#### 4.3 Analisi di un sottocampione delle serie originali

Questo paragrafo ha lo scopo di condurre una breve analisi di una parte dei dati relativi alle due serie del *Consumer Price Index* di Italia e Stati Uniti.

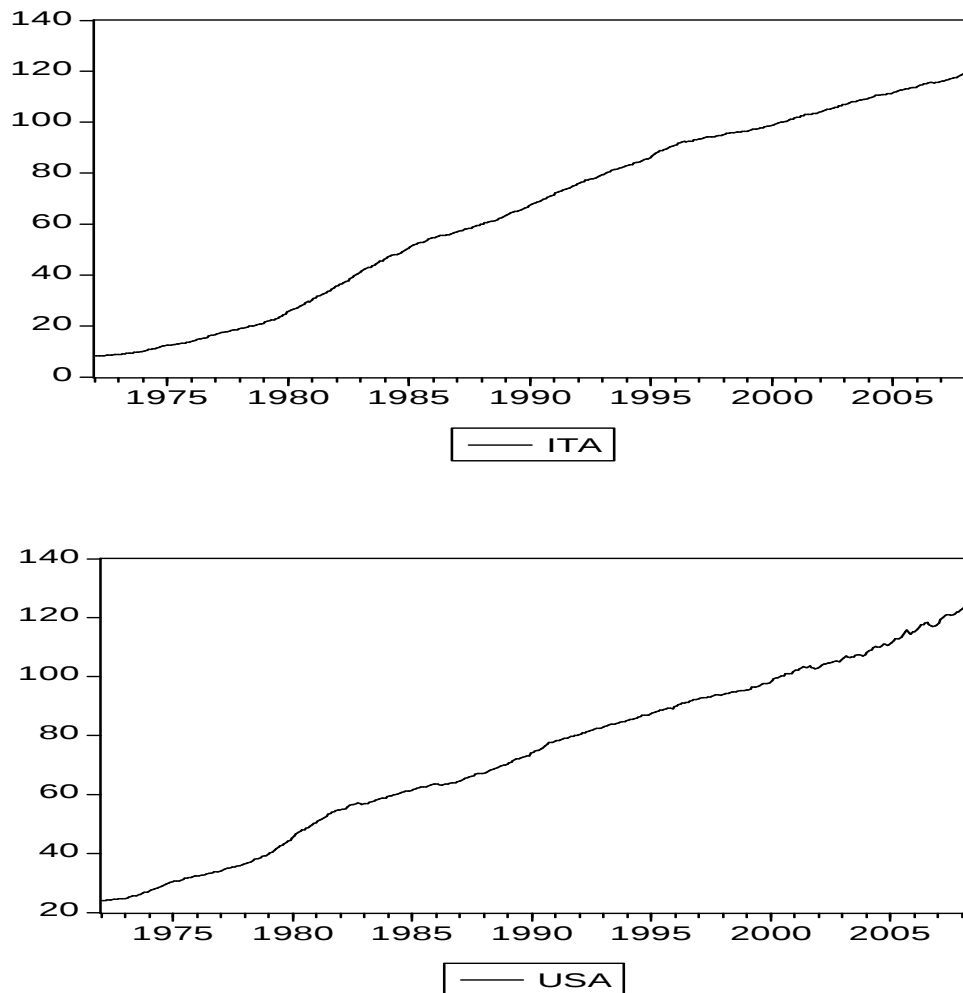
I dati, così come sono stati considerati precedentemente, mostrano un comportamento anomalo nella prima parte della serie. In particolare per quanto riguarda le osservazioni che vanno dal 1961 al 1972.

E' probabile che, durante tutto questo decennio, le osservazioni siano state rilevate a cadenza trimestrale anziché a cadenza mensile, e che quindi, in un secondo momento questi dati siano stati trasformati in modo tale da poter essere affiancati a dati mensili.

Il nostro "nuovo" campione copre un periodo di tempo che va da gennaio 1972 ad aprile 2008 (anno base 2000) con un totale di 435 osservazioni.

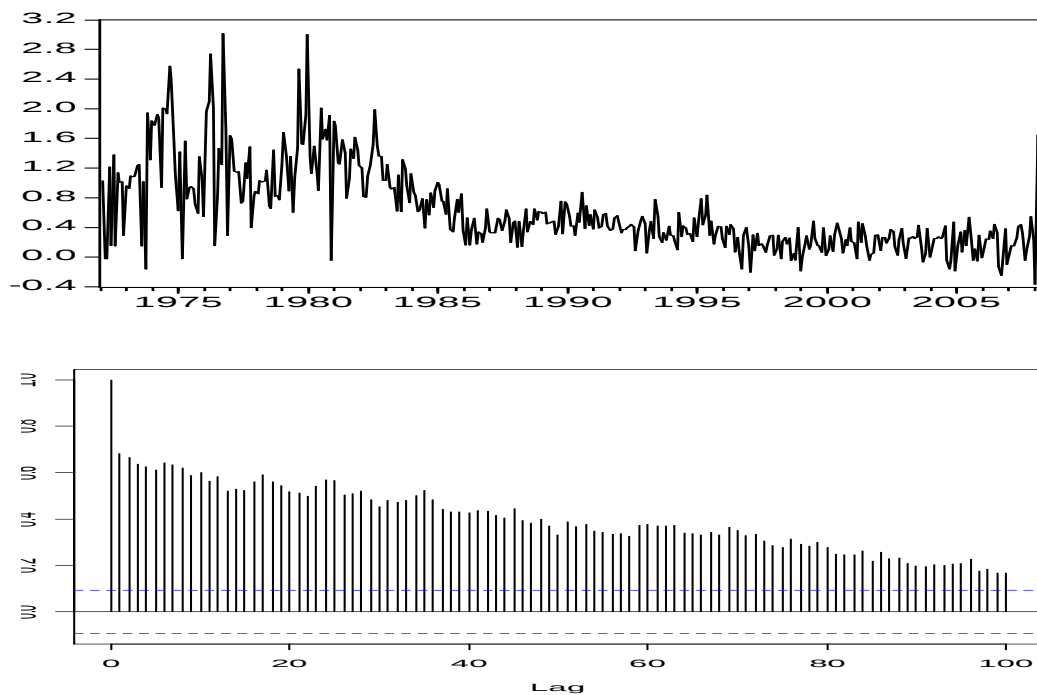
La strategia empirica applicata a queste nuove serie è la medesima considerata nel paragrafo precedente.

Le due serie si presentano come in figura 9:

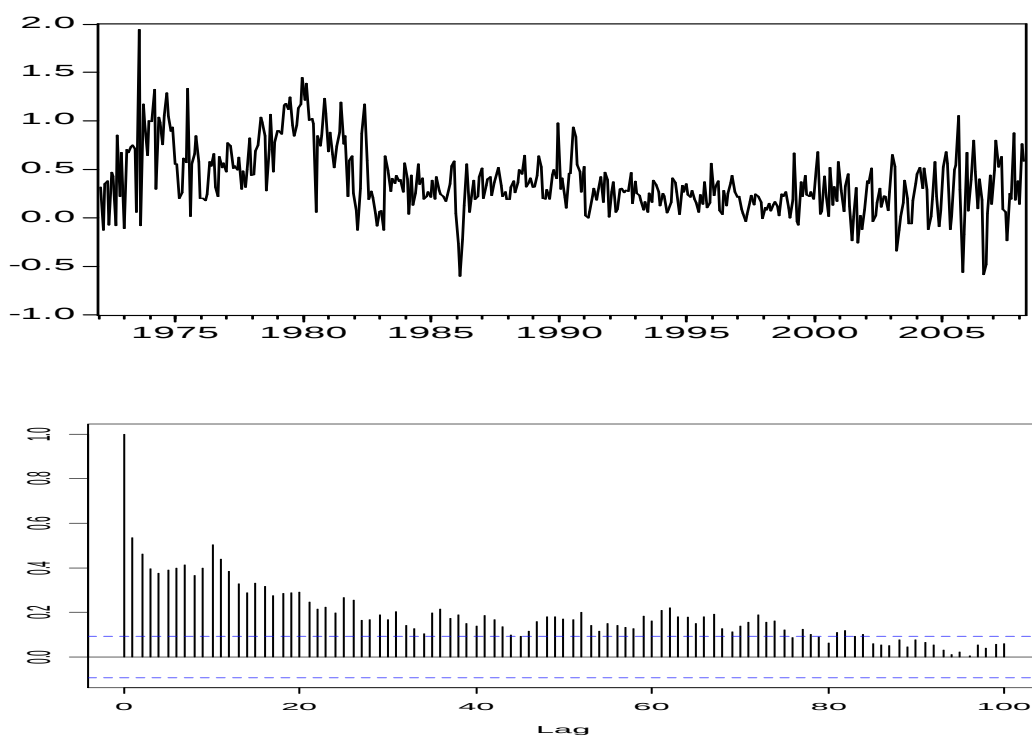


**Figura 9:** Italia e USA, indice dei prezzi al consumo

Le figure 10 e 11 riportano i grafici delle serie dell'inflazione italiana e statunitense rispettivamente, dove per serie dell'inflazione intendiamo  $100 \cdot \log(C_t - C_{t-1})$ , con  $C_t$  la serie dell'Indice dei Prezzi al Consumo. Le serie sono state, come nel paragrafo precedente, destagionalizzate usando il metodo delle medie mobili.



**Figura 10:** Italia, serie dell'inflazione



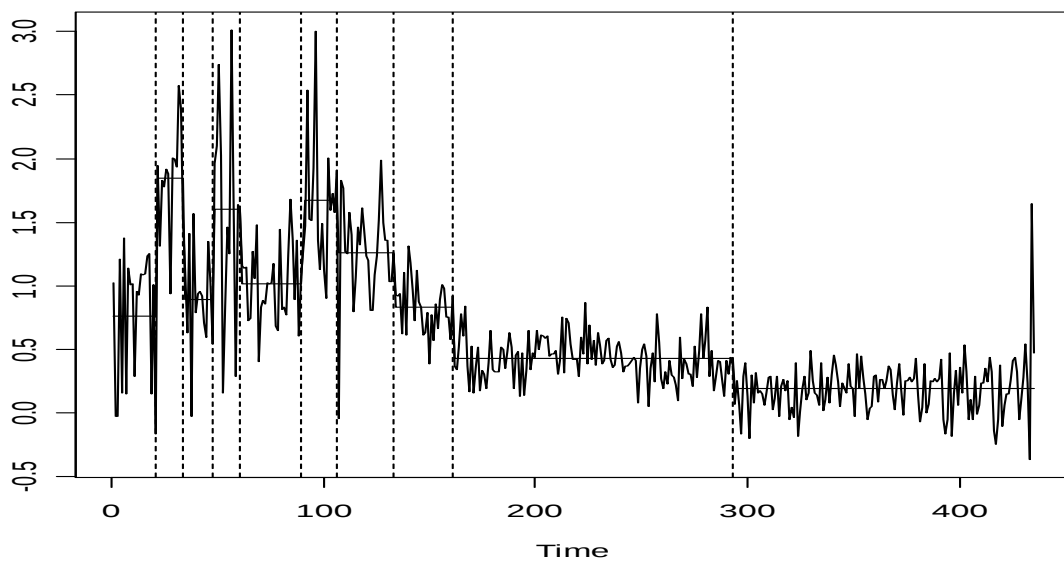
**Figura 11:** USA, serie dell'inflazione

Le funzioni di autocorrelazioni decadono lentamente a zero.

Il nostro compito è quello di verificare se questo comportamento è frutto di processi caratterizzati da memoria lunga o se, invece, si tratta di un fenomeno caratterizzato da *break* strutturali.

Le figure 12 e 13 presentano i *break* ottenuti con la procedura BP.

Le tabelle 9 e 10 indicano le date in corrispondenza delle quali si sono verificati i *break*.

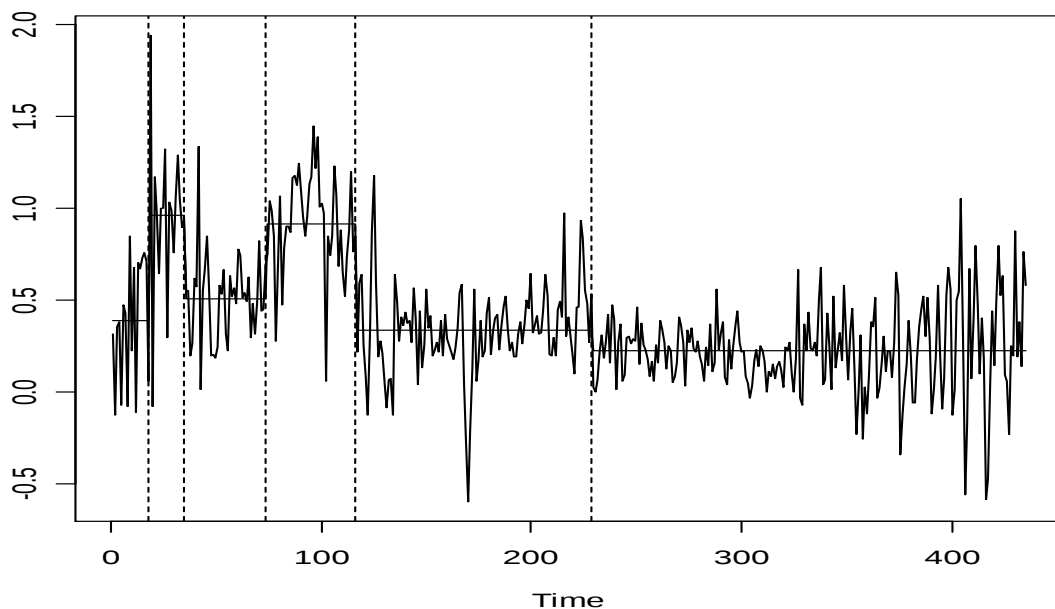


**Figura 12:** Italia, serie dell'inflazione con i rispettivi *break* stimati

| Serie                                | Numero dei <i>break</i> | Date dei <i>break</i>                                                                                    |
|--------------------------------------|-------------------------|----------------------------------------------------------------------------------------------------------|
| Indice dei prezzi al consumo, Italia | 9                       | 1973(09)<br>1974(10)<br>1975(12)<br>1977(01)<br>1979(06)<br>1980(10)<br>1983(01)<br>1985(05)<br>1996(05) |

**Tabella 9:**  
Italia, numero dei *break* e relative date

Le date in corrispondenza dei *break* sono 9, come nel caso della serie originale e non variano significativamente.



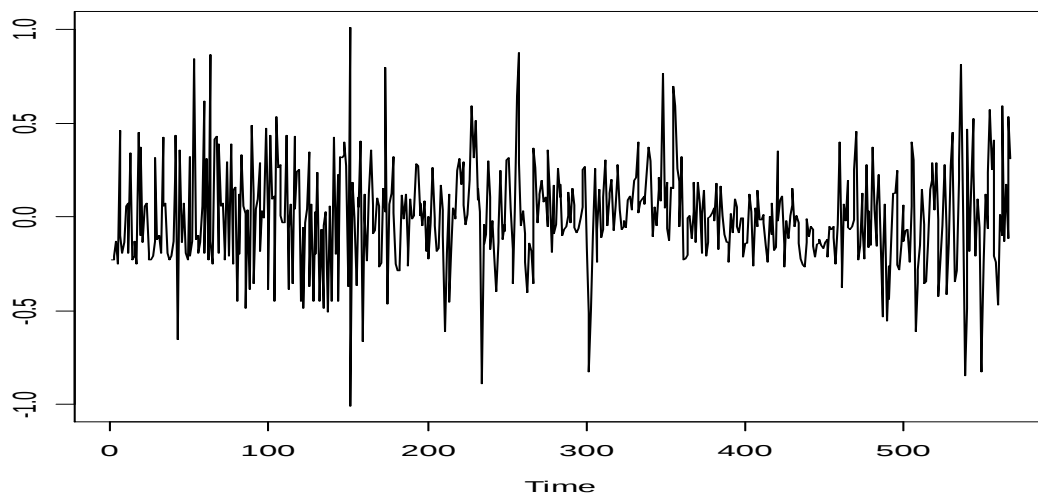
**Figura 13:** USA, serie dell'inflazione con i rispettivi *break* stimati

| Serie                             | Numero dei <i>break</i> | Date dei <i>break</i>                                    |
|-----------------------------------|-------------------------|----------------------------------------------------------|
| Indice dei prezzi al consumo, USA | 5                       | 1973(06)<br>1974(11)<br>1978(02)<br>1981(08)<br>1990(12) |

**Tabella 10:**  
USA, numero dei *break* e relative date

Anche nel caso della serie americana il numero di date e il loro avvenimento è uguale al caso della serie originaria.

Otteniamo il grafico, in figura 14 e 15, delle serie di Italia e Stati Uniti rispettivamente, prive dei *break* precedentemente trovati e stimiamo il parametro di memoria lunga (tabella 11 e 12) nell'intero campione prima e dopo la stima dei punti di *break*.

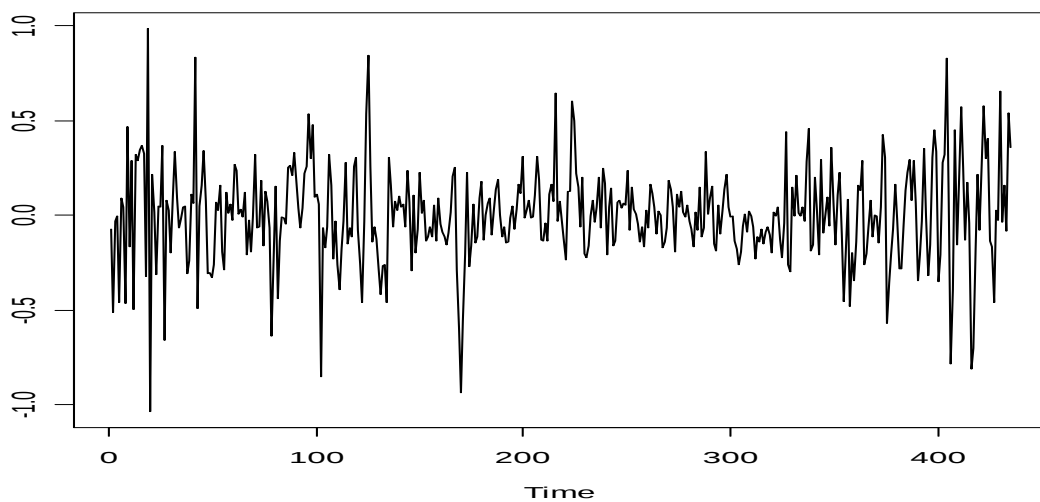


**Figura 14:** Italia, serie dell'inflazione depurato dei *break* stimati

|            | Whittle | GPH     | Local Whittle |
|------------|---------|---------|---------------|
| $\hat{d}$  | 0.4018  | 0.6439  | 0.7388        |
| $\hat{d}'$ | 0.000   | -0.0245 | 0.000         |

**Tabella 11:**

Italia, tre stime di  $\hat{d}$  nella serie originale e in quella depurate dei *break*



**Figura 15:** USA, Consumer Price Index depurato dei *break* stimati

|            | Whittle | GPH     | Local Whittle |
|------------|---------|---------|---------------|
| $\hat{d}$  | 0.3529  | 1.0853  | 0.9108        |
| $\hat{d}'$ | 0.0947  | -0.2050 | 0.000         |

**Tabella 12:**

Italia, tre stime di  $\hat{d}$  nella serie originale e in quella depurate dei *break*

Anche questa volta le stime, nonostante differiscano, portano alle stesse conclusioni. In entrambe le serie sembra evidente che l'apparenza di memoria lunga sia causata dalla presenza di *break* trascurati nella serie.

Nell'ultima parte della nostra analisi vengono prese in considerazione le sottoserie ottenute nei punti di “rottura” delle serie originali, affinché possiamo capire, con maggiore evidenza, se le conclusioni date nel paragrafo precedente sono le medesime.

| Sottoserie     | Date               | Whittle<br>$\hat{d}''$ | Local Whittle<br>$\hat{d}''$ | GPH<br>$\hat{d}''$ |
|----------------|--------------------|------------------------|------------------------------|--------------------|
| <b>1:20</b>    | 1972:01 - 1973:08  | 0.000                  | 0.8282                       | 1.061              |
| <b>21:33</b>   | 1973:09 - 1974:09  | 0.000                  | 0.000                        | 0.0160             |
| <b>34:47</b>   | 1974:10 - 1975:11  | 0.000                  | 0.3771                       | 0.7283             |
| <b>48:60</b>   | 1975:12 - 1976:09  | 0.000                  | 0.000                        | -0.4944            |
| <b>61:89</b>   | 1977:01 - 1979:05  | 0.000                  | 0.000                        | -0.1430            |
| <b>90:105</b>  | 1979:06 – 1980:09  | 0.000                  | 0.000                        | -0.0807            |
| <b>106:132</b> | 1980:10- 1982:09   | 0.000                  | 0.000                        | -0.1966            |
| <b>133:160</b> | 1983:01 – 1985:04  | 0.0622                 | 0.3144                       | 0.3856             |
| <b>161:292</b> | 1985:05 – 1996: 04 | 0.1374                 | 0.1955                       | 0.2547             |
| <b>293:435</b> | 1996:05 – 2008:04  | 0.000                  | 0.2917                       | -0.0423            |

**Tabella 13:**

Italia, stima di  $\hat{d}''$  nelle sott'orie con i tre metodi: *Whittle*, *local Whittle*, GPH

| Sottoserie     | Date              | Whittle<br>$\hat{d}''$ | Local Whittle<br>$\hat{d}''$ | GPH<br>$\hat{d}''$ |
|----------------|-------------------|------------------------|------------------------------|--------------------|
| <b>1:17</b>    | 1972:01 – 1973:05 | 0.000                  | 0.8884                       | 1.2621             |
| <b>18:34</b>   | 1973:06 – 1974:10 | 0.000                  | 0.000                        | -0.1305            |
| <b>35:73</b>   | 1974:11 – 1978:01 | 0.000                  | 0.000                        | 0.0222             |
| <b>74:115</b>  | 1978:02 – 1981:07 | 0.2886                 | 0.2529                       | 0.3723             |
| <b>116:227</b> | 1981:08 – 1990:11 | 0.3658                 | 0.1605                       | 0.2155             |
| <b>228:435</b> | 1990:12 – 2008:04 | 0.1263                 | 0.2752                       | 0.3067             |

**Tabella 14:**

USA, stima di  $\hat{d}''$  nelle sott'orerie con i tre metodi: *Whittle*, *local Whittle*, GPH

Le serie italiana sembra caratterizzata da *break* strutturali, come precedentemente concluso. Nella serie americana i due fenomeni, memoria lunga e *break* strutturali, coesistono, come riscontrato nell'analisi nel paragrafo precedente.

In conclusione possiamo dire che, nel caso della serie considerata nel suo intero e nel caso della serie considerata a partire dal 1972, i risultati portano alle medesime conclusioni, e che quindi non c'è differenza nel considerare l'una o l'altra serie.

# Capitolo 5

## Conclusioni

La definizione dei processi a memoria lunga e *break* occasionali è stato il punto di partenza di questa tesi. Il passo successivo è stato cercare di distinguere la presenza di questi fenomeni in due serie storiche: l'inflazione italiana e statunitense. L'idea di realizzare una ricerca su questo argomento è stata suggerita da studi precedenti, condotti su campioni reali, da altri ricercatori.

Inizialmente abbiamo trattato i processi a memoria lunga, soffermandoci in particolare sui processi ARIMA frazionari e tre metodi (Whittle, local Whittle e GPH), per stimare il parametro di memoria lunga.

Nel capitolo 2, sono stati introdotti i *break* strutturali e il metodo Bai- Perron per stimarli.

Nel capitolo 3 è stato esaminato il tasso d'inflazione e il suo comportamento nelle serie dell'inflazione di Italia e Stati Uniti.

Infine, nel capitolo 4, sono state prese in considerazione le due serie reali dell'inflazione di Italia e Stati Uniti, sulle quali è stata applicata una procedura empirica per verificare se l'impressione di memoria lunga, mostrata dalle loro funzioni di autocorrelazione che decrescevano lentamente a zero all'aumentare del numero di ritardi, è effettivamente dovuto a persistenza nei dati o se, invece, ciò è provocato da *break* strutturali trascurati.

Le conclusioni, a cui siamo giunti nel caso italiano, sono che la serie presenta memoria corta ed è caratterizzata da *break* strutturali, quest'ultimi causa dell'iniziale impressione di memoria lunga. L'analisi dell'inflazione statunitense, invece, ci ha portato a conclusioni differenti, in quanto nelle serie i due fenomeni coesistono. La verifica è stata resa possibile analizzando le sottoserie ottenute dividendo la serie di partenza nei punti in cui abbiamo riscontrato dei *break*.

Un campo che può ancora riservare sviluppi, e che sarebbe interessante indagare, è lo sfruttamento di modelli che includano processi con *break* strutturali e processi FI(d) in un unico algoritmo. Questi modelli potrebbero risultare molto utili, in quanto permetterebbero di studiare entrambi i fenomeni all'interno di una serie che li presenta congiuntamente, come quella dell'inflazione statunitense.

# Riferimenti bibliografici

- Andrews, D. W. K., 1993, Tests for Parameter Instability and Structural Change with Unknown Change Point, *Econometrica*, 61, 821- 856.
- Andrews, D. W. K., e W. Ploberger, 1994, Optimal Tests when a Nuisance Parameter is Present Only Under the Alternative, *Econometrica*, 62, 1383- 1414.
- Bai, J., 1994 a, Least Squares Estimation of a shift in Linear Processes, *Journal of the Time Series Analysis*, 15 (5), Settembre, 453-72.
- Bai, J., 1994 b, Estimation of Structural Change Based on Wald Type Statistics, Working Paper No. 94-6, Department of Economics, M.I.T., forthcoming in *Review of Economics and Statistics*.
- Bai, J. e Perron, P., 1998, Estimating and testing linear models with multiple structural changes. *Econometrica*, 66, 47–78.
- Bai, J. e Perron, P., 2003, Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18, 1–22.
- Bisaglia, L. e Gerolimetto, M., 2007a, An empirical strategy to detect spurious effects in long memory and occasional-break processes. Department of Statistics, University of Padova.
- Bisaglia, L. e Gerolimetto, M., 2007b, Testing structural breaks vs. long memory with the Box-pierce statistics: a Monte Carlo study. Department of Statistical Sciences, University of Padua, *Working Paper Series*, 12, 1-10.
- Bos, C., P.H. Franses & M. Ooms, 2001, Inflation, Forecast Intervals and Long Memory Regression Models, *International Journal of Forecasting*, to appear.
- Box, G.E.P. e G.M. Jenkins, 1970, Time series analysis: Forecasting and control, San Francisco: Holden-Day.
- Cecchetti, S. G., P. Hooper, B.C. Kasman, K. L. Schoenholtz, and M. W. Watson, 2007. Understanding the evolving inflation process. *Report U.S. Monetary Policy Forum*.

- Diebold, F.X. e G.D. Rudebusch, 1989, Long memory and persistence in aggregate output, *Journal of Monetary Economics* 24, 189-209.
- Fox, R. e Taqqu, M.S., 1986, Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, 14, 517–532.
- Geweke, J. e Porter-Hudak S., 1983, The Estimation and Application of Long-Memory Time Series Models. *J.T.S.A.*, 4, (1983) 221-237.
- Granger, C.W.J. e Joyeux, R., 1980, An introduction to long-range time series models and fractional differencing. *Journal of Time Series Analysis*, 1, 15-30.
- Hosking, J.R.M, 1981, Fractional differencing. *Biometrika* 68, 165-176.
- Hurst, H.E., 1951, Long term storage capacity of reservoirs (con discussione). *Trans. Am. Soc. Civ. Eng.*, 116, 770-808.
- Kunsch, H., 1987, Statistical aspects of self-similar processes. In *Proceedings of the First World Congress of the Bernoulli Society* (Yu. Prokhorov e V. V. Sazanov, eds.) 1, 67–74. VNU Science Press, Utrecht.
- Mandelbrot, B.B., 1969, Long-run linearity, locally Gaussian processes, H-spectra and infinite variances. *International Economic Review* 10, 82-111.
- Robinson, P.M., 1995, Log-periodogram regression of time series with long range dependence. *Ann. Statist.* 23 1048-1072.
- Robinson, P. M., 1995b, Gaussian semiparametric estimation of long range dependence. *Annals of Statistics*, 23, 1630–1661.
- Sowell, F.B., 1992, Modelling long run behaviour with fractional ARIMA model, *Journal of Monetary Economics* 29, 277-302.
- Whittle, P., 1951, Hypothesis Testing in time Series Analysis, (Hafner, New York).
- Zivot, E. e Andrews, D. W. K., 1992, Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis, *Journal of Business and Economic Statistics*, 10 (3), July, 251-70.