



Università  
Ca' Foscari  
Venezia

Corso di Laurea magistrale in  
Economia e Finanza

Tesi di Laurea

**Analisi empirica della relazione lead-lag tra il  
mercato spot e futures: un'applicazione delle  
strategie di trading**

**Relatore**

Prof. Davide Raggi

**Correlatore**

Ch. Roberto Casarin

**Laureando**

Leonardo Mazzon  
875111

**Anno Accademico**

2023 / 2024

# INDICE

INTRODUZIONE.....	3
<b>CAPITOLO I - EFFICIENZA DEL MERCATO E RELAZIONI DI LEAD-LAG TRA INDICI SPOT E FUTURES .....</b>	<b>5</b>
1.1 L'indice FTSE100.....	5
1.1.1 Cenno di trading sul FTSE100.....	6
1.2 Introduzione ai Futures.....	8
1.3 Teorie dell'efficienza di mercato.....	10
1.3.1 Indagine sull'efficienza di mercato.....	13
1.4 Rassegna della letteratura.....	14
1.5 Dinamiche e sincronie tra mercati Spot e Futures: un'analisi del trading ad alta frequenza e della scoperta dei prezzi.....	17
1.6 Analisi e possibili spiegazioni della relazione lead-lag.....	20
1.7 Impatto del trading sugli Spot e Futures.....	22
<b>CAPITOLO II - FONDAMENTI TEORICI DEI MODELLI ECONOMETRICI APPLICATI .....</b>	<b>25</b>
2.1 La cointegrazione.....	25
2.2 Il Test Argumented-Dickey-Fuller.....	27
2.2.1 Il test KPSS.....	28
2.3 Il test Breusch-Godfrey.....	28
2.4 Introduzione al modello ECM.....	29
2.4.1 Introduzione al modello ECM-COC.....	31
2.5 Introduzione al modello Arma.....	33
2.6 Introduzione al modello VAR.....	37
<b>CAPITOLO III - L'ANALISI ECONOMETRICA .....</b>	<b>41</b>
3.1 Metodologia adottata.....	41

3.2	Selezione e trattamento dei dati .....	42
3.3	Statistiche descrittive di base.....	43
3.4	Test di cointegrazione e Granger-Casuality .....	47
3.5	Costruzione del modello ECM .....	55
3.5.1	Costruzione modello ECM-COC.....	58
3.6	Costruzione modello Arma (1,0,0).....	61
3.7	Costruzione modello VAR.....	64
<b>CAPITOLO IV - PREVISIONI UN PASSO IN AVANTI E DEFINIZIONE DELLE STRATEGIE DI TRADING.....</b>		<b>71</b>
4.1	Prefazione previsione statica e dinamica.....	71
4.2	Creazione e valutazione delle previsioni .....	73
4.3	Integrazione delle previsioni statistiche nelle strategie di trading.....	76
4.3.1	Liquid trading strategy .....	77
4.3.2	Buy and hold strategy.....	78
4.3.3	Better predicted return than average.....	78
4.3.4	Better predicted return than first decile.....	78
4.3.5	High arbitrary cut off.....	79
4.4	Risultati empirici derivanti dall'attuazione delle strategie di trading .....	79
4.5	Liquid Strategy.....	83
4.6	Buy and Hold .....	86
4.7	Better predicted return than average .....	88
4.8	Better predicted than first decile.....	90
4.9	High arbitrary cut off.....	93
4.10	Sintesi dei risultati.....	95
<b>CONCLUSIONI .....</b>		<b>99</b>
<b>BIBLIOGRAFIA .....</b>		<b>103</b>

## Introduzione

L'analisi dei mercati finanziari rappresenta una sfida costante per economisti e analisti finanziari poiché le dinamiche che governano l'andamento dei prezzi degli asset sono influenzate da un grande numero di fattori interconnessi tra di loro. I mercati finanziari sono ecosistemi nei quali si mescolano una miriade di componenti, tra cui variabili di politica economica dei diversi Paesi, variabili macroeconomiche, innovazioni nel comparto tecnologico ed eventi geopolitici di caratura mondiale. In questo scenario in cui tutti questi elementi si combinano e si intrecciano in modo tale da influenzare i movimenti dei prezzi di ogni bene, azione o strumento finanziario, è importante comprendere a fondo le relazioni temporali tra diversi prodotti finanziari per padroneggiarli e creare delle strategie di investimento efficaci considerando anche i relativi rischi connessi. In questo contesto, un aspetto fondamentale e di rilievo è ricoperto dalla relazione che intercorre tra indici e strumenti derivati le cui implicazioni sono particolarmente di rilievo soprattutto durante i periodi di volatilità nei quali le discrepanze temporali nei movimenti dei prezzi tendono ad essere amplificate. Quindi, i mercati finanziari possono rispondere in modo non allineato alle nuove informazioni creando, così, nuove opportunità di profitto.

Questa tesi si focalizza sulle relazioni esistenti di *lead-lag* tra l'indice dei prezzi spot e quello future, cercando di intuire e capire come i movimenti di prezzo degli strumenti derivati regolamentati possano anticipare quelli dell'indici spot o viceversa. L'analisi empirica si baserà sull'indice FTSE100 e l'approfondimento di queste relazioni non solo fornisce un beneficio per una comprensione maggiormente completa delle dinamiche dei mercati, ma ha anche delle implicazioni pratiche per la formulazione e implementazione di strategie di trading comprensive di fattori per la gestione del rischio. Le motivazioni alla base di questo lavoro sono molteplici. Innanzitutto, il mercato dei futures è spesso considerato un precursore delle aspettative del mercato rispetto ai movimenti futuri dei prezzi. Questo lo rende un campo di studio ideale per investigare l'efficienza del mercato, una teoria che postula che i prezzi riflettono tutte le informazioni disponibili ad un dato istante temporale.

L'obiettivo principale è identificare le relazioni *lead-lag* tra gli indici spot e futures dell'indice britannico FTSE100 e, quindi, se tale indice riflette tutte le informazioni disponibili. Il secondo obiettivo è quello di definire delle strategie profittevoli di trading basate sulle previsioni di alcuni modelli econometrici. Studi precedenti hanno dimostrato che le variazioni ritardate nei prezzi futures possono aiutare a prevedere le variazioni nei prezzi spot. Utilizzando dati ad alta frequenza, è stato dimostrato che i cambiamenti nei prezzi dei derivati precedono in numerose situazioni i movimenti dei prezzi dell'indice a pronti, suggerendo delle opportunità di trading profittevoli.

La tesi si articola in quattro sezioni. Si inizia con un primo capitolo introduttivo nel quale si esamina l'efficienza del mercato e delle relazioni tra gli indici spot e futures, con una rassegna della letteratura. In aggiunta si cerca di ampliare anche il tema dell'impatto del trading ad alta frequenza in questo settore. Seguirà una discussione sui modelli econometrici utilizzati e le loro caratteristiche base in modo da rendere maggiormente comprensibili i passaggi successivi. La parte centrale dell'elaborato sarà dedicata all'analisi empirica: dapprima con una rapida visione delle statistiche di base delle serie storiche per poi entrare nel vivo della ricerca con una prima verifica della non stazionarietà delle serie per poi proseguire con lo studio di eventuali relazioni di cointegrazione. In seguito, verrà considerato un modello a correzione dell'errore nel quale il disequilibrio è stimato mediante la procedura di Engle-Granger, per poi passare sempre a un modello a correzione dell'errore ma che considera altri fattori come il *risk-free rate*, i dividendi e il tempo rimanente alla scadenza al contratto futures. Si aggiunge che verranno studiati anche i modelli ARMA e VAR. Come accennato in precedenza, i suddetti modelli verranno utilizzati per fare previsioni sul mese fuori dal campione e il migliore di questi verrà usato per fare sviluppare e attuare le strategie di trading in un contesto pensato per essere il più realistico possibile. I risultati ottenuti sembrano in prima battuta molto promettenti e si dimostrano nella maggioranza dei casi di successo anche per l'aspetto relativo alla gestione del rischio. Queste conclusioni non si confermano una volta considerati i costi di transazione che erodono la quasi totalità delle disponibilità economiche investite, a conferma che ciascuna strategia studiata sia inadatta in un'ambiente come quello replicato in tale studio.

# Capitolo I

## Efficienza del mercato e relazioni di lead-lag tra indici Spot e Futures

### 1.1 L'indice FTSE100

L'indice FTSE100 è stato introdotto il 3 gennaio del 1984 ed è l'acronimo di *Financial Times and Stock Exchange*, il quale rappresenta l'indice azionario della borsa londinese che include le 100 aziende del Regno Unito con maggior capitalizzazione di mercato presenti all'interno del *London Stock Exchange*. Affinché, una società sia ammessa nell'indice, è essenziale che la stessa dimostri di aver conseguito un turnover minimo delle proprie azioni pari allo 0,5% mensile in almeno dieci dei dodici mesi precedenti alla revisione dei componenti dell'indice. Queste revisioni sono effettuate con cadenza programmata e, nello specifico, nei mesi di: Marzo, Giugno, Settembre e Dicembre (medesimi mesi per la scadenza dei contratti futures). È stilata anche una lista di riserva che include 6 società che potranno subentrare all'interno del FTSE100 qualora un'impresa contenuta in esso venga eliminata. Ovviamente la capitalizzazione sarà il criterio al quale fare riferimento per ponderare la scelta della nuova entrata. Affinché le società all'interno dell'indice non cambino troppo frequentemente si è adottata la regola del 90 e 110, quindi: qualora una società, non presente nel paniere dell'indice, aumenti il proprio valore oltre la novantesima posizione allora essa sostituirà di diritto la società che occupa l'ultimo gradino della classifica. Al contrario, una società sarà cancellata dall'indice se presenta una capitalizzazione di mercato inferiore alla cento decima posizione. L'indice è determinato attraverso una media ponderata di prezzi delle 100 società selezionate nel paniere, la ponderazione dei titoli è proporzionale alla loro capitalizzazione di mercato.

Per valutare la salute del sistema finanziario del Regno Unito, investitori e analisti esaminano l'andamento dell'indice FTSE100, il quale assume un ruolo di rilievo essendo esso sensibile ad eventi globali di notevole impatto come la Brexit, il Covid o, riportando eventi più recenti, guerra energetica ed inflazione. Individuano così rischi che minacciano determinati settori e regioni o, al contrario, delle opportunità

di investimento colte tramite l'indagine dell'esposizione geografica del FTSE100 e dei pesi settoriali. Il FTSE100 costituisce un cardine di riferimento per l'allocazione di investimenti in fondi e portafogli riferiti al Regno Unito. Un connotato distintivo è la distribuzione settoriale che emerge dalla valutazione dell'indice: l'ambito relativo alle risorse naturali e il ramo finanziario detengono una quota rilevante nella capitalizzazione complessiva del mercato totale dello stesso e quindi, queste due macroaree hanno un peso significativo. Si riporta la Figura 1.1 la quale ritrae le prime dieci società del paniere per capitalizzazione.

**Figura 1.1:** Prime 10 società dell'indice FTSE 100 elencate per capitalizzazione

NAME	PRICE	CHANGE	% CHANGE	MARKET CAP ↓
Shell plc	GBp2,564.50	↑ 16.00	↑ 0.63%	165.12B
AstraZeneca PLC	GBp10,260.00	↓ -34.00	↓ -0.33%	160.72B
HSBC Holdings plc	GBp602.10	↑ 4.10	↑ 0.69%	114.43B
Unilever PLC	GBp3,825.96	↓ -25.04	↓ -0.65%	96.70B
BP p.l.c.	GBp494.75	↑ 3.75	↑ 0.76%	83.73B
Rio Tinto Group	GBp4,906.50	↑ 44.50	↑ 0.92%	79.34B
GSK plc	GBp1,655.40	↑ 5.80	↑ 0.35%	68.74B
Diageo plc	GBp2,898.00	↓ -9.00	↓ -0.31%	64.72B
RELX PLC	GBp3,390.00	↑ 15.00	↑ 0.44%	64.06B
British American Tobacco p.l.c.	GBp2,369.00	↑ 28.50	↑ 1.22%	53.89B

**Fonte:** <https://www.directorstalkinterviews.com/ftse-100-by-market-capitalisation>

### 1.1.1 Cenno di trading sul FTSE100

In questo sottoparagrafo verrà descritta una delle strategie di trading sull'indice, distinguendo tra traders che operano nel lungo termine e quelli che operano nel breve termine. Nei capitoli successivi si enunceranno altre strategie operative di stampo completamente differente. Una delle modalità attraverso cui i traders possono operare è mediante i CFD<sup>1</sup> (Contratti per Differenza) la cui flessibilità e i

<sup>1</sup> CFD: Contratti per Differenza, sono strumenti finanziari derivati il cui prezzo deriva dall'asset sottostante che può essere un'azione, un indice, materie prime, ecc... Questi strumenti vengono utilizzati per fare attività di trading sui movimenti di prezzo del sottostante in questione.

relativi costi lo rendono uno strumento conveniente. Essi permettono di fare trading su diversi asset senza effettivamente comprare o vendere l'asset reale. Difatti, fare trading con i CFD significa comprare o vendere un numero di contratti di un mercato e, di conseguenza, il guadagno o la perdita dell'investimento è data dalla differenza di prezzo di un asset dal momento in cui il contratto viene aperto fino alla sua chiusura. Il prezzo di questo strumento riflette fedelmente il prezzo del sottostante, quindi, avranno un andamento identico a quello del sottostante. La differenza sostanziale nell'utilizzare i CFD è che sono prodotti a leva che offrono un'elevata esposizione sul mercato a fronte di un investimento relativamente modesto. Questi contratti non sono regolamentati dalle regole e dagli organismi vigenti, ma sono prodotti dai broker, quindi, sono strumenti derivati negoziati *over the counter* (OTC)<sup>2</sup>. Uno degli aspetti più importanti dei CFD è che essendo prodotti dallo stesso broker, possono avere regolamentazioni molto diverse, una di queste è rappresentata dallo *spread*. Nella borsa regolamentata è il mercato che fa sì che vengano determinati il miglior compratore e il miglior venditore, mentre nel momento in cui si crea un contratto CFD è il broker che determina i suddetti aspetti. Con questa decisione si aumenta quella che è la distanza tra il miglior venditore e compratore ed è la cifra che il broker stesso guadagna, ossia lo *spread* che si menzionava prima. Per negoziare mediante questi derivati è richiesto il versamento di un margine di garanzia, che rappresenta la frazione del valore dell'operazione, necessaria per garantire l'adempimento del contratto a termine. I margini di garanzia variano a seconda del sottostante scelto e dalla frequenza operativa stabilita (*intraday* o *multiday*). Ad esempio, nel caso degli indici, il margine per il trading giornaliero varia tra il 5% e il 10% mentre per la posizione *multiday*, il margine può raggiungere anche il 100%. D'altro canto, i CFD sulle azioni del Regno Unito, operabili solo *intraday*, richiedono un margine che oscilla dal 20% al 100%. Quindi riassumendo i punti chiave di questi strumenti si afferma che la forza dei CFD risiede principalmente nell'utilizzo di bassi capitali grazie alla leva finanziaria e quindi il costo del derivato in questione è diverso rispetto a quello dello strumento finanziario in sé che sia esso un'azione, un indice o materia prima. Ma allo stesso

---

<sup>2</sup> Transazioni di titoli che avvengono direttamente tra due parti senza l'intermediazione di un mercato regolamentato come una borsa valori. Quindi, il mercato OTC offre flessibilità e opportunità per negoziare una vasta gamma di strumenti finanziari, ma comporta anche rischi maggiori rispetto ai mercati regolamentati.

tempo questo fattore, se non utilizzato adeguatamente, può condurre alla sovraesposizione. Il secondo punto è la facilità della gestione del conto e la possibilità di operare sia con posizioni *long* che *short*.

I broker, solitamente, collocano dei CFD che rispecchiano l'indice spot e un altro che, invece, riflette il contratto futures scritto sull'indice. Nell'ambito della negoziazione di indici attraverso i suddetti strumenti online, gli operatori ipotizzano la direzione dei futuri movimenti del FTSE100, eludendo il possesso effettivo dell'indice o dei componenti stessi. In tal senso, l'operatività dei traders è più ampia grazie anche all'utilizzo di strumenti di leva finanziaria che consentono di intraprendere posizioni sia rialziste che ribassiste con maggiore "aggressività". Emerge come l'impiego di Contratti per Differenza, in un contesto ribassista, si riveli di importanza strategica. In questo frangente di mercato, ristrutturare e riorganizzare il portafoglio potrebbe risultare dispendioso in termini economici e presenta delle insidie a livelli di tempistica per sincronizzare le proprie mosse con il mercato. Quindi, anziché liquidare una porzione del portafoglio per apportare una correzione, si può speculare ottenendo una capitalizzazione da movimenti di prezzi ribassisti attraverso l'uso dei CFD.

Operare nel mercato spot o nel mercato futures è a discrezione del trader: qualora egli operi nel breve termine sarà preferibile il mercato spot per via degli spread bassi. Diversamente, se egli mantiene posizioni di lungo termine, forte del fatto che i costi di swap non sono addebitati, il mercato futures sarà la scelta auspicabile. Un aspetto da non sottovalutare è presentato dai rollover su CFD futures, ovvero l'eseguire uno slittamento della posizione del trader su un contratto con scadenza futura quando il contratto in essere è vicino al suo termine. Questa procedura si attua per evitare di sostenere costi e obblighi derivanti dalla liquidazione del contratto alla scadenza T.

## **1.2 Introduzione ai Futures**

Gli strumenti derivati futures sono accordi con i quali due o più parti si impegnano ad acquistare o vendere una determinata attività sottostante che può essere sia

reale<sup>3</sup> che finanziaria, a una precisa data futura e ad un prezzo stabilito a priori. Diversamente da altri strumenti finanziari derivati quali i forward, i futures sono scambiati in borsa la quale introduce e puntualizza determinati cardini standard dell'accordo al fine di semplificare la procedura di negoziazione. La stessa borsa interviene per garantire ai contraenti il buon esito della procedura e quindi tutela l'onorabilità dell'intesa. Nella quasi totalità dei casi, i contratti futures non vengono tenuti dall'investitore fino alla data di consegna dell'oggetto sottostante e il motivo risiede nella fattispecie che colui che sottoscrive l'accordo chiude la propria posizione prima del periodo in cui avviene la consegna. Chiudere una posizione significa assumere un contratto di segno opposto rispetto all'originale. Un aspetto importante in questo caso è la dimensione del contratto che va ad indicare il preciso importo dell'oggetto del contratto che si dovrà consegnare. Qualora ci fossero traders che esigono entrare in una posizione con fine speculativo di piccole dimensioni o coprirsi da un'esposizione altrettanto piccola, la dimensione del contratto dovesse essere eccessiva allora non potranno utilizzarlo. D'altro canto, dato che ad ogni contratto corrisponde un costo e se il contratto stesso dovesse presentare una dimensione esigua allora assumere una posizione in esso risulterebbe sproporzionato.

Per le contrattazioni dei futures, le borse hanno definito dei limiti di prezzo, ovvero un intervallo entro il quale il prezzo durante la giornata può variare. Qualora la variazione sia pari al limite inferiore o al limite superiore le negoziazioni vengono di norma sospese. Questi limiti sono specifici per evitare che si verifichino degli eccessi speculativi che comporterebbero dei cambiamenti di prezzo anomali.

I contratti futures fanno riferimento al mese di consegna e il prezzo futures tende al prezzo spot man mano che il mese di consegna si avvicina. Infatti, il prezzo futures dovrebbe essere uguale o, comunque, molto vicino al prezzo spot durante il *delivery month*. Per dare evidenza a tale concetto si può presumere che, per il mese di spedizione, il prezzo futures sia superiore al prezzo spot creando così un'opportunità di guadagno sicuro per gli arbitraggisti che opererebbero in questo modo:

---

<sup>3</sup> Con il termine reale si fa riferimento a sottostanti tangibili come possono essere le materie prime: oro, rame, grano e altro.

- I. Si liquida un contratto futures;
- II. Viene eseguita la consegna;
- III. Si acquista il sottostante;

Quindi procedendo con questo schema il guadagno è pari alla differenza del prezzo futures sullo spot. Il caso contrario prevede che il prezzo futures sia minore dello spot e di conseguenza coloro che sono interessati al sottostante non lo acquireranno a pronti ma sottoscrivendo una posizione lunga sui futures, provocando un aumento del prezzo del derivato.

Entrambe le casistiche riportate garantiscono l'equivalenza, o la quasi totalità della stessa, del prezzo futures e del prezzo spot durante il mese di consegna.

### **1.3 Teorie dell'efficienza di mercato**

Le fondamenta del mercato efficiente risiedono nel concetto tale per cui il mercato riflette tutte le informazioni disponibili in quel momento. Seguendo questo filone di pensiero si dovrebbe escludere ogni possibilità di trarre dei rendimenti sistematici superiori alla media tramite delle previsioni sull'andamento del mercato. Il principio di questa teoria si lega alla nozione di camminata casuale, meglio conosciuta come *random walk*, la cui formulazione risale alla seconda metà dell'Ottocento. Infatti, l'idea da parte di Jules Regnault è che i prezzi debbano essere non predicibili sulla base dell'osservazione storica. Se lo fossero, vorrebbe dire che i mercati non processano correttamente le informazioni disponibili. Seguendo le sue orme, Bachelier formulò il pensiero che le traiettorie dei movimenti dei prezzi fossero esenti da ogni possibile previsione in quanto esse seguono un moto browniano. Alla fine della Seconda guerra mondiale, molte pubblicazioni sottolinearono come il consulto dei dati storici al fine di prevedere i prezzi fosse alquanto inutile. Quindi, secondo la teoria del mercato efficiente ogni strategia di investimento fondata su previsioni o sullo studio della stagionalità risultano improduttive nell'ottenere dei vantaggi o degli extraprofitti.

Fama (1970) fonda la definizione dell'efficienza del mercato su quattro aspetti: il gioco equo, il processo submartingala, il modello *random walk* e le condizioni di equilibrio.

Nella teoria del gioco equo i rendimenti attesi possono essere descritti da:

$$E(X_{j,t+1}|\Phi_t) = [1 + E(r_{j,t+1}|\Phi_t)]X_{j,t} \quad (1.3.1)$$

L'operatore di sinistra è il valore atteso dell'azione nel periodo  $t + 1$  mentre la componente di destra riflette il valore atteso del rendimento percentuale un passo in avanti condizionato al fatto che il mercato rifletta ogni informazione ( $\Phi_t$ ). Partendo dall'assunto che l'efficienza di mercato sia specificata dai rendimenti attesi e che le aspettative degli operatori del mercato siano formate sulla base che l'informazione sia disponibile e completa, allora:

$$\epsilon_{j,t+1} \equiv X_{j,t+1} - E(X_{j,t+1}|\Phi_t) \quad (1.3.2)$$

Costituisce la differenza tra i rendimenti del presente e i rendimenti previsti, in concomitanza del termine  $\Phi_t$ , al periodo successivo. Quindi se  $\{\epsilon_j, t\}$  può essere inteso come un gioco equo in confronto al dataset informativo disponibile rappresentato da  $\Phi_t$ , allora si avrà:

$$E(\epsilon_{j,t+1}) = 0 \quad (1.3.3)$$

considerando ora un sistema di trading fondato sulle informazioni disponibili  $\Phi_t$ :

$$\alpha(\Phi_t) = [a_1(\Phi_t), a_2(\Phi_t), \dots, a_n(\Phi_t)] \quad (1.3.4)$$

tale per cui esso fornisce informazioni all'operatore per l'ammontare di fondi da allocare in ogni  $n$  asset disponibile. Quindi, la somma dei rendimenti in eccesso nel portafoglio è rappresentata da:

$$V_{t+1} = \sum_{j=1}^n a_j(\Phi_t)[r_{j,t+1} - E(r_{j,t+1}|\Phi_t)] \quad (1.3.5)$$

si conclude che, la condizione di gioco equo è soddisfatta da tale condizione:

$$E(V_{t+1}|\Phi_t) = 0 \quad (1.3.6)$$

Quindi, non ci si dovrebbero attendere dei rendimenti maggiori rispetto a quelli normali se si utilizzano le informazioni riflesse nel mercato.

Il secondo aspetto è la proprietà di submartingala la quale implica che  $E(X_{j,t+1}|\Phi_t) \geq X_{j,t}$  rispetto al set informativo precedentemente anticipato, cioè la previsione del prezzo atteso condizionato alle informazioni deve essere quantomeno uguale o, ancor meglio, maggiore del prezzo corrente, ma mai inferiore ad esso. Da questa teoria, Fama, annuncia che in ambito di trading la strategia *Buy and Hold*<sup>4</sup> dovrebbe avere rendimenti superiori o equivalenti alla strategia *one Security and Cash*<sup>5</sup>. Quest'ultima non potrà dominare la prima.

Il modello “camminata casuale”: come detto, tutte le informazioni devono essere riflesse dai prezzi attuali e quindi, ne consegue, che i prezzi futuri cambieranno in modo imprevedibile. Inoltre, è comune accettare l'ipotesi per cui i rendimenti futuri sono identicamente distribuiti. Il quadro finale, considerando quanto detto, forma il modello *random walk*:

$$f(r_{j,t+1}|\Phi_t) = f(r_{j,t+1}) \quad (1.3.7)$$

in questo caso la funzione  $f$  rappresenta una funzione di densità uguale per tutti i periodi. Se si assume che il valore atteso sia costante nel tempo, si ottiene:

$$E(r_{j,t+1}|\Phi_t) = E(r_{j,t+1}) \quad (1.3.8)$$

il che suggerisce che la variazione del valore atteso dei prezzi è indipendente dal set informativo  $\Phi_t$ .

Le condizioni di equilibrio: prendendo in rilievo un mondo finanziario composto da tre vincoli: in ogni momento le informazioni sono reperibili e disponibili ad ogni partecipante del mercato, non sono presenti costi di trading per le azioni e, infine, investitori informati e razionali che formano aspettative reali e concrete sui rendimenti futuri in base allo studio delle informazioni disponibili. Fama ha identificato queste regole sufficienti e non necessarie, anche se molto stringenti, per avere un mercato efficiente. Infatti, alcuni aspetti non possono configurarsi come inefficienze di mercato vere e proprie, bensì, possono rappresentare delle potenziali

---

<sup>4</sup> Questa strategia verrà discussa e applicata successivamente

<sup>5</sup> Si indica un insieme di regole di trading che dettano le condizioni in base alle quali un operatore dovrebbe detenere il titolo oppure venderlo allo scoperto, o al contrario di tutto, mantenere i contanti senza investire.

violazioni. In effetti lo stesso autore individua tre possibili forme di mercato efficiente: forma forte, che presuppone tutte le informazioni siano riflesse nel mercato anche se una parte di esse siano ad accesso privilegiato ad alcuni operatori; forma semi-forte in cui i prezzi passati, anche le informazioni disponibili al pubblico debbano essere riflesse nei cambiamenti dei prezzi; la forma debole sotto intende che non vi è possibilità alcuna di prevedere i prezzi futuri avendo come base i prezzi passati dato che questi contengono tutte le informazioni.

### 1.3.1 Indagine sull'efficienza di mercato

Per quanto affermato nel paragrafo precedente, sembrerebbe che non vi siano evidenze per poter prevedere i rendimenti futuri degli asset avendo come base di partenza i prezzi passati; quindi, avendo a disposizione solamente le serie storiche osservate fino al presente,  $t$ . In questo sottoparagrafo si desidera esclusivamente dare una panoramica generale dell'ipotesi su cui si basa il lavoro, ossia la serie storica *Random Walk*.

Si è notato che il processo stocastico della camminata casuale, è applicato per dare una rappresentazione e una previsione di quelli che saranno i prezzi futuri delle azioni. A seconda delle ipotesi sulla distribuzione dei termini d'errore, rappresentati da  $\{\varepsilon_t\}$ , si distinguono tre versioni di modelli differenti di tale processo.

- I. La prima versione, la meno sofisticata, prevede di avere il prezzo dell'azione  $X_t$ , al periodo presente, a sinistra dell'equazione e quindi funge da variabile dipendente. Per cui assumerebbe tale forma:

$$X_t = \mu + X_{t-1} + \varepsilon_t \quad \varepsilon_t \sim IID(0, \sigma^2) \quad (1.3.1.1)$$

gli errori si ipotizza che siano distribuiti in modo identico ed indipendente con una media nulla e una varianza prefissata. Sotto tale formulazione gli incrementi si caratterizzano per la loro mancata correlazione e, questa indipendenza coinvolge anche ogni funzione non lineare associata agli stessi.

- II. Le premesse iniziali del modello sono percepite come troppo limitanti applicazioni su asset finanziari del mondo reale, in quanto le distribuzioni di

rendimento non sono costanti per tutto il periodo temporale. Perciò, la seconda formulazione del medesimo modello presuppone che la componente  $\{\varepsilon_t\}$  segua una distribuzione indipendente, ma non omogenea (INID). Questo modello preserva l'imprevedibilità delle variazioni di prezzo e introduce l'eteroschedasticità non condizionata degli  $\{\varepsilon'_t\}$ , adeguata alla variabilità temporale degli asset.

- III. L'ultima modifica considerata assume che  $\{\varepsilon_t\}$  sia trattato come una sequenza di incrementi non correlati, sebbene non in tutto indipendenti.

#### 1.4 Rassegna della letteratura

L'introduzione dei contratti futures e lo sviluppo di attività di trading mediante questi strumenti, agli inizi, hanno destato un'estesa preoccupazione relativamente ai rischi aggiuntivi che i derivati stessi potessero apportare per le aziende, oltre che per il mercato in sé. In parallelo, però, hanno anche contribuito ad indirizzare le ricerche accademiche sugli impatti del trading su derivati e sui mercati sottostanti. L'espansione dei mercati derivati si configura come una risposta alle esigenze di mitigazione del rischio economico legato alle transazioni di strumenti finanziari e delle materie prime. L'aumento della volatilità e dei rischi di mercato ha determinato una maggiore richiesta di strumenti di copertura. Di conseguenza, il mercato dei derivati ha visto una crescita significativa. Quest'ultimi nascono per trasferire il valore e il rischio da una controparte ad un'altra. Più precisamente: *«i derivati sono contratti il cui valore dipende (ossia deriva) da una o più variabili sottostanti. Molto spesso le variabili sottostanti sono rappresentate dal prezzo di attività negoziabili»* John C. Hull, Opzioni, Futures e altri derivati, Pearson, p. 1.

Sono state mosse molte obiezioni ai mercati dei derivati sulla base che l'attività di trading su tali strumenti, soprattutto quella esercitata da speculatori, potesse destabilizzare i prezzi spot e portare ad un aumento della volatilità di mercato. Questo potrebbe accadere a causa dell'operato di speculatori non informati che agiscono sia sul mercato spot che futures per trarre un guadagno. In tal modo le informazioni che i prezzi riflettono diminuiscono e, contrariamente, aumenta l'incertezza. Inoltre, forti oscillazioni dei prezzi in una finestra temporale ridotta

possono essere causate dall'attività di arbitraggio. A supporto di tale tesi relativa al connubio volatilità dei prezzi nel mercato spot e volume del mercato futures, uno studio eseguito sull'indice S&P500 da Schwert (1990) riporta una certa correlazione tra l'indice americano e il volume del mercato futures e azionario. Infatti, quando l'indice S&P500 presenta un'alta volatilità, anche il volume degli altri due mercati, precedentemente menzionati, è alta. Sulla stessa linea si trova anche il lavoro condotto da Darat e Rehman (1995).

Si aggiunge che, in tema di partecipazione dell'attività di trading su strumenti derivati (futures) essa può concorrere a influenzare la volatilità del mercato sottostante svolgendo un ruolo di catalizzatore per l'aumento o diminuzione; in questo ambito sono stati riscontrati dei risultati contrapposti. Riportando uno degli studi di Board, condotto nel 2001, nel quale sono stati analizzati i prezzi di chiusura delle azioni giornaliere della borsa di Londra e dei contratti del FTSE100, utilizzando un modello di volatilità stocastica. Il prodotto della ricerca porta alla luce delle evidenze secondo le quali il volume di trading dei futures non incide in modo negativo sul mercato spot. L'autore evidenzia che l'attività di trading simultaneo di futures corretti per le tendenze temporali e per le nuove informazioni, non destabilizzano il mercato spot.

Santoni (1987) ottiene dei risultati che presentano una relazione negativa tra volume dei futures dell'indice americano S&P500 e i prezzi del mercato spot. Oltre a ciò, basandosi sulle informazioni di mercato e i relativi scambi, i mercati di futures e opzioni possono aumentare il volume di trading. Questo è reso possibile dalle asimmetrie informative che popolano le aziende. Quindi, se gli scambi informativi aumentano, il mercato è più efficiente e favorisce un'incorporazione più rapida e dettagliata delle informazioni nei prezzi, i quali, essendo maggiormente precisi, rispecchiano meglio il valore reale dei beni. Di conseguenza, la volatilità del mercato sottostante può ridursi, quindi in linea teorica, il trading su strumenti derivati ha un influsso stabilizzante o meno sulla volatilità del mercato sottostante in dipendenza dell'approccio teorico assunto. Infatti, vi dovrebbe essere perfetta correlazione simultanea tra i rendimenti dell'indice di mercato e il relativo contratto futures qualora i mercati fossero privi di deficit informativi. Per l'appunto i prezzi del contratto futures e dell'indice spot dovrebbero coincidere negli spostamenti e non presentare forme di co-correlazione nel tempo. Perciò i prezzi delle attività

sopracitate dovrebbero riflettere immediatamente le informazioni di mercato anche perché, la teoria del mercato efficiente<sup>6</sup> contempla che i prezzi coincidano, tuttavia, se si dovessero verificare delle discrepanze tra i prezzi, queste dovrebbero essere eliminate. Ciò è dettato dal fatto che gli arbitraggisti operano acquistando il bene ad un prezzo basso e lo rivendono ad un prezzo alto affinché i prezzi non si riallineino. Altri studi condotti sul tema trovano delle soluzioni interessanti, ad esempio, Kyriacou e Sarno (1999) nel loro lavoro intitolato “La relazione temporale tra volatilità dei rendimenti azionari e volatilità dei rendimenti futures sui titoli azionari” per stimare la volatilità condizionale utilizzano modelli ARCH e GARCH e i risultati ottenuti dimostrano che sia il mercato azionario che futures mostrano una volatilità condizionale, e gli shock del passato influenzano la volatilità dei rendimenti. A tal proposito, notano come esista una relazione bidirezionale tra la volatilità dei rendimenti futures e volatilità dei rendimenti azionari, ma allo stesso tempo, gli effetti della volatilità legata all’azionario sulla volatilità futures è più consistente per gli shock negativi che per quelli positivi. Quindi concludono che la volatilità dei rendimenti dei due mercati sono interconnesse.

L’analisi della correlazione tra indici azionari e derivati assume un ruolo di primaria importanza nell’ambito della finanza. Tale interdipendenza consente di far luce sul processo di scoperta del prezzo nei citati mercati spot e strumenti derivati. I benefici che si traggono dal commensurare questa correlazione si riversano su molteplici figure del panorama finanziario, tra cui investitori e regolatori. L’analisi mirata conduce ad una gestione del rischio strategica per lo sviluppo di migliori strategie di investimento derivanti dalla conoscenza dei prezzi.

A tal fine, l’articolo di Balyeat, Yang e Leatham (2005) si concentra sull’impatto che le disparità informative hanno sui prezzi e sui volumi dei mercati. Quindi essi studiano 1344 aziende americane prendendo in considerazione tre variabili: i prezzi, i volumi di negoziazione e le asimmetrie informative. I risultati mostrano che quest’ultime hanno un effetto positivo sui volumi di negoziazione e sui prezzi e precisamente che per la maggior parte dei futures selezionati di materie prime si osserva una correlazione positiva tra la differenza del volume di trading previsto e quello realizzato e la volatilità dei prezzi del sottostante. In parole semplici, dato che

---

<sup>6</sup> Vale quanto discusso nel paragrafo 1.3

il volume di trading inaspettato può essere un segnalatore di nuove informazioni, allora quando il volume di trading dei futures è superiore alle previsioni si verifica un aumento della volatilità dei prezzi nel mercato sottostante.

La letteratura empirica che tratta il tema dell'impatto del trading di opzioni e derivati sulla volatilità del mercato spot presenta due approcci distinti e differenti: una metodologia consiste nell'analizzare l'effetto dell'ammissione di contratti derivati alla negoziazione sulla volatilità dell'attività sottostante. La ricerca si basa sul misurare la volatilità in due momenti: prima e dopo l'elenco dei derivati. Generalmente, questo approccio è maggiormente propenso ad affermare che, dopo l'elenco, la volatilità diminuisca. Ciò non deve stupire dato che, quando la varianza sottostante è elevata, allora i derivati hanno una probabilità maggiore di essere elencati. Questo approccio è il più diffuso tra gli autori della materia, si guardi Testler e Higginbotham (1977), Skinner (1989) e Damodaran e Lim (1991).

Il secondo approccio diretto si concentra, invece, sulla relazione che intercorre tra attività di trading, derivati e come terzo componente prende in esame la volatilità del mercato sottostante. Chan e Chung (1993) tramite i loro modelli, utilizzati per esaminare il legame tra: differenziali di arbitraggio, volatilità dei prezzi *cash* e futures e volume di trading, traggono delle conclusioni omettendo una considerazione che potrebbe aver causato un'interpretazione erronea se non ambigua dei dati. Infatti, non hanno preso in considerazione che le variabili sopracitate potessero esporre delle influenze reciproche simultanee. Quindi le loro strategie econometriche potrebbero soffrire di specificazione.

### **1.5 Dinamiche e sincronie tra mercati Spot e Futures: un'analisi del trading ad alta frequenza e della scoperta dei prezzi**

Continuando con la letteratura relativa agli studi condotti sulla dinamica dei prezzi spot e futures, dovrebbe esistere un sottile equilibrio tra derivati e indici azionari tale per cui questi ultimi non dovrebbero svolgere né la funzione di "guida" né la funzione di "inseguitrice" rispetto ai primi. Perciò, stando alle ipotesi del mercato efficiente (Fama, 1991, Shao, 2019) ci si aspetta che gli indici e i derivati siano sincronizzati nei loro movimenti in modo da rispecchiare le condizioni reali del

mercato sottostante. Ma, come ampiamente anticipato, intervengono dei fattori sia legati alle frizioni di borsa sia per la complessità reale dello stesso mercato (liquidità, e shock economici, geopolitici e altro) che danno luogo ad uno schema che si potrebbe nominare di “guida e ritardo”, nel senso che non è presente una relazione univoca per capire se i prezzi degli indici guidano i movimenti dei derivati o viceversa. Però, l’ascesa recente di sistemi trading azionario automatizzati ha innescato innovative riflessioni sulla natura causale tra gli indici azionari e i derivati su base temporale infragiornaliera. L’ elevata frequenza di transazioni eseguita da tali sistemi di trading automatizzati può orientare e influenzare le dinamiche di mercato e dar vita a nuove relazioni casuali tra indici e derivati. In sintesi, la costante e rapida operatività offerta da tali sistemi è in grado di modificare l’equilibrio tra domanda e offerta coinvolgendo, quindi, la relazione causa effetto sugli strumenti finanziari, in virtù di fattori come:

- anticipazione delle tendenze grazie all’individuazione dei pattern nei dati;
- amplificando le fluttuazioni;
- creando distorsioni informative.

Nell’ultima decade diversi studi (si veda Ryu, 2012; Kim & Ryu, 2014) si sono concentrati sul tema e le conclusioni sono eterogenee e, in alcuni casi, discordanti. Infatti, alcune ricerche hanno posto l’attenzione sul fatto che il ruolo di guida venga esercitato dagli indici a discapito dei derivati (Cao et al., 2014; Gaul & Theissen, 2015). In altri approfondimenti si evince il contrario, in particolare un’influenza inversa (Tse, 2001; Fung et al., 2015) o una relazione di cointegrazione. Una spiegazione alla varietà di risultati che si sono raggiunti può essere data attraverso diversi fattori tra cui: campioni di dati non omogenei, metodologie di analisi eterogenee e la scelta dello strumento finanziario derivato (opzione o futures, ecc.). È altresì vero che tale relazione non è stabile nel tempo ma subisce delle variazioni anche a causa di fattori esterni che possono essere rappresentati dall’ambiente politico o delle regolamentazioni finanziarie. Nonostante tutto è innegabile che i mercati future e spot, rimangono intrinsecamente connessi non solo per il lato della correlazione de rendimenti ma anche per dipendenza della volatilità.

Si è notato che la frequenza con cui si delineano relazioni di anticipo e di ritardo tra le serie di prezzi di strumenti finanziari a loro modo correlati, ad esempio derivati e

azioni, è maggiore se l'intervallo considerato è breve nonostante i coefficienti di correlazione di anticipo e ritardo tra le variazioni di prezzo siano piccole. Si parla quindi di dati ad alta frequenza e la discrepanza simultanea dei movimenti degli strumenti finanziari suggerisce come, pur essendo correlate queste serie di prezzi, alcuni guidano gli altri in quanto riescono ad inglobare velocemente le notizie. Gli studi sulle relazioni temporali intraday tra indice e futures suggeriscono determinati risultati, come:

- I rendimenti dei futures precedono o influenzano i rendimenti degli indici cash e tale effetto è amplificato quando la coerenza del movimento tra le azioni è grande;
- La capacità dei futures di anticipare i movimenti del mercato azionario è variabile in quanto il potere informativo non è uniforme e, di conseguenza, l'ampiezza del vantaggio informativo si traduce in un range che parte da 5 minuti fino ai 45 minuti;
- Il nesso di causalità non è univoco e, quindi, anche l'indice può assumere il ruolo primario per la scoperta del prezzo e influenzare i futures. Questo effetto è ampiamente marginale e tipicamente con una durata inferiore.

Le ipotesi per cui il mercato dei futures guidi la scoperta dei prezzi delle azioni e quindi abbia un peso consistente nell'informare il mercato può essere riassunto dai costi di transazione e dall'uso di una leva finanziaria estremamente elevata qualora si usassero i future. Quindi, per confermare ulteriormente questi concetti, si evidenzia come nella realtà il mercato non sia mai efficiente in senso stretto dato che intervengono anche fattori come i costi di transazione e, in aggiunta, i trader sono orientati a prediligere il mercato con maggior leva finanziaria e meno dispendioso a livello economico visto che operano in un contesto di razionalità economica per totalizzare un profitto maggiore. Pertanto, i futures richiedono un ammontare *cash* inferiore rispetto a comprare direttamente le azioni che compongono l'indice di riferimento, è bene sottolineare che solitamente i depositi a margine rappresenta una porzione inferiore rispetto al valore di mercato delle azioni. Un altro punto a favore dei future è che l'operazione in tale mercato può essere attuata nell'immediato a differenza dell'acquisto del paniere azionario che, richiedendo un ammontare maggiore di denaro potrebbe necessitare di maggior tempo per essere

implementato. Sintetizzando, la preferenza di agire in questo mercato per via dei minori e più efficienti costi potrebbe condurre all'effetto anticipazione.

## 1.6 Analisi e possibili spiegazioni della relazione lead-lag

Da una prospettiva pratica, è comunemente accettato che i due coefficienti che influenzano maggiormente la correlazione che intercorre tra i futures degli indici azionari e il mercato stesso siano l'arbitraggio e il *sentiment* di mercato. Per quanto concerne l'arbitraggio, gli operatori cercano di trarre dei profitti sfruttando le inefficienze tra due mercati e tale attività di trading mira a far convergere i prezzi dei mercati coinvolti. Invece, le aspettative degli investitori alterano le dinamiche di mercato, le quali poi si ripercuotono nei titoli quindi la positività e la fiducia degli investitori farà aumentare i prezzi mentre un sentimento negativo tenderà a farli scendere. L'opinione diffusa tra i traders suggerisce che eventuali movimenti nel prezzo dei futures rivelino aspettative sui movimenti futuri del mercato spot. Se non altro tutte le informazioni che interessano l'attività sottostante dovrebbero essere assorbite immediatamente dagli strumenti derivati ed essere pronti ad accogliere i nuovi eventi. Anche l'indice stesso dovrebbe comportarsi nella stessa misura e quindi, al fine di reagire alle nuove informazioni, ogni azione che compone l'indice deve riflettere le novità e deve essere rivalutata di conseguenza. Di fatto non tutte le azioni che costituiscono il paniere sono soggette a continui *trade* a intervalli di 10 minuti e da qui si evince il motivo per il quale l'indice si adatterà con ritardo all'inglobazione dei nuovi dati. Il vantaggio del mercato futures su quello *cash* è attribuito a differenti forze, tra cui: i valori dell'indice sono obsoleti per via della non trattazione continua delle azioni, i costi di transazione sono troppo elevati per le azioni e quindi si predilige contrattare nel mercato futures. Questi elementi hanno contribuito a trarre le conclusioni dello studio "Intraday Lead-Lag Relationships between the Futures-, Options and Stock Market" di Jong (1998), nel quale si è scoperto che i mutamenti del valore del prezzo futures anticipano i cambiamenti del prezzo *cash* con una tempistica media che va dai 5 ai 10 minuti.

Si immagini ad esempio di essere un trader che operi in un mercato nel quale sono giunte delle notizie positive; dunque, le alternative in vostro possesso sono essenzialmente un paio: acquistare lo strumento derivato e quindi comprare i

futures indicizzato all'indice oppure rivolgersi direttamente al mercato azionario e acquistare un paniere di azioni che compongono l'indice di riferimento. La prima opzione si prospetta, in linea teorica, maggiormente soddisfacente per una questione economica dato che con un esborso minore si può procedere all'acquisto del derivato che si ricorda essere uno strumento con leva finanziaria. Diversamente l'acquisto di azioni richiede un maggior tempo di esecuzione conseguente alle tempistiche da dedicare per un'attenta selezione degli strumenti finanziari, dalla quantità di transazioni e un capitale iniziale elevato. Inoltre, vi sono determinate caratteristiche degli strumenti derivati, quali futures, che contribuiscono a far sì che questi ultimi si muovano anticipatamente rispetto all'indice e quindi siano in grado di prevederlo. Queste proprietà sono ad esempio: margini bassi, la leva finanziaria, possibilità di assumere posizioni short con maggior facilità e, infine, il mercato è estremamente liquido. Pertanto, possono fungere da segnale del *sentiment* per le fluttuazioni dei prezzi delle azioni e, di conseguenza, nell'indice dato che l'operatività degli investitori che non utilizzano i futures, per determinate ragioni, assimilano queste informazioni nelle transazioni che effettuano nel mercato. L'insieme di tali ragioni e, appunto, l'inclinazione a prediligere i futures rispetto alle azioni potrebbe assumere un ruolo importante nella spiegazione della relazione di ritardo che si viene a creare in molteplici mercati. È altresì vero che, essendo l'indice un segmento delle informazioni che influenzano i futures, i movimenti dei valori nell'indice possano anch'essi anticipare i mutamenti di prezzo degli altri strumenti tanto discussi finora. Si precisa che eventuali fluttuazioni del *sentiment* del mercato andranno a impattare sia il prezzo dell'indice che dei futures verso la medesima direzione finché la base<sup>7</sup> sarà compresa nel margine di negoziazione che esclude l'attività di arbitraggio. Qualora tale intervallo venga meno allora gli arbitraggisti saranno nella posizione di conseguire dei guadagni privi di rischio fino a quando i prezzi non si riadattino nuovamente all'interno dei confini di arbitraggio.

---

<sup>7</sup> Con il termine base si indica la variazione in termini assoluti tra il prezzo futures e quello spot che necessita di rimanere nell'ambito dei limiti dell'arbitraggio definiti dall'equazione del costo di mantenimento:

$F_t = S_t \exp[(r - d)(T - t)]$ . Nella quale  $F_t$  rappresenta il Futures al tempo  $t$ ,  $S_t$  è il prezzo spot,  $r$  indica il tasso free risk,  $d$  il rendimento dei dividendi e  $T$  è la data di scadenza del contratto.

## 1.7 Impatto del trading sugli Spot e Futures

In questo paragrafo si desidera ampliare come e quanto l'attività di trading possa impattare sui mercati spot e futures, mantenendo sempre il *focus* sulla dinamica *lead-lag* tra i due mercati. Come già sottolineato, l'importanza di conoscere questa relazione di ritardo aiuta i soggetti dello scenario finanziario come i gestori di portafogli e investitori ad adottare strategie e tecniche per ridurre il rischio e massimizzare i profitti, ma è di fondamentale importanza anche per coloro che hanno l'obiettivo di mantenere inalterate l'efficienza e l'integrità del mercato. Esiste una forte connessione tra i due mercati, infatti molti studi evidenziano come i prezzi futures spesso tendono ad anticipare i movimenti futuri del mercato spot, consolidando gli indicatori degli investitori circa l'evoluzione dei prezzi.

La relazione di *lead-lag* tra i mercati spot e futures può essere influenzata da vari fattori, tra cui il *sentiment* di mercato e le strategie di arbitraggio. Il *sentiment* di mercato può causare variazioni significative nei prezzi, specialmente nel breve termine, dove l'euforia o il panico possono distorcere i prezzi dal loro valore fondamentale. Inoltre, il trading di arbitraggio, sfruttando le discrepanze di prezzo tra i mercati spot e futures, contribuisce a riallineare i prezzi, garantendo che le differenze siano temporanee e che i mercati rimangano efficienti.

La dinamica che intercorre tra i due mercati e la direzione delle influenze che essi esercitano l'uno sull'altro sono state oggetto di numerose ricerche accademiche che hanno investigato la natura empirica del fenomeno. Ad esempio, Kyle (1985) esamina il modo in cui l'*insider* trading effettuato mediante ordini dinamici si traduca nelle asimmetrie informative e quindi comporta variazioni di prezzo. In aggiunta, l'esame congiunto della volatilità e dei volumi di trading supportano la tesi per cui il trading influenza le dinamiche di mercato: l'*high frequency trading*, in italiano il trading ad alta frequenza, ha trasformato completamente i mercati finanziari e contribuisce insieme alla speculazione ad accentuare e intensificare le oscillazioni del mercato spot; contrariamente l'arbitraggio armonizza i prezzi tra i due mercati (Black, Scholes, 1973; Jegadeesh, Titman, 1993). Infatti, questo sistema di trading ha portato delle novità relative alla relazione di *lead-lag*. Per l'appunto, "The Flash Crash: High-frequency Trading in an Electronic Market" degli autori Menkveld e Yueshen (2019) trova come l'alta frequenza del trading possa impattare

e modificare la rapidità con cui i prezzi dei mercati possano integrare informazioni modificando la dinamica di *lead-lag* e riducendo lo spread *bid-ask*, migliorando quindi l'efficienza di mercato. Difatti, gli algoritmi sviluppati in questa nicchia sono in grado di reagire ed analizzare un'elevata quantità di dati di natura economica in frazioni di secondo, semplificando l'aggiustamento dei prezzi e svolgendo un ruolo di ponte per la trasmissione delle informazioni tra i due mercati, rafforzandone la relazione.

Oltre a stabilizzare la volatilità ed assicurare la liquidità, che sono due capisaldi del mercato, le differenti pratiche di trading fungono da veicolo per il flusso e l'integrazione di informazioni nei diversi settori del mercato, incidendo non solo nei prezzi ma anche nelle metodologie di investimento che sono adottate dai diversi investitori. In conclusione, per correggere le proprie decisioni e quantomeno assodare di anticipare le variazioni dei prezzi, la conoscenza della correlazione di anticipazione e ritardo che viene a crearsi nei movimenti di mercato risulta cruciale per i gestori e gli operatori economici. Questo permette di modulare le strategie per aumentare il volume dei guadagni ma, contemporaneamente limitando i rischi derivanti dalle oscillazioni dei valori.



## Capitolo II

### Fondamenti teorici dei modelli econometrici applicati

#### 2.1 La cointegrazione

Nel modello di regressione che coinvolge dati di serie storiche è fondamentale la stazionarietà delle serie. Se le serie non sono stazionarie e seguono un percorso casuale, di tipo *random walk*, la regressione lineare non è uno strumento appropriato e può portare a risultati ingannevoli. Il *random walk* è un processo detto anche integrato perciò l'osservazione di oggi può essere scritta come la somma dei termini d'errore del passato.

Tuttavia, esiste un'eccezione alla regola, ossia: se le due serie oggetto di studio seguono un percorso casuale ma sono cointegrate, allora possono essere analizzate congiuntamente nonostante la non stazionarietà.

Si parla di cointegrazione quando si studia la combinazione di serie integrate e la cointegrazione ipotizza che il livello di instabilità delle serie che si considerano è il medesimo. Di conseguenza si sta affermando che ci sono delle serie che condividono la stessa componente di *shock*. Nel quadro di un modello economico, la cointegrazione è particolarmente idonea all'analisi poiché consente di formalizzare relazioni significative tra variabili a lungo termine.

Prendendo in esame due serie non stazionarie, *random walk*:

$$Y_t = Y_{t-1} + \epsilon_{y,t} = \sum_{j=1}^t \epsilon_{y,j} \quad (2.1.1)$$

$$X_t = X_{t-1} + \epsilon_{x,t} = \sum_{j=1}^t \epsilon_{x,j} \quad (2.1.2)$$

se si volesse calcolare la regressione sul primo intervallo temporale si otterrebbe una stima di  $\alpha$  e di  $\beta$  che dipende fortemente dal periodo storico che si sta considerando. In aggiunta, il fatto che le due sequenze temporali non siano stabili fa presumere che la precedente stima possa, anch'essa, essere instabile nel tempo e, quindi, non c'è la presenza di equilibrio tra  $Y$  e  $X$ .

Di conseguenza vi è una problematica dal punto di vista inferenziale dato che la combinazione lineare tra due serie non stazionarie è tipicamente anch'essa non stazionaria.

Si ipotizzi però lo scenario in cui  $\epsilon_{y,j} = \theta\epsilon_{x,j}$ . La conseguenza condurrebbe ad ottenere  $Y_t = \theta X_t$ . Da quest'ultimo punto si deduce che  $Y_t$  e  $X_t$  condividono lo stesso trend stocastico in quanto:

$$Y_t = \theta \sum_{j=1}^t \epsilon_{x,j} \quad (2.1.3)$$

$$X_t = \sum_{j=1}^t \epsilon_{x,j} \quad (2.1.4)$$

da tale risultato si trae la conclusione che la relazione  $Y - \theta X = 0$  rappresenta un equilibrio tra le variabili, per  $\forall t$ .

Un'ipotesi meno restrittiva può essere formalizzata presupponendo che gli *shock* siano molto simili, per cui l'equilibrio viene rappresentato da:

$$\epsilon_t = Y_t - \theta X_t \quad (2.1.5)$$

nella quale il termine  $\epsilon_t$  è stazionario.

Ragion per cui si afferma che  $Y$  e  $X$  siano cointegrate qualora esista una loro combinazione lineare che sia stazionaria. La presenza di cointegrazione consente di effettuare inferenze accurate nel modello di regressione, anche quando le variabili considerate sono non stazionarie. Pertanto, è fondamentale determinare quando si verifica questa relazione in situazioni in cui le variabili non sono stazionarie. Prendendo in esame la relazione precedente, se le variabili  $X$  ed  $Y$  condividono un medesimo andamento stocastico, ciò determinerebbe la stazionarietà della loro discrepanza  $t$ , portando quindi ad un rifiuto dell'ipotesi nulla mediante un test di radice unitaria. Difatti, un metodo per accertare la presenza di cointegrazione potrebbe basarsi sulla verifica dell'assenza di radice unitaria per la serie temporale della discrepanza  $t$ .

Il termine  $Y_t - \theta X_t$  è denominato termine di correzione dell'errore, poiché questa differenza consente di eliminare il trend stocastico condiviso. Si deve ricordare poi

che, se consideriamo due variabili  $X$  e  $Y$  integrate di ordine 1 e  $X$  è cointegrata con  $Y$ , lo stimatore  $\hat{\theta}$  convergerà più rapidamente al suo reale valore rispetto al caso in cui le due variabili siano stazionarie. Inoltre, la matrice di covarianza degli stimatori sarà corretta utilizzando il metodo HAC o Newey-West che apprezzano l'eteroschedasticità e l'autocorrelazione. Tale procedura è richiesta, poiché, in generale i residui del modello sono autocorrelati.

Sfruttando la superconsistenza di  $\hat{\theta}$ , l'approccio proposto da Engle e Granger (1987) per verificare la presenza di cointegrazione consiste nel condurre un test di radice unitaria sulla serie dei residui stimati  $\hat{\epsilon}_t$ , come ad esempio il test di Dickey-Fuller per la cointegrazione (CDF). Come si vedrà nel capitolo successivo, affinché ci sia cointegrazione le variabili devono essere integrate nello stesso ordine.

## 2.2 Il Test Argumented-Dickey-Fuller

I test di Dickey-Fuller (1979) vengono impiegati per rilevare la presenza di una radice unitaria in un modello autoregressivo. E permette di distinguere se si tratta di un processo stazionario alle differenze o stazionario con *trend*. In principio tale test si fondava su modelli AR(1) i quali catturano il primo ordine di autocorrelazione. Successivamente l'applicazione del test si estese anche a modelli di ordine maggiore al primo o a modelli ARMA ( $p, q$ ), giungendo al Argumented Dickey Fuller test. Un aspetto critico dell'utilizzo dell'Argumented Dickey Fuller test riguarda la selezione del grado  $p$  del processo di selezione. Infatti, se il valore di  $p$  selezionato fosse eccessivamente elevato si avrebbe come conseguenza, una ridotta capacità del test. Al contrario, un valore flebile di  $p$  potrebbe causare correlazione seriale residua negli errori provocando distorsione. A tal proposito si fa riferimento alla procedura introdotta da Ng e Perron (1995) per determinare e identificare l'ordine del modello.

$$\Delta y_t = \alpha_0 + \alpha_1 t + \beta y_{t-1} + \sum_{j=1}^{p-1} \gamma_j y_{t-j} + \varepsilon_t \quad \varepsilon_t \sim WN \quad (2.2.1)$$

Pertanto, le ipotesi del test sono le seguenti:

$$\begin{cases} H_0: \beta = 0 \\ H_1: \beta < 0 \end{cases}$$

### 2.2.1 Il test KPSS

Il test ideato da Kwiatkowski (1993) è un test sulla stazionarietà e si fonda sul modello descritto qui:

$$\gamma_t = \beta' D_t + \mu_t + \varepsilon_t \quad (2.2.1.1)$$

$$\mu_t = \mu_{t-1} + \varepsilon_t \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2) \quad (2.2.1.2)$$

nella quale la componente  $D_t$  individua il fattore deterministico il quale può essere formato solo dalla costante o dalla costante più trend. Mentre  $\mu_t$  è un *random walk* e, infine, il termine  $\beta$  lo si deve stimare.

Dunque, il test di ipotesi che viene definito presenta come  $H_0$  e  $H_1$ :

$$\begin{cases} H_0: \sigma_\varepsilon^2 = 0 \\ H_1: \sigma_\varepsilon^2 > 0 \end{cases}$$

per valutare il sistema di ipotesi, il test KPSS utilizza il moltiplicatore di Lagrange come statistica test. Nel proseguo del loro studio gli autori dimostrano che il test KPSS converge verso una funzione del moto browniano, la quale è condizionata dalla forma dei termini  $D_t$  ma non dai valori dei coefficienti  $\beta'$ . Tale test è orientato sulla coda destra della sua distribuzione; perciò, l'ipotesi nulla di stazionarietà viene rigettata con un livello di confidenza del 100  $\alpha\%$  se la statistica del test supera il quantile 100(1- $\alpha$ ) % della sua distribuzione.

### 2.3 Il test Breusch-Godfrey

Per verificare la presenza di autocorrelazione nei residui di un modello o di una regressione si possono utilizzare diversi test come la procedura statistica di Breusch-Godfrey. È particolarmente importante nei modelli di serie storiche, nei quali l'autocorrelazione dei residui può condurre alla conclusione che il modello non sia ben specificato o siano state omesse delle variabili significative che influenzano i

risultati. È stato introdotto per superare i limiti del Durbin-Watson test<sup>8</sup>, difatti il BG può essere usato per rilevare le autocorrelazioni di ordine superiore al primo.

Il test si fonda sul concetto di regredire i residui del modello originale su una versione ritardata, fino ad un ordine  $L$  di ritardi. Il test assume la seguente formulazione:

$$e_t = a + \sum_{i=1}^L \phi_i e_{t-i} + v_t \quad v_t \sim WN(0, \sigma_v^2) \quad (2.3)$$

nella quale gli  $e_t$  rappresentano i residui del modello originale, la costante è data da  $a$ , i coefficienti per i ritardi sono raffigurati da  $\phi_i$  mentre il termine d'errore è indicato con  $v_t$ .

La determinazione dell'ordine di ritardi da considerare è un passaggio cruciale. Un numero basso potrebbe impedire la raccolta di tutta l'autocorrelazione presente. Dall'altra parte un valore elevato potrebbe portare, a causa del sovradimensionamento del modello, a risultati non efficaci.

Quindi si regrediscono i residui calcolati sulle loro versioni ritardate, fino all'ordine  $L$  stabilito. La statistica del test è calcolata come  $nR^2$ , dove  $n$  è il numero di osservazioni e  $R^2$  è il coefficiente di determinazione del modello di regressione dei residui. Il sistema di ipotesi adottato è configurato come:

$$\begin{cases} H_0 : \phi_1 = \dots = \phi_L = 0 \\ H_1 : \exists i \in \{1, 2, \dots, L\} | \phi_i \neq 0 \end{cases}$$

## 2.4 Introduzione al modello ECM

Il concetto espresso precedentemente al paragrafo (2.1) relativo alla cointegrazione è strettamente connesso al tema dell'ECM: *Error - Correction Mechanism* che in gergo è noto come modello a correzione dell'errore, il quale rappresenta un approccio per spiegare come alcune deviazioni possano essere aggiustate e riassorbite nel corso del tempo. Adottando un approccio generale, si riportano due serie non stazionarie  $X_t$  e  $Y_t$  che inserite nel modello di regressione daranno come esito dei risultati spuri.

---

<sup>8</sup> Si limita alla verifica dell'autocorrelazione di primo ordine.

Quindi, se entrambe le serie:  $X_t$  e  $Y_t$  contengono una radice unitaria  $I(1)$ , se si calcolasse una regressione:

$$Y_t = \beta_1 + \beta_2 X_t + u_t \quad (2.4.1)$$

non si potrebbero ottenere delle stime soddisfacenti per  $\beta_1$  e  $\beta_2$ .

Una soluzione per ovviare a tale problematica è indicata dalla differenziazione dei dati al fine di ottenere la stazionarietà delle variabili stesse; quindi, si deve calcolare la differenza tra i valori successivi della serie temporale. Il prodotto di tale operazione è  $\Delta Y_t \sim I(0)$  e  $\Delta X_t \sim I(0)$ , e il modello di regressione sarà così sintetizzato:

$$\Delta Y_t = \alpha_1 + \alpha_2 \Delta X_t + \Delta u_t \quad (2.4.2)$$

in questo modo il problema legato all'equazione spuria è stato risolto e, in aggiunta, si avranno delle stime precise e corrette dei parametri  $\alpha_1$  e  $\alpha_2$ . Si sottolinea come il risultato ottenuto dalla precedente equazione rappresenti solo la relazione di breve periodo tra le variabili, mentre la relazione di lungo periodo è definita da:

$$Y_t^* = \beta_1 + \beta_2 X_t \quad (2.4.3)$$

Quindi la serie  $\Delta Y_t$  non fornisce alcuna informazione circa il comportamento nel lungo periodo del modello, perciò, si configura un problema poiché in economia si è interessati alle relazioni a lungo termine. Però il concetto di cointegrazione e il modello ECM svolgono un ruolo primario in queste circostanze.

Come anticipato, entrambe le serie  $X_t$  e  $Y_t$  sono integrate  $I(1)$  e c'è una loro combinazione lineare  $I(0)$  tale per cui si può affermare che le due serie siano cointegrate. Concludendo, la serie non sarà più identificata come spuria e, inoltre, si otterrà una combinazione lineare data da:

$$\hat{u}_t = Y_t - \hat{\beta}_1 - \hat{\beta}_2 X_t \quad (2.4.4)$$

Quindi se entrambe le serie sono cointegrate e, per specificazione, il termine d'errore stimato  $\hat{u}_t$  segue un processo stazionario integrale di ordine zero,  $I(0)$ . Pertanto, è

fattibile rappresentare la connessione tra le due serie utilizzando il modello sopracitato: il modello ECM:

$$\Delta Y_t = \alpha_0 + \beta_1 + \Delta X_t - \pi \hat{u}_{t-1} + \varepsilon_t \quad (2.4.5)$$

Come si nota, sono incluse nel modello sia la componente di breve che di lungo termine. Il modello ECM è possibile esclusivamente se il disequilibrio è una funzione di cointegrazione. In questo contesto, il termine  $\beta_1$  indica il moltiplicatore di impatto ovvero l'effetto del cambiamento in  $X_t$  su  $Y_t$ . Invece, il termine  $\pi$  è l'aggiustamento, cioè la velocità di adattamento dell'equilibrio nel tempo, quindi, indica quanto il disequilibrio nel periodo precedente influenza l'aggiustamento attuale in  $Y_t$ . Si aggiunge che  $\hat{u}_{t-1} = Y_{t-1} - \hat{\beta}_1 - \hat{\beta}_2 X_{t-1}$  può essere stimato tramite i residui ed è una componente stazionaria.

Si conclude che il problema legato all'utilizzo di dati non stazionari e quindi alla regressione spuria è superato in quanto le componenti dell'equazione precedente (2.3.5) sono stazionarie, compresi i residui.

Il modello ECM è diffuso perché presenta delle peculiarità, quali:

- Si presenta come una soluzione per determinare e valutare la correzione della componente di disequilibrio riscontrata nel periodo antecedente;
- Successivamente ad aver verificato la cointegrazione, il modello ECM si presta come modello efficiente, ed essendo formulato in termini di differenze prime ammettono di eliminare le tendenze presenti nelle variabili e, quindi, di esaminare con un buon fine il tema delle regressioni spurie;
- Questo modello si integra bene nell'approccio generale-specifico in modo tale che il modello si adatti bene ai dati;
- Il termine di disequilibrio è stazionario e, quindi, si deriva il fatto che essendo le due variabili cointegrate allora vi è un processo di aggiustamento che evita uno squilibrio crescente nella relazione di lungo termine.

#### **2.4.1 Introduzione al modello ECM-COC**

Congiuntamente a quanto appena illustrato, viene qui descritto anche il modello ECM-COC ovvero un ulteriore modello econometrico per analizzare e modellare le relazioni a lungo termine di serie temporali integrate di ordine I(1). A differenza del

modello ECM, viene integrato un concetto cruciale per il trading, ossia il *cost of carry* o costo di mantenimento, il quale quantifica le spese da sostenere per “trasportare” o detenere un bene durante un determinato periodo di tempo. Sulla base delle assunzioni del modello, il prezzo dei futures di un asset non comprende solamente la valutazione prevista per il sottostante nel mercato a pronti alla scadenza del contratto, ma tiene conto anche dei costi intrapresi per mantenere o trasferire il sottostante dal momento iniziale al momento finale in  $T$ . È fondamentale porre in evidenza la categoria d’appartenenza del sottostante ossia che natura abbia lo stesso (se beni di consumo o di investimento). Si deve ammettere una distinzione qualora l’attività sottostante sia rappresentato da un bene di investimento che offre reddito o meno. Infatti, se si rientra nella prima casistica allora l’equazione che modella il *cost fo carry* è costituita da:

$$F_t = S_t e^{rT} \quad (2.4.1.1)$$

nella quale si ha il valore del contratto futures al tempo  $t$  e scadenza  $T$ . La medesima scadenza è associata al contratto spot  $S_t$  in cui si tiene conto anche del tasso *free risk*,  $r$ .

Qualora, invece, si faccia riferimento alla seconda casistica annunciata in precedenza, il costo di mantenimento incorpora anche il reddito noto e quindi si avrà:

$$F_t = (S_t - I) e^{rT} \quad (2.4.1.2)$$

la quale presenta la medesima struttura della formula (2.3.1.1) con l’inclusione del valore certo,  $I$ .

La terza situazione da evidenziare è costituita dal pagamento del dividendo; pertanto, se il titolo conferisce una percentuale del valore dello stesso si dovrà studiare la seguente formula:

$$F_t = S_t e^{(r-q)T} \quad (2.4.1.3)$$

in tal caso il nuovo elemento, rappresentato da  $q$ , è il dividendo.

Esaminando anche i beni di consumo come possono essere le *commodities* e le materie prime come il petrolio, allora la formula cambia leggermente e diventa:

$$F_t = (S_t + U)e^{rT} \quad (2.4.1.4)$$

In questo caso la componente dello stoccaggio è inclusa con la lettera  $U$ . Se quest'ultima rappresenta una proporzione del valore della merce allora sarà individuata con la lettera maiuscola e verrà sommata al tasso privo di rischio.

Il fatto di possedere il bene di consumo comporta anche il vantaggio di avere l'immediata disponibilità del sottostante ed è definito dal *convenience yield* che individua la differenza tra i futures attuale nel mercato e quello derivante dal modello del *cost of carry*. Maggiore è il valore della differenza e minore sarà il valore attribuito dal mercato al sottostante. Quindi prendendo in rassegna anche il *convenience yield* per i beni di consumo, si otterrà come formula la seguente:

$$F_t = S_t e^{(r+u-y)T} \quad (2.4.2.5)$$

Quindi esaminare anche l'aspetto del beneficio netto derivante dal detenere un asset fino alla scadenza del future è significativo perché considera delle dinamiche rilevanti per i beni di investimento e consumo. Per i primi, il *cost of carry* è la porzione di prezzo in eccesso che rende il prezzo futures maggiore di quello spot. Per i secondi, tenendo presente il *convenience yield*, il prezzo futures e spot non sono equivalenti per un ammontare pari al costo di mantenimento.

## 2.5 Introduzione al modello Arma

Nell'analisi del comportamento delle serie storiche si fa largo uso dei modelli ARMA (Autoregressive Moving Average) sviluppati da Moran & Whittle (1951). Questo modello combina due fattori, ossia: una componente autoregressiva (AR) e una componente a media mobile (MA), entrambe le componenti permettono al modello di catturare sia il breve termine che il lungo termine nelle serie storiche di conseguenza un processo ARMA( $p, q$ ) è definito come:

$$A(L)\gamma_t = C(L)\varepsilon_t \quad (2.5.1)$$

In questo contesto la dicitura  $A(L)$  sta a indicare il polinomio al quale corrisponde la lettera  $p$  che sta a indicare l'ordine del polinomio stesso. Mentre al polinomio  $C(L)$  è associata la lettera  $q$ .

Si potrebbe esprimere  $\gamma_t$  in forma MA qualora tutte le radici del polinomio  $A(L)$  siano maggiori di 1, in tal caso diventerebbe:

$$\gamma_t = A(L)^{-1}C(L)\varepsilon_t = C^*(L)\varepsilon_t \quad (2.5.2)$$

Se l'ordine del polinomio  $C^*(L)$  fosse maggiore di zero allora il polinomio è di ordine infinito. La condizione su  $A(L)$  è necessaria e sufficiente affinché il processo sia stazionario.

Invece, qualora il polinomio  $C(L)$  risulta invertibile, si deduce che la serie temporale  $\gamma_t$  può essere rappresentata da un processo autoregressivo, con ordine infinito se il termine d'errore  $q$  è maggiore di zero.

$$C(L)^{-1}A(L)\gamma_t = A^*(L)y_t = \varepsilon_t \quad (2.5.3)$$

Se una costante venisse considerata e, quindi, aggiunta al modello ARMA(p,q) si dimostra che non varia nulla dal punto di vista della media la quale rimane la medesima:  $\frac{\mu}{A(1)}$ . Infatti, tale risultato deriva dalla seguente dimostrazione:

$$A(L)\gamma_t = \mu + C(L)\varepsilon_t \quad (2.5.4)$$

$$E|A(L)\gamma_t| = \mu + E|C(L)\varepsilon_t| \quad (2.5.5)$$

Ora, per la linearità degli operatori  $E$  e  $L$  si ottiene la seguente:

$$A(L)E|\gamma_t| = \mu + C(L)E|\varepsilon_t| = \mu \quad (2.5.6)$$

Dato che  $\gamma_t$  è stazionario avremo che l'elemento  $E|\gamma_t|$  esiste costante e finito, arrivando, alla fine ad ottenere:

$$A(L)E[\gamma_t] = A(1)E[\gamma_t] \quad (2.5.7)$$

Dalla quale si ricava:

$$E[\gamma_t] = \frac{\mu}{A(1)} \quad (2.5.8)$$

Nonostante il teorema di rappresentazione di Wold<sup>9</sup> affermi la possibilità di esprimere qualsiasi processo stazionario mediante un processo a medie mobili MA, occorre notare che, dal punto di vista applicativo, sorge l'ostacolo della stessa rappresentazione di Wold, che generalmente assume una forma infinita. La serie in osservazione viene concepita come una realizzazione di un processo stocastico, il cui andamento è delineato dai coefficienti dei polinomi nell'operatore  $L$ . Questi coefficienti svolgono un ruolo chiave nella definizione delle caratteristiche di persistenza del processo, oltre alla considerazione della varianza del rumore bianco associato. Infatti, si presenta una sfida nel momento in cui si cerca di stimare un numero potenzialmente infinito di parametri facendo affidamento a un processo MA per sintetizzare le caratteristiche di media e covarianza. Questa sfida diventa evidente quando ci si trova di fronte a una serie osservata che è il risultato di un processo stazionario.

Scindendo nel dettaglio i due elementi si spiega che la componente AR è idonea a catturare la relazione che intercorre tra l'osservazione attuale e quella precedente alla stessa nelle serie storiche. In particolare, la componente AR conforma il valore corrente della serie come una somma pesata o combinazione lineare dei valori passati, in cui i pesi o coefficienti di ciascun valore passato sono determinati da un parametro di ritardo noto come ordine del modello AR ( $p$ ).

Poiché la serie temporale è registrata per  $t = 1, \dots, n$ , l'effetto ritardato della variabile esplicativa  $\gamma_{t-k}$  è osservabile solo a partire dal momento  $t = p + 1$  in avanti.

Quindi il processo stocastico definito da:

$$\gamma_t = \alpha + \phi_1\gamma_{t-1} + \phi_2\gamma_{t-2} + \dots + \phi_p\gamma_{t-p} + \varepsilon_t \quad t = p + 1, 2, 3, \dots, n \quad (2.5.9)$$

---

<sup>9</sup> Il Teorema di Wold afferma che un processo debolmente stazionario  $\gamma_t$  può essere riformulato come:  $\gamma_t = W_t + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$  nel quale  $W_t$  è la componente deterministica,  $\varepsilon_t$  è un WN e  $\psi_0 = 1, \sum_{i=1}^{\infty} \psi_i^2 < +\infty$

È un processo autoregressivo di ordine  $p$  in cui il numero di ritardi  $p$  rappresenta l'ordine massimo di ritardo dell'autoregressione. Di conseguenza per poter modellare la sequenza storica si utilizzano i ritardi.

In alternativa, si può utilizzare il *lag operator* ossia l'operatore lineare  $L$  per determinare l'operatore ritardo, in tal caso si definirebbe come:

$$L\gamma_t = \gamma_{t-1} \quad (2.5.10)$$

E generalizzando, quindi operando in modo iterativo cosicché da ripetere  $k$  volte, diventerebbe:

$$L^k \gamma_t = \gamma_{t-k} \quad (2.5.11)$$

Dalla formula (2.5.9) precedentemente illustrata si nota il termine  $\varepsilon_t$  che viene definito termine d'errore, questo è un processo *white noise* e gode della proprietà tale per cui:

$$E[\varepsilon_t \gamma_{t-k}] = 0 \quad \forall k \geq 1 \quad (2.5.12)$$

con i regressori  $\gamma_{t-k}$  sono esogeni con  $k = 1, \dots, p$ .

Il processo stocastico *white noise*, generalizzando, è strutturato da un numero infinito di variabili casuali tra di esse incorrelate a media zero e varianza costante. Il fatto che, però, siano incorrelate non guida alla conclusione che esse siano indipendenti a meno che non si stia parlando di *white noise* gaussiano.

Nei processi AR si può notare come il termine d'errore figuri come il disturbo di un modello di regressione ovvero come la differenza fra  $\gamma_t$  e la sua media condizionale. La componente di media mobile MA del modello descrive l'effetto degli errori o residui passati sul valore corrente della sequenza temporale. Nello specifico, tale componente rappresenta il valore attuale come una somma ponderata degli errori precedenti, dove i pesi o coefficienti associati a ciascun errore sono parametrizzati da un indice di ritardo denominato ordine del modello di media mobile ( $q$ ).

Quindi è l'opposto di quanto descritto precedentemente del processo AR in cui gli errori sono delineati dai valori correnti e passati della serie storica.

Il modello a media mobile può essere rappresentato dalla seguente equazione:

$$\gamma_t = \alpha + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (2.5.13)$$

Nella quale si ritrova il termine  $\varepsilon_t$  che è il rumore bianco. Tale processo si afferma essere stazionario qualora si rispettino le seguenti tre condizioni:

$$\mu = E(\gamma_t) = \alpha \quad (2.5.14)$$

$$\gamma_0 = Var(\gamma_t) = \sigma^2 \left( 1 + \sum_{j=1}^q \theta_j^2 \right) \quad (2.5.15)$$

$$\gamma_k = Cov(\gamma_k, \gamma_{t-k}) = \sigma^2 \left( \theta_k + \sum_{j=k+1}^q \theta_j \theta_{j-k} \right) \quad (2.5.16)$$

## 2.6 Introduzione al modello VAR

Avendo in mente il presupposto per il quale nello studio di fenomeni economici non ci si limita soltanto allo studio di una variabile, ma invece, si prendono in considerazione anche altre grandezze per poter descrivere accuratamente e nel dettaglio determinati scenari, si prendono in rassegna i modelli VAR che altro non sono che la generalizzazione multivariata dei processi AR.

Quindi il modello VAR non stabilisce connessioni causali esplicite tra le variabili, ma piuttosto delinea come uno *shock* causale su una variabile influisca sia su di essa che sulle altre variabili coinvolte nel sistema.

Perciò un modello VAR può essere riportato come segue:

$$\gamma_t = \nu + A_1 \gamma_{t-1} + \dots + A_p \gamma_{t-p} + u_t \quad t = 0, \pm 1, \pm 2, \dots \quad (2.6.1)$$

Nella quale l'elemento  $\gamma_t = (\gamma_{1t}, \dots, \gamma_{Kt})'$  è un vettore casuale ( $K \times 1$ ), gli  $A_i$  sono matrici di coefficienti fisse di dimensione ( $K \times K$ ), mentre l'elemento  $\nu = (\nu_1, \dots, \nu_K)'$  è un termine o più termini di intercetta, espresso come un vettore, che tiene conto della possibilità di una media non nulla  $E(\gamma_t)$ . Quindi si ottiene che il termine  $\nu = (\nu_{1t}, \dots, \nu_{Kt})'$ , ossia il *white noise*, è di dimensione  $K$ , in cui  $K$  rappresenta il numero di variabili o componenti del processo stocastico. Di conseguenza la media del termine d'errore è pari a zero:  $E(\mu_t) = 0$ , il valore atteso del prodotto esterno del termine d'errore con se stesso è pari alla matrice di covarianza:  $E(\mu_t \mu_t') = \Sigma$ , infine i termini d'errore sono incorrelati quando si confrontano in tempi differenti:  $E(\mu_t \mu_{t's}') = 0$  per  $s \neq t$ .

Prendendo in rassegna un processo VAR(1) si può dimostrare che la stazionarietà del suddetto processo è verificata quando il controllo degli autovalori della matrice  $A_1$  conduce alla conclusione che essi siano minori dell'unità. Quanto appena affermato equivale alla condizione qui riportata:

$$\det(I_K - A_1 z) \neq 0, \quad |z| < 1 \quad (2.6.2)$$

Nella quale si ha che  $I_K$  è la matrice di identità di dimensione  $K$  e,  $z$  rappresenta il vettore degli autovalori di dimensione  $K$ . In conclusione, si può riepilogare che il processo VAR( $p$ )<sup>10</sup> è stazionario qualora tutti i valori della matrice siano inferiori a uno, questa è una condizione necessaria e sufficiente.

I VAR, al pari dei modelli univariati, sono estensivamente impiegati per differenti scopi come per analizzare le caratteristiche delle serie temporali che includono e fare previsioni. Le applicazioni che trovano una maggiore applicazione comune in ambito macroeconomico sono principalmente tre:

- I. Analisi di causalità;
- II. Previsione;
- III. Analisi dinamica.

Solitamente, è piuttosto complicato definire le connessioni di causa-effetto all'interno di un'analisi empirica dei dati economici. Quando si osserva un'alta correlazione tra due variabili,  $X$  e  $Y$ , si consente di concludere che esiste una notevole tendenza da parte delle variabili a variare congiuntamente ma questo non permette di esplicitare che esiste una relazione di nessi casuali tra di loro. Potrebbero sussistere diverse spiegazioni per il manifestarsi dell'alta correlazione tra le due variabili precedentemente menzionate:  $Y$  potrebbe causare  $X$  o  $X$  potrebbe causare  $Y$  o, ancora, una terza variabile che non è stata presa in considerazione,  $Z$ , potrebbe rappresentare la causa di entrambe. La teoria statistica che studia quanto enunciato trova il suo fondamento nella seguente proporzione: la causa precede sempre l'effetto. Quindi si necessita di un periodo minimo entro il quale si osservi lo spostamento di  $Y$  e, solo in un secondo momento, si riscontra lo spostamento di  $X$ .

---

<sup>10</sup> Si può rappresentare un modello VAR di ordine  $p$  come un modello VAR(1) che prende il nome di *companion form*.

Puntualizzando si può esprimere la Granger-Casuality nella seguente scrittura:

$$X \text{ GC } Y \iff E(y_t | y_{t-1}, y_{t-2}, \dots, x_{t-1}, x_{t-2}, \dots) \neq E(y_t | y_{t-1}, y_{t-2}, \dots) \quad (2.6.3)$$

Quindi si ammette che le osservazioni passate della variabile  $X$  svolgano un ruolo fondamentale per prevedere la variabile  $Y$ <sup>11</sup>.

Tale esplicazione, in un VAR bivariato si attua semplicemente, difatti se il vettore  $z_t = (y_t, x_t)$  può assumere la forma del VAR  $\begin{bmatrix} A_{11}(L) & A_{12}(L) \\ A_{21}(L) & A_{22}(L) \end{bmatrix} \begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$  allora si ottiene che  $x$  non Granger-causa  $y \iff A_{12}(L) = 0$ . Qualora il VAR fosse stazionario non si può attuare un test  $F$  in cui  $x$  non Granger-causa  $y$ . Su questo concetto di casualità sono state mosse diverse critiche, ovvero il fatto che una variabile possa essere Granger-casuale per un'altra o meno in relazione alle altre variabili che compongono il sistema. Quindi in un sistema di ipotesi si potrebbe accettare l'ipotesi nulla di assenza di casualità in presenza di un VAR bivariato, ma allo stesso tempo, accettare la medesima ipotesi in presenza di un VAR trivariato. La seconda critica è relativa al fatto che il principio di causa effetto opera in un contesto di indipendenza degli eventi che si susseguono nel tempo fisico.

Per quanto concerne il secondo punto, ossia le previsioni, valgono le considerazioni relative ai processi AR. Dal momento che si fa uso del valore atteso condizionale come previsore, si deve porre l'attenzione sul fatto che il dataset informativo incorpori tutte le sequenze temporali che prendono parte nel VAR e non una singola serie. Quindi, stimate le matrici  $\hat{A}_1, \dots, \hat{A}_p$  la previsione dell'elemento  $\hat{y}_{T+k}$  risulterà un'estensione del caso univariato:

$$\hat{y}_{T+k} = \hat{A}_1 \hat{y}_{T+k-1} + \dots + \hat{A}_p \hat{y}_{T+k-p} \quad (2.6.4)$$

si specifica che  $\hat{y}_{T+k} = y_{T+k}$  per  $k \leq 0$ .

L'ultimo aspetto da sottoporre ad esame è l'analisi dinamica del processo VAR ossia la risposta di impulso. In questo contesto si presenta una complicazione maggiore data dalle  $n$  variabili e dagli  $n$  shock, quindi, viene il caso complesso viene della risposta d'impulso è raffigurato da una matrice  $n \times n$ . Quindi si pone l'attenzione sul valore atteso condizionale e su come questo si leghi alle variabili tramite una

---

<sup>11</sup> Se  $Y$  Granger-causa  $X$ , non si esclude a priori che  $X$  non Granger-causa  $Y$ . Oppure se  $Y$  non Granger-causa  $X$  non si respinge la possibilità che  $X$  Granger-causa  $X$ .

matrice, la quale mette in relazione errori di previsioni e *shock* strutturali. Essendo quadratica ed invertibile, la suddetta matrice, permette di ricostruire la storia degli *shock* strutturali e calcolare le risposte di impulso strutturali che offrono maggiori informazioni in quanto legano gli *shock* direttamente alle variabili osservabili. Ad ogni modo, non è agevole stimare la matrice dato che dipende dalla matrice di varianze e covarianze osservata, indi per cui si impone la stessa come una triangolare bassa.<sup>12</sup>

---

<sup>12</sup> Ossia tutti gli elementi della matrice, che in questo caso chiameremo  $C_{i,j}$ , sono nulli per  $j > i$ .

## Capitolo III

### L'analisi econometrica

#### 3.1 Metodologia adottata

Per la stesura dell'elaborato, l'analisi si è concentrata sulla letteratura disponibile nell'ambito della relazione di *lead-lag* tra i mercati spot e futures. Questi studi interessano una varietà di mercati spesso selezionati in base a criteri geografici. Questa ricerca ha permesso di comprendere approfonditamente il comportamento di *lead-lag* in diversi mercati. Si è cercato di scoprire quali modelli econometrici meglio spieghino la relazione tenendo conto anche delle principali fonti consultate per condurre lo studio. Una menzione particolare va riservata all'articolo "*A trading strategy based on the lead-lag relationship between the spot index and futures contract for the FTSE 100*" (Brooks, Rew e Ritso 2001). La replica della ricerca ha rappresentato la base portante di tutto il lavoro svolto permettendo di consolidare la conoscenza dei principali temi economici trattati, di sviluppare capacità di analisi dei dati e sviluppo di competenze base nell'ambito della programmazione in *Python*. Per quanto concerne la fase pratica di raccolta dei dati sulle serie storiche si è fatto affidamento al laboratorio Bloomberg che mette a disposizione una vasta gamma di dati finanziari ed economici su periodi temporali differenti consentendo così l'analisi di dati di tipo *intraday*. La lavorazione degli stessi è stata implementata attraverso tecniche statistiche e di programmazione; quindi, per apprendere ed esaminare le serie storiche, per poi implementare i modelli econometrici appropriati, si sono attuate analisi statistiche-descrittive di base, analisi di cointegrazione tra le variabili, trasformazioni logaritmiche, regressioni e studio dei residui delle regressioni e modelli.

Successivamente si sono sviluppati i modelli ECM, ECM-COC, ARMA e VAR con dati *in-sample* nel *training period*. In un secondo momento, con ognuno dei modelli presentati, si è passati a fare previsione sulla variazione dei rendimenti del mercato spot nel periodo *out of sample*<sup>13</sup>. Il modello che ha registrato la miglior percentuale

---

<sup>13</sup> Ossia il periodo non considerato nella costruzione dei modelli.

di corretta previsione di direzione dei rendimenti è stato impiegato per sviluppare le diverse strategie di *trading*. Inoltre, l'efficacia e la capacità dei modelli è stata soppesata attraverso l'interpretazione delle misure di valutazione quali RMSE e MAE<sup>14</sup>.

La classe proposta di strategie di trading è la medesima in Brooks et al, (2001), quindi si considerano: *Liquid trading strategy*, *Buy and hold strategy* e tre strategie di filtro: *Better predicted return than average*, *Better predicted return than first decile*, *High arbitrary cut off*. Le peculiarità di questi metodi ed i risultati verranno esaminati successivamente.

In sintesi, il metodo di lavoro adottato si è sviluppato e articolato in quattro fasi: un approccio teorico alla materia basato sulla ricerca e lettura degli articoli disponibili, selezione, studio ed elaborazione dei dati, realizzazione dei modelli e delle strategie di trading che potessero risultare profittevoli se confrontate con un investimento passivo nell'indice di riferimento.

### 3.2 Selezione e trattamento dei dati

I dati utilizzati in questo studio sono relativi ai prezzi di apertura, chiusura, massimo e minimo dell'indice FTSE100 e ai futures scritti nello stesso indice con scadenza marzo 2024. Il *database* si compone quindi di di 10.039 osservazioni<sup>15</sup> con un intervallo temporale di 10 minuti che fanno capo a tutti i giorni di negoziazione del mercato spot, ovvero del FTSE100 nel periodo che va da marzo 2023 a dicembre 2023. Nello stesso periodo di campionamento, si sono utilizzati il medesimo numero di osservazioni per i prezzi dei futures. Anche se i due strumenti finanziari appena menzionati sono correlati, si differenziano per le tempistiche di aggiornamento dei prezzi: mentre l'indice FTSE100 viene scambiato continuamente e ha una frequenza di aggiornamento di 1 minuto, i futures per loro natura non vengono scambiati continuamente e quindi i prezzi delle transazioni non sono aggiornati nel tempo in maniera sincrona. Per far fronte a questo problema si sono utilizzati i dati di

---

<sup>14</sup> Sia il Root Mean Squared Error che il Mean Absolute Error verranno discussi nel capitolo successivo.

<sup>15</sup> La statistica fa riferimento solo ai dati di spot e futures ogni 10 minuti tralasciando i dati di apertura, minimo, massimo, rendimenti, tasso privo di rischio ecc..

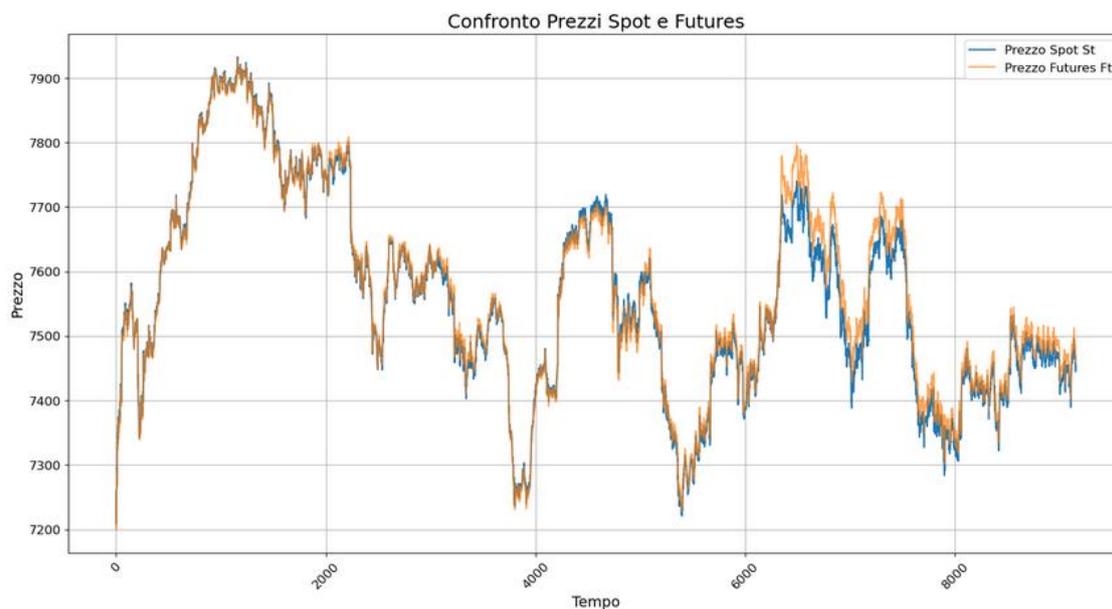
entrambe le serie storiche con un'ampiezza di intervallo di 10 minuti. Inoltre, si è considerato come tasso privo di rischio il Uk T-Bill a 3 mesi con cadenza mensile. Pertanto, per garantire coerenza con l'analisi si è convertita la media dei tassi senza rischio mensili in intervalli di 10 minuti, mantenuti costanti per l'intero *dataset*. Quindi si fa un'assunzione molto forte ossia che il tasso sia costante per tutta l'ampiezza del campione dato che l'intervallo temporale sul quale si basa l'intero studio è troppo ristretto. Lo stesso metodo di esecuzione e la medesima ipotesi sono stati replicati per i rendimenti dell'indice FTSE250 che viene usato come proxy di mercato dell'indice FTSE100 per il calcolo del modello a correzione dell'errore con integrazione del *cost of carry*.

Dato che le serie storiche del mercato spot e futures non erano allineate per diversi motivi, come, ad esempio la chiusura del mercato spot e non dei futures o viceversa; mancanza di dati reperibili in determinati orari per una delle due serie, o la maggior liquidità di uno rispetto all'altro, si sono palesati dei problemi di sincronizzazione. Di conseguenza, per ovviare alla problematica si è optato per predisporre una prima fase di pulizia dei dati, essenziale per il successo dello studio econometrico. Dunque, si è deciso di eseguire il *merge* dei dati al fine di adottare la stessa frequenza per le serie e, quindi procedere con l'eliminazione degli orari che non combaciavano, giungendo così ad avere un set di dati completo e con un indice temporale ordinato e consecutivo. In secondo luogo, si sono standardizzati i dati e, affinché il lavoro fosse svolto correttamente, si sono suddivise le osservazioni in due set: *training set* e *testing set*. Il primo utilizzato per la stima dei modelli, il secondo viene utilizzato per validare le previsioni prodotte dai modelli stimati e, in un secondo momento, per attuare le strategie di trading.

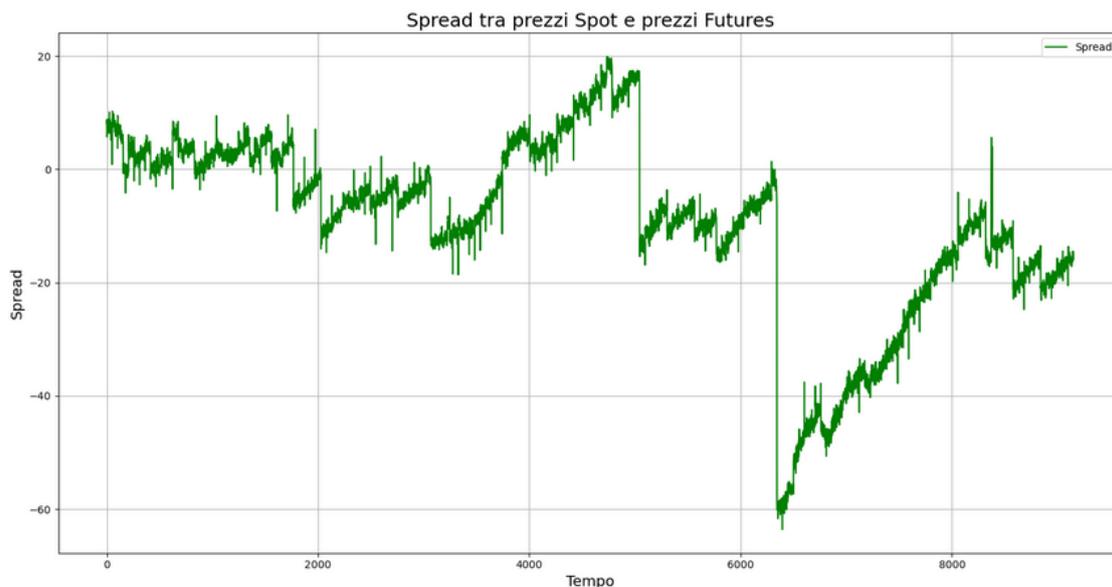
### 3.3 Statistiche descrittive di base

Come anticipato il periodo a cui si riferisce questo studio va dal 20/03/2023 al 29/12/2023 con frequenza 10 minuti, di conseguenza nelle figure seguenti si può notare il confronto dell'andamento nel tempo del prezzo spot e futures con annessa analisi visiva dello spread, ovvero della differenza di valori presentata dalle due serie.

**Grafico 3.1:** Confronto prezzo FTSE 100 e Futures



**Grafico 3.2:** Spread tra FTSE 100 e Futures



Dall'analisi dei Grafici 3.1 e 3.2 si possono interpretare diverse dinamiche che intercorrono tra le due serie: i prezzi tendono a muoversi in sincrono e nella medesima direzione ed è quello che ci si può aspettare dalla teoria dei mercati efficienti. Infatti, i valori dei prezzi a pronti e quelli futuri non dovrebbero mai discostarsi troppo l'uno dall'altro essendo i futures una misura delle aspettative sui prezzi spot. Comunque, vi sono dei momenti dove la differenza di prezzo è più marcata il che potrebbe indicare dei mutamenti del *sentiment* di mercato e delle aspettative degli investitori.

Si prosegue con l'indagine delle principali statistiche descrittive dei prezzi logaritmici e dei rendimenti, riportate nella Tabella 3.1.

**Tabella 3.1** Dettaglio delle principali statistiche descrittive per i logaritmi dei prezzi

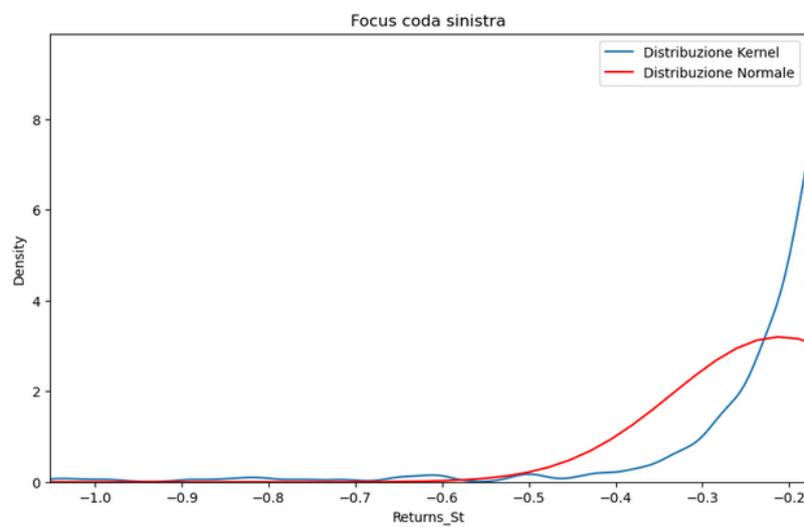
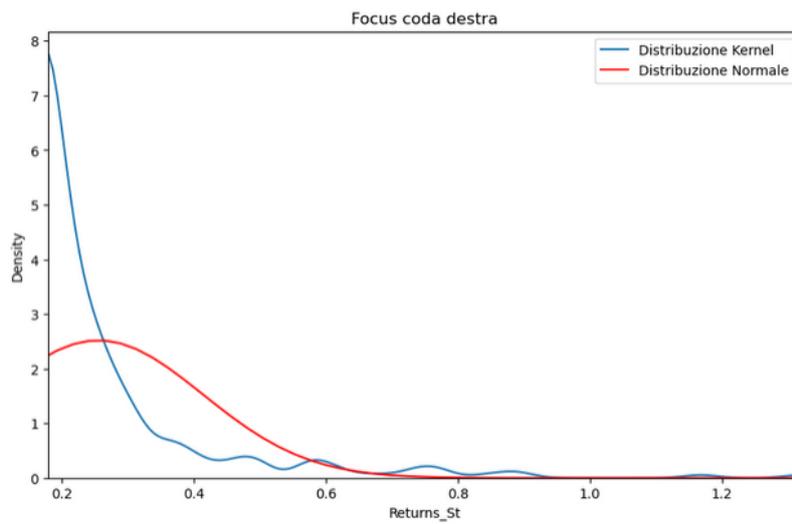
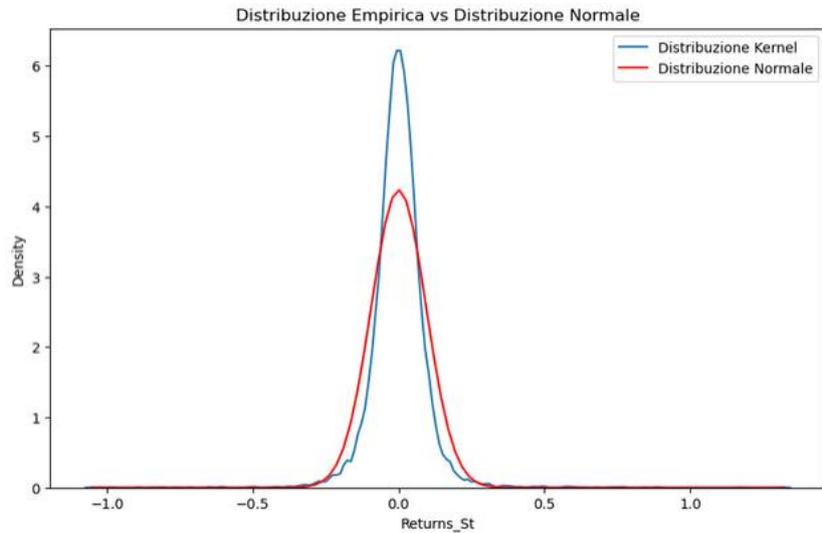
<b>Pannello A</b>	<b><math>\text{Ln } S_t</math></b>	<b><math>\text{Ln } F_t</math></b>
Osservazioni	9152	9152
Media semplice	8.930412	8.931635
Varianza	0.000424	0.000418
Standard deviation	0.020601	0.020439
Skewness	0.290086	0.217838
Kurtosis	-0.530809	-0.589662
Jarque-Bera	235.880799	205.075829
P-value	6.013611e-52	2.940015e-45

<b>Pannello B</b>	<b>Rendimenti <math>S_t</math></b>	<b>Rendimenti <math>F_t</math></b>
Osservazioni	9151	9151
Media semplice	0.000331	0.000370
Varianza	0.008872	0.009662
Standard deviation	0.09418	0.09829
Skewness	0.335334	0.246580
Kurtosis	23.293061	21.715246
Jarque-Bera	207047	179891
P-value	0.0	0.0

Si evidenzia come i valori riportati siano molto simili tra di loro a conferma che le due serie siano profondamente connesse. Si riporta per completezza anche il test di Jarque-Bera con il relativo p-value associato, convalidando il fatto che le serie in esame non possiedono una distribuzione di frequenza che segue una normale standard in conseguenza al rifiuto dell'ipotesi nulla  $H_0$ . Ciò è evidenziato anche dalla fattispecie della Skewness e Kurtosis che presentano dei valori differenti rispetto a quelli della normale che, per simmetria e curtosi, dovrebbe avere i valori rispettivamente di 0 e 3. Si noti come il valore del momento terzo, soprattutto per i

rendimenti, sia estremamente lontano rispetto alla normale e l'ispezione grafica per i rendimenti spot conferma quanto annunciato.

**Grafico 3.3** Distribuzione empirica dei rendimenti del FTSE100

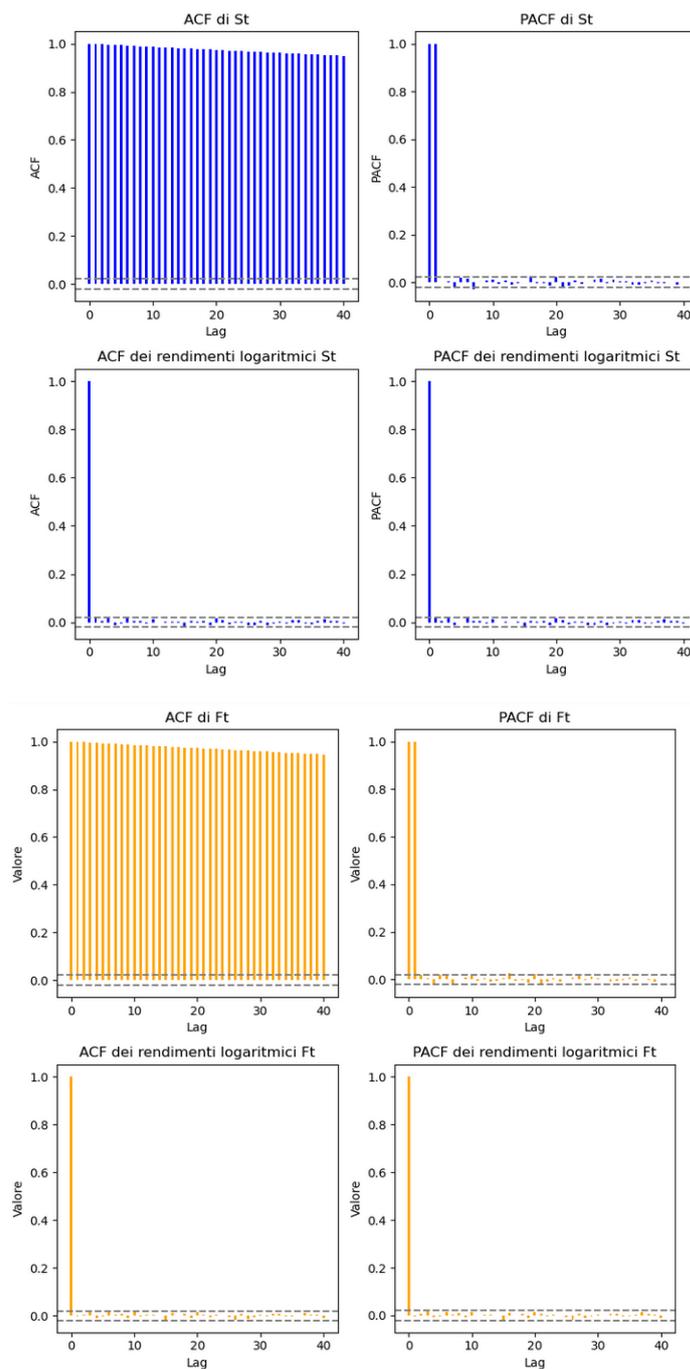


Si nota come i dati non seguano perfettamente una distribuzione normale. Infatti, il picco dei rendimenti è notevolmente più accentuato e dalle “spalle” più ristrette. Inoltre, si pone un’attenzione speciale alla distribuzione dei rendimenti agli estremi, dove il grafico della densità empirica rileva la tipica proprietà leptocurtica su entrambe le code.

### **3.4 Test di cointegrazione e Granger-Casuality**

Come già confermato in precedenza, la teoria dell’efficienza di mercato supporta la tesi che i prezzi a pronti e i prezzi *futures* siano connessi e mostrano una certa convergenza l’uno verso l’altro, riflettendo così l’efficacia delle dinamiche di mercato ad aggiornare i prezzi in virtù delle nuove informazioni acquisite. Questo aspetto suggerisce che possa esistere una relazione di cointegrazione tra le serie, (si veda Gosh 1993). Una condizione necessaria affinché si possa svolgere un test di cointegrazione è che le due serie siano integrate di ordine  $I(1)$  e che presentino stazionarietà. Tale condizione può essere stabilita in prima battuta attraverso l’analisi della funzione di autocorrelazione ACF e della funzione di autocorrelazione parziale PACF. Questa misura quantifica il grado di correlazione tra le osservazioni correnti della serie temporale e quelle precedenti, identificando quali dati storici sono più utili per prevedere con efficacia i valori futuri. Quindi se esiste una relazione tale per cui i valori presenti hanno una dipendenza in forma forte dai valori antecedenti allora si afferma che la sequenza temporale ha una correlazione robusta. Di conseguenza, si pone in esame prima l’ACF e PACF dei livelli delle serie e poi sulle trasformate logaritmiche dei rendimenti. I risultati sono riportati nel Grafico 3.4.

**Grafico 3.4** Confronto tra i grafici ACF e PACF per le serie dei prezzi e rendimenti logaritmici di  $S_t$  e  $F_t$



Dall'osservazione dei correlogrammi dei livelli delle serie spot e futures, rispettivamente  $S_t$  e  $F_t$  si nota come il valore dell'ACF sia molto elevato e positivo e la sua decrescita sia notevolmente lenta rispetto all'aumentare dei ritardi. Questo comportamento conferma il fatto che le serie presentano una forte autocorrelazione e di conseguenza i dati passati influenzano quelli presenti. Perciò, è possibile che le serie siano non stazionarie. Al contrario, i correlogrammi che prendono in esame le serie dei rendimenti logaritmici descrivono il comportamento opposto. Del resto, si

nota come il valore del correlogramma sia positivo ed elevato solamente per il primo ritardo per poi decrescere all'istante, nei limiti dell'intervallo di confidenza del 95%<sup>16</sup>, per i successivi ritardi. Si afferma quindi che non si denota un'autocorrelazione forte per la serie logaritmica dei rendimenti.

Si ipotizza la presenza di cointegrazione tra le due sequenze logaritmiche dei prezzi per le assunzioni teoriche fatte in precedenza. Considerando che la cointegrazione implica che entrambe le serie presentino il medesimo ordine di non stazionarietà e che la loro combinazione lineare conduca a stazionarietà, si procede con il test ADF – Augmented Dickey Fuller con un intervallo di confidenza del 95% per verificare tale ipotesi. I risultati sono riportati nella Tabella 3.2:

**Tabella 3.2** Risultati test ADF Rendimenti  $S_t$  e  $F_t$

<b>Pannello C</b>	<b><math>\ln S_t</math></b>	<b><math>\ln F_t</math></b>
Test ADF	-2.680320	-2.896916
P-value	0.077505949	0.0457077
Valore critico 1%	-3.431065	-3.4310648
Valore critico 5%	-2.861856	-2.8618559
Valore critico 10%	-2.566938	-2.566938
KPSS	5.039022	4.139503
P-value	0.0	0.01

<b>Pannello D</b>	<b>Rendimenti <math>S_t</math></b>	<b>Rendimenti <math>F_t</math></b>
Test ADF	-53.342353	-96.268834
P-value	0.0	0.0
Valore critico 1%	-3.431065	-3.431064
Valore critico 5%	-2.861855	-2.861855
Valore critico 10%	-2.566938	-2.566938
KPSS	0.155895	0.155802
P-value	0.1	0.1

<sup>16</sup> L'intervallo di confidenza è determinato dalle bande di oscillazione chiamate bande di Barlett. Esse delineano l'area entro la quale non si rifiuta l'ipotesi di assenza di correlazione. La costruzione di questi intervalli di confidenza si basa su un'intuizione di Bartlett, il quale ha dimostrato che un *White Noise* ha un coefficiente di autocorrelazione a cui distribuzione è riconducibile ad una normale con media zero e varianza  $1/n$ , con  $n$  numerosità campionaria

Approfondendo i risultati ottenuti per i logaritmi dei livelli del prezzo si osserva che le statistiche del test ADF per  $S_t$  ed  $F_t$  forniscono dei valori rispettivamente di -2.680320 e -2.896916 con un p-value associato alto. Perciò non si rifiuta l'ipotesi nulla di presenza di radice unitaria con un livello di confidenza del 95%. Quindi non siamo in possesso di prove sufficienti per affermare che le serie logaritmiche dei prezzi siano stazionarie ad un livello di confidenza del 5%. Di conseguenza si conferma che le serie si rivelano integrate di ordine  $I(1)$  e quindi necessitano di essere differenziate perché vi sia stazionarietà.

Invece, per quanto concerne l'analisi del test ADF sui rendimenti, si nota come il valore della statistica test sia anche in questo caso negativo per entrambe e precisamente di -53.342353 per la serie spot e di -96.268834 per la serie futures e il p-value associato è molto piccolo tanto da essere approssimato allo 0. Perciò, si conclude che l'ipotesi nulla, la quale indica la presenza di radice unitaria, viene rifiutata con estrema confidenza e perciò vi sono forti evidenze che i rendimenti siano stazionari. Queste deduzioni non variano anche considerando l'aggiunta di un numero elevato di *lags* per il calcolo del test, e comprovano quanto riscontrato in precedenza con l'esaminazione dei correlogrammi. È stato svolto, per accertarsi dei risultati, anche il test KPSS il quale è costruito diversamente rispetto all'ADF in quanto mette a sistema l'ipotesi nulla di stazionarietà e l'ipotesi alternativa di non stazionarietà (presenza di radice unitaria). Il test KPSS conferma quanto detto finora, dato che la combinazione tra i valori dei log livelli delle serie del test KPSS sono significativamente maggiori rispetto ai valori critici e i p-value sono molto bassi ci suggerisce il rifiuto dell'ipotesi nulla di stazionarietà. Al contrario, le statistiche del test per i logaritmi dei rendimenti suggeriscono la stazionarietà delle stesse convalidando le aspettative.

Quindi, visto che si ha dimostrato che le sequenze che si stanno studiando presentano lo stesso ordine di non stazionarietà, si procede con il test di Engle-Granger per verificare la cointegrazione tra il logaritmo dei prezzi a pronti e quello dei prezzi futuri. Si preferisce utilizzare tale metodologia e non quella proposta da Johansen (1988) per un ovvio motivo di semplicità e, in secondo luogo, per il fatto che vi sono solo due variabili coinvolte. La messa in atto del processo Engle-Granger si può riassumere in due passaggi:

- I. Si utilizzano le serie storiche integrate, cioè non stazionarie, per stimare il modello di regressione;
- II. Se l'esito conferma la cointegrazione e l'analisi di integrazione dei residui conferma quanto detto allora si procede con un modello dinamico ECM.

Si riporta anche la Tabella 3.3 con l'ipotesi nulla e i relativi esiti con differenti livelli di ritardi:

**Tabella 3.3** Granger Casuality tra i logaritmi delle serie

<b>Ipotesi Nulla</b>	<b>Lags</b>	<b>F_Statistic</b>	<b>P-Value</b>
LN_SPOT non Granger Cause LN_FUTURES	1	3.655436	0.055918
LN_SPOT non Granger Cause LN_FUTURES	2	9.056275	0.000118
LN_SPOT non Granger Cause LN_FUTURES	3	6.094454	0.000388
LN_SPOT non Granger Cause LN_FUTURES	4	5.356994	0.000263
LN_SPOT non Granger Cause LN_FUTURES	5	5.015263	0.000137
LN_SPOT non Granger Cause LN_FUTURES	6	4.121522	0.000388
LN_FUTURES non Granger Cause LN_SPOT	1	6.079510	0.013694
LN_FUTURES non Granger Cause LN_SPOT	2	7.864759	0.000387
LN_FUTURES non Granger Cause LN_SPOT	3	6.639429	0.000178
LN_FUTURES non Granger Cause LN_SPOT	4	4.761535	0.000776
LN_FUTURES non Granger Cause LN_SPOT	5	4.115099	0.000982
LN_FUTURES non Granger Cause LN_SPOT	6	3.396202	0.002390

L'indagine si concentra sull'ipotesi nulla, secondo la quale una delle serie temporali non causa nel senso di Granger l'altra serie, considerando diversi intervalli di ritardo. L'obiettivo è di investigare se le informazioni storiche dei rendimenti logaritmici possano essere impiegati per prevedere i rendimenti futuri. Dall'analisi emerge che nonostante il primo ritardo non sia significativo, i rendimenti spot possono esercitare un potere predittivo sulla componente futures a partire dal secondo ritardo e persistendo nel tempo. Invertendo il punto di osservazione e analizzando la casualità tra i logaritmi futures su quelli spot, emerge che la relazione casuale è significativa fin dal primo ritardo e persiste per tutti i ritardi successivi, come evidenziato dai p-value molto bassi. Quindi per entrambe le variazioni delle

serie vi è una relazione statisticamente significativa e non casuale e si estende non solo nel breve periodo ma è prolungata anche a periodi più lunghi.

Di conseguenza, prima di procedere con la stima del modello dinamico, si sono calcolate le cross-correlazioni esistenti tra i rendimenti spot dell'indice FTSE100 e quelli *lead* e *lag* dei futures per avere una panoramica generale da confrontare con i risultati della Tabella 3.3. La correlazione è molto elevata e significativa in assenza di ritardi ed è pari a 0.979 e ciò indica una fortissima relazione immediata tra i rendimenti correnti dell'indice spot e i rendimenti correnti dei futures. Mentre va a scemare in modo consistente, mantenendo sempre la significatività, al ritardo uno (0.026) suggerendo che i rendimenti passati dei futures hanno una leggera influenza sui rendimenti correnti dell'indice spot. Anche la cross-correlazione al ritardo tre (0.023) è positiva e bassa indicando una debole influenza dei rendimenti passati dello strumento derivato sui rendimenti correnti dell'indice spot. Queste evidenze suggeriscono che i rendimenti passati dei futures hanno una leggera influenza sui rendimenti correnti dell'indice spot. Al contrario le correlazioni tra i rendimenti correnti dell'indice e rendimenti futuri dei futures al contrario, non vi è nessuna correlazione significativa tra i rendimenti correnti dell'indice e i rendimenti futuri dei futures e quindi non vi sono prove sufficienti per affermare che vi sia una relazione tra i rendimenti futuri dei futures e quelli attuali spot, confermando la complessità e l'asimmetria delle dinamiche tra i due mercati. Questo suggerisce che il mercato dei futures può anticipare i movimenti dell'indice spot, sebbene l'effetto sia relativamente debole sebbene ci sia una forte sincronizzazione tra i rendimenti correnti dei futures e dell'indice spot.

**Tabella 3.4:** Correlazione fra i rendimenti dell'indice FTSE100 e il future sull'indice

	<b>Lags</b>	<b>Corr.</b>	<b>P-value</b>
$\rho(FTSE\ 100_t, Futures_{t+k})$	3	0.023	0.022
	2	0.005	0.609
	1	0.026	0.012
	0	0.979	0.000
	-1	0.006	0.75
	-2	0.007	0.49
	-3	0.01	0.063

Sulla base di queste evidenze si procede nello stimare una regressione tra i rendimenti dei due mercati, il modello è stato stimato utilizzando solamente tre ritardi:

$$S_t = a + \sum_{k=-2}^3 F_{t+k} + \varepsilon_t \quad (3.2)$$

Analogamente con quanto visto in precedenza sulla base delle semplici correlazioni, il coefficiente contemporaneo assume un valore di rilievo pari 0.938 indicando che in questa linea temporale i due mercati reagiscono simultaneamente a quasi la totalità delle informazioni. Si nota come solamente il primo coefficiente lag sia significativo ma abbastanza basso e pari a 0.0065 e ciò persegue la teoria che i futures inglobino le nuove informazioni più rapidamente. Per quanto concerne il primo coefficiente lead si nota che è significativo ma assume un valore leggermente più alto ossia 0.026.

**Tabella 3.5:** Regressione dei rendimenti del FTSE100 sui rendimenti lead-lag dei futures

	<b>Coefficiente</b>	<b>P-value</b>
$\beta_{-2}$	0.0014	0.307
$\beta_{-1}$	0.0065	0.493
$\beta_0$	0.939	0.001
$\beta_1$	0.026	0.00
$\beta_2$	-0.0014	0.00

In conclusione, si rileva che esiste una forte relazione lead-lag fra i due mercati anche se essa presenta una notevole asimmetria con i futures che anticipano i movimenti dell'indice. Come si ha ipotizzato nel primo capitolo, questo fenomeno può essere ricondotto al fatto che le informazioni sono inglobate con maggior rapidità nei prezzi degli strumenti finanziari derivati, durante la loro negoziazione. Si è notata la differenza che intercorre tra l'analisi dei prezzi con il metodo di Granger e tra i rendimenti con le cross-correlazioni mostrando un'asimmetria tra le conclusioni.

Proseguendo con lo studio della cointegrazione si riporta la formula, la quale assume questa forma:

$$\ln S_t = \beta_0 + \beta_1 \ln F_t \quad (3.3)$$

si aggiunge che per la regressione si è utilizzato uno stimatore per le varianze e covarianze robusto Newey-West<sup>17</sup> il quale è uno *standard error* robusto che tiene conto della potenziale autocorrelazione dei residui nel tempo dato che la correlazione può cambiare. I risultati sono riportati nella Tabella 3.6:

**Tabella 3.6:** Coefficienti della cointegrazione e statistiche ADF dei residui

Pannello E	Valore coefficiente
$\hat{\beta}_0$	5.787e-16
$\hat{\beta}_1$	0.9668
ADF sui residui della cointegrazione	-3.128474
P-value	0.024516
1%	-3.431065
5%	-2.861856
10%	-2.566938

Come da aspettative, esiste una forte relazione tra i logaritmi delle variabili considerate e, questo, è evidenziato anche dal coefficiente di pendenza quasi pari all'unità. Nella Tabella 3.6 sono stati aggiunti anche i valori derivanti dal test di Dickey-Fuller sui residui della regressione in quanto per determinare se due variabili sono cointegrate è necessario siano stazionari. Si nota come ad un livello di significatività dell'1% non vi siano evidenze sufficienti per respingere l'ipotesi nulla, ma con i restanti valori critici si rifiuta l'ipotesi di presenza di radice unitaria. Per cui, si evince che esiste una relazione di cointegrazione tra  $\ln S_t$  e  $\ln F_t$ .

<sup>17</sup> Se la varianza condizionale dell'errore risulta eteroschedastica non rende sospetta la consistenza degli stimatori dei parametri del modello, nonostante ciò, potrebbe condurre a stimatori non appropriati degli *standard error*. Indi per cui, nella regressione si devono stimare gli stessi in modo robusto.

### 3.5 Costruzione del modello ECM

In conseguenza ai risultati ottenuti precedentemente, nei quali si ha appurato che esiste una relazione di cointegrazione, si implementa un modello dinamico di correzione degli errori e quindi sussiste un equilibrio di lungo periodo. Ragion per cui il modello è stato sviluppato basandosi sulla seguente formula:

$$\Delta \ln S_t = \beta_0 + \delta \hat{z}_{t-1} + \sum_{i=1}^r \beta_i \Delta \ln S_{t-1} + \sum_{j=1}^s \alpha_j \Delta \ln F_t + \varepsilon_t \quad (3.4)$$

nella quale si esplicita come la variazione dei prezzi spot ( $\Delta \ln S_t$ ) sia data dalla costante ( $\beta_0$ ), dalla componente di equilibrio ovvero residui ritardati di un periodo derivati dalla stima della precedente regressione di cointegrazione ( $\delta \hat{z}_{t-1}$ ), dalla variazione dei prezzi spot ritardati di un periodo ( $\ln S_{t-1}$ ) e dalla variazione dei prezzi futures ( $\Delta \ln F_t$ ). Per determinare il numero di ritardi ottimali da considerare si è fatto ricorso al criterio di Schwarz Bayesian<sup>18</sup> con il quale si è individuato un numero di ritardi ottimali per prezzi spot e futures pari a 1. I risultati sono esplicitati nella seguente Tabella 3.7:

**Tabella 3.7:** Coefficienti modello ECM

Pannello F	Valore coefficiente
$\hat{\beta}_0$	-1.51e-07
$\hat{\delta}$	-0.0071
$\hat{\beta}_1$	0.0140
$\hat{\alpha}_1$	0.9390

Dalla lettura dei risultati ottenuti si mette in evidenza come il coefficiente stimato della costante ( $\hat{\beta}_0$ ) presenti un valore molto basso, in prossimità dello 0, e soprattutto non significativo in quanto il p-value associato è di 0.938 e quindi estremamente elevato. Il coefficiente disequilibrio ( $\hat{\delta}$ ) ha un valore negativo pari a -0.0071 ed è significativo come i restanti coefficienti. Essendo negativo suggerisce

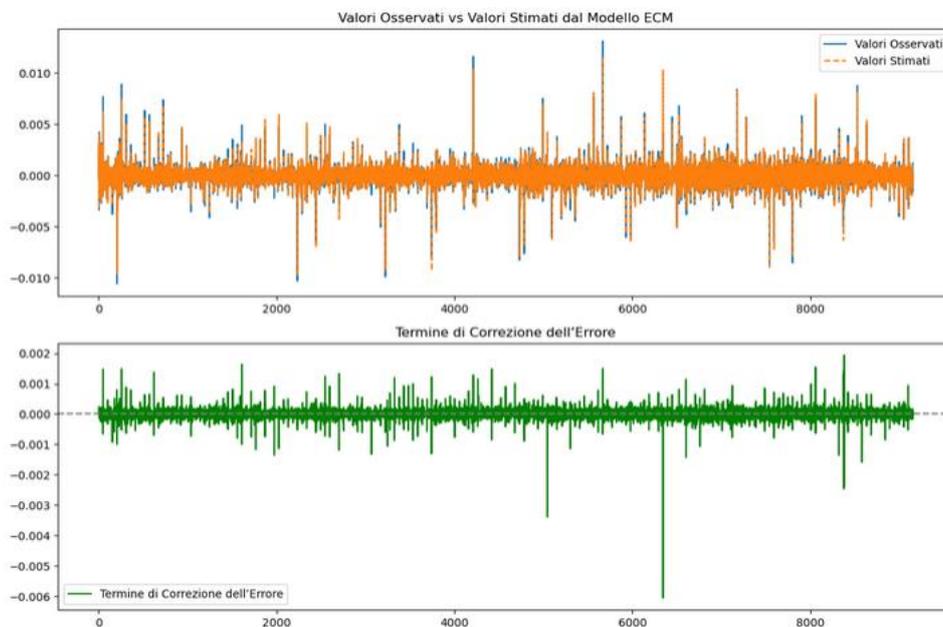
<sup>18</sup> La metodologia del BIC consiste nello stimare differenti modelli aumentando progressivamente il numero di ritardi utilizzati, per ogni modello si calcola il BIC che informa sulla complessità e sulla bontà di adattamento del modello stesso. Per determinare quale di questi è il migliore si sceglie quello che presenta il valore minore associato al criterio.

che i residui partecipano con un effetto negativo sulla variazione del prezzo. Dunque, rafforza il concetto tale per cui vi è una correzione verso un equilibrio a lungo termine. Analizzando il coefficiente relativo alla differenza dei prezzi passati di un periodo ( $\hat{\beta}_1$ ) questo è positivo e indica che le variazioni del passato influenzano le variazioni del presente in modo positivo. L'ultimo coefficiente rimanente è rappresentato dalla differenza dei prezzi futures ( $\hat{\alpha}_1$ ), anch'esso positivo e con il coefficiente maggiore marcando lo stretto legame con la variazione del prezzo a pronti.

Sintetizzando gli aspetti rilevanti dell'analisi si evidenzia come i tassi di variazione dell'indice spot siano dipendenti da tre fattori: dall'errore di cointegrazione, dalle variazioni passate dell'indice e, infine, dalle variazioni passate dei prezzi futures. Il coefficiente positivo dei futures conferma che le dinamiche dell'indice spot tendono a seguire i movimenti precedenti dei prezzi futures, rivelando quindi la capacità di questi di definire e scoprire i prezzi del mercato a pronti. Tale effetto prova quanto rivisto nella letteratura del capitolo 1. Un ultimo dettaglio riguarda il coefficiente negativo del termine di equilibrio, il quale denota che, quando  $S_t$  è consistente superando così il livello di equilibrio al tempo  $t - 1$ , le previsioni sostengono che esso diminuirà nel periodo successivo.

Dopo aver commentato i regressori non si deve trascurare la valutazione dei residui del modello per validare lo stesso. Quindi si svolge, in un primo momento, un'analisi grafica sulla relazione tra la curva stimata e quella effettiva e, in un secondo, i test di stazionarietà ADF come di consueto e, il test di LM Breusch-Godfrey Serial Correlation, per verificare la presenza o meno di autocorrelazione seriale nei residui.

**Grafico 3.5** Confronto tra curva effettiva e curva stimata e i relativi residui



Dalla Figura 3.5, si osserva come il grafico superiore mostri dei valori stimati che fluttuano intorno a zero e seguono i valori osservati in modo piuttosto preciso. Questo comportamento è ciò che ci si attenderebbe da un modello a correzione dell'errore che prova a catturare le dinamiche di breve periodo intorno ad un equilibrio di lungo termine. Le deviazioni dallo zero sono gli errori di predizione di ciascun punto temporale e non è insolito visualizzare picchi nei dati di serie temporali finanziari dato che essi presentano volatilità e sono influenzate da eventi non previsti dal modello. Si segnala come tra le deviazioni tra le due curve sono rappresentate da picchi sia positivi che negativi infatti il modello non riesce a coglierli perfettamente. Invece, dalla visione del secondo grafico, si percepisce come il termine di correzione dell'errore dimostri una certa stazionarietà essendo una serie che fluttua intorno ad una media che, nel caso specifico sembra essere zero. Quindi non vi sono presenze di tendenze o pattern nei residui.

In seguito al commento grafico si approfondisce la stazionarietà e l'autocorrelazione dei residui in modo di essere certi sulle stime efficienti offerte dal modello in questione.

**Tabella 3.8:** Statistiche test ADF e Breusch-Godfrey per i residui

Pannello F.2	Valore Statistica
Test ADF	-57.066089

1%	-3.431
5%	-2.862
10%	-2.567
Breusch-Godfrey test statistic	0.048756
P-value	0.825241

La stazionarietà dei residui è fortemente consolidata dalla statistica del test ADF che presenta un valore di -57.06 che è molto superiore ai valori critici e quindi non vi è sufficiente evidenza per accettare l'ipotesi nulla di presenza di radice unitaria. Il secondo test svolto per accertare la presenza di autocorrelazione è strutturato con un test di ipotesi che prevede l'assenza di autocorrelazione come ipotesi nulla  $H_0$ . Nel caso specifico si nota come la probabilità associata riferisce che i residui siano effettivamente incorrelati.

Quindi, il modello sembrerebbe ben specificato e stabile che cattura tutte le informazioni.

### 3.5.1 Costruzione modello ECM-COC

Un secondo modello di correzione dell'errore ma in questo caso si integra anche il costo di mantenimento del portafoglio, ossia si aggiungono alla relazione anche altri fattori come il tasso privo di rischio, i dividendi dell'indice e la differenza tra la scadenza del contratto futures e il tempo corrente di ogni singolo contratto. In primo luogo, si stima la relazione di cointegrazione come fatto per il modello ECM. In questo caso il modello varia leggermente per via dall'aggiunta dei fattori di cui sopra, e quindi si ha:

$$\ln S_t = y_0 + y_1 \ln F_t + y_2(r - d)(T - t) \quad (3.5)$$

Nella fattispecie, le componenti integrate sono: il tasso privo di rischio ( $r$ ), i dividendi ( $d$ ), il tempo a scadenza ( $T$ ) e, infine, il tempo di ogni contratto indicizzato da ( $t$ ). Rispetto all'approccio adottato precedentemente, questa equazione applica anche a relazione teorica che potrebbe incidere nei prezzi delle variabili, più precisamente potrebbe causare un potenziale contrasto tra il valore del prezzo spot e di quello futures. Si riportano i valori nella Tabella 3.9.

Anche in questo caso viene svolto il test per la verifica della stazionarietà dei residui derivanti dalla formula (3.5.1) per assicurare la presenza di cointegrazione. I residui sono individuati da:

$$\hat{z}_t = \ln S_t - \hat{y}_0 - \hat{y}_1 \ln F_t - \hat{y}_2(r - d)(T - t) \quad (3.5.1)$$

**Tabella 3.9:** Coefficienti della cointegrazione, statistiche test ADF sui residui e coefficienti modello ECM-COC

Pannello G	Valori coefficienti
$\hat{y}_0$	0.2925
$\hat{y}_1$	0.9668
$\hat{y}_2$	0.0016
ADF sui residui della cointegrazione	-3.128474
P-value	0.024516
1%	3.431065
5%	-2.861856
10%	-2.566938
$\hat{\beta}_0$	-2.264e-07
$\hat{\delta}$	-0.0070
$\hat{\beta}_1$	0.0138
$\hat{\alpha}_1$	0.9386

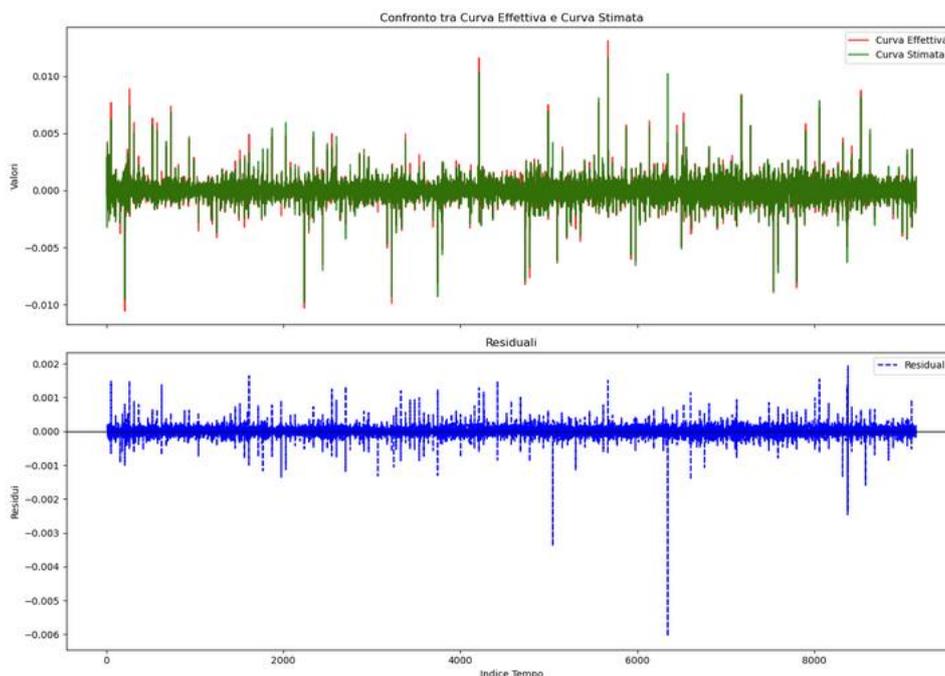
I risultati della regressione da un punto di vista analitico sono equivalenti a quelli osservati in un primo momento con la regressione (3.3). Una menzione la merita il coefficiente del *cost of carry* ( $\hat{y}_2$ ) pari a 0.0016 il quale è statisticamente significativo e positivo. Questo significa che, ad ogni incremento unitario del costo di mantenimento, vi è un aumento dello 0.16% del prezzo spot. L'incorporazione dei costi per detenere e mantenere una posizione spot costa una modifica dei valori dei coefficienti anche se la suddetta alterazione è sommamente piccola. Gli errori della cointegrazione, ossia i residui, sono stazionari ad un livello critico del 5% come dimostrato dal test ADF.

Relativamente al modello ECM-COC stimato si puntualizza come l'unico coefficiente non statisticamente significativo sia la costante. Il disequilibrio ha un coefficiente ( $\hat{\delta}$ ) negativo, come da aspettative, e descrive la rapidità con cui la variabile dipendente

si corregge e ritorna al suo stato di equilibrio persistente. Inoltre, il livello dei prezzi futures ( $\hat{\alpha}_1$ ) spicca rispetto agli altri e indica come ad un aumento dei prezzi futures ne consegua un aumento dei prezzi spot. Si rimarca il fatto che i coefficienti sono alquanto simili ai valori del modello ECM; pertanto, l'introduzione del costo di mantenimento ha un impatto complessivamente significativo, ma non è sufficientemente marcato da alterare drasticamente le osservazioni.

Anche in questo caso si sono approfondite le caratteristiche dei residui del modello stimato per verificare la correttezza dei coefficienti, pertanto si propone sia il Grafico 3.6 della serie stimata e dei valori osservati che la Tabella 3.9.1 raffigurante il test ADF con i relativi valori critici.

**Grafico 3.6** Confronto tra curva effettiva e curva stimata e i relativi residui



Come si nota, il grafico è molto simile a quello precedente quindi si rimanda al commento fatto in merito al modello ECM. Per quanto concerne l'analisi numerica sui residui si allega la Tabella 3.9.1 che riassume il test per la stazionarietà degli stessi:

**Tabella 3.9.1** Statistiche test ADF sui residui del modello

Pannello G.2	Valore Statistica
Test ADF	-57.066089
P-value	0.0

1%	3.431065
5%	-2.861856
10%	-2.566938

I dati presenti nella tabella confermano la stazionarietà dei residui in conformità con i risultati ottenuti dal test di Augmented Dickey-Fuller che presenta una statistica molto inferiore ai valori critici.

### 3.6 Costruzione modello Arma (1,0,0)

In questo paragrafo si stima un modello della famiglia ARMA al fine di offrire una base di confronto con i modelli a correzione dell'errore finora studiati. In questa sede si stima un modello ARMA( $p, q$ ) nel quale la numerosità dei termini autoregressivi è indicata dalla lettera  $p$ , mentre la numerosità che indica i termini di media mobili è individuata dalla lettera  $q$ . Anche in questo caso per stabilire il miglior numero associato ai parametri  $p$  e  $q$  si utilizza il criterio di BIC tramite la quale si trovano  $p$  e  $q$  che minimizzano l'AIC. In questo caso le grandezze rilevate sono rispettivamente di 1 e 0.

Il modello prevede di porre il rendimento spot,  $s_t$ , come variabile dipendente poiché la previsione della serie  $s_t$  è la motivazione del modello stesso. In particolare, si ipotizza:

$$s_t = \alpha_0 + \sum_{i=1}^p \alpha_i s_{t-i} + \sum_{j=1}^q \beta_j u_{t-j} + u_t \quad (3.6)$$

dalla quale si ha che  $s_t$  è la variabile dipendente che rappresenta i rendimenti della serie spot,  $\alpha_0$  è l'intercetta del modello ossia la costante, la componente autoregressiva del modello è rappresentata da  $\sum_{i=1}^p \alpha_i s_{t-i}$  nella quale si attua la somma dei prodotti degli alfa autoregressivi per le quantità passate di rendimenti logaritmici  $s_{t-1}$ , mentre la componente che assume il ruolo di media mobile è costituita da  $\sum_{j=1}^q \beta_j u_{t-j}$  anche in tal caso i risultati dei prodotti dei coefficienti di media mobile  $\beta_j$  per gli errori passati rappresentati da  $u_{t-j}$ , vengono sommati.

L'effetto sulla serie temporale attuale di un disturbo passato, nello specifico  $t - j$ , è espresso dal parametro  $\beta_j$ .

Come anticipato la componente di media mobile è nulla per via dello Schwarz Bayesian Information Criterion e quindi non viene presa in considerazione.

In sintesi, questo modello viene utilizzato per descrivere come i valori antecedenti e gli shock del passato possano incidere sulla sequenza temporale attuale.

Si riportano i risultati ottenuti nella seguente Tabella 3.10:

**Tabella 3.10:** Coefficienti modello ARMA

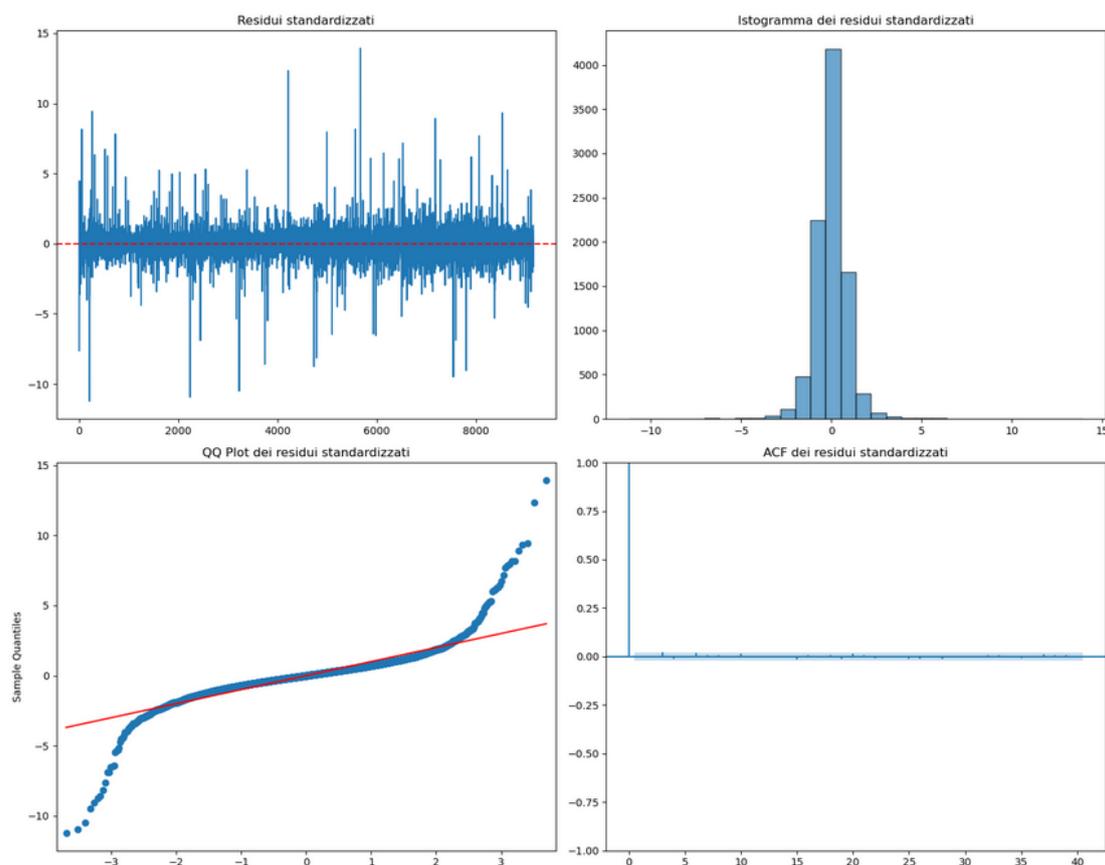
Pannello H	Valore coefficiente
$\hat{\alpha}_0$	0.0003
$\hat{\alpha}_1$	0.0170
$\sigma^2$	0.0089

Dalla lettura dei risultati si evince come, anche in questa occasione, la costante non sia statisticamente significativa con un p-value eccessivamente elevato, il termine  $\hat{\alpha}_1$ , che indica la stima del primo ritardo autoregressivo, presenta un valore modesto ma statisticamente significativo. Perciò, si afferma che esiste, tra il valore corrente e il valore istantaneamente precedente, una correlazione positiva.

Il sigma al quadrato ( $\sigma^2$ ) è la varianza del termine d'errore del modello, ossia rivela la parte dei rendimenti logaritmici spot che il modello non riesce a catturare e, in tal senso, sono una stima di 0.0089.

Per la validità del modello si riporta anche il Grafico 3.7 contenente le principali diagnosi dei residui del modello come di consuetudine per i modelli autoregressivi ARMA.

**Grafico 3.7** Principali grafici dei residui del modello ARMA



Il primo quadrante descrive l'andamento nel tempo dei residui, si osserva come fluttuino intorno allo zero e non sia presente nessun pattern il che suggerisce che la struttura temporale dei dati è stata catturata in modo ottimale. Nonostante questo, si riscontra come vi siano alcuni picchi che indicano la presenza di *outlier*. Il secondo quadrante riguarda la distribuzione dei residui, la forma assunta è simmetrica e centrata nel valore zero suggerendo che la media degli stessi è correttamente specificata. Si intuisce, tuttavia, che non sia una distribuzione normale per via dell'altezza delle barre. Il terzo quadrante confronta i quantili dei residui con quelli della distribuzione normale e si evince come i residui seguono la parte centrale della line rossa in modo pressoché impeccabile ma si discostano in entrambe le estremità il che fa intuire che hanno code pesanti. L'ultimo grafico è l'ACF nel quale si presuppone che la maggior parte delle barre sia compresa nell'intervallo di confidenza. Infatti, si nota come i residui non soffrono di autocorrelazione seriale evidenziando come il modello catturi la struttura di dipendenza nella serie. A conferma di quanto detto si allega anche la Tabella 3.11 contenente le statistiche relative ai principali test condotti sui residui.

**Tabella 3.11:** Statistiche test ADF sui residui del modello

Pannello H.2	Valore Statistiche
Test ADF	-53.338415
P-value	0.0
1%	3.431065
5%	-2.861855
10%	2.566938
Breusch-Godfrey test statistic	0.000085
P-value	0.992635

Anche in questo caso dalla lettura della statistica ADF si hanno dei residui stazionari e il test di Breusch-Godfrey non trova evidenze di autocorrelazione residua, indicando un buon adattamento del modello ARMA per quanto riguarda l'autocorrelazione dei residui e quindi il modello ha catturato in modo adeguato la dinamica temporale nella serie dei rendimenti.

### 3.7 Costruzione modello VAR

Si procede con la stima di un modello VAR (Vector Autoregressive) non vincolato, al fine di esplorare e comprendere l'effetto aggiuntivo in termini di capacità esplicativa e predittiva del termine di cointegrazione all'interno del modello ECM. Offrendo, quindi, un quadro complessivo per esaminare come la variabile dipendente sia influenzata dai ritardi della stessa più i ritardi della variabile integrata nel sistema. Nel contesto del vettore autoregressivo i rendimenti logaritmici del mercato spot sono le variabili dipendenti in modo da ampliare lo studio delle interazioni dinamiche esistenti tra i prezzi spot e futures. Il modello è:

$$\begin{cases} s_t = \theta_0 + \sum_{i=1}^p \theta_i s_{t-i} + \sum_{j=1}^q \phi_j f_{t-j} + u_t \\ f_t = \beta_0 + \sum_{i=1}^p \beta_i s_{t-i} + \sum_{j=1}^q p_j f_{t-j} + \varepsilon_t \end{cases} \quad (3.7)$$

Quindi nell'equazione come variabile dipendente si hanno i rendimenti logaritmici ( $s_t$ ) che dipendono rispettivamente dai rendimenti logaritmici spot ritardati ( $s_{t-i}$ ) e

dai futures ritardati ( $f_{t-j}$ ). Il modello permette di studiare come i valori passati della serie influenzino gli attuali valori delle stesse. Nella Tabella 3.12 seguente si esplicitano i risultati della stima:

**Tabella 3.12:** Coefficienti modello VAR

Pannello I	Valore coefficienti
$\hat{\theta}_0$	0.000355
$\hat{\theta}_1$	-0.175844
$\hat{\phi}_1$	0.188494

Dall'osservazione della tabella si rimarca come la costante abbia un valore piccolo e non statisticamente significativo. Per quanto concerne i logaritmi dei rendimenti spot ritardati ( $\hat{\theta}_1$ ) si evidenzia come questi hanno un impatto significativo negativo sui logaritmi dei rendimenti spot attuali. La conclusione relativa all'analisi dell'ultimo coefficiente stimato ( $\hat{\phi}_1$ ), attinente ai rendimenti logaritmici futures ritardati di un periodo, risulta statisticamente significativo e positivo. La matrice di correlazione evidenzia come i residui delle due equazioni siano estremamente correlati implicando, quindi, un forte movimento congiunto dei rendimenti spot e futures.

Riassumendo, l'analisi suggerisce che esistono intensi rapporti sia all'interno delle serie dei rendimenti spot e futures e sia tra di esse. Il segno opposto dei coefficienti dimostra come i rendimenti possono condizionarsi a vicenda in modo contrario ai periodi passati, evidenziando una dinamica di reversione alla media. La forte correlazione che emerge tra i residui del modello testimonia una relazione integrata tra i due rendimenti delle sequenze che il VAR prova a catturare.

Per validare il modello e, se questo sia ben specificato, si ispezionano i residui. Di conseguenza è riportata la Tabella 3.13 riassuntiva delle rilevazioni.

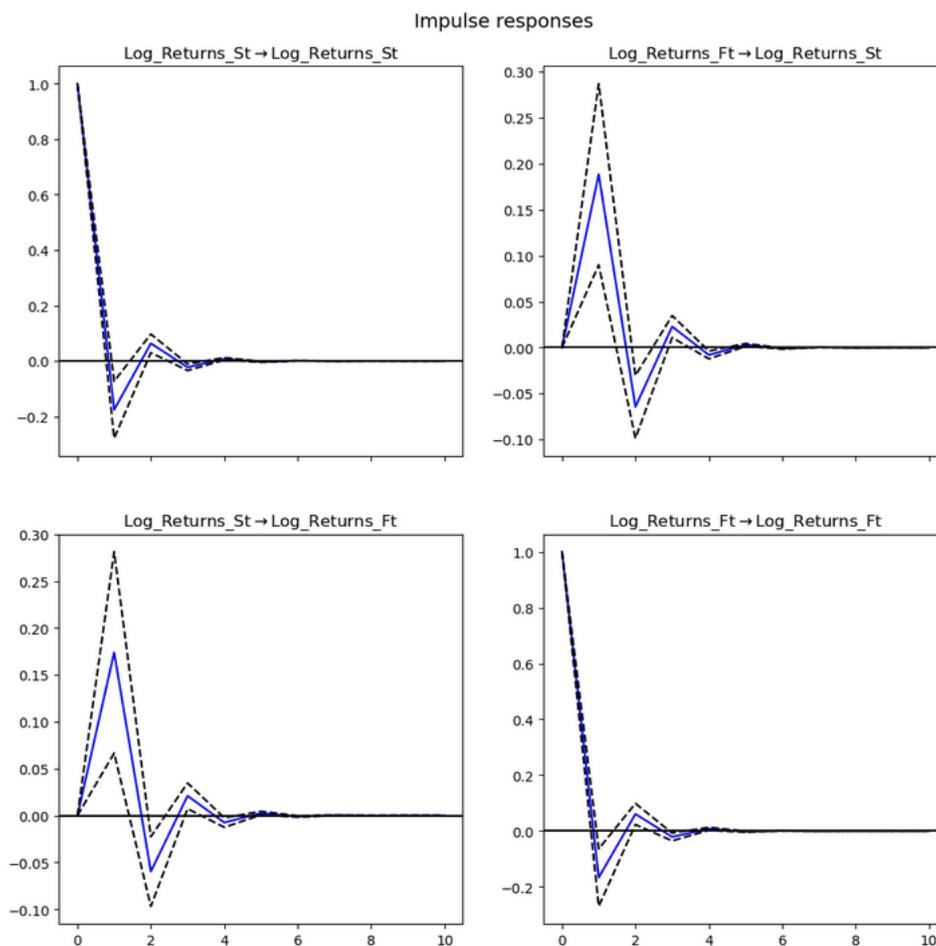
**Tabella 3.13:** Statistiche test ADF sui residui del modello

Pannello I.2	Valore Statistica
Test ADF	-53.6456
P-Value	0.0
1%	-3.43106

5%	-2.86185
10%	-2.56693
Ljung-Box test per Log_Returns_St	0.10995
P-value	0.74020
Ljung-Box test per Log_Returns_Ft:	0.08512
P-value	0.77047

In sintesi, i risultati ottenuti dall'analisi dei residui indicano che il modello VAR è ben specificato in termini di cattura delle autocorrelazioni nelle serie dei rendimenti logaritmici e quindi risulta robusto e adeguato. Inoltre, si sottolinea come, il test ADF fornisce una lettura che permette di affermare la stazionarietà dei residui stessi.

Si riportano per completezza di comprensione in termini di scoperta della relazione tra le due serie, anche i Grafici 3.8 inerenti alle risposte impulsive rappresentanti le dinamiche temporali tra i rendimenti  $S_t$  ed  $F_t$ . Questo è particolarmente rilevante quando le variabili sono integrate, cioè quando i loro valori sono influenzati dalle loro storie passate. La funzione di risposta impulsiva delinea il comportamento e l'articolazione di una variabile appartenente al sistema quando vi è uno shock nei termini d'errore dell'altra variabile integrata. Quindi, nel modello VAR, il coefficiente nella funzione di risposta d'impulso quantifica la reazione di cambiamento di una variabile stocastica ad uno shock di deviazione standard nell'altra variabile.

**Grafico 3.8** Risposta impulsiva


**Risposta di  $\text{Log\_Returns\_St}$  ad uno *shock* in  $\text{Log\_Returns\_St}$ :** mostra come la risposta sia positiva, significativa e immediata, esortando il fatto che un aumento nei rendimenti spot comporta un ulteriore aumento negli stessi in un orizzonte temporale ristretto. Lo *shock* non è duraturo, infatti la risposta termina rapidamente fornendo la dinamica di *mean-reversion* ossia i rendimenti spot tendono a ritornare, con rapidità, alla loro media.

**Risposta di  $\text{Log\_Returns\_St}$  a uno *shock* in  $\text{Log\_Returns\_Ft}$ :** segue la medesima evoluzione; infatti, si percepisce lo *shock* nei rendimenti futures sui rendimenti spot abbia un effetto positivo ma anch'esso presenta una peculiare velocità a ritornare verso la media. Confermando la breve durata degli eventi nel mercato futures sui rendimenti spot anche se quest'ultimi sono influenzati dai movimenti o aspettative dei rendimenti dei derivati.

**Risposta di  $\text{Log\_Returns\_Ft}$  a uno *shock* in  $\text{Log\_Returns\_St}$ :** nel breve termine i rendimenti futures subiscono una suggestione degli *shock* dei rendimenti spot. Gli

*shock* vengono assorbiti celermente dal mercato futures che si adegua alle informazioni.

**Risposta di Log\_Returns\_Ft a uno *shock* in Log\_Returns\_Ft:** un aumento nei rendimenti causa un ulteriore aumento degli stessi nel breve periodo il che conferma l'autocorrelazione positiva nel mercato. Tuttavia, anche questa volta si ritorna velocemente verso la media in quanto i rendimenti si stabilizzano.

Quindi si evince come il mercato abbia la capacità di ritornare ad uno stato di equilibrio nonostante gli shock possano condurre a brevi periodi di incertezza e quindi a volatilità in un primo momento. Inoltre, a convalidare le supposizioni e le teorie esposte durante il primo capitolo di questo scritto, si aggiungono anche le argomentazioni dedotte da tali esiti come, appunto, i movimenti del mercato futures influenzino e guidino quello spot.

A questo punto, si è svolto il test di casualità di *Engle-Granger* per verificare ulteriormente se la variabile indipendente eserciti degli effetti sulla variabile dipendente. L'ipotesi nulla assume la congettura che la variabile di impulso non abbia nessun tipo di capacità di influenzare la variabile dipendente; quindi, i coefficienti dei valori *lagged* della variabile impulso siano pari a zero. In calce si riportano i risultati.

**Tabella 3.14:** Granger causality per il modello VAR

<b>Pannello I.2</b>	
<i>H</i> <sub>0</sub> : Log_Returns_Ft non Granger-Causa Log_Returns_St	
Statistica Test	14.05
Valore Critico	3.842
P-value	0.0
<i>H</i> <sub>0</sub> : Log_Returns_St non Granger-Causa Log_Returns_Ft	
Statistica Test	10.08
Valore Critico	3.842
P-value	0.002

L'interpretazione della tabella delinea come si rifiuti l'ipotesi nulla con una significatività al 95% in ambo i casi, quindi, implica una relazione bidirezionale di casualità di Granger tra i rendimenti. Perciò entrambe le serie contengono

informazioni che potrebbero consentire di prevedere l'altra e, quindi, i mercati sono integrati a conferma di quanto già ampiamente discusso nei paragrafi e capitoli precedenti.



## Capitolo IV

### Previsioni un passo in avanti e definizione delle strategie di trading

In questa parte del lavoro ci si focalizza su due temi centrali: sul realizzare e valutare, mediante un insieme di indicatori di *performance*, le previsioni dei modelli discussi nel Capitolo III e sulla costruzione e verifica delle strategie di trading applicate in un contesto realistico.

Si ricordi che al principio dell'analisi si è suddiviso il *dataset* in due sottoperiodi: il primo, nel quale sono state svolte tutte le analisi descrittive e la stima dei modelli, che va dal 20 marzo 2023 al 30 novembre 2023, e il secondo periodo che fa riferimento al futuro virtuale, che prende in esame tutti i giorni di negoziazione del mese di dicembre dello stesso anno. La definizione data poc'anzi deriva dal fatto che si hanno a disposizione le osservazioni di tutte le variabili di interesse che sono coinvolte nello studio<sup>19</sup>. Ovviamente questo lasso temporale risulta di cruciale importanza per lo studio, dato che ogni singola previsione sarà confrontata con l'osservazione corrispondente.

Il secondo punto trattato in questo capitolo è lo sviluppo di tecniche di trading basate sulle previsioni dei rendimenti spot stimate dai modelli, e le stesse strategie sono tratte da Brooks et al(2001) e presentano le stesse regole e medesimi vincoli che verranno discussi in modo approfondito in seguito dato che uno degli obiettivi di questa tesi è quella di attuare delle strategie di investimento che risultino economicamente vantaggiose e redditizie, sfruttando le potenzialità del miglior modello stimato.

#### 4.1 Prefazione previsione statica e dinamica

L'analisi delle serie temporali mira primariamente ad elaborare proiezioni sull'andamento futuro della variabile d'interesse, rappresentata con  $Y$ , proiettata per un tempo futuro  $Y_{t+m}$ , nel quale l'indice  $m$  deve essere maggiore o uguale di 1 e rappresenta l'orizzonte futuro. Avendo a disposizione una serie di stime come

---

<sup>19</sup> Per quanto si può affermare sul futuro reale, vero e proprio, questo avrà luogo dal momento in cui non si dispone dei dati relativi alle variabili, quindi sia alla variabile esogena che endogena.

$\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ , la proiezione futura di  $Y$ , per  $m$  periodi successivi, effettuati al tempo  $t$ , e denota nella seguente notazione:  $Y_{t+m|t}$ . Nel corso del capitolo si discuterà anche dei diversi criteri presenti per valutare l'efficienza e la precisione della stessa previsione rispetto a un quadro reale. Si anticipa che, una previsione a  $m$  periodi futuri, si valuta come ideale qualora essa minimizzi il *Mean Squared Prediction Error* il quale è espresso nella forma di  $E[\epsilon_{t+m}^2] = E[(Y_{t+m} - Y_{t+m|t})^2]$ . Presumendo l'ipotesi tale per cui l'aspettativa condizionale dell'errore atteso  $E[\epsilon_t | Y_t, Y_{t-1}, Y_{t-2}, \dots]$  sia considerata nulla, si potrebbe concludere che la previsione ottimale e maggiormente accurata si ottenga calcolando l'aspettativa condizionata  $E[Y_{t+m} | Y_t, Y_{t-1}, Y_{t-2}, \dots]$ .

Le previsioni si dividono in due categorie:

- **Previsione Statica:** questo approccio prevede un continuo aggiornamento delle informazioni storiche della variabile  $Y_t$  ogni qualvolta che una nuova informazione si aggiunge al set esistente. Quindi qualora si volesse prevedere il periodo quattro, ossia  $Y_{t+4}$ , si potrebbero incorporare dati fino al periodo precedente  $Y_{t+3}$ , adottando quindi un meccanismo di previsione che avanza un passo alla volta. Per cui, allegare la variabile endogena in tal procedura permette un di ottenere delle previsioni meno robuste ma migliori e questo lo si deve al fatto che si avvale della tendenza auto-esplicativa della variabile.
- **Previsione Dinamica:** questo metodo consiste nel prevedere comportamento futuro della variabile di interesse  $Y_t$  utilizzando unicamente i dati reperibili fino al momento  $t$ . Quindi, per prevedere il periodo quattro, indicato come  $Y_{t+4|t}$ , si escludono dal calcolo le informazioni che emergono dopo il periodo corrente, ossia posteriormente a  $Y_t$ , perciò sono respinte le informazioni ai periodi  $Y_{t+1}, Y_{t+2}, Y_{t+3}$ . In tal modo, la previsione, conseguentemente al fatto che si fa riferimento alla relazione con i dati storici dell'endogena, sarà meno precisa.

Per valutare in modo corretto quanto una previsione può essere interpretata come corretta ed efficace, si è fatto uso di un insieme di misure atte a valutare la qualità delle stesse. Quindi, si riportano le grandezze che si sono adottate con il relativo metodo di calcolo:

L'Errore Assoluto Medio o *Mean Absolute Error* che segue la seguente formulazione:

$$\frac{1}{n} \sqrt{\sum_{t=T+1}^{T+h} |\hat{x}_t - x_t|} \quad (4.1)$$

L'Errore Quadratico Medio o *Root Mean Squared Error* che segue la seguente formulazione:

$$\sqrt{\frac{\sum_{t=T+1}^{T+h} (\hat{x}_t - x_t)^2}{h}} \quad (4.2)$$

Infine, si fa uso della Percentuale di Direzione Corretta ossia quante volte la previsione è concorde in segno con i valori realmente osservati in quel preciso istante.

## 4.2 Creazione e valutazione delle previsioni

Sono state calcolate le previsioni un passo in avanti per i rendimenti dell'indice spot utilizzando le 888 osservazioni a 10 minuti per il mese di dicembre 2023 che, come anticipato, non erano incluse nell'insieme dei dati originali usati per le stime. Quindi sfruttando ogni modello stimato si è proceduto a prevedere il valore successivo in ordine temporale. Tale valore futuro è stato quindi incluso per effettuare la previsione successiva. Perciò, ogni osservazione stimata è stata inclusa nel processo per stimare la previsione un periodo in avanti. Come confermato da analisti ed economisti, il mese di dicembre è spesso considerato un mese rialzista per via del così detto "rally di dicembre" e per le chiusure contabili delle società per l'inizio del nuovo anno; pertanto, i prezzi sono in aumento più di quanto non succeda negli altri mesi dell'anno. Essendo il mercato in aumento, le strategie di trading che si andranno a sviluppare nelle prossime pagine, si confrontano con una verifica più stringente rispetto ad uno scenario di mercato statico o regressivo, essendo che il mercato è crescente.

La Tabella 4.1 riporta le statistiche discusse nel paragrafo 4.1 per la validazione delle previsioni. Si nota fin da principio che i modelli non presentano una percentuale di

direzione corretta alta in quanto si hanno solamente due modelli con un valore dell'indicatore superiore al 50%. Questa percentuale non elevata potrebbe sorprendere, anche se i modelli sono stati implementati per fare previsione su una grandezza che è incorrelata nel tempo, e quindi imprevedibile, e che subisce delle variazioni notevoli nonostante il lasso temporale considerato sia effettivamente ridotto. L'incorrelazione è assunta come un dato di fatto nella teoria dei mercati efficienti, nel quale i prezzi riflettono le nuove informazioni che giungono in modo casuale e imprevedibile. Perciò, una percentuale superiore, seppur leggermente, indica che i modelli riescono a cogliere una struttura nei dati storici con una certa prevedibilità. Se quest'ultima è sistematica allora la si può sfruttare per attuare delle tattiche di trading che condurranno ad un rendimento atteso positivo anche se questo non sarà verificabile ad ogni singola operazione.

D'altra parte, la statistica non deve trasmettere sfiducia nell'efficacia dei sistemi di trading basati sui modelli. Anche se i modelli statistici possono teoricamente avere performance non ottimali, nella pratica possono comunque generare guadagni significativi. Studi come quelli condotti da Gencay et al. (2003) e Thomaidis et al. (2011) evidenziano come modelli di trading in tempo reale e strategie di arbitraggio statistico possano effettivamente produrre rendimenti annualizzati positivi, nonostante le complessità e le dinamiche intrinseche dei mercati finanziari. A supporto di quanto detto si riporta lo studio condotto da Liu, Zeng e Duan (2013)<sup>20</sup> nel quale si svolge un'analisi attraverso l'adozione di un approccio Bayesiano basato sulle indicazioni degli analisti di mercato per prevedere con successo i rendimenti azionari successivamente all'annuncio degli esperti. La media delle previsioni corrette che si ottiene si assesta in un range che va dal 58,6% al 60,3% quindi leggermente superiore a quanto ottenuto qui. In aggiunta, anche le raccomandazioni sui mercati finanziari degli esperti, anch'esse fondate su risultati derivanti da processi di previsione, non sono esaltanti. Le molteplici motivazioni risiedono nei medesimi fattori che si sono riscontrati nello svolgersi di tale elaborato, ossia: la complessità del mercato azionario, le tendenze del mercato stesso e il periodo preso in esame. Ma nonostante questo si sottolinea come il metodo proposto per svolgere

---

<sup>20</sup> Duan, J., Liu, H., & Zeng, J. (2013). Posterior probability model for stock return prediction based on analyst's recommendation behavior. *Knowledge-Based Systems*, 50, 151–158.

le previsioni abbia ottenuto in due casi su cinque una precisione superiore alla situazione di mera casualità nel prevedere correttamente l'orientamento dei cambiamenti dei rendimenti dell'indice. Quindi, anche la relativamente bassa precisione delle previsioni potrebbe, come già enunciato, possedere la potenzialità per sviluppare delle strategie promettenti e redditizie se applicate ad un orizzonte temporale di medio-lungo termine.

**Tabella 4.1:** Risultati previsioni

	ECM	ECM-COC	ARIMA	VAR
%Direzione Corretta	53,83%	53,94%	48,99%	45,05%
RMSE	0,1211	0,1211	0,1210	0,1211
MAE	0.064168	0.064175	0.064392	0.064164

La percentuale di direzione corretta indica quanto frequentemente il modello ha previsto correttamente la direzione di cambiamento dei rendimenti dell'indice spot. In tal senso i modelli che integrano la componente di correzione dell'errore sono migliori rispetto agli altri due e mostrano delle percentuali lievemente superiori al comportamento casuale. Inoltre, l'aggiunta della componente del costo di trasporto evidenzia una maggiore capacità di catturare la dinamica del mercato rispetto al modello *standard* di correzione dell'errore. L'ARIMA e VAR si definiscono come dei modelli non in grado di fornire delle performance degne di nota in quanto la percentuale di direzione corrette catturate da esse sono inferiori alla situazione di mera casualità, anche se per l'ARIMA è leggermente al disotto della soglia media. Quindi, esortano la loro non efficacia nel cogliere i cambiamenti dei rendimenti a 10 minuti.

Analizzando le misure di *performance* delle previsioni effettuate si nota come l'RMSE sia pressoché identico a tutti i modelli suggerendo che tutti possiedono una precisione simile in termini di magnitudine degli errori su base quadratica. Quindi nessun modello proposto può considerarsi migliore dalla lettura di questa statistica. La visione dell'errore medio assoluto ossia l'errore medio in un'unità della variabile prevista senza tener conto la direzione dell'errore stesso, riflette valori molto simili

anche in questo frangente. Perciò anche considerando questa misura non si elegge vincitore nessun modello tra quelli presenti.

Come si è visto nessun modello eccelle sugli altri, anche se i modelli con la componente di correzione dell'errore risultano migliori in termini di previsione dell'orientamento dei rendimenti, tutti i modelli hanno statistiche di prestazione affini. Questo potrebbe convincere che l'ECM ed ECM-COC siano maggiormente affidabili per sviluppare delle strategie di trading che mirano a trarre profitti sulla direzionalità; dunque, viene meno l'accuratezza delle previsioni in sé. L'attenzione si focalizza sulle previsioni della direzione di mercato poiché si presume che i costi di mantenimento ricoprano un ruolo fondamentale nel guidare i mercati spot e futures e, di conseguenza, nell'influenzare le decisioni economiche degli investitori.

Quantificare l'accuratezza dei risultati previsionali prodotti dai modelli addestrati con metodi statistici tradizionali non necessariamente garantiscono un profitto quando esse sono applicate in ambienti di trading. In effetti, modelli che statisticamente risultano inadeguati o con basse prestazioni dal punto di vista statistico, possono condurre a guadagni significativi quando sono impiegati in operazioni di trading. Tanner e Leitch (1991) hanno dimostrato che i modelli che riescono a prevedere in modo quantomeno efficace la direzione futura dei rendimenti o di individuare i momenti di inversione delle serie si sono rivelati essere più remunerativi. A questo si affianca anche una pubblicazione relativa alle strategie di trading casuali ovvero si riporta come anche le strategie frutto di studi, esperienze e conoscenze possono risultare compatibili in termini di efficacia con l'aleatorietà « le strategie di trading e i loro algoritmi, basati sulla storia passata delle serie temporali, sebbene occasionalmente abbiano la possibilità di essere di successo in piccole finestre temporali, su larga scala temporale si comportano in media non meglio della strategia puramente casuale, che, d'altra parte, è anche molto meno volatile » (A. E. Biondo, 2013).

### **4.3 Integrazione delle previsioni statistiche nelle strategie di trading**

In seguito all'analisi per stabilire quale sia il modello migliore nel compiere le previsioni e, dato che nessuno dei modelli proposti sia maggiormente accurato in

termini assoluti rispetto agli altri, si è optato di prediligere il modello con una percentuale migliore di segni previsti: ECM-COC. Quest'ultimo prenderà parte e sarà testato in diverse strategie di trading i cui risultati saranno confrontati con un investimento passivo, della medesima durata e stesso periodo, nell'indice FTSE100. Verranno inclusi anche i costi da sostenere per ogni trade effettuato e le tasse su ogni profitto derivante dall'attività, questo per ricreare una situazione quantomeno comparabile con un'ambiente realistico e per consentire un giudizio essenzialmente ragionevole e severo sulla bontà di guadagno di ciascuna tecnica presa in esame in questo elaborato.

Quindi l'ECM-COC fornisce previsioni un passo in avanti per ogni intervallo di 10 minuti presente nei giorni di trading nel mese di dicembre, ossia il periodo di trading vegliato. L'approccio adottato si avvale delle previsioni sull'indice spot e di implementazioni di regole precise sulla metodologia di esecuzione e decisione di ogni passaggio della strategia.

L'ammontare disponibile pensato per iniziare è di 1000 euro, i rendimenti sono cumulativi e qualora l'indice sia stabile e non presenta variazioni, l'investimento aumenta per una porzione pari al tasso privo di rischio e l'importo che verrà investito in ogni operazione terrà conto delle variazioni sia dell'indice che del tasso privo di rischio. Le strategie che verranno attuate sono cinque: *Liquid trading strategy*, *Buy and hold strategy*, *Better predicted return than average*, *Better predicted return than first decile*, *High arbitrary cut off*.

#### **4.3.1 *Liquid trading strategy***

Secondo le regole di questa strategia si prevede di effettuare un'operazione completa di acquisto e vendita, un *trade*, ogni volta che la previsione del modello identifica un rendimento positivo dell'indice. Se, al contrario, la previsione fosse negativa allora non viene eseguita nessuna operazione e l'ammontare fino a quel momento posseduto guadagna il tasso *risk free*. In tal senso, si impone il vincolo di non eseguire operazioni *short* ma ci si limita ad operare mediante l'assunzione di posizioni neutre o *long* sull'indice. Questa peculiarità non è totalmente inusuale in quanto mantenere delle posizioni corte sull'indice risulta economicamente molto dispendioso.

### **4.3.2 Buy and hold strategy**

Lo scopo di questa strategia è quello di effettuare un numero di operazioni di *trade* minore al fine di minimizzare i costi derivanti dalle transazioni. Perciò, se la previsione del modello risulta positiva anche per il periodo successivo a quello nel quale si ha assunto una posizione nell'indice, allora la si mantiene aperta senza doverne aprire un'altra. Quindi, le posizioni si mantengono aperte fin quando le previsioni continuano ad individuare dei rendimenti positivi, nel momento in cui la previsione individua un rendimento negativo allora si chiude la posizione. In tal modo si evita di effettuare un *trade* completo ad ogni periodo di 10 minuti e si risparmiano i costi di transazione e si massimizzano i rendimenti positivi consecutivi previsti dal modello.

### **4.3.3 Better predicted return than average**

In questo caso si applica una strategia con filtro: si acquista solo se la media dei rendimenti previsti è inferiore alla previsione. Di conseguenza, si escludono dal calcolo della media tutte quelle previsioni negative. Il filtro, o soglia, rappresentato dalla media, in questo caso, è pari a 0.24%. Quindi si entra nella posizione e la si mantiene fino a quando la previsione sarà al disotto del filtro stabilito e, in questo caso, si chiuderà la posizione. La definizione di questa strategia devia da quella precedente solamente per l'introduzione del filtro ma nell'essere applicata è identica alla *buy and hold*.

### **4.3.4 Better predicted return than first decile**

La strategia qui proposta è pensata per limitare l'operatività e ridurre il rischio tramite dei periodi di raffreddamento e la soglia dinamica basata sulla volatilità storica. Si introduce la soglia del decimo percentile per permettere di selezionare solo i segnali di trading più promettenti.

Tale strategia prevede l'utilizzo di un filtro/soglia pari al decimo percentile delle previsioni positive ed è utilizzato come soglia statica per decidere se effettuare un *trade*. In aggiunta si fa uso anche della deviazione standard mobile delle previsioni, usando una finestra dinamica di 20 periodi e di un periodo di stallo pari a 8 periodi. In ogni iterazione si valuta se la previsione è all'interno del periodo di

“raffreddamento”, se è escluso dallo stesso e se è maggiore alla soglia del primo decile allora si entra nel mercato. Se non si detiene nessuna posizione o se la differenza tra la previsione attuale e quella dell’ultimo scambio supera la soglia dinamica, si assume una posizione. Il filtro, in questo segmento, è di 0.48%.

Le regole delineate da questa strategia sono uguali a quelle precedenti ma differiscono solamente per il filtro applicato. In questo caso non ci si limita ad imporre come ostacolo la media, bensì si utilizzano solamente i rendimenti positivi del primo decile. I decili sono organizzati in ordine decrescente quindi il primo decile indica l’insieme dei rendimenti con maggior valore.

#### 4.3.5 High arbitrary cut off

Per impostare una regola di filtro eccessivamente stringente viene predisposta una barriera di vaglio in libero arbitrio pari allo 0,80%. Quindi si utilizza un *cut off* per decidere quando aprire e chiudere posizioni di mercato garantendo al contempo un ritorno pari al tasso privo di rischio quando non sono presenti segnali di trading sufficientemente forti.

#### 4.4 Risultati empirici derivanti dall’attuazione delle strategie di trading

Per l’attuazione delle strategie si sono considerati anche le spese che si devono sostenere per ogni operazione di acquisto e vendita, quindi un singolo *trade*. Non avendo avuto accesso diretto alla maggior parte delle voci dei costi per ambo i mercati, si è fatto uso delle stime di commissioni proposte dallo stesso Brooks nel *paper* di riferimento.

**Tabella 4.2:** Stima dei costi per transazione con aggiunta delle tasse

	Asset FTSE 100	
	Spot (%)	Futures (%)
Bid-ask spread	0.8	0.083
Stamp duty	0.5	0
Commissioni	0.4	0.033
Tasse sul guadagno	0.2	0

Tasse sul risk-free	0.1	0
Costi totali	2.0	0.116

**Fonte:** *A trading strategy based on the lead-lag relationship between the spot index and futures contract for the FTSE 100 (Brooks et al. 2001), più elaborazione propria.*

Per garantire una simulazione realistica, si sono introdotte due tipologie di tasse. La prima tassa si applica sui guadagni derivanti dalle operazioni sul FTSE100, mentre la seconda è definita esclusivamente sui guadagni ottenuti dall'investimento nel tasso privo di rischio. Questo approccio assicura che i risultati riflettano accuratamente le condizioni fiscali ipotetiche che gli investitori incontrerebbero. Nel mercato dei derivati, le transazioni tendono ad essere più economiche rispetto a quelle del mercato spot. Questo vantaggio economico proviene dal fatto che, nei futures, si opera su indici che non sono tangibili. Al contrario nel mercato spot è necessario negoziare ciascun componente dell'indice individualmente, il che comporta costi ovviamente più elevati.

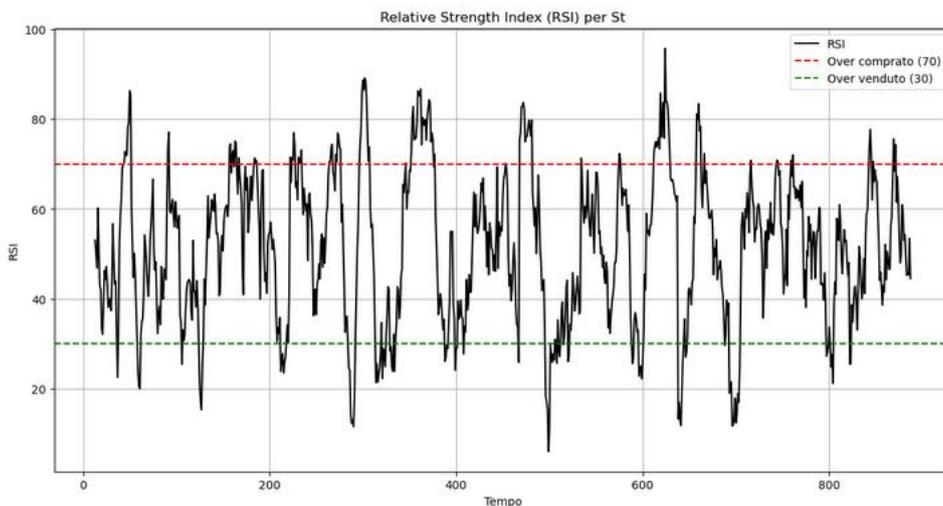
Per validare ogni singola strategia si è costruito un *performance report* al fine di valutare e analizzare le prestazioni della strategia e i risultati ottenuti, individuando le aree di successo e quelle suscettibili di miglioramenti. Gli indicatori principali presentati sono:

- I. Numero totale di trade effettuati;
- II. *Win Rate*: il rapporto tra il numero di operazioni vincenti e il numero totale di operazioni effettuate. È utile per comprendere l'affidabilità di una strategia nel tempo e in diverse situazioni di mercato oltre che per gestire il rischio in modo adeguato. È espressa in percentuale;
- III. *Max Drawdown*: è una misura di rischio che, in riferimento ad un determinato periodo temporale, indica la massima perdita cumulativa subita da un investimento. Esprime la riduzione percentuale massima dell'investimento da un massimo e un minimo;
- IV. *Profit Loss Ratio*: è il rapporto tra i potenziali profitti e perdite;
- V. *Average Win* e *Average Loss*: misurano rispettivamente il guadagno medio e la perdita media per ogni operazione. Il primo indica il guadagno medio ottenuto da trade che hanno prodotto un risultato positivo. La seconda indica la perdita media delle operazioni che hanno prodotto un risultato negativo;

- VI. Guadagno netto e lordo: è importante in quanto consente di analizzare quanto le tasse e commissioni considerate incidano sul guadagno o perdita generate dalla strategia;
- VII. Guadagno medio e perdita media: sono indicatori cruciali per valutare la regolarità delle prestazioni di una strategia, forniscono una visione generale di quanto sia consistente la strategia nel generare profitti e sostenere perdite nel tempo.

Prima di procedere con l'analisi e il commento delle strategie, viene qui esaminato graficamente il periodo nel quale sono state calcolate, che in questo caso è il mese di dicembre. A tal fine, sono stati scelti due grafici che rappresentano indicatori tecnici comunemente utilizzati nell'ambito dell'analisi dei mercati. Ci si avvale di tali strumenti per valutare molteplici aspetti come la volatilità, la direzione dei prezzi e i potenziali punti di inversione del mercato.

**Grafico 4.1:** *Relative Strength Index del mese di dicembre*



**Grafico 4.2:** Bande di boolinger del mese di dicembre



Il Grafico 4.1 è relativo all'indicatore RSI ossia l'Indice di Forza Relativa che è un oscillatore di *momentum* il cui fine è quello di misurare la velocità e il cambiamento di movimento dei prezzi degli asset. Il grafico assume valori che vanno da 0 a 100 e include delle linee orizzontali fissate ai livelli di 30 e 70. Le linee di soglia rappresentano rispettivamente i livelli di *oversold* (30) e *overbought* (70). Quindi, se una parte del grafico supera la soglia superiore, suggerisce che l'asset potrebbe essere sopravvalutato e con un potenziale successivo ribasso. Contrariamente, se la porzione di grafico interessata è al di sotto della soglia inferiore, suggerisce che l'asset è sottovalutato e, di conseguenza, si potrebbe assistere a un rimbalzo. La prudenza è essenziale quando si svolge l'analisi grafica con l'ausilio di questo indicatore, visto che movimenti di mercato rialzisti o ribassisti possono innescare avvisi prematuri di condizioni di ipercomprato o ipervenduto. Anche se un trend rialzista può far apparire l'indicatore frequentemente nell'area di ipercomprato, superare il valore di 70 sull'oscillatore non deve tradursi nell'assumere decisioni di vendere o liquidare posizioni *long*, al fine di prevenire gravi perdite finanziarie. Solitamente un primo movimento in una delle due zone d'interesse rappresenta un segnale che il mercato sta mutando, mentre un secondo segnale rafforza tale avvertimento fornendo una segnalazione più forte.

Il secondo indicatore grafico, Grafico 4.2, fa riferimento alle Bande di Boolinger le quali segnalano i livelli di volatilità. Consistono in tre linee: la banda superiore, la banda centrale e la banda inferiore. La banda centrale, in questo caso quella blu, è una media semplice e traccia la media del prezzo dell'asset nel tempo e funge da

base per la stima della volatilità. Mentre le bande superiori e inferiori sono calcolate come la deviazione standard dal valore medio. Le bande si espandono quando la volatilità è alta e si contraggono quando questa è bassa. In tal modo si percepisce come la volatilità varia rispetto al prezzo dell'asset in questione. Funzionando come indicatori dinamici di supporto e resistenza, le bande superiori e inferiori segnalano condizioni di ipercomprato e ipervenduto quando vengono oltrepassate.

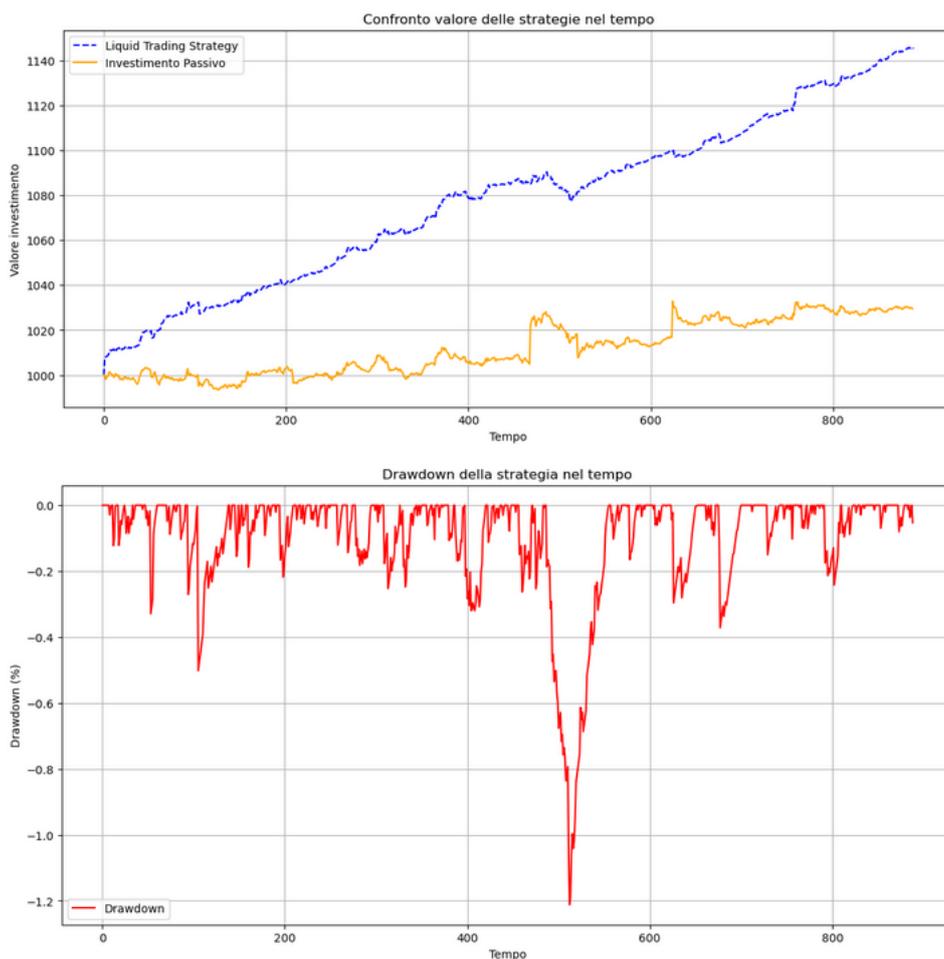
Le Bande di Bollinger sono strumenti efficaci per identificare opportunità di acquisto e vendita nel mercato. Un segnale di acquisto si presenta quando il prezzo raggiunge o esce dalla banda inferiore per poi far ritorno al suo interno. Invece, un segnale di vendita appare quando il prezzo si avvicina o esce dalla banda superiore.

#### 4.5 Liquid Strategy

Si procede con la valutazione grafica tra i rendimenti cumulati della strategia e quelli frutto di un investimento passivo nell'indice. La strategia in essere, ossia la *Liquid trading strategy*, delineata dal colore blu mostra una tendenza complessiva al rialzo, indicando che la stessa ha generato un aumento complessivo e progressivo del valore iniziale dell'investimento. La pendenza della curva è quanto meno costante anche se mostra dei periodi di fluttuazione indicando che le performance non sono sempre state ottimali ma vi sono stati anche delle perdite dovute a previsioni non corrette. La linea arancione denota, invece, l'investimento passivo e il suo incremento è moderato rispetto alla linea blu.

Il grafico inferiore è relativo al *drawdown* della strategia di trading e rappresenta la perdita massima dall'ultimo picco, mostrando quanto l'investimento è sceso durante i periodi di ribasso prima di recuperare. Si noti come siano presenti molti piccoli e brevi ribassi e un *drawdown* significativo superiore all'1% dall'intervallo di osservazioni posizionate dal 400 al 600. Nonostante i rendimenti superiori vi sono numerosi periodi di volatilità e quindi di rischio. Di conseguenza, si potrebbe optare per l'inserimento di misure di gestione del rischio come strumenti stop-loss o *hedging* per fornire una protezione ulteriore al capitale durante momenti di incertezza.

**Grafico 4.3:** Liquid strategy vs investimento passivo e Drawdown della strategia



**Tabella 4.3:** Riassuntiva delle statistiche ottenute in seguito all'applicazione della strategia Liquid

Statistica	Valore
Valore lordo finale	1145.18
Valore netto finale	1125.48
Numero trade	446
Rendimento mensile	14.51
Win rate (%)	54.71
Loss rate (%)	45.29
Average win (%)	0.072
Average loss (%)	-0.067
Profit-loss ratio	1.08
Max Drawdown (%)	-1.21
Volatilità per trade	0.10
Volatilità annua	1.51

Aspettativa media di guadagno per trade (%)	0.01
Miglior rendimento in una singola operazione (%)	0.78
Peggior rendimento in una singola operazione (%)	-0.50

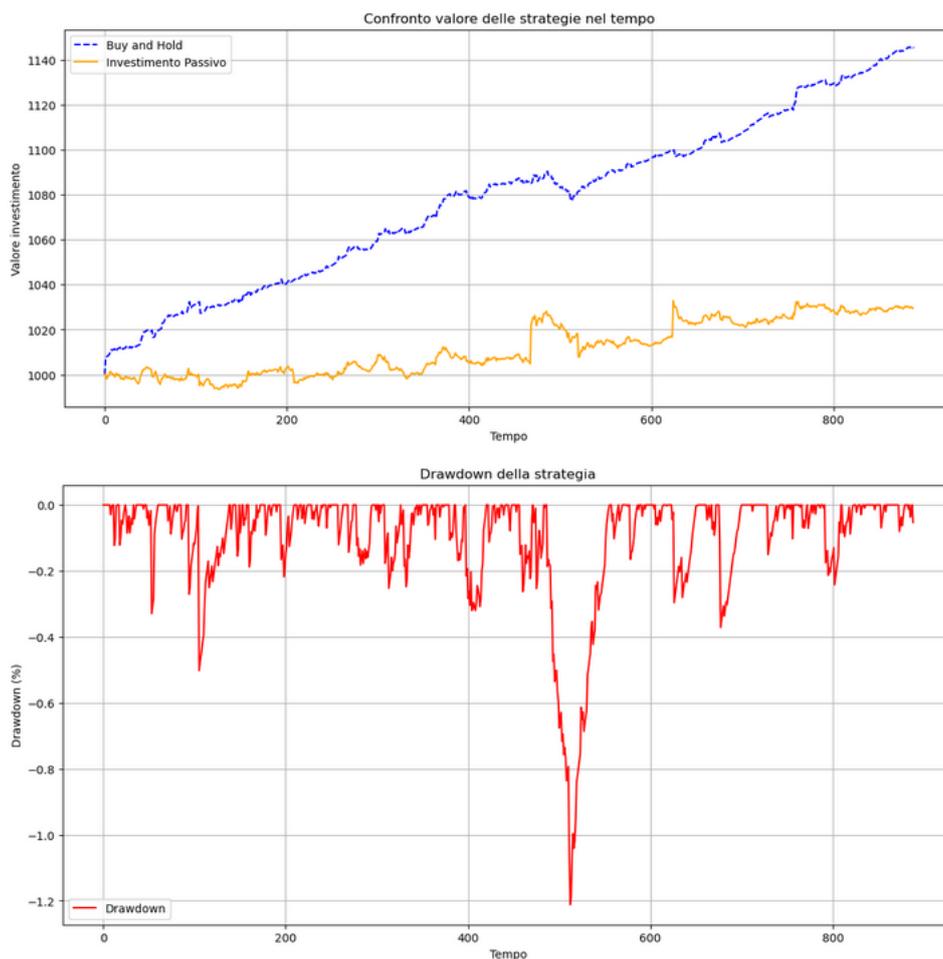
Le statistiche presentate nella Tabella 4.3 si riferiscono alle prestazioni della strategia durante la sua fase di attuazione. Si noti come il capitale iniziale di 1000 sia aumentato fino a raggiungere la quota di 1145.18. Tuttavia, considerando le tasse presentate nella Tabella 4.1, si osservi come il guadagno sia eroso a tal punto da diminuire fino ad un importo pari a 1134.24. In seguito, è stata condotta un'analisi accurata per valutare se la strategia implementata avesse conseguito un rendimento superiore rispetto ad un investimento passivo nell'indice stesso, a parità di capitale iniziale. Quindi i rendimenti cumulativi di ambo le strategie sono stati confrontati al fine di verificare se la strategia di trading risultasse vantaggiosa rispetto ad un approccio passivo di lungo termine, utilizzando le variazioni percentuali dei prezzi come criterio di valutazione.

Analizzando i dettagli della strategia si nota come il guadagno medio per le operazioni vincenti è dello 0.78 mentre la perdita media per le operazioni perdenti è del -0.50. Il rapporto tra il guadagno medio e la perdita media mostra quanto siano efficaci le operazioni vincenti rispetto a quelle perdenti. In tal senso un valore maggiore all'unità è auspicabile e denota una certa efficacia della strategia. A supporto si veda come la massima perdita subita dall'investimento, in termini percentuali, prima di una nuova crescita sia relativamente bassa e quindi la componente di gestione del rischio è relativamente buona. La variabilità dei rendimenti che la strategia può generare è mostrata dal miglior e peggior rendimento in una singola operazione che, appunto, mostrano l'ampiezza dei risultati per le operazioni prese individualmente.

## 4.6 Buy and Hold

La linea tratteggiata blu rappresenta i profitti cumulati della strategia lungo il periodo di esecuzione. Questa strategia prevede l'acquisto ogni qualvolta la previsione di rendimento fornita dal modello sia positiva e la si mantiene fino a quando la previsione non diventi negativa, in tal senso si chiude la posizione e si guadagna il *risk free rate*. I risultati sono ben visibili e mostrano come la crescita sia pronunciata e costante nel tempo non evidenziando una volatilità estrema indice di rischi relativamente bassi. Anche in questo caso la strategia ha riportato un valore finale superiore all'investimento passivo. Il grafico sottostante, relativo al drawdown, mostra un andamento simile in toto a quello della strategia liquida e, anche in questa situazione, il picco viene raggiunto a una quota pari a -1.2%. Si noti come i due Grafici 4.4 qui riportati siano uguali ai due precedenti sottolineando come le due strategie portino al medesimo risultato ma con un numero di trade inferiore.

**Grafico 4.4:** Buy & Hold vs investimento passivo e Drawdown della strategia



**Tabella 4.4:** Riassuntiva delle statistiche ottenute in seguito all'applicazione della strategia Buy & Hold

Statistica	Valore
Valore lordo finale	1145.18
Valore netto finale	1126.66
Numero trade	254
Rendimento mensile	14.51
Win rate (%)	55.12
Loss rate (%)	44.88
Average win (%)	0.09
Average loss (%)	-0.07
Profit-loss ratio	1.18
Max Drawdown (%)	-1.21
Volatilità per trade (%)	0.13

Volatilità annua (%)	2.07
Aspettativa media di guadagno per trade (%)	0.02
Miglior rendimento in una singola operazione (%)	0.80
Peggior rendimento in una singola operazione (%)	-0.50

Risalta immediatamente il valore lordo finale che è pressoché identico a quello relativo alla strategia precedente.

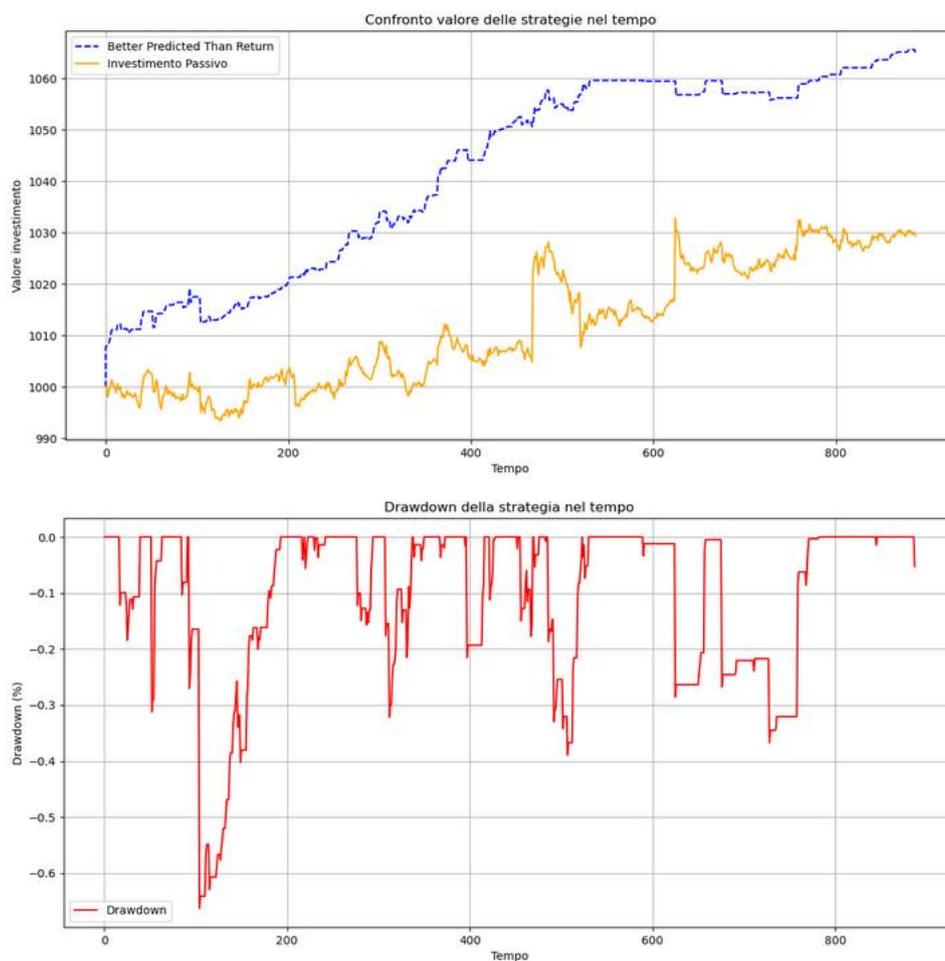
Un dato importante viene offerto dal numero di trade effettuati, si ricordi che l'obiettivo di tale strategia era quello di diminuire il numero di scambi effettuati al fine di risultare meno esposti alle commissioni da corrispondere. In definitiva, si conteggia un numero minore di trade pari a 254. Nonostante un ribasso degli scambi, si osservi come il valore finale netto sia inferiore e questo potrebbe derivare da una molteplicità di aspetti, come il fatto che la strategia in questione mantenga aperta la posizione fino a quando non si verifichi una previsione negativa e per questo potrebbe non sfruttare tutte le fluttuazioni di mercato. In secondo luogo, la volatilità per trade è più marcata e, mantenendo le posizioni aperte anche durante i periodi di flessione può incidere negativamente sui rendimenti. Infine, se i guadagni derivante da ogni singola operazione vincente fosse maggiore rispetto a quelli derivanti dall'applicazione della strategia liquida, allora le commissioni incidono con un grado di erosione superiore. Il *win rate* indica come più della metà delle operazioni svolte abbia ottenuto un guadagno. Sia l'*average win* che il *profit-loss ratio* mostrano dei miglioramenti rispetto alla strategia precedente mentre il *maxdrawdown* è rimasto invariato. In sintesi, il valore finale lordo è identico alla strategia precedente ma le differenze sono da rilevare in ogni singolo *trade* attuato.

#### 4.7 Better predicted return than average

La strategia in questione ha mostrato una performance migliore rispetto all'investimento passivo anche se questo margine viene deteriorato una volta considerati i costi relativi e tasse. Nonostante questo, il valore finale rimane pur

sempre superiore al guadagno associato alla linea arancione. La linea tratteggiata blu mostra una crescita nel tempo sebbene vi siano dei periodi di stabilità e di appiattimento dovuti ovviamente dalle regole di trading assunte. Il grafico del *drawdown* mostra una certa moderazione il che indica come la strategia sia contenuta a livello di rischi ammessi. Si mostra come la strategia sia in grado di recuperare con una certa rapidità in seguito ai picchi di *drawdown* nel tempo.

**Grafico 4.5:** Filter I vs investimento passivo e Drawdown della strategia



**Tabella 4.5:** Riassuntiva delle statistiche ottenute in seguito all'applicazione della strategia Filter I

Statistica	Valore
Valore lordo finale	1065.02
Valore netto finale	1055.10
Numero trade	134
Rendimento mensile	6.50
Win rate (%)	64.93

Loss rate (%)	35.07
Average win (%)	0.09
Average loss (%)	-0.09
Profit-loss ratio	1.77
Max Drawdown (%)	-0.66
Volatilità per trade (%)	0.13
Volatilità annua (%)	2.04
Aspettativa media di guadagno per trade (%)	0.02
Miglior rendimento in una singola operazione (%)	0.78
Peggior rendimento in una singola operazione (%)	-0.50

Con un aumento del 6.50% rispetto al valore iniziale si ha che la prima strategia di filtro ha prodotto un guadagno di 66.05 lordi. A questa cifra si devono decurtare l'ammontare da imputare alle tasse, pari a 10.95, che portano a un valore netto dell'investimento a 1055.10. Notevolmente più basso rispetto alle due strategie presentate finora. Il numero di trade effettuati si è ridotto progressivamente e conta un totale di scambi effettuati pari a 134 il che riflette l'approccio selettivo della strategia di filtro, la quale esegue le operazioni esclusivamente quando il rendimento previsto è superiore alla media dei rendimenti positivi previsti. Il *win rate* si alza notevolmente, infatti circa due terzi delle operazioni effettuate hanno conseguito dei guadagni. Nonostante questo, si deve ricordare che i rendimenti medi per operazione sono molto limitati e pari al 0.09% quindi la crescita del capitale investito non è esplosiva ma lenta e costante. In questo caso, il *max drawdown* presenta una statistica di gran lunga inferiore se confrontata con il resto delle strategie.

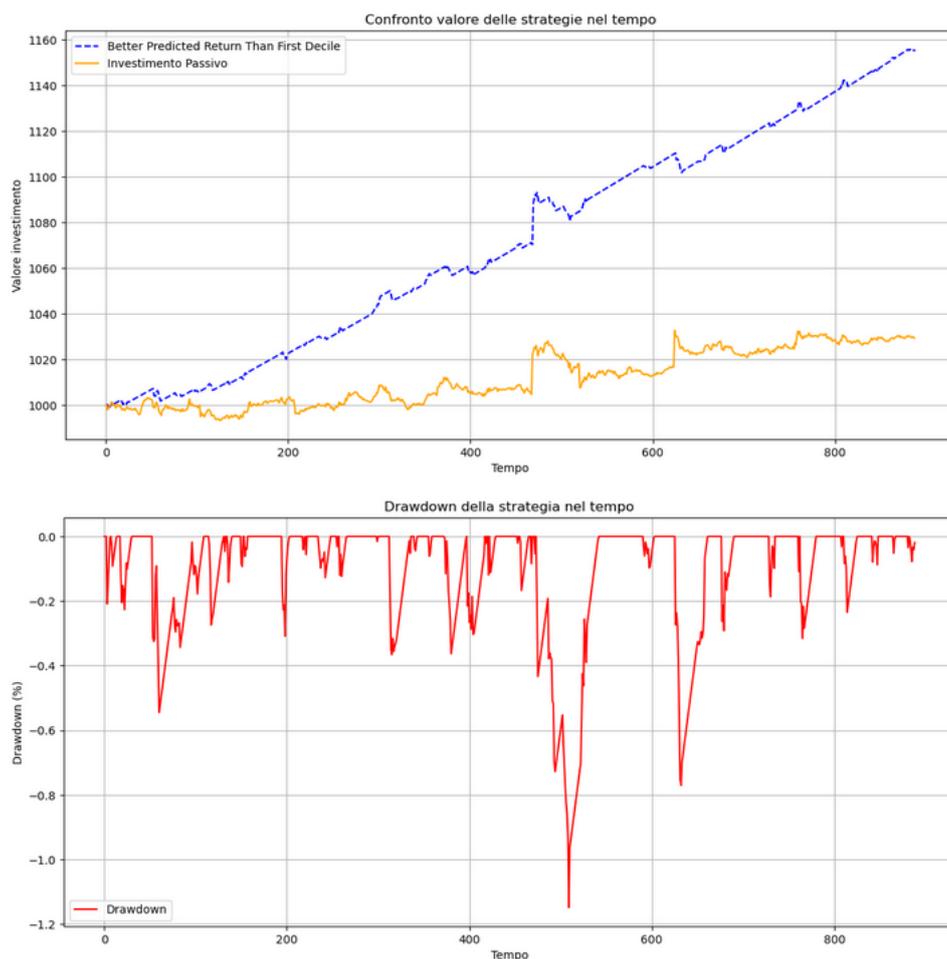
#### 4.8 Better predicten than first decile

Proseguendo con lo stesso filo conduttore, anche in questo caso la linea rappresentativa del valore dell'investimento tramite la strategia evidenzia un

incremento costante e progressivo del valore del capitale, ciò indica una performance robusta della strategia attuata. Quest'ultima si dimostra fin dall'inizio migliore rispetto all'investimento passivo nell'indice dato che è sempre superiore alla stessa e questo grazie alla capacità di sfruttare le previsioni migliori del primo decile per ottenere rendimenti più elevati.

Si registra però che la strategia in questione ha subito diversi *drawdown* lungo il periodo di investimento, a differenza della prima strategia di filtro, con il massimo picco che raggiunge la quota del -1.15%. Tuttavia, la strategia è stata in grado di recuperare le perdite e continuare a crescere.

**Grafico 4.6:** Filter II vs investimento passivo e Drawdown della strategia



**Tabella 4.6:** Riassuntiva delle statistiche ottenute in seguito all'applicazione della strategia Filter II

Statistica	Valore
Valore lordo finale	1155.57
Valore netto finale	1137.90
Numero trade	68
Rendimento mensile	15.57
Win rate (%)	52.94
Loss rate (%)	47.06
Average win (%)	0.28
Average loss (%)	-0.25
Profit-loss ratio	1.12
Max Drawdown (%)	-1.15
Volatilità per trade (%)	0.08
Volatilità annua (%)	1.23
Aspettativa media di guadagno per trade (%)	0.03
Miglior rendimento in una singola operazione (%)	1.60
Peggior rendimento in una singola operazione (%)	-0.70

Il valore finale complessivo dell'investimento senza considerare le tasse è pari a 1157.57 quindi dimostra quanto la strategia sia profittevole nel lasso temporale considerato. Il filtro selezionato è relativamente elevato e questo consente di effettuare un numero di *trade* inferiore senza intaccare la profittabilità della strategia stessa dato che si considera di attuare un'operazione solamente se si soddisfano esigenti parametri. Di questi *trade*, più della metà sono risultati vincenti e questo è evidenziato dalla percentuale di *win rate* riportata. I *trade* vincenti hanno portato dei guadagni leggermente superiori rispetto alle operazioni che hanno conseguito una perdita, valori rispettivamente di 0.28% e -0.25%. Un *profit-loss ratio* di 1.12 indica che i guadagni medi per *trade* vincenti sono superiori alle perdite medie per *trade* perdenti, suggerendo un bilancio complessivo positivo della strategia. Si sottolinea come si sia registrato un *best trade* abbastanza elevato pari a 1.60%.

## 4.9 High arbitrary cut off

Diversamente da quanto visionato in precedenza, la strategia proposta in questo ultimo caso risulta la meno efficiente di tutte e, difatti, non riesce ad *over* performare l'investimento passivo nell'indice. La linea raffigurante la strategia adottata evidenzia un andamento alquanto piatto con alcuni salti di discreto valore. Si evince, quindi, che la strategia proposta sia molto conservativa e meno volatile, consentendo una gestione del rischio più accurata e mirata, ma questo intacca ovviamente le performance, le quali delineano una crescita molto limitata del capitale investito. Il *drawdown* si mantiene allineato ad un livello di zero per la maggior parte del tempo, indicando che la strategia ha evitato perdite significative fino alla metà del periodo considerato. In seguito, si registra un primo picco intorno all'0.8% e successivamente un altro intorno al 0.28% per poi stabilizzarsi intorno ad un valore di 0.25%.

**Grafico 4.7:** Filter III vs investimento passivo e Drawdown della strategia



**Tabella 4.7:** Riassuntiva delle statistiche ottenute in seguito all'applicazione della strategia Filter III

Statistica	Valore
Valore lordo finale	1007.78
Valore netto finale	1006.29
Numero trade	8
Rendimento mensile	0.77
Win rate (%)	75.00
Loss rate (%)	25.00
Average win (%)	0.18
Average loss (%)	-0.16
Profit-loss ratio	1.13
Max Drawdown (%)	-0.26
Volatilità per trade (%)	0.17
Volatilità annua (%)	2.80
Aspettativa media di guadagno per trade (%)	0.1
Miglior rendimento in una singola operazione (%)	0.38
Peggior rendimento in una singola operazione (%)	-0.25

Come da aspettative, questa strategia riporta un valore finale estremamente esiguo se confrontato tutte le altre strategie viste precedentemente. La nota di interesse è rappresentata dal numero di *trade* eseguiti che è diminuito drasticamente ed è pari a 8. La maggior parte di questi scambi, seppur minimi, è risultata vincente, come lo denota anche la statistica dei *win rate*. Nonostante questo, la differenza tra l'*average win* e l'*average loss* è limitata e questo comporta che i guadagni derivanti dai trade vincenti sono quasi annullati dalle operazioni con esito negativo. Tuttavia, in media i guadagni dagli scambi vincenti sono stati leggermente superiori alle perdite dagli scambi perdenti. In sintesi, la strategia ha prodotto un leggero guadagno con una gestione del rischio relativamente buona, come indicato dal basso *max drawdown* e dalla volatilità. Tuttavia, il rendimento mensile è piuttosto basso a dispetto del fatto che la maggior parte dei *trade* sono stati vincenti, ma con un margine di guadagno medio relativamente basso.

#### 4.10 Sintesi dei risultati

Esaminando nel complesso i guadagni ottenuti applicando le strategie combinate con il tasso *free risk* nel mese di dicembre si può concludere affermando che queste hanno generato dei risultati profittevoli e favorevoli soprattutto grazie alla componente di guadagno derivante dal tasso privo di rischio. La seconda strategia di filtro risulta essere la più conveniente tra quelle presentate, con un numero ridotto di scambi e un guadagno lordo di 1155.17. Tuttavia, il quadro delle strategie cambia drasticamente nel momento in cui si applicano i costi di transazione che erodono quasi completamente il guadagno. Pertanto, nessuna strategia può potenzialmente battere l'investimento passivo. Il miglior rendimento registrato in un singolo *trade* è di 1.60% ed è relativo alla strategia *Better Return Than First Decile*, mentre tutte le altre mostrano dei risultati al di sotto dell'unità percentuale. Si ricordi che i costi di transazione da effettuare per l'acquisto e la vendita dell'asset è di 1.70% il che rende inefficace ogni strategia applicata per via dei numerosi *trade* eseguiti nel periodo di riferimento, conferma di ciò che si diceva poc'anzi. Un'altra componente che incide pesantemente nei guadagni è il *timing* in cui si esegue l'operazione, il modello migliore prevede correttamente il 53.94% delle volte e di conseguenza commette degli errori di previsione che possono condurre ad un esito negativo o positivo a seconda della singola operazione. Questo aspetto fa sì che, se il modello ha previsto un rendimento positivo, quindi si prosegue con un acquisto, ma in realtà il rendimento effettivo è negativo allora l'operazione risulterà fallimentare e si perde parte del capitale investito. La medesima situazione si può verificare anche con i termini invertiti.

In sintesi, il modello di previsione non si dimostra eccellente o quanto meno apprezzabile nel prevedere rendimenti dell'indice FTSE100 e, per ovvie ragioni, non riesce a generare dei rendimenti di ammontare significativo o semplicemente positivo, una volta introdotte le varie voci di spese da sostenere oltre alla tassazione. Questo risultato evidenzia le sfide inerenti alla previsione dei mercati finanziari e l'importanza di fattori esterni che possono influenzare le performance delle strategie di trading. Le inefficienze del modello si sono manifestate principalmente nell'incapacità di anticipare con precisione i movimenti del mercato, portando a

previsioni che, se implementate in strategie di trading, non risultano sufficientemente redditizie dopo aver considerato le spese operative.

In questo frangente è interessante riepilogare e confrontare i risultati ottenuti in questo elaborato e quelli, invece, proposti da Brooks et al. (2001). Tale confronto permette di evidenziare le differenze dei risultati tra i due studi, cercando di capire al contempo le ragioni sottostanti a tali discrepanze. Questo è utile anche per individuare le aree che possono essere ottimizzate e migliorate nel presente scritto.

**Tabella 4.8:** Confronto dei risultati ottenuti con l'articolo di Brooks (2001)

	<b>ECM</b>	<b>ECM-COC</b>	<b>ARIMA</b>	<b>VAR</b>		
<b>%</b>	53.83%	53.94%	48.99%	45.05%		
<b>Direzione Corretta</b>	{67.69%}	{68.750}	{64.360}	{66.80}		
<b>RMSE</b>	0.1211 {4.382e-04}	0.1211 {4.350e-04}	0.1210 {4.531e-04}	0.1211 {4.510e-04}		
<b>MAE</b>	0.064168 {0.426}	0.064175 {0.426}	0.064392 {0.438}	0.064164 {0.438}		
	<b>Inv. Passivo</b>	<b>Liquid Strategy</b>	<b>Buy&amp;Hold</b>	<b>Filter I</b>	<b>Filter II</b>	<b>Filter III</b>
<b>Rendimento</b>	3.74 {4.09}	14.51 {15.62}	14.51 {15.62}	6.50 {14.45}	15.57 {10.0}	0.77 {1.98}
<b>#Trade</b>	1 {1}	446 {583}	254 {383}	134 {135}	68 {65}	8 {8}

La comparazione dei risultati ottenuti è presentata dalla Tabella 4.8, nella quale i valori tra parentesi rappresentano gli *output* conseguiti da Brooks et al. (2001). La percentuale di direzione corretta è un indicatore chiave della capacità predittiva di ciascun modello e come si nota, questa statistica è maggiore in tutti i modelli presentati da Brooks. Nella presente tesi il modello di correzione dell'errore che

include il costo di mantenimento ha presentato il valore maggiore, con una percentuale appena inferiore al 54%. Il medesimo modello è stato identificato come il più performante anche nell'articolo di riferimento ma la percentuale è molto più elevata e di poco inferiore al 70%. Analizzando le metriche di valutazione delle previsioni si osservano delle differenze in quanto l'RMSE è nettamente maggiore nel caso riportato in questo elaborato ma al tempo stesso il MAE è inferiore rispetto a quanto calcolato dagli studiosi. In definitiva le previsioni ottenute non sono soddisfacenti né precise quanto quelle presentate dagli accademici. La medesima considerazione si estende anche ai rendimenti delle strategie, in cui quattro su cinque situazioni presentano rendimenti superiori rispetto a quelli ottenuti in questo elaborato. Solamente una strategia, la Filter II, risulta superiore rispetto al valore indicato tra parentesi. Questo può essere spiegato dal fatto che, anziché limitarsi a replicare la strategia come fatto per le altre, si sono apportate delle modifiche alle regole e vincoli di applicazione della strategia in modo tale da ottimizzarla ulteriormente. Unicamente questa strategia è stata integrata poiché si prestava meglio a tali adattamenti.

Nonostante la tesi poggia sulle stesse basi e analisi dell'articolo, si sono verificate notevoli differenze, le quali potrebbero essere spiegate da differenti fattori. La qualità dei dati di input o le specifiche tecniche di ottimizzazione sottostante ai modelli possono variare i risultati finali, o semplicemente, il periodo di studio da loro utilizzato, che potrebbe essere caratterizzato da minore volatilità, e le condizioni di mercato maggiormente stabili potrebbero aver contribuito alla variazione delle previsioni prodotte dai modelli.

Malgrado queste limitazioni e inefficienze osservate durante lo studio, si possono comunque trarre degli aspetti a supporto della validità del modello ECM-COC proposto. Innanzitutto, i costi di transazione delle banche di investimento, grazie alla loro capacità operativa e risorse, sono estremamente inferiori a quelli applicati in questi casi studio. Possono negoziare grandi volumi di azioni e strumenti derivati con maggior efficienza e a un costo minore grazie a tre fattori: accesso diretto ai mercati, miglior relazioni di trading ed economie di scala. Quindi si pone la possibilità di costruire delle strategie profittevoli in cui i rendimenti siano superiori ai costi di transazione applicati. In più il modello può essere utilizzato e sviluppato ulteriormente per migliorare i tempi di ingresso nel mercato per coloro che

desiderano operare con alte frequenze e quindi una maggiore tempestività nell'esecuzione delle regole di trading. Infatti, si presuppone che il modello possa essere migliorato ulteriormente al fine di rilevare dei segnali in modo più accurato. In aggiunta, si potrebbe pensare di adottare un approccio di investimento che pone l'obiettivo di replicare le prestazioni di un indice più ampio, impiegando solamente una sottocategoria di azioni presenti nell'indice stesso. In tal senso, si potrebbero utilizzare come proxy solo le migliori società per capitalizzazione e quindi, si cerca di catturare la tendenza generale o il movimento dell'indice sulla base di un ristretto gruppo di azioni che idealmente rappresentano le società più influenti a causa del loro maggior impatto sull'indice. Infatti, la scelta di utilizzare solo le azioni di maggior capitalizzazione è ponderata per raggiungere due scopi: semplificazione della gestione e riduzione dei costi derivanti dalle transazioni.

## Conclusioni

La presente tesi ha affrontato un'analisi approfondita della relazione *lead-lag* che intercorre tra l'indice britannico e il prezzo future iscritto sullo stesso su una scala temporale breve. A tal fine, si sono utilizzati l'indice azionario FTSE100 e i futures sull'indice FTSE100 su un intervallo temporale di 10 minuti per un periodo di studio complessivo di circa 10 mesi. Attraverso un approccio sistematico e dettagliato, sono stati esplorati vari aspetti cruciali del funzionamento dei mercati finanziari, cercando di combinare la teoria economica con la prassi per offrire una visione esaustiva e integrata delle potenzialità offerte dalle strategie di trading basate su modelli econometrici. Inizialmente, sono state esaminate le fondamenta teoriche e le principali teorie economiche riguardanti l'efficienza del mercato. Questo ha permesso di stabilire una base solida da cui partire per l'analisi empirica, confermando che, nonostante le teorie classiche, esistono opportunità per ottenere rendimenti superiori mediante strategie ben strutturate. La revisione della letteratura ha evidenziato il ruolo centrale dell'analisi delle serie temporali e dei modelli di previsione, illustrando come questi strumenti possano contribuire a sviluppare strategie di investimento più performanti e mirate. Nel primo stadio della metodologia adottata, è stato sviluppato di un modello statico tra due variabili non stazionarie al fine di verificare se una loro combinazione potesse condurre a stazionarietà. Questa fase è utile per studiare una relazione di lungo periodo tra le due serie storiche e, i risultati, fanno evincere una forte relazione tra le serie dei logaritmi, spot e futures, oggetto di studio. Il modello si poggia sulle ipotesi lineari nella relazione dinamica tra prezzi delle due variabili. Dall'equazione statica di equilibrio sono stati derivati i residui, che coerentemente con le previsioni della teoria econometrica, sono stati sottoposti a test per identificarne la loro natura. I risultati emersi hanno confermato che i residui erano stazionari.

Successivamente si è stimata l'equazione dinamica, o di breve periodo, utilizzando la componente dei residui dell'equazione statica, ritardata di un periodo. Durante questo processo è stato necessario assicurarsi che tale fattore, oltre ad essere significativo, rimanesse compreso in un *range* tra -2 e 0 al fine di prevenire l'instabilità del modello. Il passaggio successivo è stato implementare tale modello

con il *Cost of Carry* ossia il costo di mantenimento. Anche in questo caso, si sono svolte tutte le analisi preliminari necessarie partendo dalla costruzione di un modello statico per ricavare la componente residuale. Questa componente, ritardata di un periodo, è stata poi integrata nel modello a correzione dell'errore che include il COC. Questo studio ha adottato una metodologia che ha incluso l'impiego di un modello ARMA (1,0,0) e un modello VAR; anch'essi sottoposti alle analisi sui residui per verificare l'adeguatezza e la precisione delle stime prodotte e la stabilità dei modelli stessi. I risultati ottenuti dalla direzione di casualità, dalla correlazione e dai modelli evidenziano un feedback bidirezionale tra i due mercati, con una leadership predominante del mercato dei futures sui rendimenti spot. Quindi emerge come i rendimenti dei futures tendono a precedere quelli spot anche se questo si manifesta in una forma abbastanza debole rispetto a quanto riscontrato da altri studi in merito. Questo supporta la tesi che le informazioni vengano diffuse e assorbite prima nel mercato futures e successivamente l'informazione viene trasmessa nel mercato spot, a sostegno del fatto che i futures "giocano" un ruolo di scoperta dei prezzi. Questo risultato mostra che la speculazione è un determinante di rilievo dei prezzi. I contratti futures su indici dovrebbero essere in grado di fornire informazioni sulla direzione futura dell'indice azionario sottostante e questo è intuitivo per effetto di determinati vantaggi tra cui: bassi costi nel mercato futures ed effetto leva generalmente più elevato. Le evidenze di questo genere sono viste come indicative di relazioni di causa effetto tra i mercati, suggerendo che i movimenti di prezzo di un mercato possano prevedere quelli dell'altro. Tuttavia, la relazione tra rendimenti non simultanei dei due mercati non implica necessariamente una casualità. Essa dovrebbe essere interpretata come un indicatore di diversa velocità di incorporazione delle informazioni nei prezzi, Amihud-Mendelson (1988). Altrimenti potrebbe riflettere diverse preferenze degli investitori nell'utilizzo di questi strumenti.

I modelli precedentemente menzionati sono stati applicati per fare previsioni un passo in avanti sui rendimenti dell'indice spot per il mese di dicembre. L'analisi ha dimostrato che i modelli ECM, in particolare l'ECM-COC, offrono una maggiore capacità di prevedere la direzione dei rendimenti futuri con una corretta predizione della direzione dei rendimenti del 53,94% dei casi. Sebbene nessun modello si sia dimostrato nettamente superiore in termini di accuratezza statistica, l'ECM-COC ha

mostrato un vantaggio significativo nella previsione delle direzioni di mercato, anche se si necessita di rimarcare il fatto che le performance predittive per tutti i modelli sono modeste e mostrano dei limiti evidenti, con percentuali di direzione corretta che si collocano intorno ad un intervallo del 50%. In merito a questo punto ci si riserva la facoltà di indagare ulteriormente al fine di migliorare i risultati in futuro. Le strategie di trading sviluppate sono state validate considerando un ambiente di mercato che potesse quantomeno avvicinarsi alla realtà, includendo le tasse da applicare sui profitti derivanti da ogni operazione per simulare le condizioni operative effettive. I risultati empirici hanno evidenziato che alcune strategie, come la *Better Predicted Return than First Decile*, hanno ottenuto rendimenti lordi notevoli con un rendimento mensile del 15,57% rispetto al rendimento mensile del *benchmark* pari a 3,74%. La stessa ha comunque dimostrato una robustezza notevole, con un valore netto finale significativo, evidenziando il potenziale delle strategie basate su previsioni accurate. Tuttavia, i costi di transazione hanno avuto un impatto significativo sulla riduzione dei profitti netti, sottolineando l'importanza cruciale di ottimizzare le strategie di trading per minimizzare tali costi. La performance delle strategie implementate hanno rivelato che, nonostante i rendimenti positivi al netto delle tasse, nessuna strategia è riuscita a superare consistentemente l'investimento passivo nell'indice FTSE100 una volta considerati i costi di transazione fissati all'1.7% da pagare ogni volta che si effettua un *trade* completo. Questo risultato conferma la necessità di sviluppare e perfezionare ulteriormente le strategie per migliorare l'efficienza operativa e massimizzare i rendimenti netti. L'adozione di modelli di previsione accurati e strategie di trading ottimizzate rappresenta un'area di grande interesse sia per gli investitori che per gli accademici. Si desidera sottolineare come le banche di investimento e gli operatori di mercato con accesso a costi di transazione inferiori possono trarre maggiori benefici dalle strategie analizzate, e infatti potrebbero esistere potenziali circostanze in cui i risultati ottenuti possono essere utilizzati, come il *timing* ottimale per le operazioni o da un trader con costi di transazione significativamente inferiori, come un *market maker*.

In conclusione, questa tesi ha dimostrato che, nonostante le sfide associate ai costi di transazione e alla precisione nelle previsioni dei rendimenti, le strategie di trading basate su modelli di previsione mantengono un potenziale significativo per

incrementare i rendimenti degli investitori. Con ulteriori ottimizzazioni e sviluppi, queste strategie possono contribuire a una gestione del rischio più efficace e a rendimenti più stabili nel lungo periodo. La ricerca futura dovrebbe concentrarsi sull'ulteriore perfezionamento dei modelli di previsione e sulla riduzione dei costi operativi, per garantire che le strategie di trading possano essere implementate in modo più efficiente e redditizio.

## Bibliografia

- Alemaný, N., Arag3, V., & Salvador, E. (2020). Lead-lag relationship between spot and futures stock indexes: Intraday data and regime-switching models. *International Review of Economics & Finance*, 68, 269–280.
- Antoniou, A., & Holmes, P. (1995). Futures trading, information and spot price volatility: Evidence for the FTSE-100 stock index futures contract using GARCH. *Journal of Banking & Finance*, 19(1), 117–129.
- Antonopoulou, H., Theodorakopoulos, L., Halkiopoulos, C., & Mamaloukou, V. (2023). Utilizing Machine Learning to Reassess the Predictability of Bank Stocks. *Emerging Science Journal*, 7(3), 724–732.
- ATARAJAN, V. K., ABRAR UL HAQ, M., AKRAM, F., & SANKAR, J. P. (2021). Dynamic Relationship between Stock Index and Asset Prices: A Long-run Analysis. *The Journal of Asian Finance, Economics and Business*, 8(4), 601–611.
- Biondo, A. E., Pluchino, A., Rapisarda, A., & Helbing, D. (2013). Are Random Trading Strategies More Successful than Technical Ones? *PLoS ONE*, 8(7): e68344.
- Bormetti, G., Casarin, R., Corsi, F., & Livieri, G. (2020). A Stochastic Volatility Model With Realized Measures for Option Pricing. *Journal of Business & Economic Statistics*, 38(4), 856–871.
- Brooks, C., Rew, A. G., & Ritson, S. (2001). A trading strategy based on the lead–lag relationship between the spot index and futures contract for the FTSE 100. *International Journal of Forecasting*, 17(1), 31–44.
- Duan, J., Liu, H., & Zeng, J. (2013). Posterior probability model for stock return prediction based on analyst’s recommendation behavior. *Knowledge-Based Systems*, 50, 151–158.
- Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383–417.
- Garrett, I. (1992). *The pricing relationship between the FTSE 100 stock index and FTSE 100 stock index futures contract* (Doctoral dissertation, Brunel University). Centre for Empirical Research in Finance, Department of Economics.
- Gottardo, P. (s.d.). *Dinamica dei rendimenti e relazioni lead-lag fra il FIB30 e l’indice MIB30*.
- Guirrerri, S. S. (2018). *Analisi econometrica delle serie storiche con R*.

- Hull J., (2009). *Opzioni, futures e altri derivati - Decima edizione*. Pearson-Prentice Hall, Milano
- Jong, F. D., & Donders, M. W. M. (1998). Intraday Lead-Lag Relationships Between the Futures-, Options and Stock Market. *Review of Finance*, 1(3), 337–359.
- Khan, S. U., & Hijazi, S. T. (2009). Single Stock Futures Trading and Stock Price Volatility: Empirical Analysis. *The Pakistan Development Review*, 48(4), 553–563.
- Kirilenko, A., Kyle, A., SAMADI, M., & Tuzun, T. (2017). The Flash Crash: High Frequency Trading in an Electronic Market. *The Journal of Finance*, 72.
- Kyriacou, K. and Sarno, L. (1999), The temporal relationship between derivatives trading and spot market volatility in the U.K.: Empirical analysis and Monte Carlo evidence. *J. Fut. Mark.*, 19: 245-270.
- Liu, J. (2023). Research on Quantitative Trading Strategies Based on the Turtle Trading Rule. *Highlights in Business, Economics and Management*, 10, 72-80.
- Lucchetti, J. (2015). *Appunti di analisi delle serie storiche*.
- Mykland, P. A. (2010). CHAPTER 15—Option Pricing Bounds and Statistical Uncertainty: Using Econometrics to Find an Exit Strategy in Derivatives Trading. In Y. AÏT-SAHALIA & L. P. HANSEN (A c. Di), *Handbook of Financial Econometrics: Applications* (Vol. 2, pp. 135–195). Elsevier.
- Ray, K., & Panda, A. (2011). *The Impact of Derivative Trading on Spot Market Volatility: Evidence for Indian Derivative Market*, *Interdisciplinary Journal of Research in Business*. Vol. 1, 117-131.
- Ren, F., Cai, ML., Li, SP. (2023). *et al.* A Multi-market Comparison of the Intraday Lead–Lag Relations Among Stock Index-Based Spot, Futures and Options. *Comput Econ* **62**, 1–28.
- Ren, F., Cai, ML., Li, SP. *et al.* A Multi-market Comparison of the Intraday Lead–Lag Relations Among Stock Index-Based Spot, Futures and Options. *Comput Econ* **62**, 1–28 (2023)
- Stock, J. H. (1987). Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors. *Econometrica*, 55(5), 1035–1056.
- Stoll, H. R., & Whaley, R. E. (1990). Program Trading and Individual Stock Returns: Ingredients of the Triple-Witching Brew. *The Journal of Business*, 63(1), S165–S192.

Tao, R., & Brooks, C. (2019). *Introductory Econometrics for Finance*.

The Effect of Index Futures Contracts on Indices in Futures Markets. (2019).  
*International Journal of Scientific and Technological Research*.

Zemkoho, A. (2022). *A basic time series forecasting course with Python*

Zhang, Y., Zeng, Q., Ma, F., & Shi, B. (2019). Forecasting stock returns: Do less powerful predictors help? *Economic Modelling*, 78, 32–39.