

UNIVERSITÀ POLITECNICA DELLE MARCHE

Facoltà di Economia "G. Fuà"

Dottorato in Economia Politica XV ciclo

Dipartimento DISES



## Issues in Nonlinear Cointegration Modelling

Supervisor: Prof. Giulio Palomba

Prof. Riccardo Lucchetti

Tesi di Dottorato di: Emanuele Armillotta



# Indice

<b>Introduzione</b>	<b>5</b>
<b>1 Letteratura</b>	<b>9</b>
1.1 Struttura a termine tassi di interesse . . . . .	9
1.2 Teoria sulla struttura a termine dei tassi di interesse . . . . .	11
1.3 Cointegrazione e teoria delle aspettative . . . . .	13
1.4 Fallimento della teoria delle aspettative . . . . .	20
1.4.1 Premio per il rischio variabile . . . . .	20
1.4.2 Aspettative del mercato . . . . .	24
1.4.3 Politiche monetarie . . . . .	25
1.5 Modelli successivi . . . . .	27
1.5.1 Modelli lineari . . . . .	27
1.5.2 Modelli con aggiustamento a soglia . . . . .	36
1.5.3 Modelli con polinomi cubici . . . . .	40
1.5.4 Modello con aggiustamento approssimato da polinomio di Taylor	43
<b>2 Rassegna Metodologica</b>	<b>47</b>
2.1 Introduzione . . . . .	47
2.2 Proprietà dei modelli di cointegrazione lineare . . . . .	50
2.3 Proprietà dei modelli non lineari . . . . .	54
2.3.1 Metodo minimi quadrati . . . . .	54
2.3.2 Metodo massima verosimiglianza . . . . .	58
2.4 Procedura bootstrap . . . . .	65
2.4.1 Metodologie bootstrap per le serie storiche . . . . .	69

2.4.2	Bootstrap per modelli di cointegrazione . . . . .	74
2.4.3	Bootstrap per modelli non lineari . . . . .	81
2.5	Considerazioni sul modello NEC . . . . .	81
2.6	Conclusioni . . . . .	83
<b>3</b>	<b>Analisi Empirica</b>	<b>86</b>
3.1	Introduzione . . . . .	86
3.2	I dati . . . . .	89
3.2.1	Analisi del contesto storico . . . . .	91
3.3	Analisi preliminare . . . . .	97
3.4	Modelli parametrici . . . . .	104
3.4.1	Stima di modelli univariati . . . . .	104
3.4.2	Analisi <i>out-of-sample</i> univariata . . . . .	110
3.4.3	Stima modello multivariato . . . . .	111
3.4.4	Analisi <i>out-of-sample</i> multivariata . . . . .	117
3.5	Interpretazione dei risultati . . . . .	120
3.6	Conclusioni . . . . .	126
	<b>Bibliografia</b>	<b>129</b>
	<b>Allegati</b>	<b>142</b>
	Stime modelli univariati . . . . .	142
	Stime modelli multivariati . . . . .	154

# Introduzione

Nel corso degli ultimi anni la letteratura si è incentrata su diversi aspetti riguardanti l'analisi empirica di alcune delle più importanti teorie economiche attraverso l'uso di modelli econometrici non lineari. Un esempio sono i lavori sulla *power purchasing parity* di Lo e Zivot (2001), Taylor e Siklos (2001), Kapetanios e Shin (2002), Lundberg e Teräsvirta (2006), Nam (2011) e Beckmann (2012) o sul tasso di disoccupazione di Skalin e Teräsvirta (2002), Caner e Hansen (2003), Akram (2005) e Pérez-Alonso e Di Sanzo (2010).

Nel contesto della cointegrazione i modelli non lineari permettono di ovviare a tutti quei elementi che ostacolano l'aggiustamento verso l'equilibrio come ad esempio le frizioni di mercato (costi di transazione, premi per il rischio variabili, ecc.), il ciclo economico e/o gli interventi dei *policy maker*.

Nonostante in letteratura siano presenti innumerevoli modelli con diverse tipologie di meccanismi di aggiustamento non lineare c'è però da sottolineare che il loro ruolo è ancora del tutto marginale. La spiegazione è che i modelli non lineari spesso possono essere difficili da gestire soprattutto in ambito multivariato, inoltre non è presente un'unica teoria asintotica per tutti e a volte risulta necessario svilupparne un'apposita per singola classe di modelli. In aggiunta la complessità del modello richiede algoritmi complessi per la stima. In alcune circostanze l'intero modello non è consistente con la teoria economica o i singoli parametri non trovano una diretta spiegazione economica.

Nell'ambito della struttura a termine delle aspettative razionali fra i modelli non lineari che hanno avuto un maggior riscontro sia pratico che teorico è il modello

*threshold* di Balke e Fomby (1992). Questo modello approssima il meccanismo di aggiustamento con funzione a soglia per catturare gli interventi di natura straordinaria dei *policy maker*. Nonostante il suo largo utilizzo in letteratura, questo modello impone diverse imposizioni riguardo al meccanismo di aggiustamento non lineare e risulta poco gestibile in ambito multivariato. Tenuto conto di questi aspetti, un modello che consente una maggiore flessibilità è il modello NEC (*Nonlinear Error Correction*) di Lucchetti e Palomba (2009). Questi autori propongono l'utilizzo di un polinomio di terzo grado per approssimare un meccanismo di aggiustamento non lineare incognito anche come quello rappresentato da una funzione *threshold*.

La prima parte di questa tesi ha come obiettivo di evidenziare i principali contributi riguardanti la teoria delle aspettative sulla struttura a termine dei tassi di interesse. Partendo da lavoro di Campbell e Shiller (1987) vengono elencate tutte le problematiche inerenti la mancata evidenza empirica di questa teoria. In seguito sono esposte alcune fra le innumerevoli metodologie che sono state sviluppate successivamente e che tengono in considerazione questi problemi quali il premio per il rischio variabili, gli interventi economici e il ciclo economico. In particolare ci si sofferma sui modelli non lineari di Balke e Fomby (1992) e di Lucchetti e Palomba (2009).

Nel secondo capitolo vengono discussi alcuni dei più importanti teoremi riguardanti i modelli non lineari. L'obiettivo è quello di poter trovare una qualche teoria asintotica attinente al modello NEC e di come si possa risolvere una situazione in cui non sia possibile determinare la distribuzione limite. In particolare Kristensen e Rahbek (2013) forniscono alcuni punti su cui è importante riflettere. Infatti essi dimostrano come la distribuzione asintotica per determinati modelli possa non essere standard, sebbene il modello sotto l'ipotesi nulla sia lineare. Le possibili soluzioni a questo genere di problemi possono essere ovviate con opportuni metodi di simulazione.

Nella seconda parte del capitolo vengono discusse alcune delle più importanti procedure bootstrap utilizzate in ambito delle serie storiche. Data la complessità del modello che considera sia serie integrate che relazioni di cointegrazione e in

aggiunta la natura dei dati, su cui successivamente verranno effettuate le stime, la procedura più consona risulta essere il *wild bootstrap*.

Le distribuzioni asintotiche dei test di cointegrazione di Johansen (1995) e di stazionarietà possono essere alterate se in presenza di eteroschedasticità condizionale e non (Kim *et al.*, 2002a; Busetti e Taylor, 2003; Cavaliere, 2005; Cavaliere *et al.*, 2010b). Il contributo di questa seconda parte è duplice. Da una parte vengono elencate tutte le possibili metodologie più recenti che possono essere utilizzate nell'analisi preliminare e che risultano essere robuste all'eteroschedasticità. D'altra viene opportunamente scelta la procedura bootstrap più adatta per la verifica del modello NEC nella conduzione dei test.

Nella prima parte dell'ultimo capitolo viene effettuata un'analisi approfondita riguardo la verifica delle implicazioni della teoria delle aspettative derivante da Campbell e Shiller (1987). Rispetto al lavoro di Lucchetti e Palomba (2009) l'analisi preliminare è stata condotta in modo differente sia da un punto di vista metodologico che teorico. Infatti vengono verificate sia le implicazioni di lungo termine che di breve termine (le "high frequency") della teoria delle aspettative che riguardano più da vicino la struttura dinamica dell'aggiustamento. Inoltre tutti questi test vengono effettuati sia in modo standard che nella loro versione *wild bootstrap*. L'obiettivo è quello di escludere la possibilità che il meccanismo di aggiustamento non lineare possa essere una qualche sorta di sottoprodotto di "*unaccounted heteroskedasticity*".

Successivamente vengono effettuate le stime su un campione più aggiornato rispetto a quello di Lucchetti e Palomba (2009). Lo scopo è quello di verificare se effettivamente la dinamica dei tassi di interesse possa essere descritta da un modello di cointegrazione che ha al suo interno un meccanismo di aggiustamento non lineare tenendo però anche in considerazione gli ultimi interventi di politica monetaria straordinari come il *Quantitative Easing*.

Sebbene venga adottato lo stesso procedimento di verifica di Lucchetti e Palomba (2009), il modello NEC che viene qui stimato non utilizza alcuna specificazione per la varianza. L'obiettivo è quello di utilizzare procedure come lo stimatore HAC e il *wild bootstrap* che non impongono alcuna restrizione parametrica sul processo della volatilità e nel contempo sono robuste sia all'eteroschedasticità condizionale

che non condizionale.

Infine è stato preso in considerazione anche un data set con frequenza mensile. Gran parte dei modelli utilizzati per l'evidenza empirica della teoria delle aspettative sulla struttura a termine presenti in letteratura (Campbell e Shiller (1987), Balke e Fomby (1992), Seo (2003), ecc.) è stata condotta sui tassi di interesse americani con frequenza mensile. Lo scopo è quello di verificare se anche su dati a diversa frequenza l'applicazione del modello NEC riesce a replicare gli stessi risultati ottenuti con i tassi a frequenza settimanale.

# Capitolo 1

## Letteratura

### 1.1 Struttura a termine tassi di interesse

I mercati finanziari offrono sia una molteplicità di prodotti in cui l'investitore può scegliere come meglio gestire i suoi risparmi in base alle sue prospettive, sia una molteplicità di informazioni riguardo agli agenti che ci operano in esso. I movimenti dei tassi di interesse dei titoli obbligazionari sono sensibili agli shock finanziari e alle notizie macroeconomiche, così gli operatori riformulano continuamente le loro aspettative non appena l'informazione diviene disponibile. In particolar modo la struttura a termine dei tassi di interesse, o *yield curve*, che rappresenta la relazione esistente tra il tasso di rendimento offerto da un titolo finanziario a una propria scadenza e la scadenza stessa, fornisce informazioni sulle aspettative degli operatori sui futuri scenari economici. Gli investitori e le autorità prestano estrema attenzione alla *yield curve*, in quanto in ambito decisionale risulta essere uno strumento fondamentale riguardo alle azioni da intraprendere. Infatti, gli operatori valutano le loro strategie di investimento o di indebitamento in base alla conformazione della struttura a termine, mentre le autorità istituzionali ponderano e valutano le loro decisioni in merito alla conduzione della politica monetaria.

L'importanza della *yield curve* come strumento informativo risiede principalmente nella forma che questa assume. Graficamente viene tracciata su un sistema di assi cartesiani che ha sull'asse delle ordinate i rendimenti e su quello delle ascisse le diverse scadenze. Come si può notare dalla Fig.1.1, la curva che si ottiene può

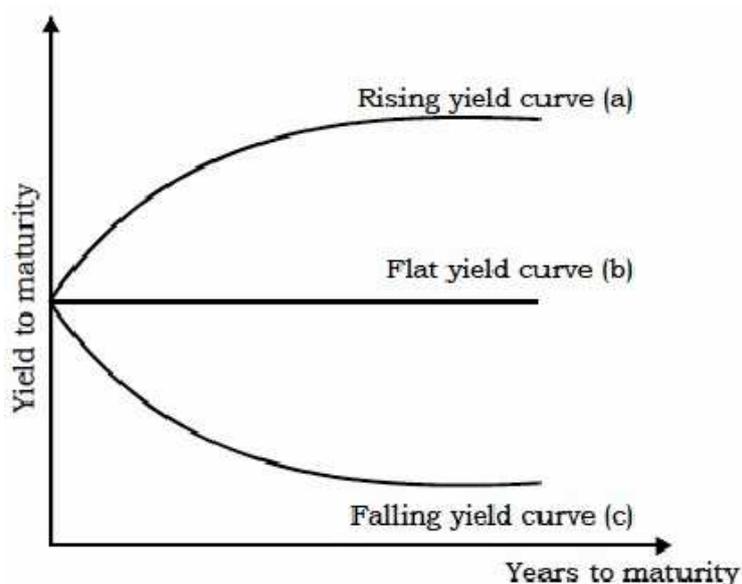


Figura 1.1: Yield Curve

avere diverse inclinazioni: crescente, piatta e decrescente.

Infatti è proprio l'inclinazione della *yield curve* che fornisce informazioni circa le aspettative degli operatori sulla situazione economica futura. Nella fattispecie, se i tassi di interesse di lungo periodo sono maggiori rispetto ai tassi a breve, quindi la curva ha pendenza positiva, allora gli operatori si attenderanno una stabile crescita economica seguita da un innalzamento dei tassi a breve effettuati dalla banca centrale. In un'ottica finanziaria un investitore prediligerà l'acquisto di titoli azionari, aumenterà il numero dei titoli obbligazionari a tasso variabile e diminuirà il tempo medio del proprio portafoglio. In caso ci si debba indebitare, la scelta per gli operatori ricadrà in modo del tutto palese verso gli strumenti a tasso fisso. Se invece la curva assume una pendenza negativa, allora gli operatori prospettano una recessione economica. In questo caso, si aspettano una riduzione dei tassi di interesse da parte della banca centrale. Gli investitori saranno orientati ad aumentare il tempo medio del portafoglio, quindi aumenteranno il numero dei titoli obbligazionari a tasso fisso, mentre coloro che dovranno indebitarsi utilizzeranno strumenti a tasso variabile. Le aspettative di rialzo o di ribasso dei tassi di interesse sono collegate a loro volta a:

- inflazione attesa: un suo aumento implica un maggior tasso di interesse che

compensi la perdita sul valore reale del titolo; viceversa una sua diminuzione implica un minor tasso di interesse,

- ciclo economico: durante la fase espansiva ci si attende un aumento dei tassi di interesse; viceversa in caso di recessione,
- conduzione della politica monetaria strettamente collegata ai primi due punti di sopra.

Questo ambito è vasto e molti autori si sono prodigati nella ricerca di un modello che meglio si presti a descrivere l'andamento dei tassi di interesse secondo la teoria delle aspettative.

## 1.2 Teoria sulla struttura a termine dei tassi di interesse

Sul piano teorico la *yield curve* rappresenta uno strumento di notevole interesse. La teoria economica introduce la dinamica della struttura a termine dei tassi di interesse in diversi modi. Tre sono le principali teorie che cercano di modellare la dinamica nel tempo della *yield curve* e la relazione esistente tra i tassi con diverse scadenze, e sono:

1. La teoria delle aspettative pure;
2. La teoria della segmentazione dei mercati;
3. La teoria del premio per la liquidità.

Introdotta da Fisher (1896), la teoria delle aspettative pure ipotizza che i tassi di interesse di lungo termine siano determinati da una media ponderata delle aspettative dei tassi a breve futuri. Questa teoria riveste una notevole importanza, in quanto fornisce informazioni sulle attese future degli operatori sul mercato. Le principali assunzioni affinché sia verificata questa relazione sono:

- la neutralità al rischio degli agenti (gli operatori agiscono in modo da massimizzare il rendimento senza tener conto del rischio);

- l'efficienza dei mercati;
- l'assenza di asimmetria informativa;
- la perfetta sostituibilità dei titoli, cioè la possibilità di poter investire in un determinato periodo in un'unica soluzione con l'acquisto di un unico titolo oppure poterlo frazionare in più titoli a scadenza più corta.

Affinché questa ipotesi sia verificata deve essere implicita l'assenza di oneri fiscali e di costi di transazione.

Se i tassi a diversa scadenza sono uguali la curva risulterà piatta, quindi, secondo la teoria delle aspettative, gli operatori si attenderanno che i tassi a breve futuri restino invariati. Al contrario, se gli agenti si aspettano che i tassi a breve futuri tendano ad aumentare allora la curva avrà pendenza positiva. In caso contrario la curva avrà pendenza negativa.

Il discorso è ben diverso per quanto concerne la teoria di segmentazione dei mercati. Questa teoria afferma che i rendimenti dei titoli con diverse scadenze siano determinati in modo esclusivo dalla domanda e dall'offerta degli investitori. Di conseguenza a fronte di una maggior quantità di domanda di titoli a lungo termine, i rendimenti di questi ultimi tenderanno a decrescere e la curva avrà inclinazione negativa. Viceversa, se la domanda per i titoli di breve tende a crescere la curva avrà pendenza positiva. Per questo motivo la forma della *yield curve* è determinata dalle sole preferenze degli operatori e ciò esclude che vi sia una qualche connessione che lega tra loro i tassi di breve e di lungo periodo.

Secondo questa teoria del premio per la liquidità gli operatori sono avversi al rischio, e per questo motivo investono su orizzonti temporali molto brevi in modo tale da ridurre il rischio di insolvenza. Al fine di poter essere competitivi, i titoli di lungo termine dovranno offrire un maggior rendimento rispetto a quelli di breve. Questa maggiorazione viene tradotta in termini di premio per la liquidità, cioè un incentivo agli operatori che investono nei titoli a più lunga scadenza. Tale teoria si presta a spiegare una delle caratteristiche più salienti della *yield curve*: la sua pendenza tendenzialmente è positiva. Poiché i tassi dei titoli con più lunga scadenza

saranno maggiori rispetto a quelli a breve.

La teoria delle aspettative razionali è quella che viene maggiormente supportata per le sue implicazioni sul comportamento degli operatori. La teoria a cui farò maggiormente riferimento è un'estensione di quella delle aspettative pure che propone un punto di conciliazione con la teoria del premio di liquidità. Infatti, Macaulay (1938), Hickman (1942) e Meiselman (1962), che hanno condotto numerosi lavori riguardo a tale argomento, sostengono la tesi che la teoria delle aspettative e del premio di liquidità possono essere viste come spiegazioni complementari del termine di struttura dei tassi di interesse.

### 1.3 Cointegrazione e teoria delle aspettative

Come specificato in precedenza, la teoria delle aspettative descrive la relazione che intercorre tra i tassi di interesse di lungo termine e i tassi futuri attesi di breve. Questa relazione viene ottenuta prendendo in considerazione due strategie di investimento: investire su un titolo di lunga durata (pari a  $k$  periodi) o investire di volta in volta in titoli di durata unitaria.

La decisione di investire al tempo  $t$  in un titolo di lungo termine a tasso fisso pari a  $rl_{t,k}$  per  $k$  periodi garantisce, per ogni unità monetaria investita, un rendimento complessivo alla scadenza pari a  $(1 + rl_{t,k})^k$ . Alternativamente, in assenza di costi di transazione, un investitore può optare di investire un'unità monetaria al tempo  $t$  in un titolo di durata di un periodo  $k$  ottenendo un rendimento di  $(1 + rs_{t,1})$  e rinnovarlo a scadenza per i successivi  $k - 1$  periodi ottenendo un rendimento complessivo pari a  $(1 + rs_{t,1})(1 + rs_{t+1,1})(1 + rs_{t+2,1}) \dots (1 + rs_{t+k,1}) = \prod_{i=0}^{k-1} (1 + rs_{t+i,1})$ .

In condizione di equilibrio le due strategie di investimento si equivalgono, quindi  $(1 + rl_{t,k})^k = \prod_{i=0}^{k-1} (1 + rs_{t+i,1})$ . Applicando i logaritmi e tenendo conto che per valori piccoli di  $(r)$  vale la relazione  $\ln(1 + r) \cong r$ , si ottiene:

$$rl_{t,k} = \frac{\sum_{i=0}^{k-1} rs_{t+i,1}}{k} \quad (1.1)$$

Questa relazione è valida se si suppone che siano noti con certezza i tassi di interesse futuri. Nella realtà gli operatori non conoscono i tassi di interesse futuri quindi formulano le proprie aspettative sui rendimenti futuri. Tenendo conto dell'ipotesi delle aspettative razionali, gli operatori prendono le loro decisioni in modo oggettivo e cioè in base all'informazione a loro nota nel periodo precedente. Quindi la formulazione secondo la teoria delle aspettative razionali assumerà la seguente formulazione:

$$rl_{t,k} = \frac{rs_t + \sum_{i=1}^{k-1} E_t(rs_{t+i})}{k} + c \quad (1.2)$$

dove  $rl_{t,k}$  rappresenta i rendimenti di lungo termine di  $k$  periodi,  $rs_t$  quelli di breve corrente di un solo periodo,  $\sum_{i=1}^{k-1} E_t(rs_{t+i})$  quelli attesi e  $c$  è il premio per il rischio. Naturalmente  $E_t(\cdot) = E(\cdot|I_t)$  è il valore atteso condizionale al set informativo  $I_t$ . Nella fattispecie la teoria afferma che i tassi di interesse debbano muoversi nella stessa direzione per eguagliare i rendimenti attesi dalle strategie di breve e di lungo termine. L'interpretazione di questa relazione è che il rendimento di lungo periodo di un titolo è uguale alla somma ponderata del rendimento del titolo corrente e di quelli attesi più un premio per il rischio che rimane costante nel tempo. Per cui, secondo questa concezione, le forze portanti per la determinazione dei tassi di interesse risultano essere le aspettative degli operatori sul mercato.

Al fine di verificare la validità della teoria delle aspettative razionali, Campbell e Shiller (1987) concentrano la loro attenzione sul comportamento nel tempo di una semplice misura della *yield curve*: lo spread tra i tassi di lungo e breve periodo. L'inclinazione della curva della struttura a termine è di certo influenzata dal segno dello spread; un valore positivo di quest'ultimo indica che i tassi di lungo periodo sono maggiori rispetto a quelli di breve per cui la curva ha un'inclinazione positiva, quindi gli agenti si aspettano un aumento dei tassi futuri a breve. Al contrario, uno spread con valore vicino allo zero o negativo sta ad indicare che gli agenti si attenderanno una riduzione dei tassi di interesse futuri.

Da questo punto di vista lo spread può essere presentato come un previsore ottimo delle variazioni future dei tassi di interesse. Campbell e Shiller (1987) rimodellano

l'equazione (1.2) in modo da ottenere due equazioni del tipo:

$$S_t = \sum_{i=1}^{k-1} \left( \frac{k-i}{k} \right) E_t(\Delta r_{s_{t+i}}) + c \quad (1.3)$$

dove  $S_t$  rappresenta lo spread mentre  $\left(\frac{k-i}{k}\right)E_t(\Delta r_{s_{t+i}})$  è una media ponderata delle variazioni dei rendimenti attesi dei titoli di breve termine e  $c$  una qualche sorta di premio per il rischio costante nel tempo, e

$$\left( \frac{1}{k-1} \right) S_t = E_t(\Delta r_{l_{t+1}}) + c \quad (1.4)$$

dove  $E_t(\Delta r_{l_{t+1}})$  rappresenta la variazione attesa dei rendimenti dei titoli di lungo termine. Entrambe le equazioni (1.3) e (1.4) mettono in evidenza come lo spread possa essere un previsore ottimale sia per le variazioni dei tassi a breve, sia per quelli a lungo termine.

Nel loro lavoro, Campbell e Shiller (1987) giungono ad un'importante conclusione riguardo alle serie storiche dei tassi di interesse. Essi osservano che le serie dei tassi nei livelli sono chiaramente non stazionarie, ma sembrano esibire una chiara tendenza a seguire un trend comune. Ciò comporta che le serie siano cointegrate. Utilizzata in molti lavori empirici, la cointegrazione è un'idea basata sul fatto che variabili non stazionarie possono avere una relazione di equilibrio di lungo periodo e, pertanto, hanno la tendenza a muoversi assieme (Engle e Granger, 1987). Il modello utilizzato in fase di stima è l'ECM (*Error Correction Mechanism*). Infatti la sua formulazione permette di visualizzare come le variabili nel breve periodo rispondono alle deviazioni dall'equilibrio di lungo periodo.

La serie  $Y_t$  è un vettore di dimensione  $n \times 1$  avente al suo interno i rendimenti di  $n$  titoli a diversa scadenza ordinati in modo decrescente rispetto alla loro durata. Se si parte da un VAR (*Vector Auto-Regressive*) il modello Vector ECM (VECM) si presenta in questa forma:

$$\Delta Y_t = \mu_t + \alpha \beta' Y_{t-1} + \Gamma(L) \Delta Y_{t-1} + \varepsilon_t. \quad (1.5)$$

In esso le variazioni di  $Y_t$  sono determinate da un nucleo deterministico  $\mu_t = \delta_0 + \delta_1 t + \delta_2 t^2 + \dots + \delta_p t^p$ , da un termine di correzione dell'errore, dalle sue variazioni passate e dall'innovazione  $\varepsilon_t$ . In particolare  $\Gamma(L)$  rappresenta il polinomio nell'operatore ritardo,  $\alpha$  è il parametro della velocità di aggiustamento, cioè l'intensità con cui  $\Delta Y_t$  risponde alle deviazioni dall'equilibrio, mentre il termine  $\beta' Y_{t-1}$  indica le relazioni lineari di cointegrazione.

Per quanto riguarda la componente di aggiustamento, se la matrice rettangolare  $\beta'$  di dimensione  $r \times n$  con  $r < n$  assume la seguente forma:

$$\begin{pmatrix} 1 & 0 & \dots & 0 & -1 \\ 0 & 1 & \dots & 0 & -1 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & -1 \end{pmatrix} \quad (1.6)$$

essa permette di calcolare gli spread fra i diversi titoli e quello a più breve scadenza. In questo caso la matrice  $\beta'$  gode di un'importante caratteristica affinché sia possibile verificare la teoria delle aspettative, cioè che il rango dovrà essere pari a  $(n - 1)$ , quindi le serie dei tassi devono seguire un unico trend comune. Il modello assumerà la seguente connotazione:

$$\Phi(L)\Delta Y_t = \mu + \alpha S_{t-1} + \varepsilon_t, \quad (1.7)$$

dove  $S_t$  rappresenta il vettore contenente gli spread tra i tassi di interesse cioè  $\beta' Y_t$ .

I risultati ottenuti da Campbell e Shiller (1987), applicando questo modello, non sembrano confermare la teoria, anzi, evidenziano un paradosso. Come sottolineato in un loro lavoro successivo (Campbell e Shiller, 1991), l'inclinazione della struttura a termine, cioè lo spread, prevede quasi sempre in modo errato la direzione dei rendimenti di lungo periodo, mentre dà la giusta direzione nei tassi di breve periodo. Quindi le stime mostrano che il coefficiente relativo allo spread ha il segno opposto a quello dettato dalla teoria delle aspettative nel caso dei titoli di lungo periodo, mentre per il caso dei titoli di breve periodo, il coefficiente risulta essere coerente

con la teoria.

Sebbene il VECM risulti una buona rappresentazione della teoria delle aspettative della struttura a termine ciò non basta per la verifica empirica di quest'ultima. Infatti Campbell e Shiller (1987) descrivono anche le condizioni necessarie e sufficienti affinché i dati possano confermare la teoria. In sostanza le implicazioni che derivano dalla teoria delle aspettative sono essenzialmente di due tipi. La prima definisce in concreto la forma della matrice di cointegrazione  $\beta$ , nella fattispecie i tassi devono seguire un unico trend stocastico, quindi il rango di cointegrazione deve essere pari a  $(n - 1)$  ed inoltre devono essere rispettate le restrizioni descritte nell'equazione (1.6). Mentre quest'ultime verificano le implicazioni di lungo periodo quindi di "low frequency", le altre condizioni necessarie riguardano l'aspetto della struttura dinamica dell'aggiustamento e sono le condizioni di "high frequency". Ritornando all'equazione (1.3), questa rappresenta il teorema delle aspettative razionali della struttura a termine che è stata ottenuta da Campbell e Shiller (1987) dalla relazione del *present value*. Tramite l'equazione (1.3), ottenuta da Campbell e Shiller (1987) dalla relazione del *present value*, Shiller (1979) è giunto ad ottenere una formulazione alternativa con una sostituzione ricorsiva in modo da ottenere

$$S_t = \sum_{i=1}^{\infty} \delta^i E_t(\Delta r_{s_{t+i}}) + c, \quad (1.8)$$

dove  $\delta$  rappresenta un fattore di sconto o un parametro di linearizzazione<sup>1</sup>.

Un modo per verificare il modello *present value*, di conseguenza anche la teoria delle aspettative razionali, è di testare l'equazione (1.8) utilizzandola per limitare il comportamento della variabile  $\xi_t = rl_t - (1/\delta)[rl_{t-1} - (1 - \delta)rs_{t-1}]$  che, tolta la costante, rappresenta la *true innovation* al tempo  $t$  in  $rl_t$ . Infatti sostituendo con l'equazione (1.1) si arriva ad ottenere

$$\xi_t = rl_t - E_{t-1}(rl_t) - c(1 - 1/\delta). \quad (1.9)$$

In termini economici questa variabile rappresenta l'*excess return* dei bond di

---

<sup>1</sup>Campbell e Shiller (1987) pongono  $\delta = 1/(1 + R)$  dove  $R$  è la media del tasso di più lungo periodo calcolato sul campione.

lungo periodo rispetto ai titoli di breve durata che risulta non prevedibile quando si ha a disposizione la sola informazione al tempo  $t - 1$ . In termini matematici le variabili all'interno dell'equazione (1.9) risultano tutte ortogonali al set informativo  $I_{t-1}$ . Un modo per verificare questa condizione di ortogonalità è quello di eseguire la regressione di  $\xi_t$  rispetto alle variabili  $rs_{t-1}$  e  $rl_{t-1}$  e controllare che i coefficienti siano congiuntamente uguali a zero. Questa procedura risulta problematica a causa della non stazionarietà dei regressori e di alcune implicazioni di carattere teorico. Un modo per aggirare questi ostacoli è quello di utilizzare una combinazione lineare e stazionaria delle variabili cioè lo spread ( $S_t$ ). Infatti quest'ultimo può essere riscritto come

$$S_t = \frac{1}{1-\delta} \Delta r l_t - \Delta r s_t - \frac{\delta}{1-\delta} \xi_t.$$

Risolvendo per  $\xi_t$  si ha

$$\xi_t = S_t - \frac{1}{\delta} S_{t-1} + \Delta r s_t. \quad (1.10)$$

Quindi un'ulteriore implicazione del modello della teoria delle aspettative razionali può essere verificata regredendo  $\xi_t$ , costruita tramite l'equazione (1.10), sulle variabili  $\Delta r s_{t-1}$  e  $\Delta r l_{t-1}$  e verificare che i coefficienti relativi a questi ultimi siano a zero. In alternativa, partendo dalle Equazioni (1.3) e (1.4), è possibile riscrivere il modello VECM in forma VAR. Infatti prendendo le variabili stazionarie  $\Delta r s_t$  e  $S_t$ , a cui sono state sottratte ad ognuna la propria media, si ottiene

$$\begin{bmatrix} \Delta r s_t \\ S_t \end{bmatrix} = \begin{bmatrix} a(L) & b(L) \\ c(L) & d(L) \end{bmatrix} \begin{bmatrix} \Delta r s_{t-1} \\ S_{t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \quad (1.11)$$

dove  $a(L), b(L), c(L)$  e  $d(L)$  sono i polinomi nell'operatore ritardo di ordine  $p$ . Riscrivendo l'equazione sotto forma di VAR(1) si ottiene  $z_t = Az_{t-1} + v_t$ , dove  $A$  è la *companion matrix* mentre  $z_t = [\Delta r s_t, \dots, \Delta r s_{t-p+1}, S_t, \dots, S_{t-p+1}]$ . Definiti poi due vettori  $g'$  e  $h'$  tale che  $g'z_t = S_t$  e  $h'z_t = \Delta r s_t$  ci si riconduce all'equazione (1.8)

$$g'z_t = \sum_{i=1}^{\infty} \delta^i h' A^i z_t \quad (1.12)$$

Semplificando per  $z_t$  si ottiene

$$g' = \sum_{i=1}^{\infty} \delta^i h' A^i = \delta h' A (I - \delta A)^{-1} \quad (1.13)$$

post moltiplicando entrambi i membri per  $(I - \delta A)^{-1}$  si hanno infine le *cross-equation constrains*

$$g'(I - \delta A) = \delta h' A. \quad (1.14)$$

Quest'ultima espressione si ricongiunge all'equazione (1.10). Infine con l'equazione (1.14) si ottiene le restrizioni per il test di Wald simile al test di azzeramento congiunto per la regressione di  $\xi_t$  sui valori  $\Delta r_{s_{t-1}}$  e  $S_{t-1}$  dell'equazione (1.10).

In conclusione per verificare le implicazioni “high frequency” della teoria delle aspettative si può ricorrere a due metodologie la prima rappresentata dall'equazione (1.10) e la seconda dall'equazione (1.14).

I risultati derivanti dall'applicazione di queste implicazioni di Campbell e Shiller (1987) evidenziano ancora una volta un'incongruenza sia tra loro che rispetto la teoria delle aspettative. Infatti i test svolti sulle radici unitarie indicano che il termine  $\xi_t$ , costruito secondo l'equazione (1.10) è stazionario; ciò implica, seppur in modo indiretto, la stazionarietà degli spread. Invece il test sulle *cross-equation constrains* dell'equazione (1.14) rifiuta fortemente l'ipotesi nulla che gli *excess return* dei bond di lungo periodo siano non prevedibili ciò si discosta da quanto previsto dalla teoria.

Un ultimo test per la verifica empirica della teoria è il confronto tra lo spread effettivo e quello teorico. Lo spread teorico è costruito con la media ponderata delle variazioni dei futuri tassi di interesse e risulta correlato con lo spread effettivo, la differenza tra i tassi di interesse, con un valore molto vicino ad uno. Quest'ultimo risultato, al contrario del precedente, suggerisce che gli spread effettivi si comportano in modo simile al modello delle aspettative.

## 1.4 Fallimento della teoria delle aspettative

L'evidente paradosso ottenuto da Campbell e Shiller (1987), con l'applicazione della metodologia ECM, ha condotto diversi autori ad investigare sulle cause che possono condurre a rifiutare la teoria delle aspettative pure. Un'estensiva letteratura che comprende, tra gli altri, Campbell e Shiller (1991), Hardouvelis (1992), Mankiw e Miron (1986), Fama (1984), Daniel (2004), Hall *et al.* (1992), giunge alla conclusione che le motivazioni alla base del mancato supporto dell'evidenza empirica risalgono principalmente a tre distinti motivi:

- premio per il rischio non costante nel tempo;
- aspettative non razionali;
- implementazione di politiche monetarie.

### 1.4.1 Premio per il rischio variabile

Come già sottolineato, il premio per il rischio (o a termine) rappresenta l'incentivo per gli investitori che acquistano titoli di durata più lunga. In altri termini, rappresenta un rendimento extra necessario per compensare gli operatori che sopportano un rischio maggiore dovuto dalla detenzione di titoli nel lungo periodo.

La teoria presuppone stringenti ipotesi riguardo alle proprietà del premio per il rischio. Infatti, secondo la teoria delle aspettative pure, le aspettative sono l'unica componente a determinare i tassi di interesse. Secondo questa impostazione non vi è previsto alcun termine costante, quindi viene imposto come vincolo che il premio sia pari a zero. Invece, nella versione allargata della teoria delle aspettative, viene supposto l'esistenza di un premio che rimane costante nel tempo. La validità di queste ipotesi viene investigata da diversi autori tra i quali Fama (1984), Tzavalis e Wickens (1998), Mankiw e Miron (1986) i quali asseriscono che il premio non rispetti tali vincoli ma, al contrario, possa variare nel tempo.

Già nel lavoro di Campbell e Shiller (1991) viene evidenziato come una delle possibili spiegazioni delle deviazioni dal modello descritto dalla teoria possa essere causato dal premio. In particolar modo, gli autori ipotizzano la possibilità che il

premio per il rischio sia correlato con gli aumenti nei tassi di interesse di breve periodo.

Nel lavoro di Tzavalis e Wickens (1998) viene analizzato il paradosso legato all'evidenza empirica della teoria delle aspettative. Nello specifico viene argomentato che questo paradosso possa dipendere dalla presenza del premio a termine che varia nel tempo. La loro idea principale è che il premio possa essere costruito come la differenza tra il rendimento atteso nel tenere per un periodo un bond di  $k$  periodi e il tasso dell'attività priva di rischio. Al fine di spiegare la costruzione di questo indice sintetico, gli autori si sono avvalsi di concetti basilari utilizzati in finanza. Il rendimento di un titolo, indicato con  $r$ , è pari al tasso di sconto che eguaglia il prezzo del titolo al valore attuale del suo pagamento finale. Prendendo in considerazione un titolo obbligazionario con *face value* pari a 1 e con durata pari a  $k$  periodi, allora il suo prezzo  $p$ , calcolato su base continua, sarà uguale a

$$p_{t,k} = \frac{1}{(1 + r_{t,k})^k} \quad (1.15)$$

Applicando il logaritmo e tenendo conto che, per valori piccoli di  $r$ , abbiamo  $\ln(1 + r) \cong r$  si ottiene

$$\ln(p_{t,k}) = -kr_{t,k} \quad (1.16)$$

Dalla teoria delle aspettative razionali può essere derivata la relazione di arbitraggio secondo cui il rendimento atteso nel tenere per un periodo un titolo che ha  $k$  periodi alla scadenza, aggiustato per il rischio, è pari al rendimento di un titolo di un periodo. In formule

$$E_t h_{t+1,k} = r s_{t,1} + \mu_{t,k} \quad (1.17)$$

dove  $r s_{t,1}$  rappresenta il rendimento di un titolo a un periodo, mentre  $\mu_{t,k}$  è il premio per il rischio. A sinistra dell'equazione abbiamo, invece, l'aspettativa del guadagno in conto capitale che è calcolato come  $h_{t+1,k} = \ln(p_{t+1,k-1}) - \ln(p_{t,k})$ . Riprendendo in considerazione l'equazione (1.16) otteniamo  $h_{t+1,k} = (k)r l_{t,k} - (k-1)r l_{k-1,t+1}$ . Nell'analisi di lungo periodo, attraverso un modello VECM, Tzavalis e Wickens (1998) utilizzano il vettore  $x_t = (h_{t+1,k}, r l_{t,k}, r s_{t,1})'$  in modo da ottenere come rela-

zioni di lungo periodo  $h_{t+1,k} - rs_{t,1}$  e  $rl_{t,k} - rs_{t,1}$  che rappresentano rispettivamente il premio a termine e lo spread. Quindi l'obiettivo è la verifica della stazionarietà del premio per il rischio: se non ci fosse cointegrazione, allora il premio per il rischio risulterebbe non stazionario.

I risultati evidenziano che il premio è variabile nel tempo. Sebbene questo risultato sia a sfavore della teoria delle aspettative, Tzavalis e Wickens (1998) trovano che la componente principale che influenza la volatilità dei rendimenti *ex post* nel detenere un titolo per un periodo siano le 'news' sui tassi interesse a breve. Gli autori quindi concludono che, nonostante il premio per il rischio sia variabile, la struttura a termine ha un singolo fattore di rappresentazione dato dalle aspettative dei tassi futuri a breve termine.

Mankiw e Miron (1986) mostrano formalmente come le variazioni del premio a termine possano condurre ad una distorsione verso il basso del coefficiente dello spread.

Il loro modello parte prendendo in considerazione la relazione tra il rendimento di due titoli, uno da un periodo ( $rl_t$ ) e l'altro da due ( $rs_t$ ), secondo la teoria delle aspettative:

$$rl_t = \mu + \frac{1}{2}(E_t rs_{t+1} + rs_t)$$

dove  $\mu$  è il premio per il rischio e  $E_t rs_{t+1}$  è il rendimento atteso al tempo  $t$  del titolo a un periodo. L'equazione può essere facilmente riscritta come

$$E_t rs_{t+1} + rs_t = -2\mu + 2(rl_t - rs_t) \quad (1.18)$$

Se si considera che gli operatori formulano in modo del tutto razionale le loro aspettative allora risulterà

$$rs_{t+1} = E_t rs_{t+1} + \varepsilon_{t+1}, \quad (1.19)$$

dove il termine  $\varepsilon_{t+1}$  è ortogonale all'informazione al tempo  $t$ . Sostituendo  $E_t rs_{t+1}$

dalla (1.19) nell'equazione (1.18) otteniamo

$$rs_{t+1} - rs_t = \alpha + \beta(rl_t - rs_t) + \varepsilon_{t+1}, \quad (1.20)$$

dove in accordo con la teoria descritta dall'equazione (1.18), risulta  $\alpha = -2\mu$  e  $\beta = 2$ . Le stime condotte da Mankiw e Miron (1986) mostrano che il valore di  $\beta$  assume valori vicino allo zero e i test rifiutano l'ipotesi che il parametro sia uguale a quello dettato dalla teoria. Gli autori inoltre dimostrano come queste distorsioni verso il basso del coefficiente dello spread siano dovute alla presenza di un premio a termine non costante ( $\mu_t$ ). Infatti, secondo l'equazione (1.20) la stima di  $\beta$  è

$$\hat{\beta} = \frac{2Var(E_t\Delta r_{t+1}) + 4cov(E_t\Delta r_{t+1}, \mu_t)}{Var(E_t\Delta r_{t+1}) + 4Var(\mu_t) + 4cov(E_t\Delta r_{t+1}, \sigma(\mu_t))}$$

dove  $Var(\cdot)$  indica la varianza e  $cov(\cdot)$  indica la covarianza delle variazioni dei rendimenti attesi e del premio per il rischio non costante. Per questo motivo se la volatilità dei tassi di interesse è molto elevata e il premio a termine rimane costante, quindi varianza e covarianza riferite a  $\mu_t$  è uguale a zero, allora l'ipotesi che  $\beta = 2$  non è rifiutata. Viceversa, se  $Var(E_t\Delta r_{t+1})$  assume valori molto piccoli e il premio a termine non rimane costante, allora il test rifiuterà l'ipotesi che  $\beta = 2$  e il coefficiente assumerà valori molto vicini allo zero.

In un lavoro antecedente Fama (1984) indaga, attraverso una regressione lineare, se i tassi *forward* possano essere indicativi per comprendere i movimenti del premio a termine e dei tassi di interesse futuri attesi. I risultati ottenuti sembrano confermare l'idea dell'autore, il quale porta a concludere che il tasso *forward* possa essere un utile previsore per entrambe le variabili (premio e tassi di interesse) e che sia più consapevole utilizzare il fattore premio come una variabile anziché come una costante.

### 1.4.2 Aspettative del mercato

Non tutti i risultati provenienti dall'evidenza empirica sono discordanti con la teoria delle aspettative. Infatti i coefficienti reattivi allo spread assumono segno opposto a quello previsto della teoria mentre per i tassi a breve i coefficienti risultano congruenti. Alternativamente alle ipotesi riguardanti il premio per il rischio, questo paradosso potrebbe essere spiegato dal fatto che i tassi di lungo periodo reagiscono troppo lentamente agli spostamenti improvvisi degli spread le cui variazioni sono principalmente determinate dai movimenti dei tassi di interesse a breve. I movimenti dello spread risultano essere troppo ampi per poter essere in linea con la teoria delle aspettative.

Mankiw e Summers (1984) indagano principalmente sulle aspettative degli operatori del mercato, in particolare sul fatto che se il tasso di lungo sia costruito in modo coerente alla teoria o meno e cioè come media ponderata dei tassi correnti e futuri. Essi riscontrano che vi è un eccesso di sensibilità; nella determinazione dei tassi di lungo periodo gli agenti danno un peso maggiore ai tassi futuri attesi anziché al tasso corrente. Per cui un modello coerente con la teoria delle aspettative potrebbe “sottostimare” l'effetto in quanto il tasso di interesse di lungo periodo *underreact* ai tassi di breve.

Anche Hardouvelis (1994) si concentra sulle stesse possibili cause del paradosso riscontrato da Campbell e Shiller (1987). Nel dettaglio, si sofferma sul fatto che, nonostante lo spread contenga informazioni riguardo alle aspettative dei tassi di interesse futuri, le aspettative di mercato violano l'ipotesi di aspettative razionali. Per cui giunge alla stessa conclusione che gli agenti danno maggior peso alle variazioni future dei rendimenti a breve termine. Inoltre egli aggiunge che un'ulteriore spiegazione possa essere dovuta dal fatto che, anche se le aspettative del mercato possano essere considerate razionali, l'informazione contenuta nello spread sia troppo complessa per determinare la variazione dei tassi futuri attesi e il premio per il rischio.

L'ipotesi di eccesso di sensibilità implica che il tasso di interesse di lungo periodo

diminuisce quando i tassi a breve salgono. A sostegno di questa ipotesi troviamo il lavoro di Froot (1989). Egli conduce un'analisi sulle possibili cause della correlazione negativa tra i tassi di lungo periodo e lo spread. I risultati mostrano che la motivazione principale sia dovuta non tanto a un premio per il rischio non costante, ma dalla violazione dell'ipotesi delle aspettative razionali.

### 1.4.3 Politiche monetarie

La relazione che intercorre tra i tassi a breve e i tassi a lungo periodo gioca un ruolo centrale nella determinazione degli obiettivi di trasmissione della politica monetaria. Nella fattispecie, lo spread dei tassi di interesse a diverse scadenze, come si può anche vedere nelle equazioni (1.3) e (1.4), costituisce uno strumento di analisi delle aspettative del mercato, in quanto fornisce indicazioni sulle attese degli operatori sull'andamento futuro dei tassi di interesse a breve e, di conseguenza, anche sulle prospettive economiche e di inflazione. Per questo motivo è stata concentrata l'attenzione sulle possibili influenze da parte delle banche centrali sulla *yield curve*, nella fattispecie, se l'adozione di nuove strategie dei *policy maker* possano essere una motivazione del rifiuto dell'analisi empirica della teoria delle aspettative pure.

Nel lavoro di Mankiw e Miron (1986) viene messo a confronto il modello di cointegrazione in differenti regimi monetari. Nel dettaglio, il campione viene suddiviso in due distinti sotto campioni: prima (1890-1914) e dopo la fondazione della Fed (1915-1979). Dalle stime risulta che, nel secondo sotto campione, le aspettative non hanno un ruolo chiave nella determinazione della *yield curve* a causa dell'implementazione del nuovo regime monetario imposto dalla Fed. Gli autori asseriscono inoltre che una delle possibili cause della mancata evidenza empirica della teoria sia che le politiche monetarie, mirate ad aumentare i futuri tassi di interesse a breve, possano comportare un effetto immediato sul tasso a lungo corrente. Questo effetto, indotto dalla Fed, causa una maggior persistenza dei tassi di interesse di lungo periodo su livelli più alti.

Una diversa analisi sull'impatto delle modifiche degli obiettivi dalla Fed viene effettuata da Hall *et al.* (1992). I quali analizzano i rendimenti dei bond americani di breve periodo a diverse scadenze in distinti sotto campioni. L'ampiezza di ogni singolo sotto campione è determinato in base ai periodi in cui la Fed ha adottato nuove procedure operative. La modalità con cui viene condotta questa ipotesi è quella di testare se il rango di cointegrazione dettato dalla teoria, cioè pari al numero dei titoli considerati meno uno, sia consistente in ogni nuovo regime. Le stime portano a concludere che, in ogni regime, i rendimenti dei bond sono cointegrati, ma il rango di cointegrazione non sempre rispecchia quello imposto dalla teoria ( $n - 1$ ).

Di differente portata è il lavoro di McCallum (1994) il quale presuppone che le banche centrali, in particolare quella americana, possano seguire una regola semplice che incorpori lo spread tra i tassi di interesse. Nel suo articolo evidenzia come gli interventi dei *policy maker* sui tassi di interesse a breve termine, che si propagano lungo tutta la struttura a termine, siano mirati ad "appiattare" la curva dei rendimenti. La giustificazione risiede nel fatto che la banca centrale, conducendo una politica lungimirante anticiclica, consideri lo spread come un indicatore globale dello stato dell'economia e che, quindi, valori inusuali di questo differenziale richiedano un'azione correttiva.

Sulla base delle ipotesi e del modello di McCallum, Hsu e Kugler (1997) conducono la loro analisi su un campione americano più aggiornato. I risultati mostrano che nei primi sotto campioni la teoria delle aspettative non trova supporto empirico, mentre nel periodo che va dal 1987 al 1995 la stessa teoria non può essere rifiutata. Quest'ultimo risultato, secondo Hsu e Kugler (1997), è dovuto dal fatto che proprio la banca centrale ha adottato strategie che erano in linea con la teoria delle aspettative. In termini pratici ciò si traduce nell'inserire nel modello il ruolo che gioca l'autorità monetaria e quindi la sua funzione di reazione basata sullo spread. I risultati che giungono dall'adozione di questa modifica portano alla verifica empirica della teoria delle aspettative razionali.

## 1.5 Modelli successivi

Il rifiuto parziale della teoria delle aspettative e le deviazioni transitorie dal modello sono state al centro di numerosi dibattiti per quanto riguarda l'evidenza empirica della teoria delle aspettative. Inoltre le ipotesi, verificate statisticamente, della presenza di un premio a rischio non costante, delle aspettative non razionali e dei mutamenti delle politiche monetarie - che possono giustificare la discrepanza dei risultati - hanno condotto verso l'introduzione di alcune modifiche alle condizioni imposte dalla teoria delle aspettative razionali.

A fronte di ciò la letteratura si è mossa in diverse direzioni. Da una parte troviamo modelli che utilizzano la stessa metodologia di Campbell e Shiller (1987), ma ampliano il set delle variabili esplicative. In questo filone di letteratura troviamo Ang e Piazzesi (2003), Valente *et al.* (2004), Carriero *et al.* (2004) inseriscono variabili aggiuntive come inflazione e indicatori dell'attività economica. Altri autori, come Longstaff (2000), Della Corte *et al.* (2007), conducono invece le loro analisi su un set di dati diversamente composto, cioè utilizzando dati con frequenza molto alta o su titoli con date di scadenza molto brevi.

Da l'altra parte troviamo autori come Balke e Fomby (1992), Enders e Granger (1998), Bec e Rahbek (2004), Seo (2003), Krishnakumar e Neto (2009), Cai *et al.* (2015) che utilizzano modelli non lineari. Questi autori ipotizzano che la relazione tra lo spread e i tassi di interesse segua una dinamica complessa che non può essere descritta con modelli lineari. Nei prossimi paragrafi elencherò alcuni dei lavori più significativi, soffermandomi in particolare sui modelli non lineari.

### 1.5.1 Modelli lineari

Nella letteratura sono presenti diversi lavori che mettono in evidenza le relazioni che intercorrono tra i tassi di interesse e le variabili macroeconomiche. In particolare, la *yield curve* sembra essere un utile strumento informativo riguardo alle aspettative sull'attività economica. La ricerca empirica si è concentrata principalmente sulle

capacità previsive dell'inclinazione della curva. In questo filone della letteratura troviamo Estrella e Hardouvelis (1991) i quali mostrano che il differenziale a termine è un ottimo previsore dell'attività economica reale e Mishkin (1990) che espone la connessione tra lo spread e i futuri tassi d'inflazione. In un altro lavoro di Estrella e Mishkin (1996) mostrano come il differenziale tra i tassi sia un utile strumento per la previsione di eventi economici come le recessioni. Nel lavoro di McCallum (2005) invece viene presupposto che le banche centrali, in particolare quella americana, possano seguire una regola semplice che incorpora lo spread tra i tassi di interesse.

Se da un lato troviamo analisi sulla capacità previsiva della *yield curve* sull'inflazione e sull'attività economica, dall'altra parte troviamo un filone di letteratura che si è invece incentrato sulle dinamiche in senso opposto: cioè su come le variabili macroeconomiche influenzino la struttura a termine dei tassi di interesse nominali.

In questo secondo filone della letteratura troviamo i lavori di Ang e Piazzesi (2003), Valente *et al.* (2004), Carriero *et al.* (2004) in cui si investiga, tramite l'utilizzo del VAR, sulle relazioni tra la *yield curve* e le variabili macroeconomiche. Questi autori presuppongono che il modo con cui viene condotta la politica monetaria sia la possibile spiegazione per cui gli shock macroeconomici influenzino i tassi di interesse. Nello specifico, i *policy maker* reagiscono secondo il principio di Taylor, quindi le loro decisioni sono ponderate su variabili dell'output e dell'inflazione. Se si considera poi che le azioni dei *policy maker* hanno conseguenze sia sui tassi di interesse di breve termine, sia sui tassi a lungo termine tramite il canale delle aspettative, allora queste variabili macroeconomiche hanno un ruolo decisivo nel determinare la forma della *yield curve*. Tutto ciò ha condotto all'inserimento di variabili esogene come l'inflazione e il reddito nel set informativo.

Di particolare rilevanza sono anche le indagini di Diebold e Li (2003), Ang e Piazzesi (2003), Giese (2008), Joslin *et al.* (2014), Boswijk *et al.* (2016), ecc. sui fattori che possono contribuire a spiegare le dinamiche della struttura a termine dei tassi di interesse. Il modello a fattori latenti, su cui si basa questo filone di letteratura, parte dal presupposto che i tassi di rendimento correnti sono generati da un'equazione differenziale e che, secondo la definizione della teoria delle aspettative

razionali, i tassi futuri rappresentano la soluzione del sistema. Considerando che i rendimenti dei tassi di interesse futuri  $r(t)$  al tempo  $t$  sono rappresentati dalla soluzione di un'equazione di secondo grado con radici uguali e reali avremmo che

$$r(t) = \beta_0 + \beta_1 \cdot e^{-\lambda t} + \beta_2 \lambda t \cdot e^{-\lambda t}, \quad (1.21)$$

dove  $\lambda, \beta_0, \beta_1, \beta_2, \beta_3$  sono costanti.

Riprendendo l'equazione (1.2) secondo cui il rendimento di un titolo a lunga scadenza  $R(t)$  sarà uguale alla media dei tassi correnti e futuri otteniamo

$$R(t) = 1/t \int_0^t r(x) dx. \quad (1.22)$$

Andando ad integrare l'equazione (1.21) si avrà

$$R(t) = \beta_0 + \beta_1 \left( \frac{1 - e^{-\lambda t}}{\lambda t} \right) + \beta_2 \left( \frac{1 - e^{-\lambda t}}{\lambda t} - e^{-\lambda t} \right) \quad (1.23)$$

questa rappresentazione di Nelson e Siegel (1987) ci permette di catturare le dinamiche tra i rendimenti e le loro rispettive scadenze ovvero la *yield curve*.

Successivamente Diebold e Li (2003) inseriscono la formulazione di Nelson e Siegel (1987) in un contesto dinamico. Infatti, se si considera l'ipotesi che i parametri possano variare nel tempo, allora l'equazione (1.23) può essere interpretata come un modello a fattori latenti in cui  $\beta_0, \beta_1, \beta_2$  sono le variabili che influenzano rispettivamente il livello ( $L_t$ ), l'inclinazione  $S_t$  e la curvatura  $C_t$  della *yield curve* quindi, si avrà

$$R(t) = L_t + S_t \left( \frac{1 - e^{-\lambda t}}{\lambda t} \right) + C_t \left( \frac{1 - e^{-\lambda t}}{\lambda t} - e^{-\lambda t} \right). \quad (1.24)$$

Ipotizzando che le variabili  $L_t, S_t, C_t$  seguano un processo autoregressivo di primo ordine, il modello può essere riscritto in forma di spazio degli stati. Dove l'equazione di transizione è rappresentata dal vettore contenente le variabili che definiscono il livello, l'inclinazione e la curvatura mentre l'equazione (1.23) rappresenta l'equazione di misurazione. Al fine di caratterizzare la relazione tra la *yield curve* e le variabili macroeconomiche Diebold *et al.* (2006) inseriscono nel set delle variabili dipendenti i *federal funds rate*, l'inflazione e un indicatore del livello dell'attività economica. Le

stime indicano una forte influenza delle variabili macroeconomiche sulla struttura a termine dei tassi interessi e una più debole per quanto riguarda gli effetti della *yield curve* sulle dinamiche future dei fattori macro. Il principale vantaggio nell'utilizzare questi modelli è quello di imporre la condizione di assenza di arbitraggio tra i diversi titoli ed inoltre permette di comprendere e caratterizzare le dinamiche del premio per il rischio non costante nel tempo.

Tra i modelli i fattori latenti (il livello, l'inclinazione e la curvatura della *yield curve*) che utilizzano anche variabili macroeconomiche troviamo Ang e Piazzesi (2003), Piazzesi e Cochrane (2009), Joslin *et al.* (2014). Questi autori modellano la struttura a termine dei tassi sempre attraverso un modello VAR ma inserendo le variabili macroeconomiche e i fattori latenti all'interno di un fattore di sconto stocastico (o *price kernel*). Formalmente il loro modello è un caso particolare della versione dei modelli affini in tempo discreto di Duffie e Kan (1996) in cui i prezzi dei bond sono funzioni esponenziali affini delle variabili di stato. Il modello parte considerando la regola di Taylor (1993) in cui i rendimenti di breve periodo siano una funzione affine delle variabili macroeconomiche ( $X_t^0$ ) e un termine di errore ortogonale alle variabili indipendenti ( $v_t$ )

$$r_t = \alpha_0 + \alpha' X_t^0 + v_t. \quad (1.25)$$

Secondo gli autori il termine di errore, che non può essere spiegato dalle variabili macroeconomiche, può essere rappresentato dai fattori latenti ( $X_t^u$ ) quindi, si ottiene

$$r_t = \delta_0 + \delta_1' X_t^0 + \delta_2' X_t^u. \quad (1.26)$$

Si suppone inoltre che la dinamica delle variabili di stato  $X_t' = (X_t^0, X_t^u)'$  segua un processo VAR con errori gaussiani

$$X_t = c + \rho X_{t-1} + \Theta \varepsilon_t \quad (1.27)$$

dove  $\varepsilon_t \sim N(0, I)$  e  $c$  è una costante.

Se il vettore  $X_t$  racchiude tutte le variabili che potrebbero interessare agli investitori, allora il prezzo di un bond al tempo  $t$  sarà una funzione di delle variabili di

stato  $P_t(X_t)$  e soddisferà la seguente equazione

$$P_t^n(X_t) = E_t(m_{t+1}P_t^n(X_{t+1})) \quad (1.28)$$

dove  $m_{t+1}$  è il *price kernel*. Nei modelli di Ang e Piazzesi (2003) e di Joslin *et al.* (2014) il fattore di sconto stocastico viene definito come

$$m_{t+1} = \exp\left(-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'u_{t+1}\right) \quad (1.29)$$

dove  $\lambda_t$  sono i prezzi di mercato del rischio che variano nel tempo associati all'incertezza  $u_t$  e caratterizzano il comportamento degli investitori di fronte al rischio. I modelli affini della struttura a termine postulano inoltre che il parametro  $\lambda_t$  sia una funzione affine delle variabili di stato

$$\lambda_t = \lambda_0 + \lambda_1 X_t.$$

Tenendo in considerazione l'equazione (1.28), le dinamiche dei tassi di interesse (Eq.(1.26)) e del *price kernel* (Eq.(1.29)) si avrà che i prezzi dei bond saranno dati dalla seguente relazione

$$P_t^n(X_t) = \exp(A_n + B_n'X_t) \quad (1.30)$$

dove  $A_n$  e  $B_n$  sono equazioni ricorrenti pari a

$$A_n = A_{n-1} + B_{n-1}'(c - \Theta\lambda_0) + \frac{1}{2}B_{n-1}'\Theta\Theta'B_{n-1} - \delta_0 \quad (1.31)$$

$$B_n' = B_{n-1}'(\rho - \Theta\lambda_1) - \delta_{12}' \quad (1.32)$$

dove  $\delta_{12}' = (\delta_1, \delta_2)'$ ,  $A_1 = -\delta_0$  e  $B_1 = -\delta_1^2$ . Infine il rendimento di un titolo zero coupon  $r_t^n$  con scadenza pari a  $n$  periodi è calcolato in modo continuo

$$r_t^n = -\frac{\log(P_t^n(X_t))}{n} = \bar{A}_n + \bar{B}_n'X_t \quad (1.33)$$

---

<sup>2</sup>Per la dimostrazione completa si veda l'Appendice A di Ang e Piazzesi (2003).

dove  $\bar{A}_n = -A_n/n$  e  $\bar{B}'_n = -B'_n/n$ . I principali vantaggi derivanti dall'applicazione di questo modello sono che

1. si può osservare come l'intera *yield curve* risponda agli shock macroeconomici anziché ai singoli rendimenti dei bond,
2. tramite la scomposizione della varianza si può stimare la proporzione dei movimenti della struttura a termine attribuibile agli shock macro o alle variabili latenti,
3. si considera il premio per il rischio variabile nel tempo e come la sua dinamica sia influenzata dalle variabili macro e i fattori latenti,
4. le variabili macroeconomiche non sono influenzate dalle dinamiche della *yield curve*.

Su quest'ultimo punto Ang e Piazzesi (2003) evidenziano i principali problemi di carattere economico. Infatti l'assunzione di indipendenza contraddice l'evidenza empirica dei lavori di Estrella e Hardouvelis (1991), Mishkin (1990) e Estrella e Mishkin (1996) in cui la struttura a termine dei tassi di interesse predice i movimenti dell'attività economica. Inoltre nel modello è stato presupposto che la politica economica della banca centrale segua la regola di Taylor in un contesto in cui la sua politica impatta solo sui rendimenti e non sulle altre variabili macroeconomiche.

Jardet *et al.* (2013) si concentrano sulla relazione tra il premio a termine e l'attività economica avvalendosi dell'utilizzo di un modello Near-Cointegrated. I risultati presenti in letteratura evidenziano che i movimenti della *yield curve* di lungo periodo, ovvero la parte finale della curva, sono guidati principalmente dalle aspettative degli operatori e che il premio a termine è anticiclico ed è piuttosto stabile. Inoltre l'applicazione di un modello di cointegrazione impone l'ipotesi che i tassi siano non stazionari sebbene gran parte della letteratura Stock e Watson (1988), Campbell e Shiller (1991), indicano che siano molto persistenti ma stazionari. Su questi presupposti Jardet *et al.* (2013) utilizzano uno stimatore della media tra i coefficienti di un VAR e un modello di cointegrazione (CVAR).

Viene supposto che il processo generatore dei dati (DGP) per le variabili di stato sia

$$X_t = \mu^0 + \sum_{j=1}^p \Phi_j^0 X_{t-j} + \eta_t \quad (1.34)$$

dove  $X_t = (r_t, s_t, g_t)'$  è un vettore contenente i rendimenti dei bond di breve periodo  $r_t$ , lo spread  $s_t = R_t - r_t$  e  $g_t$  è il logaritmo del GDP infine  $\eta_t \sim N(0, \Omega^0)$  rappresenta il termine di errore.

Sulla base dell'equazione (1.34) vengono poi stimati un modello VAR e un modello VECM. La costante  $\mu$  e i coefficienti  $\Phi_j$  relativi ai ritardi di  $X_t$  di ogni modello vengono poi raccolti rispettivamente in un vettore  $\theta$ . Quindi  $\hat{\theta}_{t,cvar}$  è il vettore contenente le stime dei parametri del modello CVAR mentre  $\hat{\theta}_{t,var}$  è il vettore ottenuto con l'applicazione del VAR. Lo stimatore  $\hat{\theta}_t^{nc}(\lambda)$  del Near-Cointegrated VAR (NCVAR) si ottiene con la media degli stimatori del VAR e del CVAR nel seguente modo:

$$\hat{\theta}_t^{nc}(\lambda) = \lambda \hat{\theta}_{t,var} + (1 - \lambda) \hat{\theta}_{t,cvar} \quad (1.35)$$

dove  $\lambda \in [0, 1]$  è ottenuto minimizzando l'errore di previsione associato alla parte delle aspettative. Preso in considerazione il rendimento di un bond con  $h$  periodi di scadenza al tempo  $t$ , indicato con  $R_t(h)$ , la sua componente descritta dalle aspettative sarà definita da  $EX_t(h) = -1/h \log \tilde{P}_t(h)$  dove  $\tilde{P}_t(h) = E_t[\exp(-r_t - r_{t+1} \dots - r_{t+h-1})]$ . Il parametro  $\lambda$  sarà quindi la soluzione del seguente problema

$$\lambda^* = \arg \min_{\lambda \in [0,1]} \sum_{i=1}^t [P_t^{obs} - \tilde{P}_t(\lambda, h)]^2 \quad (1.36)$$

dove  $P_t^{obs}$  sono le osservazioni ottenute con  $\exp(-\sum_{i=0}^{h-1} r_{t+i})$ , mentre  $\tilde{P}_t(\lambda, h)$  sono le previsioni ottenute con l'applicazione del NCVAR usando i parametri dell'equazione (1.35).

Riprendendo il modello VAR in assenza di arbitraggio quindi, considerando la presenza di un fattore di sconto stocastico come nell'equazione (1.29) e l'equazione (1.33), avremo

$$R_t(h) = -\frac{\log(P_t(h)(X_t))}{n} = \bar{A}_n + \bar{B}'_n X_t \quad (1.37)$$

dove la variabile di stato  $X_t$  segue il processo descritto nell'equazione (1.34) e i coefficienti  $\bar{A}_n$  e  $\bar{B}'_n$  sono equazioni alle differenze composte dai parametri ottenuti con il modello NCVAR. Tenendo conto che i rendimenti dei bond di lungo periodo ( $R_t(h)$ ) sono composti dalle aspettative sui tassi di interesse a breve termine futuri ( $EX_t(h)$ ) e da un premio a termine ( $TP_t(h)$ ) e che lo spread ( $S_t(h)$ ) è dato dalla differenza tra i rendimenti di lungo periodo e quello a breve ( $r_t$ ) otteniamo

$$S_t(h) = EX_t(h) - r_t + TP_t(h) \quad (1.38)$$

dove  $EX_t(h) = EX_t(h) - r_t$  rappresenta le aspettative sui tassi di interesse futuri. Le stime di questo modello evidenziano che il premio a termine è guidato principalmente dalla variabilità dello spread, la componente delle aspettative decresce debolmente durante i periodi di recessione, quindi la sequenza dei tassi di interesse futuri non è influenzata dalle condizioni economiche correnti. Infine gli autori evidenziano che il premio a termine non è anticiclico, come ipotizzato in precedenza, bensì risulta seguire il ciclo economico.

Di particolare rilevanza sono i risultati ottenuti da Giese (2008) la quale, effettuando i test sul rango di cointegrazione, conclude che gli spread non risultano stazionari. Inoltre, eseguendo test sul rango di cointegrazione, le serie dei rendimenti hanno mostrato una chiara tendenza a seguire due trend comuni: ciò risulta in pieno contrasto con la teoria delle aspettative. Per ovviare a questo problema, Giese (2008) propone l'utilizzo di diverse combinazioni lineari quali le differenze pesate degli spread, o "butterfly" spread, che risultano essere stazionarie e sono utili per riuscire a comprendere meglio la formazione delle aspettative sui tassi di interesse.

Altri lavori hanno analizzato la valenza della teoria delle aspettative sui tassi di interesse con scadenze giornaliere o settimanali. Le motivazioni che hanno guidato Longstaff (2000), Della Corte *et al.* (2007), Gunnella (2015) ad analizzare la *yield curve* nel solo tratto iniziale sono principalmente due. La prima è che i tassi a brevissima scadenza sono più rappresentativi rispetto ai tassi dei titoli di stato poiché presentano un minor grado di rischio, se non addirittura nullo. La seconda

motivazione è che se la teoria non si dimostra valida sul tratto iniziale della curva allora non potrà certo essere valida sugli interessi a più lunga scadenza.

Infine, alcuni autori si sono invece incentrati sulla verifica empirica delle implicazioni della teoria delle aspettative attraverso le *cross-equation constrains*. Gonzalo e González (2000) ipotizzano che il premio a termine sia variabile nel tempo e che ciò consenta di spiegare le deviazioni transitorie dalle condizioni di equilibrio. Infatti se si impone che il premio a termine  $\mu_t$  segua una differenza di martingala, ciò non comporta alcuna modifica alle condizioni di lungo periodo ma solo sulle restrizioni “high frequency”. Diversamente dall’equazione (1.10), che rappresenta l’*excess return*  $\xi_t$ , si ottiene

$$\xi_t = S_t - \frac{1}{\delta}S_{t-1} + \Delta r s_t + \mu_t. \quad (1.39)$$

Questa formulazione alternativa del *present value*, è detta *inexact present value model* (Gonzalo e González, 2000); da questa formulazione è possibile verificare alternativamente le implicazioni della teoria delle aspettative utilizzando lo stesso modello VAR dell’equazione (1.11), ma imponendo le seguenti restrizioni

$$g'(I - \delta A)A = \delta h' A^2. \quad (1.40)$$

Sullo stesso argomento Cao *et al.* (2014) ricavano le *cross-equation constrains* da un modello VAR Markov-switching (MS-VAR). Dato che una delle possibili motivazioni del fallimento della teoria delle aspettative è dovuta dal cambio di regimi, questi ricercatori assumono che il processo generatore dei dati sia *regime-switching*.

I risultati ottenuti da Gonzalo e González (2000), Cao *et al.* (2014) e Gunnella (2015) mostrano che mentre le implicazioni della teoria delle aspettative sui tassi di interesse non trovano validità empirica con l’applicazione delle restrizioni del modello *present value*, con la versione più “allargata” le implicazioni vengono supportate empiricamente.

### 1.5.2 Modelli con aggiustamento a soglia

Spesso nell'analisi delle serie macroeconomiche come i tassi di cambio o i tassi di crescita della produzione industriale vengono adottati modelli più sofisticati. Limitarsi al solo uso di modelli lineari spesso risulta una strategia non ottimale e può condurre a problemi di diverso genere, in quanto non si tiene conto di fattori come i costi di transazione, il ciclo economico e gli interventi di politica economica. In questo filone della letteratura troviamo Skalin e Terasvirta (1999), Lo e Zivot (2001), Bec e Rahbek (2004), Kapetanios e Shin (2008), Beckmann (2013), che dimostrano empiricamente che gli strumenti più efficaci a cogliere le dinamiche di queste variabili macroeconomiche sono i modelli non lineari. Balke e Fomby (1992) sono i primi autori ad utilizzare la cointegrazione in ambito non lineare, quindi abbandonare l'ipotesi che l'aggiustamento sia sempre presente e che l'equilibrio sia unico.

Anche nella struttura a termine dei tassi di interesse risultano presenti elementi che possono giustificare l'utilizzo di tecniche più complesse. Seo (2003) spiega che la presenza dei costi di transazione nei mercati finanziari compromette il corretto funzionamento del modello ipotizzato da Campbell e Shiller (1987). Infatti, nel momento in cui i tassi divergono dalla condizione di equilibrio, gli operatori valutano la possibilità di poter sfruttare o meno le opportunità di arbitraggio in base alla presenza di questi costi. Si viene così a creare una "banda" al cui interno gli operatori rimangono pressoché inattivi, quindi le traiettorie dei tassi divergono da quelle dettate dalla teoria delle aspettative; al di fuori di questa "banda", invece, gli operatori agiscono riportando il sistema attorno all'equilibrio.

Balke e Fomby (1992) sostengono che la non linearità del meccanismo di correzione dell'errore sia causata dall'intensità delle azioni della politica monetaria. Infatti, sono i *policy maker* a condurre interventi sul differenziale a termine dei tassi di interesse, ma la loro azione è mirata non tanto a riportare questa variabile ad un preciso valore quanto ad una "zona obiettivo". Quindi la banca centrale interviene solo quando il differenziale a termine dei tassi di interesse assume valori molto ampi che possono compromettere l'efficacia delle sue manovre. Un valore dello spread ampio può condurre gli agenti ad operare in modo del tutto opposto a quello ideale

per una corretta politica monetaria. Lo spread viene lasciato fluttuare liberamente all'interno di una "banda" di valori prefissati e quando invece supera valori che sono al di fuori di questo range, i *policy maker* intervengono riportando il differenziale a termine all'interno della zona obiettivo. Balke e Fomby (1992) sostengono che una funzione a soglia sia ideale per approssimare il comportamento della banca centrale di cercare di "appiattire" ("*smoothing*") la *yield curve*.

I modelli lineari come l'ECM non si prestano ad approssimare le dinamiche di un aggiustamento che avvengono in modo discontinuo poiché impongono restrizioni condizioni sul meccanismo di correzione dell'errore. In particolare, come descritto da Escribano (2004), un modello di cointegrazione lineare impone che:

1. l'equilibrio di lungo periodo sia unico;
2. l'aggiustamento attorno l'equilibrio sia simmetrico;
3. la correzione sia sempre presente ed avvenga in proporzione costante al precedente discostamento dal equilibrio.

Lo strumento che più si presta a delineare le caratteristiche di questi processi è il modello a soglia. Introdotto per la prima volta da Balke e Fomby (1992) nella struttura a termine dei tassi di interesse, essi utilizzano una funzione a soglia per approssimare un meccanismo di aggiustamento che si attiva solo quando l'errore di equilibrio del periodo precedente assume valori che sono al di fuori di una data banda. Per cui le deviazioni dall'equilibrio siano *mean reverting* fuori da un determinato intervallo di valori, mentre possiedono radice unitaria se sono all'interno dello stesso intervallo.

Ad esempio, partendo da un sistema cointegrato bivariato in cui  $x_t$  e  $v_t$  sono le variabili che consideriamo, la relazione di equilibrio è descritta dal seguente processo stazionario:

$$z_t = v_t - \beta' x_t$$

dove  $z_t$  è la relazione di equilibrio di lungo periodo con vettore di cointegrazione pari a  $(1, \beta)'$ . Assumendo che tale relazione segua un processo SETAR (Self-Exciting

Threshold Autoregressive) allora:

$$z_t = \rho^{(i)} z_{t-1} + \epsilon_t$$

il cui coefficiente  $\rho^{(i)}$  dipende a sua volta dalle passate realizzazioni di  $z_t$ , in particolare,

$$\rho^{(i)} = \begin{cases} 1 & \text{se } |z_{t-1}| \leq \phi \\ |\rho| < 1 & \text{se } |z_{t-1}| > \phi \end{cases}$$

dove  $\phi$  rappresenta il valore soglia.

Fintanto che  $z_{t-1}$  assumerà valori che sono all'interno della banda  $[-\phi; \phi]$  il processo sarà un *random walk*, quindi non vi è alcuna relazione di equilibrio. Quando  $z_{t-1}$  assume valori che stanno al di fuori dell'intervallo, il processo sarà *mean reverting*. Saremo quindi in presenza di cointegrazione e le variabili  $x_t$  e  $v_t$  si muoveranno attorno all'equilibrio. Il modello sopra descritto impone che il valore dell'equilibrio sia unico. Se invece fossimo più interessati ad un modello che riporti il sistema attorno ad una zona obiettivo allora la formulazione assumerà una diversa connotazione

$$z_t = \begin{cases} z_{t-1} + \epsilon_{1t} & \text{se } |z_{t-1}| \leq \phi \\ \phi^{(i)}(1 - \rho) + \rho z_{t-1} + \epsilon_{2t} & \text{se } |z_{t-1}| > \phi, \end{cases}$$

dove  $\phi^{(i)} = \phi$  se  $s_{t-1} > \phi$  mentre  $\phi^{(i)} = -\phi$  se  $s_{t-1} < -\phi$ . In questo caso la serie tenderà a ritornare nell'intervallo  $[-\phi; \phi]$  anziché verso la media. Il meccanismo di errore risponde solo quando la variabile assume valori al di fuori della banda e la forza di correzione dipenderà da quanto la variabile è lontana dagli estremi dell'intervallo.

Riprendendo la struttura a termine dei tassi di interesse e considerando l'ambito multivariato, quest'ultimo meccanismo può essere sostituito nell'equazione (1.5) in modo da ottenere il modello Band-TVECM che si presenta

$$\Phi(L)\Delta Y_t = g(S_{t-1}) + \epsilon_t$$

dove  $Y_t$  è un vettore contenente i tassi di interesse a diverse scadenze,  $\Phi(L)$  rappresenta

il polinomio nell'operatore ritardo,  $S_t$  è la combinazione lineare dei tassi di interesse ( $\beta'Y_t$ ), cioè lo spread, mentre  $g(\cdot)$  rappresenta la funzione di aggiustamento a soglia data da

$$g(S_t) = \begin{cases} \phi_1(1 - \rho_1) + \rho_1 S_t & \text{se } S_t < \phi_1 \\ 0 & \text{se } \phi_1 < S_t < \phi_2 \\ \phi_2(1 - \rho_2) + \rho_2 S_t & \text{se } S_t > \phi_2, \end{cases}$$

dove  $\phi_1$  e  $\phi_2$  rappresentano rispettivamente il limite inferiore e superiore della banda.

Innumerevoli sono i lavori che confermano a livello empirico la presenza di un meccanismo a soglia (Balke e Fomby (1992), Enders e Granger (1998), Bec e Rahbek (2004), Seo (2003), Krishnakumar e Neto (2009), Cai *et al.* (2015), ecc.). In tutti i lavori condotti sui dati statunitensi, i test per verificare l'effetto soglia supportano l'ipotesi che vi sia *mean reversion* non lineare nella struttura a termine.

Inoltre i risultati ottenuti applicando il modello a soglia sono piuttosto consistenti con la teoria delle aspettative. Da sottolineare che la teoria delle aspettative fornisce una descrizione ben specifica riguardo la relazione di lungo periodo, ma non dice nulla riguardo al meccanismo di aggiustamento. Con l'applicazione di un modello di cointegrazione lineare, Campbell e Shiller (1987) si erano trovati di fronte al fatto che: la teoria si dimostrava valida solo per i tassi di interesse a breve, ma non per quelli a lungo termine. L'applicazione del modello a soglia fornisce risultati in linea con quelli dettati dalla teoria. Infatti, i coefficienti del modello relativi al regime esterno sono coerenti con la teoria delle aspettative: lo spread prevede la giusta direzione dei movimenti dei tassi di interesse sia a breve che a lungo periodo, cioè i movimenti dei tassi di interesse risultano entrambi correlati positivamente con lo spread.

Sebbene il modello di cointegrazione a soglia fornisca una buona descrizione empirica della teoria delle aspettative, non è esente da problemi di natura tecnica. Quando si ha a che fare con modelli non lineari, i problemi principali sorgono nella

trattazione delle proprietà degli stimatori e della stazionarietà. In ambito lineare è possibile determinare se un processo sia stazionario o meno; lo stesso non si può dire dei modelli non lineari. Determinare la stazionarietà di questi ultimi richiede una metodologia alternativa. Bec e Rahbek (2004) hanno collegato questi processi alla teoria riguardante le catene di Markov al fine di poter utilizzare i risultati di Meyn e Tweedie (1993). Nello specifico gli autori si avvalgono dell'ergodicità geometrica per sostituire la definizione classica di stazionarietà e stabiliscono le condizioni per cui un processo possa essere definito geometricamente ergodico e, quindi, assicurare l'esistenza di momenti e la sua stazionarietà. Altre problematiche scaturiscono dal fatto che i coefficienti stimati hanno distribuzioni non standard e la teoria asintotica non fornisce una distribuzione limite soprattutto quando il vettore di cointegrazione è incognito.

L'inserimento di una funzione continua a tratti provoca sostanziali problemi nelle stime. Infatti, non è possibile utilizzare algoritmi convenzionali a causa della presenza di una funzione discontinua all'interno del modello. Una soluzione riguardo a questo problema viene fornita da Hansen e Seo (2002) che descrivono una metodologia di ricerca a griglia per ottenere una stima con il metodo della massima verosimiglianza solo per il caso di un modello di cointegrazione a soglia bivariato. Nel caso di un sistema con più di due variabili questa tecnica risulta meno efficace, per cui si richiedono algoritmi ancora più complessi<sup>1</sup>.

### 1.5.3 Modelli con polinomi cubici

Come descritto nel paragrafo precedente, nell'analisi delle serie economiche, spesso ci si imbatte in fenomeni (come i costi di aggiustamento non simmetrici, la dipendenza dal ciclo economico, i cambiamenti di regime associati ad interventi politici) che non possono essere approssimati da un modello lineare. Gran parte della letteratura è incentrata quasi esclusivamente sulla famiglia dei modelli a soglia. Questa scelta deriva dal fatto che questa tipologia di modelli consente una più facile interpretazione

---

<sup>1</sup>Per una trattazione più approfondita dell'utilizzo di algoritmi genetici per le stime multivariate si veda Yang *et al.* (2007) e El-Shagi (2011).

a livello economico e permette una più concreta definizione dei regimi in cui scatta il meccanismo di aggiustamento. Allo stesso tempo, i modelli a soglia soffrono di una mancata flessibilità per quanto riguarda il meccanismo di aggiustamento che impone una struttura rigida riguardo alla regione di aggiustamento. Altre difficoltà si riscontrano soprattutto nell'ambito della cointegrazione, nell'analisi di due o più processi cointegrati e nel determinare le distribuzioni dei parametri e le loro proprietà asintotiche.

Una strada alternativa viene esplorata da Escribano (2004) per lo studio della domanda di moneta attraverso l'utilizzo di una funzione polinomiale. La giustificazione dell'uso di queste funzioni risiede nel fatto che le forme polinomiali cubiche sono molto flessibili ed utili nell'approssimare forme funzionali incognite. Inoltre, i polinomi cubici permettono di individuare possibili asimmetrie, punti di soglia, o la presenza eventuale di uno o più equilibri. In secondo luogo questi modelli possono essere utilizzati per verificare empiricamente le condizioni di stabilità dell'aggiustamento non lineare. Infine le approssimazioni polinomiali cubiche possono essere utilizzate come strumento decisionale per selezionare il modello parametrico non lineare più adeguato.

Anche Escribano e Mira (2002) analizzano il problema della stazionarietà in un contesto di non linearità. Gli autori utilizzano le definizioni di  $\alpha$ -mixing e di *near epoch dependence* (NED) per estendere il concetto  $I(0)$  anche ai modelli non lineari.

**Definizione 1** ( $\alpha$ -mixing). *Sia  $\{v_t\}$  una sequenza di variabili casuali e sia  $\mathcal{F}_s^t \equiv \sigma(v_s, \dots, v_t)$  una sigma algebra, si definisce i coefficienti  $\alpha$ -mixing come*

$$\alpha_m \equiv \sup_t \sup_{A \in \mathcal{F}_{-\infty}^t, B \in \mathcal{F}_{t+m}^{+\infty}} |P(A \cap B) - P(A)P(B)| \quad (1.41)$$

*Il processo  $\{v_t\}$  è detto  $\alpha$ -mixing (o strong mixing) se  $\alpha_m \rightarrow 0$  quando  $m \rightarrow \infty$ . Se  $\alpha_m = O(m^\lambda)$  per ogni  $\lambda < -a$  allora  $\alpha_m$  si dice di ampiezza  $-a$ .*

Il coefficiente  $\alpha_m$  misura la dipendenza tra due eventi del processo separati da almeno  $m$  periodi di tempo.

Collegata a questa definizione troviamo il concetto di NED.

**Definizione 2** (NED). Sia  $\{w_t\}$  una sequenza di variabili casuali con  $E(w_t^2) < \infty$  per ogni  $t$ .  $\{w_t\}$  è near epoch dependent (NED) su una sequenza strong mixing  $\{v_t\}$  di ampiezza  $-a$  se  $\phi(n)$  è di ampiezza  $-a$ , dove  $\phi(n)$  è dato da

$$\sup_t \|w_t - E_{t-n}^{t+n}(w_t)\|_2 \equiv \phi(n) \quad (1.42)$$

dove  $E_{t-n}^{t+n}(w_t) = E(w_t | v_{t-n}, \dots, v_{t+n})$  e  $\|\cdot\|_2$  è la norma  $L_2$  di una variabile casuale definita come  $E^{1/2}|\cdot|^2$ .

Queste due definizioni sono collegate alle nozioni di “short memory in distribution” e di “short memory in mean” (Granger e Terasvirta, 1993) e permettono quindi l’utilizzo del teorema del limite centrale e della legge dei grandi numeri.

Da questo presupposto Escribano e Mira (2002) forniscono un teorema per un modello non lineare che implichi la cointegrazione. In particolare definiscono le condizioni sufficienti sia per le serie che per la funzione di aggiustamento. Il teorema infatti stabilisce che la stabilità di un processo viene rispettata se la funzione (differenziabile) soddisfa la condizione generale di Lipschitz ( $-2 < dg(z_t; \gamma)/dz_t < 0$ ).

In modo del tutto analogo a quanto descritto nei paragrafi precedenti, un modello NEC (*Nonlinear Error Correction*) secondo Escribano (2004) è definito in questo modo:

$$\Phi(L)\Delta Y_t = g(z_{t-1}; \gamma) + \epsilon_t, \quad (1.43)$$

dove  $Y_t$  è il vettore delle variabili osservate,  $\Phi(L)$  rappresenta il polinomio nell’operatore ritardo,  $g(\cdot)$  è una funzione polinomiale del tipo  $\sum_{i=1}^3 c_i z^i$ ,  $z_t$  è il vettore contenente le combinazioni lineari tra le variabili cioè  $\beta'Y_t$ , mentre  $\gamma$  è un vettore contenente tutti i coefficienti relativi la funzione non lineare ed infine  $\epsilon_t \sim i.i.d.(0, \sigma^2)$  rappresenta il vettore dei disturbi.

Nell’applicazione empirica sull’analisi della moneta in UK, Escribano argomenta che le forme funzionali razionali sono le più idonee nell’approssimare funzioni incognite soprattutto perché è facile verificare la condizione di stabilità dettate dal suo teorema. Quindi egli si avvale dell’approssimante di Padé<sup>3</sup> per la sua analisi che si

<sup>3</sup>Metodo di approssimazione di funzioni analitiche con funzioni polinomiali razionali e può essere vista come una generalizzazione della serie di Taylor.

presenta come

$$g(z, \gamma) = \sum_{i=1}^n c_i z^i = \frac{d_0 + d_1 z + d_2 z^2 + \dots + d_L z^L}{1 + e_1 z + e_2 z^2 + \dots + e_M z^M}$$

Un altro vantaggio che deriva dall'utilizzo di forme funzionali cubiche è che possono essere facilmente stimate con il metodo dei minimi quadrati (OLS) o con il metodo dei minimi quadrati non lineari (NLS) ed inoltre consente di verificare la presenza di non linearità attraverso un test con i moltiplicatori di Lagrange (LM). Come dimostrato da Escribano (2004) nella sua ricerca sulla domanda di moneta, questo modello in ambito univariato risulta notevolmente più adatto nell'analisi dei dati che presentano segni di non linearità. Le difficoltà più importanti sopraggiungono quando si vuole contestualizzare il modello nell'ambito del multivariato. Uno dei problemi maggiori è chiaramente la difficoltà di generalizzare il polinomio con l'approssimante di Padé in ambito multivariato. Nonostante il fatto che si possa utilizzare tecniche di stima semplici come il metodo dei minimi quadrati, questo modello richiede un elevato numero di parametri da stimare.

#### 1.5.4 Modello con aggiustamento approssimato da polinomio di Taylor

I lavori condotti con i modelli a soglia mostrano chiaramente che il meccanismo di cointegrazione più adatto a rappresentare a livello empirico la teoria delle aspettative sia quello non lineare. Tra i lavori condotti sull'analisi dei tassi di interesse americani attraverso un meccanismo di aggiustamento non lineare simile a quello di Escribano, troviamo il lavoro di Lucchetti e Palomba (2009). Il loro obiettivo è quello catturare gli interventi dei *policy maker* attraverso una forma più flessibile rispetto ai modelli a soglia poiché uno dei problemi principali dei modelli con aggiustamenti a tratti sono le condizioni che si impongono sulla regione in cui non avviene l'aggiustamento. Infatti, per ragioni legate soprattutto agli aspetti computazionali, una funzione di aggiustamento univariata a soglia impone che la regione sia un segmento, mentre,

nel caso multivariato, potrebbe essere un parallelogramma, un cerchio o un'ellisse. Per questo motivo utilizzare la metodologia di Escribano (2004) risulta consono allo scopo dell'analisi sebbene sia molto più appetibile estendere questa metodologia nel ambito multivariato. Per questo Lucchetti e Palomba (2009) hanno deciso di utilizzare come funzione polinomiale la serie di Taylor per approssimare il meccanismo di aggiustamento.

L'espansione della serie di Taylor rappresenta un metodo di approssimazione per funzioni incognite non lineari e, sebbene la metodologia introdotta da Balke e Fomby (1992) sia idonea a rappresentare il comportamento dei policy maker, il modello di Lucchetti e Palomba non si discosta del tutto da questo filone di letteratura. Infatti, una forma polinomiale cubica consente di approssimare in modo consistente una funzione a soglia - come mostrato nella Figura (1.2) - permette di individuare possibili asimmetrie e i punti di equilibrio.

I vantaggi che si possono trarre con l'utilizzo dell'espansione della serie di Taylor sono principalmente di carattere analitico. Il modello gode di una maggiore flessibilità in quanto, non impone alcuna restrizione sulla funzione di aggiustamento. La formulazione del meccanismo di aggiustamento può essere facilmente implementata in un contesto multivariato. Quindi è possibile analizzare un numero di processi cointegrati maggiore di due. Inoltre il modello rimane lineare nei parametri e mantiene in modo ridotto il numero dei coefficienti da stimare.

Riprendendo la formulazione per un modello NEC (*Nonlinear Error Correction*) (1.43), il modello di Lucchetti e Palomba (2009) presenta in modo del tutto generale come:

$$\Phi(L)\Delta Y_t = g(S_{t-1}) + \varepsilon_t \quad (1.44)$$

dove il vettore dove  $Y_t$  è composto da  $(rs_t, rm_t, rl_t)'$  i cui elementi rappresentano rispettivamente i tassi a breve, medio e lungo periodo. Il vettore  $S_t$  è costituito da due spread  $(sm_t, sl_t)'$  che sono definiti rispettivamente come la differenza tra i rendimenti di medio e breve termine ( $sm_t = rm_t - rs_t$ ) e tra i rendimenti tra il lungo e breve termine ( $sl_t = rl_t - rs_t$ ); infine  $\varepsilon_t \sim N(0, \Sigma)$  è l'innovazione. La funzione  $g(\cdot)$  è la serie di potenze di Taylor, quindi

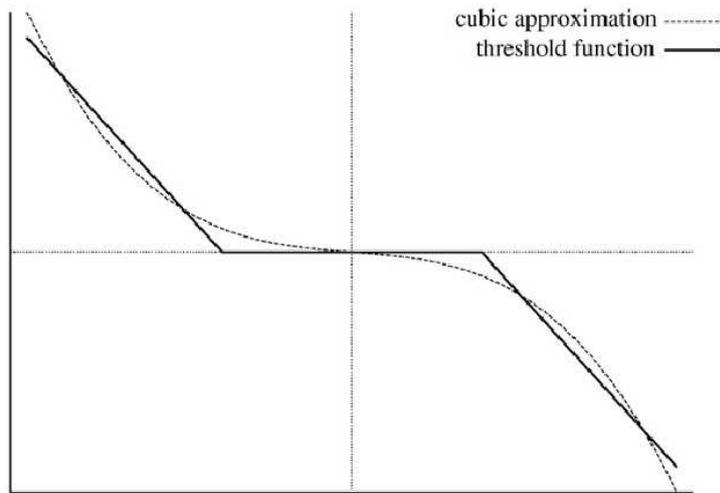


Figura 1.2: Approssimazione cubica di una funzione a soglia

$$g(x) = \frac{\sum_{i=0}^n f^i(x)}{i!} (x - x_0)^i$$

Applicato nel contesto multivariato alle serie dei tassi di interesse si otterrà

$$g(S_t) = \mu + \alpha' S_t + \theta' (S_t \otimes S_t)_+ + \lambda' (S_t \otimes S_t \otimes S_t)_+$$

dove  $\otimes$  è il prodotto di Kronecker mentre  $(\cdot)_+$  è una simbologia adottata dagli autori che sta ad indicare l'inverso di Moore-Penrose della matrice di duplicazione (Magnus e Neudecker, 1988). Nello specifico  $(\cdot)_+$  è un operatore che elimina tutti gli elementi ripetuti. Applicato al terzo termine dell'equazione  $(S_t \otimes S_t)$ , l'operatore sarà uguale alla matrice

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

quindi in modo compatto si otterrà

$$S_t = [sm_t \quad sl_t]'$$

$$Q_t = (S_t \otimes S_t)_+ = [sm_t^2 \quad sm_t \cdot sl_t \quad sl_t^2]'$$
(1.45)

$$C_t = (S_t \otimes S_t \otimes S_t)_+ = [sm_t^3 \quad sm_t^2 \cdot sl_t \quad sm_t \cdot sl_t^2 \quad sl_t^3]'$$

$S_t$ ,  $Q_t$  e  $C_t$  sono rispettivamente i termini di primo, secondo e terzo termine della serie di Taylor.

I termini  $Q_t$  e  $C_t$  non trovano una diretta interpretazione economica ma sono necessari per approssimare la funzione a soglia. Un'altra tipologia di problemi da considerare riguarda le proprietà asintotiche del modello. Gli usuali teoremi riguardanti le distribuzioni asintotiche e la stazionarietà del processo non possono essere applicati. Sebbene in letteratura siano presenti diversi teoremi per alcuni modelli non lineari, come ad esempio quello di Escribano e Mira (2002), molti non possono essere applicati al modello NEC perché non rispetta alcune delle condizioni.

## Capitolo 2

# Rassegna Metodologica

### 2.1 Introduzione

La teoria delle aspettative sulla struttura a termine dei tassi di interesse è stata al centro di molti dibattiti. Sebbene la letteratura econometrica si sia mossa in diverse direzioni per confermare o per confutare tale teoria, tra i modelli che hanno avuto un maggior impatto sulla sua evidenza sono quelli a soglia (Balke e Fomby (1992), Enders e Granger (1998), Bec e Rahbek (2004) Seo (2003), Krishnakumar e Neto (2009), Cai *et al.* (2015)).

Come accennato in precedenza, le limitazioni più evidenti dei modelli a soglia si riscontrano in ambito inferenziale e in quello analitico. Infatti, quando si considerano più di due variabili, il modello richiede tecniche piuttosto complesse in sede di stima e diventa più difficile da generalizzare. Un altro limite che si riscontra con i modelli a soglia è come definire la forma della regione in cui non avviene l'aggiustamento. Naturalmente nell'ambito univariato questa regione è rappresentata da una porzione di retta mentre, in ambito multivariato non può essere dato per scontato che sia un parallelogramma, un'ellisse oppure da un qualche insieme chiuso e limitato.

Rimanendo sempre nel contesto della non linearità, un modello che permette un più alto grado di flessibilità e facilità di stima è il modello NEC di Lucchetti e Palomba (2009). L'ipotesi principale da cui partono questi autori è che un polinomio di terzo grado inserito nel meccanismo di aggiustamento fornisce una buona

approssimazione di una funzione di aggiustamento a tratti come quella del modello a soglia. I vantaggi che derivano dall'applicazione sono:

1. assenza di un vincolo o di una soglia sulla forma funzionale dell'aggiustamento della zona in cui non avviene l'aggiustamento;
2. Il modello è lineare nei parametri per cui è possibile applicare i metodi standard di stima (OLS, ML, ecc.);
3. Nel contesto multivariato l'aggiustamento avviene sfruttando le potenze degli spread ed i loro prodotti incrociati.

Sebbene la scelta di utilizzare il modello NEC per approssimare le politiche monetarie più incisive dell'ultimo trentennio risulti ottimale, ci sono alcune considerazioni da fare.

Nonostante la maggior capacità dei modelli non lineari, rispetto a quelli lineari, di catturare dinamiche complesse presenti in economia, la teoria asintotica per questi modelli non è completa, in quanto la natura complessa, sia della componente non lineare che della non stazionarietà delle serie, ha condotto ad ampliare o a sviluppare nuovi risultati sulla legge dei grandi numeri (LLN) e sul teorema del limite centrale (CLT). Inoltre non c'è un teorema generale valido per tutti i modelli non lineari. I principali risultati asintotici presenti in letteratura sono stati sviluppati sotto diverse condizioni sulla funzione di aggiustamento, sull'ipotesi che il vettore di cointegrazione sia noto oppure no, sulla presenza o sull'assenza di un trend lineare o sul metodo di stima. In particolare molti dei teoremi asintotici presenti in letteratura sono stati ideati per la classe dei modelli a soglia (Hansen e Seo (2002), Bec e Rahbek (2004), Kristensen e Rahbek (2013) ecc.).

Nell'ambito delle serie macroeconomiche e/o finanziarie un fattore da considerare è l'eteroschedasticità condizionale e non spesso rilevata empiricamente (si veda ad esempio Hamori e Tokihisa (1997)).

Hamori e Tokihisa (1997), Kim *et al.* (2002b), Cavaliere (2005), Cavaliere *et al.* (2010b) e Boswijk *et al.* (2016) hanno dimostrato l'impatto di tale fattore sulle di-

stribuzioni asintotiche di alcuni dei più consueti test come quelli per la stazionarietà e per il rango di cointegrazione. Una possibile soluzione a questo problema è l'implementazione della procedura *wild bootstrap* nei test come suggerito da Cavaliere *et al.* (2010b).

Questo capitolo è incentrato su alcuni dei più importanti risultati teorici che possono trovare una qualche applicazione al modello NEC e su alcune tecniche di simulazione utili in questo contesto. Lo scopo è quello di verificare e di ricercare la metodologia più adatta quando si utilizza il modello NEC sia in un contesto generale che nell'applicazione ai tassi di interesse americani. Ciò comprende l'utilizzo di strumenti più accurati sia nella fase preliminare che in quella di specificazione e che tengano in considerazione sia la natura dei dati che il modello scelto.

Nella prima parte di questa sezione verranno riportati alcuni dei più importanti risultati asintotici presenti in letteratura riguardo ai modelli non lineari, con particolare riguardo all'analisi delle proprietà asintotiche degli stimatori per i modelli di cointegrazione con aggiustamento non lineare e alla distribuzione dei test di significatività, dei moltiplicatori di Langrange (LM), di Wald (W) e del rapporto di verosimiglianza (LR).

Dato che in alcuni contesti non è corretto utilizzare le usuali distribuzioni standard, nella seconda parte verranno elencate alcune delle tecniche di simulazione che permettono una più accurata analisi riguardo all'esecuzione dei test.

In prima istanza farò un'ampia panoramica sulle più importanti tecniche bootstrap utilizzate in letteratura. Poi mi soffermerò sulle procedure applicabili nell'ambito delle serie storiche che tengono in considerazione sia la non linearità dell'aggiustamento che la cointegrazione. In seguito mi soffermerò sui metodi di simulazione che possono essere applicati in un contesto in cui vi sia eteroschedasticità (condizionale e non). Infine farò un elenco delle procedure applicabili al modello NEC che poi utilizzerò nel capitolo successivo per le stime.

Prima di proseguire farò un breve elenco di alcune notazioni che utilizzerò nel

resto del capitolo:  $\|\cdot\|$  indica la norma Euclidea e  $B(\cdot)$  indica il moto browniano che si distribuisce come  $\int_0^1 B(r)dr \sim N(0, 1/3)$ . Per quanto riguarda le convergenze “ $\xrightarrow{W}$ ” sta ad indicare quella debole nello spazio di funzioni, “ $\xrightarrow{P}$ ” la convergenza in probabilità mentre “ $\xrightarrow{d}$ ” in distribuzione. Data una matrice  $X$  di dimensione  $n \times k$  di rango  $k$ , si definisce il suo complemento ortogonale la matrice  $X_\perp$  di dimensione  $n \times (n - k)$  di rango  $n - k$  tale per cui  $X'X_\perp = 0$ .

## 2.2 Proprietà dei modelli di cointegrazione lineare

Nei modelli di cointegrazione lineare l’inferenza statistica è piuttosto complicata in quanto bisogna considerare le variabili non stazionarie. Con i processi integrati le condizioni del teorema del limite centrale non sono più valide, per cui è necessario un nuovo approccio per poter esplorare le proprietà asintotiche. Il più importante risultato è quello ottenuto da Donsker (1951) che collega i moti browniani ai processi non stazionari ottenendo una versione *ad hoc* per questi ultimi del teorema del limite centrale.

Infatti considerato un processo integrato di ordine uno

$$x_t = x_{t-1} + u_t = \sum_{s=1}^t u_s \quad (2.1)$$

con  $x_0 = 0$  e  $u_t \sim i.i.d.(0, \sigma_u^2)$ , può essere dimostrato che una sua trasformazione converge a un moto browniano.

**Teorema 1** (Functional Central Limit Theorem (FCLT)). *Dato il processo stocastico  $X_T$  definito da*

$$X_T(r) = \frac{1}{\sigma_u \sqrt{T}} x_{[Tr]} = \frac{1}{\sigma_u \sqrt{T}} \sum_{t=1}^{[Tr]} u_t \quad (2.2)$$

con  $[Tr]$  che indica la più grande parte intera che non supera  $Tr$ , dove  $r$  è compreso nell’intervallo  $[0, 1]$ .

Allora  $X_T(r) \xrightarrow{d} B(r)$ .

Inoltre è possibile dimostrare che per una funzione  $h(\cdot)$  continua nell'intervallo  $[0, 1]$  otteniamo

$$h(X_T(r)) \xrightarrow{d} h(B(r)). \quad (2.3)$$

Quest'ultima equazione deriva dal teorema della mappatura continua di Billingsley (1968). Quindi la FCLT collegata all'equazione (2.3) ci assicura che i processi integrati abbiano come limite una distribuzione (o una funzione della) Gaussiana.

La differenza più evidente tra la CLT per i processi stazionari e la FCLT per le serie integrate è che nei primi i momenti campionari convergono ad una costante mentre per i secondi convergono a variabili casuali. Esistono diverse versioni del FCLT sotto condizioni meno forti dell'indipendenza. Nel corso del capitolo verranno enunciate alcune estensioni di questo teorema alle sequenze delle differenze di martingala e ai processi *strong mixing*.

Riguardo invece alle proprietà asintotiche dei processi cointegrati lineari è importante subito stabilire alcuni risultati utili della regressione statica. Considerando il seguente processo cointegrato bivariato per le due variabili I(1)  $x_t$  e  $y_t$

$$\begin{aligned} y_t &= \beta_0 x_t + u_t \\ \Delta x_t &= \epsilon_t \\ u_t &\sim N(0, \sigma_u^2), \quad \epsilon_t \sim N(0, \sigma_\epsilon^2), \quad \mathbb{E}(u_t \epsilon_s) = \delta_{ts} \sigma_{u\epsilon} \end{aligned} \quad (2.4)$$

dove  $\delta_{ts}$  è il delta di Kronecker. Ciò a cui siamo interessati è la distribuzione limite dello stimatore dei minimi quadrati  $\hat{\beta}$  per il parametro  $\beta_0$ . Andando ad analizzare la seguente equazione

$$T(\hat{\beta} - \beta_0) = \left( T^{-1} \sum_{t=1}^T x_t u_t \right) \left( T^{-2} \sum_{t=1}^T x_t^2 \right)^{-1} \quad (2.5)$$

per il Teorema 1 sappiamo che

$$T^{-2} \sum_{t=1}^T x_t^2 \xrightarrow{W} \sigma_\epsilon^2 \int_0^1 B_s(r)^2 dr. \quad (2.6)$$

Al fine di determinare la distribuzione di  $T^{-1} \sum_{t=1}^T x_t u_t$  è conveniente riformulare

$u_t$  e  $\epsilon_t$  nel seguente modo:

$$\begin{aligned} u_t &= \gamma\epsilon_t + v_t \\ \gamma &= \frac{\sigma_{u\epsilon}}{\sigma_\epsilon^2} \\ \sigma_v^2 &= \sigma_u^2 - \frac{\sigma_{u\epsilon}^2}{\sigma_\epsilon^2} \end{aligned} \quad (2.7)$$

Per costruzione avremmo  $\mathbb{E}(\epsilon_t v_s) = 0 \quad \forall t \neq s$ . Definiti i moti browniani  $B_\epsilon(r)$  e  $B_v(r)$  ottenuti dalle rispettive successioni di  $\epsilon_t$  e  $v_t$ , utilizziamo le equazioni (2.4) e (2.7)

$$\begin{aligned} T^{-1} \sum_{t=1}^T x_t u_t &= T^{-1} \sum_{t=1}^T x_t (\gamma\epsilon_t + v_t) \\ &= \gamma \left( T^{-1} \sum_{t=1}^T x_{t-1} \epsilon_t \right) + \gamma \left( T^{-1} \sum_{t=1}^T \epsilon_t^2 \right) + \left( T^{-1} \sum_{t=1}^T x_{t-1} v_t \right) + \left( T^{-1} \sum_{t=1}^T v_t \epsilon_t \right) \end{aligned} \quad (2.8)$$

dove

$$T^{-1} \sum_{t=1}^T x_{t-1} \epsilon_t \xrightarrow{W} \frac{\sigma_\epsilon^2}{2} (B_\epsilon(1) - 1) \quad (2.9)$$

per le proprietà della serie  $\epsilon_t$

$$T^{-1} \sum_{t=1}^T \epsilon_t^2 \xrightarrow{p} \sigma_\epsilon^2 \quad (2.10)$$

e

$$T^{-1} \sum_{t=1}^T x_{t-1} v_t \xrightarrow{W} \sigma_\epsilon \sigma_v \int_0^1 B_\epsilon(r) dB_v(r), \quad (2.11)$$

$$T^{-1} \sum_{t=1}^T v_t \epsilon_t \xrightarrow{p} 0. \quad (2.12)$$

Quindi sotto l'ipotesi nulla risulta  $\beta = 0$

$$T\hat{\beta} \xrightarrow{W} \left[ \gamma \frac{\sigma_\epsilon^2}{2} (B_\epsilon(1) - 1) + \sigma_\epsilon \sigma_v \int_0^1 B_\epsilon(r) dB_v(r) \right] \left[ \sigma_\epsilon^2 \int_0^1 B_\epsilon(r)^2 dr \right]^{-1} \quad (2.13)$$

(Phillips e Hansen, 1988)

È importante notare che questo risultato contiene tutte le proprietà asintotiche della

regressione di cointegrazione. Lo stimatore  $\hat{\beta}$  è superconsistente perché il tasso di convergenza è  $T$  e, in secondo luogo, la distribuzione limite non è standard. Su quest'ultimo punto c'è però da notare che, quando  $\gamma = 0$  significa ipotizzare che c'è esogeneità per la stima di  $\beta_0$ , e sapendo inoltre che

$$\int_0^1 B_\epsilon(r) dB_v(r) \xrightarrow{W} N\left(0, \int_0^1 B_s(r)^2 dr\right) \quad (2.14)$$

allora  $\hat{\beta}$  avrà distribuzione normale standardizzata.

Col modello NEC non è possibile applicare in modo diretto questi risultati. Quindi è necessario ricercare tra i più importanti teoremi presenti in letteratura quello più adeguato per il NEC. In particolar modo lo scopo ultimo è quello di poter individuare la distribuzione limite dei parametri di questo modello.

Prima di passare ad alcuni tra i più importanti risultati presenti in letteratura, in merito ai modelli non lineari, è necessario fare una premessa riguardo al contesto empirico preso in esame. La teoria economica delle aspettative razionali sulla struttura a termine ci fornisce una descrizione ben precisa della relazione di lungo periodo, cioè che lo spread fra due tassi di interesse deve risultare stazionario. Ciò implica che i tassi cointegrino a coppie, quindi che la matrice di cointegrazione possa essere considerata nota come il suo rango. Per quanto riguarda il meccanismo di aggiustamento la questione è completamente opposta. La teoria non fa alcun riferimento esplicito riguardo alla forma della funzione di aggiustamento per cui è possibile fare solo alcune ipotesi sulla sua forma. Balke e Fomby (1992) utilizzano un modello a soglia e la ragione di questa scelta deriva dal fatto che gli aggiustamenti sono condotti in modo discontinuo dai *policy maker*. Il modello di Lucchetti e Palomba (2009) approssima questo comportamento con un polinomio di terzo grado e ciò permette anche un'analisi multivariata del meccanismo di aggiustamento più semplice da stimare. L'obiettivo è quello di verificare se una funzione di aggiustamento non lineare continua sia capace di rappresentare in modo appropriato gli interventi dei *policy maker*.

L'analisi sull'aggiustamento è incentrata su una funzione non lineare e sulla sua

rilevanza all'interno della struttura a termine dei tassi di interesse. Per cui è importante verificare il comportamento asintotico dei test e anche dei parametri stimati. Data la complessità del modello, un'importante questione su cui concentrarsi è la distribuzione limite dei parametri stimati e dei test LM, LR e W. Infatti è necessario verificare che il modello NEC abbia i requisiti necessari per poter essere collegato ad un qualche risultato asintotico e per poter poi derivare le diverse distribuzioni limite.

## 2.3 Proprietà dei modelli non lineari

### 2.3.1 Metodo minimi quadrati

Come già notato nel primo capitolo, nell'utilizzare un modello di cointegrazione non lineare, in cui il meccanismo di correzione dell'errore è approssimato da un polinomio razionale cubico (approssimante di Padé). Escribano (2004) impone alcune condizioni sia sulla stazionarietà delle serie che sulla funzione di aggiustamento. In particolare, queste condizioni fanno riferimento ai risultati ottenuti da Álvaro e Santiago (1996) riguardo alle proprietà asintotiche degli stimatori dei modelli non lineari attraverso il metodo dei minimi quadrati.

Questi autori partono con l'estendere il concetto di stazionarietà e di integrazione per poter dimostrare le proprietà asintotiche dei processi integrati. Tra le definizioni riguardo alla dipendenza seriale limitata delle serie storiche troviamo quella di *strong mixing* che afferma:

**Definizione 3** ( $\alpha$ -mixing). *Sia  $\{v_t\}$  una sequenza di variabili casuali e sia  $\mathcal{F}_s^t \equiv \sigma(v_s, \dots, v_t)$  una sigma algebra, si definiscono i coefficienti  $\alpha$ -mixing come*

$$\alpha_m \equiv \sup_t \sup_{A \in \mathcal{F}_{-\infty}^t, B \in \mathcal{F}_{t+m}^{+\infty}} |P(A \cap B) - P(A)P(B)| \quad (2.15)$$

Per  $m = 1, 2, \dots$  la sequenza  $\{\alpha_m\}$  rappresenta la misura del tasso al quale gli eventi passati diventano indipendenti con quelli futuri come la distanza tra gli stessi aumenta. Una sequenza è detta di ampiezza  $-a$  se  $\alpha_m = O(m^{-\lambda})$  per  $\lambda < -a$ .

Anche questa definizione è troppo stringente per cui viene combinata con un'altra chiamata *near epoch dependent* (NED).

**Definizione 4** (NED). *Sia  $\{w_t\}$  una sequenza di variabili casuali con  $E(w_t^2) < \infty$  per ogni  $t$ .  $\{w_t\}$  è near epoch dependent (NED) su una sequenza strong mixing  $\{v_t\}$  di ampiezza  $-a$  se  $\phi(n)$  è di ampiezza  $-a$ , dove  $\phi(n)$  è dato da*

$$\sup_t \|w_t - E_{t-n}^{t+n}(w_t)\|_2 \equiv \phi(n) \quad (2.16)$$

dove  $E_{t-n}^{t+n}(w_t) = E(w_t | v_{t-n}, \dots, v_{t+n})$  e  $\|\cdot\|_2$  è la norma  $L_2$  di una variabile casuale definita come  $E^{1/2}|\cdot|^2$ .

Quando  $\phi(n)$  tende a zero con un appropriato tasso, allora  $w_t$  dipende essenzialmente dai più recenti valori di  $v_t$ . Uno dei vantaggi nell'utilizzo di queste definizioni è quello di essere completamente non parametrico, quindi di non imporre alcuna restrizione sulla forma del processo. Inoltre un'altra importante caratteristica delle sequenze NED è che, sotto opportune condizioni, le funzioni di queste serie sono ancora considerate NED.

Arrivati a questo punto è possibile definire una versione del FCLT per le variabili NED, dimostrata da Davidson (1994), per cui

**Teorema 2** (FCLT per NED). *Considerate le seguenti assunzioni:*

1. *sia  $\{w_t\}$  una sequenza di variabili casuali a media zero, uniformemente limitata  $L_r$  e NED di ampiezza  $-1/2$  su un processo  $\alpha$ -mixing di grandezza pari a  $-r/(r-2)$  con  $r \in [0; 1]$*
2.  *$T^{-1}\mathbb{E}(\sum_{t=1}^T w_t)^2 \rightarrow \sigma^2$  dove  $0 < \sigma^2 < \infty$ .*

Allora  $W_T(r) = T^{-1/2} \sum_{i=1}^{\lfloor Tr \rfloor} w_t \xrightarrow{d} B(r)$ .

Da questo teorema derivano due importanti implicazioni:

**Definizione 5.** *Una sequenza  $\{w_t\}$  è  $I(0)$  se è NED su un sequenza  $\alpha$ -mixing sottointesa  $\{v_t\}$ , ma la sequenza  $\{x_t\}$  data da  $x_t = \sum_{s=1}^t w_s$  non è NED su  $\{v_t\}$ . In questo caso  $x_t$  sarà  $I(1)$ .*

Da questa definizione otteniamo che  $x_t$ , vista come una somma di serie NED, è I(1) mentre la sua differenza prima  $\Delta x_t$  è I(0).

Ritornando ai concetti di cointegrazione è possibile dunque affermare che

**Definizione 6.** *Due sequenze  $\{y_t\}$  e  $\{x_t\}$  sono cointegrate con vettore di cointegrazione normalizzato pari a  $\beta^* = [1, -\beta_{12}^*]'$  se  $y_t - \beta_{12}^* x_t$  è NED su una sequenza  $\alpha$ -mixing  $\{v_t\}$  ma non lo è per  $y_t - \delta_{12} x_t$  con  $\delta_{12} \neq \beta_{12}^*$*

Questa definizione può essere estesa anche alla situazione in cui si ha  $g(y_t - \beta_{12}^* x_t)$ .

Escribano e Mira (2002) forniscono anche un teorema di rappresentazione per i modelli di cointegrazione non lineari che rispettano determinate condizioni:

**Teorema 3** (Teorema di Rappresentazione). *Sia  $X_t = (y_t, x_t)'$  un vettore  $N \times 1$  dove  $y_t$  è uno scalare e  $x_t$  è di dimensione  $(N - 1) \times 1$  generati da*

$$\Delta y_t = B(L)\Delta x_{t-1} + A(L)\Delta y_{t-1} + g(z_{t-1}, \gamma) + \varepsilon_t \quad (2.17)$$

con  $\Delta x_t = \varepsilon_t$  e  $z_t = y_t - \beta' x_t$ .

Siano rispettate le seguenti condizioni

1.  $\varepsilon_t$  è una differenza di martingala con valore atteso zero e varianza costante pari a  $\sigma_\varepsilon$  e  $\varepsilon_t$  è  $\alpha$ -mixing con matrice delle varianze e covarianze pari a  $\Sigma$
2. il polinomio  $B(L)$  ha tutte le radici al di fuori del cerchio unitario
3. la funzione  $g(z_{t-1}, \gamma)$  è continua e differenziabile su  $z$
4. condizione di stabilità:  $-2 < dg(z_{t-1}, \gamma)/dz_{t-1} < 0$ , dove  $d(\cdot)$  rappresenta la derivata prima
5.  $\Delta x_t$ ,  $\varepsilon$  e i prodotti incrociati hanno momenti secondi finiti

Allora

- $X_t$  è un vettore di componenti I(1) e
- $z_t$  e  $\Delta X_t$  sono NED e quindi  $y_t$  e  $x_t$  cointegrano e il loro vettore di cointegrazione è  $(1, \beta')$ .

Come si può notare, il teorema considera la sola ipotesi che la relazione di lungo periodo sia lineare. Inoltre le condizioni che impongono sulla funzione di aggiustamento, differenziabilità e stabilità, restringe considerevolmente l'ambito delle forme funzionali a cui si può applicare.

Sulla base dei risultati precedenti Álvaro e Santiago (1996) trovano la distribuzione asintotica dei coefficienti ottenuti con il metodo dei minimi quadrati di un modello di cointegrazione non lineare.

**Teorema 4** (Stimatore a due passi del NEC). *Considerati il Teorema 2 e le Condizioni 1 - 5 del Teorema 3 per l'equazione (2.17) che ha come relazione di cointegrazione*

$$y_t = \beta' x_t + z_t \quad (2.18)$$

Le Equazioni (2.18) e (2.17) possono essere stimate attraverso il metodo OLS e NLS con le seguenti proprietà:

- *Lo stimatore OLS del vettore di cointegrazione  $\beta$ , nell'Eq.(2.18), è superconsistente e la sua distribuzione asintotica è non standard.*
- *Gli stimatori ottenuti con l'OLS e il NLS dei parametri  $A(1)$ ,  $B(1)$  e  $\gamma$  dell'equazione (2.17) hanno la stessa distribuzione limite, non importa quale valore di  $z_{t-1}$  si usi.*
- *Gli stimatori OLS o NLS dei parametri  $A(1)$ ,  $B(1)$  e  $\gamma$  dell'equazione (2.17), dove  $z_{t-1}$  era stata precedentemente stimata attraverso il metodo OLS nell'equazione (2.18), sono  $\sqrt{T}$ -consistenti e asintoticamente normali.*

Con questi risultati a disposizione è possibile definire alcune proprietà asintotiche dei modelli. Infatti, dopo aver ottenuto le distribuzioni asintotiche dei parametri, è possibile costruire le opportune statistiche test. In particolare nel nostro caso avremmo che i test classici LM, W e LR seguono le usuali proprietà, quindi è possibile verificare le ipotesi usando le distribuzioni standard.

Inoltre questi risultati sono validi nel caso in cui la relazione di lungo periodo è lineare e non vi è distinzione se la matrice di cointegrazione sia nota oppure incognita.

Se si considera il modello di Lucchetti e Palomba (2009) la condizione di stabilità per la funzione di aggiustamento non è verificata. Infatti un polinomio di terzo grado è una funzione illimitata per valori di  $z_t \rightarrow \infty$ .

### 2.3.2 Metodo massima verosimiglianza

Nell'analisi delle proprietà asintotiche degli stimatori dei modelli di cointegrazione non lineari ottenuti con il metodo della massima verosimiglianza, Kristensen e Rahbek (2013) mostrano che le distribuzioni limite di questi non sono standard. Il motivo è dovuto dal fatto che le distribuzioni asintotiche dei parametri dei coefficienti di breve e di lungo periodo non sono asintoticamente indipendenti. Ciò compromette l'inferenza soprattutto per quanto riguarda i test classici come LM, LR e W. Nel loro lavoro, Kristensen e Rahbek (2013) verificano un'ampia classe d'ipotesi, in particolare quella della presenza della non linearità, definendo anche le distribuzioni asintotiche sia degli stimatori che dei test. Nei casi in cui l'inferenza non è standard, gli autori propongono l'utilizzo della procedura bootstrap per effettuare i test.

In un contesto generale i modelli ECM, in cui le  $X_t \in \mathbb{R}^p$  (con  $t = 1, \dots, T$ ) sono le osservazioni, possono essere rappresentati dalla seguente formulazione

$$\Delta X_t = g(\beta' X_{t-1}) + \Phi_1 \Delta' X_{t-1} + \dots + \Phi_k \Delta' X_{t-k} + \varepsilon_t \quad (2.19)$$

dove il termine di errore  $\varepsilon_t$  è una sequenza di differenze di martingala (MDS) e la funzione  $g(\cdot)$  descrive il meccanismo di aggiustamento (lineare o no) attorno all'equilibrio di lungo periodo  $\beta' X_{t-1}$ , con  $\beta \in \mathbb{R}^{p \times r}$  dove  $p$  è il numero delle serie prese in esame e  $r$  è il rango di cointegrazione.

Una rappresentazione della funzione di aggiustamento  $g(\cdot)$  che permette di raccogliere i più noti modelli non lineari di correzione dell'errore presenti in letteratura può essere descritta nella seguente forma:

$$g(\beta' X_{t-1}) = \alpha \beta' X_{t-1} + \delta f(\beta' X_{t-1}; \xi) \quad (2.20)$$

in cui il parametro  $\delta$  ha dimensione pari a  $p \times r_\xi$ , mentre il parametro  $\xi$  ha dimen-

sione  $d_\xi$ .

Spesso nell'utilizzo dei modelli ECM vengono fatte ipotesi sulla forma del vettore di cointegrazione o perché imposte *a priori* dalla teoria o per facilitare i calcoli. Una rappresentazione normalizzata di  $\beta$  in funzione del parametro reale  $\beta_0$  può essere espressa come

$$\beta = \beta_0 + \kappa_0 b \quad (2.21)$$

dove la matrice  $\kappa_0$  di dimensione  $p \times (p-r)$  è nota, mentre  $b$  di dimensione  $(p-r) \times r$  deve essere stimata. Nel caso dello spread tra due variabili avremmo  $\beta = (1, \beta_2)$  quindi, dovremmo porre  $\kappa = (0, 1)$  in modo tale che il primo parametro di  $\beta$  viene fissato, mentre  $\beta_2$  deve essere stimato.

Riprendendo il modello generale (Eq.(2.20)) e tenendo in considerazione la formulazione nell'equazione (2.21), il modello di regressione sarà

$$\Delta X_t = g(Z_{0,t-1} + b' Z_{1,t-1}) + \Phi Z_{2,t-1} + \varepsilon_t \quad (2.22)$$

dove

$$\begin{aligned} Z_{0,t} &= \beta_0' X_t \in \mathbb{R}^r, \\ Z_{1,t} &= \kappa_0' X_t \in \mathbb{R}^{p-r}, \\ Z_{2,t} &= (\Delta X_t', \dots, \Delta X_{t-k+1}') \in \mathbb{R}^{pk}. \end{aligned} \quad (2.23)$$

Per ottenere la distribuzione asintotica degli stimatori e dei test statistici, è necessario prima stabilire una versione più allargata del teorema del limite centrale. Per far ciò Kristensen e Rahbek (2013) estendono i risultati di Caner e Hansen (2001).

Definiti i seguenti due processi

$$\begin{aligned} \phi_T(s, \xi) &= \frac{1}{\sqrt{T}} \sum_{t=1}^{[Ts]} f(Z_{0,t}, \xi) \varepsilon_t \\ x_T(s) &= K_T^{-1} Z_{1, Ts} \end{aligned} \quad (2.24)$$

dove  $s \in [0, 1]$  e  $\xi \in \Xi$  per qualche insieme compatto  $\xi \subseteq K_T^{-1}$  è una matrice dei

pesi tale per cui una sua sequenza  $K_T^{-1}Z_{1,Ts}$  converge debolmente (Condizione 7). Per poter stabilire la convergenza debole per questi processi è necessario in prima istanza stabilire delle condizioni di regolarità:

**Condizione 1.** *La sequenza  $(\varepsilon_t, Z_{0,t})$ :*

1. *è stazionaria ed ergodica*
2.  *$\varepsilon_t$  è una differenza di martingala rispetto a  $\mathcal{F}_{T,t}$  tale per cui  $\mathbb{E}[\varepsilon_t|\mathcal{F}_{T,t-1}] = 0$  e  $\mathbb{E}[\varepsilon_t\varepsilon_t'|\mathcal{F}_{T,t-1}] = \Omega_\varepsilon$  per qualche matrice costante e finita  $\Omega_\varepsilon \in \mathbb{R}^{d_\varepsilon \times d_\varepsilon}$*

**Condizione 2.** *Le sequenze  $f(Z_{0,t}, \xi)$  e  $\varepsilon_t$  per qualche costante  $m, n > 0$  soddisfano:*

1.  $\mathbb{E}[\sup_{\xi \in \Xi} \|f(Z_{0,t}, \xi)\|^m] < \infty$  e  $\mathbb{E}[\|\varepsilon_t\|^m] < \infty$
2.  $\|f(Z_{0,t}, \xi) - f(Z_{0,t}, \xi')\| < B(Z_{0,t}) \|\xi - \xi'\|$ , per tutti gli  $\xi, \xi' \in \Xi$  e  $\mathbb{E}[B(Z_{0,t})^n] < \infty$ .

Secondo Kristensen e Rahbek (2013) la Condizione 1.2 può essere estesa anche ai processi con eteroschedasticità condizionale, sebbene abbiano tralasciato la dimostrazione, se si assume che  $\mathbb{E}[\varepsilon_t\varepsilon_t'|\mathcal{F}_{T,t-1}] = \Omega_{\varepsilon,t}$  e tale per cui  $\Omega_{\varepsilon,t}$  sia stazionaria e i suoi momenti primi esistano.

Inoltre la Condizione 2.2 può essere facilmente generalizzabile imponendo che la funzione  $f(\cdot, \cdot)$  sia continuamente differenziabile in  $\Xi$ .

**Teorema 5 (FCLT).** *Siano soddisfatte le Condizioni 1 e 2 con  $m, n \geq 2$ , il processo della somma parziale definito dall'equazione (2.24) soddisfa*

$$\phi_T(s, \xi) \xrightarrow{W} \phi(s, \xi) \tag{2.25}$$

dove  $\phi_T(s, \xi)$  è un processo Gaussiano.

La “ $\xrightarrow{W}$ ” sta ad indicare la convergenza debole nello spazio di funzioni. Questo teorema è ottenuto dall'estensione degli argomenti di Escanciano (2004).

La definizione degli stimatori e dei test statistici richiede di individuare la funzione score e la matrice Hessiana. Dall'equazione (2.22) e usando la definizione

di Magnus e Neudecker (1988), definiamo il differenziale di primo ordine della log-verosimiglianza rispetto ai parametri  $L_T(\theta, b, \xi)$  come

$$dL_T(\theta, b, \xi; d\theta, db, d\xi) = \mathbb{S}_T(\theta, b, \xi)' \text{vec}(d\theta, db, d\xi) \quad (2.26)$$

dove  $\theta = (\alpha, \delta, \Phi)$  e  $\mathbb{S}_T(\cdot)$  rappresenta il vettore score. Il procedimento è identico per l'Hessiana  $\mathbb{H}_T(\theta, b, \xi)$  tranne per il fatto che si utilizza il differenziale di secondo ordine.

Riscrivendo la log-verosimiglianza in modo da raggruppare le derivate parziali dei singoli parametri e esprimendo tutto in funzione di  $\xi$ , otteniamo:

$$dL_T(\theta, b, \xi; d\theta, db, d\xi) = (\text{vec}(d\theta))' \sum_{t=1}^T u_t(\xi) + (\text{vec}(d\xi))' \sum_{t=1}^T v_t(\xi) + \sum_{t=1}^T Z'_{1,t-1}(db) w_t(\xi) \quad (2.27)$$

dove  $u_t(\xi), v_t(\xi), w_t(\xi)$  raccolgono le derivate parziali dei rispettivi parametri.

Nel lavoro di Kristensen e Rahbek (2013) vengono riportate tutte le tipologie di test (risposta simmetrica, numero dei ritardi, ecc.). Nel proseguo riporterò solo i casi di interesse cioè i test riferiti ai parametri  $\delta$  e  $\xi$ . Si noti che tutti i risultati sono stati ottenuti dagli autori nella condizione che  $\beta$  sia identificato sotto l'ipotesi nulla.

Al fine di poter verificare se la componente non lineare è rilevante nel rappresentare il meccanismo di aggiustamento dell'errore, è necessario costruire un opportuno test. Riprendendo il modello dell'equazione (2.22) l'ipotesi che andremo a verificare sarà

$$H_0 : R'_\delta \text{vec}(\theta, b, \xi) = \text{vec}(\delta) = 0. \quad (2.28)$$

Ora nel caso in cui il parametro  $\xi$  non sia identificato sotto l'ipotesi nulla

dovremmo avvalerci dei sup-test. Quindi i tre test classici che otterremo saranno:

$$\begin{aligned}
\sup_{\xi \in \Xi} LR_T(\xi) &= \sup_{\xi \in \Xi} 2[L_T(\hat{\theta}(\xi), b, \xi) - L_T(\tilde{\theta}, b, \xi)], \\
\sup_{\xi \in \Xi} LM_T(\xi) &= \sup_{\xi \in \Xi} \mathbb{S}_T(\tilde{\theta}(\xi), b, \xi)' \mathbb{H}_T^{-1}(\tilde{\theta}(\xi), b, \xi) \mathbb{S}_T(\tilde{\theta}(\xi), b, \xi), \\
\sup_{\xi \in \Xi} W_T(\xi) &= \sup_{\xi \in \Xi} \text{vec}(\hat{\delta}(\xi))' R_\delta \left[ (0, R'_\delta) \mathbb{H}_T^{-1}(\hat{\theta}(\xi), b, \xi) (0, R'_\delta)' \right]^{-1} R'_\delta \text{vec}(\hat{\delta}(\xi)),
\end{aligned} \tag{2.29}$$

dove  $\hat{\theta}$  sono i parametri ottenuti con la massimizzazione del modello libero e  $\tilde{\theta}$  sono i parametri ottenuti col modello vincolato.

Kristensen e Rahbek (2013) impongono alcune condizioni di regolarità sui processi stocastici in modo da definire la distribuzione asintotica dello score e dell'Hessiana.

**Condizione 3.** *Lo spazio dei parametri  $\Xi$  per  $\xi$  è compatto e  $\vartheta_0 = (\theta_0, \beta_0, \xi_0)$  giace all'interno di  $\Theta \times \Xi$*

**Condizione 4.** *La funzione  $f(z, \xi)$  è tre volte differenziabile in  $z$  e  $\xi$ .*

*La funzione e le sue derivate sono polinomialmente limitate di ordine  $\rho \geq 1$  in  $z$  uniformemente su  $\xi$ ,  $\|f(z, \xi)\| \leq C(1 + |z|^\rho)$  per  $C > 0$ .*

**Condizione 5.** *Il termine di errore  $\varepsilon_t$  è una differenza di martingala rispetto a  $\mathcal{F}_{t-1} = \mathcal{F}(\varepsilon_s, X_s : s \leq t-1)$ . Inoltre,  $\Omega_0 \equiv \mathbb{E}[\varepsilon_t \varepsilon_t' | \mathcal{F}_{t-1}] > 0$  e  $\mathbb{E}[\|\varepsilon_t\|^{q_\varepsilon}] < \infty$  per qualche  $q_\varepsilon \geq 2$ .*

**Condizione 6.** *Il processo  $(Z'_{0,t}, Z'_{2,t})'$  è stazionario ed ergodico con  $\mathbb{E}[\|Z_{0,t}\|^{q_0}] < \infty$  e  $\mathbb{E}[\|Z_{2,t}\|^{q_2}] < \infty$  per qualche  $q_0, q_2 \geq 1$ .*

**Condizione 7.** *Con  $\kappa_0$  definito nell'Eq.(2.21) e per qualche sequenza della matrice diagonale  $K_T \in \mathbb{R}^{(p-r) \times (p-r)}$  che soddisfa  $K_T^{-1} \rightarrow 0$  per  $T \rightarrow \infty$ , il processo non stazionario  $Z_{1,t} = \kappa'_0 X_t$ :*

1.  $K_T^{-1} Z_{1,[T]} \xrightarrow{W} F(\cdot)$  su  $D([0, 1])$ , per un processo stocastico continuo  $F(s)$  che soddisfa  $\int_1^0 F(s) F(s)' ds > 0$
2.  $\sup_{T \geq 1} \sup_{t \leq T} \mathbb{E} \|K_T^{-1} Z_{1,t}\|^{q_1} < \infty$  per  $q_1 > 0$

dove  $D([0, 1])$  indica lo spazio delle funzioni càdlàg.

Da notare che la Condizione 4 ci assicura che le prime tre derivate della funzione di verosimiglianza rispetto ai parametri siano definite inoltre, la Condizione 6 con il limite polinomiale, ci permette di verificare l'esistenza dei loro momenti. La restrizione  $\int_1^0 F(s)F(s)'ds > 0$  è usata per assicurare che la matrice di informazione associata ad  $F$  non sia singolare.

Al fine di poter stabilire la distribuzione asintotica per i test è necessario prima procedere all'analisi del comportamento delle sequenze che costituiscono lo score e l'Hessiana della log-verosimiglianza. Applicando il Teorema 5 otteniamo

**Teorema 6.** *Supposto che le Condizioni 3 - 6 e la 7.1 siano soddisfatte con  $q_2, q_\varepsilon = 2$  e  $q_0 = 2\rho$ . Allora  $u_t(\xi), v_t(\xi), w_t(\xi)$  definite dall'equazione (2.27) e  $Z_{1,t}$*

$$\left( \frac{1}{\sqrt{T}} \sum_{t=1}^{[Ts]} u_t(\xi)', \frac{1}{\sqrt{T}} \sum_{t=1}^{[Ts]} v_t(\xi)', \frac{1}{\sqrt{T}} \sum_{t=1}^{[Ts]} w_t(\xi)', (K_T^{-1} Z_{1,[Ts]})' \right) \quad (2.30)$$

$$\xrightarrow{W} (B_u(s, \xi)', B_v(s, \xi)', B_w(s, \xi)', F(s))$$

dove  $B_u, B_v, B_w$  sono processi Gaussiani.

Siano  $V_{\theta, \beta, T}$  e  $V_{\xi, T}$  i vettori contenenti i tassi di convergenza dei parametri della QMLE.

**Teorema 7.** *Per  $\delta_0 = 0$ , supposto che le Condizioni 3 - 7 siano soddisfatte con  $q_1, q_\varepsilon = \max(4, 3d_\varepsilon)$ ,  $q_0 = \rho \max(4, 3d_\varepsilon)$  e  $q_2 = 2$ . Assumendo inoltre che la matrice delle covarianze  $\Sigma(\xi_1, \xi_2) > 0$  per tutti gli  $\xi_1, \xi_2 \in \Xi$ . Allora i seguenti risultati sono dimostrati uniformemente su  $\xi$ . Con probabilità tendente a uno, esiste un solo punto di minimo  $(\hat{b}(\xi)', \hat{\theta}(\xi))$  di  $L_T(b, \theta, \xi)$  nell'intorno  $\{\|\theta - \theta_0\| < \varepsilon, \|K_T b\| < \varepsilon\}$  per  $\varepsilon > 0$ . Allora*

$$T^{1/2} V_{b, \theta, T}^{-1/2} \text{vec}(\hat{\theta}(\xi) - \theta_0) \xrightarrow{W} \mathbb{H}_{\theta, b, \theta, b}^{-1}(\xi) \mathbb{S}_{\theta, b}(\xi) \quad (2.31)$$

dove

$$\mathbb{H}_{\theta, b, \theta, b} = \begin{pmatrix} \int_0^1 F(s)F(s)'ds \otimes \Sigma_{w, w} & \int_0^1 F(s)ds \otimes \Sigma_{w, u}(\xi, \xi) \\ \int_0^1 F(s)'ds \otimes \Sigma_{u, w}(\xi, \xi) & \Sigma_{u, u}(\xi, \xi) \end{pmatrix} \quad (2.32)$$

$$\mathbb{S}_{\theta,b} = \left( \text{vec} \left( \int_0^1 F(s) dB'_w(s) \right)', B_u(1, \xi)' \right)' \quad (2.33)$$

Possiamo notare che il DGP sotto l'ipotesi nulla è un modello di cointegrazione lineare quindi, sotto le condizioni di Johansen (1995), si dimostra che  $F(s)$  è un moto browniano. Sebbene la componente non lineare scompaia sotto l'ipotesi nulla possiamo subito notare che la distribuzione asintotica dipende ancora dal parametro  $\xi$ .

**Teorema 8.** *Si supponga che le Condizioni 3 - 7 siano rispettate e si assuma che  $H_0 = R' \text{vec}(\theta, \xi) = 0$  con  $R$  ha rango pieno e  $\beta$  sia identificato.*

*Se  $\xi$  non è identificato sotto l'ipotesi nulla, allora con  $q_0 = \rho \max(4, 3d_\varepsilon)$ ,  $q_1, q_\varepsilon = \max(4, 3d_\varepsilon)$ , e  $q_2 = 2$ , i test*

$$\begin{aligned} \sup LM_T &\xrightarrow{d} \sup_{\xi \in \Xi} \mathbb{V}_\theta(\xi) \mathbb{V}_\theta(\xi)' \\ \sup LR_T &\xrightarrow{d} \sup_{\xi \in \Xi} \mathbb{V}_\theta(\xi) \mathbb{V}_\theta(\xi)' \\ \sup W_T &\xrightarrow{d} \sup_{\xi \in \Xi} \mathbb{V}_\theta(\xi) \mathbb{V}_\theta(\xi)' \end{aligned} \quad (2.34)$$

dove, con  $\mathbb{S}_{\theta,b}(\xi)$  e  $\mathbb{H}_{\theta,b\theta,b}(\xi)$  dati dal Teorema 3,

$$\mathbb{V}_\theta(\xi) = \left[ (M_\theta)'_{\perp} \mathbb{H}_{\theta,b\theta,b}^{-1}(\xi) (M_\theta)_{\perp} \right]^{-1} (M_\theta)'_{\perp} \mathbb{H}_{\theta,b\theta,b}^{-1}(\xi) \mathbb{S}_{\theta,b}(\xi) \quad (2.35)$$

con  $M_\theta = \text{diag} \left( I_{(p-r)r}, (R_\theta)_{\perp} \right)$ .

Nel caso in cui  $\mathbb{E}[f(Z_{0,t-1}, \xi)] = 0$ , come ad esempio  $f(Z_{0,t-1}, \xi)$  sia simmetrica rispetto a zero, otteniamo lo stesso risultato di Hansen e Seo (2002). In questo frangente viene evitato il contributo della componente non lineare quindi, la distribuzione di  $\mathbb{V}_\theta(\xi)$  è un processo Gaussiano in quanto  $\Sigma_{w,u} = 0$  e otteniamo

$$\mathbb{H}_{\theta,b\theta,b} = \begin{pmatrix} \int_0^1 F(s) F(s)' ds \otimes \Sigma_{w,w} & 0 \\ 0 & \Sigma_{u,u}(\xi, \xi) \end{pmatrix} \quad (2.36)$$

Nel caso in cui  $\mathbb{E}[f(Z_{0,t-1}, \xi)] \neq 0$  il risultato che otteniamo diverge da quello di Hansen e Seo (2002). La distribuzione limite ha due componenti: la prima è composta dai processi stazionari  $(Z_{0,t-1}, Z_{2,t-1}, f(Z_{0,t-1}, \xi))$  con i rispettivi score  $(\mathbb{S}_\alpha(\xi), \mathbb{S}_\Phi(\xi), \mathbb{S}_\delta(\xi))$ , mentre la seconda, non stazionaria è costituita dal processo  $(Z_{1,t-1})$  con il vettore dello score  $\mathbb{S}_b(\xi)$ .

In conclusione Kristensen e Rahbek (2013) mettono in evidenza due risultati.

- Il primo afferma che, se il vettore di cointegrazione  $\beta$  è noto, i test che escludono la componente non lineare sono asintoticamente distribuiti in modo standard.
- Il secondo afferma che, se il vettore di cointegrazione è incognito, non può essere considerato come se fosse noto poiché le distribuzioni limite dei test sono influenzate dalla componente non stazionaria. L'unica eccezione si presenta quando  $\mathbb{E}[f(Z_{0,t-1}, \xi)] = 0$  in cui si evita il contributo della componente non stazionaria e la distribuzione limite è quella riportata da Hansen e Seo (2002).

Per ovviare a questi problemi e per valutare le distribuzioni limite dei vari test Kristensen e Rahbek (2013) propongono l'utilizzo del bootstrap. Tra le diverse procedure bootstrap gli autori considerano quella wild perché robusto all'eteroschedasticità condizionale. Attraverso le simulazioni dimostrano che questa classe di test fornisce ottimi risultati sia in termini di ampiezza che di potenza del test.

## 2.4 Procedura bootstrap

Alcuni tra i modelli non lineari non godono delle usuali proprietà asintotiche come quelli lineari e spesso c'è una naturale incertezza riguardo alle loro distribuzioni limite. Di fronte a queste problematiche la metodologia bootstrap è un ottimo strumento sia in termini di accuratezza che di potenza nella verifica dei test.

La maggior parte dei test statistici utilizzati in econometria sono asintoticamente pivotali, cioè hanno una distribuzione asintotica nota che non dipende dai parametri della popolazione incognita. I test più noti sono le statistiche  $t$  e  $F$ , il test di Dickey-Fuller e quello di Durbin-Watson. Un'approssimazione dei  $p$ -value delle sta-

tistiche test che sono asintoticamente pivotali o meno può essere ottenuta attraverso il bootstrap.

Il bootstrap, il cui nome deriva da Efron (1979), è una tecnica di ricampionamento che non richiede assunzioni particolari sulla forma del DGP. L'idea di base di questo strumento è che, anziché utilizzare una qualche forma funzionale della DGP incognita, viene presa al suo posto una sua stima. Questa procedura è molto usata per fare inferenza nell'analisi dei parametri incogniti e ciò comprende la loro distorsione, la stima della varianza, la verifica delle ipotesi e la costruzione degli intervalli di confidenza. La metodologia bootstrap è molto utile quando l'inferenza è basata su procedure complesse per le quali i risultati asintotici non sono disponibili o non sono utili a causa della ridotta dimensione del campione.

Un altro motivo perché è meglio utilizzare il bootstrap anziché l'inferenza asintotica è l'*asymptotic refinement* riferito al tasso di convergenza alla distribuzione vera. Sebbene una statistica sia asintoticamente pivotale, cioè la sua distribuzione limite è indipendente dai valori campionari, la differenza tra la vera distribuzione della statistica e la distribuzione del bootstrap diminuisce, al crescere della numerosità campionaria, più velocemente rispetto alla differenza tra la vera distribuzione e quella asintotica (Beran, 1988).

Supponiamo di avere un campione i.i.d.  $X_n = \{x_1, \dots, x_n\}$  estratto da una distribuzione  $F_0 = F(x, \theta_0)$ : la nostra statistica di interesse sarà  $\hat{\theta} = T(X_n)$ . Dovendo fare inferenza è necessario fornire un'approssimazione di  $G_n(x, F_0) = P(\hat{\theta} \leq x)$  che è una funzione di ripartizione di  $T(X_n)$ . Un metodo per stimare  $G_n(x, F_0)$  è la teoria asintotica. Infatti, sotto stringenti ipotesi, è possibile derivare la distribuzione limite  $G(x, F_0)$  e, centralizzando e normalizzando, è possibile ottenere una statistica asintoticamente pivotale che non dipende da  $F_0$ .

Un secondo metodo è il bootstrap. Questo fornisce un'approssimazione alternativa delle distribuzioni dei campioni finiti delle statistiche di interesse  $T(X_n)$ . In base alle condizioni di partenza è possibile scegliere se applicare il bootstrap non parametrico o parametrico. Nel caso in cui la funzione di distribuzione  $F_0$  sia incognita, il bootstrap non parametrico la rimpiazza con un suo stimatore, la funzione

di distribuzione empirica cumulata (EDF),

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(x_i \leq x) \quad (2.37)$$

dove  $\mathbb{1}(\cdot)$  rappresenta la funzione indicatrice.

Una piccola annotazione sulla simbologia: d'ora in poi le osservazioni, le statistiche o i parametri contrassegnati dall'asterisco (come ad esempio  $x^*$ ,  $\theta^*$ ) stanno ad indicare i valori simulati ottenuti con i metodi bootstrap.

Per ricavare la stima della distribuzione della statistica d'interesse  $G(x, F_0)$  si utilizza la seguente procedura:

1. si estraggono casualmente con ripetizione  $n$  campioni  $X_n^* = \{x_1^*, \dots, x_n^*\}$  da  $X_n = \{x_1, \dots, x_n\}$  cioè  $F_n(x)$ ,
2. si calcola la statistica bootstrap  $\hat{\theta}^* = T^*(X_n^*)$  dal campione simulato,
3. si ripetono i punti 1 e 2  $B$  volte in modo da ottenere le statistiche bootstrap  $\hat{\theta}_1^*, \hat{\theta}_2^*, \dots, \hat{\theta}_B^*$  e si calcola poi la EDF della statistica come

$$G(x, F_n) = \frac{1}{B} \sum_{b=1}^B \mathbb{1}(\hat{\theta}_b^* \leq x).$$

Da questo algoritmo è possibile poi proseguire al fine di ottenere un intervallo di confidenza o condurre un test d'ipotesi. Se ad esempio si vuole condurre un test ad una coda con livello di significatività pari ad  $\alpha$ , in cui la regione di rifiuto viene posta a sinistra, il  $p$ -value bootstrap sarà calcolato dalla EDF come

$$p(\hat{\theta}) = \frac{1}{B} \sum_{b=1}^B \mathbb{1}(\hat{\theta}_b^* \leq \hat{\theta}).$$

Quindi  $p(\hat{\theta})$  sarà la frazione delle statistiche bootstrap che non superano la statistica ottenuta dal campione osservato  $\hat{\theta}$ . Naturalmente l'ipotesi nulla viene rifiutata quando il valore di  $p(\hat{\theta})$  è inferiore ad  $\alpha$ .

Davidson (2009) elenca due “golden rule” da rispettare quando si usa il bootstrap. La prima afferma che le replicazioni bootstrap devono essere fatte con il

modello dell'ipotesi nulla, mentre l'altra indica che il bootstrap deve essere basato sulle stime efficienti. La prima regola deve essere soddisfatta al fine di ottenere un test asintoticamente valido, mentre la seconda consente al bootstrap di ottenere la probabilità di rifiutare l'ipotesi nulla vera più vicina al livello di significatività.

Nel caso in cui la funzione  $F_0$  appartenga ad una famiglia di distribuzioni note  $F(\cdot, \theta)$ , allora è possibile applicare il bootstrap parametrico. La procedura per calcolare rimane pressoché invariata tranne che per il punto 1. Infatti il campione bootstrap viene generato estraendo con ripetizione dalla distribuzione  $F(\cdot, T(X_n))$  (Horowitz, 2003).

Per quanto concerne la scelta del numero di ripetizioni da effettuare col bootstrap Davidson (2009) consiglia di utilizzare  $\alpha(B + 1)$  in modo tale che sia un intero.

Il metodo bootstrap fornisce una stima consistente della distribuzione di una statistica sotto determinate condizioni, come dimostrato da Beran e Ducharme (1991).

Uno dei maggiori limiti di questa versione generale del bootstrap è che non può essere applicato ai dati dipendenti. Infatti nel momento in cui viene meno la condizione di indipendenza, il bootstrap non fornisce più una stima consistente, come mostrato da Singh (1981). Infatti i dati ricampionati i.i.d. del bootstrap non replicano la struttura di dipendenza del processo generatore dei dati.

Anche nell'ambito delle serie storiche il bootstrap, così come rappresentato in precedenza, non consente una replicazione ottimale dei dati come discusso da Li e Maddala (1996). Per ottenerla è necessaria una metodologia alternativa che permetta di catturare la struttura di dipendenza temporale presente nei dati. Questo argomento viene affrontato da Hall (1992), Efron e Tibshirani (1993), Davison e Hinkley (1997), Härdle *et al.* (2003) e Lahiri (2003) che descrivono diversi aspetti della metodologia bootstrap in ambito delle serie storiche.

I modelli più noti nella letteratura delle serie storiche sono il *sieve bootstrap*, il *block bootstrap* e il *wild bootstrap*. Il primo modella la dipendenza temporale usando un modello parametrico, mentre il secondo cerca di preservarla all'interno dei blocchi delle osservazioni (MacKinnon, 2006). Il terzo infine è una procedura utilizzata nel caso in cui vi sia la presenza di eteroschedasticità.

## 2.4.1 Metodologie bootstrap per le serie storiche

### Sieve bootstrap

Il *sieve bootstrap* ideato da Swanepoel e van Wyk (1986) è basato sulla proprietà di invertibilità dei processi stazionari. Prendendo in considerazione un processo univariato lineare omoschedastico

$$y_t = \beta' X_t + \varepsilon_t \quad (2.38)$$

dove  $\beta$  e  $X_t$  sono vettori di dimensione  $k \times 1$ ,  $\varepsilon_t$  è il termine di errore la cui media condizionata è  $E(u_t|X_t) = 0$ . Il *sieve bootstrap* cerca di approssimare il processo esplicitato rispetto a  $\varepsilon_t$  con un AR(p).

La procedura parte dalla stima dell'equazione (2.38) da cui si ricavano i residui  $\hat{\varepsilon}_t$ . Successivamente viene stimato

$$\hat{\varepsilon}_t = \sum_{i=1}^p \rho_i \hat{\varepsilon}_{t-i} + u_t$$

dove  $p$  è scelto utilizzando i criteri informativi, mentre  $u_t$  rappresenta il termine di errore. Dopo aver ricampionato con reinserimento i residui  $u_t$ , viene generato il campione bootstrap in modo ricorsivo dalla seguente equazione

$$\varepsilon_t^* = \sum_{i=1}^p \hat{\rho}_i \varepsilon_{t-i}^* + u_t^*$$

dove  $\hat{\rho}_i$  sono i parametri stimati mentre  $t = -k, \dots, 0, 1, \dots, n$ . Il valore  $k$  è scelto in modo arbitrario, mentre i parametri dei valori iniziali che vanno da  $-k$  a  $-k + p - 1$  sono posti uguali alla media campionaria della serie analizzata, in questo caso pari a zero. I dati bootstrap ottenuti precedentemente vengono poi utilizzati nell'equazione

$$y_t^* = X_t \hat{\beta} + \varepsilon_t^*.$$

Questa procedura viene ripetuta  $B$  volte al fine di ottenere la distribuzione o la

statistica di interesse su  $\hat{\beta}$ .

L'applicazione del *sieve bootstrap* risulta piuttosto limitata ad alcuni determinati processi. Infatti una delle supposizioni su cui si basa questo metodo è il fatto che gli errori  $u_t$  siano indipendenti e identicamente distribuiti. Inoltre la dinamica di un AR(p) non fornisce una buona approssimazione dei processi di serie storiche che si presentano nella pratica (Davidson, 2009). Sebbene sia possibile utilizzarlo in alcuni processi non stazionari (Park, 2003), il *sieve bootstrap* non può essere applicato ai processi non lineari e questo esclude anche quelli eteroschedastici (Bühlmann, 1997).

### **Block bootstrap**

Il block bootstrap è un metodo non parametrico originariamente proposto da Künsch (1989). Lo scopo è di dividere le osservazioni in blocchi al fine di replicare la struttura di dipendenza delle serie temporali. Questa metodologia consiste praticamente nel costruire blocchi con osservazioni contigue del campione osservato e poi ricampionare questi blocchi. In base a come si vogliono suddividere i dati si può scegliere fra due diverse metodologie di block bootstrap: a blocchi sovrapposti (*moving block bootstrap*) o non sovrapposti (*non-overlapping block bootstrap*).

Prendendo un campione costituito da  $n$  osservazioni di  $y_t$ , nel caso dei blocchi sovrapposti di lunghezza  $l$ , il primo blocco è composto dalle osservazioni  $\{y_1, \dots, y_l\}$ , il secondo da  $\{y_2, \dots, y_{l+1}\}$  e così via fino ad ottenere  $n - l + 1$  blocchi. Nel caso dei blocchi non sovrapposti, sempre di lunghezza  $l$ , il primo blocco è uguale ed è composto da  $\{y_1, \dots, y_l\}$  mentre il secondo da  $\{y_{l+1}, \dots, y_{2l}\}$  e si ottengono  $n/l$  blocchi. Dopo aver raggruppato i dati in blocchi, questi vengono poi ricampionati in modo casuale e con reinserimento per costruire il campione bootstrap.

Un fattore cruciale quando si utilizza il block bootstrap è la scelta della lunghezza del blocco  $l$ . Infatti se  $l$  risulta troppo piccolo, i blocchi non riescono a replicare la struttura di dipendenza poichè viene interrotta dall'inizio di un nuovo blocco. Se  $l$  è troppo grande il campione bootstrap potrebbe essere non sufficientemente casuale (MacKinnon, 2006) e tenderà ad essere influenzato dalle caratteristiche casuali del campione osservato. Per Bühlmann (2002) la lunghezza ottimale di  $l$  dipende dallo

scopo per il quale il bootstrap è utilizzato: se è usato per la distorsione, la varianza o la stima della distribuzione. Nel caso della varianza e della distorsione la lunghezza asintoticamente ottimale, ottenuta minimizzando l'errore quadratico medio asintotico (AMSE) dello stimatore del block bootstrap, è

$$l_{opt} = c \cdot n^{1/3}$$

dove  $c$  è un fattore di proporzionalità incognito.

Una variante del modello block bootstrap è lo *stationary bootstrap* (Politis e Romano, 1993). In questo modello la lunghezza di ogni blocco è scelto in modo casuale in quanto viene estratta da una distribuzione geometrica. Lo *stationary bootstrap* produce serie di dati stazionari (mentre con i metodi sovrapposti e non, le serie sono non stazionarie) e riduce la distorsione della media e della varianza delle stime bootstrap. Nonostante queste proprietà questo metodo risulta il meno efficiente. Infatti Lahiri (1999) dimostra che il AMSE del *stationary bootstrap* risulta sempre maggiore rispetto ai metodi che utilizzano i blocchi a lunghezza fissa.

Il *moving block bootstrap* è il più usato in letteratura in quanto è un metodo non parametrico per cui non impone alcuna assunzione sul processo generatore di dati e non richiede la stima dei parametri incogniti. Inoltre può essere utilizzato in qualsiasi modello dinamico anche nel caso in cui vi sia eteroschedasticità e correlazione seriale (MacKinnon, 2006). Sebbene il *moving block bootstrap* offre un'accuratezza di ordine superiore rispetto ai metodi di inferenza asintotica, il tasso di miglioramento è piuttosto ridotto (Hall *et al.*, 1995). Un'altra problematica di questo metodo è la diversa frequenza nel campione bootstrap. Infatti gli estremi del campione risultano avere una frequenza ridotta rispetto a quelli centrali: la prima e l'ultima osservazione appaiono solo in un blocco mentre la seconda e la penultima in due blocchi e così via.

### **Wild bootstrap**

Un metodo particolare di bootstrap robusto all'eteroschedasticità condizionale e non condizionale è il *wild bootstrap*. La prima versione è di Wu (1986), ridefinita poi da Liu (1988), in cui viene mostrata come questa procedura non parametrica risulta

efficace quando i termini dell'errore sono eteroschedastici.

Riprendendo in considerazione il modello nell'equazione (2.38), escludendo che gli errori siano omoschedastici, il *wild bootstrap* prende in considerazione la presenza di eteroschedasticità applicando dei disturbi sui residui della stima. In pratica, dopo aver ottenuto i residui  $\hat{\varepsilon}_t$  dalla stima, i residui bootstrap vengono ottenuti dalla seguente equazione

$$\varepsilon_t^* = \hat{\varepsilon}_t u_t^* \quad (2.39)$$

dove  $u_t^*$  rappresenta una variabile casuale univariata i.i.d. che viene generata dai dati. Per quanto riguarda  $u_t^*$  ci sono determinate condizioni che deve rispettare. Le condizioni riguardano i momenti, in particolar modo  $E(u_t^*) = 0$  e  $E(u_t^{*2}) = 1$ , più una condizione extra sul momento terzo,  $E(u_t^{*3}) = 1$  (Liu, 1988).

Queste proprietà ci assicurano che l'eteroschedasticità non condizionale è asintoticamente e correttamente simulata nel campione bootstrap come dimostrato da Cavaliere *et al.* (2010b). Per quanto riguarda la scelta di come la variabile casuale  $u_t^*$  sia distribuita ci sono diverse opzioni. Liu (1988) suggerisce di utilizzare la variabile Rademacher definita come

$$u_t^* = \begin{cases} 1 & \text{con probabilità } \frac{1}{2} \\ -1 & \text{con probabilità } \frac{1}{2}, \end{cases} \quad (2.40)$$

mentre Mammen (1993) propone l'utilizzo della seguente variabile casuale

$$u_t^* = \begin{cases} (1 - \sqrt{5})/2 & \text{con probabilità } (\sqrt{5} + 1)/(2\sqrt{5}) \\ (\sqrt{5} - 1)/2 & \text{con probabilità } (\sqrt{5} - 1)/(2\sqrt{5}). \end{cases} \quad (2.41)$$

Davidson e Flachaire (2001) mostrano come i test basati con il *wild bootstrap* con la distribuzione Rademacher hanno prestazioni migliori se la distribuzione del termine di errore è simmetrica. Quando la distribuzione è asimmetrica allora il bootstrap di Mammen (1993) è migliore nell'approssimare la probabilità di rifiuto.

Cavaliere *et al.* (2010b) utilizzano invece la normale standardizzata per  $u_t$ .

Il *wild bootstrap* permette di rimuovere la distorsione presente nelle probabilità

di rifiuto dei test, che altrimenti avverrebbe con i valori critici calcolati con i metodi asintotici, ma come tutti i metodi ha una limitazione. Infatti, per costruzione, questo metodo assume che vi sia indipendenza seriale nel termine di errore, per cui esso non è robusto se vi è autocorrelazione nei dati. Quindi, se la statistica per cui il *wild bootstrap* è stato usato non può essere corretta per l'autocorrelazione, questa procedura produrrà risultati inferenziali asintoticamente non validi (Smeekes e Urbain, 2015).

### Procedure miste

In letteratura sono presenti anche modelli bootstrap alternativi a quelli standard. Shao (2010) propone un'estensione della procedura del *wild bootstrap* di Wu (1986) nelle serie storiche permettendo alle variabili ausiliari di essere dipendenti in modo tale che il campione bootstrap sia capace di simulare la struttura di dipendenza delle serie originali.

Riprendendo l'equazione (2.39)

$$\varepsilon_t^* = \hat{\varepsilon}_t u_t^* \quad (2.42)$$

che mostra come viene generato il campione nel *wild bootstrap*, si può notare che la correlazione contemporanea viene catturata dal termine  $u_t^*$  che è scalare, quindi qualsiasi autocorrelazione all'interno di  $\hat{\varepsilon}_t$  sarebbe persa. Ciò implica che la matrice delle covarianze di lungo periodo non viene correttamente replicata dalla procedura. Tuttavia, a meno che la statistica, per cui il *wild bootstrap* è usato, può essere corretta per l'autocorrelazione, per esempio aumentando i numeri dei ritardi o stimando non parametricamente la matrice delle covarianze di lungo periodo, la procedura restituirà un'inferenza asintoticamente non valida. Smeekes e Urbain (2015) consigliano di modificare il *wild bootstrap* in modo da considerare anche la presenza dell'autocorrelazione piuttosto che affidarsi ad una correzione della statistica.

Tra i modelli presenti in letteratura troviamo: il *block-wise wild bootstrap* e il *dependent wild bootstrap*. Il *block-wise wild bootstrap* proposto da Shao (2011) viene costruito partendo da un *block bootstrap* non sovrapposto. Infatti come prima cosa

viene scelta la lunghezza dei blocchi  $l$ , tale per cui  $l \leq T$  poi il loro numero attraverso  $n = T/l$ , che per semplicità viene supposto intero. Viene generata una serie  $\{v_i^*\}_1^n$  estraendo  $n$  osservazioni i.i.d. da una normale standardizzata  $N(0, 1)$ . Riprendendo l'equazione (2.39), il campione bootstrap sarà costituito ponendo  $u_t^* = v_t^*$  in modo che il primo blocco verrà ottenuto come  $\varepsilon_t^* = \hat{\varepsilon}_t v_1^*$  per  $t = 1, \dots, l$ , il secondo sarà  $\varepsilon_t^* = \hat{\varepsilon}_t v_2^*$  per  $t = l + 1, \dots, 2l$ , e così via. Quindi la dipendenza all'interno del blocco è preservata, ma tra i blocchi viene persa.

Il *dependent wild bootstrap*, sempre di Shao (2010), è un'estensione del *block bootstrap* con blocchi sovrapposti. La procedura è simile a quella precedente: viene scelto  $l$  per determinare l'estensione della dipendenza, dopo di che viene generata una serie  $u_t^*$  tale per cui la  $Cov(u_s^*, u_t^*) = K(\frac{s-t}{l})$  dove  $K(\cdot)$  è una funzione kernel con  $K(0) = 1$  e  $K(x) = 0$  per tutte le  $x \geq 1$ . Quindi le osservazioni  $\{u_t^*\}_1^T$  sono estratte da una  $N(0, \Sigma)$ , dove  $\Sigma = (K(\frac{i-j}{l}))_{i,j=1}^T$ . Per quanto concerne la scelta della funzione kernel può essere fatta tra quella di Bartlett, di Parzen o di una kernel spettrale quadratica. Il *dependent wild bootstrap* diversamente dal *wild bootstrap* utilizza variabili casuali  $\{u_t^*\}_1^T$  dipendenti che sono capaci di catturare la dipendenza del campione osservato.

Per la validità asintotica le condizioni richieste per entrambi i modelli affinché la procedura sia asintoticamente valida sono prettamente due: quella di *strong mixing* e la condizione 9 di Cavaliere *et al.* (2010b) per la serie  $\varepsilon_t$ .

## 2.4.2 Bootstrap per modelli di cointegrazione

Nel corso degli ultimi anni la metodologia bootstrap è diventata uno strumento utile per risolvere alcuni dei problemi di inferenza in ambito delle serie storiche.

Un esempio è l'applicazione del bootstrap nei processi integrati. Paparoditis e Politis (2005), Parker *et al.* (2006) e Phillips (2010) mostrano che calcolare i valori critici dei test per le radici unitarie con questa metodologia risulti più accurato rispetto ai metodi asintotici e permette di eliminare eventuali distorsioni provocate dal campione.

Inoltre Goncalves e Kilian (2004), Kim *et al.* (2002a), Buseti e Taylor (2003), Cavaliere (2005), Cavaliere e Taylor (2008) e Cavaliere *et al.* (2012) analizzano gli effetti della volatilità variabile sui test univariati per la presenza di radici unitarie.

Questi autori dimostrano che alcuni dei più comuni test mostrano significanti *size distortion* nel momento in cui viene meno l'ipotesi di volatilità costante nel tempo. L'applicazione del *wild bootstrap* al test ADF (Augmented Dickey-Fuller) (Cavaliere e Taylor, 2008) consente di ovviare a eventuali problemi di distorsione nei test di radice unitaria provocati dalla varianza dell'innovazione non costante nel tempo.

La letteratura si è incentrata anche nell'uso di metodi di simulazioni nel caso in cui vi sia la presenza di cointegrazione lineare. Formalmente i ricercatori hanno utilizzato le usuali procedure bootstrap seppur con alcune modifiche dovute principalmente alla presenza di processi non stazionari.

Chang *et al.* (2006) propongono l'utilizzo del *sieve bootstrap* per ricampionare da un modello di cointegrazione. Se si considera una versione generale della relazione di cointegrazione, questa può essere rappresentata da

$$\begin{aligned} y_t &= \theta x_t + v_t \\ x_t &= x_{t-1} + w_t \end{aligned} \tag{2.43}$$

dove  $y_t$  ha dimensione  $r \times 1$  mentre  $x_t$  è  $(n - r) \times 1$ . La matrice di cointegrazione  $\beta = (-\theta, I_r)'$  è composta dalla matrice dei parametri  $\theta$  di dimensione  $r \times (n - r)$  e  $I_r$  è la matrice identità  $r \times r$ . Infine  $\varepsilon_t = (v_t', w_t')'$  è un vettore di processi stazionari.

Il metodo *sieve bootstrap* richiede di approssimare  $\varepsilon_t$  con un VAR di ordine  $q$  in modo da avere

$$\varepsilon_t = \Phi_1 \varepsilon_{t-1} + \Phi_2 \varepsilon_{t-2} + \dots + \Phi_q \varepsilon_{t-q} + u_t; \tag{2.44}$$

naturalmente la scelta di  $q$  avviene con i criteri informativi.

Le fasi dell'algoritmo prevedono:

1. di stimare la prima equazione della (2.43) in modo da ottenere i residui  $\hat{v}_t$  e calcolarsi  $w_t = \Delta x_t$  in modo da avere  $\hat{w}_t$ ;
2. di stimare il VAR dell'equazione (2.44) per ottenere  $\hat{u}_t$  e i parametri  $\hat{\Phi}_i$  con  $i = 1, \dots, q$ ;

3. di ricampionare ( $u_t^*$ ) dai residui centrati con

$$\hat{u}_t - \frac{1}{T} \sum_{t=1}^T \hat{u}_t$$

e successivamente di ricostruire il campione bootstrap ( $\varepsilon^*$ ) ricorsivamente da

$$\varepsilon_t^* = \hat{\Phi}_1 \varepsilon_{t-1}^* + \dots + \hat{\Phi}_q \varepsilon_{t-q}^* + u_t^* \quad (2.45)$$

dove i valori iniziali di  $\varepsilon_t^*$  per  $t = 0, \dots, 1 - q$  sono presi dai dati osservati ( $\varepsilon_t$ );

4. ed infine ottenere  $x_t^*$  da  $\varepsilon_t^* = (v_t^{*'}, w_t^{*'})'$  con

$$x_t^* = x_0 + \sum_{k=1}^t w_k^*$$

e poi  $y_t^*$  con

$$y_t^* = \hat{\theta} x_t^* + v_t^*.$$

Due importanti osservazioni devono essere fatte. In primo luogo, la scelta della stima di  $\theta$  nel punto 1 deve essere effettuata tenendo conto del metodo più efficiente possibile, in quanto questa decisione influenza tutta la procedura bootstrap. Chang *et al.* (2006) consigliano di stimare il parametro considerando ed inserendo tutte le restrizioni del test d'ipotesi. Un secondo punto concerne la dipendenza delle serie generate  $x_t^*$  e  $\varepsilon_t^*$  dai rispettivi valori iniziali ( $\varepsilon_0, \dots, \varepsilon_{1-q}$  e  $x_0$ ). Per risolvere questo problema si può scegliere se generare un numero considerevolmente alto di  $\varepsilon_t^*$  o se introdurre una costante in modo da rendere il campione bootstrap indipendente.

Chang *et al.* (2006), assumendo l'esistenza dei momenti per  $u_t$  fino al quarto ordine e l'invertibilità del processo di  $\varepsilon_t$ , dimostrano la consistenza del bootstrap per le statistiche test utilizzando il metodo dei minimi quadrati. In aggiunta evidenziano come questa procedura permette di migliorare l'efficienza asintotica delle statistiche riducendo la distorsione da campioni finiti. Un ultimo punto sul quale bisogna soffermarsi è il modello di riferimento su cui è stata svolta la procedura e la dimostrazione della consistenza del bootstrap. Sebbene Chang *et al.* (2006) si siano soffermati solo sulla relazione di cointegrazione, questa metodologia può essere facil-

mente estesa, con dovute modifiche ed estensioni, anche ai casi più particolari come l'inserimento di un nucleo deterministico all'interno della relazione e/o a modelli più sofisticati come l'ECM, come mostrato da Palm *et al.* (2010)

Sullo stesso modello di Chang *et al.* (2006), anche Jentsch *et al.* (2014) estendono il *block bootstrap* ai modelli di cointegrazione. Riprendendo la relazione di cointegrazione descritta dall'equazione (2.43) e considerando un campione composto da  $T$  osservazioni, l'algoritmo è composto da 4 diverse fasi:

1. viene stimato  $\hat{v}_t$  e  $\hat{\theta}$  poi si calcola  $w_t = \Delta x_t$  in modo da ottenere il vettore  $\hat{\varepsilon}_t = (\hat{v}_t', w_t')'$ .
2. Si sceglie la lunghezza dei blocchi  $l$  e un numero di blocchi  $k$  tale per cui  $l \cdot k \geq T$ . Si estraggono in modo i.i.d.  $k$  volte da una variabile casuale uniforme discreta  $U \sim (0, T - l)$  in modo da ottenere  $\{i_0, i_1, \dots, i_{k-1}\}$ .
3. Dopo aver suddiviso  $\hat{\varepsilon}_t$  in  $k$  blocchi sovrapposti viene generato il campione bootstrap  $\varepsilon_1^*, \dots, \varepsilon_{kl}^*$  da

$$\varepsilon_{ml+j}^* = \hat{\varepsilon}_{i_m+j} - \frac{1}{T-l+1} \sum_{t=0}^{T-l} \hat{\varepsilon}_{t+j}$$

con  $j = 1, 2, \dots, l$  e  $m = 0, 1, \dots, k-1$

4. Da  $\varepsilon_t^* = (v_t^{*'}, w_t^{*'})'$  si genera il campione bootstrap  $(y_t^{*'}, x_t^{*'})'$  con

$$x_t^* = x_0 + \sum_{k=1}^t w_k^*$$

$$y_t^* = \hat{\theta} x_t^* + v_t^*$$

dove  $t = 1, \dots, kl$ .

La pseudo serie storica generata da questa procedura conserva la struttura di cointegrazione del processo sottostante.

Per quanto i risultati asintotici Jentsch *et al.* (2014) prendono in considerazione la stima effettuata con il metodo dei minimi quadrati. Essi dimostrano che sotto

l'ipotesi che la serie  $\varepsilon_t$  sia *strong mixing* (si veda la Definizione 1) il *block bootstrap* risulta consistente nell'approssimare la matrice di cointegrazione. Inoltre la teoria asintotica, descritta per questa procedura, è valida non solo per il caso lineare, ma anche per tutte le tipologie di termini di errore che soddisfano la condizione *strong mixing*.

Cavaliere *et al.* (2010b) invece mostrano come la procedura *wild bootstrap* può essere applicata ai modelli di cointegrazione concentrandosi in particolar modo sul test del rango di cointegrazione di Johansen (1995).

Riprendendo in esame un modello VECM come nell'equazione (1.5)

$$\Delta y_t = \mu_t + \alpha\beta'y_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t \quad (2.46)$$

dove  $y_t$  è un processo multivariato di dimensione  $n \times 1$ ,  $\mu_t$  è il nucleo deterministico che può essere composto da una costante e/o da un trend lineare.  $\Gamma_i$  sono matrici  $n \times p$  contenenti i coefficienti riferiti a  $\Delta y_{t-i}$  mentre,  $\alpha$  e  $\beta$ , di dimensione  $n \times r$ , rappresentano rispettivamente il *loading factor* e la matrice di cointegrazione. Il termine di errore  $\varepsilon_t$  è un vettore di differenze di martingala di dimensione pari a  $n \times 1$  la cui varianza condizionale è variabile nel tempo.

L'algoritmo per generare un campione con il *wild bootstrap* è così composto:

1. vengono stimati i residui  $\hat{\varepsilon}_t$  e i parametri  $\hat{\beta}, \hat{\alpha}, \hat{\Sigma}_i, \hat{\mu}_t$  dall'equazione (2.46) secondo il modello imposto dall'ipotesi nulla.
2. Si controlla che l'equazione  $|\hat{A}(z)| = 0$ , con

$$\hat{A}(z) = (1 - z)I_n - \hat{\alpha}\hat{\beta}'z - \sum_{i=1}^{k-1} \hat{\Gamma}_i(1 - z), \quad (2.47)$$

abbia tutte le radici al di fuori del cerchio unitario.

3. Dai residui centrati

$$\tilde{\varepsilon}_t = \hat{\varepsilon}_t - \frac{1}{T} \sum_{i=1}^T \hat{\varepsilon}_i$$

viene costruito il campione bootstrap da  $\varepsilon_t^* = \tilde{\varepsilon}_t u_t^*$ , dove  $\{u_t^*\}_0^T$  è una sequenza di osservazioni i.i.d. estratte da una Normale standardizzata.

4. Da  $\varepsilon_t^*$  viene costruito il campione bootstrap ricorsivamente con

$$\Delta y_t^* = \tilde{\mu}_t + \tilde{\alpha} \tilde{\beta}' y_{t-1}^* + \sum_{i=1}^{k-1} \tilde{\Gamma}_i \Delta y_{t-i}^* + \varepsilon_t^* \quad (2.48)$$

per i valori iniziali di  $y_t^*$  con  $t = -k+1, \dots, 0$  vengono presi i dati effettivamente osservati da  $y_t$ .

Il punto 2 dell'algoritmo, che verifica che la serie sia I(1) con rango di cointegrazione  $r$ , serve anche a controllare se è verificata la Condizione 8 per la validità asintotica della procedura bootstrap. Per quanto riguarda la scelta della distribuzione di  $u_t$  nel punto 3, questa non è vincolata alla sola distribuzione normale standardizzata ed infatti è possibile sostituirla con la distribuzione Rademacher o quella consigliata da Mammen (1993) (si veda l'equazione (2.41)). Da un punto di vista asintotico questa scelta è ininfluente sui test bootstrap ed inoltre non sussistono sostanziali differenze tra le proprietà nei campioni finiti sui test basati sulle tre distribuzioni (Cavaliere *et al.*, 2010b).

Le assunzioni con cui Cavaliere *et al.* (2010b) dimostrano la validità asintotica del *wild bootstrap* sono:

**Condizione 8.**  $|A(z)| = 0$  (equazione (2.47)) abbia tutte le radici al di fuori del cerchio unitario e che le matrici  $\alpha$  e  $\beta$  abbiano rango pieno pari a  $r$ .

**Condizione 9.** Il processo  $\varepsilon_t$  può essere scritto (rappresentato) come  $\varepsilon_t = \sigma_t z_t$  dove:

(a)  $\sigma_t = \sigma(t/T)$ , dove  $\sigma(\cdot)$  è un elemento non stocastico di  $\mathcal{D}_{\mathbb{R}^{p \times p}}[0, 1]$  tale per cui  $\sigma(u)\sigma(u)' > 0$  per tutte le  $u \in [0, 1]$ .

(b)  $z_t$  è un vettore di dimensione  $p \times 1$  di differenze di martingala con varianza condizionale pari a  $E(z_t z_t' | \mathcal{I}_{t-1})$  e soddisfa  $E\|z_t\|^{4r}$  per  $r > 2$ .

Un'implicazione della Condizione 8 è che  $\Delta y_t$  e  $\beta' y_t$  possono essere riscritti come processi lineari espressi in termini di  $\varepsilon_t$ . Cavaliere *et al.* (2010b) mostrano come

questi processi sono “stabili”, in quanto i coefficienti decadono esponenzialmente. Infatti nel caso in cui la varianza non condizionale di  $\varepsilon_t$  fosse costante, questi processi “stabili” possono essere considerati stazionari.

La Condizione 9 implica che  $\varepsilon_t$  sia una sequenza di differenze di martingala e la sua dinamica è rappresentata da un modello molto generale. Tutto ciò permette di considerare una vasta classe di modelli che descrive il comportamento della matrice delle covarianze, anche quando si considerano salti singoli o multipli nella varianza o variazioni in linea con i modelli GARCH e ARCH. Infatti uno dei vantaggi più importanti del *wild bootstrap* è che non impone nessun modello parametrico per la volatilità.

Cavaliere *et al.* (2010b) evidenziano come per una più corretta analisi per la scelta del rango di cointegrazione in un contesto in cui la varianza dell’errore non è costante sia utile implementare la procedura bootstrap. Infatti essi dimostrano che la distribuzione asintotica delle statistiche PLR (*pseudo likelihood ratio*) sul rango di cointegrazione di Johansen (1995) non sono pivotali in presenza di eteroschedasticità non condizionale e dipendono dai parametri di quest’ultimo. Cavaliere *et al.* (2010a) forniscono anche una trattazione a parte nell’ipotesi in cui gli errori siano condizionalmente eteroschedastici ma non condizionalmente omoschedastici. Essi evidenziano come i test PLR con un normale ricampionamento bootstrap, in questo contesto, siano asintoticamente validi per il test di Johansen (1995), ma non lo sia in presenza di eteroschedasticità non condizionale.

Per quanto riguarda le stime con la MLE, Boswijk *et al.* (2016) mostrano come la distribuzione asintotica dei test sui parametri di lungo periodo e/o sui coefficienti di aggiustamento, quindi la matrice di cointegrazione e dei *loading factor*, non è pivotale quando la volatilità dei dati non è costante nel tempo. Per far fronte a questo problema questi autori propongono l’utilizzo del *wild bootstrap* nei test, W e LR, dimostrandone anche la sua validità asintotica.

Nelle simulazioni di Monte Carlo e sui dati reali si osserva in modo marcato le differenze tra i risultati condotti con il test *wild bootstrap* e quelli ottenuti con i metodi asintotici. Infatti nelle simulazioni effettuate su una varietà di modelli sia

eteroschedastici che omoschedastici da Cavaliere *et al.* (2010b) e da Boswijk *et al.* (2016) i risultati suggeriscono che i test bootstrap, sia per il rango di cointegrazione che per i coefficienti, hanno buone prestazioni nei campioni finiti ed evitano il problema “over-size” che possono avvenire con i metodi standard.

### 2.4.3 Bootstrap per modelli non lineari

Generalmente nella trattazione dei modelli non lineari si riscontrano diverse tipologie di difficoltà, tra cui la natura asintotica dei test. Come mostrato dai Kristensen e Rahbek (2013) in alcune situazioni può essere difficile derivare la distribuzione asintotica di un test o perché non è asintoticamente pivotale o se lo è, non converge ad una distribuzione nota. Le procedure bootstrap sono spesso utilizzate in ambito non lineare proprio per risolvere problemi di questo genere; alcuni esempi in letteratura, per quanto riguarda la verifica della presenza di non linearità per i modelli a soglia, sono Seo (2006), Hurn e Becker (2007), Nedeljkovic (2008), Dong *et al.* (2014).

Occorre fare un’importante riflessione a livello computazionale riguardo all’utilizzo del bootstrap nei modelli non lineari. Spesso nella simulazione delle serie possono presentarsi problemi con la stima dei modelli non lineari. Infatti, sebbene una procedura di stima funzioni con i dati osservati potrebbe non riuscire con uno o più campioni simulati con il bootstrap (Davidson, 2009). Ad esempio l’algoritmo per la massimizzazione potrebbe non convergere.

Per quanto riguarda la scelta dell’ipotesi nulla, l’idea è quella di utilizzare un modello lineare. La ricostruzione di un campione simulato potrebbe essere piuttosto difficoltosa se viene impostato su un modello non lineare. Inoltre ricampionare con un modello lineare risulta meno dispendiosa in termini di calcoli e quindi richiede una minor quantità di tempo.

## 2.5 Considerazioni sul modello NEC

Il modello preso in considerazione è il NEC che nel caso multivariato si presenta in questo modo:

$$\Phi(L)\Delta Y_t = \gamma' d_t + \alpha' S_{t-1} + \theta' Q_{t-1} + \lambda' C_{t-1} + \varepsilon_t \quad (2.49)$$

dove  $Y_t$  è un vettore che contiene le variabili di interesse,  $\Phi(L)$  è il polinomio dell'operatore ritardo,  $d_t$  è il nucleo deterministico, i termini  $S_t$ ,  $Q_t$  e  $C_t$  contengono i termini del polinomio di Taylor (Eq. 1.45) e  $\varepsilon_t$  è il termine di errore. Sebbene il NEC possa essere stimato attraverso il metodo della massima verosimiglianza, ciò che risulta problematico sono le distribuzioni asintotiche degli stimatori e delle statistiche test.

In un contesto generale con il modello NEC non possono essere fatte assunzioni certe riguardo alle distribuzioni limite dei test, o meglio non è possibile calcolarli a meno che non vengano imposte delle assunzioni piuttosto stringenti. Nel caso preso in esame cioè la struttura a termine dei tassi di interesse, la teoria delle aspettative razionali fornisce un'indicazione precisa riguardo alla relazione di lungo termine tra i tassi. Questo ci permette di lavorare in un contesto in cui il vettore di cointegrazione può essere considerato noto.

Come dimostrato da Kristensen e Rahbek (2013), lavorare con vettore di cointegrazione noto o incognito nei modelli di cointegrazione non lineare è una condizione determinata nella scelta di quale sia la distribuzione asintotica più idonea da utilizzare. Sebbene nel contesto analizzato si possano sfruttare i risultati asintotici standard, sarebbe opportuno fare un'ulteriore verifica empirica sulla validità dei risultati ottenuti precedentemente. Il metodo bootstrap ci offre una via alternativa ed inoltre permette alcuni raffinamenti asintotici inferenziali. In aggiunta questa procedura è capace di tenere in considerazione l'impatto di diversi fattori come le condizioni iniziali, l'ampiezza del campione e in particolar modo la distribuzione dell'errore che, sebbene generalmente non influiscono sulle distribuzioni asintotiche, possono influire sui risultati.

Per quanto riguarda la scelta della metodologia bootstrap più idonea da utilizzare per le serie storiche, si deve tener conto di diversi aspetti tra i quali l'autocorrelazione, la stazionarietà del processo e la presenza di eteroschedasticità. Dato che il *sieve bootstrap* non può essere applicato nei processi non lineari, quindi anche quel-

li eteroschedastici, la scelta può ricadere sui metodi del *block bootstrap* e del *wild bootstrap*. Inoltre, sebbene il *wild bootstrap* potrebbe non funzionare bene come gli altri metodi bootstrap sui residui quando gli errori sono in realtà omoschedastici, il costo di assicurarsi contro l'eteroschedasticità risulta essere piuttosto piccolo (Davidson, 2009). Data la semplicità di calcolo di quest'ultima procedura rispetto al *block bootstrap*, che richiede di suddividere il campione in determinati intervalli, la scelta ricade sul *wild bootstrap*.

In alcuni lavori empirici viene verificata l'ipotesi di variabilità nella volatilità condizionale e non nei dati. In un contesto di dati macroeconomici e finanziari non può essere scartata l'ipotesi che vi sia eteroschedasticità condizionale o volatilità condizionale non costante, come verificato da Hamori e Tokihisa (1997) sulle serie univariate. Hansen (1992) analizza le proprietà dei modelli di cointegrazione lineari più noti in letteratura macroeconomica. Anche lui riscontra la presenza di eteroschedasticità condizionale nelle serie nei tassi a lungo e a breve termine dei titoli di stato americani. Tenendo conto di questo aspetto Hansen (1995), Xu e Phillips (2008), Cavaliere *et al.* (2010b), Cavaliere *et al.* (2014) e Boswijk *et al.* (2016) analizzano l'impatto che la volatilità non stazionaria può avere nei metodi di inferenza asintotica standard quando si applica a questi dati.

In particolare il lavoro di Cavaliere *et al.* (2010b) si sofferma sul test per il rango di cointegrazione per i titoli di stato americani a diversa scadenza. L'analisi empirica ha messo in luce i differenti risultati che si possono ottenere con l'applicazione del test di Johansen (1995) rispetto a quello del *wild bootstrap*.

## 2.6 Conclusioni

In questa sezione sono state analizzate alcune delle teorie asintotiche che possono in qualche modo essere collegate al modello NEC.

Diversamente dai modelli non lineari con variabili stazionarie, la dimostrazione della consistenza per i modelli di cointegrazione non lineare è più difficile da stabilire in modo generale. Infatti dipende essenzialmente dalla forma della funzione di aggiustamento non lineare, in particolare quando le variabili assumono valori infini-

tesimamente grandi (Park e Phillips, 2001). Gran parte della letteratura si è mossa sui modelli di cointegrazione a soglia (Hansen e Seo (2002), de Jong (2002), Bec e Rahbek (2004), Seo (2009), ecc.), in quanto le funzioni che approssimano il meccanismo di aggiustamento sono limitate. Nel primo lavoro di Kristensen e Rahbek (2010) vengono considerate funzioni illimitate ma, per valori dei termini di correzioni tendenti all'infinito, la funzione diventa lineare. Anche Escribano e Mira (2002), che considerano modelli alternativi e molto simili al NEC, impongono delle condizioni sulla funzione di aggiustamento in particolare che la sua derivata prima sia limitata all'interno di specifici valori (condizione di stabilità).

Fra i diversi teoremi presenti in letteratura quello che più si adatta al modello NEC è il teorema di Kristensen e Rahbek (2013). Infatti nel loro lavoro derivano la distribuzione asintotica dei test LM, LR e Wald per i modelli in cui la funzione di aggiustamento è differenziabile e polinomialmente limitata. Inoltre questi autori distinguono in modo particolare la situazione in cui si lavori nel caso in cui la matrice di cointegrazione ( $\beta$ ) sia nota oppure no. Infatti Kristensen e Rahbek (2013) dimostrano come le distribuzioni limite siano influenzate da una componente non stazionaria e che quest'ultima è presente nel caso in cui si lavori in un contesto in cui  $\beta$  venga stimato. In caso contrario le distribuzioni limite sono uguali a quelle del caso standard.

Nella seconda parte ho elencato alcuni dei metodi più comuni utilizzati in letteratura sulle serie storiche per le simulazioni bootstrap. Questa metodologia è molto utile soprattutto quando l'inferenza è piuttosto complessa e i risultati asintotici non sono disponibili. In ambito delle serie storiche le usuali procedure bootstrap non possono essere applicate in quanto i campioni simulati non replicano la stessa struttura di dipendenza dei dati di provenienza. Inoltre in base alle caratteristiche del campione (presenza eteroschedasticità e/o autocorrelazione) e del modello utilizzato (lineare o no) è necessario applicare la giusta metodologia bootstrap.

Per il caso empirico preso in esame, i rendimenti dei titoli di stato americani, la letteratura si è mossa in direzione del *wild bootstrap* (Cavaliere e Taylor (2008),

Cavaliere *et al.* (2010b), Cavaliere *et al.* (2012) e Boswijk *et al.* (2016)). La scelta è dipesa dalla natura finanziaria dei dati caratterizzati da volatilità non stazionaria per cui questa metodologia di simulazione risulta la più idonea in questo contesto. C'è infine da sottolineare che l'applicazione del metodo *wild bootstrap* non deve essere limitato ai soli test di specificazione, ma deve essere anche estesa nell'analisi preliminare (test radici unitarie, test rango cointegrazione, ecc.).

## Capitolo 3

# Analisi Empirica

### 3.1 Introduzione

Come già visto nel primo capitolo, i risultati ottenuti in precedenza con un modello di cointegrazione lineare sono discordanti con la teoria delle aspettative razionali (Campbell e Shiller, 1987). Infatti, il coefficiente relativo allo spread assume segno opposto rispetto a quello imposto dalla teoria ed inoltre, in alcuni casi, gli spread risultano essere non stazionari e i test condotti sul rango della matrice di cointegrazione rifiutano l'ipotesi dettata dalla teoria (Giese, 2008; Cavaliere *et al.*, 2010b; Boswijk *et al.*, 2016). Autori come Ang e Piazzesi (2003), Valente *et al.* (2004), Carriero *et al.* (2004) hanno suggerito che i problemi nell'applicazione dei modelli lineari provengono dal fatto che la dinamica della *yield curve* sia influenzata da fattori macroeconomici. Quindi, per una corretta specificazione, si suppone che un modello deve considerare anche la relazione esistente tra le variabili macroeconomiche e i tassi di interesse.

Un altro filone della letteratura presume che l'aggiustamento verso l'equilibrio di lungo periodo, implicito nel modello di cointegrazione, non avvenga in modo continuo (Bec e Rahbek, 2004; Seo, 2003; Krishnakumar e Neto, 2009; Cai *et al.*, 2015). L'ipotesi principale è che le serie dei tassi di interesse siano cointegrate solo quando questi sono molto distanti dalla relazione di equilibrio mentre, se sono molto vicine all'equilibrio, non vi è alcuna presenza di un meccanismo di aggiustamento. In altri

termini, quando lo spread è compreso in un intervallo di valori, non vi è cointegrazione, ma il contrario accade quando lo spread assume valori al di fuori di questa banda. Tale ipotesi trova supporto teorico nel fatto che spesso nell'analisi economica non si tiene conto di elementi che possono condurre verso frizioni nel meccanismo di aggiustamento come i costi di transizione, l'andamento del ciclo economico e gli interventi di politica monetaria. Le stime condotte con un modello a soglia da Balke e Fomby (1992) mostrano risultati soddisfacenti sia da un punto di vista teorico, sia da quello metodologico. Infatti, i coefficienti dello spread nel regime in cui si verifica l'aggiustamento risultano congruenti con quelli previsti dalla teoria delle aspettative cioè un aumento dello spread provoca un innalzamento dei tassi di interesse futuri sia di breve che di lungo periodo. Inoltre, i test di specificazione evidenziano chiaramente la presenza di una relazione non lineare all'interno del modello.

Questo terzo capitolo raccoglie i risultati derivanti dall'applicazione empirica del modello NEC alla struttura a termine dei tassi di interesse secondo la teoria delle aspettative razionali. L'obiettivo principale è di verificare se questo modello non lineare sia in grado di descrivere la dinamica della *yield curve* e se sia capace di catturare i principali interventi di politica monetaria, le crisi e gli shock di questi ultimi trent'anni nell'economia statunitense.

Nella prima parte presenterò una breve descrizione dei dati a disposizione seguita da un'analisi storica sugli eventi che hanno più influenzato l'economia americana e le seguenti azioni di politica monetaria messe in atto per fronteggiarli. Successivamente verificherò la teoria delle aspettative sia attraverso le implicazioni "low frequency" (restrizioni sulla matrice di cointegrazione) che quelle di breve termine (*cross-equation constraints*) le quali sono più strettamente legate alla dinamica della struttura di aggiustamento.

In seguito utilizzerò lo stesso approccio adottato da Lucchetti e Palomba (2009), ma su un data set che tiene conto degli ultimi avvenimenti come la crisi dei mutui *sub-prime*. Quindi applicherò un modello di cointegrazione non lineare in cui la funzione di aggiustamento è approssimata da un polinomio cubico. Verranno effettuate anche un'analisi comparativa tra i modelli che possono presentarsi come i maggio-

ri competitor. Infine nell'ultimo paragrafo verrà effettuata un'analisi generale sui risultati ottenuti con le stime. In particolare verrà discusso il comportamento del meccanismo di aggiustamento non lineare di fronte agli eventi più importanti dell'economia americana.

Gli obiettivi di questo capitolo sono:

1. di verificare che il modello NEC riesca a replicare gli stessi risultati di Lucchetti e Palomba (2009) in un contesto che tenga in considerazione gli ultimi eventi economici e i relativi interventi di politica monetaria. Il data set originario di questi autori arriva a coprire solo il periodo iniziale della crisi economica, per cui non è stato possibile osservare come il modello si comporti in questa situazione. Infatti in questi ultimi anni la politica monetaria statunitense, per far fronte alla recessione, ha utilizzato strumenti non convenzionali, quindi è interessante verificare se il meccanismo non lineare sia capace di catturare tutte le manovre che la banca centrale americana ha svolto in questo periodo.
2. Attraverso l'analisi dei tassi di interesse mensili, un ulteriore obiettivo è quello di esaminare se il modello possa sinteticamente descrivere gli interventi di politica monetaria anche nei dati aventi diversa frequenza. La motivazione dell'utilizzo dei dati mensili replica essenzialmente la scelta seguita da gran parte della letteratura (Campbell e Shiller (1987), Balke e Fomby (1992), ecc.) che hanno condotto le loro verifiche sulla struttura a termine dei tassi di interesse.
3. La letteratura fornisce diverse evidenze empiriche riguardo alla presenza di eteroschedasticità (condizionale e non) nelle serie finanziarie, soprattutto nell'ambito dei tassi di interesse dei bond americani, come ad esempio Hamori e Tokihisa (1997) e Goncalves e Kilian (2004). Quindi è necessario ricorrere a metodologie robuste ai cambi della varianza e che tengano conto anche della non linearità del modello. Inoltre l'evidenza empirica diventa più favorevole alla teoria se utilizzati test e modelli robusti all'eteroschedasticità (Cao *et al.*, 2014; Gunnella, 2015). Un altro obiettivo è di utilizzare metodi più aggiornati sia nella fase di verifica della teoria delle aspettative che in quella di specificazione rispetto al lavoro precedente di Lucchetti e Palomba (2009). Diver-

samente da loro ho utilizzato procedure che non impongono alcuna dinamica alla volatilità nel modello NEC. Per far ciò ho effettuato l'analisi seguendo le procedure di Cavaliere *et al.* (2010b) e di Boswijk *et al.* (2016).

## 3.2 I dati

La struttura a termine dei tassi di interesse descrive la relazione che intercorre tra i rendimenti dei titoli e la loro rispettiva scadenza per cui le obbligazioni che meglio si prestano a rappresentare questa relazione sono quelle del debito sovrano. Infatti, quando viene presa in considerazione la relazione tra la scadenza di un titolo e il suo rendimento, possono subentrare elementi di rischio o clausole contrattuali che possono influenzare tale relazione. Per esempio, i titoli del tesoro americani presentano un alto grado di liquidità, quindi anche un livello molto basso di rischio. Inoltre, da un punto di vista prettamente econometrico, il numero delle osservazioni è molto elevato ed è possibile scegliere anche la loro frequenza.

Per quanto concerne la scelta delle diverse scadenze dei titoli governativi, ho seguito Lucchetti e Palomba (2009). L'utilizzo di variabili rappresentative del breve, medio e lungo periodo ci permette di avere una chiara visione dell'efficacia degli interventi dei *policy maker* sui rendimenti dei titoli governativi e dell'influenza sulle aspettative degli operatori. Quindi, come variabile di breve periodo, è stato scelto il rendimento dell'obbligazione avente scadenza di tre mesi ( $rs_t$ ), in quanto questo titolo, come dimostrato da Cook e Hahn (1989), risponde in modo immediato ai movimenti del tasso dei *federal funds*. Per le variabili di medio e lungo periodo sono stati scelti rispettivamente i bond con scadenza a due anni ( $rm_t$ ) e a dieci anni ( $rl_t$ ). Infine, come variabili dello spread sono state prese in considerazione la differenza tra il tasso di rendimento del bond a dieci anni e quello a tre mesi ( $sl_t = rl_t - rs_t$ ) e la differenza tra i tassi di interesse a medio termine e a breve termine, cioè tra i rendimenti a due e tre mesi ( $sm_t = rm_t - rs_t$ ).

La scelta di procedere con un'analisi su dati a frequenza sia settimanale che mensile è dovuta dall'esigenza di poter meglio catturare gli aggiustamenti condotti dalla

politica monetaria. Infatti, i dati con frequenza più alta sono influenzati da ulteriori fattori come gli effetti del fine settimana, gli annunci di dividendi e la chiusura del mercato che possono avere effetto sull'analisi in quanto l'informazione non arriva in modo uniforme nel tempo (Ghysels *et al.*, 1996). Inoltre, gran parte della letteratura (Campbell e Shiller (1987), Balke e Fomby (1992), Giese (2008), ecc) ha condotto le sue analisi su un data set con frequenza mensile.

Il campione copre il periodo che va dal 12 Ottobre 1982 fino al 23 Febbraio 2016. I dati hanno frequenza settimanale e sono  $T = 1742$ , mentre le serie mensili contengono 401 osservazioni. La scelta del periodo di campionamento segue una motivazione ben precisa: in primo luogo è necessario precisare che, al fine di assicurare l'omogeneità del campione, le osservazioni sono state prese in un periodo in cui gli obiettivi di politica monetaria negli USA non sono stati cambiati radicalmente. Infatti, Hall *et al.* (1992) e Hsu e Kugler (1997) hanno mostrato come le stime e i test effettuati in ogni sotto campione, che corrispondono ai periodi in cui la Fed ha adottato nuove procedure, non conducono agli stessi risultati. Il campione parte dal 12 Ottobre 1982, subito dopo che la *Federal Open Market committee* (FOMC) ha adottato una diversa procedura operativa. In definitiva la FOMC cambia il suo obiettivo principale passando dall'aggregato monetario M1 ai *federal funds rate*. In secondo luogo, il periodo preso in esame comprende i principali interventi effettuati dalla autorità monetaria durante uno dei periodi più turbolenti dell'economia americana. Infatti, come accennato in precedenza, gli interventi dei *policy maker* avvengono in modo sporadico e solo quando gli spread assumono valori inconsueti. Quindi, considerare una finestra temporale che comprende eventi come il crollo dello stock market dell'87, la crisi asiatica, lo scoppio della bolla *dot.com* e della bolla immobiliare che hanno richiesto un massiccio intervento dell'autorità monetaria, ci permette di catturare meglio questi interventi e di identificarli in modo più concreto attraverso gli effetti del meccanismo di aggiustamento non lineare.

Al fine di verificare la validità previsiva dell'analisi proposta è stato tenuto fuori dalla stima un anno di osservazioni. Per l'esattezza le ultime 12 osservazioni per le serie mensili e 52 per le serie con frequenza settimanale. Si tenga presente che

nella trattazione dei modelli non lineari, spesso si riscontrano problemi quando si valutano le capacità previsive rispetto ai modelli lineari. Infatti nonostante la stima fornisca ottime approssimazioni di alcune serie macroeconomiche, quando si passa alla verifica *out-of-sample*, Stock e Watson (1998) e Granger (2001) evidenziano come i modelli lineari si dimostrino più robusti e posseggano una maggiore o uguale capacità previsiva rispetto ai modelli non lineari. Con ciò si vuole sottolineare la difficoltà nel trovare un modello non lineare adatto a descrivere i dati. Un esempio è il lavoro di Marcellino (2002) nell'analisi di alcune serie come il tasso di disoccupazione, un modello non lineare produce previsioni migliori rispetto al suo concorrente lineare. In conclusione, in alcune serie macroeconomiche vi sono chiari elementi di non linearità, ma è piuttosto difficile trovare il modello più adatto e i risultati dall'analisi *out-of-sample* forniscono una prova più che consistente di questo fatto.

### 3.2.1 Analisi del contesto storico

In questo paragrafo presenterò una breve trattazione sui principali avvenimenti storici ed economici che coprono il periodo che va da ottobre del 1982 sino a oggi. In particolare, mi soffermerò sugli eventi di maggior importanza per l'economia americana e sugli interventi effettuati dalla Fed che hanno avuto una qualche relazione con il mercato dei bond governativi.

Considerando la Figura 3.1 è possibile notare come i tassi seguano un trend generalmente ribassista.

Senza dubbio l'economia americana sembra essere intrisa di avvenimenti che hanno condotto ad un lento, ma marcato declino dei tassi di interesse.

Nel 1982 l'economia americana era in piena recessione e inoltre l'impennata dei prezzi del petrolio e l'inflazione, che negli anni precedenti si attestava intorno al 10%, tenevano alti i tassi di interesse. Per poter uscire dalla recessione il programma economico di Reagan prevedeva di combinare un regime monetario restrittivo e una politica fiscale espansionistica. Per stimolare la ripresa fu effettuata una massiccia spesa da parte del governo seguita da un netto taglio delle tasse per favorire gli investimenti dei privati. Sebbene si registrò un aumento del reddito nazionale lordo (4,2% annuo dall'82 all'88) e un basso tasso di disoccupazione (da 10,8% nel

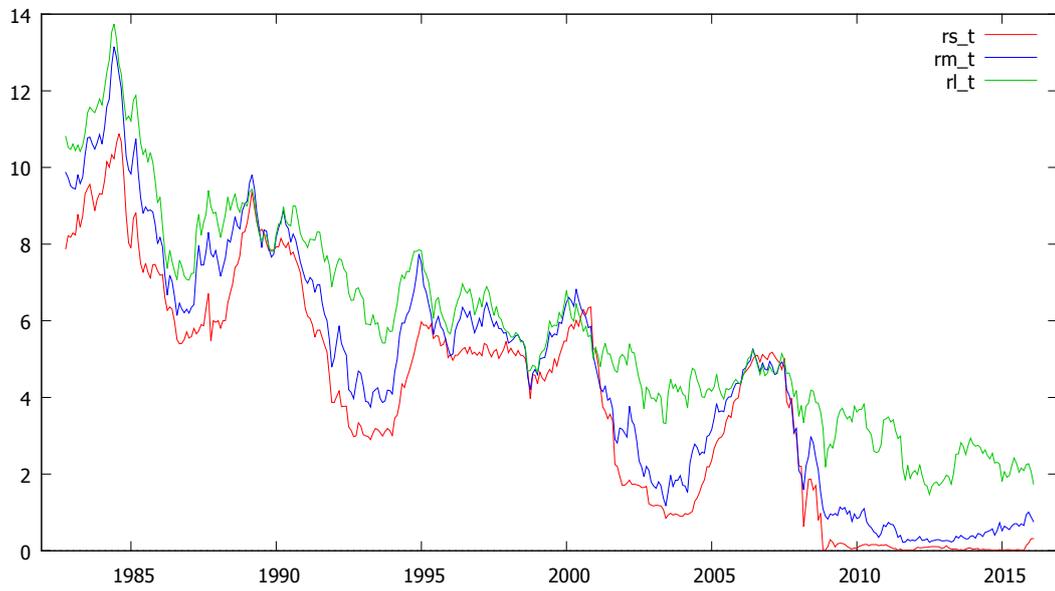


Figura 3.1: Treasury Bond

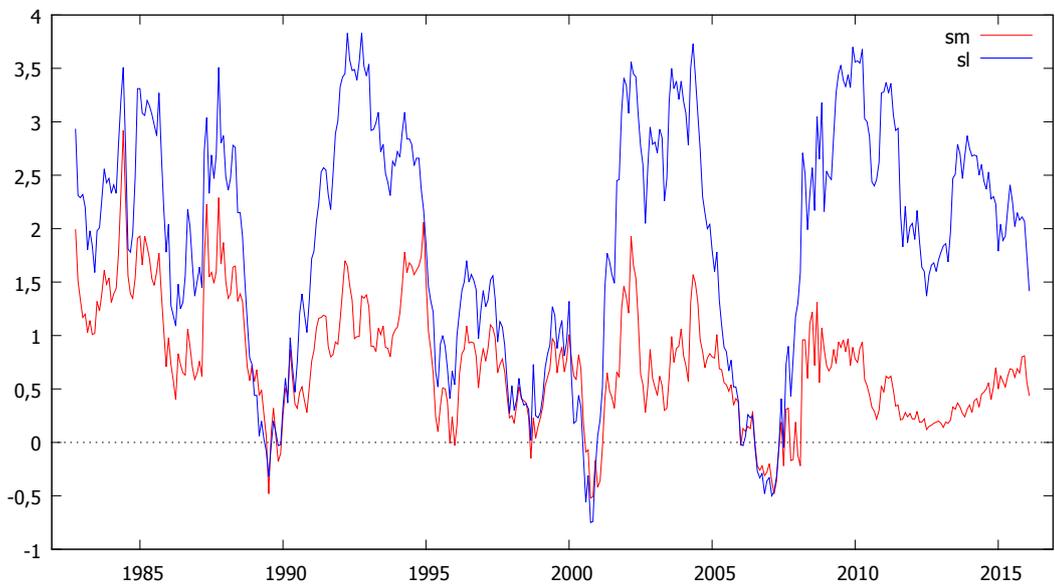


Figura 3.2: Spread

1982 a 5,4% nel 1988), il debito pubblico raddoppiò (passando da 749 a 1.746 di miliardi di dollari) (Magazzino, 2010). Gran parte di questo debito era coperto da prestiti provenienti dall'estero. Infatti, durante la legislazione di Reagan, il dollaro si apprezzò rispetto alle valute estere grazie anche ai suoi alti tassi di interesse (con picchi dell'11% nell'estate dell'1984). Quindi la domanda dei titoli americani era guidata principalmente dagli investitori esteri. La politica monetaria era mirata a mantenere ferma l'inflazione portandola al di sotto del 5%, quindi la stretta monetaria fu abbastanza ripida con l'aumento della spesa pubblica (Goodfriend, 2002). Solo intorno al 1985, cioè quando il prezzo di petrolio si stabilizzò su livelli molto più bassi rispetto al 1982, la politica monetaria poté far scendere i tassi di interesse, come si può anche notare nel grafico 3.1.

Il 19 Ottobre 1987 i prezzi delle azioni crollarono rapidamente. Il sistema finanziario sperimentò una fase recessiva in cui tutte le azioni persero gran parte del loro valore, in particolare l'indice Dow Jones perse più di 500 punti base e lo S&P più del 20% del suo valore (Carlson, 2006). Le autorità monetarie agirono tempestivamente sul sistema finanziario per prevenire conseguenze negative nell'economia reale. Infatti nella settimana successiva la Fed intervenne iniettando liquidità nel sistema attraverso operazioni di mercato aperto. Nelle settimane a venire la Fed allentò la presa monetaria abbassando i tassi di interesse. Questa manovra ebbe ripercussioni su tutti i tassi a breve termine, comportando una loro globale diminuzione su tutto il mercato. Nel triennio successivo l'economia statunitense si riprese nonostante la crescita industriale in media annua del 1%, l'inflazione si è attestata in tutto il periodo intorno al 5% e il tasso di disoccupazione scese del 2%. Come conseguenza la Fed decise di aumentare i tassi e portandoli dal 6% al 10% nel marzo del 1989 (Goodfriend, 2002).

Nell'Agosto del 1990 l'America entrò in guerra e di conseguenza il prezzo del greggio crebbe più del doppio. Si intuì che si sarebbe entrati in recessione durante tutto il periodo della guerra del Golfo, quindi fece scendere i tassi gradualmente. Le ripercussioni di questa guerra si propagarono fino al 1994. Durante questi ultimi anni il tasso di disoccupazione continuava a crescere mentre l'inflazione saliva in modo molto lieve, per cui la Fed decise di continuare a ridurre i *federal fund rate*

portandoli fino ad un minimo del 3% nel Febbraio del 1994 (Goodfriend, 2002).

Da Febbraio del '94 a febbraio del '95 l'economia statunitense era piuttosto stabile, ma la Fed dovette aumentare i tassi di interesse per far fronte alla crescente inflazione. A metà del 1995 l'America, la Germania e il Giappone firmarono il “*Reverse Plaza Accord*”. Questo patto prevedeva il salvataggio dell'economia giapponese dal continuo apprezzamento dello yen. L'obiettivo era di stabilizzare il cambio tra il dollaro e lo yen, quindi il Giappone doveva portare i tassi di interesse su livelli molto bassi rispetto a quelli americani. Per far ciò la banca centrale giapponese dovette acquistare i bond in valuta americana, mentre gli altri due paesi dovettero acquistare valuta americana (Frankel, 2015). Dalla Figura 3.1 è possibile vedere il picco intorno al 1995.

Dal 1996 fino al 1999 l'economia americana ebbe un lunghissimo periodo di *boom* governato da un tasso di produzione pari al 2% annuo (Goodfriend, 2002). In questo periodo la Fed non dovette compiere sostanziali interventi, per cui lasciò inalterati i tassi di interesse per quasi tutto l'arco temporale. Tutto questo fu possibile grazie alla credibilità che l'autorità monetaria godeva in quel periodo e alle aspettative degli operatori fortemente ancorate. L'unica eccezione in cui la Fed dovette intervenire fu l'estate del '98. La causa fu la crisi russa. Con la riduzione del prezzo del petrolio, la situazione fiscale della Russia peggiorò ulteriormente. Inoltre gli investitori spostarono i loro investimenti e quindi i tassi di interesse dei titoli russi si impennarono. Nell'estate del '98 la Russia fu costretta a passare da un sistema di cambio fisso a un sistema flessibile e bloccò i pagamenti dei propri debiti. Questi eventi ebbero importanti ripercussioni sia a livello internazionale che sulla politica economica statunitense. In primo luogo gli operatori decisero di riversare i loro investimenti nel mercato americano. Infatti, i rendimenti dei titoli americani con scadenze pluriennali caddero drasticamente. In secondo luogo il tasso di cambio si apprezzò dopo la crisi Russa, quindi la FOMC (*Federal Open Market Committee*) dovette abbassare i tassi di interesse di 75 punti base in vista anche delle turbolenze nei mercati finanziari.

Due eventi caratterizzarono l'inizio del XXI secolo: lo scoppio della bolla “dot.com” e l'attacco alle Torri Gemelle. Alla fine degli anni '90 le società appartenenti al setto-

re internet segnarono forti tassi di crescita sui corsi azionari, in media del 10% annuo, principalmente dovuto dall'euforia degli operatori. Lo scoppio della bolla avvenne all'inizio di Marzo del 2000, quando le società mostrarono bilanci molto deludenti. Tra il 2000 e 2002 si registrarono il fallimento di oltre la metà delle società nel settore informatico. La Fed cominciò ad agire solo agli inizi del 2001, cominciando a tagliare gradualmente i tassi di interesse portandoli dal 6,5% al 3,5% (Goodfriend, 2002). A Settembre del 2001 si registrò il crollo più intensivo a causa dell'attacco terroristico. Le conseguenze dell'accaduto si ripercossero immediatamente sul mercato causando un suo arresto totale per una settimana. Dopo l'apertura dei mercati americani l'indice Dow Jones perse 685 punti base. L'attacco aveva mostrato la fragilità dei mercati finanziari: l'interruzione dei servizi e dei pagamenti. Le conseguenze si riversarono anche sulle aspettative dei consumatori, l'atmosfera di quel periodo aveva ridotto sia i consumi che gli investimenti (Neely, 2004). La Fed dovette intervenire immediatamente aumentando la liquidità nel mercato attraverso l'acquisto di REPO, prestando direttamente alle banche attraverso "discount window" ed estensioni del "float". Inoltre durante i successivi tre mesi la Fed portò il tasso di interesse all'1,75%. Il mercato dei bond governativi subì un forte calo su tutti i titoli a diversa scadenza.

Negli anni successivi la Fed decise di mantenere i tassi a livelli molto bassi per stimolare l'economia americana. Questa decisione ha favorito in particolar modo l'andamento del mercato immobiliare. Infatti, i bassi tassi di interesse stimolarono la domanda degli immobili che fece salire il loro valore del 15% annuo. Dal 2003 si verificò un netto aumento dei mutui a persone con merito creditizio molto basso, denominati mutui *sub-prime*. Merito di questo aumento oltre alle dinamiche del mercato immobiliare fu lo sviluppo delle pratiche di cartolarizzazione, operazione che consentiva la vendita a terzi del prestito recuperando parte del credito. Con la pratica della cartolarizzazione e con l'assenza di regolamentazione sul grado di *leverage* per gli istituti bancari assisterono alla formazione di un'altra bolla speculativa. Questa bolla si manteneva in piedi principalmente dal fatto che il valore delle case era maggiore al valore del rimborso del prestito. Così, quando la Fed decise di aumentare i tassi di interesse, mentre il valore delle case cominciò a scendere, tutto

il sistema crollò. I primi segnali apparirono nella seconda metà del 2007 quando gli istituti finanziari che erogavano i mutui *sub-prime* cominciarono a registrare pesanti perdite dovute al mancato rimborso delle rate. Inoltre i titoli, ormai sparsi ovunque, contenenti questi mutui persero il loro valore. Le società dovettero affrontare prima una crisi di liquidità seguita da una crisi di fiducia. Il sistema finanziario registrò notevoli vendite di titoli seguite da riduzione di prestiti. L'economia reale subì la stretta creditizia, per cui si ridussero i consumi e ci fu un crollo della fiducia. I provvedimenti della Fed cominciarono nell'Agosto del 2007 con una riduzione del tasso di interesse dal 5,25% fino al 2% nell'aprile del 2008. Insieme ai soliti interventi furono istituiti programmi innovativi come il TAF (*Term Auction Facility*) per facilitare i prestiti alle banche e ridurre così lo spread nel mercato monetario (Cecchetti, 2008).

La crisi ebbe il suo apice nel Settembre del 2008 con il fallimento del colosso Lehman Brothers. Questo evento ebbe un impatto considerevole sui mercati, tanto che furono necessari ulteriori interventi da parte dei policy maker. Dato il basso margine di manovra sul tasso di interesse, portato dal 2% allo 0,25%, la FOMC decise di ricorrere a strumenti non convenzionali come il *Quantitative Easing* (QE). Questo programma prevedeva l'acquisto massiccio dei titoli del tesoro e dei titoli cartolarizzati sul mercato da parte della banca centrale. Nel Novembre del 2008 la Fed lanciò il primo programma di *Quantitative Easing* (QE1) con l'acquisto di 600 miliardi di dollari di titoli "tossici" seguita poi a Marzo con l'acquisto di titoli del tesoro (Guillén, 2015). Dato l'impatto poco incisivo sull'economia americana la FOMC decise quindi di intervenire nuovamente con un secondo programma (QE2). La FOMC annunciò a novembre del 2010 l'acquisto di 600 miliardi di dollari dei titoli del tesoro con un acquisto razionalizzato a 75 miliardi al mese (Labonte, 2016). Il programma terminò a Giugno del 2011. Il 21 Settembre del 2011 la FOMC comunicò l'implementazione del programma denominata *Operation Twist*. Questo piano era finalizzato a ridurre i rendimenti dei titoli a più lunga durata attraverso lo scambio di titoli del tesoro a breve termine in possesso della banca centrale con i titoli a più lunga scadenza. A Settembre del 2012 la Fed attuò il terzo *Quantitative Easing* volto all'inizio all'acquisto dei titoli MBS (*mortgage backed securities*) per 40 miliardi al mese, poi nel Dicembre del 2012 fino a Ottobre 2014 fu rivolto all'acquisto dei titoli del tesoro a lungo termine per 45 miliardi di dollari al mese (Labonte, 2016).

Per tutto il 2015 la FOMC ha tenuto i *federal funds rate* su livelli vicini allo zero.

### 3.3 Analisi preliminare

Per poter giustificare l'utilizzo del modello NEC sui dati americani è necessario verificare le implicazioni della teoria delle aspettative sul data set con la stessa metodologia di Campbell e Shiller (1987).

Data la natura del campione, è opportuno procedere con una metodologia robusta all'eteroschedasticità condizionale e non, come opportunamente descritto nel capitolo precedente. Liu (1988), Paparoditis e Politis (2005), Goncalves e Kilian (2004), ecc., dimostrano come la distribuzione asintotica nulla dei test non sia pivotale e dipenda dal processo della volatilità nel caso in cui quest'ultima non sia costante nel tempo. La procedura *wild bootstrap* consente di eliminare eventuali distorsioni dei test provocate dalla presenza di eteroschedasticità e/o dal campione.

Diversamente dal lavoro di Lucchetti e Palomba (2009) in questo paragrafo i test sono stati condotti sia nella loro versione originale che con la procedura bootstrap. Per quanto riguarda l'esecuzione di quest'ultima, in tutti i test ho seguito l'iter descritto da Liu (1988) e da Cavaliere *et al.* (2010b), paragrafo (2.4.2). Quindi tutte le simulazioni sono state ottenute replicando i residui ottenuti dal modello descritto dall'ipotesi nulla. I residui simulati sono stati ottenuti con la distribuzione normale standardizzata, la scelta è stata casuale poichè la distribuzione su cui eseguire il *wild bootstrap* è indifferente ai fini asintotici (Cavaliere *et al.*, 2010b).

Come già discusso nel primo capitolo, Campbell e Shiller (1987) elencano le diverse procedure per condurre i test. Una prima condizione è che i tassi di interesse possono variare nel tempo senza dover necessariamente tornare in media, mentre la relazione tra i rendimenti dei titoli con diversa scadenza dovrebbe essere stazionaria. L'analisi sulla presenza o meno di radici unitarie evidenzia che i rendimenti possono essere rappresentati da un processo integrato di ordine uno, mentre gli spread relativi ai rendimenti di breve periodo dovrebbero seguire un processo *mean reverting* o  $I(0)$ .

Per identificare l'ordine di integrazione ho utilizzato il test ADF (Dickey e Fuller, 1981), il test DF-GLS di Elliott *et al.* (1996), il test KPSS di Kwiatkowski *et al.* (1992) e il test PP di Phillips e Perron (1988). I test sono stati condotti prima considerando la sola presenza di una costante e poi anche la presenza di un trend. I risultati sono riportati nella Tabella 3.1.

Tabella 3.1: Test sulla stazionarietà dei rendimenti

Tassi di interesse con frequenza settimanale				
Test con costante				
Serie / Test	ADF	DF-GLS	KPSS	PP
$r_t^{3month}$	-1,04260	0,75871	9,79952***	-0,86380
$r_t^{2year}$	-1,72356	0,51752	10,6922***	-1,67561
$r_t^{10year}$	-1,72983	0,71961	11,4658***	-1,71908
$sm_t$	-4,69225***	-3,40464***	4,06891***	-5,85071***
$sl_t$	-2,87155**	-2,29501**	0,46054**	-3,12512**
Test con costante e trend				
$r_t^{3month}$	-2,38247	-2,14161	0,236138***	-2,03628
$r_t^{2year}$	-3,45805**	-2,20606	0,26438***	-2,77614
$r_t^{10year}$	-3,88486**	-2,35934	0,61309***	-3,65962**
$sm_t$	-5,51463***	-0,02260	0,18927**	-6,34051***
$sl_t$	-3,10290	-3,09486**	0,45595***	-3,13728*
Tassi di interesse con frequenza mensile				
Test con costante				
Serie / Test	ADF	DF-GLS	KPSS	PP
$r_t^{3month}$	-1,03186	0,40801	4,9510***	-1,05688
$r_t^{2year}$	-1,80118	0,51328	5,4092***	-1,75767
$r_t^{10year}$	-1,76128	0,75508	5,8093***	-1,77313
$sm_t$	-6,43675***	-0,9815	2,2280***	-6,64808***
$sl_t$	-3,95936***	-1,1559	0,23770	-3,52652***
Test con costante e trend				
$r_t^{3month}$	-4,37025***	-0,01519	0,1199*	-2,50990
$r_t^{2year}$	-3,57627**	-0,01330	0,1344*	-3,12288
$r_t^{10year}$	-4,00435***	-1,87152	0,3157***	-3,94552**
$sm_t$	-6,99929***	-1,90622	0,10158	-7,40511***
$sl_t$	-3,96402***	-1,94959	0,23535***	-3,52956**

Il numero degli asterischi posti di fianco ai numeri indicano rispettivamente il rifiuto dell'ipotesi nulla al 10%, 5% e 1%

Il test ADF con costante conferma per tutti i tassi di interesse, settimanali e mensili, la presenza di una radice unitaria mentre, se si considera la presenza di un trend temporale, le serie possono essere considerate processi I(0). I test DF-GLS e PP in entrambi i casi confermano che le serie dei rendimenti dei bond siano processi

integrati di ordine uno, l'unica eccezione è per i tassi di interesse a 10 anni dove il test PP accetta l'ipotesi che la serie sia stazionaria attorno a un trend temporale. Il test KPSS conferma i risultati ottenuti con gli altri metodi rifiutando l'ipotesi di stazionarietà in tutti i casi. Per quanto riguarda gli spread tra i rendimenti i risultati sono discordi tra i vari test. Il test KPSS rifiuta ampiamente l'ipotesi che gli spread siano stazionari tranne per il caso della serie  $sl_t$  mensile e  $sm_t$  nel caso in cui si consideri la presenza di un trend temporale. Caso opposto per i test ADF e i test PP che invece rifiutano l'ipotesi che le serie degli spread siano non stazionari, tranne per il caso dello spread di lungo periodo se si considera la presenza di un trend temporale. Per quanto riguarda il test DF-GLS la situazione è controversa fra le diverse frequenze delle serie degli spread. Infatti il test accetta l'ipotesi di stazionarietà per i dati a frequenza settimanale mentre quelli mensili accetta l'ipotesi che siano processi I(1).

Tabella 3.2: Test sulla stazionarietà dei rendimenti con wild bootstrap

<b>Tassi di interesse con frequenza settimanale</b>		
<b>Test con costante</b>		
Serie / Test ADF	p-value	bootstrap p-value
$r_t^{3month}$	0,7401	0,8361
$r_t^{2year}$	0,4193	0,7164
$r_t^{10year}$	0,4162	0,4314
$sm_t$	0,0000	0,2541
$sl_t$	0,0268	0,3444

<b>Tassi di interesse con frequenza mensile</b>		
<b>Test con costante</b>		
Serie / Test ADF	p-value	bootstrap p-value
$r_t^{3month}$	0,7441	0,8197
$r_t^{2year}$	0,3959	0,3579
$r_t^{10year}$	0,3988	0,3061
$sm_t$	0,0000	0,1334
$sl_t$	0,0016	0,4153

Bootstrap effettuato con 999 replicazioni e utilizzando come distribuzione la Gaussiana.

Come descritto del paragrafo (2.4.2) utilizzare la procedura bootstrap nei test per la stazionarietà ci consente di ovviare a problemi di distorsioni provocate dal campione (Paparoditis e Politis, 2005; Parker *et al.*, 2006; Phillips, 2010) e dalla presenza di eteroschedasticità (Goncalves e Kilian, 2004; Kim *et al.*, 2002a; Busetti e Taylor, 2003; Cavaliere, 2005; Cavaliere e Taylor, 2008; Cavaliere *et al.*, 2012).

Nella Tabella 3.2 sono stati riportati i risultati ottenuti con il *wild bootstrap* ADF (Cavaliere e Taylor, 2008). Per tutti i tassi di interesse la procedura bootstrap replica gli stessi risultati ottenuti precedentemente cioè la presenza di radici unitarie nei livelli delle serie. Caso opposto per quanto riguarda gli spread. Infatti i test bootstrap confermano l'ipotesi di non stazionarietà degli spread.

Al fine di valutare la validità delle implicazioni “low frequency” della teoria della struttura a termine ho eseguito fra i test preliminari anche quello sul rango di cointegrazione. Questa condizione di lungo termine, secondo Campbell e Shiller (1987), impone che i tassi di rendimento siano guidati da un solo trend stocastico. Nel caso preso in esame, tre tassi a diversa scadenza, il rango della matrice di cointegrazione dovrebbe essere pari a  $n - 1 = 2$ . Per verificare questa ipotesi mi sono avvalso del test di Johansen (1995).

Diversamente da lavoro di Lucchetti e Palomba (2009), ho eseguito anche il test *wild bootstrap* sul rango di cointegrazione per ottenere risultati robusti alla volatilità non stazionaria ed evitare problemi di *size distortion* (paragrafo 2.4.2).

Per quanto riguarda il nucleo deterministico dell'equazione (1.5) la scelta non appare semplice. In primo luogo c'è da sottolineare che, secondo la teoria del premio di liquidità, bisogna imporre la presenza di un termine costante all'interno della relazione di lungo periodo. Infatti le serie degli spread presentano valore atteso diverso da zero facilmente visibile nella Figura 3.2. In secondo luogo c'è da considerare il periodo temporale preso in esame. I tassi di interesse mostrano un andamento decrescente in questi ultimi 34 anni, per cui non è possibile escludere la possibilità che le serie seguono un trend lineare. Per cui la scelta può ricadere sia nel caso in cui la costante sia vincolata che non. Dato che nessuna delle due ipotesi può essere scartata, ho condotto il test su tutti e due i casi. Per quanto concerne l'ordine dei ritardi la scelta è stata effettuata considerando i risultati forniti dai criteri informativi di

Schwarz (1978) (BIC) e Hannan e Quinn (1979) (HQC). Quando i dati presentano eteroschedasticità i test sul rango di cointegrazione risultano consistenti se basati su questi due criteri informativi (Kapetanios, 2000), mentre con l’AIC (Kapetanios, 2004) ciò non accade.

I risultati riportati nella Tabella 3.3 sono discordanti con la teoria. La condizione di “low frequency” della teoria delle aspettative non viene rispettata sebbene i test indichino la presenza di cointegrazione. Infatti in tutti i casi e su entrambi i data set il test di Johansen (1995) e il *wild bootstrap* rifiutano l’ipotesi secondo cui il rango di cointegrazione sia pari a  $n - 1 = 2$ . Alla luce di ciò, i risultati sembrano confermare le ipotesi proposte da Giese (2008) e da Diebold e Li (2003), secondo cui una combinazione degli spread risulta preferibile rispetto all’utilizzo dei singoli spread.

Affinché la teoria delle aspettative possa essere confermata empiricamente tutte le ipotesi devono essere verificate. Un’altra condizione necessaria, che riguarda il meccanismo di aggiustamento, viene imposta tramite la *cross-equation constraints*. Quest’ultima rappresenta la restrizione “high frequency” e viene verificata tramite l’equazione (1.14). La procedura che ho adottato segue quella di Campbell e Shiller (1987), Gonzalo e González (2000) e Gunnella (2015) quindi, le implicazioni sono state verificate sia singolarmente, con un modello VAR bivariato (singolo spread e  $\Delta r_t^{3month}$ ) sia congiuntamente con il modello multivariato (entrambi gli spread e il tasso di interesse). Le statistiche test sono state condotte sia con lo stimatore HAC che con la procedura bootstrap.

Come si può notare dalla tabella (3.4), in tutti i casi le *cross-equation restriction* vengono rifiutate ad un livello minimo di significatività pari al 10%. Nei VAR bivariati le implicazioni di breve periodo della teoria delle aspettative sono fortemente rifiutate per il test Wald corretto per l’eteroschedasticità, eccezione fatta per  $sm_t$  mensili, mentre con il *wild bootstrap* vengono rifiutate, ma ad un livello più basso. Nella verifica congiunta e in entrambi i test le restrizioni sono fortemente rifiutate.

Tabella 3.3: Test Johansen

<b>Test sul rango matrice cointegrazione settimanale</b>						
Test con costante vincolata e ordine dei ritardi pari a 2						
Periodo di stima: 12/10/1982 - 24/02/2015 (T = 1690)						
Rango	Autovalore	Traccia	<i>p-value</i>	<i>bootstrap p-value</i>	$\lambda$ -max	<i>p-value</i>
0	0,052617	104,73	[0,0000]	[0,0000]	91,347	[0,0000]
1	0,0053587	13,387	[0,3414]	[0,5391]	9,0805	[0,4375]
2	0,0025451	4,3067	[0,3803]	[0,5571]	4,3067	[0,3795]
<b>Test sul rango matrice cointegrazione mensile</b>						
Test con costante vincolata e ordine dei ritardi pari a 5						
Periodo di stima: 10/1982 - 02/2015 (T = 389)						
Rango	Autovalore	Traccia	<i>p-value</i>	<i>bootstrap p-value</i>	$\lambda$ -max	<i>p-value</i>
0	0,21430	105,41	[0,0000]	[0,0000]	93,821	[0,0000]
1	0,019365	11,584	[0,4952]	[0,5125]	7,6067	[0,6036]
2	0,010172	3,9772	[0,4276]	[0,6486]	3,9772	[0,4267]
Test con costante non vincolata e ordine dei ritardi pari a 5						
Rango	Autovalore	Traccia	<i>p-value</i>	<i>bootstrap p-value</i>	$\lambda$ -max	<i>p-value</i>
0	0,12484	64,543	[0,0000]	[0,0010]	51,872	[0,0000]
1	0,026508	12,672	[0,1280]	[0,4034]	10,451	[0,1871]
2	0,0056928	2,2208	[0,1362]	[0,5305]	2,2208	[0,1362]

Bootstrap effettuato con 999 replicazioni e utilizzando come distribuzione la Gaussiana.

Tabella 3.4: Test sulle *cross-equation restrictions*

<b>Tassi di interesse con frequenza settimanale</b>			
<b>Cross-equation test</b>			
Serie / Test	Wald test (HAC)	p-value	bootstrap p-value
$sm_t, \Delta r_t^3$	34,1638	0,0000	0,0000
$sl_t, \Delta r_t^3$	15,2167	0,0042	0,0000
$sm_t, sl_t, \Delta r_t^3$	173,4914	0,0000	0,0000
<b>Cross-equation test</b>			
<b>Test</b>			
Serie / Test	Wald test (HAC)	p-value	bootstrap p-value
$sm_t, \Delta r_t^3$	4,7379	0,0935	0,0080
$sl_t, \Delta r_t^3$	9,8928	0,0071	0,0000
$sm_t, sl_t, \Delta r_t^3$	201,6544	0,0000	0,0000

Bootstrap effettuato con 999 replicazioni e utilizzando come distribuzione la Gaussiana.

Seppur con un orizzonte temporale differente i risultati, finora esposti, sui tassi di interesse americani con frequenza mensile ricalcano alcune delle più recenti ricerche, sullo stesso tema, presenti in letteratura (Cavaliere *et al.*, 2010b), (Boswijk *et al.*, 2016) e Cavaliere *et al.* (2016).

In questa prima analisi è possibile evidenziare i contrasti con la teoria.

1. Non tutti i test sembrano confermare l'ipotesi che gli spread siano stazionari, in particolar modo per la differenza tra i tassi di rendimento a 2 anni e a 3 mesi per cui la maggior parte dei test sono a favore della non stazionarietà.
2. Il test di Johansen (1995) confuta l'ipotesi che le serie seguano un unico trend.
3. Le *cross-equation constrains* non sono verificate.
4. Anche effettuando i test con la procedura bootstrap i risultati non confermano la teoria delle aspettative. I test ADF *wild bootstrap* non rifiutano l'ipotesi di non stazionarietà per gli spread, i test sul rango di cointegrazione sono a favore dell'ipotesi che i tassi di interesse siano governati dalla presenza di due trend stocastici e, infine, le implicazioni "high frequency" sono rifiutate.

In conclusione proseguire con lo stesso l'approccio di Campbell e Shiller (1987) non sembra giustificato da questa prima analisi, in quanto le implicazioni sia di lungo termine che di breve non vengono rispettate. Sembra chiaro che la normale metodologia di cointegrazione non ci permetta di ottenere risultati congruenti con la teoria per cui è necessario passare ad una modellazione non standard come quella del NEC.

## 3.4 Modelli parametrici

### 3.4.1 Stima di modelli univariati

In questa prima parte analizzerò il meccanismo di aggiustamento non lineare inserito all'interno di un modello univariato come descritto nel paragrafo 1.5.4.

L'equazione in forma generale si presenta in questo modo:

$$\Delta r_t^i = d_t + \Gamma_i(L)' \Delta R_{t-1} + \alpha_i' S_{t-1} + \theta_i' Q_{t-1} + \lambda_i' C_{t-1} + \varepsilon_t \quad (3.1)$$

dove  $r_i$  con  $i = s, m, l$  indica rispettivamente i rendimenti di breve, di medio o di lungo periodo,  $R_t$  è un vettore contenente i tre tassi di interesse considerati, infine i vettori  $S_t$ ,  $Q_t$ ,  $C_t$  sono i vettori che contengono i termini della serie di Taylor definiti dall'equazione (1.45). Infine  $d_t$  è la componente deterministica che include una costante e 3 variabili dummy per il “*Black Monday*” del 19 ottobre del 1987, l'attacco alle Torri Gemelle nell'11 Settembre 2001 e l'annuncio ufficiale del fallimento della Lehman Brothers avvenuto il 15 Settembre 2008.

Prima di passare alla stima dell'equazione (3.1), ho differenziato le serie dei tassi dei rendimenti dei singoli tassi di interesse. In secondo luogo sono andato a calcolare l'ordine dei ritardi di  $\Gamma_i(L)$ . La scelta è avvenuta comparando i criteri informativi BIC e HBQ di ogni singolo modello. Infine ho stimato l'equazione (3.1) per ognuno dei tre tassi sui dati a frequenza settimanale e mensile con il metodo OLS

e con l'utilizzo dello stimatore HAC (Newey e West, 1987) per ottenere varianze e covarianze consistenti anche nel caso in cui vi sia la presenza di eteroschedasticità e autocorrelazione.

Le stime dei coefficienti del meccanismo di aggiustamento sono riportate nelle Tabelle 3.5 e 3.7 mentre i risultati completi delle stime sono riportati all'interno della sezione degli Allegati.

Tabella 3.5: Coefficienti del meccanismo di aggiustamento univariato settimanale

Var.	Var. Dip. $\Delta r s_t$			Var. Dip. $\Delta r m_t$			Var. Dip. $\Delta r l_t$		
	Coeff.	S.E.	<i>t</i> -stat.	Coeff.	S.E.	<i>t</i> -stat.	Coeff.	S.E.	<i>t</i> -stat.
$sm_t$	0,139	0,033	4,198	0,017	0,035	0,498	0,002	0,026	0,095
$sl_t$	-0,062	0,024	-2,610	0,006	0,025	0,225	-0,001	0,021	-0,061
$sm_t^2$	-0,164	0,059	-2,776	-0,130	0,057	-2,279	-0,067	0,045	-1,505
$sl_t^2$	0,020	0,017	1,159	-0,020	0,019	-1,057	-0,003	0,018	-0,170
$sm_t sl_t$	0,044	0,046	0,960	0,106	0,048	2,231	0,062	0,043	1,434
$sm_t^3$	0,009	0,022	0,430	-0,060	0,025	-2,357	-0,040	0,023	-1,792
$sl_t^3$	-0,001	0,004	-0,022	0,009	0,005	1,865	0,003	0,005	0,580
$sm_t^2 sl_t$	0,043	0,039	1,112	0,114	0,046	2,452	0,067	0,041	1,640
$sm_t sl_t^2$	-0,026	0,020	-1,257	-0,063	0,024	-2,621	-0,036	0,022	-1,657

Campione: 12/10/1982 - 24/02/2015 (T = 1690). Errori standard HAC (larghezza della banda 5)

Da questa prima tabella emerge che:

- per quanto riguarda le variazioni di rendimento di breve periodo le componenti che maggiormente ne influenzano l'andamento sono gli spread  $sm_t$  e  $sl_t$  e l'unica componente statisticamente significativa è lo spread tra i rendimenti di medio e di breve periodo al quadrato  $sm_t^2$ .
- Caso opposto per il tasso di rendimento a 2 anni: tutti i coefficienti relativi al meccanismo di aggiustamento non lineare risultano significativi almeno al 10%, tranne per  $sl_t^2$ , mentre i coefficienti relativi agli spread assumono valori molto vicini allo zero e statisticamente non sono significativi.
- Le variazioni di rendimento dei bond decennali non sembrano influenzate dagli spread dei rendimenti. Per quanto riguarda i coefficienti riferiti alla componente non lineare  $sm_t^2, \dots, sm_t sl_t^2$ , solo due risultano statisticamente significativi al 10%; c'è però da sottolineare che per altri tre coefficienti il test di significatività risulta molto vicino al 10%.

Tabella 3.6: Test Diagnostici sui dati settimanali

Variabile	$\Delta r_{st}$	$\Delta rm_t$	$\Delta rl_t$
<b>Analisi corretta specificazione</b>			
<i>Wald test per i coefficienti di <math>sm_t^2, sl_t^2, \dots, sm_t sl_t^2</math></i>			
	$F(7, 1672) = 4,1066$ p-value = 0,0002	$F(7, 1674) = 5,3112$ p-value = 0,000	$F(7, 1676) = 12,3045$ p-value = 0,000
<i>Test RESET per i quadrati e i cubi</i>			
	$F(2, 1670) = 2,8786$ p-value = 0,0565	$F(2, 1672) = 1,1233$ p-value = 0,3254	$F(2, 1674) = 0,9260$ p-value = 0,3960
<b>Analisi Residui</b>			
<i>Test di Breusch-Godfrey per l'autocorrelazione del prim'ordine</i>			
	LMF = 20,9691 p-value = 0,000	LMF = 4,2852 p-value = 0,0386	LMF = 3,7199 p-value = 0,0539
<i>Test di Ljung-Box per l'autocorrelazione del prim'ordine</i>			
	Q' = 1,9440 p-value = 0,163	Q' = 0,4441 p-value = 0,505	Q' = 0,2755 p-value = 0,600
<i>Test ARCH(2)</i>			
	LM = 42,2524 p-value = 0,000	LM = 18,9164 p-value = 0,000	LM = 56,1922 p-value = 0,000
<i>Test CUSUM per la stabilità dei parametri</i>			
	$t(1671) = 3,3236$ p-value = 0,0009	$t(1673) = -0,4528$ p-value = 0,6507	$t(1675) = -0,1056$ p-value = 0,9159

I test di azzeramento congiunto (Tabella 3.6) dei soli coefficienti riferiti all'aggiustamento non lineare rifiuta l'ipotesi nulla per tutti e tre i tassi di rendimento. In tutti e tre i test il modello libero è preferibile statisticamente a quello vincolato. Il test di Wald conferma ulteriormente che la presenza di un meccanismo di aggiustamento non lineare aiuta meglio a comprendere la dinamica dei rendimenti dei bond governativi.

L'analisi dei test diagnostici sui dati a frequenza settimanale sono riportati nella Tabella 3.6.

Il test RESET (*Regression Specification Error Test*) condotto con i quadrati e i cubi mostra che solo l'equazione relativa al tasso di interesse di breve periodo presenta una scorretta specificazione.

Per quanto riguarda la presenza di autocorrelazione di prim'ordine nei residui, i test che ho condotto non forniscono un risultato univoco. Il test di Breusch-Godfrey segnala la presenza di autocorrelazione, mentre il test di Ljung-Box accetta l'ipotesi che non vi sia autocorrelazione tra i residui di primo ordine. La scelta dell'ordine dei ritardi è stata effettuata considerando i risultati ottenuti dai criteri informativi BIC e HQC.

Il test ARCH condotto su 2 ritardi mostra chiaramente che non si può rifiutare la presenza di eteroschedasticità condizionale dei residui, anche qui la scelta dei ritardi è stata dettata dai criteri informativi. Data la natura finanziaria dei dati e la loro frequenza è chiaro che i modelli necessitano di una specificazione per la dinamica della varianza condizionale.

Infine ho effettuato il test di CUSUM per verificare la presenza di cambiamenti strutturali nei parametri. Naturalmente solo nel caso dei rendimenti di breve periodo viene rifiutata la nulla. Una spiegazione di questo risultato potrebbe provenire dal fatto che i rendimenti di breve periodo sono più sensibili all'impatto di nuove notizie sul mercato.

I coefficienti relativi dall'applicazione del modello univariato ai dati mensili sono riportati nella Tabella 3.7. Da una prima analisi su questa tabella è possibile affermare che

Tabella 3.7: Coefficienti del meccanismo di aggiustamento univariato mensile

Var.	Var. Dip. $\Delta rs_t$			Var. Dip. $\Delta rm_t$			Var. Dip. $\Delta rl_t$		
	Coeff.	S.E.	t-stat	Coeff.	S.E.	t-stat.	Coeff.	S.E.	t-stat.
$sm_t$	0,527	0,181	2,906	-0,037	0,214	-0,175	-0,093	0,157	-0,589
$sl_t$	-0,233	0,125	-1,865	0,054	0,137	0,393	0,048	0,109	0,436
$sm_t^2$	-0,436	0,244	-1,790	-0,431	0,175	-2,469	-0,345	0,202	-1,710
$sl_t^2$	0,088	0,086	1,025	-0,090	0,109	-0,828	-0,040	0,097	-0,413
$sm_t sl_t$	0,090	0,242	0,373	0,431	0,219	1,966	0,307	0,223	1,376
$sm_t^3$	-0,074	0,108	-0,685	-0,217	0,101	-2,151	-0,143	0,104	-1,378
$sl_t^3$	0,002	0,021	0,090	0,036	0,025	1,442	0,013	0,025	0,497
$sm_t^2 sl_t$	0,242	0,163	1,489	0,408	0,169	2,411	0,274	0,189	1,449
$sm_t sl_t^2$	-0,120	0,086	-1,389	-0,238	0,091	-2,628	-0,147	0,106	-1,388

Campione: 10/1982 - 02/2015 (T = 389). Errori standard HAC (larghezza della banda 5)

Tabella 3.8: Test Diagnostici sui dati mensili

Variabile	$\Delta rs_t$	$\Delta rm_t$	$\Delta rl_t$
<b>Analisi corretta specificazione</b>			
<i>Wald test per i coefficienti di <math>sm_t^2, sl_t^2, \dots, sm_t sl_t^2</math></i>			
	$F(7, 373) = 1,8570$	$F(7, 373) = 13,5944$	$F(7, 373) = 8,8802$
	p-value = 0,0755	p-value = 0,000	p-value = 0,000
<i>Test RESET per i quadrati e i cubi</i>			
	$F(2, 371) = 5,7524$	$F(2, 371) = 0,6948$	$F(2, 371) = 0,1809$
	p-value = 0,0035	p-value = 0,4998	p-value = 0,8346
<b>Analisi Residui</b>			
<i>Test di Breusch-Godfrey per l'autocorrelazione del prim'ordine</i>			
	LMF = 0,1205	LMF = 0,02805	LMF = 0,1423
	p-value = 0,729	p-value = 0,867	p-value = 0,706
<i>Test di Ljung-Box per l'autocorrelazione del prim'ordine</i>			
	Q' = 0,0158	Q' = 0,0043	Q' = 0,0088
	p-value = 0,90	p-value = 0,948	p-value = 0,925
<i>Test ARCH(2)</i>			
	LM = 24,1245	LM = 4,7221	LM = 1,9158
	p-value = 0,000	p-value = 0,0943	p-value = 0,3837
<i>Test CUSUM per la stabilità dei parametri</i>			
	$t(372) = 1,6023$	$t(372) = -1,4601$	$t(372) = -0,8366$
	p-value = 0,1099	p-value = 0,1451	p-value = 0,4033

- le variazioni di rendimento di breve periodo sono principalmente determinate dagli spread di breve e medio periodo come nel caso dei dati settimanali. L'unica componente che statisticamente è significativa al 10% risulta essere lo spread tra i rendimenti di medio e breve periodo al quadrato  $sm_t^2$ .
- Nel caso delle variazioni dei rendimenti di medio termine la situazione è diversa: 5 coefficienti su 7 della componente non lineare del meccanismo di aggiustamento risultano significativi almeno al 10%, mentre i coefficienti relativi agli spread sono statisticamente non significativi.
- Le variazioni di rendimento dei bond decennali non sono influenzati dagli spread lineari dei rendimenti. Solo il coefficiente relativo alla variabile  $sm^2$  risulta statisticamente significativo al 10%, mentre gli altri test portano ad accettare l'ipotesi di azzeramento sebbene con un p-value molto vicino al valore soglia.

Nella Tabella 3.8 sono riportati tutti i test diagnostici. Il test di Wald sull'azzeramento congiunto di tutti i parametri relativi all'aggiustamento non lineare ( $sm_t^2$ ,  $sl_t^2$ , ... ,  $sm_tsl_t^2$ ) rifiuta la nulla in tutte e tre le equazioni, sintomo che questi parametri contribuiscono a specificare la media condizionata.

Il test RESET evidenzia che solo il modello riferito ai rendimenti dei bond di breve periodo necessita di un più corretta specificazione, mentre le altre 2 equazioni non ne hanno bisogno.

Nei test per la verifica della presenza di autocorrelazione, in tutti e due i test viene accettata l'ipotesi che i residui non siano correlati con il loro passato, quindi tutti e tre i modelli risultano adeguatamente specificati.

Il test ARCH indica che nel modello riferito ai rendimenti di breve periodo siamo chiaramente di fronte ad un modello con eteroschedasticità condizionale, contrariamente a quello di lungo periodo. Per quanto riguarda il modello riferito ai tassi di medio periodo, l'ipotesi che vi sia eteroschedasticità condizionale viene rifiutata ad un livello molto vicino a quello della soglia del 10%, per cui possiamo non rifiutare l'ipotesi che il processo sia omoschedastico.

Per il test CUSUM sulla stabilità dei parametri in tutti e tre i casi l'ipotesi nulla viene accettata.

Paragonando i risultati per tutti e due i data set possiamo notare che i risultati sono simili e sono a favore del modello non lineare. Infatti il test di Wald, che può essere visto come un test tra il modello lineare ECM e quello non lineare, rivela che quest'ultimo meccanismo di aggiustamento opera significativamente per entrambi i data set. I test diagnostici non mostrano segni di misspecificazione del modello, a parte il fatto che i dati a frequenza settimanale necessitano di una specificazione anche per la dinamica della varianza condizionata. Inoltre c'è da sottolineare che, mentre gli spread risultano altamente significativi per la dinamica dei rendimenti a 3 mesi lo stesso non è valido per i rendimenti di lungo periodo, che però vengono meglio descritti dalla componente non lineare.

### 3.4.2 Analisi *out-of-sample* univariata

La verifica con i test diagnostici di possibili misspecificazioni ci permette di validare in parte il modello. Questi test sono stati utilizzati non per valutare il modello come strumento di previsione, quanto per verificare la sua robustezza. Spesso i modelli non lineari forniscono buone approssimazioni all'interno del campione, ma hanno scarsa capacità previsiva. Infatti un fenomeno che potrebbe riscontrarsi è quello dell'“*overfitting*”. Con questo termine si indica la situazione in cui un modello tenta di descrivere la componente idiosincratICA o casuale anziché soffermarsi sulla sola relazione sistematica sottostante ai dati, come dovrebbe fare.

Per valutare l'abilità previsiva ho utilizzato lo stesso approccio di Lucchetti e Palomba (2009), cioè il RMSE (*Root Mean Square Error*) e il test di Diebold e Mariano (1995). Il test di Diebold e Mariano (1995) è stato condotto con una funzione di perdita simmetrica, la *V-shaped*, e una asimmetrica, la *Linex*. Per la varianza di lungo periodo è stata utilizzata una *Bartlett window* di ampiezza pari a 3 per tutte e due le tipologie di funzioni. Il campione *out-of-sample* è composto da un anno di osservazioni, quindi 52 osservazioni per i dati settimanali e 12 per quelli mensili. Il modello non lineare (NEC) è stato confrontato con i migliori *competitor* lineari:

- Random Walk (RW): ottenuto da una differenziazione delle serie, la logica dietro questo modello è che la dinamica dei tassi di interesse sia governata solo dalla componente casuale.
- Autoregressivo sulle differenze (DAR): differenziando le serie e utilizzando come regressori esogeni i ritardi degli altri tassi di rendimento. Il modello presuppone che non vi sia alcun attrattore di lungo periodo.
- Modello di correzione dell'errore con rango di cointegrazione pari a 1 (ECM1): il rango del vettore di cointegrazione è quello dettato dal test di Johansen.
- Modello di correzione dell'errore con rango di cointegrazione pari a 2 (ECM2): replicazione del modello di Campbell e Shiller (1987).
- Autoregressivo sui livelli (LAR): autoregressivo stimato senza differenziare le serie.

La tabella 3.9 mostra i risultati ottenuti applicando il test DM e il RMSE sui dati settimanali. È possibile notare che le previsioni per tutti i modelli sono pressoché simili. I valori del RMSE sono molto simili tra il NEC e i modelli lineari, leggermente più bassi nei tassi a breve e medio termine, e il test di Diebold e Mariano (1995) non rifiuta l'ipotesi nulla che i modelli abbiano la stessa capacità previsiva.

Anche per quanto riguarda il test condotto sui dati mensili i risultati sono molto simili. Come si può osservare nella Tabella 3.11 non c'è una sostanziale differenza tra le previsioni fatte con il modello non lineare e quelli lineari. In conclusione è possibile sostenere che il modello univariato non lineare non mostra segni di mispecificazione o di “*overfitting*”.

### 3.4.3 Stima modello multivariato

Analogamente all'equazione (1.5), il modello multivariato da stimare ha equazione:

$$\Phi(L)\Delta R_t = \gamma' d_t + \alpha' S_{t-1} + \theta' Q_{t-1} + \lambda' C_{t-1} + \varepsilon_t \quad (3.2)$$

Tabella 3.9: RMSE su dati settimanali

Modello / Serie	$rs_t^{3mesi}$	$rs_t^{2anni}$	$rl_t^{10anni}$
NEC	0,0780*	0,1232	0,2450
RW	0,2391	0,3074	0,2685
DAR	0,1890	0,2070	0,1876
ECM1	0,1263	0,3558	0,3299
ECM2	0,1272	0,3577	0,2783
LAR	0,2520	0,1197*	0,1501*

I valori contrassegnati con l'asterisco indicano i valori RMSE più bassi.

Tabella 3.10: Diebold-Mariano sui dati settimanali

Funzione V-shape			
Modello / Serie	$rs_t^{3mesi}$	$rs_t^{2anni}$	$rl_t^{10anni}$
RW	-0,0697	-0,1110	-0,0143
DAR	-0,0567	-0,0786	0,0125
ECM1	-0,1960	-0,1356	-0,0452
ECM2	-0,3350	-0,1364	-0,0190
LAR	-0,5713	0,0021	0,1000
Funzione Linex			
Modello / Serie	$rs_t^{3mesi}$	$rs_t^{2anni}$	$rl_t^{10anni}$
RW	-0,1643	-0,0881	0,0007
DAR	-0,1632	-0,0832	0,0003
ECM1	-0,1403	-0,1006	-0,0072
ECM2	-0,2912	-0,1011	0,0000
LAR	-0,3553	0,0094	0,0094

Test a due code: i valori negativi del test indicano una miglior accuratezza previsiva per il modello non lineare viceversa per i valori positivi. Il numero degli asterischi posti di fianco ai numeri indicano rispettivamente il rifiuto dell'ipotesi nulla al 10%, 5% e 1%

Tabella 3.11: RMSE su dati mensili

Modello / Serie	$rs_t^{3mesi}$	$rs_t^{2anni}$	$rl_t^{10anni}$
NEC	0,2218	0,1294*	0,2914
RW	0,2218	0,3425	0,2927
DAR	0,1255*	0,1957	0,8487
ECM1	0,1680	0,3529	0,3594
ECM2	0,3513	0,9347	1,0740
LAR	0,2609	0,1793	0,2430*

I valori contrassegnati con l'asterisco indicano i valori RMSE più bassi.

Tabella 3.12: Diebold-Mariano sui dati mensili

Funzione V-shape			
Modello / Serie	$rs_t^{3mesi}$	$rs_t^{2anni}$	$rl_t^{10anni}$
RW	0,2044	-0,3733	-0,0530
DAR	0,7136	-0,0784	0,0349
ECM1	0,0154	-0,4010	-0,1703
ECM2	0,0484	-0,4126	-0,0812
LAR	-0,0826	0,0250	0,0101
Funzione Linex			
Modello / Serie	$rs_t^{3mesi}$	$rs_t^{2anni}$	$rl_t^{10anni}$
RW	0,0235	-0,3282	0,0512
DAR	1,2207	-0,1150	0,3151
ECM1	0,0492	-0,3413	-0,0038
ECM2	0,0494	-0,3459	0,0209
LAR	-0,1228	-0,0586	0,2044

Test a due code: valori negativi del test indicano una miglior accuratezza previsiva per il modello non lineare viceversa per i valori positivi. Il numero degli asterischi posti di fianco ai numeri indicano rispettivamente il rifiuto dell'ipotesi nulla al 10%, 5% e 1%

dove  $R_t$  è un vettore di dimensione pari a  $n = 3$  che contiene i tassi di interesse a breve, medio e lungo periodo,  $d_t$  è il nucleo deterministico che contiene al suo interno le variabili dummy definite nel modello precedente e la costante. I termini  $S_t$ ,  $Q_t$  e  $C_t$  contengono i termini del polinomio di Taylor.

Per la scelta dei ritardi ho seguito le indicazioni fornite dai criteri informativi e i test di autocorrelazione. La scelta è ricaduta su 2 ritardi per i dati a frequenza settimanale e 1 ritardo per i dati a frequenza mensile. In particolare i residui del modello NEC stimato con solo ritardo sui dati settimanali risultavano fortemente autocorrelati quindi sono passato alla stima del modello

Il modello è stato stimato con il metodo OLS cioè come una VAR in cui sono state inserite come variabili esogene tutte le componenti del polinomio di Taylor. Le stime sono riportate nelle Tabelle 3.13 e 3.14.

Tabella 3.13: Coefficienti del meccanismo di aggiustamento multivariato settimanale

Var.	Var. Dip. $\Delta rs_t$			Var. Dip. $\Delta rm_t$			Var. Dip. $\Delta rl_t$		
	Coeff.	S.E.	t-stat.	Coeff.	S.E.	t-stat.	Coeff.	S.E.	t-stat.
$sm_t$	0,146	0,036	4,056	0,031	0,033	0,954	0,005	0,027	0,182
$sl_t$	-0,066	0,025	-2,693	-0,002	0,023	-0,092	-0,003	0,020	-0,126
$sm_t^2$	-0,164	0,066	-2,483	-0,131	0,050	-2,641	-0,062	0,045	-1,378
$sl_t^2$	0,023	0,017	1,297	-0,016	0,019	-0,834	-0,002	0,018	-0,091
$sm_t sl_t$	0,041	0,052	0,794	0,100	0,046	2,145	0,057	0,045	1,251
$sm_t^3$	0,009	0,028	0,332	-0,062	0,031	-2,009	-0,044	0,025	-1,739
$sl_t^3$	-0,000	0,005	-0,102	0,009	0,005	1,735	0,003	0,005	0,506
$sm_t^2 sl_t$	0,044	0,045	0,974	0,118	0,049	2,416	0,069	0,043	1,622
$sm_t sl_t^2$	-0,026	0,023	-1,094	-0,064	0,024	-2,658	-0,036	0,023	-1,592

Test di Wald per i coefficienti di  $sm_t^2$ ,  $sl_t^2$ , ... ,  $sm_t sl_t^2$   
 $W = 106,468$  con p-value = 0,0000 e p-value bootstrap = 0,0201

Test di Ljung-Box per l'autocorrelazione di prim'ordine  
 $Q' = 2,26227$                        $Q' = 0,0265155$                        $Q' = 0,131517$   
p-value = 0,133                      p-value = 0,871                      p-value = 0,717

Test ARCH(2)  
 $LM = 47,23$                        $LM = 25,8786$                        $LM = 55,5477$   
p-value = 0,000                      p-value = 0,000                      p-value = 0,000

Campione: 12/10/1982 - 24/02/2015 (T = 1690). Errori standard HAC (larghezza della banda 5) e wild bootstrap effettuato con 999 replicazioni con distribuzione Gaussiana.

I risultati ottenuti ricalcano circa quelli ottenuti con l'applicazione del modello univariato.

Osservando la Tabella 3.13 è possibile notare che solo per il tasso di rendimento di breve periodo gli spread lineari ( $sm_t$  e  $sl_t$ ) risultano essere statisticamente

significativi. Per quanto riguarda le componenti non lineari per  $\Delta rs_t$  e  $\Delta rl_t$  solo un componente risulta statisticamente significativo. Tutti gli altri vengono rifiutati anche se di poco come nel caso del tasso di rendimento di lungo periodo. Invece per  $\Delta rm_t$  la situazione è opposta: tutti i coefficienti riferiti alla componente non lineare risultano essere significativi ad un livello di confidenza di almeno il 10%, solo uno non risulta significativo.

Sul test di Wald per l'azzeramento congiunto dei parametri dell'aggiustamento non lineare ci sono due precisazioni da fare. In primo luogo l'ipotesi di partenza secondo la teoria delle aspettative sulla struttura a termine è che la matrice di cointegrazione sia nota e assuma una forma ben definita. Come dimostrato da Kristensen e Rahbek (2013) (paragrafo 2.3.2) ciò ci consente di lavorare con le distribuzioni asintotiche dei test standard per cui la procedura bootstrap non sarebbe necessaria. In secondo luogo c'è da considerare che l'ipotesi di eteroschedasticità non può essere esclusa (si veda anche il test ARCH). Ciò implica che seppur le stime dei parametri siano consistenti, le distribuzioni limite dei test possono dipendere da parametri che provengono dal processo di volatilità sottostante (Boswijk *et al.*, 2016). Per ovviare a questo problema una soluzione sarebbe l'utilizzo di procedure bootstrap. Inoltre l'utilizzo di tecniche di simulazione permette di evitare l'impatto di fattori come l'ampiezza del campione e la distribuzione dell'errore che possono influenzare i risultati dei test (paragrafo 2.4.1). La scelta di procedere con il *wildbootstrap* mi è sembrata opportuna soprattutto alla luce di questi ultimi fattori.

I risultati dei test di Wald sia con lo stimatore HAC sia con i residui bootstrap indicano che l'effetto dei parametri dell'aggiustamento non lineare non possono essere ignorati. Infatti l'ipotesi di azzeramento congiunto di questi ultimi parametri viene rifiutata in entrambi i casi.

Il test di Ljung-Box accetta l'ipotesi nulla di assenza di autocorrelazione tra i residui e il loro passato in tutte e 3 le equazioni, quindi il modello è correttamente specificato. Il test ARCH invece indica che la variabilità dei coefficienti è fortemente correlata col passato. Infatti tutti e 3 i test confermano ad un livello di confidenza dell'1% la presenza di eteroschedasticità condizionale come nel caso univariato.

Tabella 3.14: Coefficienti del meccanismo di aggiustamento multivariato mensile

Var.	Var. Dip. $\Delta rs_t$			Var. Dip. $\Delta rm_t$			Var. Dip. $\Delta rl_t$		
	Coeff.	S.E.	t-stat	Coeff.	S.E.	t-stat.	Coeff.	S.E.	t-stat.
$sm_t$	0,527	0,172	3,069	-0,039	0,185	-0,209	-0,099	0,163	-0,607
$sl_t$	-0,233	0,130	-1,792	0,056	0,131	0,429	0,057	0,123	0,465
$sm_t^2$	-0,436	0,287	-1,523	-0,423	0,203	-2,086	-0,345	0,232	-1,488
$sl_t^2$	0,088	0,098	0,897	-0,088	0,107	-0,816	-0,054	0,106	-0,510
$sm_t sl_t$	0,090	0,292	0,309	0,420	0,238	1,762	0,320	0,254	1,256
$sm_t^3$	-0,074	0,113	-0,658	-0,219	0,101	-2,171	-0,157	0,107	-1,473
$sl_t^3$	0,002	0,025	0,078	0,036	0,026	1,359	0,017	0,028	0,618
$sm_t^2 sl_t$	0,242	0,181	1,338	0,408	0,176	2,323	0,296	0,196	1,511
$sm_t sl_t^2$	-0,120	0,104	-1,152	-0,236	0,101	-2,329	-0,159	0,114	-1,396

Test di Wald per i parametri  $sm_t^2, sl_t^2, \dots, sm_t sl_t^2$   
 $W = 68,0365$  con p-value = 0,0000 e p-value bootstrap = 0,00102

Test di Ljung-Box per l'autocorrelazione di prim'ordine  
 $Q' = 0,015842$                        $Q' = 0,004336$                        $Q' = 0,008814$   
p-value = 0,900                      p-value = 0,948                      p-value = 0,925

Test ARCH(2)  
 $LM = 24,1245$                        $LM = 4,72214$                        $LM = 1,91581$   
p-value = 0,000                      p-value = 0,0943                      p-value = 0,3837

Campione: 10/1982 - 02/2015 (T = 389). Errori standard HAC (larghezza della banda 5) e wild bootstrap effettuato con 999 replicazioni con distribuzione Gaussiana.

Dai risultati della stima del NEC sui dati mensili si può notare che:

- gli spread  $sm_t$  e  $sl_t$  hanno valori considerevoli e sono statisticamente significativi solo nel caso del tasso di interesse di breve periodo, mentre per i tassi di rendimento di medio e lungo termine assumono valori molto vicini allo zero e non sono statisticamente significativi.
- I coefficienti relativi all'aggiustamento non lineare non risultano significativi ai livelli standard sia nei tassi di rendimento di breve che di lungo periodo.
- Per quanto concerne il tasso di rendimento di medio periodo ( $\Delta rm_t$ ) 5 coefficienti su 7 risultano statisticamente significativi.

Anche per i dati mensili ho utilizzato la versione *wild bootstrap* del Wald test in quanto anche su queste serie non può essere esclusa totalmente l'ipotesi che non vi sia eteroschedasticità condizionale. Entrambi i test di Wald indicano che l'effetto dell'aggiustamento è statisticamente significativo sebbene quasi tutti i singoli test dei coefficienti dei parametri non lineari del meccanismo di aggiustamento accettano l'ipotesi nulla di azzeramento.

Il secondo test di mispecificazione indica che i residui non sono autocorrelati, infatti il test di Ljung-Box accetta in tutti e tre i casi l'ipotesi nulla.

Infine nella serie dei tassi di rendimento di breve periodo non può essere rifiutata la presenza di eteroschedasticità, caso contrario per il tasso di lungo termine che accetta l'ipotesi che la varianza condizionale sia costante nel tempo. Per il tasso di rendimento a medio termine la situazione è uguale al caso univariato: il p-value è molto vicino al valore soglia del 10% per cui si potrebbe accettare l'ipotesi di omoschedasticità.

#### 3.4.4 Analisi *out-of-sample* multivariata

I risultati ottenuti con l'applicazione del modello multivariato sembrano confermare gli stessi ottenuti nel caso univariato. Anche per il modello NEC multivariato è stato calcolato il RMSE ed è stato effettuato il test di Diebold e Mariano (1995) per valutarne capacità previsiva. L'ampiezza del campione *out-of-sample* è la stessa, quindi sono 52 osservazioni nel caso dei dati settimanali e 12 per i dati mensili.

Dai risultati riportati nella tabella 3.15 il modello non lineare ha la radice dell'errore medio quadratico nel complesso minore rispetto alle alternative lineari. Le differenze di valore non sono sostanziali, anzi divergono di poco per cui è possibile ipotizzare che abbia migliore o eguale capacità previsiva.

Nella Tabella 3.16 sono riportati i test di Diebold e Mariano. Anche qui non c'è una distinzione statisticamente significativa tra i valori previsti dal modello non lineare e quelli lineari. In sostanza il modello NEC multivariato non ha migliore capacità predittiva rispetto a quelli lineari quindi è possibile escludere l'ipotesi che ci sia "*overfitting*".

Il modello nel caso dei dati mensili si comporta a grandi linee allo stesso modo; il RMSE risulta inferiore rispetto ai modelli competitivi tranne nel caso del tasso di interesse di lungo periodo. Il test DM accetta l'ipotesi nulla in tutti i casi, eccezione fatta solo con il modello VAR sui livelli per il tasso di interesse a breve, quindi le previsioni fatte con il modello NEC e quelli lineari sono statisticamente uguali. In conclusione i risultati ottenuti applicando il modello NEC e i modelli lineari nel campione *out-of-sample* sono generalmente uguali, un'ulteriore evidenza che il modello non è mispecificato e non c'è alcun segno di "*overfitting*".

Tabella 3.15: RMSE su dati settimanali

Modello / Serie	$rs_t^{3mesi}$	$rs_t^{2anni}$	$rl_t^{10anni}$
NEC	0,0219*	0,0476	0,0802
RW	0,2391	0,3074	0,2685
DVAR	0,0241	0,0472*	0,0785*
VECM1	0,1263	0,3558	0,3299
VECM2	0,1272	0,3577	0,2783
LVAR	0,2520	0,1197	0,1501

I valori contrassegnati con l'asterisco indicano i valori RMSE più bassi.

Tabella 3.16: Diebold-Mariano sui dati settimanali

Funzione V-shape			
Modello / Serie	$rs_t^{3mesi}$	$rs_t^{2anni}$	$rl_t^{10anni}$
RW	-0,1695	-0,1675	-0,4525
DVAR	-0,2968	-0,0165	0,1112
VECM1	-0,6096	-0,1917	-0,4778
VECM2	-0,4541	-0,1929	-0,4708
LVAR	-1,0003	-0,2886	-0,2878
Funzione Linex			
Modello / Serie	$rs_t^{3mesi}$	$rs_t^{2anni}$	$rl_t^{10anni}$
RW	-0,1658	-0,0872	-0,3344
DVAR	-0,2258	0,0357	0,1112
VECM1	-0,3009	-0,0945	-0,3228
VECM2	-0,1009	-0,0949	-0,3484
LVAR	-0,3521	-0,2168	-0,1538

Test a due code: valori negativi del test indicano una miglior accuratezza previsiva per il modello non lineare viceversa per i valori positivi. Il numero degli asterischi posti di fianco ai numeri indicano rispettivamente il rifiuto dell'ipotesi nulla al 10%, 5% e 1%

Tabella 3.17: RMSE su dati mensili

Modello / Serie	$rs_t^{3mesi}$	$rs_t^{2anni}$	$rl_t^{10anni}$
NEC	0,0534*	0,1145*	0,1914
RW	0,2391	0,3074	0,2685
DVAR	0,0596	0,1151	0,1830*
VECM1	0,1263	0,3558	0,3299
VECM2	0,1272	0,3577	0,2783
LVAR	0,2609	0,1793	0,2430

I valori contrassegnati con l'asterisco indicano i valori RMSE più bassi.

Tabella 3.18: Diebold-Mariano sui dati mensili

Funzione V-shape			
Modello / Serie	$rs_t^{3mesi}$	$rs_t^{2anni}$	$rl_t^{10anni}$
RW	-1,0322	-0,6795	-0,4333
DVAR	-0,4401	0,0609	-0,0204
VECM1	-0,7405	-0,7139	-0,5971
VECM2	-1,0139	-0,7246	-0,4817
LVAR	-3,8021***	-0,2779	-0,3384
Funzione Linex			
Modello / Serie	$rs_t^{3mesi}$	$rs_t^{2anni}$	$rl_t^{10anni}$
RW	-0,5422	-0,3675	-0,3789
DVAR	-0,4133	0,0781	0,1169
VECM1	-0,6822	-0,3791	-0,5851
VECM2	-0,9442	-0,3788	-0,4635
LVAR	-2,2640**	-0,2944	-0,1801

Test a due code: valori negativi del test indicano una miglior accuratezza previsiva per il modello non lineare viceversa per i valori positivi. Il numero degli asterischi posti di fianco ai numeri indicano rispettivamente il rifiuto dell'ipotesi nulla al 10%, 5% e 1%

### 3.5 Interpretazione dei risultati

I risultati finora esposti evidenziano che, da un punto di vista statistico, il meccanismo di aggiustamento, approssimato da un polinomio di terzo grado, permette di descrivere l'effetto soglia presente nella relazione dei tassi di interesse e non inserirlo nel modello può condurre a problemi di misspecificazione. Da un punto di vista economico invece, la componente non lineare del modello non trova un'immediata interpretazione, ma è necessaria per descrivere tutti quegli elementi come i costi di transazione, i premi per il rischio variabili e gli interventi dei *policy maker* che impediscono all'equilibrio.

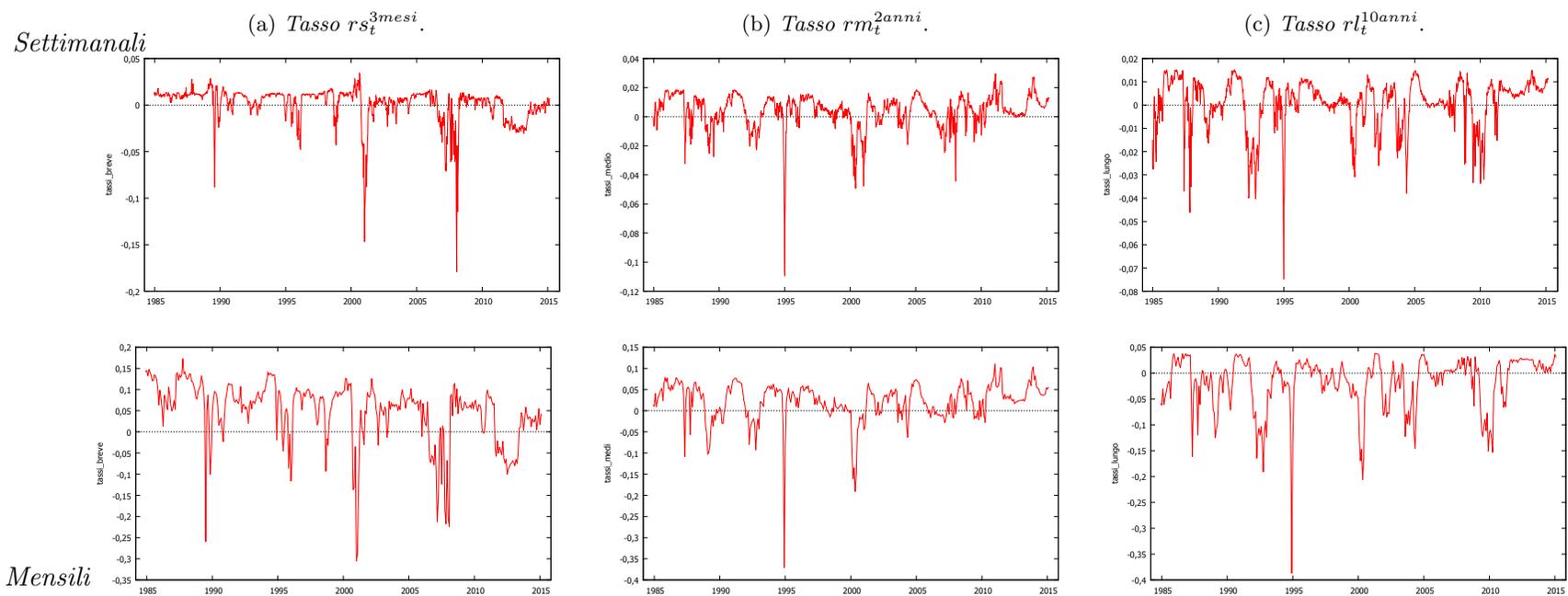
In questo lavoro si vuole dimostrare che il meccanismo di aggiustamento non lineare è in grado di prevedere gli eventi occasionali come gli interventi di politica monetaria per cui le serie dei tassi di interesse dei bond americani degli ultimi trent'anni si prestano allo scopo preposto. Infatti l'economia statunitense ha subito diversi shock nel corso di questi ultimi tempi (la crisi del 1987, lo scoppio della bolla "dot.com" e la crisi dei mutui sub-prime) quindi questo data set contiene al suo interno diversi eventi che hanno portato i *policy maker* ad agire in modo marcato sul mercato dei titoli governativi.

Per verificare la capacità del modello multivariato NEC di catturare effettivamente i maggiori interventi di politica monetaria e di capire quando il meccanismo di aggiustamento ha giocato un ruolo significativo nel guidare i tassi di rendimento dei bond mi sono avvalso di un'analisi grafica.

Quindi nella Figura 3.3 sono riportati i grafici ottenuti attraverso la stima degli effetti combinati dei coefficienti relativi agli spread sui rispettivi tassi di interesse.

Nel complesso si può immediatamente notare che i picchi più ampi coincidono con gli avvenimenti che hanno causato le più grandi perturbazioni sul mercato obbligazionario. Andando però a fare un confronto tra i singoli grafici, è possibile constatare che il meccanismo di aggiustamento svolge un ruolo lievemente diverso per ogni singolo tasso. Infatti nei rendimenti dei bond di breve termine l'impatto del meccanismo di correzione dell'errore ha un peso di maggior intensità mentre nei

Figura 3.3: Aggiustamenti storici



tassi di più lungo periodo è minore. La spiegazione può essere dovuta dal fatto che gli interventi di politica monetaria agiscono in modo diretto sui tassi di breve periodo per cui quelli maggiormente influenzati sono i bond di breve durata.

Scorrendo nei grafici dei tassi di breve periodo, le date in cui il meccanismo di aggiustamento ha avuto più influenza su tali bond coincidono con:

- l'implementazione del *Brady Plan* nel luglio del 1989,
- lo scoppio della bolla “dot.com” nel Marzo del 2000 seguita dall'attacco delle torri gemelle (settembre 2001),
- la crisi dei mutui sub-prime (a partire da agosto 2007).

Andando ad osservare invece i tassi di medio e lungo termine le date in cui si registrano i maggiori picchi corrispondono:

- alla crisi del 1987, in cui la Fed iniettò liquidità nel sistema attraverso operazioni sul mercato dei buoni del tesoro,
- alla crisi del peso messicano nel dicembre del 1994, in cui il governo degli Stati Uniti intervenne poiché alcuni dei titoli di stato messicani erano collegati ai treasury bond americani attraverso il Brady Plan,
- allo scoppio della bolla “dot.com” (2000) e all'attacco alle torri gemelle (2001), in cui la banca centrale americana intervenne sia con riduzioni sostanziali del tasso di interesse sia con operazioni di mercato aperto permanenti (POMO), cioè attraverso l'acquisto nei mercati secondari dei titoli di stato,
- la recente crisi dei mutui sub-prime, in cui la FOMC decise di intervenire con strumenti e con programmi di politica monetaria straordinari: *Quantitative Easing* e *Operation Twist*.

In particolare, gli eventi dell'ultima crisi ha catturato in modo particolare la mia attenzione. Il meccanismo di aggiustamento dei tassi di lungo periodo sembra seguire in modo dettagliato tutti gli avvenimenti di politica monetaria di quel periodo. Infatti nel 2008 la FOMC, per stimolare l'attività economica, decise di adottare

strumenti non convenzionali come il *Quantitative Easing* (QE). Questo programma prevede l'acquisto sul mercato da parte della banca centrale di diversi attività finanziarie e titoli governativi in modo da ridurre i rendimenti di lungo termine e aumentare l'efficacia della politica monetaria (Thornton, 2012). Il primo *Quantitative Easing* (QE1) venne effettuato nel marzo del 2009, il secondo (QE2) tra novembre del 2010 e il giugno del 2011 ed infine il terzo (QE3) a settembre del 2012. Oltre a queste operazioni venne implementato nel settembre del 2011 l'*Operation Twist*; un programma di scambio di titoli governativi per stimolare la domanda aggregata.

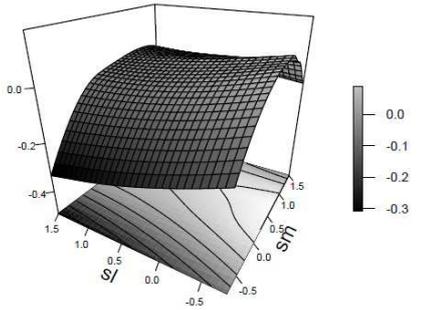
Nei grafici c'è una sostanziale differenza di come la politica monetaria abbia agito tra i diversi tassi di breve, di medio e di lungo termine. Infatti, nei bond a breve termine l'impatto del meccanismo di aggiustamento è più immediato e decisivo, mentre per il tasso di più lungo termine l'impatto è più lieve, ma più duraturo. Inoltre anche fra i tassi di medio e lungo termine c'è una sostanziale differenza, riguardo soprattutto il periodo della crisi di mutui sub-prime. Una possibile spiegazione può essere ricondotta al fatto che gran parte dei programmi di stimolo dell'economia avevano come scopo quello di ridurre i tassi di lungo periodo. Infatti i diversi piani di azione della Fed, come il *Quantitative Easing* e l'*Operation Twist*, agivano in modo esclusivo sui titoli a più lunga scadenza ed avevano durata mensile se non semestrale. Infatti, gli effetti degli ultimi interventi straordinari effettuati dalla Fed sono maggiormente visibili sui rendimenti dei bond con scadenza decennale.

Nella Figura 3.4 sono riportate le superfici di aggiustamento per ogni singolo tasso di interesse. I grafici sono stati ottenuti attraverso un metodo di approssimazione non parametrica locale tra i singoli tassi di rendimento e gli spread con un ritardo pari a uno ( $sm_{t-1}$  e  $sl_{t-1}$ ). Lo scopo è quello di verificare come i tassi a diversa scadenza rispondo ai singoli spread.

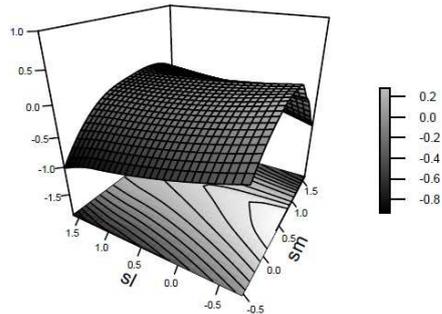
Se il meccanismo avesse seguito una funzione lineare, come nel modello VECM, allora sarebbe emersa una superficie in cui le linee sarebbero state equidistanti e parallele tra di loro e ci sarebbero state ampie regioni di aggiustamento. Invece, come si può notare dai grafici, ci sono ampie zone in cui l'aggiustamento rimane pressochè inattivo, mentre al di fuori di queste regioni il meccanismo non lineare

Figura 3.4: Superfici di aggiustamento.

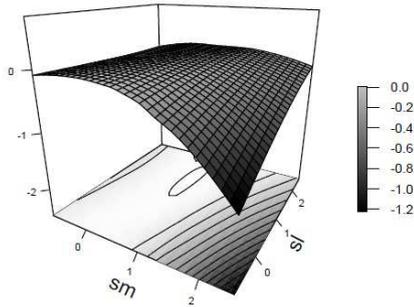
Tassi settimanali  
(a)  $Tasso\ r_s^3mesi$ .



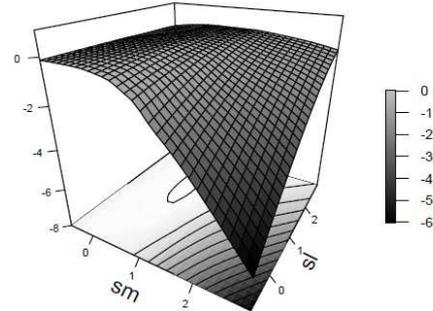
Tassi mensili  
(b)  $Tasso\ r_s^3mesi$ .



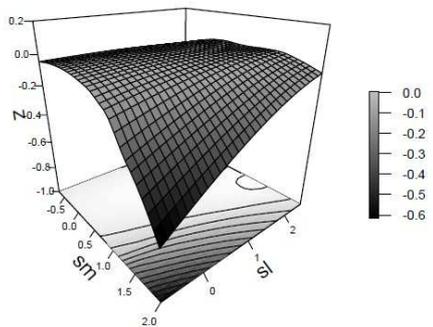
(c)  $Tasso\ r_m^2anni$ .



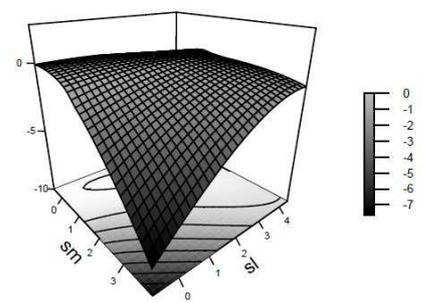
(d)  $Tasso\ r_m^2anni$ .



(e)  $Tasso\ r_t^{10anni}$ .



(f)  $Tasso\ r_t^{10anni}$ .



opera in modo discreto e sostanziale.

Andando ad osservare in primo luogo i grafici dei movimenti dei rendimenti dei bond di breve periodo, le zone in cui scatta l'aggiustamento sono più ampie rispetto ai bond di più lungo termine. Infatti i tassi di breve periodo rispondono principalmente ai valori che assumono gli spread; se questi ultimi assumono valori inconsueti, o meglio estremi, il meccanismo di correzione entra in funzione altrimenti non avviene nulla.

Nei tassi di più lungo termine il meccanismo ha una forma un pò più particolare; i rendimenti sembrano rispondere maggiormente alle forme più inconsuete della *yield curve*. Infatti i grafici mostrano che il momento in cui si attiva la correzione verso l'equilibrio avviene quando i valori degli spread si muovono in direzione opposte. Quindi, quando la *yield curve* ha un'inclinazione ed una curvatura "standard" i tassi si muovono liberamente mentre, quando la curva assume forme "inusuali" il meccanismo entra in gioco.

L'analisi empirica della teoria delle aspettative eseguita con un modello di cointegrazione non lineare come il NEC può evidenziare le modalità con cui viene condotta la politica monetaria. Se si suppone che la relazione di equilibrio di lungo periodo sia una regione di spazio anziché su un solo punto, ciò ci permette di valutare l'ipotesi che gli interventi condotti dai *policy maker* abbiano avuto come scopo quello di far rientrare gli spread in un range di valori. Le stime e i test condotti confermano tale ipotesi ed evidenziano che nel momento in cui l'inclinazione o la curvatura della *yield curve* presenta valori inconsueti, cioè quando gli operatori formulano le loro aspettative non in linea con quelli della banca centrale o con le situazioni di normalità, scattano gli interventi di politica monetaria. Ulteriori prove a sostegno di questa ipotesi possono essere facilmente viste nei grafici in cui il meccanismo di aggiustamento non lineare, che connota gli interventi della banca centrale, riveste un ruolo cruciale nel guidare i tassi di interesse soprattutto durante i periodi in cui sono avvenute le maggiori crisi.

## 3.6 Conclusioni

Il contributo principale di questo capitolo è rivolto principalmente a verificare se un meccanismo di aggiustamento non lineare è un'importante caratteristica dei titoli di stato americani. Nello specifico se un meccanismo non lineare come un polinomio di terzo grado sia capace di visualizzare il comportamento e l'influenza dei policy maker sul mercato dei titoli obbligazionari sovrani americani nel corso di questi ultimi anni.

Nel primo capitolo sono state elencate tutte le implicazioni della teoria delle aspettative che derivano dall'applicazione del modello di Campbell e Shiller (1987). L'analisi preliminare ci permette verificare quale di queste ipotesi trovano riscontro empirico nel data set. Questa prima verifica è stata condotta in modo differente dal lavoro di Lucchetti e Palomba (2009) sia da un punto di vista metodologico che teorico. Infatti alla luce dagli ultimi risultati di Cao *et al.* (2014) e di Gunnella (2015), riguardanti sempre la struttura a termine secondo la teoria delle aspettative, l'evidenza empirica risulta più favorevole se vengono utilizzati test e modelli più robusti all'eteroschedasticità. Inoltre, accanto alle implicazioni "low frequency", sono state considerate anche le implicazioni "high frequency" della teoria delle aspettative che riguardano più da vicino la struttura dinamica dell'aggiustamento come descritto nel primo capitolo.

L'implementazione del *wild bootstrap* nei test consente di ovviare ai problemi di *size distortion* che la volatilità variabile ha sui test. Infatti i lavori di Cavaliere e Taylor (2008), Cavaliere *et al.* (2010b), Cavaliere *et al.* (2012) e Boswijk *et al.* (2016) mostrano come le distribuzione limite dei test non sono asintoticamente pivotali nel momento in cui viene meno l'ipotesi di volatilità costante nel tempo sia nei test di stazionarietà che nei test per il rango di cointegrazione di Johansen (1995). Per cui i test sono stati condotti sia normalmente che nella loro versione *wild bootstrap*. L'obiettivo di tutta questa analisi è di fornire sufficiente evidenza che ci permette di poter passare ad una modellazione non standard eliminando ogni dubbio sul fatto che la componente non lineare possa essere un potenziale sottoprodotto di eteroschedasticità *unaccounted*.

Nonostante l'utilizzo dei metodi di simulazione tutte le implicazioni introdotte

da Campbell e Shiller (1987) sono rifiutate dai test. Infatti alle condizioni di breve e di lungo termine della teoria delle aspettative non trovano alcun supporto empirico sia per i dati a frequenza settimanale che mensile. Questi risultati mostrano che, sebbene ci sia una lieve tendenza dei tassi di interesse a muoversi nella direzione necessaria per riportare gli spread al loro valore di equilibrio, questa non può essere catturata con modellazione lineare.

Le stime di Lucchetti e Palomba (2009) sono state ottenute utilizzando un modello per la specificazione della varianza condizionale, il GARCH nel caso univariato e il BEKK nel multivariato. In questo lavoro gli effetti dell'eteroschedasticità condizionale che non condizionale sono stati considerati utilizzando procedure che non impongono alcuna specificazione parametrica per la volatilità come lo stimatore HAC e il *wild bootstrap*. Come già discusso nel capitolo 2 la scelta di utilizzare quest'ultima procedura deriva essenzialmente dalla natura dei dati presi in esame e dalla semplicità di calcolo. Infatti il *wild bootstrap* risulta uno strumento adatto per il modello NEC sia nel caso in cui la distribuzione asintotica dei test non sia standard sia nel caso in cui i dati mostrano considerevoli cambi di volatilità.

Il campione sui cui è stato svolto il lavoro di Lucchetti e Palomba (2009) non tiene conto di tutto il periodo che concerne la crisi dei mutui *sub-prime* e le conseguenti azioni di politica monetaria. La recente crisi ha costretto la FED ad utilizzare strumenti non convenzionali come il *Quantitative Easing*. Alla luce di ciò risulta interessante verificare come il meccanismo di aggiustamento lineare si comporti.

L'analisi grafica ha mostrato come il modello NEC sia capace di individuare e di distinguere fra le diverse azioni di politica monetaria che hanno avuto differenti obiettivi. Infatti i grafici ottenuti con le stime degli effetti combinati dei coefficienti relativi agli spread sui singoli tassi (Figura 3.3) mostrano come l'impatto del meccanismo di aggiustamento sia generalmente prossimo allo zero tranne nei periodi di turbolenza economica dove, invece, il meccanismo assume valori piuttosto ampi.

Il NEC riesce a distinguere i diversi programmi operativi non convenzionali adottati dalla FED dopo la recente crisi. Infatti, nei bond a breve termine l'impatto del meccanismo di aggiustamento è più immediato e decisivo, mentre per il tasso di

più lungo termine l'impatto è più lieve, ma più persistente. Ciò risulta congruente con le azioni adottate dai *policy maker* che in un primo momento hanno ridotto i tassi di interesse a breve e successivamente hanno adottato programmi più longevi come il *Quantitative Easing* e il *Operation Twist* per ridurre i tassi di lungo termine.

Infine un ulteriore contributo fornito da questo capitolo è stato quello di verificare se il meccanismo di aggiustamento non lineare sia un'importante caratteristica empirica anche per i dati mensili. Infatti in letteratura, partendo da Campbell e Shiller (1987), Balke e Fomby (1992), Seo (2003), ecc., i modelli sono stati utilizzati sui tassi di interesse dei bond americani ma con frequenza mensile. I risultati del modello NEC applicato a questo data set risultano pressoché simili a quelli settimanali. Infatti i diversi test di specificazione confermano le stesse ipotesi, le prestazioni previsive sono le stesse e le superfici di aggiustamento sono perfettamente uguali. L'unica differenza sostanziale emerge dal confronto dei tassi con scadenza a 2 anni nella Figura 3.3. L'impatto delle più recenti crisi non sono del tutto visibili nei dati mensili mentre nei dati settimanali, seppur con minore intensità, l'effetto degli interventi viene catturato dal modello.

# Bibliografia

- ANG, A. E. PIAZZESI, M. (2003). *A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables*. Journal of Monetary Economics, 50(4): 745–787.
- BALKE, N. S. E. FOMBY, T. B. (1992). *Threshold cointegration*. Research paper, Federal Reserve Bank of Dallas.
- BEC, F. E. RAHBEK, A. (2004). *Vector equilibrium correction models with non-linear discontinuous adjustments*. Econometrics Journal, 7(2): 628–651.
- BECKMANN, J. (2013). *Nonlinear adjustment, purchasing power parity and the role of nominal exchange rates and prices*. The North American Journal of Economics and Finance, 24(C): 176–190.
- BERAN, R. (1988). *Prepivoting test statistics: a bootstrap view of asymptotic refinements*. Journal of the American Statistical Association, 83: 687–697.
- BERAN, R. E. DUCHARME, G. (1991). *Asymptotic Theory for Bootstrap Methods in Statistics*. Publications CRM, Centre de recherches mathématiques.
- BILLINGSLEY, P. (1968). *Convergence in Probability Measures*. John Wiley.
- BOSWIJK, H. P., CAVALIERE, G., RAHBEK, A. E. TAYLOR, A. R. (2016). *Inference on co-integration parameters in heteroskedastic vector autoregressions*. Journal of Econometrics, 192(1): 64–85.
- BÜHLMANN, P. (1997). *Sieve bootstrap for time series*. Bernoulli, 3: 123–148.
- (2002). *Bootstraps for Time Series*. Statistical Science, 17(1): 52–72.

- BUSETTI, F. E TAYLOR, A. M. R. (2003). *Variance Shifts, Structural Breaks, and Stationarity Tests*. Journal of Business and Economic Statistics, 21(4): 510–531.
- CAI, B., GAO, J. E TJOSTHEIM, D. (2015). *A New Class of Bivariate Threshold Cointegration Models*. Monash econometrics and business statistics working papers, Monash University, Department of Econometrics and Business Statistics.
- CAMPBELL, J. Y. E SHILLER, R. J. (1987). *Cointegration and tests of present value models*. Journal of Political Economy, 94: 1062 – 1088.
- (1991). *Yield spreads and interest rate movements: A bird’s eye view*. Journal of Political Economy, 95.
- CANER, M. E HANSEN, B. (2001). *Threshold Autoregression with a Unit Root*. Econometrica, 69: 1555–1596.
- CAO, R., ZHU, X. E RAHMAN, S. (2014). *Yield Factors, the Expectations Hypothesis and Regime Shifts*. ANNALS OF ECONOMICS AND FINANCE, 15(2): 411–431.
- CARLSON, M. (2006). *A Brief History of the 1987 Stock Market Crash*. Discussion of the federal reserve response, Federal Reserve Response.
- CARRIERO, A., FAVERO, C. E KAMINSKA, I. (2004). *Financial Factors, Macroeconomic Information and the Expectations Theory of the Term Structure of Interest Rates*. Working Papers 253, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University.
- CAVALIERE, G. (2005). *Unit Root Tests under Time-Varying Variances*. Econometric Reviews, 23(3): 259–292.
- CAVALIERE, G., DE ANGELIS, L., RAHBEK, A. E TAYLOR, A. M. R. (2016). *Determining the Cointegration Rank in Heteroskedastic VAR Models of Unknown Order*. Essex Finance Centre Working Papers 17454, University of Essex, Essex Business School.
- CAVALIERE, G., PHILLIPS, P. C., SMEEKES, S. E TAYLOR, A. R. (2012). *Lag Length Selection for Unit Root Tests in the Presence of Nonstationary Volati-*

- lity*. Cowles foundation discussion papers, Cowles Foundation for Research in Economics, Yale University.
- CAVALIERE, G., RAHBEK, A. E TAYLOR, A. M. R. (2014). *Bootstrap Determination of the Co-Integration Rank in Heteroskedastic VAR Models*. *Econometric Reviews*, 33(5-6): 606–650.
- CAVALIERE, G., RAHBEK, A. E TAYLOR, A. R. (2010a). *Cointegration Rank Testing Under Conditional Heteroskedasticity*. *Econometric Theory*, 26(06): 1719–1760.
- (2010b). *Testing for co-integration in vector autoregressions with non-stationary volatility*. *Journal of Econometrics*, 158(1): 7–24.
- CAVALIERE, G. E TAYLOR, A. R. (2008). *Bootstrap Unit Root Tests For Time Series With Nonstationary Volatility*. *Econometric Theory*, 24(01): 43–71.
- CECCHETTI, S. G. (2008). *Crisis and Responses: the Federal Reserve and the Financial Crisis of 2007-2008*. Nber working papers, National Bureau of Economic Research, Inc.
- CHANG, Y., PARK, J. Y. E SONG, K. (2006). *Bootstrapping cointegrating regressions*. *Journal of Econometrics*, 133(2): 703–739.
- COOK, T. E HAHN, T. (1989). *The effect of changes in the federal funds rate target on market interest rates in the 1970s*. *Journal of Monetary Economics*, 24(3): 331–351.
- DANIEL, L. T. (2004). *Tests of the expectations hypothesis: resolving the Campbell-Shiller paradox*. Working Papers 2003-022, Federal Reserve Bank of St. Louis.
- DAVIDSON, J. (1994). *Stochastic Limit Theory: An Introduction for Econometricians*. Oxford University Press.
- DAVIDSON, R. (2009). *Bootstrapping econometric models*. Working papers, HAL.
- DAVIDSON, R. E FLACHAIRE, E. (2001). *The wild bootstrap, tamed at last*. Rap. techn., London School of Economics and Political Science, LSE Library.

- DAVISON, A. E HINKLEY, D. (1997). *Bootstrap Methods and Their Application*. Cambridge University Press, Cambridge UK.
- DE JONG, R. M. (2002). *Nonlinear minimization estimators in the presence of cointegrating relations*. Journal of Econometrics, 110(2): 241–259.
- DELLA CORTE, P., SARNO, L. E THORNTON, D. L. (2007). *The Expectation Hypothesis of the Term Structure of Very Short-Term Rates: Statistical Tests and Economic Value*. CEPR Discussion Papers 6445, C.E.P.R. Discussion Papers.
- DICKEY, D. A. E FULLER, W. A. (1981). *Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root*. Econometrica, 49(4): 1057–72.
- DIEBOLD, F. X. E LI, C. (2003). *Forecasting the term structure of government bond yields*. Cfs working paper series, Center for Financial Studies (CFS).
- DIEBOLD, F. X. E MARIANO, R. S. (1995). *Comparing Predictive Accuracy*. Journal of Business & Economic Statistics, 13(3): 253–63.
- DIEBOLD, F. X., RUDEBUSCH, G. D. E BORAG[CARON]AN ARUOBA, S. (2006). *The macroeconomy and the yield curve: a dynamic latent factor approach*. Journal of Econometrics, 131(1-2): 309–338.
- DONG, C., GAO, J., TJOSTHEIM, D. E YIN, J. (2014). *Specification Testing for Nonlinear Multivariate Cointegrating Regressions*. Monash Econometrics and Business Statistics Working Papers 8/14, Monash University, Department of Econometrics and Business Statistics.
- DUFFIE, D. E KAN, R. (1996). *A Yield-Factor Model Of Interest Rates*. Mathematical Finance, 6(4): 379–406.
- EFRON, B. (1979). *Bootstrap methods: Another look at the jackknife*. Annals of Statistics, 7: 1–26.
- EFRON, B. E TIBSHIRANI, R. (1993). *An Introduction to the Bootstrap*. Chapman and Hall, New York.

- EL-SHAGI, M. (2011). *An evolutionary algorithm for the estimation of threshold vector error correction models*. *International Economics and Economic Policy*, 8(4): 341–362.
- ELLIOTT, G., ROTHENBERG, T. J. E STOCK, J. H. (1996). *Efficient Tests for an Autoregressive Unit Root*. *Econometrica*, 64(4): 813–36.
- ENDERS, W. E GRANGER, C. W. J. (1998). *Unit Root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates*. Staff General Research Papers 1388, Iowa State University, Department of Economics.
- ENGLE, R. F. E GRANGER, C. W. J. (1987). *Co-integration and error correction: Representation, estimation, and testing*. *Econometrica*, 55: 251 – 276.
- ESCANCIANO, J. C. (2004). *Model Checks Using Residual Marked Empirical Processes*. Faculty working papers, School of Economics and Business Administration, University of Navarra.
- ESCRIBANO, A. (2004). *Nonlinear Error Correction: The Case Of Money Demand In The United Kingdom (1878 2000)*. *Macroeconomic Dynamics*, 8(01): 76–116.
- ESCRIBANO, A. E MIRA, S. (2002). *Nonlinear error correction models*. (2001-03).
- ESTRELLA, A. E HARDOUVELIS, G. A. (1991). *The Term Structure as a Predictor of Real Economic Activity*. *Journal of Finance*, 46(2): 555–76.
- ESTRELLA, A. E MISHKIN, F. S. (1996). *The yield curve as a predictor of U.S. recessions*. *Current Issues in Economics and Finance*, 2(Jun).
- FAMA, E. F. (1984). *The information in the term structure*. *Journal of Financial Economics*, 13(4): 509–528.
- FISHER, I. (1896). *Appreciation and interests*. *Publications of the American Economic Association*, 11: 21–29.
- FRANKEL, J. (2015). *The Plaza Accord, 30 Years Later*. Nber working papers, National Bureau of Economic Research, Inc.

- FROOT, K. A. (1989). *New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates*. *Journal of Finance*, 44(2): 283–305.
- GHYSELS, E., HARVEY, A. E RENAULT, E. (1996). *Stochastic Volatility*. Cahiers de recherche 9613, Universite de Montreal, Departement de sciences economiques.
- GIESE, J. V. (2008). *Level, Slope, Curvature: Characterising the Yield Curve in a Cointegrated VAR Model*. *Economics - The Open-Access, Open-Assessment E-Journal*, 2: 1–20.
- GONCALVES, S. E KILIAN, L. (2004). *Bootstrapping autoregressions with conditional heteroskedasticity of unknown form*. *Journal of Econometrics*, 123(1): 89–120.
- GONZALO, J. E GONZÁLEZ, M. (2000). *Econometric implications of non-exact present value models*. Rap. tecn., Universidad Carlos III de Madrid Departamento de Economía.
- GOODFRIEND, M. (2002). *The phases of U.S. monetary policy : 1987 to 2001*. *Economic Quarterly*, 1(Fall): 1–17.
- GRANGER, C. W. (2001). *Overview Of Nonlinear Macroeconometric Empirical Models*. *Macroeconomic Dynamics*, 5(04): 466–481.
- GRANGER, C. W. J. E TERASVIRTA, T. (1993). *Modelling Non-Linear Economic Relationships*. Oxford University Press.
- GUILLÉN, M. F. (2015). *The Global Economic & Financial Crisis: A Timeline*. Lauder institute.
- GUNNELLA, V. (2015). *Essays in macroeometrics: methodological aspects and empirical applications*. alma.
- HALL, A. D., ANDERSON, H. M. E GRANGER, C. W. J. (1992). *A Cointegration Analysis of Treasury Bill Yields*. *The Review of Economics and Statistics*, 74(1): 116–26.
- HALL, P. (1992). *The Bootstrap and Edgeworth Expansion*. Springer, New York.

- HALL, P., HOROWITZ, J. L. & JING, B. Y. (1995). *On blocking rules for the bootstrap with dependent data*. *Biometrika*, 82: 561–574.
- HAMORI, S. & TOKIHISA, A. (1997). *Testing for a unit root in the presence of a variance shift*. *Economics Letters*, 57(3): 245–253.
- HANNAN, E. J. & QUINN, B. G. (1979). *The determination of the order of an autoregression*. *Journal of the Royal Statistical Society (B)*, 41: 190–195.
- HANSEN, B. E. (1992). *Heteroskedastic cointegration*. *Journal of Econometrics*, 54(1-3): 139–158.
- (1995). *Regression with Nonstationary Volatility*. *Econometrica*, 63(5): 1113–1132.
- HANSEN, B. E. & SEO, B. (2002). *Testing for two-regime threshold cointegration in vector error-correction models*. *Journal of Econometrics*, 110(2): 293–318.
- HARDOUVELIS, G. A. (1992). *The term structure spread and future changes in long and short rates: is there a puzzle?* Research Paper 9208, Federal Reserve Bank of New York.
- (1994). *The term structure spread and future changes in long and short rates in the G7 countries: Is there a puzzle?* *Journal of Monetary Economics*, 33(2): 255–283.
- HICKMAN, W. B. (1942). *The term structure of interest rates: An exploratory analysis*. National Bureau of Economic Research.
- HOROWITZ, J. L. (2003). *The bootstrap in econometrics*. *Statistical Science*, 18: 211–218.
- HSU, C. & KUGLER, P. (1997). *The Revival of the Expectations Hypothesis of the US Term Structure of Interest Rates*. *Economics Letters*, 55(1): 115–120.
- HURN, S. & BECKER, R. (2007). *Testing for nonlinearity in mean in the presence of heteroskedasticity*. NCER Working Paper Series 8, National Centre for Econometric Research.

- HÄRDLE, W., HOROWITZ, J. L. E KREISS, J.-P. (2003). *Bootstrap methods for time series*. International Statistical Review, 71: 435–459.
- JARDET, C., MONFORT, A. E PEGORARO, F. (2013). *No-arbitrage Near-Cointegrated VAR(p) term structure models*. Journal of Banking & Finance, 37: 389–402.
- JENTSCH, C., PAPARODITIS, E. E POLITIS, D. N. (2014). *Block Bootstrap Theory for Multivariate Integrated and Cointegrated Processes*. Working Papers 14-18, University of Mannheim, Department of Economics.
- JOHANSEN, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.
- JOSLIN, S., PRIEBSCHE, M. E SINGLETON, K. J. (2014). *Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks*. Journal of Finance, 69(3): 1197–1233.
- KAPETANIOS, G. (2000). *Information Criteria, Model Selection Uncertainty and the Determination of Cointegration Rank*. NIESR Discussion Papers 166, National Institute of Economic and Social Research.
- (2004). *The Asymptotic Distribution Of The Cointegration Rank Estimator Under The Akaike Information Criterion*. Econometric Theory, 20(04): 735–742.
- KAPETANIOS, G. E SHIN, Y. (2008). *GLS detrending-based unit root tests in nonlinear STAR and SETAR models*. Economics Letters, 100(3): 377–380.
- KIM, T.-H., LEYBOURNE, S. E NEWBOLD, P. (2002a). *Unit root tests with a break in innovation variance*. Journal of Econometrics, 109(2): 365–387.
- (2002b). *Unit root tests with a break in innovation variance*. Journal of Econometrics, 109(2): 365–387.
- KRISHNAKUMAR, J. E NETO, D. (2009). *Estimation and Testing for the Cointegration Rank in a Threshold Cointegrated System*. Research Papers by the Institute of Economics and Econometrics, Geneva School of Economics and Management,

University of Geneva 2009.01, Institut d'Economie et Econométrie, Université de Genève.

KRISTENSEN, D. E RAHBEK, A. (2010). *Testing and Inference in Nonlinear Cointegrating Vector Error Correction Models*. CREATES Research Papers 2010-68, Department of Economics and Business Economics, Aarhus University.

— (2013). *Testing And Inference In Nonlinear Cointegrating Vector Error Correction Models*. *Econometric Theory*, 29(06): 1238–1288.

KWIATKOWSKI, D., PHILLIPS, P. C. B., SCHMIDT, P. E SHIN, Y. (1992). *Testing the null hypothesis of stationarity against the alternative of a unit root : How sure are we that economic time series have a unit root?* *Journal of Econometrics*, 54(1-3): 159–178.

KÜNSCH, H. (1989). *The jackknife and the bootstrap for general stationary observations*. *The Annals of Statistics*, 17(3): 1217–1241.

LABONTE, M. (2016). *Monetary Policy and the Federal Reserve: Current Policy and Conditions*. Rap. tecn., Congressional Research Service.

LAHIRI, S. (1999). *Theoretical comparisons of block bootstrap methods*. *Annals of Statistics*, 27: 386–404.

— (2003). *Resampling Methods for Dependent Data*. Springer, New York.

LI, H. E MADDALA, G. S. (1996). *Bootstrapping Cointegrating Regressions*. *Journal of Econometrics*, 80: 297–318.

LIU, R. Y. (1988). *Bootstrap procedures under some non-I.I.D. models*. *Annals of Statistics*, 16: 1696–17081.

LO, M. C. E ZIVOT, E. (2001). *Threshold Cointegration And Nonlinear Adjustment To The Law Of One Price*. *Macroeconomic Dynamics*, 5(04): 533–576.

LONGSTAFF, F. A. (2000). *The term structure of very short-term rates: New evidence for the expectations hypothesis*. *Journal of Financial Economics*, 58(3): 397–415.

- LUCCHETTI, R. E PALOMBA, G. (2009). *Nonlinear adjustment in US bond yields: An empirical model with conditional heteroskedasticity*. *Economic Modelling*, 26(3): 659–667.
- MACAULAY, F. R. (1938). *Some theoretical problems suggested by movements of interest rates bond yields and stock prices in the united states since 1856*. *The European Journal of Finance*.
- MACKINNON, J. G. (2006). *Bootstrap Methods in Econometrics*. Rap. tecn., Queen’s University, Department of Economics.
- MAGAZZINO, C. (2010). *The economic policy of Ronald Reagan*. Mpra paper, University Library of Munich, Germany.
- MAGNUS, J. E NEUDECKER, H. (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley & Sons.
- MAMMEN, E. (1993). *Bootstrap and wild bootstrap for high dimensional linear models*. *Annals of Statistics*, 21: 255–285.
- MANKIW, N. G. E MIRON, J. A. (1986). *The Changing Behavior of the Term Structure of Interest Rates*. *The Quarterly Journal of Economics*, 101(2): 211–28.
- MANKIW, N. G. E SUMMERS, L. H. (1984). *Do Long-Term Interest Rates Overreact to Short-Term Interest Rates?* *Brookings Papers on Economic Activity*, 15(1): 223–248.
- MARCELLINO, M. (2002). *Instability and Non-Linearity in the EMU*. CEPR Discussion Papers 3312, C.E.P.R. Discussion Papers.
- MCCALLUM, B. T. (1994). *Monetary Policy and the Term Structure of Interest Rates*. Nber working papers, National Bureau of Economic Research, Inc.
- (2005). *Monetary policy and the term structure of interest rates*. *Economic Quarterly*, (Fall): 1–21.
- MEISELMAN, D. (1962). *The term structure of interest rates*. Prentice-Hall Inc.

- MEYN, S. E TWEEEDIE, R. (1993). *Markov Chains and Stochastic Stability*. Springer-Verlag, London.
- MISHKIN, F. S. (1990). *The Information in the Longer Maturity Term Structure about Future Inflation*. *The Quarterly Journal of Economics*, 105(3): 815–28.
- NEDELJKOVIC, M. (2008). *Testing for Smooth Transition Nonlinearity in Adjustments of Cointegrating Systems*. The Warwick Economics Research Paper Series (TWERPS) 876, University of Warwick, Department of Economics.
- NEELY, C. J. (2004). *The Federal Reserve responds to crises: September 11th was not the first*. *Review*, 1(Mar): 27–42.
- NELSON, C. R. E SIEGEL, A. F. (1987). *Parsimonious Modeling of Yield Curves*. *The Journal of Business*, 60(4): 473–89.
- NEWBY, W. K. E WEST, K. D. (1987). *A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix*. *Econometrica*, 55(3): 703–08.
- PALM, F. C., SMEEKES, S. E URBAIN, J.-P. (2010). *A Sieve Bootstrap Test For Cointegration In A Conditional Error Correction Model*. *Econometric Theory*, 26(03): 647–681.
- PAPARODITIS, E. E POLITIS, D. N. (2005). *Bootstrapping Unit Root Tests for Autoregressive Time Series*. *Journal of the American Statistical Association*, 100: 545–553.
- PARK, J. Y. (2003). *Bootstrap Unit Root Tests*. *Econometrica*, 71(6): 1845–1895.
- PARK, J. Y. E PHILLIPS, P. C. B. (2001). *Nonlinear Regressions with Integrated Time Series*. *Econometrica*, 69(1): 117–161.
- PARKER, C., PAPARODITIS, E. E POLITIS, D. N. (2006). *Unit root testing via the stationary bootstrap*. *Journal of Econometrics*, 133(2): 601–638.
- PHILLIPS, P. E PERRON, P. (1988). *Testing for a unit root in time series regression*. *Biometrika*, 75: 335–346.

- PHILLIPS, P. C. (2010). *Bootstrapping I(1) data*. Journal of Econometrics, 158(2): 280–284.
- PHILLIPS, P. C. E HANSEN, B. E. (1988). *Estimation and Inference in Models of Cointegration: A Simulation Study*. Cowles foundation discussion papers, Cowles Foundation for Research in Economics, Yale University.
- PIAZZESI, M. E COCHRANE, J. (2009). *Decomposing the Yield Curve*. 2009 meeting papers, Society for Economic Dynamics.
- POLITIS, D. E ROMANO, J. (1993). *The stationary bootstrap*. Journal of the American Statistical Association, 89: 1303–1313.
- SCHWARZ, G. E. (1978). *Estimating the dimension of a model*. Annals of Statistics, 6: 461–464.
- SEO, B. (2003). *Nonlinear mean reversion in the term structure of interest rates*. Journal of Economic Dynamics and Control, 27(11): 2243–2265.
- SEO, M. (2006). *Bootstrap testing for the null of no cointegration in a threshold vector error correction model*. Journal of Econometrics, 134(1): 129–150.
- SHAO, X. (2010). *The Dependent Wild Bootstrap*. Journal of the American Statistical Association, 105(489): 218–235.
- (2011). *A bootstrap-assisted spectral test of white noise under unknown dependence*. Journal of Econometrics, 162: 213–224.
- SINGH, K. (1981). *On the Asymptotic Accuracy of Efron's Bootstrap*. Ann. Statist., 9: 1187–1195.
- SKALIN, J. E TERASVIRTA, T. (1999). *Another Look at Swedish Business Cycles, 1861-1988*. Journal of Applied Econometrics, 14(4): 359–78.
- SMEEKES, S. E URBAIN, J. (2015). *A multivariate invariance principle for modified wild bootstrap methods with an application to unit root testing*. Research memorandum, Maastricht University, Graduate School of Business and Economics (GSBE).

- STOCK, J. E. & WATSON, M. (1988). *Testing for common trends*. *Journal of the American Statistical Association*, 83: 1097–1107.
- STOCK, J. H. & WATSON, M. W. (1998). *A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series*. NBER Working Papers 6607, National Bureau of Economic Research, Inc.
- SWANEPOEL, J. E. & VAN WYK, J. (1986). *The bootstrap applied to power spectral density function estimation*. *Biometrika*, 73: 135–141.
- THORNTON, D. L. (2012). *The Federal Reserve's response to the financial crisis: what it did and what it should have done*. Working Papers 2012-050, Federal Reserve Bank of St. Louis.
- TZAVALIS, E. & WICKENS, M. (1998). *A Re-examination of the Rational Expectations Hypothesis of the Term Structure: Reconciling the Evidence from Long-Run and Short-Run Tests*. *International Journal of Finance & Economics*, 3(3): 229–39.
- VALENTE, G., TAYLOR, M., SARNO, L. & CLARIDA, R. (2004). *The Role of Asymmetries and Regime Shifts in the Term Structure of Interest Rates*. Working papers, Warwick Business School, Finance Group.
- WU, C. F. J. (1986). *Jackknife, bootstrap and other resampling methods in regression analysis*. *Annals of Statistics*, 14: 1261–1295.
- XU, K.-L. & PHILLIPS, P. C. (2008). *Adaptive estimation of autoregressive models with time-varying variances*. *Journal of Econometrics*, 142(1): 265–280.
- YANG, Z., TIAN, Z. & YUAN, Z. (2007). *GSA-based maximum likelihood estimation for threshold vector error correction model*. *Computational Statistics & Data Analysis*, 52(1): 109–120.
- ÁLVARO, E. & SANTIAGO, M. (1996). *Nonlinear cointegration and nonlinear error correction*. Des - working papers. statistics and econometrics. ws, Universidad Carlos III de Madrid. Departamento de Estadística.

# Allegati

## Modello univariato

OLS, usando le osservazioni 1982/10/12–2015/02/24 ( $T = 1690$ )

**Variabile dipendente:**  $\Delta r_t^{3mesi}$

Errori standard HAC, larghezza di banda 8 (Kernel di Bartlett)

	Coefficiente	Errore Std.	rapporto $t$	p-value
const	-0,00502335	0,00670323	-0,7494	0,4537
$d_{1987}$	0,320047	0,0200910	15,9299	0,0000
$d_{2001}$	-0,471269	0,00733106	-64,2840	0,0000
$d_{2008}$	-1,00098	0,00667837	-149,8831	0,0000
$\Delta r_{t-1}^{3mesi}$	0,00448663	0,0533506	0,0841	0,9330
$\Delta r_{t-1}^{2anni}$	0,260097	0,0547447	4,7511	0,0000
$\Delta r_{t-1}^{10anni}$	-0,118900	0,0540456	-2,2000	0,0279
$\Delta r_{t-2}^{3mesi}$	-0,0762430	0,0497398	-1,5328	0,1255
$\Delta r_{t-3}^{3mesi}$	0,0470274	0,0393331	1,1956	0,2320
$sm_{t-1}$	0,138642	0,0330259	4,1980	0,0000
$sl_{t-1}$	-0,0617574	0,0236591	-2,6103	0,0091
$sm_{t-1}^2$	-0,163771	0,0590012	-2,7757	0,0056
$sl_{t-1}^2$	0,0200399	0,0172972	1,1586	0,2468
$smsl_{t-1}$	0,0443394	0,0462070	0,9596	0,3374
$sm_{t-1}^3$	0,00952470	0,0221162	0,4307	0,6668
$sl_{t-1}^3$	-9,66470e-005	0,00432521	-0,0223	0,9822
$sm^2 sl_{t-1}$	0,0435488	0,0391452	1,1125	0,2661
$smsl_{t-1}^2$	-0,0257394	0,0204716	-1,2573	0,2088

Media var. dipendente	-0,004586	SQM var. dipendente	0,114284
Somma quadr. residui	18,89644	E.S. della regressione	0,106309
$R^2$	0,143398	$R^2$ corretto	0,134688
Log-verosimiglianza	1399,010	Criterio di Akaike	-2762,020
Criterio di Schwarz	-2664,235	Hannan-Quinn	-2725,809
$\hat{\rho}$	0,033885	Durbin-Watson	1,931131

Test per i coefficienti dello spread non lineare –

Ipotesi nulla: i parametri valgono zero per le variabili

$$sm_{t-1}^2 \quad sl_{t-1}^2 \quad smsl_{t-1} \quad sm_{t-1}^3 \quad sl_{t-1}^3 \quad sm^2 sl_{t-1} \quad smsl_{t-1}^2$$

Statistica test:  $F(7, 1672) = 4,10663$

con p-value =  $P(F(7, 1672) > 4,10663) = 0,000175126$

Test RESET di specificazione –

Ipotesi nulla: la specificazione è adeguata

Statistica test:  $F(2, 1670) = 2,87859$

11 con p-value =  $P(F(2, 1670) > 2,87859) = 0,0564932$

Test LM per l'autocorrelazione fino all'ordine 1 –

Ipotesi nulla: Non c'è autocorrelazione

Statistica test: LMF = 20,9691

con p-value =  $P(F(1, 1671) > 20,9691) = 5,01269e-006$

Test per ARCH di ordine 2 –

Ipotesi nulla: non sono presenti effetti ARCH

Statistica test: LM = 42,2524

con p-value =  $P(\chi^2(2) > 42,2524) = 6,68369e-010$

Test CUSUM per la stabilità dei parametri –

Ipotesi nulla: nessun cambiamento nei parametri

Statistica test: Harvey-Collier  $t(1671) = 3,32358$

con p-value =  $P(t_{1671} > 3,32358) = 0,0009079$

OLS, usando le osservazioni 1982/09/28–2015/02/24 ( $T = 1692$ )

**Variabile dipendente:**  $\Delta r_t^{2anni}$

Errori standard HAC, larghezza di banda 8 (Kernel di Bartlett)

	Coefficiente	Errore Std.	rapporto $t$	p-value
const	-0,00678528	0,00699986	-0,9693	0,3325
$d_{1987}$	0,229114	0,0197685	11,5898	0,0000
$d_{2001}$	-0,577269	0,00639444	-90,2768	0,0000
$d_{2008}$	-0,388363	0,00682354	-56,9151	0,0000
$\Delta r_{t-1}^{3mesi}$	-0,104317	0,0488962	-2,1334	0,0330
$\Delta r_{t-1}^{2anni}$	0,330690	0,0563601	5,8674	0,0000
$\Delta r_{t-1}^{10anni}$	-0,0319767	0,0485204	-0,6590	0,5100
$\Delta r_{t-2}^{3mesi}$	0,0318445	0,0368060	0,8652	0,3871
$\Delta r_{t-3}^{3mesi}$	0,0857877	0,0347697	2,4673	0,0137
$sm_{t-1}$	0,0173749	0,0348758	0,4982	0,6184
$sl_{t-1}$	0,00566425	0,0251376	0,2253	0,8218
$sm_{t-1}^2$	-0,129984	0,0570387	-2,2789	0,0228
$sl_{t-1}^2$	-0,0203077	0,0192136	-1,0569	0,2907
$smsl_{t-1}$	0,105972	0,0475039	2,2308	0,0258
$sm_{t-1}^3$	-0,0597615	0,0253599	-2,3565	0,0186
$sl_{t-1}^3$	0,00931452	0,00499458	1,8649	0,0624
$sm^2sl_{t-1}$	0,113898	0,0464450	2,4523	0,0143
$smsl_{t-1}^2$	-0,0633252	0,0241592	-2,6212	0,0088
Media var. dipendente	-0,006702	SQM var. dipendente	0,121884	
Somma quadr. residui	21,65021	E.S. della regressione	0,113724	
$R^2$	0,138170	$R^2$ corretto	0,129418	
Log-verosimiglianza	1286,575	Criterio di Akaike	-2537,150	
Criterio di Schwarz	-2439,344	Hannan–Quinn	-2500,934	
$\hat{\rho}$	0,016187	Durbin–Watson	1,966372	

Test per i coefficienti dello spread non lineare –

Ipotesi nulla: i parametri valgono zero per le variabili

$$sm^2 sl^2 sm sl sm^3 sl^3 sm^2 sl sm sl^2$$

Statistica test:  $F(7, 1674) = 5,31121$

con p-value =  $P(F(7, 1674) > 5,31121) = 5,05943e-006$

Test RESET di specificazione –

Ipotesi nulla: la specificazione è adeguata

Statistica test:  $F(2, 1672) = 1,12334$

con p-value =  $P(F(2, 1672) > 1,12334) = 0,325436$

Test LM per l'autocorrelazione fino all'ordine 1 –

Ipotesi nulla: Non c'è autocorrelazione

Statistica test: LMF = 4,28518

con p-value =  $P(F(1, 1673) > 4,28518) = 0,0385992$

Test per ARCH di ordine 2 –

Ipotesi nulla: non sono presenti effetti ARCH

Statistica test: LM = 18,9164

con p-value =  $P(\chi^2(2) > 18,9164) = 7,80483e-005$

Test CUSUM per la stabilità dei parametri –

Ipotesi nulla: nessun cambiamento nei parametri

Statistica test: Harvey–Collier  $t(1673) = -0,45281$

con p-value =  $P(t_{1673} > -0,45281) = 0,650745$

OLS, usando le osservazioni 1982/09/28–2015/02/24 ( $T = 1692$ )

**Variabile dipendente:**  $\Delta r_t^{10anni}$

Errori standard HAC, larghezza di banda 8 (Kernel di Bartlett)

	Coefficiente	Errore Std.	rapporto $t$	p-value
const	-0,00391165	0,00595896	-0,6564	0,5116
$d_{1987}$	0,298806	0,0158815	18,8147	0,0000
$d_{2001}$	-0,250603	0,00587708	-42,6408	0,0000
$d_{2008}$	-0,129453	0,00683108	-18,9506	0,0000
$\Delta r_t^{3mesi}$	-0,106333	0,0459184	-2,3157	0,0207
$\Delta r_t^{2anni}$	0,110010	0,0569798	1,9307	0,0537
$\Delta r_t^{10anni}$	0,171243	0,0498288	3,4366	0,0006
$sm_{t-1}$	0,00249541	0,0263244	0,0948	0,9245
$sl_{t-1}$	-0,00127275	0,0210035	-0,0606	0,9517
$sm_{t-1}^2$	-0,0673762	0,0447575	-1,5054	0,1324
$sl_{t-1}^2$	-0,00308389	0,0180978	-0,1704	0,8647
$smsl_{t-1}$	0,0616203	0,0429672	1,4341	0,1517
$sm_{t-1}^3$	-0,0403579	0,0225218	-1,7919	0,0733
$sl_{t-1}^3$	0,00277815	0,00479172	0,5798	0,5621
$sm^2 sl_{t-1}$	0,0669016	0,0407833	1,6404	0,1011
$smsl_{t-1}^2$	-0,0361678	0,0218255	-1,6571	0,0977
Media var. dipendente	-0,006188	SQM var. dipendente	0,121759	
Somma quadr. residui	22,59912	E.S. della regressione	0,116120	
$R^2$	0,098542	$R^2$ corretto	0,090474	
Log-verosimiglianza	1250,285	Criterio di Akaike	-2468,570	
Criterio di Schwarz	-2381,632	Hannan–Quinn	-2436,378	
$\hat{\rho}$	0,012749	Durbin–Watson	1,973687	

Test per i coefficienti dello spread non lineare –

Ipotesi nulla: i parametri valgono zero per le variabili

$$sm_{t-1}^2 \quad sl_{t-1}^2 \quad smsl_{t-1} \quad sm_{t-1}^3 \quad sl_{t-1}^3 \quad sm^2 sl_{t-1} \quad smsl_{t-1}^2$$

Statistica test:  $F(7, 1676) = 12,3045$

con p-value =  $P(F(7, 1676) > 12,3045) = 1,98782e-015$

Test RESET di specificazione –

Ipotesi nulla: la specificazione è adeguata

Statistica test:  $F(2, 1674) = 0,92696$

con p-value =  $P(F(2, 1674) > 0,92696) = 0,395958$

Test LM per l'autocorrelazione fino all'ordine 1 –

Ipotesi nulla: Non c'è autocorrelazione

Statistica test: LMF = 3,71995

con p-value =  $P(F(1, 1675) > 3,71995) = 0,0539347$

Test per ARCH di ordine 2 –

Ipotesi nulla: non sono presenti effetti ARCH

Statistica test: LM = 56,1922

con p-value =  $P(\chi^2(2) > 56,1922) = 6,28092e-013$

Test CUSUM per la stabilità dei parametri –

Ipotesi nulla: nessun cambiamento nei parametri

Statistica test: Harvey–Collier  $t(1675) = -0,105641$

con p-value =  $P(t_{1675} > -0,105641) = 0,91588$

OLS, usando le osservazioni 1982:10–2015:02 ( $T = 389$ )

**Variabile dipendente:**  $\Delta r_t^{3mesi}$

Errori standard HAC, larghezza di banda 5 (Kernel di Bartlett)

	Coefficiente	Errore Std.	rapporto $t$	p-value
const	-0,0625288	0,0356304	-1,7549	0,0801
$d_{1987}$	-1,38187	0,0426136	-32,4279	0,0000
$d_{2001}$	-0,216011	0,0334019	-6,4670	0,0000
$d_{2008}$	-0,845420	0,0454274	-18,6104	0,0000
$\Delta r_t^{3mesi}$	0,00710354	0,105817	0,0671	0,9465
$\Delta r_t^{2anni}$	0,140372	0,122667	1,1443	0,2532
$\Delta r_t^{10anni}$	-0,0100562	0,0948677	-0,1060	0,9156
$sm_{t-1}$	0,526757	0,181278	2,9058	0,0039
$sl_{t-1}$	-0,233410	0,125182	-1,8646	0,0630
$sm_{t-1}^2$	-0,436427	0,243842	-1,7898	0,0743
$sl_{t-1}^2$	0,0877358	0,0855677	1,0253	0,3059
$smsl_{t-1}$	0,0902711	0,241803	0,3733	0,7091
$sm_{t-1}^3$	-0,0742352	0,108355	-0,6851	0,4937
$sl_{t-1}^3$	0,00192408	0,0214360	0,0898	0,9285
$sm^2sl_{t-1}$	0,241971	0,162534	1,4887	0,1374
$smsl_{t-1}^2$	-0,119708	0,0862025	-1,3887	0,1658
Media var. dipendente	-0,019280	SQM var. dipendente	0,282412	
Somma quadr. residui	25,01647	E.S. della regressione	0,258975	
$R^2$	0,191598	$R^2$ corretto	0,159089	
Log-verosimiglianza	-18,25035	Criterio di Akaike	68,50070	
Criterio di Schwarz	131,9180	Hannan–Quinn	93,64217	
$\hat{\rho}$	-0,006358	Durbin–Watson	2,011676	

Test per i coefficienti dello spread non lineare –

Ipotesi nulla: i parametri valgono zero per le variabili

$$sm_{t-1}^2 \quad sl_{t-1}^2 \quad smsl_{t-1} \quad sm_{t-1}^3 \quad sl_{t-1}^3 \quad sm^2 sl_{t-1} \quad smsl_{t-1}^2$$

Statistica test:  $F(7, 373) = 1,85696$

con p-value =  $P(F(7, 373) > 1,85696) = 0,0755292$

Test RESET di specificazione –

Ipotesi nulla: la specificazione è adeguata

Statistica test:  $F(2, 371) = 5,75243$

con p-value =  $P(F(2, 371) > 5,75243) = 0,00346499$

Test LM per l'autocorrelazione fino all'ordine 2 –

Ipotesi nulla: Non c'è autocorrelazione

Statistica test: LMF = 0,0661623

con p-value =  $P(F(2, 371) > 0,0661623) = 0,93599$

Test per ARCH di ordine 2 –

Ipotesi nulla: non sono presenti effetti ARCH

Statistica test: LM = 24,1245

con p-value =  $P(\chi^2(2) > 24,1245) = 5,77341e-006$

Test CUSUM per la stabilità dei parametri –

Ipotesi nulla: nessun cambiamento nei parametri

Statistica test: Harvey–Collier  $t(372) = 1,60233$

con p-value =  $P(t_{372} > 1,60233) = 0,109932$

OLS, usando le osservazioni 1982:10–2015:02 ( $T = 389$ )

**Variabile dipendente:**  $\Delta r_t^{2anni}$

Errori standard HAC, larghezza di banda 5 (Kernel di Bartlett)

	Coefficiente	Errore Std.	rapporto $t$	p-value
const	-0,0345223	0,0336685	-1,0254	0,3059
$d_{1987}$	-0,731439	0,0578194	-12,6504	0,0000
$d_{2001}$	0,409480	0,0325905	12,5644	0,0000
$d_{2008}$	-0,249532	0,0447656	-5,5742	0,0000
$\Delta r_t^{3mesi}$	0,0579451	0,0854719	0,6779	0,4982
$\Delta r_t^{2anni}$	0,167706	0,109313	1,5342	0,1258
$\Delta r_t^{10anni}$	0,0860205	0,0938669	0,9164	0,3600
$sm_{t-1}$	-0,0387911	0,213757	-0,1815	0,8561
$sl_{t-1}$	0,0561976	0,137042	0,4101	0,6820
$sm_{t-1}^2$	-0,422750	0,175266	-2,4120	0,0163
$sl_{t-1}^2$	-0,0875685	0,107451	-0,8150	0,4156
$smsl_{t-1}$	0,419761	0,218329	1,9226	0,0553
$sm_{t-1}^3$	-0,219128	0,0999964	-2,1914	0,0290
$sl_{t-1}^3$	0,0355267	0,0246625	1,4405	0,1506
$sm^2 sl_{t-1}$	0,408320	0,168085	2,4292	0,0156
$smsl_{t-1}^2$	-0,236220	0,0899181	-2,6271	0,0090
Media var. dipendente	-0,027918	SQM var. dipendente	0,312358	
Somma quadr. residui	31,18846	E.S. della regressione	0,289163	
$R^2$	0,176134	$R^2$ corretto	0,143002	
Log-verosimiglianza	-61,14029	Criterio di Akaike	154,2806	
Criterio di Schwarz	217,6978	Hannan–Quinn	179,4220	
$\hat{\rho}$	-0,003329	Durbin–Watson	2,005482	

Test per i coefficienti dello spread non lineare –

Ipotesi nulla: i parametri valgono zero per le variabili

$$sm_{t-1}^2 \quad sl_{t-1}^2 \quad smsl_{t-1} \quad sm_{t-1}^3 \quad sl_{t-1}^3 \quad sm^2 sl_{t-1} \quad smsl_{t-1}^2$$

Statistica test:  $F(7, 373) = 13,5944$

con p-value =  $P(F(7, 373) > 13,5944) = 1,13045e-015$

Test RESET di specificazione –

Ipotesi nulla: la specificazione è adeguata

Statistica test:  $F(2, 371) = 0,694842$

con p-value =  $P(F(2, 371) > 0,694842) = 0,499802$

Test LM per l'autocorrelazione fino all'ordine 1 –

Ipotesi nulla: Non c'è autocorrelazione

Statistica test: LMF = 0,0280546

con p-value =  $P(F(1, 372) > 0,0280546) = 0,867071$

Test per ARCH di ordine 2 –

Ipotesi nulla: non sono presenti effetti ARCH

Statistica test: LM = 4,72214

con p-value =  $P(\chi^2(2) > 4,72214) = 0,0943191$

Test CUSUM per la stabilità dei parametri –

Ipotesi nulla: nessun cambiamento nei parametri

Statistica test: Harvey–Collier  $t(372) = -1,4601$

con p-value =  $P(t_{372} > -1,4601) = 0,145106$

OLS, usando le osservazioni 1982:10–2015:02 ( $T = 389$ )

**Variabile dipendente:**  $\Delta r_t^{10anni}$

Errori standard HAC, larghezza di banda 5 (Kernel di Bartlett)

	Coefficiente	Errore Std.	rapporto $t$	p-value
const	0,00258823	0,0280868	0,0922	0,9266
$d_{1987}$	-0,484677	0,0544674	-8,8985	0,0000
$d_{2001}$	0,384210	0,0323614	11,8725	0,0000
$d_{2008}$	0,0168574	0,0419133	0,4022	0,6878
$\Delta r_t^{3mesi}$	-0,0319336	0,0717093	-0,4453	0,6563
$\Delta r_t^{2anni}$	0,141814	0,0978923	1,4487	0,1483
$\Delta r_t^{10anni}$	0,0309667	0,0905778	0,3419	0,7326
$sm_{t-1}$	-0,0988209	0,157673	-0,6267	0,5312
$sl_{t-1}$	0,0572535	0,108934	0,5256	0,5995
$sm_{t-1}^2$	-0,345457	0,203752	-1,6955	0,0908
$sl_{t-1}^2$	-0,0542581	0,0960471	-0,5649	0,5725
$smsl_{t-1}$	0,319549	0,223159	1,4319	0,1530
$sm_{t-1}^3$	-0,156920	0,104089	-1,5076	0,1325
$sl_{t-1}^3$	0,0171697	0,0253588	0,6771	0,4988
$sm^2sl_{t-1}$	0,295794	0,190315	1,5542	0,1210
$smsl_{t-1}^2$	-0,158989	0,105958	-1,5005	0,1343
Media var. dipendente	-0,025398	SQM var. dipendente	0,300121	
Somma quadr. residui	31,54197	E.S. della regressione	0,290797	
$R^2$	0,097467	$R^2$ corretto	0,061172	
Log-verosimiglianza	-63,33244	Criterio di Akaike	158,6649	
Criterio di Schwarz	222,0822	Hannan–Quinn	183,8064	
$\hat{\rho}$	-0,004752	Durbin–Watson	2,007415	

Test per i coefficienti dello spread non lineare –

Ipotesi nulla: i parametri valgono zero per le variabili

$$sm_{t-1}^2 \quad sl_{t-1}^2 \quad smsl_{t-1} \quad sm_{t-1}^3 \quad sl_{t-1}^3 \quad sm^2 sl_{t-1} \quad smsl_{t-1}^2$$

Statistica test:  $F(7, 373) = 8,8802$

con p-value =  $P(F(7, 373) > 8,8802) = 4,00704e-010$

Test RESET di specificazione –

Ipotesi nulla: la specificazione è adeguata

Statistica test:  $F(2, 371) = 0,180935$

con p-value =  $P(F(2, 371) > 0,180935) = 0,834563$

Test LM per l'autocorrelazione fino all'ordine 1 –

Ipotesi nulla: Non c'è autocorrelazione

Statistica test: LMF = 0,142309

con p-value =  $P(F(1, 372) > 0,142309) = 0,706212$

Test per ARCH di ordine 2 –

Ipotesi nulla: non sono presenti effetti ARCH

Statistica test: LM = 1,91581

con p-value =  $P(\chi^2(2) > 1,91581) = 0,383697$

Test CUSUM per la stabilità dei parametri –

Ipotesi nulla: nessun cambiamento nei parametri

Statistica test: Harvey–Collier  $t(372) = -0,83656$

con p-value =  $P(t_{372} > -0,83656) = 0,403377$

## Modello multivariato

Sistema VAR, ordine ritardi 2

Stime OLS usando le osservazioni 1982/10/12–2015/02/24 ( $T = 1690$ )

Log-verosimiglianza = 5376,97

Determinante della matrice di covarianza = 3,45971e-007

AIC = -6,2958

BIC = -6,1126

HQC = -6,2280

Test portmanteau: LB(48) = 694,195, df = 414 [0,0000]

**Equazione 1:**  $\Delta r_t^{3mesi}$

Errori standard robusti rispetto all'eteroschedasticità, variante HC1

	Coefficiente	Errore Std.	rapporto $t$	p-value
const	-0,00528000	0,00639085	-0,8262	0,4088
$\Delta r_{t-1}^{3mesi}$	-0,00355153	0,0642132	-0,0553	0,9559
$\Delta r_{t-2}^{3mesi}$	-0,0571938	0,0502468	-1,1383	0,2552
$\Delta r_{t-1}^{2anni}$	0,283246	0,0667949	4,2405	0,0000
$\Delta r_{t-2}^{2anni}$	-0,0522477	0,0687367	-0,7601	0,4473
$\Delta r_{t-1}^{10anni}$	-0,135435	0,0552936	-2,4494	0,0144
$\Delta r_{t-2}^{10anni}$	0,0444725	0,0491374	0,9051	0,3656
$d_{1987}$	0,324562	0,0184741	17,5685	0,0000
$d_{2001}$	-0,470029	0,00671917	-69,9535	0,0000
$d_{2008}$	-1,00321	0,00850741	-117,9220	0,0000
$sm_{t-1}$	0,146080	0,0360143	4,0562	0,0001
$sl_{t-1}$	-0,0661333	0,0245575	-2,6930	0,0072
$sm^2_{t-1}$	-0,164390	0,0662019	-2,4832	0,0131
$sl^2_{t-1}$	0,0226776	0,0174891	1,2967	0,1949
$smsl_{t-1}$	0,0411172	0,0518051	0,7937	0,4275
$sm^3_{t-1}$	0,00929213	0,0280137	0,3317	0,7402
$sl^3_{t-1}$	-0,000467027	0,00457712	-0,1020	0,9187
$sm^2sl_{t-1}$	0,0441999	0,0453801	0,9740	0,3302
$smsl^2_{t-1}$	-0,0256874	0,0234800	-1,0940	0,2741
Media var. dipendente	-0,004586	SQM var. dipendente	0,114284	
Somma quadr. residui	18,92850	E.S. della regressione	0,106431	
$R^2$	0,141944	$R^2$ corretto	0,132701	
$\hat{\rho}$	0,036555	Durbin-Watson	1,925796	

Test F per zero vincoli

Tutti i ritardi di $\Delta r_t^{3mesi}$	$F(2, 1671) = 0,64806$	[0,5232]
Tutti i ritardi di $\Delta r_t^{2anni}$	$F(2, 1671) = 8,99497$	[0,0001]
Tutti i ritardi di $\Delta r_t^{10anni}$	$F(2, 1671) = 3,01817$	[0,0492]
Tutte le variabili, ritardo 2	$F(3, 1671) = 1,05253$	[0,3683]

**Equazione 2:**  $\Delta r_t^{2anni}$

Errori standard robusti rispetto all'eteroschedasticità, variante HC1

	Coefficiente	Errore Std.	rapporto $t$	p-value
const	-0,00710073	0,00641917	-1,1062	0,2688
$\Delta r_{t-1}^{3mesi}$	-0,120067	0,0568633	-2,1115	0,0349
$\Delta r_{t-2}^{3mesi}$	0,0710020	0,0458383	1,5490	0,1216
$\Delta r_{t-1}^{2anni}$	0,376378	0,0689070	5,4621	0,0000
$\Delta r_{t-2}^{2anni}$	-0,0989465	0,0600243	-1,6484	0,0995
$\Delta r_{t-1}^{10anni}$	-0,0618660	0,0506628	-1,2211	0,2222
$\Delta r_{t-2}^{10anni}$	0,0810828	0,0489591	1,6561	0,0979
$d_{1987}$	0,235762	0,0182566	12,9138	0,0000
$d_{2001}$	-0,575222	0,00648009	-88,7677	0,0000
$d_{2008}$	-0,392394	0,00787756	-49,8117	0,0000
$sm_{t-1}$	0,0311498	0,0326687	0,9535	0,3405
$sl_{t-1}$	-0,00210949	0,0229307	-0,0920	0,9267
$sm^2_{t-1}$	-0,131384	0,0497431	-2,6413	0,0083
$sl^2_{t-1}$	-0,0157850	0,0189263	-0,8340	0,4044
$smsl_{t-1}$	0,0996090	0,0464445	2,1447	0,0321
$sm^3_{t-1}$	-0,0623126	0,0310153	-2,0091	0,0447
$sl^3_{t-1}$	0,00884132	0,00509521	1,7352	0,0829
$sm^2sl_{t-1}$	0,118242	0,0489480	2,4157	0,0158
$smsl^2_{t-1}$	-0,0643325	0,0242031	-2,6580	0,0079
Media var. dipendente	-0,006331	SQM var. dipendente		0,121478
Somma quadr. residui	21,73210	E.S. della regressione		0,114041
$R^2$	0,128081	$R^2$ corretto		0,118689
$\hat{\rho}$	0,003958	Durbin-Watson		1,991937

Test F per zero vincoli

Tutti i ritardi di $\Delta r_t^{3mesi}$	$F(2, 1671) = 3,26137$	[0,0386]
Tutti i ritardi di $\Delta r_t^{2anni}$	$F(2, 1671) = 16,0606$	[0,0000]
Tutti i ritardi di $\Delta r_t^{10anni}$	$F(2, 1671) = 1,84956$	[0,1576]
Tutte le variabili, ritardo 2	$F(3, 1671) = 1,50093$	[0,2125]

**Equazione 3:**  $\Delta r_t^{10anni}$

Errori standard robusti rispetto all'eteroschedasticità, variante HC1

	Coefficiente	Errore Std.	rapporto $t$	p-value
const	-0,00446371	0,00565378	-0,7895	0,4299
$\Delta r_{t-1}^{3mesi}$	-0,104406	0,0479109	-2,1792	0,0295
$\Delta r_{t-2}^{3mesi}$	0,0149354	0,0381333	0,3917	0,6954
$\Delta r_{t-1}^{2anni}$	0,121677	0,0654587	1,8588	0,0632
$\Delta r_{t-2}^{2anni}$	-0,0575228	0,0640970	-0,8974	0,3696
$\Delta r_{t-1}^{10anni}$	0,164785	0,0555138	2,9684	0,0030
$\Delta r_{t-2}^{10anni}$	0,0378261	0,0567121	0,6670	0,5049
$d_{1987}$	0,300482	0,0183205	16,4014	0,0000
$d_{2001}$	-0,250208	0,00683204	-36,6228	0,0000
$d_{2008}$	-0,130498	0,00778139	-16,7705	0,0000
$sm_{t-1}$	0,00496871	0,0273680	0,1816	0,8560
$sl_{t-1}$	-0,00257433	0,0203923	-0,1262	0,8996
$sm^2_{t-1}$	-0,0620764	0,0450368	-1,3783	0,1683
$sl^2_{t-1}$	-0,00167861	0,0184714	-0,0909	0,9276
$smsl_{t-1}$	0,0567233	0,0453415	1,2510	0,2111
$sm^3_{t-1}$	-0,0436080	0,0250713	-1,7394	0,0822
$sl^3_{t-1}$	0,00258492	0,00510773	0,5061	0,6129
$sm^2sl_{t-1}$	0,0689931	0,0425451	1,6216	0,1051
$smsl^2_{t-1}$	-0,0360839	0,0226616	-1,5923	0,1115
Media var. dipendente	-0,005722	SQM var. dipendente		0,121066
Somma quadr. residui	22,54471	E.S. della regressione		0,116154
$R^2$	0,089315	$R^2$ corretto		0,079505
$\hat{\rho}$	0,008815	Durbin-Watson		1,982179

Test F per zero vincoli

Tutti i ritardi di $\Delta r_t^{3mesi}$	$F(2, 1671) = 2,39718$	[0,0913]
Tutti i ritardi di $\Delta r_t^{2anni}$	$F(2, 1671) = 2,0888$	[0,1242]
Tutti i ritardi di $\Delta r_t^{10anni}$	$F(2, 1671) = 4,98732$	[0,0069]
Tutte le variabili, ritardo 2	$F(3, 1671) = 0,276908$	[0,8421]

Per il sistema nel complesso —

Ipotesi nulla: il ritardo maggiore è 1

Ipotesi alternativa: il ritardo maggiore è 2

Test del rapporto di verosimiglianza:  $\chi_9^2 = 53,798$  [0,0000]

Test sui coefficienti dello spread non lineare —

Ipotesi nulla: i parametri valgono zero per le variabili

$$sm_{t-1}^2 \quad sl_{t-1}^2 \quad smsl_{t-1} \quad sm_{t-1}^3 \quad sl_{t-1}^3 \quad sm^2 sl_{t-1} \quad smsl_{t-1}^2$$

Test Wald: Chi-quadro(21) = 106,468

con p-value = 2,03498e-013

con bootstrap p-value = 0,02011

Test LM per l'autocorrelazione fino all'ordine 1 —

Ipotesi nulla: Non c'è autocorrelazione

equazione 1: Ljung-Box Q' = 2,26227

con p-value =  $P(\text{Chi-quadro}(1) > 2,26227) = 0,133$

equazione 2: Ljung-Box Q' = 0,0265155

con p-value =  $P(\text{Chi-quadro}(1) > 0,0265155) = 0,871$

equazione 3: Ljung-Box Q' = 0,131517

con p-value =  $P(\text{Chi-quadro}(1) > 0,131517) = 0,717$

Test per ARCH di ordine 2 —

Ipotesi nulla: non sono presenti effetti ARCH

equazione 1: Statistica test LM = 47,23

con p-value =  $P(\text{Chi-quadro}(2) > 47,23) = 5,54787e-011$

equazione 2: Statistica test LM = 25,8786

con p-value =  $P(\text{Chi-quadro}(2) > 25,8786) = 2,40181e-0061$

equazione 3: Statistica test LM = 55,5477

con p-value =  $P(\text{Chi-quadro}(2) > 55,5477) = 8,6691e-0131$

Sistema VAR, ordine ritardi 1

Stime OLS usando le osservazioni 1982:10–2015:02 ( $T = 389$ )

Log-verosimiglianza = 209,593

Determinante della matrice di covarianza = 6,83254e-005

AIC = -0,8308

BIC = -0,3417

HQC = -0,6369

Test portmanteau: LB(48) = 593,799, df = 423 [0,0000]

**Equazione 1:**  $\Delta r_t^{3mesi}$

Errori standard robusti rispetto all'eteroschedasticità, variante HC1

	Coefficiente	Errore Std.	rapporto $t$	p-value
const	-0,0625288	0,0389811	-1,6041	0,1095
$\Delta r_t^{3mesi}$	0,00710354	0,0992220	0,0716	0,9430
$\Delta r_t^{2anni}$	0,140372	0,128350	1,0937	0,2748
$\Delta r_t^{10anni}$	-0,0100562	0,0933590	-0,1077	0,9143
$d_{1987}$	-1,38187	0,0459731	-30,0583	0,0000
$d_{2001}$	-0,216011	0,0320854	-6,7324	0,0000
$d_{2008}$	-0,845420	0,0481281	-17,5660	0,0000
$sm_{t-1}$	0,526757	0,171638	3,0690	0,0023
$sl_{t-1}$	-0,233410	0,130239	-1,7922	0,0739
$sm^2_{t-1}$	-0,436427	0,286505	-1,5233	0,1285
$sl^2_{t-1}$	0,0877358	0,0978128	0,8970	0,3703
$smsl_{t-1}$	0,0902711	0,292160	0,3090	0,7575
$sm^3_{t-1}$	-0,0742352	0,112859	-0,6578	0,5111
$sl^3_{t-1}$	0,00192408	0,0247783	0,0777	0,9381
$sm^2sl_{t-1}$	0,241971	0,180901	1,3376	0,1818
$smsl^2_{t-1}$	-0,119708	0,103928	-1,1518	0,2501

Media var. dipendente	-0,019280	SQM var. dipendente	0,282412
Somma quadr. residui	25,01647	E.S. della regressione	0,258975
$R^2$	0,191598	$R^2$ corretto	0,159089
$\hat{\rho}$	-0,006358	Durbin-Watson	2,011676

Test F per zero vincoli

Tutti i ritardi di $\Delta r_t^{3mesi}$	$F(1, 373) = 0,00512547$	[0,9430]
Tutti i ritardi di $\Delta r_t^{2anni}$	$F(1, 373) = 1,19609$	[0,2748]
Tutti i ritardi di $\Delta r_t^{10anni}$	$F(1, 373) = 0,0116026$	[0,9143]

**Equazione 2:**  $\Delta r_t^{2anni}$

Errori standard robusti rispetto all'eteroschedasticità, variante HC1

	Coefficiente	Errore Std.	rapporto $t$	p-value
const	-0,0345223	0,0354765	-0,9731	0,3311
$\Delta r_t^{3mesi}$	0,0579451	0,0945882	0,6126	0,5405
$\Delta r_t^{2anni}$	0,167706	0,114699	1,4621	0,1445
$\Delta r_t^{10anni}$	0,0860205	0,0902685	0,9529	0,3412
$d_{1987}$	-0,731439	0,0540773	-13,5258	0,0000
$d_{2001}$	0,409480	0,0339124	12,0746	0,0000
$d_{2008}$	-0,249532	0,0471563	-5,2916	0,0000
$sm_{t-1}$	-0,0387911	0,185342	-0,2093	0,8343
$sl_{t-1}$	0,0561976	0,131062	0,4288	0,6683
$sm^2_{t-1}$	-0,422750	0,202685	-2,0857	0,0377
$sl^2_{t-1}$	-0,0875685	0,107288	-0,8162	0,4149
$smsl_{t-1}$	0,419761	0,238225	1,7620	0,0789
$sm^3_{t-1}$	-0,219128	0,100915	-2,1714	0,0305
$sl^3_{t-1}$	0,0355267	0,0261419	1,3590	0,1750
$sm^2sl_{t-1}$	0,408320	0,175774	2,3230	0,0207
$smsl^2_{t-1}$	-0,236220	0,101407	-2,3294	0,0204
Media var. dipendente	-0,027918	SQM var. dipendente		0,312358
Somma quadr. residui	31,18846	E.S. della regressione		0,289163
$R^2$	0,176134	$R^2$ corretto		0,143002
$F(15, 373)$	2,63e+15	P-value( $F$ )		0,000000
$\hat{\rho}$	-0,003329	Durbin-Watson		2,005482

Test F per zero vincoli

Tutti i ritardi di $\Delta r_t^{3mesi}$	$F(1, 373) = 0,375283$	[0,5405]
Tutti i ritardi di $\Delta r_t^{2anni}$	$F(1, 373) = 2,13785$	[0,1445]
Tutti i ritardi di $\Delta r_t^{10anni}$	$F(1, 373) = 0,908094$	[0,3412]

Equazione per  $\Delta r_t^{10anni}$

Errori standard robusti rispetto all'eteroschedasticità, variante HC1

	Coefficiente	Errore Std.	rapporto $t$	p-value
const	0,00258823	0,0325083	0,0796	0,9366
$\Delta r_t^{3mesi}$	-0,0319336	0,0776106	-0,4115	0,6810
$\Delta r_t^{2anni}$	0,141814	0,105394	1,3456	0,1793
$\Delta r_t^{10anni}$	0,0309667	0,0999877	0,3097	0,7570
$d_{1987}$	-0,484677	0,0524159	-9,2468	0,0000
$d_{2001}$	0,384210	0,0347438	11,0584	0,0000
$d_{2008}$	0,0168574	0,0419658	0,4017	0,6881
$sm_{t-1}$	-0,0988209	0,162678	-0,6075	0,5439
$sl_{t-1}$	0,0572535	0,123030	0,4654	0,6419
$sm^2_{t-1}$	-0,345457	0,232165	-1,4880	0,1376
$sl^2_{t-1}$	-0,0542581	0,106348	-0,5102	0,6102
$smsl_{t-1}$	0,319549	0,254367	1,2563	0,2098
$sm^3_{t-1}$	-0,156920	0,106549	-1,4728	0,1417
$sl^3_{t-1}$	0,0171697	0,0277669	0,6184	0,5367
$sm^2sl_{t-1}$	0,295794	0,195759	1,5110	0,1316
$smsl^2_{t-1}$	-0,158989	0,113876	-1,3962	0,1635
Media var. dipendente	-0,025398	SQM var. dipendente	0,300121	
Somma quadr. residui	31,54197	E.S. della regressione	0,290797	
$R^2$	0,097467	$R^2$ corretto	0,061172	
$F(15, 373)$	5,72e+14	P-value( $F$ )	0,000000	
$\hat{\rho}$	-0,004752	Durbin-Watson	2,007415	

Test F per zero vincoli

Tutti i ritardi di $\Delta r_t^{3mesi}$	$F(1, 373) = 0,169299$	[0,6810]
Tutti i ritardi di $\Delta r_t^{2anni}$	$F(1, 373) = 1,81051$	[0,1793]
Tutti i ritardi di $\Delta r_t^{10anni}$	$F(1, 373) = 0,0959172$	[0,7570]

Per il sistema nel complesso —

Test sui coefficienti dello spread non lineare

Ipotesi nulla: i parametri valgono zero per le variabili

$$sm_{t-1}^2 \quad sl_{t-1}^2 \quad smsl_{t-1} \quad sm_{t-1}^3 \quad sl_{t-1}^3 \quad sm^2 sl_{t-1} \quad smsl_{t-1}^2$$

Test Wald: Chi-quadro(21) = 64,3492

con p-value = 6,75809e-010

con bootstrap p-value = 0,00102

Test LM per l'autocorrelazione fino all'ordine 1

Ipotesi nulla: Non c'è autocorrelazione

equazione 1: Ljung-Box Q' = 0,0158416

con p-value =  $P(\text{Chi-quadro}(1) > 0,0158416) = 0,9$

equazione 2: Ljung-Box Q' = 0,00433555

con p-value =  $P(\text{Chi-quadro}(1) > 0,00433555) = 0,948$

equazione 3: Ljung-Box Q' = 0,00881411

con p-value =  $P(\text{Chi-quadro}(1) > 0,00881411) = 0,925$

Test per ARCH di ordine 2

Ipotesi nulla: non sono presenti effetti ARCH

equazione 1: Statistica test LM = 24,1245

con p-value =  $P(\text{Chi-quadro}(2) > 24,1245) = 5,77341\text{e-}006$

equazione 2: Statistica test LM = 4,72214

con p-value =  $P(\text{Chi-quadro}(2) > 4,72214) = 0,0943191$

equazione 3: Statistica test LM = 1,91581

con p-value =  $P(\text{Chi-quadro}(2) > 1,91581) = 0,383697$