



Università
Ca' Foscari
Venezia

DIPARTIMENTO DI ECONOMIA

Corso di Laurea Magistrale in Economia e Finanza

Tesi di Laurea

**Struttura a termine dei tassi di interesse e
teoria delle aspettative: un'analisi di
cointegrazione**

Relatore:

Ch. Prof. Davide Raggi

Candidato:

Matteo Stefanini

Matricola 893648

ANNO ACCADEMICO 2022/2023

Sommario

La teoria delle aspettative della struttura a termine viene spesso verificata tramite analisi di cointegrazione, testando la stazionarietà del differenziale di rendimento tra i tassi a breve e lungo termine. Nonostante l'ipotesi di aspettative implichi che lo spread sia un processo stazionario, in quanto combinazione di serie stazionarie, in questo lavoro si dimostra come, soprattutto all'aumentare della maturity, lo spread si comporta maggiormente come un processo non stazionario, rendendo inadeguati i test di cointegrazione. La trattazione di questa tesi è costituita da una prima parte di introduzione sulla struttura a termine dei tassi di interesse e della teoria delle aspettative. Nella seconda parte vengono presentati alcuni concetti chiavi in ambito delle serie storiche, necessari a una maggiore comprensione dell'analisi empirica successivamente svolta. Quest'ultima è stata condotta utilizzando i rendimenti sui Treasury bond americani a 1,3,5,10 anni e il Federal Fund Rate, ovvero il tasso di policy della Federal Reserve. Si è innanzitutto rilevata la non stazionarietà dei rendimenti, ma la presenza di uno stesso ordine di integrazione ha permesso l'utilizzo delle tecniche di cointegrazione. Lo studio è proseguito con la costruzione degli spread tra i vari tassi di interesse, testandone la stazionarietà per verificare l'effettiva sussistenza di una relazione di cointegrazione. Il differenziale di rendimento è risultato convergere ad un processo non stazionario, in particolare all'aumentare del termine. Per verificare la robustezza dei risultati ottenuti, si sono poi rilassate alcune delle assunzioni iniziali. Si è proceduto, poi, a costruire un modello a correzione dell'errore (ECM), stimando il termine di correzione con la procedura di Engle e Granger, trovando conferme a supporto dei risultati precedenti. Oltre ad avere un

significato statistico importante, i risultati raggiunti meritano di essere tenuti in considerazione nell'ambito della politica monetaria. Se il legame tra tassi a lungo e a breve si indebolisce all'aumentare dello spread, viene minacciato un aspetto cruciale del meccanismo di trasmissione della politica monetaria. Le banche centrali possono controllare i tassi chiave, come il Federal Fund Rate nel caso della Fed, influenzando altri tassi a breve. Tuttavia, questo effetto potrebbe non trasmettersi in modo prevedibile ai rendimenti a più lunga scadenza.

Indice

1	Struttura a Termine e Teoria delle Aspettative	3
1.1	Struttura a Termine	3
1.2	Teoria delle Aspettative	7
1.2.1	Teoria dei mercati segmentati	10
1.2.2	Teoria del premio per la liquidità e dell'ambiente preferito	11
2	Introduzione alle serie storiche	12
2.1	Alcune caratteristiche delle serie storiche	12
2.1.1	Processo <i>White Noise</i>	14
2.1.2	Processo <i>Moving Average</i>	16
2.1.3	Processo <i>Autoregressivo</i>	17
2.2	Processi integrati	17
2.2.1	Test di radice unitaria	21
2.3	Cointegrazione	23
2.4	Modelli a correzione dell'errore	26
3	Implicazioni della teoria delle aspettative	29
4	Analisi empirica	33
4.1	Dati	33
4.2	Test di Cointegrazione	33
4.3	ECM non ristretto	40

INDICE

A	Proposizione Limitante	45
B	Codici Python per l'analisi empirica	46
	Bibliografia	64

Introduzione

La realizzazione di questo lavoro nasce dall'idea di mettere insieme campi diversi della scienza economica quali econometria, finanza e macroeconomia. Lo studio delle relazioni di lungo periodo permette di condurre analisi approfondite che tengono conto di diversi fattori. La struttura a termine dei tassi di interesse suggerisce come una relazione di questo tipo possa sussistere tra tassi a breve e lunga scadenza, relazione che viene spesso spiegata dalla teoria delle aspettative. Comprendere la struttura a termine è fondamentale, in quanto essa ha implicazioni non soltanto nell'ambito della politica monetaria, ma anche in termini di strategie e decisioni di investimento. L'obiettivo del lavoro è quello di condurre uno studio sul comportamento del differenziale di rendimento tra i tassi a diversa scadenza. La teoria delle aspettative della struttura a termine viene spesso sottoposta a verifica testando la stazionarietà dello spread. Se i rendimenti fossero serie non stazionarie, ed integrate di ordine uno, allora la cointegrazione è un test possibile per la teoria delle aspettative. La derivazione di quest'ultima porta a considerare lo spread come un processo stazionario, in quanto combinazione di serie stazionarie. I risultati ottenuti in questo studio giungono ad una conclusione opposta: lo spread, all'aumentare del termine, converge ad un processo non stazionario, rendendo inappropriati i test di cointegrazione. L'analisi empirica verrà condotta utilizzando 240 osservazioni mensili, 20 anni di dati da dicembre 2003 a novembre 2003, dei tassi sui Treasury americani a 1,3,5, 10 anni e del Federal Fund Rate, ovvero il tasso chiave con cui la Fed conduce la sua politica monetaria. Inoltre, verrà usato un approccio a finestre

mobili, in modo tale da generare un numero di risultati sufficiente ad avere delle frequenze di rifiuto delle ipotesi di cointegrazione; le cosiddette regressioni rolling permettono anche di considerare instabilità o cambi strutturali nella relazione di lungo periodo. Il lavoro sarà suddiviso nel seguente modo. Nei primi capitoli si introdurranno struttura a termine dei tassi di interesse e teoria delle aspettative. Prima di procedere con l'analisi empirica, procederemo a spiegare brevemente alcuni concetti importanti relativi alle serie storiche, in modo tale da rendere maggiormente comprensibili gli argomenti successivi. L'analisi empirica partirà dalla verifica della non stazionarietà e l'integrazione dello stesso ordine dei tassi di interesse, requisito necessario per l'utilizzo di tecniche di cointegrazione. Una volta fatto ciò, si passerà allo studio delle implicazioni sulla teoria delle aspettative. Come accennato in precedenza, essa prevede che lo spread tra tassi con maturity diverse sia un processo stazionario, in quanto somma di serie stazionarie. Sottoporremo quindi il differenziale di rendimento a test di stazionarietà, giungendo a risultati che sembrano propendere verso la non esistenza di cointegrazione. Successivamente, si cercherà di rilassare alcune assunzioni fatte in precedenza in modo da verificare la robustezza dei risultati ottenuti. L'ipotesi di aspettative implica che lo spread rappresenti anche il vettore di cointegrazione, basato sui coefficienti $(1,-1)$. Considereremo quindi un modello a correzione dell'errore senza questa restrizione, nel quale il disequilibrio è stimato tramite la procedura di Engle e Granger. Anche in questo caso, giungeremo a conclusioni a sfavore della cointegrazione.

Capitolo 1

Struttura a Termine e Teoria delle Aspettative

La struttura a termine dei tassi di interesse ha da sempre avuto un ruolo importante per gli agenti del sistema economico quali analisti e investitori, in quanto una sua corretta comprensione fornisce preziose indicazioni sull'andamento futuro dell'economia. Essa è influenzata da vari fattori, tra cui aspettative di inflazione, aspettative sull'economia, politiche monetarie, condizioni di liquidità e fattori globali. Essere in grado di analizzare una determinata curva osservabile sul mercato è quindi un aspetto fondamentale. Al tempo stesso risulta importante comprendere come la curva stessa si possa spiegare, e la teoria delle aspettative rappresenta uno degli approcci teorici che cerca di spiegarla.

1.1 Struttura a Termine

I titoli finanziari, a cui è associato un determinato tasso di interesse, differiscono tra loro per tre aspetti principali:

- Rischio, relativo alla probabilità di insolvenza della controparte. In particolare, se quest'ultima aumenta, aumenterà anche il tasso di interesse richiesto;

- Liquidità, ovvero quanto rapidamente un titolo può essere convertito in denaro "contante". Più il mercato su cui è scambiato quel titolo è ampio e profondo, più quel titolo sarà liquido, e più alta sarà la sua domanda.
- Scadenza (*maturity*), che comporta la classica suddivisione tra tassi a breve e lungo termine. ¹

La presenza contemporanea di titoli obbligazionari con varie maturity permette di determinare quella che viene chiamata Struttura a Termine dei Tassi di Interesse (*Term Structure*), ovvero la relazione tra il rendimento di un titolo e la sua scadenza. È importante sottolineare come tutte le altre caratteristiche di quel titolo debbano essere uguali. Prendiamo, ad esempio, il rischio. Esso viene tipicamente definito tramite il rating con cui le agenzie classificano un titolo (AA, BB+, BB). In questo modo, possiamo osservare una specifica curva dei rendimenti per ognuna delle classi di rating. In base alle condizioni del mercato, la struttura a termine può assumere diverse forme. Tipicamente, si osserva una curva crescente, caratterizzata da un rendimento maggiore per titoli a più lunga scadenza. Ciò riflette l'idea che i rendimenti dovrebbero essere più alti per titoli che hanno una scadenza più lunga, in quanto incorporano un maggior rischio.² Inoltre, in una congiuntura espansionistica, gli investitori chiedono rendimenti più elevati sulle obbligazioni con scadenze più lunghe per compensare l'inflazione e i futuri aumenti dei tassi. Va ricordato come l'inflazione, contenuta intorno al 2%, è considerata sinonimo di crescita dell'economia. Infine, se confrontiamo una struttura a termine di tassi di interesse relativa ad obbligazioni nominali con una di obbligazioni reali, ad esempio gli *index-linked bond*, si potrebbero ottenere informazioni sul tasso di inflazione atteso sul mercato. In genere, una curva dei tassi crescente è anche concava, e quasi si appiattisce per rendimenti a lunghissima scadenza. Considerando il premio al rischio marginale,

¹Sarebbe più coerente parlare di "*Time to Maturity*", ovvero la vita residua del titolo.

²Concetto differente da quello dei rating. Non viene considerato come rischio di insolvenza della controparte, piuttosto come il rischio di una variazione del tasso di interesse. Più un titolo ha vita residua, più alte sono le probabilità che il tasso possa variare.

esso risulta decrescente all'aumentare della scadenza. Vediamo un paio di esempi per capire come la forma della struttura a termine può essere giustificata e come trarre utili indicazioni.

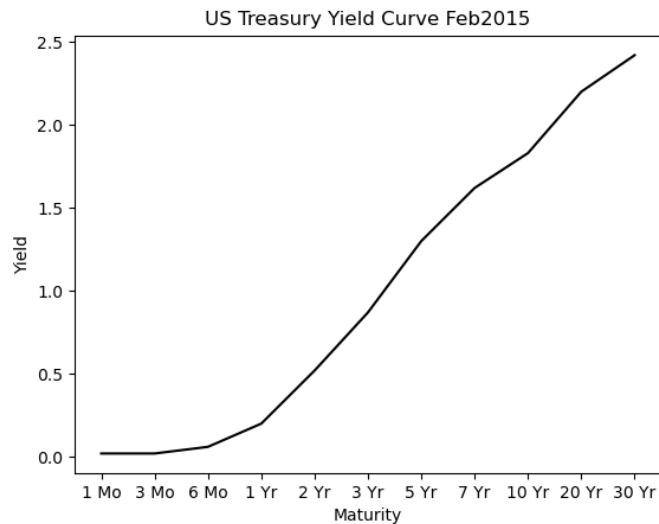


Figura 1.1: Curva dei rendimenti sui Treasury bond a Febbraio 2015

Nella Figura 1.1 è rappresentata la curva dei rendimenti americana a Febbraio 2015. Gli Stati Uniti stavano uscendo dalla crisi dei Subprime, la Fed si preparava a terminare le politiche espansive degli anni precedenti. La curva crescente riflette l'aspettativa degli investitori di un imminente rialzo dei tassi da parte della banca centrale. La curva può anche essere inclinata negativamente, situazione che si verifica meno frequentemente ma che fornisce un segnale forte sull'andamento economico futuro e, di solito, preoccupa gli analisti quando viene osservata sul mercato. Una Term Structure decrescente riflette l'aspettativa futura di un calo dei tassi di interesse, probabile segnale di politiche monetarie espansive delle banche centrali attuate per contrastare fasi di recessione.

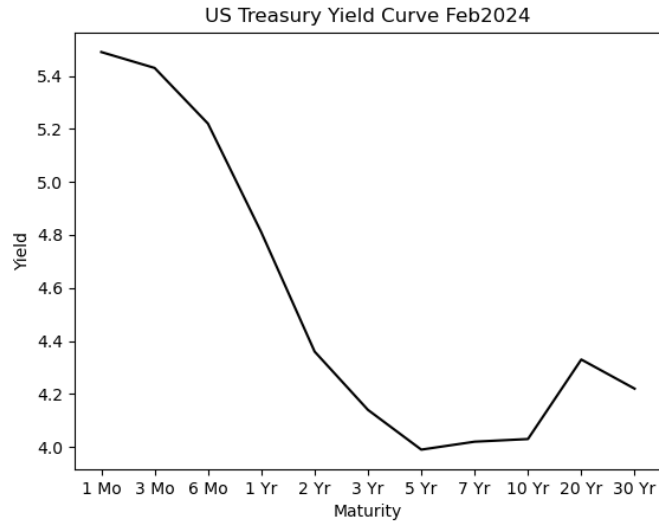


Figura 1.2: Curva dei rendimenti sui Treasury a Febbraio 2024

Nella Figura 1.2 possiamo osservare la curva dei rendimenti americana a Febbraio 2024. Notiamo come le aspettative sui tassi a breve siano decrescenti, riflettendo l'idea che la Fed procederà al taglio dei tassi una volta che l'inflazione sarà tornata a livelli normali. Questo è un caso particolare di curva decrescente, risultato di una situazione economica caratterizzata da molteplici fattori che si sono accumulati nel corso degli ultimi anni. La forma invertita è, in genere, associata a periodi di recessione quando i tassi d'interesse e l'inflazione sono ridotti o in calo. Storicamente, la curva dei rendimenti assume la posizione invertita circa 12-18 mesi prima che abbia inizio una fase di recessione.

In genere, una teoria che possa spiegare la struttura a termine dei tassi di interesse deve essere tale da riuscire a giustificare tre evidenze empiriche:

- Sincronicità dei movimenti. Caratteristica per cui una variazione dei tassi di interesse ha lo stesso effetto su tutti gli spot yield, che li porta a muoversi insieme.
- Relazione tra i tassi a breve e l'inclinazione della curva. Quando si osservano tassi a breve elevati, la curva tende ad essere inclinata negativamente, e

viceversa;

- Bassa probabilità di osservare curve decrescenti.

Tipicamente vengono presentate 3 diverse teorie che cercano di spiegare la struttura a termine dei tassi di interesse. Quella su cui ci concentreremo maggiormente è la teoria delle aspettative, che sarà centrale nella successiva analisi. Le altre due sono la teoria della segmentazione del mercato e la teoria della preferenza per la liquidità e dell'ambiente preferito. Queste ultime possono essere considerate una sorta di estensione della teoria delle aspettative, in quanto tengono conto di aspetti che invece l'ipotesi di aspettative tende a tralasciare o considerare in maniera differente.

1.2 Teoria delle Aspettative

Come accennato in precedenza, la teoria delle aspettative è un approccio teorico che cerca di spiegare la relazione tra i rendimenti di titoli con scadenze diverse. Essa parte dal presupposto che gli investitori prendano decisioni di investimento sulla base delle loro aspettative future sui tassi di interesse. Prima di procedere con la sua derivazione, si rende necessaria l'introduzione di alcuni concetti fondamentali. Innanzitutto, nel parlare di ipotesi di aspettative verranno considerati esclusivamente gli Zero Coupon Bond, ovvero quelle obbligazioni che non pagano interessi intermedi sotto forma di cedole periodiche, ma il cui unico rendimento è dato dalla differenza tra il prezzo d'acquisto e il valore nominale che verrà garantito alla scadenza. Per questa ragione, vengono tipicamente emesse ad un prezzo inferiore (sotto la pari). Il secondo aspetto che merita di essere spiegato è la differenza tra i tassi *spot* e tassi *forward*. Il tasso spot, detto anche a pronti, è quel tasso relativo ad un contratto che viene siglato oggi con scadenza t . Il tasso forward, detto anche a termine, è il tasso relativo ad un contratto che verrà firmato in futuro e scadrà al tempo $t + n$. Nel nostro caso, rappresenta quindi il tasso di uno ZCB con inizio

differito. La relazione che sussiste tra i tassi spot e forward è fondamentale, in quanto ha implicazioni importanti nella formulazione della teoria delle aspettative. In sintesi, partendo dal principio di non arbitraggio, è possibile dimostrare come i tassi forward siano impliciti nei tassi spot, poiché il loro valore viene derivato dalla relazione tra il tasso spot di un investimento con durata maggiore e quello di un investimento di durata minore. La teoria delle aspettative afferma, quindi, che i tassi a lungo termine sono una media dei futuri tassi attesi a breve. In assenza di opportunità di arbitraggio, un investitore deve, infatti, essere indifferente, tra l'investire in una obbligazione a lungo termine o scegliere un portafoglio costituito da più obbligazioni con maturity breve. L'ipotesi cruciale che stiamo facendo nello spiegare l'ipotesi di aspettative è, infatti, quella della neutralità al rischio degli investitori, i quali non hanno preferenze nell'investire in obbligazioni con scadenze diverse. Ciò porta ad affermare che obbligazioni con scadenze diverse siano perfetti sostituiti, conclusione che è stata ampiamente criticata e che rappresenta uno dei punti di partenza delle altre teorie che cercano di spiegare la struttura a termine. Quella che stiamo introducendo è chiamata, infatti, teoria delle aspettative pure (Hall et al. (1992)). Consideriamo con $i_{m,t}$ il tasso continuo di uno zero coupon bond a m periodi e $f_{m,t}$ il tasso forward, definito come il tasso contrattuale di un'obbligazione zero coupon uniperiodale acquistata al tempo t e con scadenza al tempo $t + m$, la relazione che lega $i_{m,t}$ e $f_{m,t}$ può quindi essere descritta come segue:

$$i_{m,t} = \frac{1}{m} \left(\sum_{j=1}^m f_{j,t} \right) \quad (1.1)$$

con $m = 1, 2, 3, \dots$, che ci dice come il tasso spot di uno ZCB a m periodi sia la media dei tassi forward intermedi uniperiodali. Il passo ulteriore è quello di concentrarsi sulla relazione che sussiste tra i tassi forward, impliciti nella curva dei tassi spot, e i tassi che effettivamente si osserveranno in futuro. Come detto, un

investitore dovrebbe essere in grado di ricavare i tassi forward dalla curva dei tassi spot. Bisogna chiedersi, a questo punto, se questi tassi forward coincidano con i futuri tassi attesi, ovvero è necessario introdurre la componente di aspettative. Si ipotizzi che la relazione tra i tassi a termine e i tassi attesi sia caratterizzata dall'equazione:

$$f_{j,t} = E_t (i_{m,t+j-1}) + \phi_{j,t} \quad (1.2)$$

dove E_t rappresenta l'operatore matematico di aspettativa, condizionato alle informazioni disponibili al tempo t , e $\phi_{j,t}$ è il premio per il rischio a termine. La sostituzione dell'eq.1.2 nella 1.1, porta ad un'equazione di questo tipo:

$$i_{m,t} = \frac{1}{m} \left[\sum_{j=1}^m E_t (i_{1,t+j-1}) \right] + r_{m,t} \quad (1.3)$$

dove $r_{m,t} \equiv \frac{1}{m} \sum_{j=1}^m \phi_{j,t}$ è una variabile che cattura gli effetti del premio a termine. L'equazione può essere vista come una relazione generale che lega i rendimenti a diverse scadenze e ci anticipa chiaramente come i rendimenti con scadenze simili si muovano insieme. Ciò porta ad affermare come la teoria delle aspettative riesca spiegare chiaramente la prima delle evidenze empiriche che avevamo introdotto. Possiamo, infatti, riscrivere l'equazione 1.3 in termini di spread, ovvero:

$$i_{m,t} - i_{1,t} = \frac{1}{m} \left[\sum_{s=1}^{m-1} \sum_{j=1}^s E_t \Delta i_{1,t+j} \right] + r_{m,t} \quad (1.4)$$

dove Δ è l'operatore differenza prima. L'equazione 1.4 fornisce una forte implicazione, che sarà poi la base dell'analisi empirica. Assumendo che i rendimenti a

scadenza siano realizzazione di processi stocastici integrati di ordine uno, $I(1)$, se la componente del premio a termine è costante, e quindi stazionaria, tutte le variabili a destra dell'equazione sono stazionarie. Ne consegue che anche il termine a sinistra dell'equazione (spread) è stazionario, $I(0)$. Il modello prevede, quindi, che i rendimenti a diverse scadenze siano cointegrati, con un vettore di cointegrazione $(1, -1)$. Abbiamo visto come la teoria delle aspettative si propone di spiegare la forma della struttura a termine sulle base delle aspettative sui tassi futuri. Il passo ulteriore sarebbe quello di concentrarsi maggiormente sulle proprietà del premio a termine. Come affermato in precedenza, secondo l'ipotesi delle aspettative pure, i premi a termine sono tutti identicamente uguali a zero. Un'altra versione più blanda dell'EH afferma la proposizione meno rigorosa che i premi a termine siano costanti nel tempo. Le ulteriori teorie che avevamo citato all'inizio si basano su assunzioni differenti riguardo il premio a termine, per cui si rende necessaria una loro breve introduzione.³

1.2.1 Teoria dei mercati segmentati

La teoria dei mercati segmentati nega che le obbligazioni con scadenze diverse siano perfetti sostituti, in quanto proprio la loro diversa scadenza comporta dinamiche differenti. In questa impostazione, i tassi sulle obbligazioni con lunga scadenza sono maggiori data una preferenza degli investitori verso le obbligazioni con scadenze brevi, e quindi meno rischiose. Questa teoria riuscirebbe quindi a spiegare la notevole frequenza con cui si osserva un'inclinazione positiva della curva della struttura a termine dei tassi. Tuttavia non riesce a spiegare né la sincronicità degli spostamenti della curva, né perché a spot yield a breve molto elevati corrisponda di norma una curva dei rendimenti decrescente.

³Nella derivazione dell'eq.1.3 abbiamo inserito fin dall'inizio la componente di premio a termine, nonostante la teoria delle aspettative pure la ponga uguale a zero, in quanto basata sull'assunto che gli investitori siano neutrali al rischio. Nel corso della nostra analisi faremo sempre riferimento alla forma comprendente il premio a termine.

1.2.2 Teoria del premio per la liquidità e dell'ambiente preferito

La teoria del premio di liquidità e dell'ambiente preferito rappresenta, in un certo senso, l'estensione logica della teoria delle aspettative, in quanto stabilisce che il tasso di interesse riferito a un'obbligazione a lungo termine è uguale alla media dei tassi di interesse a breve più un premio per il rischio. Il presupposto è che le obbligazioni con scadenze diverse siano sostituti imperfetti. Gli investitori tendono a preferire obbligazioni a più breve termine, perché sopportano un minore rischio di interesse. La componente di ambiente preferito assume che gli investitori abbiano preferenze diverse sulla maturity delle obbligazioni (un diverso ambiente preferito); in particolare chi acquista obbligazioni ha preferenze per maturity brevi, mentre chi le vende per maturity lunghe. In questa situazione allora per convincere gli investitori a detenere obbligazioni a lunga scadenza, il rendimento di queste deve essere superiore al rendimento delle obbligazioni a scadenza breve. Ad oggi, questa teoria sembra essere l'unica in grado di spiegare tutte e tre le evidenze empiriche. Interessante a riguardo è il lavoro di Cox et al. (1981).

Capitolo 2

Introduzione alle serie storiche

2.1 Alcune caratteristiche delle serie storiche

I dati a cui vengono applicate le tecniche inferenziali tipiche dell'econometria possono essere di due tipi: *cross-section*, nel caso in cui le osservazioni disponibili siano relative ad individui diversi, oppure *serie storiche*, quando ciò che abbiamo sono osservazioni, su una o più variabili, misurate nel tempo. Il caso delle serie storiche presenta una differenza concettuale di base che può risultare anche banale: il tempo evolve in un'unica direzione, e quindi esiste un passato. In un contesto del genere, la tendenza di molti fenomeni ad evolversi in modo più o meno regolare porta a pensare che il dato rilevato in un determinato istante t sia più simile a quello rilevato all'istante $t - 1$ piuttosto che in epoche distanti; si può dire che la serie storica che analizziamo ha "memoria di sé", caratteristica tipicamente indicata col nome di persistenza. Essa differenzia nettamente i campioni di serie storiche da quelli *cross-section*, perché nei primi l'ordine dei dati ha un'importanza fondamentale, mentre nei secondi esso è del tutto irrilevante. Se pensiamo a come gli eventi economici siano caratterizzati da dinamicità, viene naturale osservare come, nello spiegare ed interpretare determinati fatti, la componente temporale gioca un ruolo fondamentale. In ambito di serie storiche si parla di processo stocastico, definito come una

sequenza infinitamente lunga di variabili casuali. Un campione di T osservazioni consecutive nel tempo non viene, quindi, pensato come una realizzazione di T variabili casuali distinte, bensì come parte di un'unica realizzazione di un processo stocastico, la cui memoria è data dal grado di connessione fra le variabili casuali che lo compongono. Una serie storica è, quindi, una sequenza finita di variabili casuali in cui l'indice $t = 1, \dots, T$ indica il tempo. Come accade per tutti i modelli di regressione classica, anche quelli per le serie storiche descrivono l'evoluzione delle medie condizionali, con la differenza che nel caso di queste ultime si cerca di misurare anche la relazione tra passato e presente di una certa variabile. In generale, lo scopo dei modelli per le serie storiche è quello di approssimare le distribuzioni condizionali¹ dei dati, ovvero descrivere, da un punto di vista statistico, il comportamento di una certa variabile una volta che si riesce ad osservare tutto quello che è successo in precedenza. Di conseguenza, è logico aspettarsi che il valore di una data serie storica y al tempo t sia dipendente con il suo valore al tempo precedente. Si parla, allora, di autocorrelazione, ovvero la correlazione di una serie con i propri valori ritardati. Se poi consideriamo la dipendenza temporale a più ritardi, al netto dei valori intermedi, si parla di autocorrelazione parziale.² Un aspetto cruciale che permette di fare inferenza sulle caratteristiche di un determinato processo, è che esso sia stabile nel tempo, cioè che i suoi connotati probabilistici rimangano invariati, per lo meno all'interno dell'intervallo di osservazione considerato. Una serie storica si dice, quindi, stazionaria se le sue caratteristiche statistiche, come media e varianza, non dipendono dal tempo, ovvero sono costanti. Le serie storiche stazionarie sono importanti in analisi statistiche e modellazione, poiché semplificano l'analisi dei dati temporali. Possiamo vedere la stazionarietà anche come assenza di

¹Da non confondere con la distribuzione marginale, che descrive una determinata variabile al netto del fattore tempo.

²Lo studio della struttura di autocorrelazione è fondamentale, in quanto fornisce preziose informazioni per identificare un modello che descriva nella maniera più appropriata possibile il processo generatore di dati, o per verificare che un dato modello sia correttamente specificato, come accade nell'analisi dei residui.

un trend che caratterizza la serie storica; è evidente come la presenza di una componente tendenziale, che ad esempio determina una crescita costante nel tempo, porta un processo ad essere non stazionario. Un'altra caratteristica di cui vale la pena parlare è l'ergodicità, ovvero la condizione che limita la memoria del processo. Un processo non ergodico è caratterizzato da una persistenza tale che, prendendo un segmento del processo stesso, esso sia insufficiente a dire qualcosa sul modo in cui si distribuisce. In un processo ergodico, al contrario, la memoria dello stesso è debole su lunghi orizzonti, nel senso che eventi molto lontani tra loro possono essere considerati indipendenti. Osservando la serie per un periodo abbastanza lungo, siamo in grado di osservare la quasi totalità delle sottosequenze che il processo è in grado di generare. Le condizioni per le quali un processo stocastico è anche ergodico non sono oggetto di questo lavoro data la loro complessità, per cui ci limitiamo solo a introdurre il concetto. A questo punto, possiamo chiederci in che modo utilizzare il processo stocastico come idealizzazione del vero processo generatore di dati. Osservando una serie storica, a prescindere dal fatto che la si possa considerare realizzazione di un processo stazionario, ciò che vogliamo fare è trovare il processo che rappresenti la serie nel miglior modo possibile. Detto ciò, si rende necessaria una breve introduzione dei modelli maggiormente utilizzati per descrivere le serie storiche.

2.1.1 Processo *White Noise*

Il white noise, definito come ϵ_t è il processo stocastico più semplice che possiamo osservare. È un processo che possiede momenti almeno fino al secondo ordine (media e varianza), i quali sono costanti nel tempo (è quindi un processo stazionario), ma non danno al processo alcuna memoria di se, ovvero sono indipendenti. Un white noise è quindi una successione di variabili aleatorie, ϵ_t , con le seguenti caratteristiche:

- $E[\epsilon_t] = 0 \quad \forall t$

- $E(\epsilon_t^2) = V(\epsilon_t) = \sigma^2$

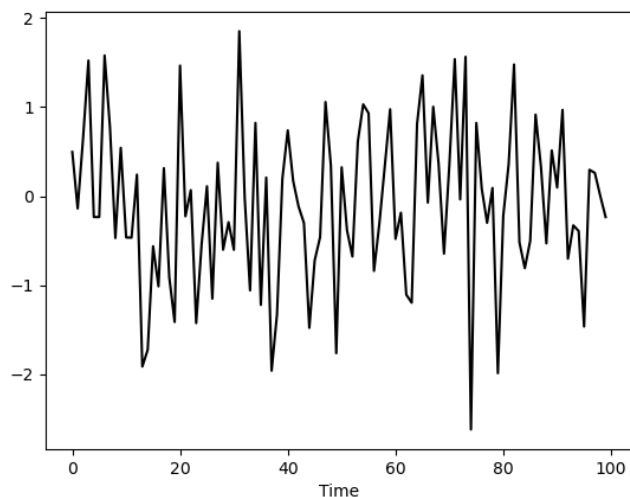


Figura 2.1: Possibile traiettoria di un processo White Noise

Un white noise, come quello in Figura 2.1 è, in sostanza, un processo composto da un numero infinito di variabili casuali a media zero e varianza costante; variabili che sono tutte incorrelate tra loro. Ne consegue che dipendenza e persistenza di cui abbiamo parlato prima non siano caratteristiche proprie del white noise. In quanto tale non sembrerebbe il processo più adeguato a modellare una serie storica, e quindi descrivere dati che invece dipendenza la presentano. Esso è comunque importante per due ragioni: la prima risiede nel fatto che il white noise viene tipicamente utilizzato per definire il termine di disturbo (errore) dei modelli di regressione; la seconda è che, applicando un polinomio nell'operatore ritardo, ovvero facendo dipendere l'osservazione presente dagli shock passati, si ottiene l'estensione classica dei white noise, denominati modelli a media mobile, o Moving Average.

2.1.2 Processo *Moving Average*

Un processo MA è definito nel seguente modo:

$$y_t = \sum_{i=0}^q \theta_i \epsilon_{t-i}, \quad (2.1)$$

dove ϵ_t è un processo white noise, e q è l'ordine del modello, ossia il numero di ritardi considerati. Un modello MA(1) sarà quindi nella forma: $Y_t = \epsilon_t + \theta\epsilon_{t-1}$. Analizzando la sua media, essa può essere scritta come:

$$E(y_t) = E \left[\sum_{i=0}^q \theta_i \epsilon_{t-i} \right] \quad (2.2)$$

$$= \sum_{i=0}^q \theta_i E(\epsilon_{t-i}) = 0. \quad (2.3)$$

Per quanto riguarda l'autocorrelazione, la sua funzione per un modello MA(q) assume valori diversi da 0 solo per valori maggiori di q . Infatti, l'autocovarianza di ordine k è data da:

$$E(y_t y_{t+k}) = E \left[\left(\sum_{i=0}^q \theta_i \epsilon_{t-i} \right) \left(\sum_{j=0}^q \theta_j \epsilon_{t-j+k} \right) \right] \quad (2.4)$$

$$= \sum_{i=0}^q \theta_i \left(\sum_{j=0}^q \theta_j E(\epsilon_{t-i} \epsilon_{t-j+k}) \right) \quad (2.5)$$

Sfruttando le proprietà del white noise, si ha che $E(\epsilon_{t-i} \epsilon_{t-j+k}) = \sigma^2$ per $j = i + k$ e 0 in tutti gli altri casi, cosicché l'espressione precedente si riduce a:

$$\gamma_k = E(y_t y_{t+k}) = \sigma^2 \sum_{i=0}^q \theta_i \theta_{i+k} \quad (2.6)$$

dove si intende che $\theta_i = 0$ per $i > q$. Possiamo notare come per $k > q$, le autocovarianze sono nulle. Un processo MA(q), quindi, è un processo ottenuto come combinazione di diversi osservazioni di uno stesso white noise, e presenta una persistenza

maggiore all'aumentare dell'ordine.

2.1.3 Processo *Autoregressivo*

I modelli autoregressivi sono forse i più adatti ed intuitivi per rappresentare le caratteristiche di persistenza. Sono infatti definiti da un modello di regressione la cui variabile osservata oggi dipende da i suoi valori passati; l'idea è quella che la variabile dipendente sia una funzione lineare dei suoi valori ritardati più una componente di white noise, che definisce il disturbo e interpretabile, quindi, come classica differenza tra y_t e la sua media condizionale. Un modello AR sarà quindi così definito:

$$y_t = \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \epsilon_t. \quad (2.7)$$

Utilizziamo la sua forma più semplice per analizzare meglio le caratteristiche, ovvero un modello AR(1):

$$y_t = \beta_1 y_{t-1} + \epsilon_t. \quad (2.8)$$

In genere, un modello AR presenta delle caratteristiche di persistenza tali che le variabili a ritardo più ravvicinato forniscono maggiori informazioni delle precedenti, rendendo decrescente la struttura di autocorrelazione. Per l'analisi dei suoi momenti rimandiamo alla sezione successiva, in cui parleremo più nel dettaglio di processi integrati e in che modo si possa verificare la stazionarietà di un processo.

2.2 Processi integrati

L'idea generale è che la variabile che osservo oggi sia funzione esclusivamente di quello che ho osservato ieri, più uno shock casuale, rappresentazione coerente con la teoria dei mercati efficienti.³ Esaminiamo l'eq.2.8 e chiediamoci sotto quali con-

³In modo più che sintetico, essa afferma come non sia possibile generare degli extra-rendimenti nel mercato. Tutte le informazioni possibili sono implicite nel prezzo di un titolo. Nel prevedere il futuro prezzo dello stesso, posso fare affidamento esclusivamente sulla sua osservazione al periodo

dizioni il processo risulta stazionario. Risolvendo il modello "all'indietro", cioè sostituendo a y_{t-1} il suo valore in funzione di y_{t-2} e così via, si ha che:

$$y_t = \sum_{j=0}^{\infty} \beta_1^j \epsilon_{t-j}$$

La media è nulla e non dipende da t , in quanto:

$$E[y_t] = \sum_{j=0}^{\infty} \beta_1^j E[\epsilon_{t-j}] = 0$$

La varianza invece è data da:

$$\text{Var}(y_t) = \text{Var}(\epsilon_t) \sum_{j=0}^{\infty} \beta_1^{2j} = \frac{\sigma^2}{1 - \beta_1^2},$$

che non dipende da t solo se $|\beta_1| < 1$. Generalizzando il modello, e sostituendo i valori come fatto in precedenza, si ha che:

$$y_t = y_{t-n} + \sum_{i=0}^{n-1} \epsilon_{t-i} \quad (2.9)$$

Ora, supponiamo che il processo abbia avuto origine in un preciso periodo, $t = 0$ in cui il valore di $y = 0$. In questo caso, la 2.9 diventa:

$$y_t = \sum_{i=1}^t \epsilon_i; \quad (2.10)$$

che possiamo considerare un rappresentazione particolare di un processo a media mobile, in cui i coefficienti sono uguali a uno e la variabile y_t è funzione degli errori passati. Analizzando i momenti di questo periodo, la media del processo è 0, che quindi renderebbe il processo stazionario. Considerando la varianza, essa non è costante in quanto y_t è somma di t variabili i.i.d con varianza σ^2 . Ne consegue che

precedente.

la varianza di y_t è $t\sigma^2$, rendendo il processo non stazionario. Quello che abbiamo presentato è detto Random Walk, che può essere quindi considerato un'estensione di un modello AR(1) in cui la persistenza è tale da modificare le sue caratteristiche; ovvero, l'effetto di uno shock al tempo t permane nel futuro. Questi processi sono anche chiamati "a radice unitaria", data la presenza di un coefficiente uguale a 1 nella serie. Un aspetto caratteristico dei processi stazionari è quello di essere mean reverting, ovvero la tendenza della serie di tornare verso il suo valore atteso⁴. L'assenza della stessa in un processo random walk provoca periodi, anche lunghi, in cui la serie presenta trend crescenti o decrescenti molto marcati. Per questo, quando facciamo riferimento a questi tipi di processi, parliamo più un generale di trend stocastici, opposti ai trend deterministici che sono semplici funzioni lineari del tempo. Vediamo che aspetto ha un Random Walk:

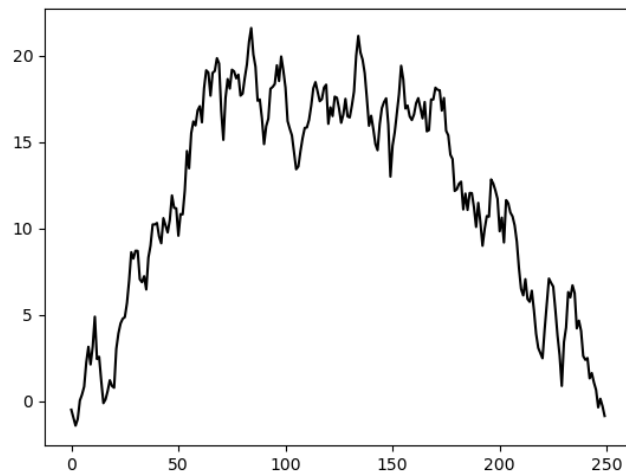


Figura 2.2: Possibile traiettoria di un processo Random Walk

In Figura 2.2 è rappresentata una simulazione di un processo Random Walk; possiamo notare l'alternanza di trend crescenti e decrescenti, e l'aumento di persistenza dovuto alla presenza della radice unitaria.

⁴Tale tendenza può essere misurata la funzione di risposta di impulso, che per un processo mean reverting tende a zero

Allo stesso tempo, nulla preclude che ci possano essere dei casi in cui un trend deterministico si sovrappone ad un trend stocastico. Essi sono definiti Random Walk con drift, ovvero dei Random Walk a cui si sovrappone una funzione lineare del tempo. Considerando la 2.9, essa diventa:

$$y_t = \sum_{i=0}^t \epsilon_i + \mu \cdot t \quad (2.11)$$

Tipicamente, quando si hanno serie storiche di variabili macroeconomiche, ciò che osserviamo sono realizzazione di processi non stazionari. Risulta così interessante capire se la non stazionarietà è causata da trend deterministici, oppure stocastici (come il random walk). Prendiamo, ad esempio, la serie storica del PIL per un periodo di tanti anni. Fattori come il progresso tecnico e l'accumulazione di capitale fanno sì che la serie presenti un trend crescente, a cui si contrappone l'effetto del ciclo economico che, essendo un fenomeno di breve periodo a media 0 per definizione, è stazionario. Possiamo quindi pensare ad una serie storica che oscilla in maniera stazionaria attorno ad un trend. L'idea potrebbe quindi essere quella di scomporre la serie, studiando distintamente il trend e la componente di ciclo. La serie potrebbe essere rappresentata in questo modo:

$$y_t = \alpha + \beta t + u_t,$$

dove u_t è un processo stocastico stazionario. Processi di questo tipo sono detti Trend Stazionari. In questi casi è possibile stimare la componente deterministica del modello (ovvero i coefficienti α e β), tramite OLS per poi descrivere la componente residuale, che rappresenta l'andamento di breve periodo della serie, tramite un modello stazionario. A questo punto, possiamo introdurre il concetto di processi integrati. Si definisce come processo $I(d)$ un processo la cui differenza d-esima è stazionaria; un processo $I(1)$ è quindi un processo la cui differenza prima è stazionaria.

Questo è il caso classico di serie storiche di prezzi, non stazionarie, che diventano stazionarie considerando i rendimenti, ovvero la loro differenza prima. Ad esempio, y_t è un Random Walk se δ_t è un White Noise. Tramite la trasformazione in $I(0)$ di un processo $I(1)$, modelleremo un tasso di crescita anziché il (logaritmo del) PIL, il tasso d'inflazione anziché (il logaritmo del) l'indice dei prezzi, e così via.

2.2.1 Test di radice unitaria

Un processo integrato è, quindi, un esempio di processo non stazionario, tipica rappresentazione delle serie storiche macroeconomiche. Di conseguenza, nel momento in cui utilizziamo una serie storica, è necessario sapere se essa sia rappresentata da un processo stazionario o meno. L'analisi del grafico può essere un buon punto di partenza; sappiamo che se una serie presenta un marcato trend, che sia crescente o decrescente, è difficile che possa essere stazionaria. Il caso del trend stocastico è già di più difficile interpretazione in quanto, come visto, se fosse solo per la media il processo potrebbe essere considerato stazionario. Si rende necessaria una regola decisionale che sia meno soggettiva ed arbitraria. Una regola di questo tipo è rappresentata dai test di radice unitaria, chiamati in questo modo perchè verificano, appunto, la presenza di una radice unitaria (componente random walk) all'interno del processo. Partendo dalla rappresentazione di un modello AR(1): $y_t = \beta_1 y_{t-1} + \epsilon_t$, l'idea del test Dickey Fuller è appunto quella di verificare la presenza di una componente RW all'interno del processo, ovvero il caso in cui β_1 sia uguale a 1. Il sistema di ipotesi sarà quindi:

- $H_0 : \beta_1 = 1$
- $H_1 : \beta_1 < 1$

in cui l'ipotesi nulla H_0 rappresenta la non stazionarietà. Spesso, risulta più intuitivo considerare il modello AR(1) riparametrizzato nel seguente modo:

$$\Delta y_t = \delta y_{t-1} + \epsilon_t, \quad (2.12)$$

in cui $\delta = \beta_1 - 1$. Di conseguenza, il sistema di ipotesi diventa:

- $H_0 : \delta = 0$
- $H_1 : \delta < 0$

Infatti, se $\delta=0$, allora $\beta_1=1$ e siamo in presenza di un processo I(1). Il sistema è quindi basato sulla solita statistica test $t = \frac{\hat{\delta}}{\sqrt{\widehat{\text{Var}}(\delta)}}$, dove " \wedge " rappresenta lo stimatore OLS. Infine, la distribuzione asintotica della statistica non è una Normale standard, e dipende dal modello che viene utilizzato per la stima di δ (ad esempio, se nel modello viene inserita o meno l'intercetta). È importante sottolineare come la distribuzione del test t si distribuisce come una variabile DF se la sequenza ϵ_t è incorrelata, ovvero ϵ_t è un white noise per cui $\text{Cov}(\epsilon_t, \epsilon_{t-j}) = 0, \forall j$. Se i residui presentassero autocorrelazione, la distribuzione della statistica t non sarebbe calcolabile. L'estensione per ovviare a questo problema è rappresentata dall'Augmented Dickey-Fuller che estende il test DF ad un modello generico di ordine p .⁵ Avremo quindi un modello di questo tipo:

$$y_t = \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \epsilon_t$$

che, riparametrizzato come fatto in precedenza, diventa:

$$\Delta y_t = \delta y_{t-1} + \beta_1 \Delta y_{t-1} + \dots + \beta_p \Delta y_{t-p+1} + u_t.$$

⁵L'idea è simile ad un modello di regressione semplice. Quando abbiamo il problema della variabile omessa, aggiungiamo regressori per far sì che il termine di errore non sia correlato con i regressori. Nelle serie storiche i regressori sono rappresentati dalle osservazioni della variabile a periodi precedenti, per questo aggiungiamo ritardi per renderli incorrelati.

seguendo lo stesso tipo di verifica di ipotesi. La distribuzione del test ADF è la stessa del DF se l'ordine di p è sufficientemente alto affinché la correzione sia efficace, ovvero i residui diventino white noise.⁶ Il caso appena esaminato è relativo al test di un processo autoregressivo che genera un Random Walk. Nel caso in cui la serie storica presenti una componente tendenziale, un drift, il test ADF non cambia i suoi connotati. L'aspetto delicato è quello di decidere se includere o meno questo trend all'interno della regressione da stimare. Spesso la componente tendenziale è talmente evidente che risulta naturale considerarla, altre un pò meno. Si potrebbe decidere di stimare la regressione che include la componente tendenziale; al massimo il coefficiente ad esso relativo risulterà zero. Il problema sussiste nel cambiamento della distribuzione della variabile, che non è insormontabile ma non è oggetto di questo lavoro.

2.3 Cointegrazione

Come abbiamo detto in precedenza, le variabili economiche sono spesso descritte da processi non stazionari, comportando problemi inferenziali e costringendoci ad eseguire trasformazioni come la differenziazione per avere delle serie stazionarie. Non è difficile dimostrare che una combinazione lineare di due processi stocastici stazionari, è stazionaria. Di conseguenza, una combinazione lineare tra due processi $I(1)$, non stazionari, dovrebbe essere ancora un processo $I(1)$. Tuttavia, potremmo osservare come una o più variabili tendano a muoversi insieme, in maniera sistematica, poiché potrebbero condividere lo stesso trend stocastico. In questo modo, se si considerasse una combinazione tra le variabili, risulterebbe un processo stazionario. Si dice, quindi, che le variabili sono cointegrate. Tale relazione, essendo stazionaria, e quindi stabile nel tempo, potrebbe essere considerata come una relazione di lungo

⁶Tipicamente si sceglie p in modo da minimizzare criteri come Akaike e Schwartz. Si può anche scegliere di eseguire un test di Philippe Perron (PP), Ng and Perron (1995)

periodo.⁷ In genere, si dice che c'è cointegrazione tra due variabili non stazionarie, $I(1)$, se esiste almeno una loro combinazione lineare che è stazionaria. Possiamo formalizzare questa affermazione nel seguente modo: Se x_t è $I(d)$ e y_t è $I(b)$, allora $z_t = x_t + \lambda y_t$ è $I(c)$, dove:

- $c = \max(d, b)$ per $d \neq b$
- $c \leq \max(d, b)$ per $d = b$

Quando la disuguaglianza vale in senso stretto si ha cointegrazione. Noi ci concentreremo maggiormente sull'ipotesi che $d = b$, cioè che la combinazione lineare di due processi integrati di ordine 1 è stazionaria. Possiamo anche considerare il caso di un processo multivariato che descrive, congiuntamente, la dinamica di più variabili. Consideriamo $\mathbf{Y}_t = (y_{1t}, \dots, y_{nt})'$ come un vettore ($n \times 1$) di serie storiche $I(1)$. \mathbf{Y}_t è cointegrato se esiste almeno un vettore ($n \times 1$), $\boldsymbol{\beta} = (\beta_1, \dots, \beta_n)'$, per cui

$$\boldsymbol{\beta}'\mathbf{Y}_t = \beta_1 y_{1t} + \dots + \beta_n y_{nt} \sim I(0)$$

Questa affermazione giustifica anche il fatto che ci potrebbero essere più vettori di cointegrazione, i quali possono essere rappresentati tramite una matrice, detta di cointegrazione, ed il cui numero di vettori linearmente indipendenti prende il nome di rango di cointegrazione. In particolare, se il vettore \mathbf{Y}_t , che ($n \times 1$), è cointegrato, ci potrebbero essere $0 < r < n$ vettori di cointegrazione linearmente indipendenti, in cui r rappresenta il rango della matrice di cointegrazione. Possiamo osservare tutto ciò in termine di trend stocastici comuni. Se il vettore di serie storiche ($n \times 1$), \mathbf{Y}_t , è cointegrato con $0 < r < n$ vettori di cointegrazione, allora ci saranno $n - r$ trend stocastici comuni $I(1)$. Consideriamo $\mathbf{Y}_t = (y_{1t}, y_{2t})' \sim I(1)$ e $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t})' \sim I(0)$ e supponiamo che \mathbf{Y}_t è cointegrato con vettore di

⁷Letteratura generale e applicazioni finanziarie possono essere osservate in Engle and Granger (1987), Hamilton (2020), Campbell and Shiller (1987), Johansen (1995), Banerjee et al. (1993)

cointegrazione $\beta = (1, -\beta_2)'$. la relazione può essere rappresentata come:

$$y_{1t} = \beta_2 \sum_{s=1}^t \varepsilon_{1s} + \varepsilon_{3t}$$

$$y_{2t} = \sum_{s=1}^t \varepsilon_{1s} + \varepsilon_{2t}$$

In questo caso, il trend stocastico comune è $\sum_{s=1}^t \varepsilon_{1s}$. La relazione $\beta'Y_t$ permette di annullare il trend stocastico, infatti:

$$\beta'Y_t = \beta_2 \sum_{s=1}^t \varepsilon_{1s} + \varepsilon_{3t} - \beta_2 \left(\sum_{s=1}^t \varepsilon_{1s} + \varepsilon_{2t} \right) = \varepsilon_{3t} - \beta_2 \varepsilon_{2t} \sim I(0)$$

Come detto, la cointegrazione è spesso associata a relazioni di lungo periodo, chiamate anche relazione di equilibrio⁸. Ad esempio, il modello del reddito permanente implica la cointegrazione tra consumo e reddito, con il consumo come tendenza comune. I modelli relativi alla teoria della crescita implicano una cointegrazione tra reddito, consumo e investimento, con la produttività come tendenza comune. La parità di potere d'acquisto implica la cointegrazione tra il tasso di cambio nominale e i prezzi esteri e nazionali. La parità del tasso di interesse coperto implica la cointegrazione tra i tassi di cambio a termine e a pronti. L'equazione di Fisher implica la cointegrazione tra tassi di interesse nominali e inflazione. Fino ad arrivare al nostro caso, in cui la teoria delle aspettative implica una relazione di cointegrazione tra tassi di interesse nominali a diverse scadenze. La logica conseguenza di relazioni di questo tipo sono i modelli a correzione dell'errore, generati appunto quando si ha una deviazione dalla relazione di cointegrazione.

⁸Perché, nonostante la dinamica delle serie Y_t sia instabile, la loro relazione risulta non dipendente da t , e quindi valida anche nel lungo periodo

2.4 Modelli a correzione dell'errore

La relazione di lungo periodo determinata dalla cointegrazione tra due variabili presuppone che esse si muovano insieme nel corso del tempo. Al tempo stesso, nulla preclude che nel breve periodo possano verificarsi deviazioni da tale relazione. Il meccanismo a correzione dell'errore rappresenta il modello che cerca di spiegare come certe deviazioni possano essere corrette, riassorbite nel tempo. L'importanza di tali modelli è, infatti, quella di coniugare variabili economiche, breve e lungo periodo. Vediamo nel dettaglio come possono essere definiti. Consideriamo un vettore bivariato, $I(1)$, $\mathbf{Y}_t = (y_{1t}, y_{2t})'$, e assumiamo che \mathbf{Y}_t è cointegrato, con un vettore di cointegrazione $\boldsymbol{\beta} = (1, -\beta_2)'$, in modo che $\boldsymbol{\beta}'\mathbf{Y}_t = y_{1t} - \beta_2 y_{2t}$ è $I(0)$. Tale relazione può essere rappresentata come segue (Engle and Granger (1987)):

$$\Delta y_{1t} = c_1 + \alpha_1 (y_{1t-1} - \beta_2 y_{2t-1}) + \sum_j \psi_{11}^j \Delta y_{1t-j} + \sum_j \psi_{12}^j \Delta y_{2t-j} + \varepsilon_{1t} \quad (2.13)$$

$$\Delta y_{2t} = c_2 + \alpha_2 (y_{1t-1} - \beta_2 y_{2t-1}) + \sum_j \psi_{21}^j \Delta y_{1t-j} + \sum_j \psi_{22}^j \Delta y_{2t-j} + \varepsilon_{2t} \quad (2.14)$$

che è appunto un modello a correzione dell'errore che spiega la dinamica delle variabili cointegrate. In pratica, la componente $(y_{1t-1} - \beta_2 y_{2t-1})$ dell'eq.2.13 rappresenta lo squilibrio al periodo precedente, e il rispettivo coefficiente α_1 definisce l'ampiezza, o la velocità di aggiustamento. Le altre componenti aggiuntive rappresentano le variazioni delle variabili dipendenti a periodi precedenti, aggiunte per rendere incorrelati i residui e ottenere un modello che funzioni meglio in termini di stima. Se un sistema cointegrato ha rango di cointegrazione r , avremmo r relazioni di lungo periodo, o di equilibrio, e quindi r processi stazionari che descrivono l'andamento nel tempo delle deviazioni dalle stesse. Nel momento in cui questi squilibri $((y_{1t-1} - \beta_2 y_{2t-1})$ siano diversi da 0, si avrà un movimento nel vettore $\mathbf{Y}_t + 1$ tale per cui lo squilibrio tende a venire riassorbito con ampiezza α_1 . Sulla base del modello precedente si possono verificare 3 possibili situazioni:

- Non si verifica disequilibrio a $t - 1$, ovvero $y_{t-1} = \beta x_{t-1}$, per cui se a $t - 1$ si osserva equilibrio, ci si aspetta che tale relazione rimanga invariata;
- $y_{t-1} > \beta x_{t-1}$, per cui ci si attende una variazione negativa di y . Infatti, per compensare ad un disequilibrio positivo occorre che decresca y_t per tornare in equilibrio.
- $y_{t-1} < \beta x_{t-1}$, ovvero ci si trova al di sotto del livello di equilibrio, di conseguenza, per compensare tale situazione, ci si aspetta che la y cresca.

Per le due situazioni in cui è necessaria una variazione della y per ritornare all'equilibrio, la velocità è sempre misurata dal coefficiente α .

Come spesso accade, nel momento in cui decidiamo di stimare dei parametri, possono sorgere dei problemi. In particolare, nel nostro caso possiamo trovarci di fronte alle seguenti situazioni. La matrice di cointegrazione è nota, e di conseguenza anche il suo rango. Questa è, ad esempio, l'ipotesi dell'implementazione pratica che esamineremo in seguito, in cui il vettore di cointegrazione è ipotizzato essere lo spread, quindi un vettore (1,-1) in quanto assumiamo che la teoria delle aspettative sia valida. Gli altri due casi sono più complicati; potremmo conoscere il rango ma non la matrice, o addirittura non conoscere entrambi. Per l'ultima situazione, che è la più generale, si fa ricordo alla procedura di Johansen, in cui si suppone che il sistema cointegrato possa essere rappresentato come un VAR di ordine finito con errori gaussiani. Questa procedura è abbastanza complicata e, non essendo oggetto del successivo lavoro, si è deciso solamente di presentarla. Per qualsiasi approfondimento si può far ricorso alla rispettiva letteratura. Esistono delle procedure alternative per stimare la matrice di cointegrazione quando questa non è nota. Una di queste è quella di Engle e Granger, esattamente quella che abbiamo utilizzato nell'analisi. Per comprendere meglio, consideriamo un semplice caso univariato $y_t = \beta x_t + \epsilon_t$. Se y e x_t sono non stazionarie, ma esiste una loro combinazione lineare che è stazionaria, allora sono cointegrate. L'idea è, quindi,

quella di estrarre i residui $\epsilon_t = y_t - \beta x_t$ e, tramite un test ADF, verificare che siano stazionari. In particolare, lo stimatore OLS di β è superconsistente, nel senso che tende al vero valore del parametro più velocemente rispetto al caso normale di OLS con variabili stazionarie. Verificata la stazionarietà dei residui, si possono inserire all'interno del modello a correzione dell'errore, ritardati di un periodo, sostituendoli alla componente $(y_{1t-1} - \beta_2 y_{2t-1})$ dell'eq.2.13. Tuttavia, la distribuzione di $\hat{\beta}$ non è gaussiana, ma dipende dalle ipotesi sul termine di errore che, comunque, deve essere stazionario, e quindi non si riesce a costruire delle statistiche test opportune. La soluzione potrebbe essere quella di usare delle regressioni aumentate per stimare il vettore di cointegrazione, in modo tale da rendere efficiente lo stimatore di β . In pratica, la distribuzione dello stimatore può essere approssimato alla normale nel caso in cui i termini di errore sono incorrelati. Per questo, si ricorre spesso allo stimatore DOLS (Dynamic OLS), proposto in Stock and Watson (1993), in cui vengono aggiunti regressori, come le variazioni delle variabili ritardate e a periodo successivo, fino a rendere incorrelati i residui. Sono esattamente le componenti aggiuntive dell'equazioni 2.13 e 2.14. Va sottolineato, come aggiungendo regressori, i parametri da stimare non cambiano; nel nostro caso rimane β . In generale, i residui del modello rimangono autocorrelati nonostante l'utilizzo del DOLS. Occorre quindi utilizzare uno stimatore consistente per gli errori standard che tenga conto dell'autocorrelazione e dell'eteroschedasticità (vedi, ad esempio, Newey and West (1986)).

Capitolo 3

Implicazioni della teoria delle aspettative

Come evidenziato nell'equazione 1.4, la forma classica della teorie delle aspettative, considerata in termini di spread, può essere scritta come segue:

$$i_{m,t} - i_{1,t} = \frac{1}{m} \left[\sum_{s=1}^{m-1} \sum_{j=1}^s E_t \Delta i_{1,t+j} \right] + r_{m,t}$$

L'eq. 1.4 implica una caratteristica importante. Se i rendimenti sono integrati di ordine uno, ovvero sono processi non stazionari, e il premio per il rischio è stazionario, allora lo spread dovrebbe essere stazionario. Il che equivale ad affermare che il rendimento di un'obbligazione a m -periodi e il rendimento di una a 1 periodo dovrebbero essere cointegrati, con il vettore di cointegrazione $(1, -1)$. Questa aspettativa potrebbe fallire per due ragioni: primo, il premio al rischio potrebbe essere non stazionario, o presentare elementi di non linearità¹; il secondo motivo nasce dalla presenza della doppia sommatoria, la quale implica che all'aumentare della maturity m , lo spread $i_{m,t} - i_{1,t}$ si comporta come un processo non stazionario.

Consideriamo innanzitutto la singola sommatoria di m -termini costanti:

¹Vedi Engle et al. (1987) e Clarida et al. (2006) per casi più generali in termini di premio al rischio

$$\frac{1}{m} \sum_{s=1}^m c = c, \quad (3.1)$$

che è finita in quanto lo scalare $1/m$ compensa l'aumento nel numeratore. Consideriamo ora lo stesso scalare applicato alla doppia sommatoria:

$$\frac{1}{m} \sum_{s=1}^{m-1} \sum_{j=1}^s c = \frac{1}{m} [c + (c + c) + \cdots + (c + c + \cdots + c)] \quad (3.2)$$

$$= \frac{c}{m} [1 + 2 + \cdots + (m - 1)] \quad (3.3)$$

$$= \frac{m(m - 1)c}{2m} \quad (3.4)$$

Notiamo come l'ultimo termine diverge quando $m \rightarrow \infty$. In questo caso, lo scalare $1/m$ non funziona, rendendone necessario uno differente.

Assumiamo ora che $\sum_{j=1}^{\infty} E_t \Delta i_{1,t+j} < \infty$.² Ovvero creiamo un sistema di aspettative che ha un termine m preciso, finito. Quando $m \rightarrow \infty$, allora:

$$\lim_{m \rightarrow \infty} i_{m,t} - i_{1,t} = \sum_{j=1}^{\infty} E_t \Delta i_{1,t+j} + r_{m,t}, \quad (3.5)$$

la cui prova è in appendice. Ciò implica che, al limite, lo spread $i_{m,t} - i_{1,t}$ ha una rappresentazione moving average di ordine infinito, $MA(\infty)$. Inoltre, i coefficienti di ogni termine di errore della serie sono uguali a uno, rendendo sempre persistente l'effetto dello shock, ovvero esso non ha la tendenza a scomparire. In altre parole la non stazionarietà è data dalla non assoluta sommabilità dei coefficienti del processo. Nel caso più semplice, in cui i termini $E_t \Delta i_{1,t+j}$ sono white noise, la non stazionarietà può essere vista osservando che la varianza di $\sum_{j=1}^{\infty} E_t \Delta i_{1,t+j}$ è illimitata. In particolare, quando il premio al rischio è costante e $E_t \Delta i_{1,t+j}$ è un white noise con varianza unitaria, la varianza dello spread $i_{1,t+j} + r_{m,t}$, per un valore di m finito

²"Proposizione Limitante" di Li and Davis (2017)

risulta essere³:

$$\text{var}(i_{m,t} - i_{1,t}) = \sum_{j=1}^{m-1} \left(1 - \frac{j}{m}\right)^2 = \frac{m}{3} - \frac{1}{2} + \frac{1}{6m} \quad (3.6)$$

Si noti come per $m \rightarrow \infty$, la varianza esplode all'infinito. Quindi, anche nel caso di un premio al rischio stazionario, il risultato non cambia. Inoltre, il termine $E_t \Delta i_{1,t+j}$ può anche non essere un white noise; può possedere un alto grado di dipendenza temporale, o essere eteroschedastico, e il risultato rimane invariato. Tutto ciò implica che, se utilizziamo lo spread come termine di correzione dell'errore, cioè testando la stazionarietà dello stesso per verificare la cointegrazione, all'aumentare di m le probabilità di rifiutare l'ipotesi nulla di non cointegrazione decrescono, diventando zero nel limite. A livello empirico, il legame tra la cointegrazione e l'EH si indebolisce quindi all'aumentare di m e, almeno per gli spread a lungo termine, ciò rende inappropriati i test di cointegrazione dell'EH. In effetti, l'incapacità di trovare cointegrazione all'aumentare della differenza di termine è prevista dall'EH, risultato che può spiegare come, in lavori precedenti, si trovava spesso cointegrazione solo quando la differenza di termine era relativamente breve.⁴ Abbiamo detto in precedenza come l'ipotesi di aspettative pure si basa su forti assunzioni in merito al premio a termine, per cui si rende necessaria una migliore interpretazione di $r_{m,t}$, che tenga conto di premi e costi associati alle ipotesi di segmentazione del mercato o di habitat preferito. Essi possono essere determinati, ad esempio, dalla preferenza per la liquidità, dalla corrispondenza con la durata o da vincoli istituzionali e riflettono la mancanza di sostituibilità tra obbligazioni a lungo e a breve termine. Per spread relativamente brevi, può prevalere il legame guidato dalla sostituibilità tra obbligazioni a lungo e a breve termine, enfatizzato dall'ipotesi di aspettative. Tuttavia, la proposizione limitante implica che questo legame si indebolisce all'au-

³ $\sum_{j=1}^{m-1} \left(1 - \frac{j}{m}\right)^2 = (m-1) - \frac{2 \sum_{j=1}^{m-1} j}{m} + \frac{\sum_{j=1}^{m-1} j^2}{m^2} = \frac{m}{3} - \frac{1}{2} + \frac{1}{6m}$ in Hamilton (2020)

⁴Campbell and Shiller (1991) e Sarno and Thornton (2003)

mentare di m , per cui alla fine dominano le forze di segmentazione. In altre parole, nel limite i legami a breve e a lungo termine non sono sostituibili.

Capitolo 4

Analisi empirica

4.1 Dati

Si è deciso di utilizzare osservazioni mensili dei rendimenti a scadenza (costanti) dei bond americani a 1-anno (T1), 3-anni (T3), 5-anni (T5), 10-anni (T10), e il tasso effettivo dei fondi federali (Effective Federal Fund Rate, FF), come proxy per il tasso a breve.¹ Il campione è costituito da 240 osservazioni, ovvero 20 anni di dati osservati da dicembre 2003 a novembre 2003. Nel corso della nostra analisi, si è preferito fare uso di regressioni basate su finestre di 72 osservazioni e aggiornate sempre di un periodo. L'utilizzo delle finestre è importante per due motivi. Primo, ci consente di avere un numero di risultati sufficienti per ottenere delle frequenze di rifiuto dell'ipotesi di cointegrazione. Inoltre, le regressioni rolling riescono a tenere conto di eventuali instabilità e cambi strutturali nelle relazioni di lungo periodo.

4.2 Test di Cointegrazione

Per verificare le implicazioni precedentemente presentate, ipotizziamo innanzitutto che il vettore di cointegrazione sia $(1, -1)$ e verifichiamo la stazionarietà dello spread

¹I dati sono stati scaricati dal sito "Federal Reserve Economic Data" della Fed di St.Louis: <http://research.stlouisfed.org/fred2/>.

a seconda della lunghezza del termine.² Indicheremo con y_t il tasso a lungo e con x_t quello a breve. Chiaramente, la classificazione è relativa, nel senso che il tasso a 5 anni, ad esempio, sarà considerato a lungo nei confronti di FF, T1 e T3, mentre è di breve termine nei confronti di T10. Lo spread è quindi definito come: $sp_t = y_t - x_t$. Innanzitutto, siamo partiti dall'analisi della serie dei diversi tassi di interesse per verificare la non stazionarietà, e quindi la presenza di radici unitarie. Abbiamo quindi proceduto ad effettuare due tipi di test: l'Augmented Dickey-Fuller test (ADF) e il test Kwiatkowski-Phillips-Schmidt-Shin (KPSS). Entrambi verificano la stazionarietà di una serie storica, con la differenza che nell'ADF l'ipotesi nulla è rappresentata dalla non stazionarietà, mentre la stessa nel KPSS rappresenta la stazionarietà. Tipicamente, è sufficiente eseguire un test ADF per capire se una serie è stazionaria o meno. L'aggiunta del KPSS è una conferma ulteriore, in quanto ci permette di tenere in considerazione sia l'errore del primo che del secondo tipo. In statistica inferenziale, l'errore del primo tipo si ha quando si rifiuta l'ipotesi nulla quando questa è vera; viceversa, l'errore del secondo tipo consiste nell'accettare l'ipotesi nulla quando questa è falsa. Ci aspetteremo, quindi, che se un test ADF rifiuta l'ipotesi nulla di non stazionarietà, il test KPSS relativo dovrebbe portare a non rifiutare l'ipotesi nulla di stazionarietà. Rappresentiamo nelle figure successive l'andamento dei rendimenti considerati; già dall'analisi dei grafici possiamo notare la probabile non stazionarietà delle serie, nonostante non siano presenti chiare componenti tendenziali.

²Simili analisi sono state effettuate da Hall et al. (1992) e Engsted and Tanggaard (1994) i quali usano, a differenza dei precedenti, la procedura di Johansen.

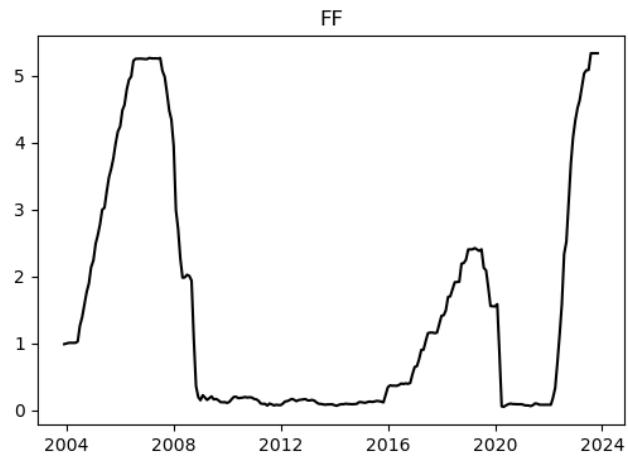


Figura 4.1: FF

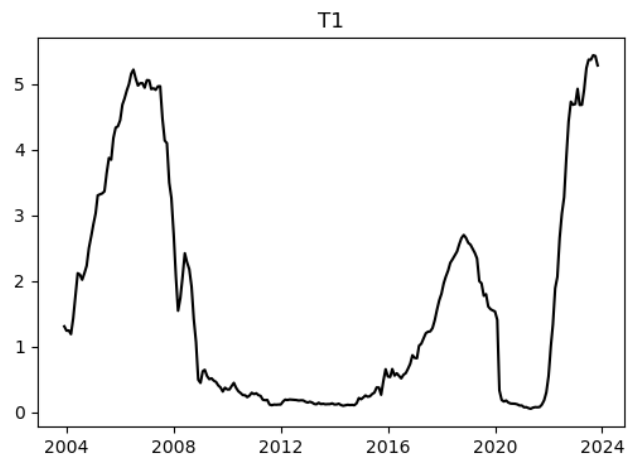


Figura 4.2: T1

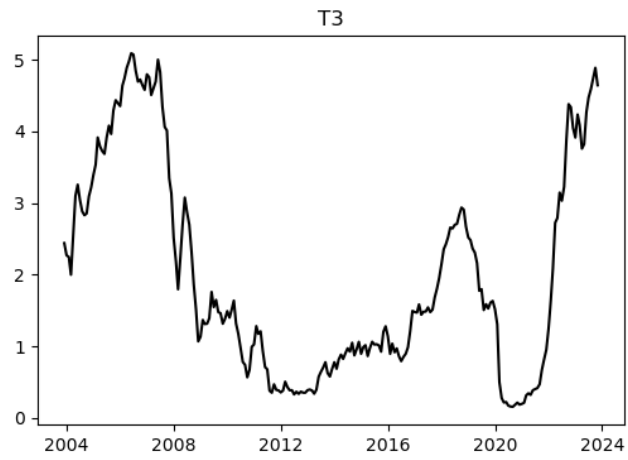


Figura 4.3: T3

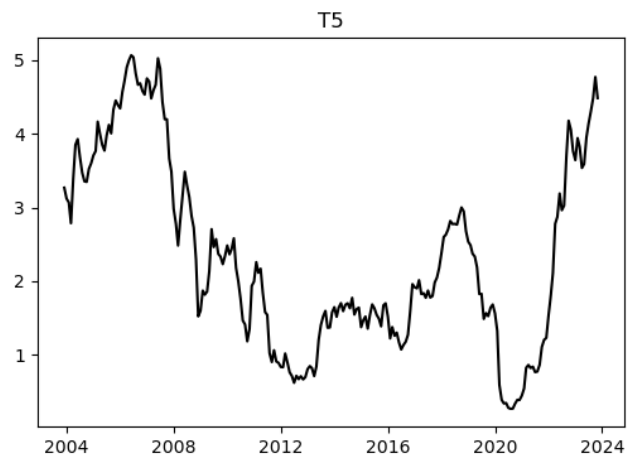


Figura 4.4: T5

Serie	n	ADF	KPSS
FF	240	0.031145	0.094947
T1	240	0.053755	0.091251
T3	240	0.682590	0.041627
T5	240	0.564345	0.013729
T10	240	0.361427	0.010000

Tabella 4.1: % Risultati Test ADF e KPSS

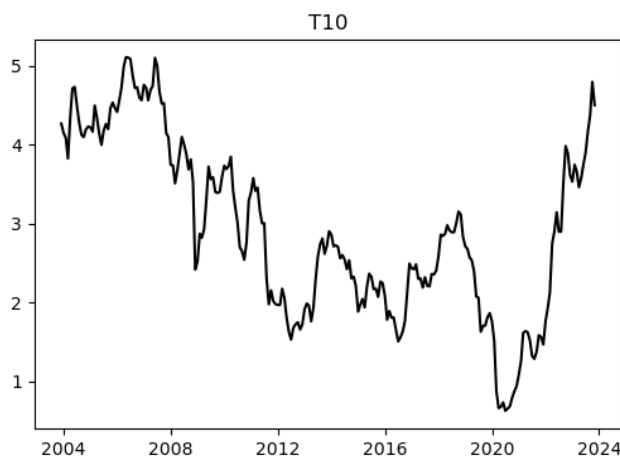


Figura 4.5: T10

Nella Tabella 4.1 sono rappresentati i principali risultati dei test sulle serie storiche dei tassi. Ricordiamo come un p-value minore di 0.05 per un test ADF ci porta a rifiutare l'ipotesi nulla di non stazionarietà della serie, viceversa per il test KPSS. Possiamo subito notare come per i tassi a più breve termine, come FF e T1, i test portano a conclusioni più incerte. Considerando un livello di confidenza del 5% dovremmo essere portati a dire che la serie FF sia stazionaria, in quanto per il test ADF dovremmo rifiutare l'ipotesi nulla di non stazionarietà, risultato condiviso dal test KPSS che ci porta a non rifiutare l'ipotesi nulla di stazionarietà. Al tempo stesso, analizzando il grafico, tendiamo ad escludere questa ipotesi. Considerando, infatti, un livello di confidenza dell'1%, la serie risulta non stazionaria. Questo ci fa capire come la potenza di questi test può talvolta essere limitata da vari fattori,

FF è tasso a breve				Treasury spreads					
T1- FF	T3- FF	T5- FF	T10- FF	T3-T1	T5-T3	T5-T1	T10- T5	T10- T3	T10- T1
85	91	96	96	93	94	97	91	95	98

Tabella 4.2: % no cointegrazione usando spread

come presenza di eteroschedasticità. Detto ciò, confrontando il grafico e i risultati del test, possiamo affermare come la serie FF, nel nostro campione di osservazione, sembra comportarsi in maniera più coerente con l'ipotesi di non stazionarietà. Anche per quanto riguarda T1 la situazione sembra essere "al limite", in quanto già ad un livello del 5% siamo in grado di non rifiutare l'ipotesi di non stazionarietà, nonostante il test KPSS ci dia un risultato contrastante. Per gli altri tassi i risultati sembrano nettamente più chiari. Concludiamo, quindi, che le serie dei rendimenti si comportino come processi non stazionari, come spesso fanno la maggior parte delle variabili economiche.

Come accennato in precedenza, assumendo che il vettore di cointegrazione sia determinato come (1,-1), il test di cointegrazione consiste in un test di radice sullo spread nell'equazione 1.4 per verificarne la stazionarietà. In pratica, la regressione risulta la seguente:

$$\Delta sp_t = \gamma_0 + \gamma_1 sp_{t-1} + v_t \quad (4.1)$$

Si è proceduto, quindi, a svolgere un classico test ADF (in modalità rolling) sulla serie storica dei vari spread.

La Tabella 4.2 riporta la percentuale di volte per le quali l'ipotesi nulla di non stazionarietà non può essere rifiutata al livello del 5%, equivalente del dire che non ci sono evidenze di cointegrazione e che quindi non vi è supporto a favore dell'ipotesi di aspettative. Possiamo notare come tutte le percentuali siano alte, sopra l'80%. Questi risultati sono in linea con ciò che prevede l'ipotesi di aspettative, anche considerando le percentuali al crescere del termine.

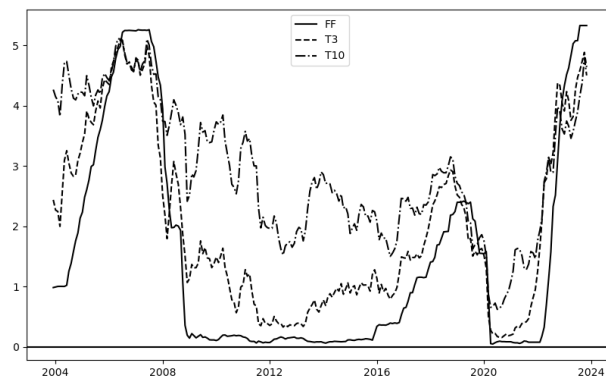


Figura 4.6: FF-T3-T10

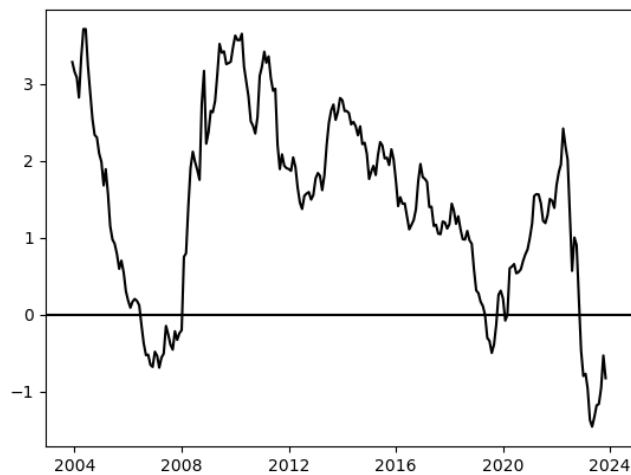


Figura 4.7: Spread T10-FF

In Figura 4.6 è rappresentato l'andamento di T10, T3 e del tasso sui fondi federali. In linea di massima, le serie sembrano muoversi insieme. Tuttavia, possiamo notare come il legame sia visibilmente debole negli anni successivi alla crisi del 2008. Il tasso di policy della Fed è stato ancorato sui livelli dello zero fino al 2016, mentre i tassi a 3 e 10 anni sembrano non seguire la stessa dinamica. Nella Figura 4.7 osserviamo, invece lo spread tra il tasso a 10 anni e il Federal Fund rate. L'alternanza di trend crescenti e decrescenti ci conferma la non stazionarietà del processo.

4.3 ECM non ristretto

La costruzione dell'EH definita in questo modo limita il vettore di cointegrazione tra i diversi tassi alla forma (1,-1). La presenza di un premio al rischio non stazionario, o comunque che possa variare col tempo, rende questa implicazione abbastanza vincolante. Si è deciso, quindi, di rilassare questa ipotesi e lasciare che il vettore di cointegrazione abbia una forma più generale. In particolare, nel nostro sistema bivariato $[i_{m,t}, i_{1,t}]$, il vettore di cointegrazione è $(1, \lambda)$, con $\lambda = \frac{\theta}{1-\theta}$, se:

$$i_{m,t} - \theta \text{spread}_{m,t} = i_{m,t} - \theta(i_{m,t} - i_{1,t}) = (1 - \theta)i_{m,t} + \theta i_{1,t}$$

è stazionario, o integrato di ordine 0. Stimare un modello a correzione dell'errore di questo tipo equivale a seguire la procedura di Engle e Granger, ovvero stimare una regressione tra il tasso a lungo e quello a breve, estrarre i residui verificandone la stazionarietà per testare la cointegrazione; successivamente, gli stessi residui (ritardati di un periodo) rappresenteranno il termine di correzione dell'errore, il cui coefficiente dovrà essere statisticamente diverso da 0 per validare l'ECM.³ In pratica, la regressione è definita come:

$$y_t = \beta_0 + \beta_1 x_t + e_t, \tag{4.2}$$

il cui residuo stimato sarà:

$$\hat{e}_t = y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t. \tag{4.3}$$

La Tabella 4.3 riporta i risultati del test ADF sui residui; è sovrapponibile alla Tabella 4.2 e le percentuali risultano ancora abbastanza alte. Tuttavia, esse sono più

³Nel nostro sistema bivariato avremo due ECM, uno per il tasso a breve e uno per il tasso a lungo. Il fatto che la Fed si muova prima del mercato nel settare i suoi obiettivi di politica monetaria, tramite una variazione di FF, il tasso a breve sarà debolmente esogeno. Consideriamo quindi il solo ECM in la variabile dipendente è il tasso a lungo.

FF è tasso a breve				Treasury spreads					
T1- FF	T3- FF	T5- FF	T10- FF	T3-T1	T5-T3	T5-T1	T10- T5	T10- T3	T10- T1
77	76	90	79	80	73	84	79	82	77

Tabella 4.3: % no cointegrazione usando residui

basse per tutte le differenze di termine e anche il loro crescere all'aumentare dello spread diventa meno evidente. I tassi di rifiuto più bassi non dovrebbero sorprendere più di tanto, in quanto la stima OLS cerca di trovare un residuo stazionario e ponendo meno restrizioni vi riesce più spesso. Nella Figura 4.8 è rappresentato il grafico con sovrapposte le serie dello Spread tra il tasso a 10 anni e il Federal Fund rate, e quella relativa ai residui del modello (regressione) creato con gli stessi tassi. L'andamento sembra essere lo stesso solamente negli ultimi anni, mentre prima e durante la crisi, le serie sembrano comportarsi in maniera opposta.

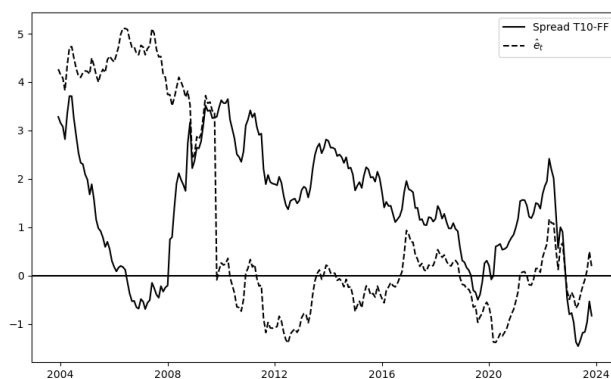


Figura 4.8: Spread T10-FF e Residui T10-FF

Una volta estratti i residui dalla regressione, essi vengono quindi posti all'interno dell'ECM come termine di correzione dell'errore. In pratica, si crea un modello simile all'Eq.2.13. Il risultato è la regressione seguente:

$$\Delta y_t = \alpha_{02} + \alpha_2^* \hat{\epsilon}_{t-1} + v_{2t}, \quad (4.4)$$

FF è tasso a breve				Treasury spreads					
T1- FF	T3- FF	T5- FF	T10- FF	T3-T1	T5-T3	T5-T1	T10- T5	T10- T3	T10- T1
63	69	66	73	88	82	79	81	78	76

Tabella 4.4: % non significatività termine di correzione

nel quale si testa $H_0 : \alpha_2^* = 0$.

La Tabella 4.4 indica, come sempre, le percentuali di non rifiuto dell'ipotesi di non cointegrazione, costruite su valori di p-value del coefficiente del termine di correzione superiori al 5%. Infatti, se tale coefficiente risulta statisticamente non significativo, un modello costruito in questo modo non è validato, mettendo in dubbio anche la cointegrazione tra le variabili considerate. Anche in questo caso, dato l'utilizzo della stima OLS, esse sono inferiori rispetto al primo test. Tuttavia, sembrano ancora sufficienti a rigettare l'ipotesi di cointegrazione prevista dalla teoria delle aspettative

Infine, una possibile estensione dell'analisi è stata proposta da Engle et al. (1987). In pratica, come abbiamo accennato all'inizio, il fallimento della cointegrazione potrebbe essere causata sia dalla presenza della doppia sommatoria nell'Eq.1.4, che dalla non stazionarietà del premio al rischio. Si potrebbe quindi pensare un premio al rischio variabile nel tempo, considerando per lo stesso un modello di tipo GARCH-in-Mean(1,1), trasformando l'Eq.4.1 nella seguente:

$$\Delta sp_t = \phi_0 + \phi_1 sp_{t-1} + \phi_2 h_t + \epsilon_t, \quad (4.5)$$

dove $h_t^2 = \pi_0 + \pi_1 h_{t-1}^2 + \pi_2 sp_{t-1}^2$, ovvero la varianza condizionale dipende sia dai suoi valori ritardati che dallo spread stesso.

Conclusione

Capire la struttura a termine dei tassi di interesse è importante; essa è una componente essenziale del meccanismo di trasmissione della politica monetaria, è fondamentale nelle previsioni macroeconomiche e nelle strategie di investimento. La teoria delle aspettative, che sia la forma pura o meno, è uno degli approcci teorici maggiormente conosciuti che cerca di spiegare la dinamica della struttura a termine. Essa è spesso testata tramite tecniche di cointegrazione, ponendo condizioni stazionarie in alcune delle sue componenti, come il premio a termine. In questo lavoro si è cercato di dimostrare come test della teoria delle aspettative che pongono restrizioni di questo tipo, e che assumono la stazionarietà dello spread, potrebbero essere inadeguati. La non stazionarietà dello spread, soprattutto all'aumentare del termine, e quindi la minor probabilità di trovare supporto alla cointegrazione, sono previste dalla stessa teoria delle aspettative. Li and Davis (2017) hanno mostrato evidenze in questo senso anche tramite simulazioni Monte Carlo. Inoltre, abbiamo cercato di osservare come alcune delle serie storiche si sono comportate a livello grafico, confrontandole soprattutto con il tasso sui fondi federali. La dinamica sembra divergere soprattutto durante la crisi del 2008, a cui la Fed ha dovuto far fronte portando avanti politiche espansive, in termini di riduzione dei tassi, senza precedenti. Anche per questo motivo, si è deciso di eseguire l'intera analisi con l'utilizzo di finestre mobili, in modo tale da considerare i cambiamenti strutturali tipici degli ultimi anni. I risultati ottenuti sono interessanti, in quanto sembrano mettere in dubbio il legame tra tassi a breve e lungo termine, minacciando una componente

chiave della trasmissione della politica monetaria. La variazione del tasso di policy potrebbe avere effetti sui tassi a più breve scadenza, mentre quelli a lungo potrebbero non seguire la stessa dinamica. Struttura a termine e teoria delle aspettative sono influenzate da molteplici fattori; abbiamo visto come ipotesi di neutralità al rischio o di preferenze per la liquidità giocano un ruolo importante nella derivazione della teoria. Inoltre, potrebbero essere presi in considerazione altri aspetti. Bulkeley et al. (2015) hanno esteso l'analisi considerando bias comportali come legge dei grandi numeri e conservatorismo, che potessero spiegare il rigetto dell'ipotesi di aspettative. Lakdawala and Wu (2017) hanno mostrato, invece, come il grado di credibilità di una banca centrale, in particolare la Fed, può influenzare il livello, la curvatura e la pendenza della struttura a termine.

Appendice A

Proposizione Limitante

Possiamo trasformare l'Equazione 1.4 nel seguente modo:

$$i_{m,t} - i_{1,t} = \frac{1}{m} \left[\sum_{s=1}^{m-1} \sum_{j=1}^s E_t \Delta i_{1,t+j} \right] + r_m \quad (\text{A.1})$$

$$= \frac{1}{m} [E_t \Delta i_{1,t+1} + (E_t \Delta i_{1,t+1} + E_t \Delta i_{1,t+2}) + \dots] + r_m \quad (\text{A.2})$$

$$= \frac{1}{m} [(m-1)E_t \Delta i_{1,t+1} + (m-2)E_t \Delta i_{1,t+2} + \dots + E_t \Delta i_{1,t+m-1}] + r_m \quad (\text{A.3})$$

$$= \sum_{j=1}^{m-1} \left(1 - \frac{j}{m}\right) E_t \Delta i_{1,t+j} + r_m \quad (\text{A.4})$$

$$= \sum_{j=1}^{m-1} E_t \Delta i_{1,t+j} - \frac{1}{m} \sum_{j=1}^{m-1} j E_t \Delta i_{1,t+j} + r_m \quad (\text{A.5})$$

Per ipotesi $\sum_{j=1}^{m-1} E_t \Delta i_{1,t+j}$ è finita, e quindi è una sequenza di Cauchy.¹ Ne consegue che:

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{j=1}^{m-1} j E_t \Delta i_{1,t+j} = 0 \quad (\text{A.6})$$

Il risultato è ottenuto prendendo il limite da entrambe le parti dell'Eq.A.5

¹Hayashi (2000)

Appendice B

Codici Python per l'analisi empirica

```
1 import pandas as pd
2 import numpy as np
3 import scipy
4 import statsmodels.api as sm
5 import seaborn as sns
6 import matplotlib.pyplot as plt
7 from statsmodels.tsa.stattools import adfuller
8
9 #Scarico le serie storiche dei tassi
10 EFFF = pd.read_excel('FF.xlsx', index_col=0) ['FF']
11 T1 = pd.read_excel('T-1.xlsx', index_col=0) ['DGS1']
12 T3 = pd.read_excel('T-3.xlsx', index_col=0) ['DGS3']
13 T5 = pd.read_excel('T-5.xlsx', index_col=0) ['DGS5']
14 T10 = pd.read_excel('T-10.xlsx', index_col=0) ['DGS10']
15
16 #Grafici
17 plt.figure(figsize=(6, 4))
18 plt.plot(T1, color='black')
19 plt.title('T1')
20 plt.savefig('T1n.png')
21 plt.show
22
```

```
23 plt.figure(figsize=(6, 4))
24 plt.plot(T3, color='black')
25 plt.title('T3')
26 plt.savefig('T3n.png')
27 plt.show
28
29 plt.figure(figsize=(6, 4))
30 plt.plot(T5, color='black')
31 plt.title('T5')
32 plt.savefig('T5n.png')
33 plt.show
34
35 plt.figure(figsize=(6, 4))
36 plt.plot(T10, color='black')
37 plt.title('T10')
38 plt.savefig('T10n.png')
39 plt.show
40
41 plt.figure(figsize=(6, 4))
42 plt.plot(EFFR, color='black')
43 plt.title('FF')
44 plt.savefig('FF1.png')
45 plt.show
46
47 #Definisco la funzione per il test ADF
48 from statsmodels.tsa.stattools import adfuller
49 def adf_test(timeseries):
50     print("Results of Dickey-Fuller Test:")
51     dftest = adfuller(timeseries, regression='c',
52                       autolag='AIC')
53     dfoutput = pd.Series(
54         dftest[0:4],
55         index=[
```

```

56         "Test Statistic",
57         "p-value",
58         "Lags Used",
59         "Number of Observations Used",
60     ],
61 )
62 for key, value in dftest[4].items():
63     dfoutput["Critical Value (%s)" % key] = value
64 print(dfoutput)
65
66 #Definisco la funzione per il test KPSS
67 from statsmodels.tsa.stattools import kpss
68 def kpss_test(timeseries):
69     print("Results of KPSS Test:")
70     kpsstest = kpss(timeseries, regression="c",
71                    nlags="auto")
72     kpss_output = pd.Series(
73         kpsstest[0:3], index=["Test Statistic",
74                             "p-value", "Lags Used"]
75     )
76     for key, value in kpsstest[3].items():
77         kpss_output["Critical Value (%s)" % key] = value
78     print(kpss_output)
79
80 #Eseguo i test per le variabili
81 adf_test(EFFR)
82 kpss_test(EFFR)
83
84 adf_test(T1)
85 kpss_test(T1)
86
87 adf_test(T3)
88 kpss_test(T3)

```

```

89
90 adf_test(T5)
91 kpss_test(T5)
92
93 adf_test(T10)
94 kpss_test(T10)
95
96 #Costruisco gli spread per ogni tasso lungo/breve
97 Spread1 = T1 - EFFR
98 Spread3 = T3 - EFFR
99 Spread5 = T5 - EFFR
100 Spread10 = T10 - EFFR
101
102 Spread3_1 = T3 - T1
103 Spread5_3 = T5 - T3
104 Spread5_1 = T5 - T1
105 Spread10_5 = T10 - T5
106 Spread10_3 = T10 - T3
107 Spread10_1 = T10 - T1
108
109 #Definisco una funzione per eseguire un test ADF rolling sullo spread.
110 def rolling_adf_test(time_series, window_size, shift_step):
111     results = []
112     for i in range(0, len(time_series) - window_size+1, shift_step):
113         window = time_series[i:i + window_size]
114         result = adfuller(window)
115         p_value = result[1]
116         results.append(p_value)
117     return results
118 window_size = 72
119 shift_step = 1
120
121 #Eseguo il test di cui salvo i pvalues.

```

```
122 p_values1 = rolling_adf_test(Spread1, 72, 1)
123 df1 = pd.DataFrame(p_values1)
124 df1.to_excel('pvalues1.xlsx', index=False)
125
126 p_values3 = rolling_adf_test(Spread3, 72, 1)
127 df3 = pd.DataFrame(p_values3)
128 df3.to_excel('pvalues3.xlsx', index=False)
129
130 p_values5 = rolling_adf_test(Spread5, 72, 1)
131 df5 = pd.DataFrame(p_values5)
132 df5.to_excel('pvalues5.xlsx', index=False)
133
134 p_values10 = rolling_adf_test(Spread10, 72, 1)
135 df10 = pd.DataFrame(p_values10)
136 df10.to_excel('pvalues10.xlsx', index=False)
137
138 p_values3_1= rolling_adf_test(Spread3_1, 72, 1)
139 df3_1 = pd.DataFrame(p_values3_1)
140 df3_1.to_excel('pvalues3_1.xlsx', index=False)
141
142 p_values5_3= rolling_adf_test(Spread5_3, 72, 1)
143 df5_3 = pd.DataFrame(p_values5_3)
144 df5_3.to_excel('pvalues5_3.xlsx', index=False)
145
146 p_values5_1= rolling_adf_test(Spread5_1, 72, 1)
147 df5_1 = pd.DataFrame(p_values5_1)
148 df5_1.to_excel('pvalues5_1.xlsx', index=False)
149
150 p_values10_5= rolling_adf_test(Spread10_5, 72, 1)
151 df10_5 = pd.DataFrame(p_values10_5)
152 df10_5.to_excel('pvalues10_5.xlsx', index=False)
153
154 p_values10_3= rolling_adf_test(Spread10_3, 72, 1)
```

```

155 df10_3 = pd.DataFrame(p_values10_3)
156 df10_3.to_excel('pvalues10_3.xlsx', index=False)
157
158 p_values10_1= rolling_adf_test(Spread10_1, 72, 1)
159 df10_1 = pd.DataFrame(p_values10_1)
160 df10_1.to_excel('pvalues10_1.xlsx', index=False)
161
162 #Procedura Engle e Granger
163 #T1-FF
164 window_size = 72
165 shift_step = 1
166 results = []
167 p_values_error_correction = []
168
169 for i in range(0, len(T1) - window_size + 1, shift_step):
170     T1_window = T1[i:i + window_size]
171     EFFR_window = EFFR[i:i + window_size]
172     model_ols = sm.OLS(T1_window, sm.add_constant(EFFR_window)).fit()
173     residlr = model_ols.resid
174
175     result = adfuller(residlr)
176     p_value = result[1]
177     results.append(p_value)
178
179     df1r = pd.DataFrame(results)
180     df1r.to_excel('pvalues1r.xlsx', index=False)
181
182
183     ecm_term1 = pd.Series(residlr).shift(1)
184     ecm_x = sm.add_constant(EFFR_window.diff().shift(1))
185     ecm_y = T1_window.diff()
186     ecm_x['error_correction'] = ecm_term1
187

```



```

188 ecm_data = pd.concat([ecm_y, ecm_x], axis=1).dropna()
189 ecm_y = ecm_data.iloc[:, 0]
190 ecm_x = ecm_data.iloc[:, 1:]
191
192 ecm_model = sm.OLS(ecm_y, ecm_x).fit()
193
194 p_values_error_correction.append(ecm_model.pvalues['
error_correction'])
195
196 df_results = pd.DataFrame({'P-value_error_correction':
p_values_error_correction})
197 df_results.to_excel('p_values_error_correction1q.xlsx', index=
False)
198
199 #T3-FF
200 window_size = 72
201 shift_step = 1
202 results = []
203 p_values_error_correction = []
204
205 for i in range(0, len(T3) - window_size + 1, shift_step):
206     T3_window = T3[i:i + window_size]
207     EFFR_window = EFFR[i:i + window_size]
208     model_ols = sm.OLS(T3_window, sm.add_constant(EFFR_window)).fit()
209     resid3r = model_ols.resid
210
211     result = adfuller(resid3r)
212     p_value = result[1]
213     results.append(p_value)
214
215 df3r = pd.DataFrame(results)
216 df3r.to_excel('pvalues3r.xlsx', index=False)
217

```

```

218 ecm_term3 = pd.Series(resid3r).shift(1)
219 ecm_x = sm.add_constant(EFFR_window.diff().shift(1))
220 ecm_y = T3_window.diff()
221 ecm_x['error_correction'] = ecm_term3
222
223 ecm_data = pd.concat([ecm_y, ecm_x], axis=1).dropna()
224 ecm_y = ecm_data.iloc[:, 0]
225 ecm_x = ecm_data.iloc[:, 1:]
226
227 ecm_model3 = sm.OLS(ecm_y, ecm_x).fit()
228
229 p_values_error_correction.append(ecm_model3.pvalues['
error_correction'])
230
231 df_results = pd.DataFrame({'P-value_error_correction':
p_values_error_correction})
232 df_results.to_excel('p_values_error_correction3q.xlsx', index=
False)
233
234 #T5-FF
235 window_size = 72
236 shift_step = 1
237 results = []
238 p_values_error_correction = []
239
240 for i in range(0, len(T5) - window_size + 1, shift_step):
241     T5_window = T5[i:i + window_size]
242     EFFR_window = EFFR[i:i + window_size]
243     model_ols = sm.OLS(T5_window, sm.add_constant(EFFR_window)).fit()
244     resid5r = model_ols.resid
245
246     result = adfuller(resid5r)
247     p_value = result[1]

```

```

248     results.append(p_value)
249
250     df5r = pd.DataFrame(results)
251     df5r.to_excel('pvalues5r.xlsx', index=False)
252
253     ecm_term5 = pd.Series(resid5r).shift(1)
254     ecm_x = sm.add_constant(EFFR_window.diff().shift(1))
255     ecm_y = T5_window.diff()
256     ecm_x['error_correction'] = ecm_term5
257
258     ecm_data = pd.concat([ecm_y, ecm_x], axis=1).dropna()
259     ecm_y = ecm_data.iloc[:, 0]
260     ecm_x = ecm_data.iloc[:, 1:]
261
262     ecm_model5 = sm.OLS(ecm_y, ecm_x).fit()
263
264     p_values_error_correction.append(ecm_model5.pvalues['
error_correction'])
265
266     df_results = pd.DataFrame({'P-value_error_correction':
p_values_error_correction})
267     df_results.to_excel('p_values_error_correction5q.xlsx', index=
False)
268
269 #T10-FF
270 window_size = 72
271 shift_step = 1
272 results = []
273 p_values_error_correction = []
274
275 for i in range(0, len(T10) - window_size + 1, shift_step):
276     T10_window = T10[i:i + window_size]
277     EFFR_window = EFFR[i:i + window_size]

```

```

278     model_ols = sm.OLS(T10_window, sm.add_constant(EFFR_window)).fit()
279     resid10r = model_ols.resid
280
281     result = adfuller(resid10r)
282     p_value = result[1]
283     results.append(p_value)
284
285     df10r = pd.DataFrame(results)
286     df10r.to_excel('pvalues10r.xlsx', index=False)
287
288     ecm_term10 = pd.Series(resid10r).shift(1)
289     ecm_x = sm.add_constant(EFFR_window.diff().shift(1))
290     ecm_y = T10_window.diff()
291     ecm_x['error_correction'] = ecm_term10
292
293     ecm_data = pd.concat([ecm_y, ecm_x], axis=1).dropna()
294     ecm_y = ecm_data.iloc[:, 0]
295     ecm_x = ecm_data.iloc[:, 1:]
296
297     ecm_model10 = sm.OLS(ecm_y, ecm_x).fit()
298
299     p_values_error_correction.append(ecm_model10.pvalues['
error_correction'])
300
301     df_results = pd.DataFrame({'P-value_error_correction':
p_values_error_correction})
302     df_results.to_excel('p_values_error_correction10q.xlsx', index=
False)
303
304 #T3-T1
305 window_size = 72
306 shift_step = 1
307 results = []

```

```

308 p_values_error_correction = []
309
310 for i in range(0, len(T3) - window_size + 1, shift_step):
311     T3_window = T3[i:i + window_size]
312     T1_window = T1[i:i + window_size]
313     model_ols = sm.OLS(T3_window, sm.add_constant(T1_window)).fit()
314     resid31r = model_ols.resid
315
316     result = adfuller(resid31r)
317     p_value = result[1]
318     results.append(p_value)
319
320     df31r = pd.DataFrame(results)
321     df31r.to_excel('pvalues31r.xlsx', index=False)
322
323     ecm_term31 = pd.Series(resid31r).shift(1)
324     ecm_x = sm.add_constant(T1_window.diff().shift(1))
325     ecm_y = T3_window.diff()
326     ecm_x['error_correction'] = ecm_term31
327
328     ecm_data = pd.concat([ecm_y, ecm_x], axis=1).dropna()
329     ecm_y = ecm_data.iloc[:, 0]
330     ecm_x = ecm_data.iloc[:, 1:]
331
332     ecm_model31 = sm.OLS(ecm_y, ecm_x).fit()
333
334     p_values_error_correction.append(ecm_model31.pvalues['
error_correction'])
335
336     df_results = pd.DataFrame({'P-value_error_correction':
p_values_error_correction})
337     df_results.to_excel('p_values_error_correction31q.xlsx', index=
False)

```

```

338
339 #T5-T3
340 window_size = 72
341 shift_step = 1
342 results = []
343 p_values_error_correction = []
344
345 for i in range(0, len(T5) - window_size + 1, shift_step):
346     T5_window = T5[i:i + window_size]
347     T3_window = T3[i:i + window_size]
348     model_ols = sm.OLS(T5_window, sm.add_constant(T3_window)).fit()
349     resid53r = model_ols.resid
350
351     result = adfuller(resid53r)
352     p_value = result[1]
353     results.append(p_value)
354
355     df53r = pd.DataFrame(results)
356     df53r.to_excel('pvalues53r.xlsx', index=False)
357
358     ecm_term53 = pd.Series(resid53r).shift(1)
359     ecm_x = sm.add_constant(T3_window.diff().shift(1))
360     ecm_y = T5_window.diff()
361     ecm_x['error_correction'] = ecm_term53
362
363     ecm_data = pd.concat([ecm_y, ecm_x], axis=1).dropna()
364     ecm_y = ecm_data.iloc[:, 0]
365     ecm_x = ecm_data.iloc[:, 1:]
366
367     ecm_model53 = sm.OLS(ecm_y, ecm_x).fit()
368
369     p_values_error_correction.append(ecm_model53.pvalues['
error_correction'])

```

```

370
371     df_results = pd.DataFrame({'P-value_error_correction':
372     p_values_error_correction})
373
374     df_results.to_excel('p_values_error_correction53q.xlsx', index=
375     False)
376
377 #T5-T1
378 window_size = 72
379 shift_step = 1
380 results = []
381 p_values_error_correction = []
382
383 for i in range(0, len(T5) - window_size + 1, shift_step):
384     T5_window = T5[i:i + window_size]
385     T1_window = T1[i:i + window_size]
386     model_ols = sm.OLS(T5_window, sm.add_constant(T1_window)).fit()
387     resid51r = model_ols.resid
388
389     result = adfuller(resid51r)
390     p_value = result[1]
391     results.append(p_value)
392
393 df51r = pd.DataFrame(results)
394 df51r.to_excel('pvalues51r.xlsx', index=False)
395
396 ecm_term51 = pd.Series(resid51r).shift(1)
397 ecm_x = sm.add_constant(T1_window.diff().shift(1))
398 ecm_y = T5_window.diff()
399 ecm_x['error_correction'] = ecm_term51
400
401 ecm_data = pd.concat([ecm_y, ecm_x], axis=1).dropna()
402 ecm_y = ecm_data.iloc[:, 0]
403 ecm_x = ecm_data.iloc[:, 1:]

```

```

401
402     ecm_model51 = sm.OLS(ecm_y, ecm_x).fit()
403
404     p_values_error_correction.append(ecm_model51.pvalues['
error_correction'])
405
406     df_results = pd.DataFrame({'P-value_error_correction':
p_values_error_correction})
407     df_results.to_excel('p_values_error_correction51q.xlsx', index=
False)
408
409 #T10-T1
410 window_size = 72
411 shift_step = 1
412 results = []
413 p_values_error_correction = []
414
415 for i in range(0, len(T10) - window_size + 1, shift_step):
416     T10_window = T10[i:i + window_size]
417     T5_window = T5[i:i + window_size]
418     model_ols = sm.OLS(T10_window, sm.add_constant(T5_window)).fit()
419     resid105r = model_ols.resid
420
421     result = adfuller(resid105r)
422     p_value = result[1]
423     results.append(p_value)
424
425     df105r = pd.DataFrame(results)
426     df105r.to_excel('pvalues105r.xlsx', index=False)
427
428     ecm_term105 = pd.Series(resid105r).shift(1)
429     ecm_x = sm.add_constant(T5_window.diff().shift(1))
430     ecm_y = T10_window.diff()

```



```

431     ecm_x['error_correction'] = ecm_term105
432
433     ecm_data = pd.concat([ecm_y, ecm_x], axis=1).dropna()
434     ecm_y = ecm_data.iloc[:, 0]
435     ecm_x = ecm_data.iloc[:, 1:]
436
437     ecm_model105 = sm.OLS(ecm_y, ecm_x).fit()
438
439     p_values_error_correction.append(ecm_model105.pvalues['
error_correction'])
440
441     df_results = pd.DataFrame({'P-value_error_correction':
p_values_error_correction})
442     df_results.to_excel('p_values_error_correction105q.xlsx', index=
False)
443
444 #T10-T3
445 window_size = 72
446 shift_step = 1
447 results = []
448 p_values_error_correction = []
449
450 for i in range(0, len(T10) - window_size + 1, shift_step):
451     T10_window = T10[i:i + window_size]
452     T3_window = T3[i:i + window_size]
453     model_ols = sm.OLS(T10_window, sm.add_constant(T3_window)).fit()
454     resid103r = model_ols.resid
455
456     result = adfuller(resid103r)
457     p_value = result[1]
458     results.append(p_value)
459
460     df103r = pd.DataFrame(results)

```

```

461     df103r.to_excel('pvalues103r.xlsx', index=False)
462
463     ecm_term103 = pd.Series(resid103r).shift(1)
464     ecm_x = sm.add_constant(T3_window.diff().shift(1))
465     ecm_y = T10_window.diff()
466     ecm_x['error_correction'] = ecm_term103
467
468     ecm_data = pd.concat([ecm_y, ecm_x], axis=1).dropna()
469     ecm_y = ecm_data.iloc[:, 0]
470     ecm_x = ecm_data.iloc[:, 1:]
471
472     ecm_model103 = sm.OLS(ecm_y, ecm_x).fit()
473
474     p_values_error_correction.append(ecm_model103.pvalues['
error_correction'])
475
476     df_results = pd.DataFrame({'P-value_error_correction':
p_values_error_correction})
477     df_results.to_excel('p_values_error_correction103q.xlsx', index=
False)
478
479 #T10-T1
480 window_size = 72
481 shift_step = 1
482 results = []
483 p_values_error_correction = []
484
485 for i in range(0, len(T10) - window_size + 1, shift_step):
486     T10_window = T10[i:i + window_size]
487     T1_window = T1[i:i + window_size]
488     model_ols = sm.OLS(T10_window, sm.add_constant(T1_window)).fit()
489     resid101r = model_ols.resid
490

```

```

491     result = adfuller(resid101r)
492     p_value = result[1]
493     results.append(p_value)
494
495     df101r = pd.DataFrame(results)
496     df101r.to_excel('pvalues101r.xlsx', index=False)
497
498     ecm_term101 = pd.Series(resid101r).shift(1)
499     ecm_x = sm.add_constant(T10_window.diff().shift(1))
500     ecm_y = T10_window.diff()
501     ecm_x['error_correction'] = ecm_term101
502
503     ecm_data = pd.concat([ecm_y, ecm_x], axis=1).dropna()
504     ecm_y = ecm_data.iloc[:, 0]
505     ecm_x = ecm_data.iloc[:, 1:]
506
507     ecm_model101 = sm.OLS(ecm_y, ecm_x).fit()
508
509     p_values_error_correction.append(ecm_model101.pvalues['
error_correction'])
510
511     df_results = pd.DataFrame({'P-value_error_correction':
p_values_error_correction})
512     df_results.to_excel('p_values_error_correction101q.xlsx', index=
False)
513
514 #Grafici
515 plt.figure(figsize=(10, 6))
516 plt.plot(Spread10, label='Spread T10-FF', color="black")
517 plt.plot(resid10r, label= r'$\hat{e}_t$', color="black", linestyle
="--")
518 plt.legend()
519 plt.axhline(y=0, color="black")

```

```
520 plt.savefig('SpRes1.png')
521 plt.show()
522
523 plt.figure(figsize=(10, 6))
524 plt.plot(EFFR, label='FF', color="black")
525 plt.plot(T3, label='T3', color="black", linestyle="--")
526 plt.plot(T10, label='T10', color='black', linestyle="-.")
527 plt.legend()
528 plt.axhline(y=0, color="black")
529 plt.savefig('Confrontotassi.png')
530 plt.show()
531
532 plt.plot(Spread10, label='Spread T10-FF', color="black")
533 plt.axhline(y=0, color="black")
534 plt.savefig('Sp10-FF.png')
535 plt.show()
```

Bibliografia

- A. Banerjee, J. J. Dolado, J. W. Galbraith, and D. Hendry. *Co-integration, error correction, and the econometric analysis of non-stationary data*. Oxford university press, 1993.
- G. Bulkley, R. D. Harris, and V. Nawosah. Can behavioral biases explain the rejections of the expectation hypothesis of the term structure of interest rates? *Journal of Banking & Finance*, 58:179–193, 2015.
- J. Y. Campbell and R. J. Shiller. Cointegration and tests of present value models. *Journal of political economy*, 95(5):1062–1088, 1987.
- J. Y. Campbell and R. J. Shiller. Yield spreads and interest rate movements: A bird’s eye view. *The Review of Economic Studies*, 58(3):495–514, 1991.
- R. H. Clarida, L. Sarno, M. P. Taylor, and G. Valente. The role of asymmetries and regime shifts in the term structure of interest rates. *The Journal of Business*, 79(3):1193–1224, 2006.
- J. C. Cox, J. E. Ingersoll Jr, and S. A. Ross. A re-examination of traditional hypotheses about the term structure of interest rates. *The Journal of Finance*, 36(4):769–799, 1981.
- R. F. Engle and C. W. Granger. Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, pages 251–276, 1987.

- R. F. Engle, D. M. Lilien, and R. P. Robins. Estimating time varying risk premia in the term structure: The arch-m model. *Econometrica: journal of the Econometric Society*, pages 391–407, 1987.
- T. Engsted and C. Tanggaard. Cointegration and the us term structure. *Journal of Banking & Finance*, 18(1):167–181, 1994.
- A. D. Hall, H. M. Anderson, and C. W. Granger. A cointegration analysis of treasury bill yields. *The review of Economics and Statistics*, pages 116–126, 1992.
- J. D. Hamilton. *Time series analysis*. Princeton university press, 2020.
- F. Hayashi. *Econometrics*, princeton university press: Princeton. 2000.
- S. Johansen. *Likelihood-based inference in cointegrated vector autoregressive models*. OUP Oxford, 1995.
- A. Lakdawala and S. Wu. Federal reserve credibility and the term structure of interest rates. *European Economic Review*, 100:364–389, 2017.
- J. Li and G. Davis. Rethinking cointegration and the expectation hypothesis of the term structure. *Journal of Empirical Finance*, 44:177–189, 2017.
- W. K. Newey and K. D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix. 1986.
- S. Ng and P. Perron. Unit root tests in arma models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association*, 90(429):268–281, 1995.
- L. Sarno and D. L. Thornton. The dynamic relationship between the federal funds rate and the treasury bill rate: An empirical investigation. *Journal of Banking & Finance*, 27(6):1079–1110, 2003.

Bibliografia

J. H. Stock and M. W. Watson. A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica: journal of the Econometric Society*, pages 783–820, 1993.