



Università  
Ca' Foscari  
Venezia  
Facoltà  
di Economia

Corso di Laurea Magistrale  
in Economia e Finanza

Prova finale di Laurea

# Copertura del Crack Spread con modelli a correlazione dinamica

**Relatori**

Prof. Domenico Sartore  
Prof. Roberto Casarin

**Correlatore**

Prof. Marcella Lucchetta

**Laureando**

Enrico Piccin  
Matricola 817087

**Anno Accademico**  
**2011 / 2012**

# Ringraziamenti

Un grazie sincero a quanti mi hanno aiutato e incoraggiato, in particolare al prof. Sartore per il prezioso lavoro svolto in questi mesi, fornendomi utili indicazioni di ricerca e spunti di analisi metodologica.

L'attività di tutor di econometria da lui proposta è stata per me un'occasione unica professionale, di esperienza didattica, di approfondimento e di collaborazione; a diretto contatto con esperienze scolastiche di altri allievi.

Non riuscirei mai ad esprimere pienamente la gratitudine per l'eccezionale collaborazione del prof. Casarin su questo progetto; mettendomi a disposizione le fonti principali della ricerca e l'aiuto nella puntuale verifica e analisi delle mie proposte.

Grazie prof. Casarin per la sensibilità dimostratami e per avermi sostenuto e stimolato a ricercare nuove soluzioni; dedicandomi preziosi pomeriggi di studio e di ricerca, sottratti al suo impegno didattico.

# Indice

<b>Introduzione .....</b>	<b>1</b>
<b>CAPITOLO 1. Copertura del crack spread .....</b>	<b>3</b>
1.1. Definizione di spread.....	3
1.1.1. Crack Spread.....	3
1.2. Copertura del crack spread .....	5
1.2.1. Crack spread semplice 1:1.....	5
1.2.2. Crack spread diversificato:3:2:1 e 5:3:2.....	7
1.2.3. Esempi.....	8
1.2.3.1. Fissare il margine di raffinazione.....	8
1.2.3.2. Raffineria con copertura diversificata.....	9
1.2.3.3. Comprare crack spread .....	10
1.2.4. Fattori che influenzano il valore di crack spread.....	12
1.3. Copertura ottimale.....	13
1.3.1. Introduzione al problema di copertura .....	13
1.3.2. Teoria della copertura ottimale .....	14
1.3.3. Formalizzazione del problema.....	17
<b>CAPITOLO 2. Modelli Autoregressivi Vettoriali.....</b>	<b>23</b>
2.1. Proprietà e assunzioni base dei processi VAR .....	23
2.1.1. Processi VAR( $p$ ) stabili.....	23
2.1.2. Processi stazionari.....	28
2.1.3. Autocovarianza di un processi VAR(1).....	29
2.1.4. Autocovarianza di processi VAR( $p$ ) stabili.....	30
2.1.5. Autocorrelazione di un processo VAR( $p$ ) stabile .....	31
2.2. Vector Error Correction Models.....	32
2.2.1. Processi integrati .....	33
2.2.2. Processi cointegrati e Vector error correction models.....	37

<b>CAPITOLO 3. Modelli GARCH Multivariati.....</b>	<b>43</b>
3.1. Introduzione ai modelli GRACH Multivariati .....	43
3.2. Il modello CCC.....	46
3.3. Il modello BEKK.....	46
3.4. Il modello DCC .....	47
3.4.1. <i>Matrice di varianze e covarianze, correlazioni e log-likelihood</i> .....	47
3.4.2. <i>Stime ed errori standard</i> .....	49
<b>CAPITOLO 4. Applicazione empirica .....</b>	<b>51</b>
4.1. Dataset .....	51
4.1.1. <i>Descrizione delle serie storiche</i> .....	51
4.1.2. <i>Analisi delle proprietà statistiche dei dati</i> .....	55
4.2. Implementazione VECM.....	60
4.2.1. <i>Struttura dei ritardi test di cointegrazione</i> .....	60
4.2.2. <i>Stima VECM</i> .....	61
4.2.3. <i>Analisi dei residui</i> .....	63
4.3. Stima della volatilità.....	68
4.3.1. <i>Stime CCC</i> .....	70
4.3.2. <i>Stime DCC</i> .....	70
4.3.3. <i>Stime BEKK</i> .....	71
4.4. Copertura del crack spread .....	73
4.4.1. <i>Copertura con CCC</i> .....	73
4.4.2. <i>Copertura con DCC</i> .....	75
4.4.3. <i>Copertura con BEKK</i> .....	78
4.4.4. <i>Coperture alternative: full-hedge e un-hegde</i> .....	80
4.5. Confronto finale.....	81
<b>Conclusioni .....</b>	<b>83</b>
<b>APPENDICE .....</b>	<b>85</b>
<b>Riferimenti bibliografici.....</b>	<b>89</b>

## Indice delle tabelle

Tabella 1: Esempio di crack spread diversificato.....	7
Tabella 2: Statistiche descrittive.....	55
Tabella 3: Test ADF su Futures crude oil .....	56
Tabella 4: Test ADF su Futures Gasoline .....	57
Tabella 5: Test ADF su Futures Heating oil.....	57
Tabella 6: Cointegration test.....	61
Tabella 7: Stima relazioni di cointegrazione VECM.....	61
Tabella 8: Stima error correction VECM .....	62
Tabella 9: Test ADF sui residui VECM .....	65
Tabella 10: Test eteroschedasticità sui residui .....	67
Tabella 11: Matrice di Correlazione CCC.....	70
Tabella 12: Confronto CCC.....	75
Tabella 13: Confronto DCC .....	78
Tabella 14: Confronto BEKK.....	80
Tabella 15: Confronto risultati .....	82

# Indice delle figure

Figura 1: Torre di distillazione del crude oil .....	6
Figura 2: Andamento prezzi Futures .....	52
Figura 3: Andamento prezzi Spot.....	52
Figura 4: Spread Crude Oil.....	53
Figura 5: Spread Heating Oil.....	53
Figura 6: Spread Gasoline .....	54
Figura 7: Andamento Futures&Spot.....	54
Figura 8: Distribuzioni di frequenza.....	55
Figura 9: Granger Causality Crude Oil.....	59
Figura 10: Granger Causality Gasoline .....	59
Figura 11: Granger Causality Heating Oil.....	60
Figura 12: Residui VECM.....	64
Figura 13: Correlogramma dei residui.....	66
Figura 14: Distribuzioni di frequenza sui residui .....	68
Figura 15: Stime di volatilità .....	69
Figura 16: Matrice di correlazioni dinamiche DCC .....	71
Figura 17: Matrice di correlazioni dinamiche BEKK .....	72
Figura 18: Rapporti di copertura ottimi CCC ( $\lambda=2$ ).....	73
Figura 19: Rapporti di copertura ottimi CCC ( $\lambda=4$ ).....	74
Figura 20: Rapporti di copertura ottimi CCC ( $\lambda=6$ ).....	74
Figura 21: CEI (CCC).....	75
Figura 22: Rapporti di copertura ottimi DCC ( $\lambda=2$ ).....	76
Figura 23: Rapporti di copertura ottimi DCC ( $\lambda=4$ ).....	76
Figura 24: Rapporti di copertura ottimi DCC ( $\lambda=6$ ).....	77
Figura 25: CEI (DCC) .....	77
Figura 22: Rapporti di copertura ottimi BEKK ( $\lambda=2$ ).....	78
Figura 23: Rapporti di copertura ottimi BEKK ( $\lambda=4$ ).....	79
Figura 24: Rapporti di copertura ottimi BEKK ( $\lambda=6$ ).....	79
Figura 25: CEI (BEKK).....	80
Figura 25: CEI Full-hedge e Un-hedge .....	81

# INTRODUZIONE

L'obiettivo del lavoro è di verificare la superiorità delle strategie di copertura che utilizzano i modelli Generalized Autoregressive Conditional Heteroskedasticity (GARCH) multivariati con correlazione dinamica (DCC) per la previsione della matrice di varianze e covarianze dei prezzi spot e futures. Ci si attende che la superiorità dei modelli DCC-GARCH rispetto ai modelli a correlazione costante sia dovuta al fatto che questi ultimi non colgono la variabilità nel tempo della correlazione. Un altro obiettivo del lavoro di tesi è di utilizzare congiuntamente i modelli GARCH Multivariati con i modelli Vector Error Correction Models (VECM). In questo modo ci si attende di avere delle previsioni più accurate rispetto a quelle ottenute in precedenti lavori sulla copertura ottimale, in cui si usavano modelli Autoregressivi Vettoriali, perdendo le relazioni di equilibrio intercorrenti tra prezzi futures e spot.

L'applicazione empirica è stata così organizzata:

1. Dalle serie storiche dei prezzi giornalieri futures e spot di crude oil, gasoline e heating oil vengono eseguite le stime attraverso un modello VECM, così da poter stimare tutte le relazioni di cointegrazione tra variabili spot e futures e verificare il numero di relazioni di cointegrazione (ciò rappresenta un elemento di novità rispetto al lavoro di Haigh e Holt (2002) i quali stimano le singole relazioni tra futures e relativi prodotti spot, non cogliendo in questo modo tutte le relazioni intercorrenti tra tre tipi di sottostanti).
2. Dai residui ottenuti con il VECM, vengono effettuate le stime di volatilità attraverso svariati modelli GARCH multivariati: constant conditional correlation di Bollerslev (CCC), dynamic conditional correlation di Engle e Sheppard (DCC), BEKK di Engle e Kroner; per ogni modello si ottengono stime di matrici di varianze e covarianze dinamiche nel tempo e correlazioni costanti/dinamiche dipendentemente dal modello utilizzato.
3. Con le stime di volatilità ottenute si è calcolato la copertura ottimale usando diversi gradi di avversione al rischio; quindi è stato condotto un confronto con altri due tipi di copertura: full-hedge (compro un futures crude oil e vendo uno sul gasoline e uno sull'heating oil) e un-hedge (non entro sul mercato futures, non coprendo in questo modo il mercato spot). Il tutto per trovare con quale modello si ottenga la migliore copertura ottimale.

La tesi è stata organizzata in quattro capitoli. Nel primo viene formalizzato il problema di copertura (hedging) ottimale. Nel secondo e nel terzo vengono definiti rispettivamente i modelli VECM e GARCH multivariati. Per quest'ultimi l'attenzione è stata concentrata sul sistema DCC di Engle e Sheppard. Nel quarto ed ultimo capitolo presenteremo il data set, le stime e i risultati dell'applicazione empirica.

# CAPITOLO 1

## Copertura del crack spread

### 1.1. Definizione di Spread

Gli spread sulle materie prime sono quantità importanti sia per gli investitori che per i produttori. Ci sono quattro tipi di spread comunemente utilizzati: gli spread tra i prezzi dello stesso prodotto in due luoghi diversi (location spreads) o tempi (calendar spreads) diversi, tra i prezzi di input e output di produzione (production spreads) o tra i prezzi dei diversi tipi dello stesso prodotto (quality spreads). In questa tesi verranno trattati i problemi legati ai differenziali di prezzo nei mercati dei prodotti petroliferi. In questo ambito, un esempio di location spread è dato dal differenziale di prezzo tra il petrolio e l'olio combustibile (crack spread) e rappresenta il valore di produzione (incluso l'utile) per una società di raffinaria. Un esempio di location spread, sempre nell'ambito della raffinazione del petrolio, si ha quando una raffineria è in grado di acquistare greggio (crude oil) necessario alla lavorazione, sia negli Stati Uniti (WTI Crude Oil) che nel Regno Unito (Brent Crude Oil). Vi è allora possiede una vera opzione di scelta legata direttamente all'andamento dei prezzi del petrolio WTI e Brent.

#### 1.1.1. Crack spread

Nel settore petrolifero, i manager delle raffinerie sono interessati alla copertura della differenza tra i costi di ingresso e prezzi d'uscita. I profitti delle raffinerie sono legati direttamente allo spread tra il prezzo del greggio (crude oil) e dei prezzi dei prodotti raffinati - gasoline e distillati (diesel fuel, jet fuel e heating oil).

Tale differenza è definita "crack spread". Il nome deriva dal processo di raffinazione (cracking) che scinde il greggio nei suoi principali prodotti raffinati.

Gli operatori di borsa hanno acquistato e venduto il crack spread - noto anche come intercommodity spread - sul New York Mercantile Exchange, Inc., per più di un decennio, usando futures sull'heating oil (olio combustibile), sul gasoline (benzina), e sul crude oil (petrolio). Negli ultimi anni, l'uso del crack spread è sempre più diffuso in risposta alle fluttuazioni dei prezzi drammatici causati da condizioni climatiche estreme o crisi politiche. L'impatto del clima molto freddo negli ultimi inverni, la crisi del Golfo Persico tra il 1990 e il 1991, hanno fatto registrare prezzi e margini bassi nel 1998 e

all'inizio del 1999, e l'aumento dei prezzi nel 2000, e altri eventi nazionali hanno talvolta generato margini elevati per le raffinerie e per gli operatori del mercato, ma altre volte fortemente compresso la loro redditività.

Altri cambiamenti delle condizioni di mercato possono avere un minore, ma pur sempre significativa impatto sui prezzi. La recente polemica sulle norme ambientali in materia di formulazione della benzina e del tenore di zolfo dei combustibili distillati si è certamente fatta sentire nei mercati.

Gli impianti di raffinazione del petrolio, come la maggior parte dei produttori, sono divisi tra due mercati: le materie prime di cui hanno bisogno da acquistare e i prodotti finiti che offrono in vendita. Il prezzo del greggio e dei suoi principali prodotti raffinati sono spesso indipendentemente soggetti a variabili dell'offerta, della domanda, della produzione economica, alle normative ambientali e altri fattori. Come tale, per i raffinatori di petrolio può essere enormemente rischioso quando il prezzo del petrolio greggio aumenta mentre i prezzi dei prodotti raffinati rimangono stabili, o addirittura scendono. Tale situazione può seriamente limitare il crack spread, che rappresenta il margine di profitto di un impianto di raffinazione che compra crude oil e contemporaneamente vende i prodotti raffinati in un mercato competitivo.

Poiché raffinatori sono su entrambi i lati del mercato, la loro esposizione al rischio può essere superiore a quella sostenuta dalle aziende che vendono semplicemente crude oil o i prodotti raffinati.

Oltre a coprire i costi operativi e fissi di esercizio di raffinazione, i manager degli impianti petroliferi sperano di ottenere un rendimento positivo dall'operazione in crack spread. Inoltre, gli investitori possono usare il crack spread come copertura contro il valore del patrimonio netto di una società di raffinazione. In altro modo, gli operatori professionali possono considerare l'utilizzo di crack spread come parte del loro portafoglio energetico, con l'aggiunta del vantaggio dei suoi bassi margini (un trade sul crack spread riceve una spread creditizio per il margine messo). Insieme con altri indicatori, quali le scorte di petrolio greggio e di raffinazione, i tassi di utilizzo, uno spostamento del crack spread o del margine di raffinazione può aiutare gli investitori a farsi un'idea migliore su dove alcune aziende e il mercato del petrolio può essere diretto nel breve termine.

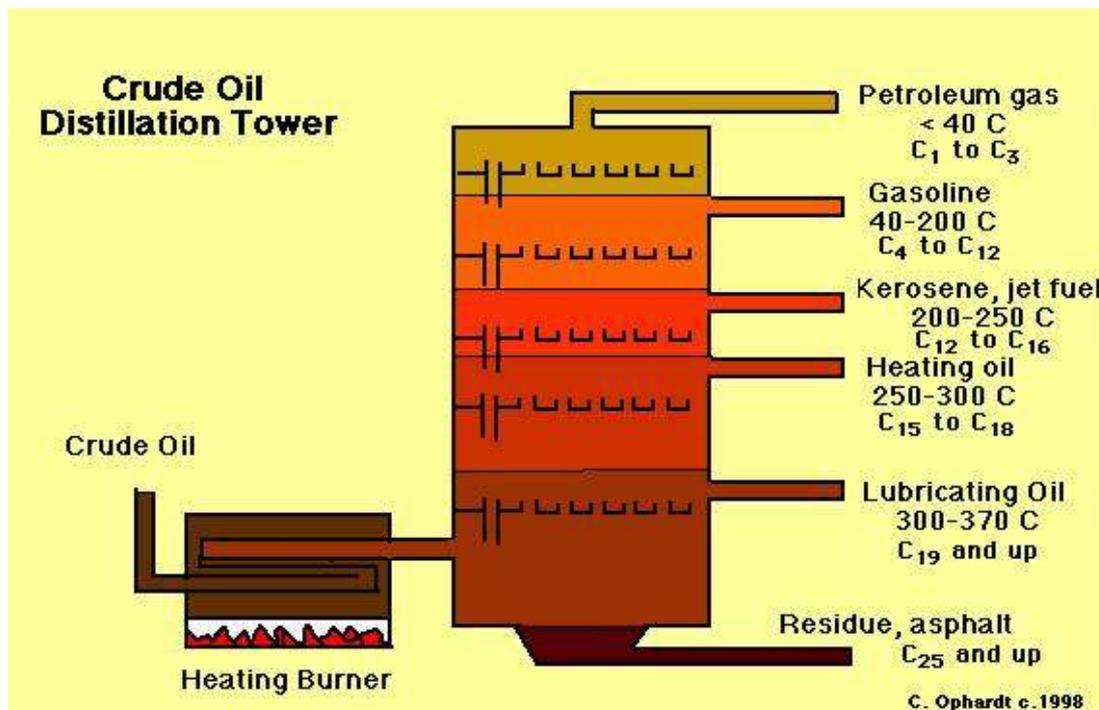
## **1.2. Copertura totale del crack spread**

Ci sono diversi modi per gestire il rischio di prezzo legato alla gestione di un raffineria. Poiché gli output di una raffineria variano in base alla configurazione dell'impianto, il suo greggio, e la sua necessità di servire le richieste di prodotti stagionali del mercato, ci sono vari tipi di crack spread per aiutare i raffinatori nella copertura di vari rapporti di greggio e prodotti raffinati. Ogni società di raffinazione deve valutare la sua particolare posizione e sviluppare una strategia di crack spread nel mercato dei futures compatibile con il suo patrimonio.

### **1.2.1. Crack spread semplice 1:1**

Il tipo più comune di crack spread è quello semplice 1:1, che rappresenta il margine di profitto della raffineria tra i prodotti raffinati (benzina o gasolio) e crude oil. Il crack spread - il margine di raffinazione teorico – si sostanzia con la vendita del futures sul prodotto raffinato (ad esempio, benzina o olio combustibile) e nel contemporaneo acquisto del futures sul crude oil, bloccando in tal modo il differenziale tra i prodotti raffinati e il greggio. Essendo il crack spread quotato in dollari al barile, e dal momento che il greggio è quotato in dollari al barile e i prodotti raffinati sono espressi in centesimi di dollaro per gallone, heating oil e gasoline devono essere convertiti in dollari per barile moltiplicando i centesimi di dollaro per gallone per 42 (ci sono 42 galloni in un barile). Se il valore di prodotto raffinato è superiore al prezzo del greggio, il margine di cracking è positivo. Viceversa, se il valore del prodotto raffinato è inferiore a quello del greggio, allora il crack spread è negativo. Questa somma viene poi divisa per il numero di barili di greggio per ridurre il valore di uno spread al barile.

**Figura 1: Torre di distillazione del crude oil**



Quando le raffinerie vogliono coprirsi dal loro rischio di crack spread, sono naturalmente lunghi sul crack spread in quanto è come se acquistassero sempre crude oil e vendessero i prodotti raffinati. Se raffinatori si aspettano che il prezzo del crude oil rimanga stabile o aumenti leggermente, mentre i prezzi dei prodotti raffinati diminuiscano bruscamente (diminuzione del crack spread), le raffinerie venderebbero il crack, cioè vendono futures su heating oil e gasoline e comprano futures sul crude oil. Il vendere o comprare lo spread, riflette ciò che viene fatto sul versante della diffusione e produzione del prodotto, tradizionalmente il premio al rischio.

A volte, però, i traders fanno l'opposto: acquistano i prodotti raffinati e vendono petrolio, trovano quindi utile una strategia buy sul crack spread. L'acquisto di un crack spread è l'opposto dell'hedging con crack spread o venderlo. Comporta vendere futures sul petrolio e acquistare, contemporaneamente dei futures sui prodotti raffinati. Quando le raffinerie sono costretti a chiudere per riparazioni o di turn-over stagionale, devono spesso entrare nel mercato del crude oil e dei prodotti raffinati per onorare gli esistenti contratti di fornitura. Gli impianti di raffinazione, non essendo in grado di produrre abbastanza prodotti per soddisfare obblighi di fornitura a lungo termine, devono acquistare prodotti a prezzi spot per la rivendita ai propri clienti a lungo termine. Inoltre,

mancando uno spazio adeguato per lo stoccaggio per le forniture in arrivo di petrolio greggio, il raffinatore deve vendere l'eccesso di crude oil sul mercato spot.

Se l'offerta della raffineria e impegni di vendita sono sostanziali e se è costretta a fare un ingresso non pianificato nel mercato a pronti, è possibile che i prezzi si muovano contro di essa. Per proteggersi da un aumento dei prezzi dei prodotti raffinati e da una diminuzione dei prezzi del greggio, la raffineria utilizza uno short hedging contro il crude oil e long hedge contro i prodotti raffinati, che è lo stesso di acquistare del crack spread.

### 1.2.2. Crack spread diversificato: 3:2:1 e 5:3:2

Ci sono strategie di copertura più complesse per crack spread che sono progettate per replicare il rendimento dei prodotti raffinati di un impianto di raffinazione. In una raffineria tipica, la produzione di benzina(gasoline) è approssimativamente il doppio di olio combustibile distillato (il taglio dei barili che contiene diesel fuel, heating oil, e jet fuel). Questo rapporto di raffinazione ha spinto molti operatori a concentrarsi su 3:2:1 spread Crack - tre contratti futures sul petrolio greggio rispetto a due futures sul gasoline e future sull'heating oil. Usando un rapporto con tre contratti sul crude oil con due sul gasoline più uno sul heating oil, il 3:2:1 margine di cracking è calcolato come segue per ottenere un 3:2:1 crack spread: (assumendo che il gasoline sia quotato 57.50\$ per gallone, heating oil 54.50\$ per gallone e il crude oil a 18.50\$ per barili).

**Tabella 1: Esempio di crack spread diversificato**

Prodotto	Conversione da galloni in barili	Peso d'investimento	Risultato
<b>Gasoline</b>	57.50 per gallone x 42	= \$ 24.15 per barili x 2	= \$ 48.30
<b>Heating oil</b>	54.50 per gallone x 42	= \$ 22.89 per barili x 1	= \$22.89
<b>Gasoline + Heating oil</b>			= \$71.19
<b>Crude oil</b>		\$ 18.50 x 3	= \$ 55.50
<b>Margine di cracking</b>	= (Gasoline + Heating oil) - Crude oil	= \$ 71.19 - \$55.50	= \$15.69
<b>Crack spread 3:2:1</b>		= \$ 15.69/3	= \$5.23 per barile

Inoltre, un impianto di raffinazione che produce crude oil con tasso di gasoline inferiore potrebbe essere interessato al trading di altre combinazioni di crack spread, come un crack spread 5:3:2 diffusione. Questo rapporto di crack spread viene eseguito dalla vendita di cinque futures dei prodotti raffinati (cioè tre futures sul gasoline e due sull'heating oil) e di acquisto di cinque futures sul crude oil bloccando così il differenziale in 5:3:2 che più da vicino replica i margini di cracking dell'impianto di raffinazione.

I traders professionali possono considerare l'utilizzo di crack spread diversificati come parte del loro portafoglio complessivo. Inoltre, la comunità degli hedge funds possono utilizzare un portafoglio diversificato da 3:2:1 o 5:3:2 crack spread come copertura contro il valore del patrimonio netto di un impianto di raffinazione.

### **1.2.3 Esempi**

#### **1.2.3.1. Fissare il margine di raffinazione**

In gennaio, il management di un impianto di raffinazione esamina la sua strategia di acquisizione del petrolio greggio e i suoi margini potenziali sul mercato della benzina per la primavera. Egli vede che i prezzi della benzina siano forti, e pianifica una strategia di spread tra crude oil e gasoline della durata di due mesi che gli permetterà di bloccare i suoi margini. Analogamente, un trader professionista può analizzare i grafici tecnici e decidere di vendere il crack spread, se il trader ha una view che gli attuali livelli di crack spread siano relativamente alti, e probabilmente a diminuirà in futuro.

Nel mese di gennaio, il differenziale tra i futures sul crude oil con scadenza aprile (\$90.00 al barile) e i futures sul gasoline RBOB con scadenza maggio (\$ 2.60 per gallone o 109,20 dollari al barile) mostra ciò che il raffinatore spera sia un favorevole crack spread 1:1 di 19,20 dollari al barile. In genere, raffinatori acquistano petrolio greggio per la lavorazione in un determinato mese, e vendono i prodotti raffinati un mese dopo.

Il raffinatore decide di vendere il crack spread vendendo futures su RBOB gasoline e acquistando dei futures sul crude oil, bloccando così il valore di crack spread di 19,20 dollari per barile. Egli esegue questo con la vendita del futures su RBOB gasoline con scadenza maggio a \$ 2.60 per gallone (o 109,20 dollari per barile), e comprando i futures sul crude oil con scadenza aprile a \$ 90.00 al barile. Due mesi dopo, a marzo, il

raffinatore acquista il petrolio greggio a 100,00 dollari al barile sul mercato spot per la raffinazione in prodotti. Allo stesso tempo, vende anche la benzina dal suo magazzino e nel mercato cash per 2,75 dollari al gallone, o 115,50 dollari al barile. Il suo valore di crack spread sul mercato spot è diminuito da gennaio, ed è ora 15,50 dollari al barile (115,50 dollari al barile il gasoline meno \$ 100.00 per barile per il petrolio greggio). Poiché il mercato dei futures riflette il mercato a pronti, i futures sul crude oil di aprile sono anche venduti a 100,00 dollari al barile a marzo - 10 dollari in più rispetto a quando il raffinatore gli ha acquistati. I futures sul gasoline RBOB sono anche venduti più alti a \$ 2,75 per gallone (115,50 dollari al barile). Per completare l'operazione di crack spread, il raffinatore riacquista il crack spread dapprima riacquistando i futures sul gasoline venduti a gennaio, e vende anche i futures sul petrolio greggio. La raffineria ha bloccato un profitto di 3,70 dollari al barile su questo crack spread.

Il raffinatore ha bloccato, con successo, un valore di crack spread pari a 19,20 \$ (Il guadagno dei futures di 3,70 dollari si aggiunge al margine di cracking nel mercato cash di \$ 15.50). Se il raffinatore non fosse stato coperto, il suo margine di cracking sarebbe stato limitato ad un guadagno di 15,50 dollari che aveva sul mercato cash. Invece, in combinazione con il guadagno sui futures, il suo margine di cracking finale netto, ottenuto con la copertura è di \$ 19.20 - il margine favorevole che originariamente veniva richiesto nel mese di gennaio.

### **1.2.3.2. Raffineria con copertura diversificata**

Un raffinatore indipendente che è esposto al rischio all'aumento dei costi del greggio e al calo dei prezzi dei prodotti raffinati corre il rischio che il suo margine di raffinazione possa essere minore del previsto. Decide così di bloccare gli attuali margini favorevoli di cracking, utilizzando una strategia di crack spread 3:2:1, che corrisponde al margine di cracking in raffineria.

Il 15 settembre, il raffinatore decide di vendere il crack spread 3:2:1, vendendo due futures sul gasoline RBOB e uno sull'heating oil, e acquistando tre futures sul crude oil, bloccando in tal modo il 3:2:1 crack spread con 27,40 dollari al barile. Egli esegue questo con la vendita di due futures sul gasoline RBOB con scadenza dicembre a \$ 3,00 al gallone (126,00 dollari per barile) e un futures dicembre sull'heating oil a \$ 3,10 per

gallone (130,20 dollari al barile), e acquistando tre crude oil futures con scadenza novembre a 100,00 dollari al barile.

Un mese dopo, il 15 ottobre, il raffinatore acquista il greggio a 110,00 dollari al barile sul mercato spot per la raffinazione in prodotti. Allo stesso tempo, vende anche gasoline della sua riserva sul mercato a pronti a \$ 3,10 per gallone (130,20 dollari al barile) e heating oil per 3,20 dollari per gallone (136,50 dollari al barile). Il valore di crack spread sul mercato a pronti è diminuito da settembre, e ora è quotato a 21,60 dollari al barile.

Poiché il mercato dei futures riflette da vicino il mercato spot, i futures sul crude oil con scadenza novembre sono anche quotati a \$ 110,00 per barile – 10 dollari in più rispetto a quando li ha acquistati. I futures sul gasoline RBOB con scadenza dicembre sono anche quotati più alti a \$ 3,10 per gallone (130,20 dollari al barile) e quelli sul heating oil con stessa scadenza vengono scambiati a \$ 3,20 per gallone (134,40 dollari al barile). Per liquidare l'operazione di crack spread 3:2:1, il raffinatore riacquista il crack spread in primo luogo riacquistando i due futures sul gasoline e quello sul heating oil venduti a gennaio, e vende anche indietro i tre futures crude oil. La raffineria ha bloccato un profitto sul crack spread di 5,80 dollari per barile di profitto su questo futures trade.

Il raffinatore ha bloccato con successo un guadagno di 27,40 \$ con il crack spread 3:2:1 (il guadagno di 5,80 dollari dei futures si aggiunge al mercato di cash con un margine di cracking di \$ 21,60). Se il raffinatore non fosse stato coperto, il margine di cracking sarebbe stato limitato ad un guadagno di 21,60 dollari che aveva sul mercato cash. Invece, in combinazione con il guadagno sui futures, il suo ultimo margine di cracking 3:2:1 con la copertura è 27,40 \$ - il margine favorevole che originariamente richiedeva nel mese di gennaio.

### **1.2.3.3. Comprare crack spread (detto anche crack spread invertito)**

Lo scopo di chi compra crack spread è l'opposto di chi si copre o vende il crack spread. Comporta vendere crude oil e comprare i prodotti raffinati. I raffinatori sono naturalmente lunghi sul crack spread in quanto continuamente comprano crude oil e vendono prodotti raffinati. A volte, però, i raffinatori fanno l'opposto: acquistano prodotti petroliferi e vendono petrolio greggio, trovano quindi una strategia utile comprare il crack spread.

Quando le raffinerie sono costretti a chiudere per riparazioni o di turn-over stagionale, devono spesso entrare nel mercato del crude oil e dei prodotti raffinati per onorare gli esistenti contratti di fornitura. Gli impianti di raffinazione, non essendo in grado di produrre abbastanza prodotti per soddisfare obblighi di fornitura a lungo termine, devono acquistare prodotti a prezzi spot per la rivendita ai propri clienti a lungo termine. Inoltre, mancando uno spazio adeguato per lo stoccaggio per le forniture in arrivo di petrolio greggio, il raffinatore deve vendere l'eccesso di crude oil sul mercato spot.

Se l'offerta della raffineria e impegni di vendita sono sostanziali e se è costretta a fare un ingresso non pianificato nel mercato a pronti, è possibile che i prezzi si muovano contro di essa. Per proteggersi da un aumento dei prezzi dei prodotti raffinati e da una diminuzione dei prezzi del greggio, la raffineria utilizza uno short hedging contro il crude oil e long hedge contro i prodotti raffinati, che è lo stesso di acquistare del crack spread.

Il "crack spread invertito" è anche un utile strategia per i trader professionali come movimento direzionale se hanno una view che gli attuali livelli di crack spread siano relativamente bassi, e probabilmente aumenteranno in futuro.

In questo esempio, il raffinatore sta programmando piani di manutenzione, e decide di comprare il semplice crack spread 1:1 crepa in gennaio con l'acquisto del futures sul gasoline RBOB, e l'acquisto dei futures sul crude oil, bloccando in tal modo il valore di crack spread a 15,00 dollari al barile. Egli esegue questo acquistando futures sul gasoline RBOB con scadenza Maggio a \$ 2.50 per gallone (o 105,00 dollari al barile), e vendendo futures sul crude oil scadenza aprile a 90,00 dollari al barile.

Due mesi dopo, a marzo, quando inizia la manutenzione dell'impianto di raffinazione, vende il petrolio greggio ad un prezzo inferiore di \$ 80.00 per barile sul mercato a pronti, poiché la raffineria chiude. Alla stesso tempo, egli acquista anche gasoline sul mercato a pronti a \$ 2.60 per gallone, o 109,20 dollari al barile. Il valore di crack spread nel mercato cash è aumentato da gennaio, e ora è quotato a 29,20 dollari al barile (109,20 dollari al barile gasoline meno 80,00 dollari al barile per il crude oil). Poiché il mercato dei futures si riflette sul mercato a pronti, i futures sul crude oil scadenza aprile sono anche venduti a 80,00 dollari al barile a marzo - \$ 10,00 meno di quando li ha acquistati a gennaio. Futures sul gasoline RBOB sono quotati più alti di \$ 2.60 per

gallone (109,20 dollari al barile). Per completare la transazione sul crack spread, il raffinatore liquida il crack spread in primo luogo con la vendita dei futures sul gasoline acquistati a Gennaio, ed egli riacquista i futures sul crude oil, ad un livello attuale di 29,20 dollari al barile. La raffineria blocca un profitto di \$ 14,20 per barile con questa operazione di crack spread sui futures (29,20 dollari al barile meno 15,00 dollari per il crack spread di gennaio).

Il raffinatore è riuscito a coprirsi perfettamente dall'aumento del crack spread (il guadagno di \$ 14,20 i futures si aggiunge al margine di cracking del mercato cash di \$ 15.50). Se il raffinatore non fosse stato coperto, il suo margine sarebbe stato limitato ad un guadagno di \$ 15,50 che ha avuto sul mercato a pronti. Invece, combinato con il guadagno futures, il suo margine finale netto di cracking con la copertura è \$ 29,20.

#### 1.2.4. Fattori che influenzano il valore di Crack Spread

Di seguito in tabella 2, verranno descritti brevemente i fattori che possono andare ad influenzare il crack spread facendo contrarre o aumentare: sono fattori di stagionalità, di tassazione, problemi geo-politici.

**Tabella 2: Fattori del Crack Spread**

Causa	Effetti Tipici	Effetto sul Crack Spread
<b>1. Problemi geo-politici: politiche, geografia, demografia, economie e politiche estere</b>	Domanda di crude oil	Inizialmente debolezza del crack spread: prezzi del crude oil alti rispetto ai prodotti raffinati. In seguito crack spread forte: a causa della diminuzione delle scorte di crude oil, diminuzione della produzione di prodotti raffinati
<b>2. Stagionalità invernale</b>	Aumento della domanda di distillati	Crack spread forte
<b>3. Rallentamento crescita economica</b>	Diminuzione della domanda di prodotti raffinati	Crack spread debole
<b>4. Forte domanda sostenuta di prodotti</b>	Elevato utilizzo delle raffinerie	Crack spread forte
<b>5. Rafforzamento della regolamentazione ambientale su specifici prodotti</b>	Riduzione della domanda di prodotto	Crack spread forte
<b>6. Scadenza del mese di trading</b>	Si opera sul mercato cash-lunghi o corti sui prodotti	I valori del crack possono variare a causa delle chiusure di posizioni
<b>7. Aumento della tassazione</b>	Aumento delle vendite di fronte a scadenze di tassazione	Crack debole vicino alla scadenza delle tasse mentre forte dopo le scadenze

<b>8. Stagionalità estiva</b>	Aumento della domanda di gasoline	Crack spread forte
<b>9. Manutenzione della raffineria</b>	Diminuzione nella produzione di prodotti	Crack spread forte
<b>10. Debolezza della valuta</b>	Debolezza del crude oil	Crack spread debole

### 1.3. Copertura ottimale

#### 1.3.1. Introduzione al problema di copertura

La non costanza della varianza delle variazioni dei prezzi per varie commodities ha ricevuto grande attenzione, soprattutto perché il problema è stato reindirizzato 35 anni fa da Samuelson (1965). Da allora, numerose pubblicazioni si sono concentrate su questioni che circondano la volatilità dei movimenti dei prezzi di varie commodities e strumenti finanziari. Per i prezzi dell'energia, in particolare, la ricerca empirica si è concentrata sulla questioni inerenti il rapporto tra il tempo-variazione (Volatilità) dei prezzi dell'energia, e il volume / maturity sui contratti di energy futures (ad esempio Serletis, 1991; Herbert, 1995). Diversi studi hanno esaminato anche l'impatto dei derivati energetici sulla volatilità dei prezzi spot, indagando se il trading sui contratti futures su energia aumenti o diminuisca il prezzo cash del sottostante (ad esempio Nainar, 1993; Fleming e Ostdiek, 1999). Altre ricerche hanno presentato evidenze empiriche sull'efficienza del mercato e la ricerca di premi del rapporto tempo-variazione per il rischio mediante variazioni di classi di modelli quali ARCH-M e GARCH-M (ad esempio Deaves e Krinsky, 1992; Moosa e Al-Loughani, 1994).

Questa tesi va a verificare l'efficacia dell'utilizzo dei futures su crude oil, gasoline e heating oil per aiutare a ridurre l'incertezza dei prezzi per gli energy traders. In particolare, ci avvaleremo di GARCH multivariati (MGARCH) modello che ci permette anche di inserire l'effetto del tempo sulla maturity, spesso già incorporato nel mercato futures. Le ricerche precedenti non hanno considerato la possibilità di copertura del crude oil, gasoline e heating oil contemporaneamente con una impostazione di varianza e correlazione non costanti nel tempo in modo da confrontare questo approccio con alternative più semplici.

I nostri risultati suggeriscono miglioramenti sostanziali rispetto a tutte le altre procedure di copertura purchè questi collegamenti informativi tra i mercati dell'energia siano direttamente computate utilizzando un modello MGARCH che incorpora l'effetto di

maturità.. I nostri risultati suggeriscono che, anche se l'aggiornamento raccomandato da modelli MGARCH risulta più costosa rispetto ad altre tecniche causa i frequenti ribilanciamenti del portafoglio, i guadagni potenziali derivanti dal frequente aggiornamento del portafoglio coperto sono superiori ai costi delle commissioni. L'analisi fornisce dunque informazioni preziose su come ridurre l'incertezza dei prezzi per l'energy trading e illustra l'importanza di come la non costanza della volatilità e soprattutto della correlazione influisca nella copertura dei prezzi sull'energy trading. Infine, i risultati sottolineano anche l'importanza spot e futures tra crude oil, gasoline e heating oil. Il resto del lavoro è organizzato come segue. Nella prossima sezione una breve rassegna sulla teoria della copertura ottimale e, pure, in dettaglio il case study dell'energy trader di un impianto di raffinazione che compra crude oil e vende i prodotti raffinati.

### **1.3.2. Teoria della copertura ottimale**

Un concetto di base utilizzati nella letteratura di copertura più moderna è la nozione che i traders selezionano in modo ottimale combinazioni di posizioni cash e futures per ridurre al minimo il rischio di portafoglio. Queste combinazioni, generalmente espressa come proporzione di contratti futures per coprire la posizione spot/cash, sono comunemente denominate rapporti di copertura (hedging). Escludendo la strategia di copertura completa (full-hedge) in cui un trader assume una esposizione sul mercato futures di quantità uguale, ma di segno opposto a quella sul mercato cash, tutte le strategie di copertura richiede l'uso di dati finanziari storici. Molti studi hanno calcolato la copertura ottimale utilizzando tecniche di OLS (semplicemente regredendo le variazioni dei prezzi cash sulle variazioni dei prezzi futures). Il risultante coefficiente di pendenza veniva quindi utilizzato come rapporto di copertura ottimale (Ederington, 1979; Kahl, 1986). Come sottolineato da Myers e Thompson (1989), tuttavia, questo metodo non riesce a spiegare correttamente per tutte le informazioni disponibili che condizionano gli hedgers al momento della loro decisione.

La metodologia tradizionale, assume anche la matrice di covarianza dei prezzi cash e futures e, quindi, i rapporti di copertura, siano costanti nel tempo. Fama (1965) ha osservato che varianze e covarianze dei rendimenti non sono, tuttavia, costanti rispetto al tempo. Di conseguenza, un ampio corpus di ricerca si è concentrata sull'utilizzo di

modelli ARCH / GARCH (Engle, 1982; Bollerslev, 1986) per modellare la dinamicità della varianza rispetto al tempo e quindi, per stimare dinamici rapporti di copertura. Ad esempio, Cecchetti et al. (1988) ha applicato un modello bivariato ARCH per i prezzi dei futures finanziari, mentre Myers (1991) e Baillie e Myers (1991) hanno applicato i modelli GARCH bivariati ai prezzi delle materie prime. Tong (1996), confrontano GARCH-generated con strategie di copertura OLS per le azioni e i tassi forward. Infine, Kroner e Sultan (1993) utilizzano un Garch bivariato per un modello di copertura sui tassi di cambio.

Esempi di modelli progettati per catturare effetti sul portafoglio del variare del tempo in cui si utilizzano diversi asset (e quindi la cattura covariabilità tra i prezzi) sono limitati. Un notevole eccezione è la ricerca di Kroner e Claessens (1991), che ha stimato l'effetto del variare del tempo su portafogli di debito. Un altro contributo recente è quello di Gagnon, Lypney e McCurdy (GLM) (1998). GLM hanno studiato gli effetti dinamici di un problema di copertura su un portafoglio multi-valuta utilizzando una GARCH (1,1) trivariato con sistema BEKK. Essi confrontano le performance di strategie statiche e dinamiche di hedging per minimizzazione del rischio e mettono in luce che i modelli a tempo variabile conducono a risultati migliori in termine di riduzione del rischio. Queste due applicazioni, come la maggior parte delle applicazioni dei modelli ARCH / GARCH in ottica multivariata, si sono concentrati su portafogli finanziari composti da futures.

Concettualmente, la copertura attraverso contratti futures è una procedura utilizzata per limitare o ridurre il rischio di variazioni sfavorevoli dei prezzi poichè prezzi dei futures e prezzi spot lo stesso sottostante tendono a muoversi allo stesso modo. Pertanto, variazioni del valore di una posizione spot sono compensate da variazioni del valore di una posizione opposta futures. Inoltre, i contratti futures sono preferiti come strumento di copertura grazie alla loro liquidità, velocità e grazie anche a minori costi di transazione.

Tra le industrie e le imprese che hanno maggiori probabilità di utilizzare una strategia di copertura sono quelle appartenenti al settore oil&gas. Le imprese si coprono solo se si aspettano che un evento sfavorevole sorgerà. Knill (2006) ha suggerito che se una compagnia petrolifera o del gas utilizza futures a copertura dei rischi, esse saranno coperte solo dal rischio di ribasso. La copertura da parte dei produttori di greggio

prevede normalmente la vendita dei futures sulle materie prime perché i produttori o raffinatori utilizzano i contratti a termine per bloccare i futuri prezzi di vendita o per fissare un prezzo floor. Così, essi tendono ad assumere posizioni corte in futures. Allo stesso tempo, gli energy trader, investitori o utilizzatori di olio combustibile focalizzandosi nel bloccare un prezzo d'acquisto o in prezzo cap tendono a prendere posizioni lunghe in futures. Daniel (2001) mostra che le strategie di copertura possono ridurre in modo sostanziale volatilità dei prezzi del petrolio senza ridurre significativamente il rendimento, e con l'ulteriore vantaggio di una maggiore prevedibilità e certezza.

Una delle strategie più diffusi di copertura si basa sulla minimizzazione della varianza del portafoglio, il cosiddetto rapporto di copertura a minima varianza. Con il criterio a varianza minima, la gestione del rischio richiede la determinazione del rapporto di copertura ottimale OHR (la quantità ottimale di futures acquistati o venduti, espressa in percentuale della posizione sul mercato cash). Al fine di stimare tale rapporto, le prime ricerche hanno semplicemente usato la pendenza della retta di regressione del classico modello lineare lineare dei prezzi degli spot sui prezzi futures, assumendo un rapporto di copertura invariante rispetto al tempo (Ederington (1979), Figlewski (1985) e Myers e Thomson (1989)). Tuttavia, è ormai ampiamente riconosciuto che la volatilità sui rendimenti degli asset, covarianze e correlazioni siano variabili nel tempo con le dinamiche persistenti, e si basano su tecniche come la volatilità condizionale (CV) e la volatilità stocastica (SV) modelli. Baillie e Myers (1991) sostengono che, se la distribuzione congiunta dei prezzi spot e dei prezzi futures si modifica nel tempo, stimare un rapporto di copertura costante potrebbe non essere appropriato.

Inoltre nel 2002 Haigh e Holt stimano dei rapporti di copertura ottimi variabili nel tempo con GARCH Multivariati con procedura BEKK, dapprima stimando regressione AR(7) tra singoli futures e relativi prodotti spot.

Il lavoro di Tansuchat, Chang e McAleer del 2010, stima con procedura DCC MV-GARCH solamente la copertura ottimale tra futures e spot sul crude oil WTI/Brent, non stimando i modelli con diversi fattori di rischio presenti nei mercati del gasoline e heating oil, così come Dunis, Laws e Evans (2005) modellizzano la copertura solamente sul mercato del gasoline.

Il nostro lavoro va oltre tutto questo: la copertura viene fatta su tutti e tre i mercati sia spot che futures (crude oil, gasoline, heating oil), le relazioni tra le variabili vengono stimate con la metodologia VECM (così da poter cogliere tutte le relazioni di cointegrazione tra le variabili, cosa non verificata in Haigh e Holt (2002)), inoltre le stime di volatilità vengono effettuate con MV-GARCH con metodologia DCC, così da poterli, in seguito, confrontare con modelli a correlazione costante (CCC, BEKK); quello che si otterrà sarà un rapporto di copertura migliore con i modelli a correlazione dinamica rispetto a quelli con correlazione costante.

Le procedure di stima sono chiaramente più complessa quando lo scenario di copertura tiene conto di molteplici rischi. Inoltre, estendendo la configurazione multipla di rischio per modelli dinamici complicano le stime. Di conseguenza, le stime di rapporti di copertura varianti rispetto al tempo con correlazioni dinamiche tra più mercati (crude oil, heating oil, gasoline) per un energy trader, in uno stato con più fattori di rischio non sono ancora state intraprese.

### **1.3.3. Formalizzazione del problema**

Il case study è quello di un tipico energy trader che deve coprire un impianto di raffinazione dai costi. Ci concentreremo su gasoline e heating oil, come componenti significativi di un barile di crude oil. Con la tecnologia di raffinazione attuale, l'output del gasoline è circa il doppio di quello del heating oil. Per questo motivo molti raffinatori si concentrano sul crack spread 3:2:1, vale a dire, tre contratti di crude oil rispetto ai due contratti di gasoline e un contratto di heating oil.

Nella successiva analisi considereremo un problema di due periodo per un raffinatore che ha impegni in corso per l'acquisto di crude oil e la vendita dei prodotti raffinati.

Il nostro rappresentante raffinatore vuole coprire il suo impegno per l'acquisto di 150.000 barili di crude oil e delle sue risultanti vendite gasoline e heating oil. La configurazione di input e output della raffineria è tale che il trader sceglie un crack spread 3:2:1. Cioè, ogni giorno il raffinatore utilizzando un strategia full-hedge sarebbe andato lungo su 150 contratti di crude oil e corto su 100 di gasoline e 50 di heating oil. Si assume che nel primo periodo (t-1) tutte le decisioni in materia di trading su futures siano fatti, (un giorno prima dell' acquisto e della vendita che ne derivano), nel secondo

periodo (t) vengono risolte tutte le incertezze, tutte le posizioni in futures vengono liquidati, e si verificano tutte le transazioni spot.

Il trader ha almeno tre contratti futures a sua disposizione per la copertura dei prezzi spot associati al crude oil, CC, e le vendite spot di gasoline, GC, e heating oil, HC. In primo luogo, il future sul crude oil verrà indicato con CF, quello sul gasoline con GF e quello sull'heating oil con HF; tutti contratti scambiati sul NYMEX.

Il profitto realizzato per barile – sulla base di un crack spread hedging 3:2:1 – al tempo t, è dato da:

$$\pi_t = -CC_t + \left(\frac{1}{3}\right)HC_t + \left(\frac{2}{3}\right)GC_t + b_1(CF_{t-1} - CF_t) + \left(\frac{1}{3}\right)b_2(HF_{t-1} - HF_t) + \left(\frac{2}{3}\right)b_3(GF_{t-1} - GF_t) - y_t \quad (1)$$

Dove  $\pi_t$  è il profitto totale (al barile),  $b_1$ ,  $b_2$  e  $b_3$  sono le posizioni in futures su crude oil, heating oil e gasoline (positivo se short, negativo se long), e  $y_t$  sono i costi di transazione al barile. Prezzi con un (t-1) al pedice sono costanti note al tempo t, mentre i prezzi con t al pedice sono variabili stocastiche. Andando lunghi sui futures al tempo t-1 al prezzo  $CF_{t-1}$  e depositando il margine richiesto in un conto di deposito si può coprire la parte d'incertezza associata con all'acquisto del crude oil sullo spot. La proporzione coperta è rappresentata da  $b_1$ . Al tempo t l'energy trader potrebbe liquidare i contratti futures al prezzo  $CF_t$ , vedendosi ritornare il margine sul futures, e acquistare la posizione sul mercato spot al prezzo  $CC_t$ . Pertanto, se  $CF_t < CF_{t-1}$  il trader avrà perso sulla transazione con i futures. In alternativa, se  $CF_t > CF_{t-1}$  o  $CF_t = CF_{t-1}$ , il trader avrà guadagnato o avrà trovato il punto di pareggio. Per l'heating oil, il trader andrà short sui futures sull'heating oil al tempo t-1 e al contrario (andrà long) sui contratti futures al tempo t. In questo caso, se  $HF_t < HF_{t-1}$  il trader avrà guadagnato sulla transazione con i futures. In alternativa, se  $HF_t > HF_{t-1}$  o  $HF_t = HF_{t-1}$ , il trader avrà perso o avrà trovato il punto di pareggio. Nel caso del gasoline, il trader andrà short sui futures sul gasoline al tempo t-1 e al contrario (andrà long) sui contratti futures al tempo t.. Le condizioni di guadagno, perdita o pareggio sono identiche a quelle appena elencate in precedenza con l'heating oil.

Per determinare il rapporto di copertura ottimale (cioè il numero ottimale di contratti da acquistare o vendere in ogni mercato), viene impiegato un approccio media-varianza. Ovvero, si suppone che l'energy trader di energia abbia come obiettivo la massimizzazione del profitto atteso aggiustato per il rischio nel periodo  $t$ , in cui il rischio è misurato come la variabilità dei profitti. In tale modo si possono aggiungere alla funzione obiettivo poche restrizione (che con altri approcci quali minimizzare la varianza non si possono fare) e togliere già dalla funzione obiettivo i costi di transazione (Kroner e Sultan, 1993; GLM, (1998)).

Quindi, supponendo che il trader voglia massimizzare una funzione di media-varianza detta Certainty Equivalent Income (CEI), l'obiettivo è quello di determinare la posizione ottimale sui futures,  $b_j$ , dove  $j$  = crude oil (1), heating oil (2), gasoline (3), per ciascun contratto, così da massimizzare la funzione CEI nel secondo periodo. Il problema del nostro trader rappresentativo è il seguente:

$$\max_{b_1, b_2, b_3} CEI_t = \left[ E(\pi_t | \Omega_{t-1}) - \frac{\lambda}{2} \{V(\pi_t | \Omega_{t-1}) - y_t\} \right] \quad (2)$$

Nell'equazione (2),  $E(\cdot | \Omega_{t-1})$  indica l'operatore atteso (condizionale);  $\pi_t$  è definito in (1);  $\Omega_{t-1}$  è il set di informazioni disponibile al tempo  $t-1$ ;  $\lambda$  è il coefficiente (grado) di avversione rischio;  $V(\cdot | \Omega_{t-1})$  rappresenta l'operatore varianza condizionale e  $y_t$  rappresentano i costi di transazione (commissioni per il ribilanciamento del portafoglio) sostenuti per unità (per barile).

Il profitto atteso al tempo  $t$  condizionale all'informazione disponibile al tempo  $t-1$ , può essere scritto come:

$$E(\pi_t | \Omega_{t-1}) = E(w' S_t + b' \delta F_t | \Omega_{t-1}) = w' E(S_t | \Omega_{t-1}) + b' \delta E(F_t | \Omega_{t-1})$$

$$\text{dove } S_t' = (CC_t, HC_t, GC_t)$$

$$F_t' = (-CF_t + CF_{t-1}, -HF_t + HF_{t-1}, -GF_t + GF_{t-1})$$

$$w' = \left( -1, +\frac{1}{3}, +\frac{2}{3} \right) \text{ è il vettore dei pesi di crack spread 3:2:1}$$

$$\delta = \text{diag}\{w\} \stackrel{\text{def}}{=} \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{2}{3} \end{pmatrix}$$

La varianza condizionale del profitto può essere scritta come

$$\begin{aligned} V(\pi_t | \Omega_{t-1}) &= V(w'S_t + b'\delta F_t | \Omega_{t-1}) \\ &= w'V(S_t | \Omega_{t-1}) + b'\delta V(F_t | \Omega_{t-1})\delta b + 2w' \text{Cov}(S_t, F_t | \Omega_{t-1})\delta b \end{aligned}$$

Dove  $\text{Cov}(\cdot, \cdot | \Omega_{t-1})$  è l'operatore covarianza condizionale.

Nelle precedenti espressioni in  $E(\pi_t | \Omega_{t-1})$  e  $V(\pi_t | \Omega_{t-1})$ , i valori di  $E(S_t | \Omega_{t-1})$ ,  $V(S_t | \Omega_{t-1})$ ,  $E(F_t | \Omega_{t-1})$  e  $V(F_t | \Omega_{t-1})$  sono ottenuti attraverso i modelli VECM e MV-GARCH.

Il problema del nostro impianto di raffinazione viene scritto come un problema di ottimizzazione alla Markowitz, in forma matriciale:

$$\begin{aligned} \max_b \quad & \left[ E(\pi_t | \Omega_{t-1}) - \frac{\lambda}{2} V(\pi_t | \Omega_{t-1}) \right] \\ \text{s. v.} \quad & b'\tau = 1 \end{aligned} \tag{5}$$

dove  $\tau' = (1,1,1)$

La soluzione ottima la si trova, risolvendo il problema di ottimizzazione vincolata.

Ricorrendo al metodo dei moltiplicatori di Lagrange ( $\theta$ ):

$$L = \left( E(\pi_t | \Omega_{t-1}) - \frac{\lambda}{2} V(\pi_t | \Omega_{t-1}) \right) - \theta(b'\tau - 1)$$

$\Leftrightarrow$

$$\begin{cases} \nabla_b L = 0 \\ \nabla_\theta L = 0 \end{cases}$$

$\Leftrightarrow$

$$\begin{cases} \nabla_b L = -\frac{\lambda}{2}[2\delta V(F_t|\Omega_{t-1})\delta'b + 2\delta Cov(S_t, F_t|\Omega_{t-1})'w] + \delta E(F_t|\Omega_{t-1}) - \theta\tau = 0 \\ \nabla_\theta L = -(b'\tau - 1) = 0 \end{cases}$$

$\Leftrightarrow$

$$\begin{cases} b = \left[-\frac{\lambda}{2}(2\delta V(F_t|\Omega_{t-1})\delta')\right]^{-1} [-2\delta Cov(S_t, F_t|\Omega_{t-1})'w - \delta E(F_t|\Omega_{t-1}) + \theta\tau] \\ b'\tau = 1 \end{cases}$$

$\Leftrightarrow$

$$\begin{cases} b = \lambda[\delta V(F_t|\Omega_{t-1})\delta']^{-1}[2\delta Cov(S_t, F_t|\Omega_{t-1})'w + \delta E(F_t|\Omega_{t-1}) - \theta\tau] \\ b'\tau = 1 \end{cases}$$

$\Leftrightarrow$

$$\begin{cases} b = \lambda[\delta V(F_t|\Omega_{t-1})\delta']^{-1}[2\delta Cov(S_t, F_t|\Omega_{t-1})'w + \delta E(F_t|\Omega_{t-1}) - \theta\tau] \\ \lambda[2\delta Cov(S_t, F_t|\Omega_{t-1})'w + \delta E(F_t|\Omega_{t-1}) - \theta\tau]'[\delta V(F_t|\Omega_{t-1})\delta']^{-1}\tau = 1 \end{cases}$$

$\Leftrightarrow$

$$\begin{cases} b = \lambda[\delta V(F_t|\Omega_{t-1})\delta']^{-1}[2\delta Cov(S_t, F_t|\Omega_{t-1})'w + \delta E(F_t|\Omega_{t-1}) - \theta\tau] \\ \lambda[2\delta Cov(S_t, F_t|\Omega_{t-1})'w + \delta E(F_t|\Omega_{t-1})]'[\delta V(F_t|\Omega_{t-1})\delta']^{-1}\tau - \lambda\theta\tau'[\delta V(F_t|\Omega_{t-1})\delta']^{-1}\tau = 1 \end{cases}$$

$\Leftrightarrow$

$$\begin{cases} b = \lambda[\delta V(F_t|\Omega_{t-1})\delta']^{-1}[2\delta Cov(S_t, F_t|\Omega_{t-1})'w + \delta E(F_t|\Omega_{t-1}) - \theta\tau] \\ \theta = \frac{1 - \lambda[2\delta Cov(S_t, F_t|\Omega_{t-1})'w + \delta E(F_t|\Omega_{t-1})]'[\delta V(F_t|\Omega_{t-1})\delta']^{-1}\tau}{-\lambda\tau'[\delta V(F_t|\Omega_{t-1})\delta']^{-1}\tau} \end{cases}$$

In questo modo abbiamo ottenuto la nostra soluzione ottima data dal vettore  $b$  di rapporti di copertura ottimali sui tre futures.

Per arrivare ad ottenere i tre rapporti di copertura, come vedremo in seguito, si è dovuto andare a stimare la volatilità delle serie storiche attraverso diversi modelli econometrici, così da poterne in seguito confrontare i risultati, in modo da accertare la superiorità nell'hedging di un modello rispetto ad un altro.

## CAPITOLO 2

### Modelli Autoregressivi Vettoriali

In questo capitolo, di base, verrà introdotta la stazionarietà del modello Vector Autoregressive di ordine finito. Alcune proprietà importanti saranno discussi. I principali usi di modelli vettoriali autoregressivi sono l'analisi di previsione e strutturale. In tutto questo capitolo, il modello di interesse si presume essere conosciuto. Presenteremo inoltre il modello Vector Error Correction Term, i processi integrati riferiti a questo e quelli cointegrati.

Per la presentazione dei modelli VAR e VECM si fa riferimento al libro di Lutkepohl "New Introduction to Multiple Time Series Analysis" e Hamilton "Time Series Analysis".

#### 2.1. Proprietà e assunzioni base di processi VAR

##### 2.1.1. Processi VAR( $p$ ) stabili

L'oggetto di interesse nel seguito è il modello VAR( $p$ ) (modello VAR di ordine  $p$ ):

$$y_t = v + A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t, \quad t = 0, \pm 1, \pm 2, \dots, \quad (2.1.1)$$

dove  $y_t = (y_{1t}, \dots, y_{kt})'$  è un vettore casuale ( $K \times 1$ ),  $A_i$  con  $i = 1, \dots, p$  sono matrici di coefficienti autoregressivi ( $K \times K$ ),  $v = (v_1, \dots, v_k)'$  è un vettore di costanti (intercette) che consente di ottenere un processo con media diversa da zero  $E(y_t)$ . Infine,  $u_t = (u_{1t}, \dots, u_{kt})'$  è un processo di tipo white noise di dimensione  $K$ , cioè  $E(u_t) = 0$ ,  $E(u_t u'_s) = \sum_u e$  e  $E(u_t u'_s) = 0$  per  $s \neq t$ . Inoltre la matrice di varianze e covarianze  $\sum_u$  si assume essere non singolare se non indicato diversamente.

Con lo scopo di comprendere alcune caratteristiche del modello appena descritto, si consideri il modello VAR(1)

$$y_t = v + A_1 y_{t-1} + u_t \quad (2.1.2)$$

con valore iniziale  $y_0$  aleatorio. Se questo meccanismo di generazione dei dati iniziasse al tempo  $t=1$ , si otterrebbe per sostituzione ricorsiva

$$\begin{aligned}
y_1 &= v + A_1 y_0 + u_1 \\
y_2 &= v + A_1 y_1 + u_2 = v + A_1(v + A_1 y_0 + u_1) + u_2 \\
&= (I_K + A_1)v + A_1^2 y_0 + A_1 u_1 + u_2 \\
&\vdots \\
y_t &= (I_K + A_1 + \dots + A_1^{t-1})v + A_1^t y_0 + \sum_{i=0}^{t-1} A_1^i u_{t-1} \\
&\vdots
\end{aligned} \tag{2.1.3}$$

Pertanto, i vettori  $y_1, \dots, y_t$  sono univocamente determinati da  $y_0, u_1, \dots, u_t$ . Inoltre, la distribuzione congiunta di  $y_1, \dots, y_t$  è determinata dalla distribuzione congiunta di  $y_0, u_1, \dots, u_t$ .

Anche se a volte conosciamo che un processo generatore di dati ha inizio in un determinato periodo, spesso è conveniente assumere che il processo inizi all'infinito nel passato. Questa ipotesi è fatta in realtà in (2.1.1). Che tipo di processo è consistente con il meccanismo (2.1.1) in quel caso? Per capire questa domanda consideriamo il nuovo processo VAR (1) processo (2.1.2). Da (2.1.3) abbiamo che

$$\begin{aligned}
y_t &= v + A_1 y_{t-1} + u_t \\
&= (I_K + A_1 + \dots + A_1^j)v + A_1^{j+1} y_{t-j-1} + \sum_{i=0}^j A_1^i u_{t-1}
\end{aligned}$$

Se tutti i valori di  $A_1$  sono in modulo inferiori a 1, la sequenza  $A_1^i$  con  $i = 0, 1, \dots$ , è assolutamente sommabile. Quindi, la somma all'infinito

$$\sum_{i=1}^{\infty} A_1^i u_{t-1}$$

Converge in media quadratica. Inoltre,

$$(I_K + A_1 + \dots + A_1^j)v \xrightarrow{j \rightarrow \infty} (I_K - A_1)^{-1}v$$

Inoltre,  $A_1^{j+1}$  converge alla matrice nulla per  $j \rightarrow \infty$  e, quindi, si ignora il termine  $A_1^{j+1} y_{t-j-1}$  al limite. Se tutti gli autovalori di  $A_1$  sono in modulo inferiori a 1 il processo è ben definito e quindi si può scrivere

$$y_t = \mu + \sum_{i=1}^{\infty} A_1^i u_{t-1}, \quad t = 0, \pm 1, \pm 2, \dots, \quad (2.1.4)$$

Dove

$$\mu := (I_K - A_1)^{-1} v$$

Le distribuzioni marginali e le distribuzioni congiunte di  $y_t$  sono univocamente determinate dalle distribuzioni del processo  $u_t$ . Il momento primo e secondo del processo  $y_t$  si rivelano essere

$$E(y_t) = \mu \quad \text{per tutti i } t \quad (2.1.5)$$

e

$$\begin{aligned} \Gamma_y(h) &:= E(y_t - \mu)(y_{t-h} - \mu)' \\ &= \lim_{n \rightarrow \infty} \sum_{i=0}^n \sum_{j=0}^n A_1^i E(u_{t-i} u_{t-h-j}') (A_1^j)' \\ &= \lim_{n \rightarrow \infty} \sum_{i=0}^n A_1^{h+i} \Sigma_u A_1^{i'} = \sum_{i=0}^{\infty} A_1^{h+i} \Sigma_u A_1^{i'} \end{aligned} \quad (2.1.6)$$

Perché  $E(u_t u_s') = 0$  per  $s \neq t$  e  $E(u_t u_t') = \Sigma_u$  per ogni  $t$ .

Poiché la condizione per i valori della matrice  $A_1$  è di notevole importanza, chiameremo un processo VAR (1) stabile se tutti gli autovalori di  $A_1$  sono inferiori in modulo ad 1. La condizione è equivalente a

$$\det(I_K - A_{1z}) \neq 0 \quad \text{per } |z| \leq 1 \quad (2.1.7)$$

Vale forse la pena sottolineare che il processo di  $y_t$  per  $t = 0, \pm 1, \pm 2, \dots$  può essere definito anche se la condizione di stabilità (1.1.7) non è soddisfatta. Noi non lo faremo qui perché assumeremo sempre la stabilità dei processi definiti per tutti  $t \in \mathbb{Z}$ .

La discussione precedente può essere estesa facilmente ai processi VAR ( $p$ ) con  $p > 1$ , perché ogni processo VAR ( $p$ ) può essere scritto nella forma di un VAR (1). Più precisamente, se  $y_t$  è un VAR ( $p$ ) come in (1.1.1), un corrispondente VAR (1) di dimensioni  $Kp$

$$Y_t = v + AY_{t-1} + U_t \quad (2.1.8)$$

Può essere definito, come di seguito

$$Y_t := \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix}, \quad v := \begin{bmatrix} v \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$A := \begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ I_K & 0 & \dots & 0 & 0 \\ 0 & I_K & & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I_K & 0 \end{bmatrix}, \quad U_t := \begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Alla luce di quanto detto nella precedente discussione,  $Y_t$  è stabile se

$$\det(I_{Kp} - A_z) \neq 0 \quad (2.1.9)$$

Il suo vettore medio è

$$\mu := E(Y_t) = (I_{Kp} - A)^{-1}v$$

e l'autocovarianza è

$$\Gamma_Y(h) = \sum_{i=0}^{\infty} A^{h+i} \Sigma_u (A^i)', \quad (2.1.10)$$

dove  $\Sigma_u := E(U_t U_t')$ . Usando la matrice  $(K_X K_p)$

$$J := [I_K : \dots : 0], \quad (2.1.11)$$

il processo  $y_t$  è ottenuto come  $y_t = J Y_t$ . Poiché  $Y_t$  è un ben definito stocastico processo stocastico ben definito, lo stesso vale per  $y_t$ . La media è  $E(y_t) = J\mu$  che è costante per tutte le  $t$  e l'autocovarianza  $\Gamma_y(h) = J \Gamma_Y(h) J'$  sono anche invarianti rispetto al tempo.

È facile vedere che

$$\det(I_{Kp} - A_z) = \det(I_K - A_{1z} - \dots - A_p z^p)$$

Data la definizione di polinomio caratteristico di una matrice, chiamiamo questo *polinomio inverso caratteristico* del processo VAR (p). Quindi, il processo (1.1.1) è stabile se il suo polinomio caratteristico inverso non ha radici all'interno del cerchio unitario. Formalmente  $y_t$  è stabile se

$$\det(I_K - A_{1z} - \dots - A_p z^p) \neq 0 \quad \text{per } |z| \leq 1 \quad (2.1.12)$$

Questa condizione è chiamata *condizione di stabilità*.

In sintesi, diciamo che  $y_t$  è un processo VAR (p) stabile se (1.1.12) vale e

$$y_t = J Y_t = J\mu + J \sum_{i=0}^{\infty} A^i U_{t-i} \quad (2.1.13)$$

Poiché  $U_t = (u_t', 0, \dots, 0)'$  coinvolge il processo white noise  $u_t$ , il processo  $y_t$  è determinato dal suo processo white noise. Spesso specifiche ipotesi riguardo a  $u_t$  sono fatte affinché determinino il processo  $y_t$  con la convenzione precedente. Un esempio importante è l'assunzione che  $u_t$  sia un processo white noise Gaussiano, cioè che  $u_t \sim \mathcal{N}$

$0, \sum u_t$ ) i.i.d per ogni  $t$  e  $u_t$  e  $u_s$  sono indipendenti per  $s \neq t$ . In tal caso, si può dimostrare che  $y_t$  è un processo Gaussiano, cioè  $y_t, \dots, y_{t+h}$  ha una distribuzione normale multivariata per tutti  $t$  e  $h$ .

### 2.1.2. Processi Stazionari

Sappiamo che un processo stocastico è stazionario se il primo e secondo momento sono invariati rispetto al tempo. In altre parole, un processo stocastico  $y_t$  è stazionario se

$$E(y_t) = \mu \quad \text{per tutti i } t \quad (2.1.14)$$

e

$$E[(y_t - \mu)(y_{t-h} - \mu)'] = \Gamma_y(h) = \Gamma_y(-h)' \quad \text{per tutti i } t \quad \text{e } h = 0, 1, 2, \dots \quad (2.1.15)$$

La condizione (2.1.14) sta a significare che tutte le  $y_t$  hanno lo stesso vettore medio finito  $\mu_t$  e richiede che le autocovarianze del processo non dipendano da  $t$  ma solamente al tempo  $h$  i due vettori  $y_t$  e  $y_{t-h}$  sono separati. Si noti che, se non diversamente indicato, tutte le quantità si presumono essere finite. Ad esempio,  $\mu$  è un vettore di termini medi finiti e  $\Gamma_y(h)$  è una matrice di covarianze finite. Altre definizioni di stazionarietà che sono spesso utilizzati in letteratura. Ad esempio, la distribuzione congiunta di  $n$  vettori consecutivi può essere assunto invariante rispetto al tempo per ogni  $n$ . Noi, tuttavia, utilizzeremo la definizione di quanto sopra esposto nel seguito. Chiameremo un processo strettamente stazionario se le distribuzioni congiunte di  $n$  consecutive variabili sono invariati rispetto al tempo e vi è un motivo per distinguere tra il nostro concetto di stazionarietà e la forma più rigorosa. Con la nostra definizione, il processo white noise  $u_t$  utilizzato in (2.1.1) è un esempio evidente di un processo stazionario. Affermiamo questo fatto come una proposizione.

#### Proposizione 1: Condizione di stazionarietà

Un processo VAR(p) stabile  $y_t$ , per  $t = \pm 0, \pm 1, \pm 2, \dots$ , è stazionario.

Poiché la stabilità implica stazionarietà, la condizione di stabilità (2.1.12) è spesso indicata con il nome di condizione di stazionarietà nella letteratura dell'analisi di serie

storiche. Il contrario della Proposizione 1 non è vero. In altre parole, un processo instabile non è necessariamente non stazionario. Poiché processi stazionari instabili non sono di nostro interesse nel seguito, non discuteremo di questo caso.

### 2.1.3. Autocovarianza di un processo VAR(1)

Con lo scopo di illustrare la computazione dell'autocovarianza quando i coefficienti del processo sono dati, si supponga che  $y_t$  sia un processo stabile stazionario VAR(1)

$$y_t = v + A_1 y_{t-1} + u_t$$

con una matrice di covarianze white noise  $E(u_t u_t') = \Sigma_u$ . Alternativamente, il processo può essere scritto in forma di media aggiustata come segue

$$y_t - \mu = A_1 (y_{t-1} - \mu) + u_t \quad (2.1.16)$$

Dove  $\mu = E(y_t)$ , come visto in precedenza. Moltiplicando in seguito per  $(y_{t-h} - \mu)'$  si ottiene

$$E[(y_t - \mu)(y_{t-h} - \mu)'] = A_1 E[(y_{t-1} - \mu)(y_{t-h} - \mu)'] + E[u_t (y_{t-h} - \mu)']$$

Così, per  $h=0$ ,

$$\Gamma_y(0) = A_1 \Gamma_y(-1) + \Sigma_u = A_1 \Gamma_y(1)' + \Sigma_u \quad (2.1.17)$$

e per  $h > 0$

$$\Gamma_y(h) = A_1 \Gamma_y(h-1) \quad (2.1.18)$$

Queste equazioni sono solitamente riferite come ad equazioni Yule-Walker. Se  $A_1$  e la matrice di covarianze  $\Gamma_y(0) = \Sigma_y$  di  $y_t$  sono note, allora  $\Gamma_y(h)$  può essere computata usando (2.1.18).

Se  $A_1$  e  $\Sigma_u$  sono date, allora  $\Gamma_y(0)$  può essere determinata come segue: per  $h = 1$ , otteniamo da (2.1.18),  $\Gamma_y(1) = A_1 \Gamma_y(0)$ . Allora sostituendo  $A_1 \Gamma_y(0)$  per  $\Gamma_y(1)$  in (2.1.17) otteniamo

$$\Gamma_y(0) = A_1 \Gamma_y(0) A_1' + \Sigma_u$$

e

$$\text{vec} \Gamma_y(0) = \text{vec}(A_1 \Gamma_y(0) A_1') + \text{vec} \Sigma_u$$

#### 2.1.4. Autocovarianza di processi VAR(p) stabili

Per processi VAR(p) di ordine elevato,

$$y_t - \mu = A_1(y_{t-1} - \mu) + \dots + A_p(y_{t-p} - \mu) + u_t, \quad (2.1.19)$$

le equazioni di Yule-Walker sono anche ottenute moltiplicando in seguito con  $(y_{t-h} - \mu)'$ .

Per  $h=0$ , usando  $\Gamma_y(i) = \Gamma_y(-i)'$ ,

$$\begin{aligned} \Gamma_y(0) &= A_1 \Gamma_y(-1) + \dots + A_p \Gamma_y(-p) + \Sigma_u \\ &= A_1 \Gamma_y(1)' + \dots + A_p \Gamma_y(p)' + \Sigma_u \end{aligned} \quad (2.1.20)$$

e per  $h > 0$ ,

$$\Gamma_y(h) = A_1 \Gamma_y(h-1) + \dots + A_p \Gamma_y(h-p) \quad (2.1.21)$$

Queste equazioni possono essere usate per computare  $\Gamma_y(h)$  per  $h \geq p$ , se  $A_1, \dots, A_p$  e  $\Gamma_y(p-1), \dots, \Gamma_y(0)$  sono conosciuti.

Le matrici di auto covarianza iniziali, per  $|h| < p$ , possono essere determinate usando il processo VAR(1) che corrisponde a (2.1.19),

$$Y_t - \mu = A(Y_{t-1} - \mu) + U_t \quad (2.1.22)$$

Dove  $Y_t$ ,  $A$ , e  $U_t$  sono come in (2.1.8) e  $\mu := (\mu', \dots, \mu')' = E(Y_t)$ . Procedendo quindi come nel caso del VAR(1), si ottiene

$$\Gamma_Y(0) = A\Gamma_Y(0)A' + \Sigma_U$$

Dove  $\Sigma_u = E(U_t U_t')$  e

$$\begin{aligned} \Gamma_Y(0) &= E \left( \begin{bmatrix} y_t - \mu \\ \vdots \\ y_{t-p+1} - \mu \end{bmatrix} [(y_t - \mu)', \dots, (y_{t-p+1} - \mu)'] \right) \\ &= \begin{bmatrix} \Gamma_y(0) & \Gamma_y(1) & \cdots & \Gamma_y(p-1) \\ \Gamma_y(-1) & \Gamma_y(0) & \cdots & \Gamma_y(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_y(-p+1) & \Gamma_y(-p+2) & \cdots & \Gamma_y(0) \end{bmatrix} \end{aligned}$$

### 2.1.5. Autocorrelazione di un processo VAR(p) stabile

Poiché le autocovarianze dipendono dall'unità di misura utilizzata per le variabili del sistema, a volte sono difficili da interpretare. Pertanto, le autocorrelazioni

$$R_y(h) = D^{-1}\Gamma_y(h)D^{-1} \quad (2.1.23)$$

è solitamente più conveniente lavorare come se fossero misure invarianti di dipendenza lineare tra le variabili del sistema. Qui  $D$  è una matrice diagonale con le deviazioni standard dei componenti  $y_t$  sulla diagonale principale. Cioè, gli elementi diagonali di  $D$  sono le radici quadrate degli elementi diagonali di  $\Gamma_y(0)$ . Indicando la covarianza tra  $y_i$  e  $y_{i,t-h}$  con  $\gamma_{ij}(h)$  (vale a dire,  $\gamma_{ij}(h)$  è il  $ij$ -esimo elemento di  $\Gamma_y(h)$ ) gli elementi della diagonale  $\gamma_{11}(0), \dots, \gamma_{KK}(0)$  di  $\Gamma_y(0)$  sono le varianze di  $y_{1t}, \dots, y_{Kt}$ . Così,

$$D^{-1} = \begin{bmatrix} 1/\sqrt{\gamma_{11}(0)} & & & 0 \\ & & \ddots & \\ & 0 & & 1/\sqrt{\gamma_{KK}(0)} \end{bmatrix}$$

e la correlazione tra  $y_{i,t}$  e  $y_{j,t-h}$  è

$$\rho_{ij}(h) = \frac{\gamma_{ij}(h)}{\sqrt{\gamma_{ii}(0)}\sqrt{\gamma_{jj}(0)}} \quad (2.1.24)$$

che è il  $ij$ -esimo elemento di  $R_y(h)$ .

## 2.2. Vector Error Correction Models (VECM)

Come definito in precedenza, un processo è stazionario se è invariante rispetto al tempo nel primo e secondo momento. In particolare, non ha trends né varianze mutevoli. Un processo VAR ha questa proprietà se il polinomio determinante dal suo operatore VAR ha tutte le radici al di fuori del cerchio unitario complesso. Chiaramente, processi stazionari non sono in grado di catturare alcune caratteristiche principali di molte serie storiche. Ad esempio, i trends sono abbastanza comuni nella pratica: gli investimenti, i redditi e il consumo hanno trend. Quindi, se l'interesse è analizzare le variabili originali (o i loro logaritmi) piuttosto che i rendimenti percentuali (o logaritmici), è necessario disporre di modelli che si adattino alla non stazionarietà dei dati. Ne risulta che un processo VAR può generare tendenze deterministiche e stocastiche se il polinomio determinante dell'operatore VAR ha radici sul cerchio unitario. In realtà, è anche sufficiente acconsentire alle radici unitarie (radici per  $z = 1$ ) per ottenere un comportamento in trend delle variabili. Nella sezione successiva, l'effetto di radici unitarie in un processo AR univariato sarà analizzato. Le variabili generate da tali processi sono chiamate variabili integrate ed i processi di generazione sottostanti sono processi integrati. Processi di vettori con radici unitarie saranno considerati nella sezione 2.2.1. In questi processi, alcuni delle variabili possono avere trend comuni così che si muovano insieme ad alcune misura. Esse sono quindi chiamate cointegrate. Questa funzione viene analizzato in dettaglio nella sezione 2.2.3 e si dimostra che i

Vector Error Correction Models (VECM) offrono un modo conveniente per parametrizzarli e specificarli. Una volta che si ha una configurazione adatta modello, può essere utilizzata a fini previsivi e per analisi di causalità.

### 2.2.1. Processi Integrati

Ricordando che un processo VAR(p),

$$y_t = v + A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t, \quad (2.2.1)$$

è stabile se il polinomio definito da

$$\det(I_K - A_1 z - \dots - A_p z^p)$$

non ha radici all'interno e sopra il cerchio unitario complesso. Per un processo AR(1) univariato,  $y_t = \alpha y_{t-1} + u_t$ , questa proprietà significa che

$$1 - \alpha z \neq 0 \quad \text{per } |z| \leq 1$$

o, equivalentemente,  $|\alpha| < 1$ .

Si consideri il caso limite, dove  $\alpha = 1$ . Il processo risultante  $y_t = y_{t-1} + u_t$  è chiamato random walk. Incominciando il processo al tempo  $t = 0$  con alcuni  $y_0$  fissati, è facile vedere con sostituzioni successive per  $y_t$  ritardati, che

$$y_t = y_{t-1} + u_t = y_{t-2} + u_{t-1} + u_t = \dots = y_0 + \sum_{i=1}^t u_i \quad (2.2.2)$$

Così,  $y_t$  consiste nella somma di tutti i disturbi o delle innovazioni del precedente periodi in modo che ogni disturbo abbia un impatto duraturo sul processo. Se  $u_t$  è un white noise con varianza  $\sigma_u^2$ ,

$$E(y_t) = y_0$$

e

$$\text{Var}(y_t) = t\text{Var}(u_t) = t\sigma_u^2$$

Quindi, la varianza di un random walk tende all'infinito. Inoltre, la correlazione

$$\begin{aligned}\text{Corr}(y_t, y_{t+h}) &= \frac{E[(\sum_{i=1}^t u_i)(\sum_{i=1}^{t+h} u_i)]}{[t\sigma_u^2(t+h)\sigma_u^2]^{1/2}} \\ &= \frac{t}{(t^2 + th)^{1/2}} \xrightarrow{t \rightarrow \infty} 1\end{aligned}$$

per ogni intero  $h$ . Quest'ultima proprietà del random walk significa che  $y_t$  e  $y_s$  sono fortemente correlati anche se sono lontani nel tempo. Può anche essere dimostrato che il tempo atteso tra due incroci di zero è infinito. Queste proprietà si riflettono spesso nei comportamenti in trend.

Se il processo ha un termine  $v$  costante diverso da zero,  $y_t = v + y_{t-1} + u_t$  viene chiamato random walk con drift e ha un andamento lineare deterministico nella media.

Per vedere questa proprietà, supponiamo ancora una volta che il processo sia iniziato in  $t = 0$  con un  $y_0$  fisso. Allora

$$y_t = y_0 + tv + \sum_{i=1}^t u_i$$

e  $E(y_t) = y_0 + tv$ .

La discussione precedente suggerisce che l'avvio di processi instabili ad un certo tempo finito  $t_0$  è utile per ottenere processi con momenti finiti. Dall'altro lato, se un processo AR inizia a un certo tempo finito, non è necessariamente stazionario, anche se è stabile. Per vedere questa proprietà, si supponga che  $y_t = v + \alpha y_{t-1} + u_t$  sia processo AR(1) univariato stabile con  $|\alpha| < 1$ . Partendo con una variabile casuale  $y_0$  al tempo  $t = 0$ , si ottiene

$$y_t = v \sum_{i=0}^{t-1} \alpha^i + \alpha^t y_0 + \sum_{i=0}^{t-1} \alpha^i u_{t-1}$$

Quindi,

$$E(y_t) = v \sum_{i=0}^{t-1} \alpha^i + \alpha^t E(y_0)$$

Generalmente è invariante rispetto al tempo se  $\alpha$  e  $v \neq 0$ . Un risultato simile è ottenuto per il secondo momento,

$$Var(y_t) = \alpha^{2t} Var(y_0) + \sigma_u^2 \sum_{i=0}^{t-1} \sigma^{2i}$$

Tuttavia, i momenti primo e il secondo approssimano valori limite come  $t \rightarrow \infty$  e si potrebbe chiamare questo processo asintoticamente stazionario. Per semplificare le cose, il termine "asintoticamente" a volte viene eliminato e tali processi sono poi chiamati semplicemente stazionari. Inoltre, se si considerano esclusivamente i processi stocastici senza termini deterministici ( $v = 0$ ), la variabile iniziale può essere scelta in modo tale  $y_t$  che sia stazionario se il processo è stabile. In particolare, se si sceglie

$$y_0 = \sum_{i=0}^{\infty} \alpha^i u_{-i}$$

Otteniamo, per  $v = 0$

$$y_t = \alpha^t \sum_{i=0}^{\infty} \alpha^i u_{-i} + \sum_{i=0}^{t-1} \alpha^i u_{t-i} = \sum_{i=0}^{\infty} \alpha^i u_{t-i} \quad t = 1, 2, \dots,$$

e, quindi, per  $t=1, 2, \dots$ ,

$$E(y_t) = 0$$

$$\text{Var}(y_t) = \sigma_u^2 / (1 - \alpha^2),$$

e anche l'autocovarianza è invariante rispetto al tempo. Così, per un processo stabile si può infatti scegliere la variabile iniziale tale che  $y_t$  sia stazionario, anche se il processo viene avviato in un dato periodo.

Un comportamento simile a quello di un random walk è anche osservato per processi AR di ordine maggiore quali

$$y_t = v + \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + u_t,$$

se  $1 - \alpha_1 z - \dots - \alpha_p z^p$  ha una radice per  $z = 1$ . Si noti che

$$1 - \alpha_1 z - \dots - \alpha_p z^p = (1 - \lambda_1 z) \dots (1 - \lambda_p z),$$

dove  $\lambda_1, \dots, \lambda_p$  sono i reciproci delle radici del polinomio. Se il processo ha solo una radice unitaria e tutti le altre radici sono al di fuori del cerchio unitario complesso, il suo comportamento è simile a quello di un random walk, vale a dire, che le sue variazioni aumentano linearmente, la correlazione tra le variabili nei periodi  $h$  tende a 1 e il processo ha un andamento lineare in media se  $v \neq 0$ . Nel caso in cui una delle radici è strettamente dentro il cerchio unitario, il processo diventa esplosivo, ciò vuol dire che le sue variazioni tendano all'infinito a un ritmo esponenziale. Molti ricercatori ritengono che tali processi siano modelli irreali per la maggior parte di dati economici. Inoltre processi con radici sopra il cerchio unitario sono spesso utile, ci concentreremo solo sul caso di radici unitarie e di tutte le altre radici al di fuori del cerchio unitario. Questa situazione è di notevole interesse pratico.

Processi univariati con radici unitarie  $d$  ( $d$  radici pari a 1) nei loro operatori AR sono chiamati integrati di ordine  $d$  ( $I(d)$ ). Se vi è solo una radice unitaria, cioè, il processo è  $I(1)$ , è abbastanza facile vedere come un processo stabile e possibilmente stazionario possa essere ottenuto: semplicemente facendo le differenze prime,  $\Delta y_t = (1 - L) y_t = y_t - y_{t-1}$ , del processo originario. Più in generale, se il processo è  $I(d)$  può essere reso stabile differenziandolo  $d$  volte, cioè  $\Delta^d y_t = (1-L)^d y_t$  è stabile e, di nuovo, i valori iniziali possono essere scelti in modo tale che esso sia stazionario.

### 2.2.2. Processi cointegrati e Vector Error Correction Models

Relazioni di equilibrio si conoscono tra molte variabili economiche come il reddito familiare e le spese o prezzi della stessa merce nei diversi mercati. Supponiamo che le variabili di interesse siano raccolte nel vettore  $y_t = (y_{1t}, \dots, y_{Kt})'$  e la loro relazione di equilibrio di lungo periodo sia  $\beta' y_t = \beta_1 y_{1t} + \dots + \beta_K y_{Kt} = 0$ , dove  $\beta = (\beta_1, \dots, \beta_K)'$ . In ogni periodo particolare, questo rapporto non può essere soddisfatto esattamente, ma si può avere  $\beta' y_t = z_t$ , dove  $z_t$  è una variabile stocastica che rappresenta le deviazioni dalla equilibrio. Se c'è davvero un equilibrio, sembra plausibile assumere che le variabili  $y_t$  si muovano insieme e che  $z_t$  sia stabile. Questa impostazione, tuttavia, non esclude la possibilità che le variabili  $y_t$  vaghino a lungo come un gruppo. Così, essi possono essere guidate da un trend stocastico comune. In altre parole, non è escluso che ciascuna variabile sia integrata, ma esiste una combinazione lineare delle variabili che è stazionaria. Variabili integrate con questa proprietà sono chiamate cointegrate.

In generale, le variabili in un processo  $y_t$  di dimensione-K sono denominate cointegrate di ordine (d,b), brevemente,  $y_t \sim CI(d,b)$ , se tutti i componenti di  $y_t$  sono  $I(d)$  ed esiste una combinazione lineare  $z_t := \beta' y_t$  con  $\beta = (\beta_1, \dots, \beta_K)' \neq 0$  tale che  $z_t$  sia  $I(d-b)$ . Ad esempio, se tutti i componenti di  $y_t$  sono  $I(1)$  e  $\beta' y_t$  è stazionario ( $I(0)$ ), allora  $y_t \sim CI(1, 1)$ . Il vettore  $\beta$  è chiamato vettore di cointegrazione. Un processo che consta di variabili cointegrate viene chiamato processo cointegrato. Questi processi sono stati introdotti da Granger (1981) e Engle e Granger (1987). Da allora sono diventati popolari nei lavori di econometria teorica ed applicata.

Nel seguito, una definizione leggermente diversa di cointegrazione sarà utilizzata al fine di semplificare la terminologia. Chiamiamo un processo integrato  $y_t$  di dimensione-K di ordine d, brevemente,  $y_t \sim I(d)$ , se  $\Delta^d y_t$  è stabile e  $\Delta^{d-1} y_t$  non è stabile. Il processo  $y_t$   $I(d)$  viene chiamato cointegrato se vi è una combinazione lineare  $\beta' y_t$  con  $\beta \neq 0$  che è integrato di ordine inferiore a d. Questa definizione differisce da quella data da Engle e Granger (1987) in quanto non si escludono componenti  $y_t$  con ordine di integrazione inferiore a d. Se c'è solo un  $I(d)$  componente in  $y_t$  e tutti gli altri componenti sono stabili ( $I(0)$ ), allora il vettore  $y_t$  è  $I(d)$  secondo la nostra definizione perché  $\Delta^d y_t$  è stabile e  $\Delta^{d-1} y_t$  non lo è. In tal caso un relazione  $\beta' y_t$  che coinvolga solo i componenti fissi è una relazione di cointegrazione nei nostri termini. Chiaramente, questo aspetto della nostra definizione non è in linea con l'idea originale di cointegrazione come una

speciale relazione tra variabili integrate con trend stocastici comuni. Nel seguito, la nostra definizione è ancora utile perché semplifica la terminologia evitando di distinguere tra variabili con diversi ordini di integrazione. Tuttavia il lettore dovrebbe tenere a mente le idee di base di cointegrazione quando si tratta di interpretare le relazioni specifiche.

Ovviamente, un vettore di cointegrazione non è unico. Inoltre, vi possono essere vari vettori di cointegrazione linearmente indipendenti. Per esempio, se ci sono quattro variabili in un sistema, i due primi possono essere collegati da una relazione di equilibrio di lungo periodo e anche gli ultimi due. Così, ci può essere un vettore di cointegrazione con zeri nelle ultime due posizioni e uno con zeri nelle prime due posizioni. Inoltre, vi può essere una relazione di cointegrazione che coinvolga tutte le quattro variabili.

Prima che il concetto di cointegrazione fosse introdotto, le teorie sugli error correction models è stata discussa nella letteratura econometria (v., ad esempio, Davidson, Hendry, Srba e Yeo (1978), Hendry & von Ungern-Sternberg (1981), Salmon (1982)). In un error correction model, i cambiamenti in una variabile dipendono dalle deviazioni di alcuni relazione di equilibrio. Supponiamo, per esempio, che  $y_{1t}$  rappresenta il prezzo di una commodity in un determinato mercato e  $y_{2t}$  è il prezzo corrispondente della stessa commodity in un altro mercato. Si supponga inoltre che la relazione di equilibrio tra le due variabili sia data da  $y_{1t} = \beta_1 y_{2t}$  e che i cambiamenti nella  $y_{1t}$  dipendano dalle deviazioni da questo equilibrio nel periodo  $t - 1$ ,

$$\Delta y_{1t} = \alpha_1 (y_{1,t-1} - \beta_1 y_{2,t-1}) + u_{1t}$$

Una relazione simile può essere ottenuta per  $y_{2t}$

$$\Delta y_{2t} = \alpha_2 (y_{1,t-1} - \beta_1 y_{2,t-1}) + u_{2t}$$

In un error correction model più generale,  $\Delta y_{it}$  può inoltre dipendere in precedenti cambiamenti di entrambe le variabili come, per esempio, nel seguente modello:

$$\Delta y_{1t} = \alpha_1 (y_{1,t-1} - \beta_1 y_{2,t-1}) + \gamma_{11,1} \Delta y_{1,t-1} + \gamma_{12,1} \Delta y_{2,t-1} + u_{1t}$$

$$\Delta y_{2t} = \alpha_2(y_{1,t-1} - \beta_1 y_{2,t-1}) + \gamma_{21,1} \Delta y_{1,t-1} + \gamma_{22,1} \Delta y_{2,t-1} + u_{2t} \quad (2.2.3)$$

Ulteriori ritardi di  $\Delta y_{it}$  possono essere inclusi.

Per vedere la stretta relazione tra gli error correction models e il concetto di cointegrazione, supponiamo che  $y_{1t}$  e  $y_{2t}$  siano entrambi variabili  $I(1)$ . In tale caso tutti i termini in (2.2.3) che coinvolgono  $\Delta y_{it}$  sono stabili. Inoltre,  $u_{1t}$  e  $u_{2t}$  sono errori white noise, che sono anche stabili. Perché un termine instabile non può essere pari a un processo stabile,

$$\alpha_i(y_{1,t-1} - \beta_1 y_{2,t-1}) = \Delta y_{it} - \gamma_{i1,1} \Delta y_{1,t-1} - \gamma_{i2,1} \Delta y_{2,t-1} - u_{it}$$

Deve essere anche stabile. Quindi, se  $\alpha_1 \neq 0$  o  $\alpha_2 \neq 0$ ,  $y_{1t} - \beta_1 y_{2t}$  è stabile e, quindi, rappresenta una relazione di cointegrazione.

In notazione vettoriale e matriciale il modello (2.2.3) può essere riscritto come

$$\Delta y_t = \alpha \beta' y_{t-1} + \Gamma_1 \Delta y_{t-1} + u_t$$

o

$$y_t - y_{t-1} = \alpha \beta' y_{t-1} + \Gamma_1 (y_{t-1} - y_{t-2}) + u_t \quad (2.2.4)$$

dove  $y_t := (y_{1t} y_{2t})'$ ,  $u_t := (u_{1t} u_{2t})'$ ,

$$\alpha := \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}, \quad \beta' := (1, -\beta_1) \quad e \quad \Gamma_1 := \begin{bmatrix} \gamma_{11,1} & \gamma_{12,1} \\ \gamma_{21,1} & \gamma_{22,1} \end{bmatrix}$$

Ridisponendo i termini in (2.2.4) si ottiene la rappresentazione VAR(2)

$$y_t = (I_K + \Gamma_1 + \alpha \beta') y_{t-1} - \Gamma_1 y_{t-2} + u_t$$

Quindi, variabili cointegrate possono essere generate da un processo VAR.

Per vedere come la cointegrazione può sorgere più in generale in un processo VAR di dimensione  $K$ , si consideri il processo VAR(2)

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + u_t \quad (2.2.5)$$

con  $y_t = (y_{1t}, \dots, y_{Kt})'$ . Si supponga che il processo non sia stabile con

$$|I_K - A_1 z - A_2 z^2| = (1 - \lambda_1 z) \dots (1 - \lambda_n z) = 0 \quad \text{per } z = 1$$

Poiché i  $\lambda_i$  sono i reciproci delle radici del polinomio determinante, uno o più di essi deve essere uguale a 1. Tutte le altre radici sono assunte essere esterne alla circonferenza unitaria, cioè, tutti i  $\lambda_i$  che non sono uno sono all'interno del cerchio unitario complesso. Poiché  $|I_K - A_1 - A_2| = 0$ , la matrice

$$\Pi := -(I_K - A_1 - A_2)$$

è singolare. Si supponga  $rk(\Pi) = r < K$ . Allora  $\Pi$  può essere decomposto come  $\Pi = \alpha\beta'$ , dove  $\alpha$  e  $\beta$  sono matrici  $(K \times r)$ . dalla discussione nella sezione precedente, sappiamo che ogni variabile diventa stazionaria semplicemente differenziandola. Supponiamo che la differenziazione sia sufficiente una sola volta, sottraiamo  $y_{t-1}$  su entrambi i lati

di (2.2.5) e riorganizzare i termini come

$$y_t - y_{t-1} = -(I_K - A_1 - A_2)y_{t-1} - A_2 y_{t-1} + A_2 y_{t-2} + u_t$$

o

$$\Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + u_t \quad (2.2.6)$$

dove  $\Gamma_1 := -A_2$ , o

$$\alpha\beta' y_{t-1} = \Delta y_t - \Gamma_1 \Delta y_{t-1} - u_t$$

Poiché il lato destro comporta termini stazionari solo, anche  $\alpha\beta' y_{t-1}$  deve essere stazionario e rimane stazionario sulla moltiplicazione per  $(\alpha'\alpha)^{-1}\alpha'$ . In altre parole,  $\beta' y_t$

è stazionario e, quindi, ogni elemento di  $\beta'y_t$  rappresenta una relazione di cointegrazione. Si noti che semplicemente prendendo le differenze prime di tutte le variabili in (2.2.5) si elimina il termine di cointegrazione che potrebbe contenere le relazioni di grande importanza per una particolare analisi. Inoltre, in generale, un processo VAR con variabili cointegrate non ammette una rappresentazione pura VAR sulle differenze prime.

Può anche essere utile sottolineare che qui abbiamo lavorato sotto il presupposto che tutte le variabili siano stazionarie dopo una differenziazione. In generale, variabili con ordini di integrazione più elevati possono anche essere presenti. In questo caso,  $\beta'y_t$  non può essere stazionario anche se  $rk(\Pi) = r < K$ . I componenti di  $y_t$  possono ancora essere cointegrate di un ordine superiore, se esistono combinazioni lineari che hanno un ordine di integrazione ridotta.

Nel seguito, saremo interessati al caso particolare in cui tutti le singoli variabili siano  $I(1)$  o  $I(0)$ . Il processo VAR (p) di dimensioni  $K$

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t \quad (2.2.7)$$

è chiamato cointegrato di ordine  $r$  se

$$\Pi := -(I_K - A_1 - \dots - A_p)$$

ha rango  $r$  e, quindi,  $\Pi$  può essere scritto come prodotto matriciale  $\alpha\beta'$  con  $\alpha$  e  $\beta$  di dimensione  $(K \times r)$  e rango  $r$ . La matrice  $\beta$  viene chiamata matrice di cointegrazione e  $\alpha$  è talvolta chiamato matrice di carico (loading matrix). Se  $r = 0$ ,  $\Delta y_t$  una rappresentazione VAR (p-1) stabile e, per  $r = K$ ,  $|I_K - A_1 - \dots - A_p| = |-\Pi| \neq 0$  e, quindi, l'operatore VAR non ha radici unitarie in modo tale che  $y_t$  è un processo VAR stabile (p). Riscrivendo (2.2.7) come in (2.2.6) otteniamo una rappresentazione vector error correction model (VECM)

$$\begin{aligned} \Delta y_t &= \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{p-1} \Delta y_{t-p+1} + u_t \\ &= \alpha\beta' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{p-1} \Delta y_{t-p+1} + u_t \end{aligned} \quad (2.2.8)$$

dove

$$\Gamma_1 := -(A_{i+1} + \dots + A_p), \quad i = 1, \dots, p-1$$

Se questa rappresentazione di un processo cointegrato è data, è facile recuperare la corrispondente forma VAR (2.2.7) notando che

$$\begin{aligned} A_1 &= \Pi + I_K + \Gamma_1 \\ A_i &= \Gamma_i - \Gamma_{i-1}, \quad i = 2, \dots, p-1, \\ A_p &= -\Gamma_{p-1} \end{aligned} \tag{2.2.9}$$

Vale la pena sottolineare che possiamo anche riorganizzare i termini in un modo diverso e ottenere una rappresentazione

$$\Delta y_t = D_1 \Delta y_{t-1} + \dots + D_{p-1} \Delta y_{t-p+1} + \Pi y_{t-p} + u_t \tag{2.2.10}$$

dove l'error correction term appare al ritardo  $p$  e

$$D_i = -(I_K - A_1 - \dots - A_i), \quad i = 1, \dots, p-1$$

Infine nello specificare il modello VECM l'ordine di ritardo, il rango di cointegrazione ed eventuali restrizioni devono essere determinati. L'ordine di ritardo (lag) e l'ordine di cointegrazione sono in genere determinati prima che ulteriori restrizioni siano imposte sui parametri delle matrici. Inoltre, la definizione di un VECM di solito inizia determinando una lunghezza di lag adatta perché, nella scelta l'ordine di ritardo, il rango di cointegrazione non è necessariamente noto, mentre molte procedure per la specificazione del rango di cointegrazione richiedono la conoscenza dell'ordine del ritardo.

## CAPITOLO 3

### Modelli GARCH Multivariati

In questo capitolo, svilupperemo le proprietà teoriche di una nuova classe di modelli GARCH-Multivariati in grado di stimare matrici di covarianze di grandi dimensioni che variano nel tempo, denominati Dynamic Conditional Correlation (DCC) Multivariate GARCH, riprendendo il modello da Engle e Sheppard (2001). Mostreremo che il problema della stima della covarianza multivariata condizionale può essere semplificato stimando modelli GARCH univariati per la varianza di ogni bene, e poi, utilizzando i residui trasformati risultanti dalla prima fase, stimando una correlazione condizionale che varia nel tempo. Gli errori standard dei parametri del primo step restano consistenti, e solo gli errori standard per i parametri di correlazione hanno bisogno di essere modificati. Saranno inoltre presentate brevemente le proprietà teoriche dei modelli CCC e BEKK, così da poterli mettere a confronto con il DCC.

#### 3.1. Introduzione ai modelli GARCH Multivariati

Mentre i modelli GARCH univariati hanno avuto ampio successo empirico, i problemi associati con la stima di modelli GARCH multivariati con correlazione dinamica hanno costretti i ricercatori a stimare modelli con portata limitata o restrizioni considerevoli. Matrici di covarianze di grandi dimensioni che variano nel tempo sono necessarie nella gestione e ottimizzazione del portafoglio, modelli della struttura a termine dei treasuries o di commodities, e autoregressioni vettoriali di grandi dimensioni. In questo capitolo descriveremo un modello che può essere usato per stimare matrici di covarianze estremamente grandi variabili nel tempo e descriveremo le proprietà teoriche dei modelli GARCH-Multivariato con correlazione dinamica condizionale (DCC), introdotti per la prima volta da Engle (2001). Questa classe di modelli MV-GARCH si differenzia da altre specifiche in quanto i modelli GARCH univariati sono stimati per ogni serie di asset, e quindi, usando i residui standardizzati risultanti dalla prima fase, una matrice di correlazione che varia nel tempo è stimata utilizzando un semplice specificazione. Questa parametrizzazione preserva la semplice interpretazione dei modelli GARCH univariati con un facile stimatore correlazione da calcolare. Questa stima multi-fase richiede la modifica degli errori standard dei parametri, ma gli

Bollerslev-Wooldridge errori standard per ogni modello GARCH univariato rimangono consistenti, e solo la covarianza asintotica dei parametri dello stimatore correlazione deve essere modificata. Bollerslev, Engle e Wooldridge (1988) originariamente proposero un modello GARCH multivariato forma vech che ha fornito un quadro molto generale per i modelli multivariati di stima della volatilità. Il modello completo senza restrizioni richiede  $O(k^4)$  parametri da stimare con la massima likelihood, dove  $k$  è il numero di serie storiche che vengono modellate. Un modello più semplice, il diagonal vech è stato anche proposto che consente, per coefficienti diversi da zero solo su propri effetti ritardati, di ridurre il numero di parametri che devono essere stimati  $O(k^2)$ . La specificazione diagonale consente una interpretazione relativamente semplice, in quanto ogni serie ha una specificazione GARCH-like. Tuttavia, derivare le necessarie restrizioni sui parametri per garantire la positiva determinatezza della covarianza condizionale diventa estremamente difficile al crescere di  $k$ .

Bollerslev (1990) ha introdotto il modello GARCH multivariato con correlazione costante condizionale (CCC), dove sono stimati modelli GARCH univariati per ogni asset e quindi la matrice di correlazione viene stimata utilizzando lo stimatore di correlazione MLE trasformando i residui stimati usando le loro deviazioni standard condizionali. L'assunzione di correlazione costante rende flessibile la stima di un grande modello e assicura che lo stimatore sia definito positivamente, semplicemente chiedendo ad ogni varianza univariata condizionale di essere diversa da zero e la matrice di correlazione sia di rango pieno. Tuttavia, lo stimatore di correlazione costante, così come proposto, non fornisce un metodo per costruire coerenti errori standard utilizzando il processo di stima multi-step. Bollerslev trova plausibile la nozione di correlazione costante, tuttavia l'opera recente di Tsui e Yu (1999) ha trovato che la correlazione costante può essere respinta per le attività certe. Bera (1996) e Tse (2000) entrambi hanno sviluppato test di correlazione costanti, il primo un test bivariato mentre l'ultima è un test più generale LM multivariato.

La formulazione BEKK, proposta in Engle e Kroner (1995), ha sviluppato una forma quadratica generale per l'equazione di covarianza condizionale che ha eliminato il problema di assicurare la positiva determinatezza della stima della covarianza condizionale del modello originale vech. Affinché il modello BEKK possa essere pienamente generale, il numero di parametri che devono essere stimati è  $O(k^4)$ , ma una

stima standard con BEKK di norma comporta  $O(k^2)$  parametri. Altre formulazioni più trattabili del modello BEKK comprendono diagonali e scalari che pongono dei limiti su alcuni parametri affinché siano pari a zero, anche se queste restrizioni sono tipicamente rifiutabili. Oltre al numero elevato di parametri che devono essere stimati per la forma generale, l'interpretazione esatta dei singoli coefficienti è difficile stimare. Recentemente, Alexander (2000) ha sostenuto l'uso di modelli GARCH fattore, come prima delineato Engle, Ng, e Rothschild (1990), per la stima delle matrici di covarianze di grandi dimensioni. Factor o Orthogonal MV-GARCH forniscono un metodo per la stima di ogni matrice di covarianze dinamica usando solo i modelli GARCH univariati. Alexander mostra come un numero limitato di fattori possa spiegare una quantità significativa della volatilità in alcuni casi. Tuttavia, questo approccio, riducendo il numero di parametri stimati a  $o(k)$ , è limitato sia dalla difficoltà di interpretare i coefficienti sui modelli GARCH univariati sia dalla scarsa performance per i sistemi meno correlati ad esempio sull'equities.

Engle (2001) ha proposto una nuova classe di stimatore che conserva sia la facilità di stima del modello di correlazione costante di Bollerslev permettendo però di avere correlazioni non costanti nel tempo. I modelli Dynamic conditional correlation MV-GARCH conservano la parsimonia dei modelli GARCH univariati delle singole classi di volatilità con un semplice correlazione dinamica GARCH-like. Inoltre, il numero di parametri stimata utilizzando la massima verosimiglianza (maximum likelihood) è  $O(k)$ , un notevole miglioramento rispetto sia vech che ai modelli BEKK. Ancora più importante, il numero di parametri che richiedono una stima simultanea è  $O(1)$ . L'obiettivo di questo capitolo è quello di esplorare teoricamente le proprietà del modello MV-GARCH DCC, quando si stimano grandi matrici di covarianza condizionali.

Tse e Tsui (1998) hanno proposto un modello GARCH multivariato con correlazione dinamica, tuttavia nessun tentativo è stato fatto per consentire la stima separata dei processi GARCH univariati e lo stimatore di correlazione dinamica. Inoltre, non consentendo il ritorno alla correlazione incondizionata nello stimatore di correlazione, il numero di parametri che deve essere contemporaneamente stimato è  $O(k^2)$ , ed è solo leggermente inferiore alla formulazione tipica di un BEKK.

### 3.2. Il modello CCC

Si consideri il modello GARCH Multivariato CCC sviluppato da Bollerslev (1990):

$$y_t = E(y_t|\Omega_{t-1}) + \varepsilon_t, \quad \varepsilon_t = D_t \eta_t \quad (3.2.1)$$

$$\text{var}(\varepsilon_t|\Omega_{t-1})$$

Dove  $y_t = (y_{1t}, \dots, y_{mt})'$ ,  $\eta_t = (\eta_{1t}, \dots, \eta_{mt})'$  sono sequenze di vettori casuali indipendentemente ed identicamente distribuiti,  $\Omega_{t-1}$  è l'informazione passata disponibile al tempo t,  $D_t = \text{diag}\left(h_{1t}^{\frac{1}{2}}, \dots, h_{mt}^{\frac{1}{2}}\right)$ , m è il numero dei rendimenti, e  $t=1, \dots, n$ . Come  $\Gamma = E(\eta_t \eta_t' | \Omega_{t-1}) = E(\eta_t \eta_t')$ , dove  $\Gamma = \{\rho_{ij}\}$  per  $i, j=1, \dots, m$ , la matrice di correlazione costante condizionale degli shocks non condizionali,  $\eta_t$ , è equivalente alla matrice di covarianza costante condizionale degli shocks non condizionali,  $\varepsilon_t$ , dalla (3.2.1),  $\varepsilon_t \varepsilon_t' = D_t \eta_t \eta_t' D_t$ ,  $D_t = (\text{diag} Q_t)^{1/2}$ , e  $E(\varepsilon_t \varepsilon_t' | \Omega_{t-1}) = Q_t = D_t \Gamma D_t$ , dove  $Q_t$  è la matrice di covarianza condizionale.

Il modello CCC di Bollerslev (1990) assume che la varianza condizionale per ogni rendimento,  $h_{it} = 1, \dots, m$ , segue un processo GARCH univariato, che è

$$h_{it} = \omega_i + \sum_{j=1}^r \alpha_{ij} \varepsilon_{i,t-j}^2 + \sum_{j=1}^s \beta_{ij} h_{i,t-j}$$

Dove  $\alpha_{ij}$  rappresenta l'effetto ARCH e  $\beta_{ij}$  rappresenta l'effetto GARCH, mentre  $\sum_{j=1}^r \alpha_{ij} + \sum_{j=1}^s \beta_{ij}$  denota la persistenza di lungo termine.

### 3.3 Il modello BEKK

Il modello BEKK sviluppato da Enlge e Kroner (1995), ha la proprietà attraente che le matrici di covarianza condizionale sono definite positivamente. Il modello BEKK per un GARCH (1,1) Multivariato è dato da:

$$H_t = C' C + A' \varepsilon_{t-1} \varepsilon_{t-1}' A + B' H_{t-1} B$$

Dove i singoli elementi delle matrici A, B e C sono dati da

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}, \quad C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

Con  $\alpha_{ii}^2 + \beta_{ii}^2 < 1$ ,  $i = 1, 2$  per la stazionarietà. In questa rappresentazione diagonale, le varianze condizionali sono funzioni dei propri valori ritardati e dei propri rendimenti degli shocks ritardati, mentre le covarianze condizionali sono funzioni delle covarianze ritardati e dei ritardi dei prodotti incrociati degli corrispondenti rendimenti degli shocks. Inoltre, questa formulazione garantisce  $H_t$  essere definita positivamente quasi certamente per ogni  $t$ .

### 3.4. Il modello DCC

#### 3.4.1. Matrice di varianze e covarianze, correlazioni e log-likelihood

Il modello GARCH multivariato proposto presuppone che i rendimenti degli assets  $k$  siano condizionalmente multivariati normali con valore atteso uguale a zero e la matrice di covarianza  $H_t$ . I rendimenti possono essere in media zero oppure i residui filtrati di serie storiche.

$$r_t | \Omega_{t-1} \sim N(0, H_t)$$

e

$$H_t \equiv D_t R_t D_t$$

Dove  $D_t$  è la matrice diagonale  $k \times k$  della deviazione standard dinamica dei modelli GARCH univariati con  $\sqrt{h_{it}}$  sulla diagonale  $i$ -esima, e  $R_t$  la matrice di correlazione variabile rispetto al tempo.

La log-likelihood di questo stimatore può essere scritta

$$L = -\frac{1}{2} \sum_{t=1}^T (k \log(2\pi) + \log(|H_t|) + r_t' D_t^{-1} R_t^{-1} D_t^{-1} r_t)$$

$$\begin{aligned}
&= -\frac{1}{2} \sum_{t=1}^T (k \log(2\pi) + \log(|D_t R_t D_t|) + r' H_t^{-1} r_t) \\
&= -\frac{1}{2} \sum_{t=1}^T (k \log(2\pi) + 2 \log(|D_t|) + \log(|R_t|) + \epsilon_t' R_t^{-1} \epsilon_t)
\end{aligned}$$

dove  $\epsilon_t \sim N(0, R_t)$  sono i residui standardizzati dalla loro deviazione standard condizionale. Si propone di riscrivere gli elementi di  $D_t$  come un modello GARCH univariato, così che

$$h_{it} = w_i + \sum_{p=1}^{P_i} \alpha_{ip} r_{it-p}^2 + \sum_{q=1}^{Q_i} \beta_{iq} h_{it-q}$$

Per  $i = 1, 2, \dots, k$  con le solite restrizioni per la non negatività e stazionarietà, come per la non negatività delle varianze e  $\sum_{p=1}^{P_i} \alpha_{ip} + \sum_{q=1}^{Q_i} \beta_{iq} < 1$ . I pedici sono presenti sui singoli P e Q per ciascuna serie per indicare che la lunghezza dei lag scelti non necessariamente deve essere la stessa. La specificazione dei modelli GARCH univariati non è limitata al normale GARCH (p,q), ma può includere qualsiasi processo GARCH distribuito normalmente con gli errori che soddisfano condizioni di stazionarietà adeguate e non-negatività vincoli. La struttura proposta a correlazione dinamica è la seguente:

$$Q_t = \left( 1 - \sum_{m=1}^M \alpha_m - \sum_{n=1}^N \beta_n \right) \bar{Q} + \sum_{m=1}^M \alpha_m (\epsilon_{t-m} \epsilon_{t-m}') + \sum_{n=1}^N \beta_n Q_{t-n}$$

$$R_t = Q_t^{*-1} Q_t Q_t^{*-1}$$

dove  $\bar{Q}$  è la covarianza non condizionale dei residui standardizzati, risultanti dal primo passo della stima, e

$$Q_t^* = \begin{bmatrix} \sqrt{q_{11}} & 0 & 0 & \dots & 0 \\ 0 & \sqrt{q_{22}} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \sqrt{q_{kk}} \end{bmatrix}$$

così che  $Q_t^*$  è una matrice diagonale composta dalla radice quadrata degli elementi diagonali di  $Q_t$ .

L'elemento tipico di  $R_t$  sarà nella forma  $\rho_{ijt} = \frac{q_{ijt}}{\sqrt{q_{ii}q_{jj}}}$

### 3.4.2 Stime ed errori standard

Il modello DCC è stato progettato per consentire la stima in due fasi, dove nella prima fase modelli GARCH Univariati sono stimati per ogni serie di residui, e nella seconda fase, i residui, trasformati dalla loro deviazione standard stimata durante la prima fase, vengono utilizzati per stimare i parametri della correlazione dinamica. La likelihood utilizzata nella prima fase comporta la sostituzione di  $R_t$  con  $I_k$ , matrice identità di dimensione  $k$ . Siano i parametri del modello,  $\theta$ , scritti in due gruppi  $(\phi_1, \phi_2, \dots, \phi_k, \Psi) = (\phi, \Psi)$  dove gli elementi di  $\phi_i$  corrispondono ai parametri del modello GARCH univariato per la serie  $i$ -esima,  $\phi_i = (\omega, \alpha_{1i}, \dots, \alpha_{pi}, \beta_{1i}, \dots, \beta_{qi})$ . La funzione quasi-likelihood risultante dal primo step:

$$\begin{aligned} QL_1(\theta|r_t) &= -\frac{1}{2} \sum_{t=1}^T (k \log(2\pi) + \log(|I_k|) + 2 \log(|D_t|) + r' D_t^{-1} R_t^{-1} D_t^{-1} r_t) \\ &= -\frac{1}{2} \sum_{t=1}^T (k \log(2\pi) + 2 \log(|D_t|) + r' D_t^{-2} r_t) \\ &= -\frac{1}{2} \sum_{t=1}^T \left( k \log(2\pi) + \sum_{n=1}^k \left( \log(h_{it}) + \frac{r_{it}^2}{h_{it}} \right) \right) \\ &= -\frac{1}{2} \sum_{t=1}^T \left( T \log(2\pi) + \sum_{n=1}^T \left( \log(h_{it}) + \frac{r_{it}^2}{h_{it}} \right) \right) \end{aligned}$$

che è semplicemente la somma dei log-likelihoods dei singoli modelli GARCH per ciascun assets. Una volta che la prima fase è stata stimata, la seconda fase viene stimata utilizzando la corretto specificata likelihood, condizionata ai parametri stimati nella prima fase di likelihood:

$$\begin{aligned}
 QL_2(\Psi|\hat{\vartheta}, r_t) &= -\frac{1}{2} \sum_{t=1}^T (k \log(2\pi) + 2\log|D_t| + \log(|R_t|) + r' D_t^{-1} R_t^{-1} D_t^{-1} r_t) \\
 &= -\frac{1}{2} \sum_{t=1}^T (k \log(2\pi) + 2\log(|D_t|) + \log(|R_t|) + \epsilon_t' R_t^{-1} \epsilon_t)
 \end{aligned}$$

Poiché siamo in un condizionamento di  $\hat{\vartheta}$ , l'unica parte della log-likelihood che influenzerà la selezione dei parametri è  $\log(|R_t|) + \epsilon_t' R_t^{-1} \epsilon_t$ , e nella stima dei parametri DCC, è spesso più facile escludere i termini costanti e massimizzare semplicemente:

$$QL_2(\Psi|\hat{\vartheta}, r_t) = -\frac{1}{2} \sum_{t=1}^T (\log(|R_t|) + \epsilon_t' R_t^{-1} \epsilon_t)$$

# CAPITOLO 4

## Applicazione empirica

Dopo aver definito, nei capitoli precedenti, i modelli VECM e DCC ed aver impostato il problema di copertura, in questa capitolo finale presenteremo l'applicazione empirica sviluppata con i software Eviews e MatLab.

Nel seguito introdurremo le serie storiche utilizzate e ne analizzeremo le principali proprietà statistiche per modellarle con la procedura VECM e stimare la volatilità dei residui attraverso GARCH-Multivariati con diverse procedure (DCC, CCC, BEKK). Attraverso le stime di volatilità calcoleremo i rapporti di copertura del crack spread per ogni singolo modello, così da verificare empiricamente la superiorità del modello DCC nelle decisioni di copertura ottimale.

### 4.1. Dataset

#### 4.1.1. Descrizione delle serie storiche

Il dataset si sostanzia di sei serie storiche: Light sweet Crude Oil WTI spot, Heating Oil Spot, RBOB Gasoline Spot e relativi contratti Futures con medesimo sottostante; tutti i contratti Futures e relativi prezzi Spot sono quotati al New York Mercantile Exchange (NYMEX).

I dati scaricati hanno periodicità giornaliera e un range temporale che va dal 03/05/2005 al 10/02/2012 con un campione pari a circa 1660 osservazioni (giorni di borsa aperta); tutte le serie storiche sono state estrapolate da Bloomberg.

Di seguito in figura 2 e 3 vengono presentati gli andamenti dei contratti futures e dei relativi Spot, così da poter evidenziare come sullo storico gli andamenti tra i tre futures e i tre Spot siano molto simili tra loro. Naturalmente i prezzi dell'Heating Oil e del Gasoline, essendo quotati in cent per gallone, sono stati convertiti in barili (cioè moltiplicati per 42) così da poterli confrontare con il Crude Oil (quotato in \$ al barile).

I grafici essendo stati ricavati con Bloomberg non presentano il nome del contratto ma il codice d'identificazione: CL1 Comdty (Crude Oil futures), HO1 Comdty (Heating Oil futures), XB1 Comdty (Gasoline futures), USCRWTIC Index (Crude Oil spot), NO2INYPR Index (Heating Oil spot), RBOB87PM Index (Gasoline spot).

**Figura 2: Andamento prezzi Futures**



**Figura 3: Andamento prezzi Spot**



Per verificare che l'andamento del derivato fosse in linea con quello del sottostante, è utile anche proiettare gli andamenti degli spread tra futures e relativi spot.

Figura 4: Spread Crude Oil

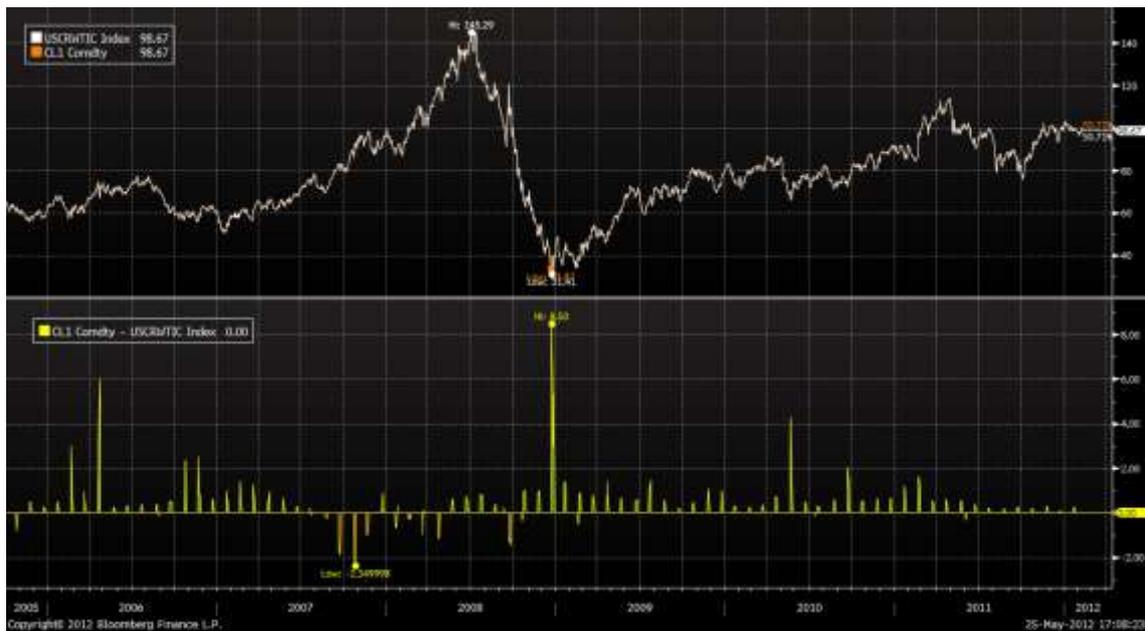


Figura 5: Spread Heating Oil



**Figura 6: Spread Gasoline**



Infine il grafico 7, che evidenzia gli andamenti molto simili, nel range temporale, di tutti e sei i contratti del settore energy.

**Figura 7: Andamento Futures&Spot**



#### 4.1.2. Analisi delle proprietà statistiche dei dati

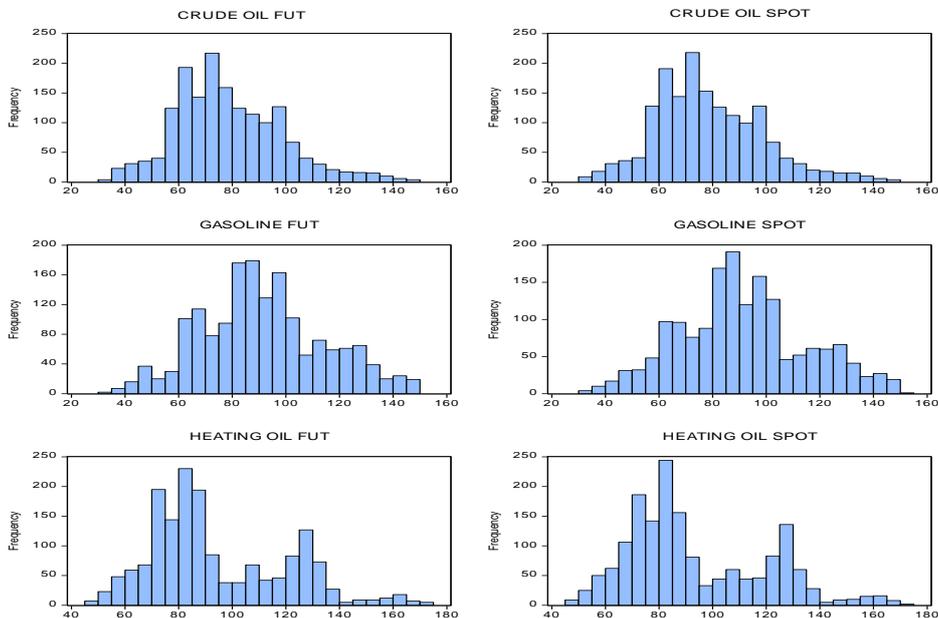
L'analisi della tabella 2, che riporta le principali statistiche descrittive sulle serie storiche, evidenzia la similitudine tra i valori di statistica tra i singoli contratti futures e relativi sottostanti, come già emerso in precedenza con grafici degli andamenti.

**Tabella 2: Statistiche descrittive**

	CF	CS	GF	GS	HF	HS
<b>Mean</b>	78.77270482	78.68608434	91.25343	91.11625	94.40389	93.68267241
<b>Median</b>	75.465	75.43	89.2269	89.0274	85.932	85.1235
<b>Maximum</b>	145.29	145.29	149.982	150.801	172.452	171.717
<b>Minimum</b>	33.87	31.41	33.2934	31.7184	47.5902	47.0652
<b>Std. Dev.</b>	20.25551334	20.32446323	23.16509	23.78458	25.59	25.88117836
<b>Skewness</b>	0.59841783	0.584712155	0.235956	0.207059	0.709524	0.707985009
<b>Kurtosis</b>	3.41810778	3.411243498	2.695005	2.675397	2.772542	2.715257447
<b>Jarque-Bera</b>	111.1667217	106.2866482	21.83753	19.14954	142.859	144.2850843
<b>Probability</b>	0	0	1.81E-05	6.95E-05	0	0
<b>Sum</b>	130762.69	130618.9	151480.7	151253	156710.5	155513.2362
<b>Sum Sq. Dev.</b>	680664.1766	685306.0337	890254.7	938506.5	1086393	1111256.917
<b>Observations</b>	1660	1660	1660	1660	1660	1660

Come si nota, dal test di Jarque-Bera e dalla relativa probabilità, nessuna distribuzione di frequenza sembra assomigliare ad una normale; lo si può anche vedere graficamente dai grafici seguenti.

**Figura 8: Distribuzioni di frequenza**



Verifichiamo, preliminarmente, la stazionarietà delle serie storiche: per poter sviluppare modelli ECM con il metodo di Engle e Granger, nel nostro caso modelli VECM, abbiamo bisogno di serie storiche stazionarie nelle differenze prime (cioè serie integrate di ordine uno I(1)) o al massimo nelle differenze seconde (I(2)) e non stazionarie nei livelli (I(0)).

Come vedremo dai test ADF seguenti, tutte e sei le nostre variabili sono I(1) quindi stazionarie nelle differenze prime, possiamo quindi sviluppare un modello VECM.

Per motivi di spazio riportiamo, in tabella 3, 4 e 5, solamente i test ADF sulle variabili futures, in quanto i test sui relativi sottostanti sono molto simili.

**Tabella 3: Test ADF su Futures crude oil**

Null Hypothesis: D(CRUDE OIL WTI FUT) has a unit root  
 Exogenous: None  
 Lag Length: 0 (Automatic - based on SIC, maxlag=24)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-42.38145	0.0001
Test critical values:		
1% level	-2.566357	
5% level	-1.941015	
10% level	-1.616571	

\*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(CRUD\_OIL\_WTI\_FUT,2)  
 Method: Least Squares  
 Date: 05/03/12 Time: 16:35  
 Sample (adjusted): 10/05/2005 2/10/2012  
 Included observations: 1658 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(CRUD_OIL_WTI_FUT(-1))	-1.040215	0.024544	-42.38145	0.0000
R-squared	0.520153	Mean dependent var		0.000241
Adjusted R-squared	0.520153	S.D. dependent var		2.738873
S.E. of regression	1.897243	Akaike info criterion		4.119284
Sum squared resid	5964.426	Schwarz criterion		4.122549
Log likelihood	-3413.886	Hannan-Quinn criter.		4.120494
Durbin-Watson stat	2.004990			

**Tabella 4: Test ADF su Futures Gasoline**

Null Hypothesis: D(GASOLINE\_FUT) has a unit root  
 Exogenous: None  
 Lag Length: 0 (Automatic - based on SIC, maxlag=24)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-41.39867	0.0000
Test critical values:		
1% level	-2.566357	
5% level	-1.941015	
10% level	-1.616571	

\*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(GASOLINE\_FUT,2)  
 Method: Least Squares  
 Date: 05/03/12 Time: 16:42  
 Sample (adjusted): 10/05/2005 2/10/2012  
 Included observations: 1658 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(GASOLINE_FUT(-1))	-1.016815	0.024562	-41.39867	0.0000
R-squared	0.508433	Mean dependent var		0.000142
Adjusted R-squared	0.508433	S.D. dependent var		3.128550
S.E. of regression	2.193485	Akaike info criterion		4.409463
Sum squared resid	7972.451	Schwarz criterion		4.412728
Log likelihood	-3654.445	Hannan-Quinn criter.		4.410673
Durbin-Watson stat	2.001863			

**Tabella 5: Test ADF su Futures Heating oil**

Null Hypothesis: D(HEATING OIL FUT) has a unit root  
 Exogenous: None  
 Lag Length: 0 (Automatic - based on SIC, maxlag=24)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-41.07628	0.0000
Test critical values:		
1% level	-2.566357	
5% level	-1.941015	
10% level	-1.616571	

\*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(HEATING\_OIL\_FUT,2)  
 Method: Least Squares  
 Date: 05/03/12 Time: 16:46  
 Sample (adjusted): 10/05/2005 2/10/2012  
 Included observations: 1658 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(HEATING_OIL_FUT(-1))	-1.009011	0.024564	-41.07628	0.0000
R-squared	0.504524	Mean dependent var		0.000122
Adjusted R-squared	0.504524	S.D. dependent var		2.816384
S.E. of regression	1.982454	Akaike info criterion		4.207151
Sum squared resid	6512.213	Schwarz criterion		4.210416
Log likelihood	-3486.728	Hannan-Quinn criter.		4.208361
Durbin-Watson stat	2.000611			

Infine l'ultima analisi svolta sulle serie storiche si sostanzia in dei test di Granger causality, con una rolling window dinamica di 250 osservazioni; lo scopo è quello di mettere in relazione i tre contratti futures con i relativi sottostanti, così da poter rilevare se talvolta sia il futures che causi lo spot, viceversa o nessuno causi l'altro.

Il test C1 è dato dall'ipotesi nulla ( $H_0$ ) che futures non causi spot, mentre  $H_1$  è futures causa spot; quindi se p-value < 15% allora rifiuto l'ipotesi nulla e quindi futures causa spot e viceversa se p-value > 15%.

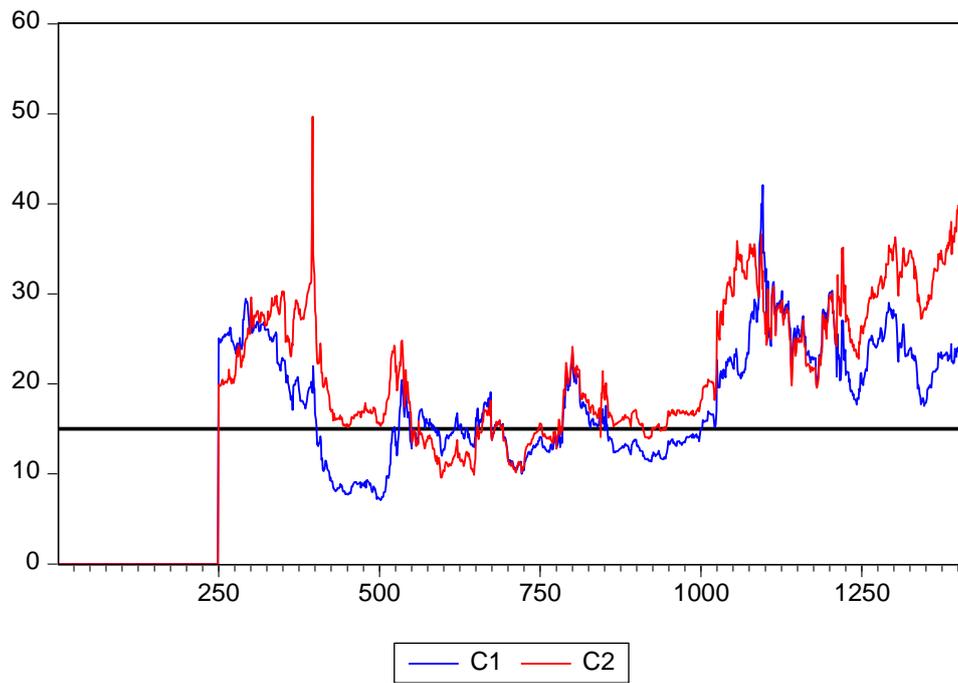
Il test C2 è il contrario del precedente: l'ipotesi nulla è spot causa futures, quindi se p-value < 15% allora Spot causa futures e viceversa.

Come si può notare nella figura 13 per il Crude Oil abbiamo che tra la 375-esima e 500-esima osservazione futures causa spot così come tra 850 e 1000, mentre Spot causa futures tra 600 e 750.

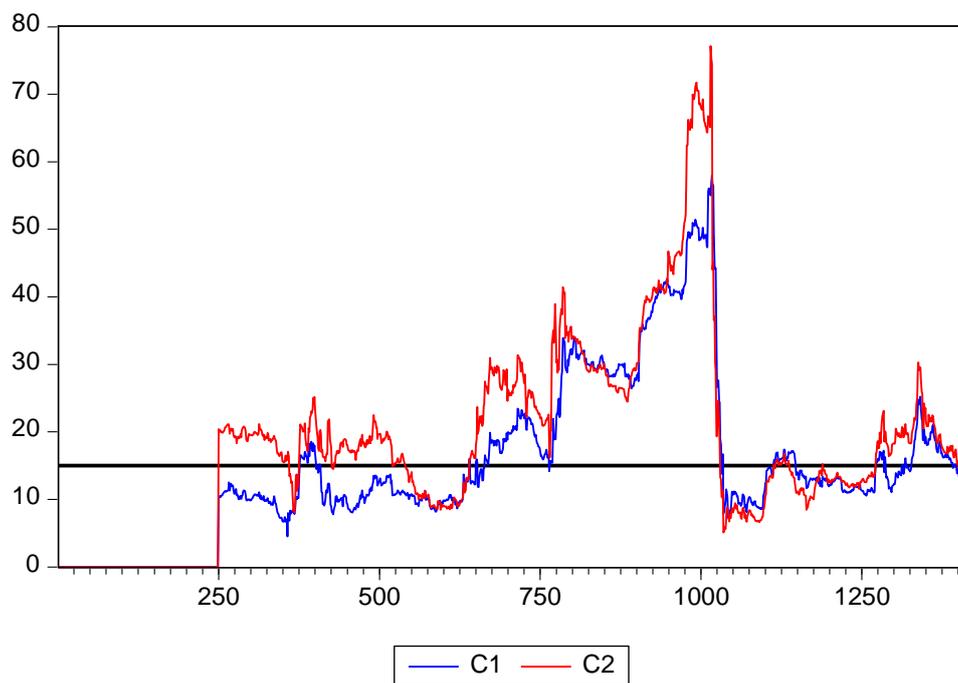
Per quanto riguarda il Gasoline (figura 14) tra 250 e 600 sembra essere il futures a causare lo spot, mentre da 1050 a 1250 si concausano.

Infine Heating Oil (figura 15) sembrano concausarsi insieme se non per il primo periodo tra 250-esima e 250-esima osservazione in lo spot causa il futures.

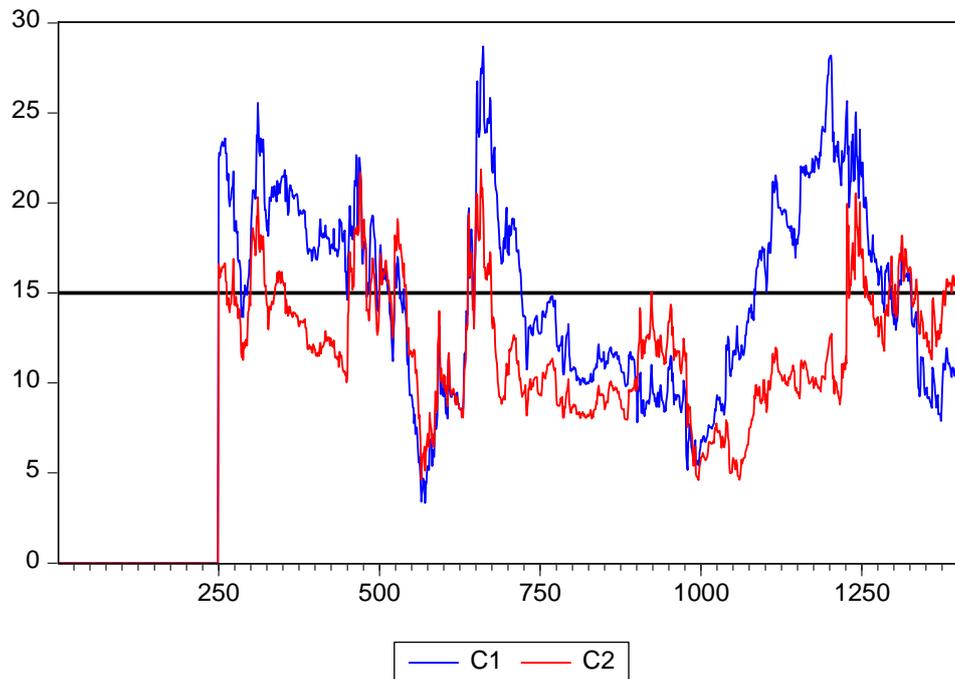
**Figura 9: Granger Causality Crude Oil**



**Figura 10: Granger Causality Gasoline**



**Figura 11: Granger Causality Heating Oil**



## **4.2. Implementazione VECM**

### **4.2.1. Struttura dei ritardi e test di cointegrazione**

Il primo passo per poter impostare la nostra stima VECM è stato quello di determinare i lags delle variabili e il grado di cointegrazione: l'analisi preliminare ha indicato che un ritardo di 15 periodi (nel nostro caso giorni) fosse sufficiente a rendere tutti i residui white noise (utilizzando Ljung-Box Q statistics). Inoltre, abbiamo verificato, avvalendoci del Johansen cointegration test, che il numero di relazioni di cointegrazioni è di almeno 3, assumendo un livello di significatività del 5% (si veda tab. 6).

## Tabella 6: Cointegration test

Sample (adjusted): 10/25/2005 2/10/2012  
 Included observations: 1644 after adjustments  
 Trend assumption: Linear deterministic trend  
 Series: CRUD\_OIL\_WTI\_FUT GASOLINE\_FUT HEATING\_OIL\_FUT  
 CRUDE\_OIL\_SPOT GASOLINE\_SPOT HEATING\_OIL\_SPOT  
 Lags interval (in first differences): 1 to 15

### Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.073821	231.9438	95.75366	0.0000
At most 1 *	0.029963	105.8691	69.81889	0.0000
At most 2 *	0.016089	55.85752	47.85613	0.0074
At most 3	0.012685	29.19156	29.79707	0.0586
At most 4	0.004041	8.204651	15.49471	0.4438
At most 5	0.000941	1.548495	3.841466	0.2134

## 4.2.2. Stima VECM

Abbiamo quindi stimato il nostro modello VECM, con 15 ritardi e 3 gradi di cointegrazione, ottenendo per la parte di cointegrazione le stime seguenti:

## Tabella 7: Stima relazioni di cointegrazione VECM

### Vector Error Correction Estimates

Sample (adjusted): 10/25/2005 2/10/2012  
 Included observations: 1644 after adjustments  
 Standard errors in ( ) & t-statistics in [ ]

Cointegrating Eq:	CointEq1	CointEq2	CointEq3
CRUD_OIL_WTI_FUT(-1)	1.000000	0.000000	0.000000
GASOLINE_FUT(-1)	0.000000	1.000000	0.000000
HEATING_OIL_FUT(-1)	0.000000	0.000000	1.000000
CRUDE_OIL_SPOT(-1)	-0.992543 (0.00307) [-323.303]	-0.082938 <b>(0.06062)</b> [-1.36814]	-0.026105 (0.02888) [-0.90391]
GASOLINE_SPOT(-1)	-0.000135 (0.00201) [-0.06721]	-1.006250 (0.03971) [-25.3419]	-0.051374 (0.01892) [-2.71583]

HEATING_OIL_SPOT(-1)	-0.003748 (0.00267) [-1.40172]	0.069195 <b>(0.05280)</b> [ 1.31044]	-0.930660 (0.02516) [-36.9961]
C	-0.311763	0.485186	-0.468137

Le tre relazioni di cointegrazioni sono definite in questo modo:

1. Coefficiente futures sul Crude oil pari a 1, gli altri due futures 0 e determino quelli spot;
2. Coefficiente futures sul Gasoline uguale a 1, gli altri due futures 0 e determino quelli spot;
3. Coefficiente futures sul Heating Oil uguale a 1, gli altri due futures 0 e determino quelli spot.

Quello che è interessante notare, di queste prime stime, è il fatto che nella seconda relazione di cointegrazione sia Heating oil spot che Crude oil spot sembrano essere non significativi (probabilità evidenziata in giallo); questo significa che non inficiano sul futures gasoline e che questo viene spiegato bene esclusivamente dal suo sottostante. Nelle altre due relazioni, invece, tutte le variabili Spot sembrano significative, cogliendo al meglio le relazioni intercorrenti tra contratti futures e Spot.

Per quanto riguarda gli error correction term, solo la seconda relazione di cointegrazione sembra risultare significativa per tutte e sei le variabili differenziate, mentre le altre due sembrano essere non significative con tutte le variabili, come si può notare dai risultati di stima seguenti. Per motivi di spazio abbiamo evitato di introdurre le stime di tutti i coefficienti delle variabili ritardate.

**Tabella 8: Stima error correction VECM**

Error Correction:	D(CF)	D(GF)	D(HF)	D(CS)	D(GS)	D(HS)
CointEq1	-0.439519 (0.26336) [-1.66886]	0.276504 (0.30746) [ 0.89932]	0.027581 (0.27796) [ 0.09922]	0.143310 (0.26307) [ 0.54475]	0.278918 (0.32300) [ 0.86353]	0.041245 (0.28012) [ 0.14724]
CointEq2	0.037952 <b>(0.02085)</b> [ 1.82028]	0.084278 <b>(0.02434)</b> [ 3.46252]	0.031842 <b>(0.02201)</b> [ 1.44700]	0.033252 <b>(0.02083)</b> [ 1.59665]	0.134605 <b>(0.02557)</b> [ 5.26413]	0.036241 <b>(0.02218)</b> [ 1.63429]
CointEq3	-0.121137	-0.136770	-0.178048	-0.140617	-0.189494	-0.132203

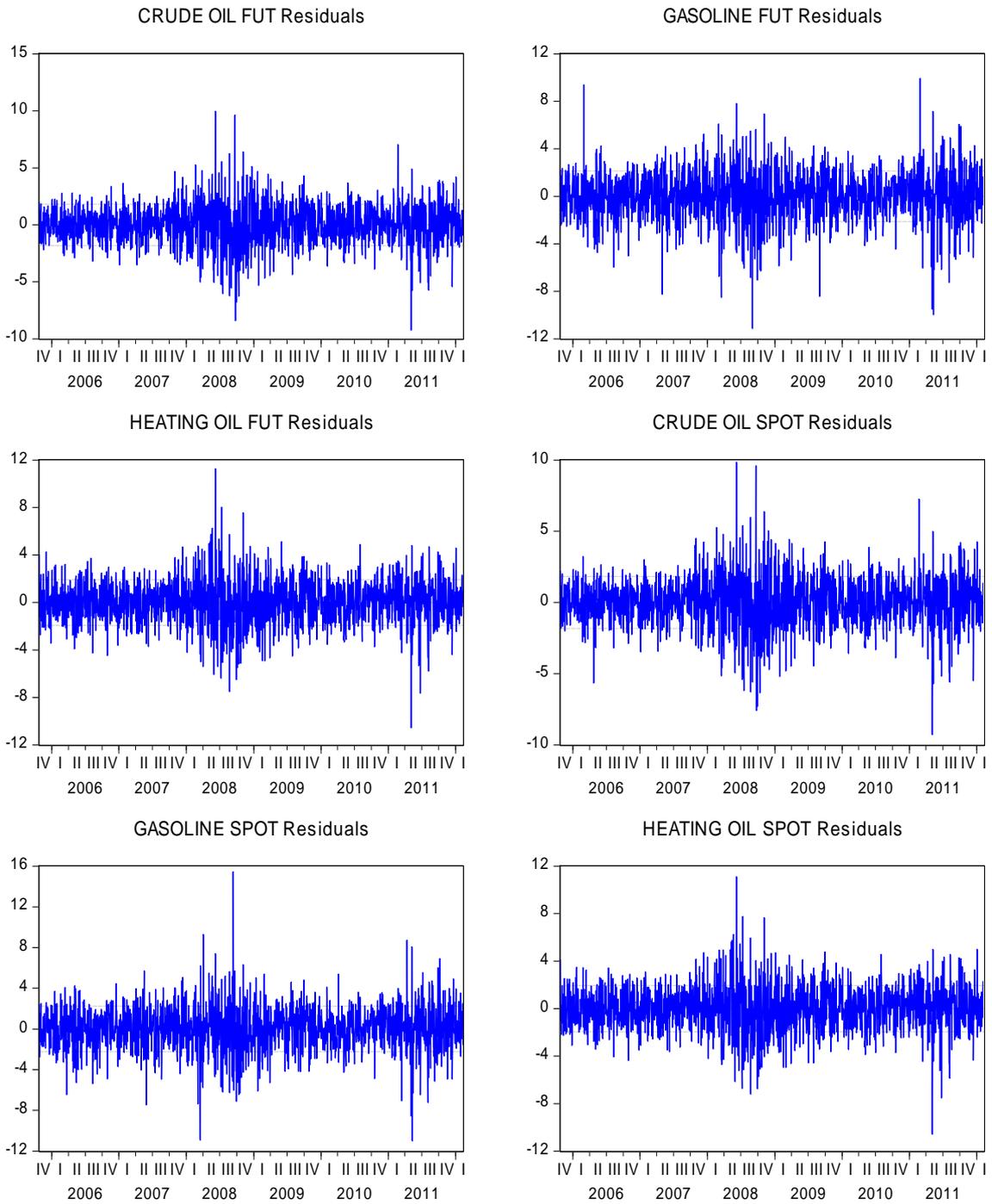
	(0.05933)	(0.06926)	(0.06262)	(0.05927)	(0.07277)	(0.06310)
	[-2.04171]	[-1.97460]	[-2.84331]	[-2.37267]	[-2.60418]	[-2.09497]
R-squared	0.131999	0.110426	0.110149	0.135731	0.125941	0.098053
Adj. R-squared	0.079919	0.057052	0.056758	0.083875	0.073498	0.043937
Sum sq. resids	5171.493	7048.179	5760.763	5160.063	7778.583	5850.337
S.E. equation	1.826594	2.132419	1.927854	1.824575	2.240188	1.942784
F-statistic	2.534551	2.068898	2.063064	2.617446	2.401468	1.811883
Log likelihood	-3274.771	-3529.269	-3363.472	-3272.952	-3610.322	-3376.155
Akaike AIC	4.098261	4.407869	4.206170	4.096049	4.506474	4.221599
Schwarz SC	4.407300	4.716908	4.515209	4.405087	4.815513	4.530638
Mean dependent	0.023327	0.034990	0.035378	0.022841	0.037575	0.036655
S.D. dependent	1.904273	2.195982	1.985010	1.906269	2.327347	1.986924
Determinant resid covariance (dof adj.)		0.072521				
Determinant resid covariance		0.050938				
Log likelihood		-11549.20				
Akaike information criterion		14.75815				
Schwarz criterion		16.67156				

#### 4.2.3. Analisi dei residui

Ultimo passo della stima VECM è verificare la stabilità del modello con l'analisi dei residui (vedi figura 12); per fare questo ci avvaleremo del test ADF, correlogramma e test di eteroschedasticità.

Si può notare, che non vi siano indizi a favore della presenza di radici unitarie. L'applicazione del test ADF (vedi tabella 9) su tutte le sei serie di residui suggerisce la stazionarietà sui livelli e quindi le variabili sembrano essere  $I(0)$ . Generalmente si suggerisce molta cautela nel decidere a favore dell'assenza di radici unitarie per non incorrere in una regressione spuria. In questo caso si è decisamente in zona di rifiuto dell'ipotesi (di presenza di radici unitarie).

**Figura 12: Residui VECM**



### Tabella 9: test ADF sui residui VECM

Group unit root test: Summary  
 Series: RESID01, RESID02, RESID03, RESID04, RESID05, RESID06

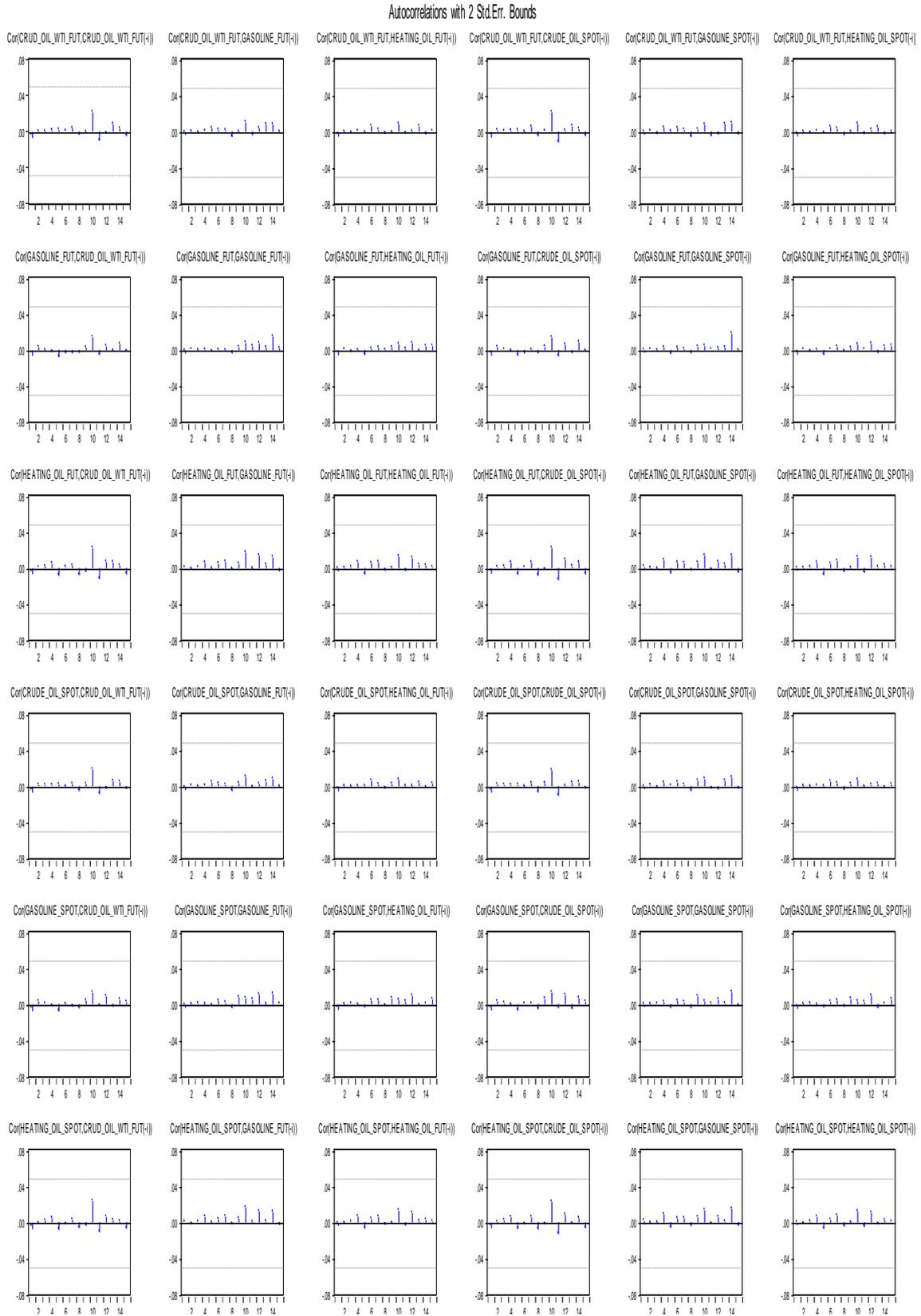
Sample: 10/03/2005 2/10/2012  
 Exogenous variables: None  
 Automatic selection of maximum lags  
 Automatic lag length selection based on SIC: 0  
 Newey-West automatic bandwidth selection and Bartlett kernel  
 Balanced observations for each test

Method	Statistic	Prob.**	Cross-sections	Obs
<u>Null: Unit root (assumes common unit root process)</u>				
Levin, Lin & Chu t*	-99.5734	0.0000	6	9858
<u>Null: Unit root (assumes individual unit root process)</u>				
ADF - Fisher Chi-square	197.084	0.0000	6	9858
PP - Fisher Chi-square	197.086	0.0000	6	9858

Per avvalorare l'ipotesi di residui conformi ad una realizzazione di un processo white noise, si mettono in evidenza, nel seguito, il correlogramma dei residui e il test di eteroschedasticità. Entrambi costituiscono un'analisi utile per fornire indicazioni sulla presenza di autocorrelazione e sull'eteroschedasticità.

Nel presente studio non vi sono indicazione di presenza di autocorrelazione ed eteroschedasticità, come si può notare dalla figura 13 e 14 e dalla tabella 10.

**Figura 13: Correlogramma dei residui**



## Tabella 10: Test eteroschedasticità sui residui

VEC Residual Heteroskedasticity Tests: No Cross Terms (only levels and squares)

Date: 05/07/12 Time: 07:06

Sample: 10/03/2005 2/10/2012

Included observations: 1644

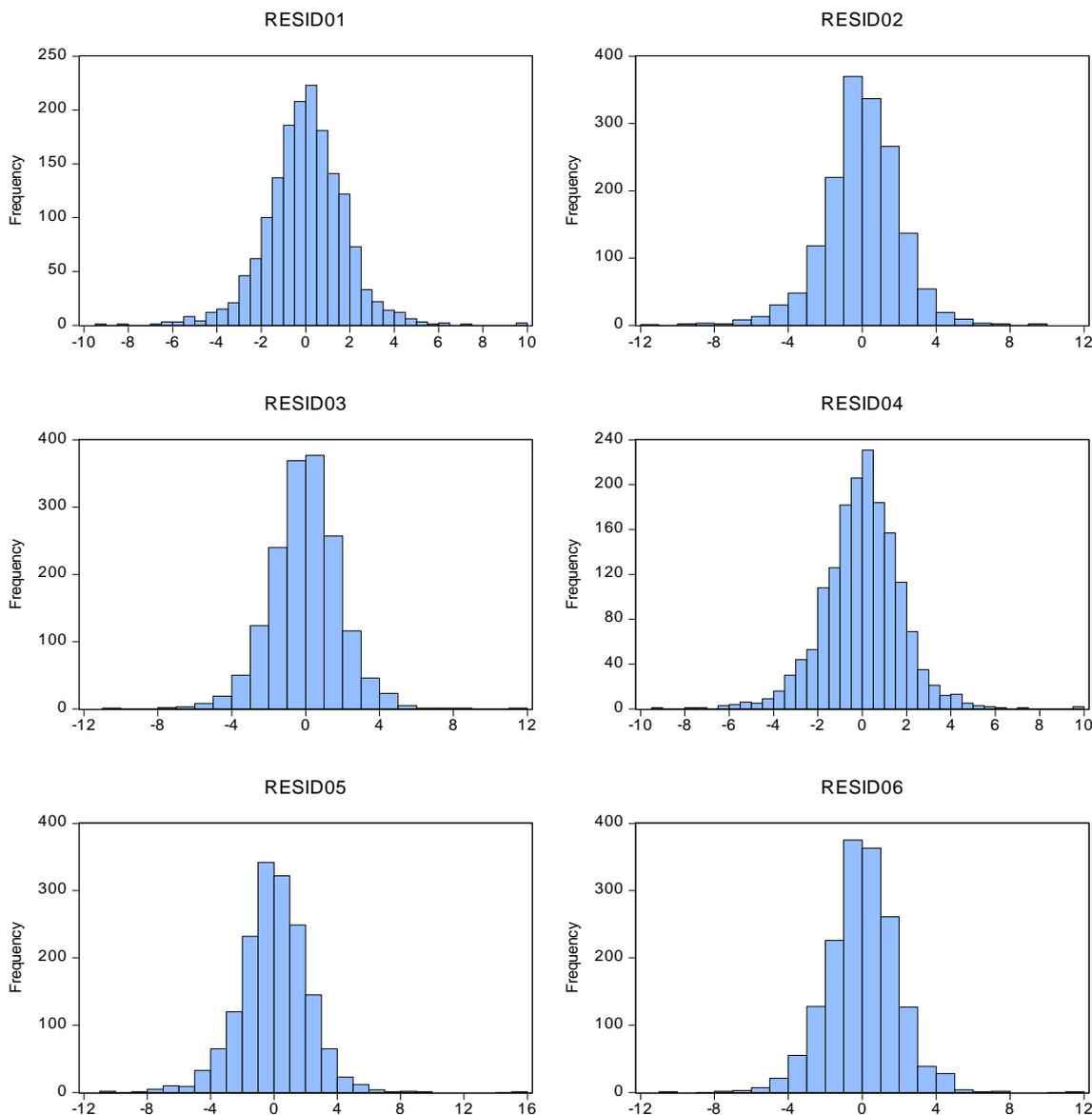
Joint test:					
Chi-sq	df	Prob.			
7841.071	3906	0.3489			

Individual components:					
Dependent	R-squared	F(186,1457)	Prob.	Chi-sq(186)	Prob.
res1*res1	0.349566	4.209903	0.0000	574.6861	0.0000
res2*res2	0.214050	2.133379	0.0000	351.8989	0.0000
res3*res3	0.212708	2.116381	0.0000	349.6915	0.0000
res4*res4	0.341615	4.064464	0.0000	561.6147	0.0000
res5*res5	0.334039	3.929111	0.0000	549.1595	0.0000
res6*res6	0.211974	2.107113	0.0000	348.4847	0.0000
res2*res1	0.248178	2.585797	0.0000	408.0044	0.0000
res3*res1	0.238000	2.446630	0.0000	391.2718	0.0000
res3*res2	0.221812	2.232792	0.0000	364.6597	0.0000
res4*res1	0.347855	4.178304	0.0000	571.8731	0.0000
res4*res2	0.254134	2.668993	0.0000	417.7955	0.0000
res4*res3	0.243360	2.519461	0.0000	400.0846	0.0000
res5*res1	0.257990	2.723573	0.0000	424.1350	0.0000
res5*res2	0.263462	2.802008	0.0000	433.1315	0.0000
res5*res3	0.226699	2.296401	0.0000	372.6932	0.0000
res5*res4	0.264957	2.823639	0.0000	435.5893	0.0000
res6*res1	0.240058	2.474467	0.0000	394.6549	0.0000
res6*res2	0.221926	2.234265	0.0000	364.8469	0.0000
res6*res3	0.212830	2.117922	0.0000	349.8919	0.0000
res6*res4	0.246627	2.564346	0.0000	405.4545	0.0000
res6*res5	0.228129	2.315173	0.0000	375.0448	0.0000

Il test di Jarque-Bera indica chiaramente il rifiuto dell'ipotesi di normalità. Tuttavia, questo è dovuto a certi residui anomali sulle code delle distribuzioni di frequenza.

**Figura 14: Distribuzioni di frequenza sui residui**



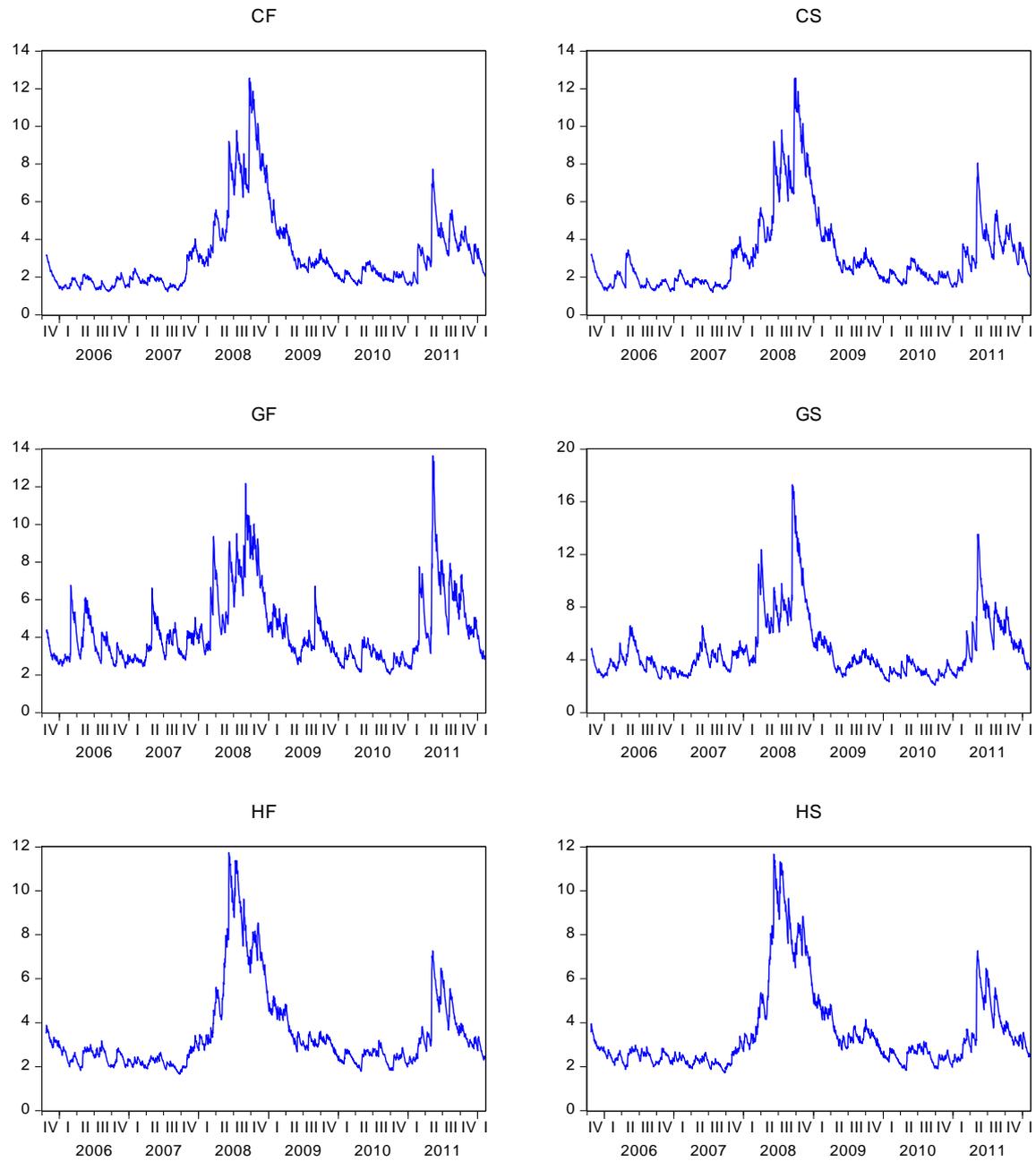
### 4.3. Stima della volatilità

Il punto successivo è stato quello di stimare la volatilità con svariati modelli GARCH-Multivariati (CCC, DCC, BEKK,) presentati nel capitolo 3, in modo tale che l'analisi di confronto possa evidenziare con quale di questi si ottenga la migliore copertura. La rappresentazione utilizzata è quella di un GARCH(1,1).

Nel seguito, quindi, si presenteranno stime di matrici di varianze e covarianze dinamiche nel tempo e, in base al modello, correlazioni invariante o variante rispetto al tempo.

Come vedremo nel seguito le stime di volatilità sono uguali sia che si usi un sistema CCC che un DCC o BEKK, quello che cambia è la stima della correlazione che modificherà i risultati dell'hedging.

**Figura 15: Stime di volatilità**



Nella volatilità (vedi figura 15) sono presenti, in tutti i contratti, due principali outliers: il primo tra il 3° e 4° trimestre 2008 e il secondo a metà 2011.

Questi “picchi” di volatilità sono dovuti nel 2008 alla crisi dei mutui subprime e al conseguente fallimento della banca d'affari Lehman Brothers e nel 2011 alla crisi del debito sovrano Europeo, in particolare del quasi default della Grecia.

Questi eventi hanno causato una forte esplosione della volatilità sui mercati finanziari, dovuta al cosiddetto “panic selling”.

#### 4.3.1. Stime CCC

Come è logico che sia, le correlazioni tra questi tipi di contratti (tutti appartenenti al settore Energy) sono molto alte, tutte maggiori di 0.70, e addirittura quelle tra i futures e relativi spot superiori a 0.90; questa alta correlazione la si vede anche solamente graficamente dall'andamento molto simile dei contratti, visto in precedenza.

Come vedremo, nel seguito con i risultati dell'hedging, il fatto che non vengano stimate correlazioni dinamiche nel tempo, ma costanti, non tenendo quindi conto che in momenti della storia ci siano diminuzioni delle relazioni intercorrenti tra variabili, produrrà esiti che andranno ad inficiare negativamente sui risultati della copertura.

**Tabella 11: Matrice di Correlazione CCC**

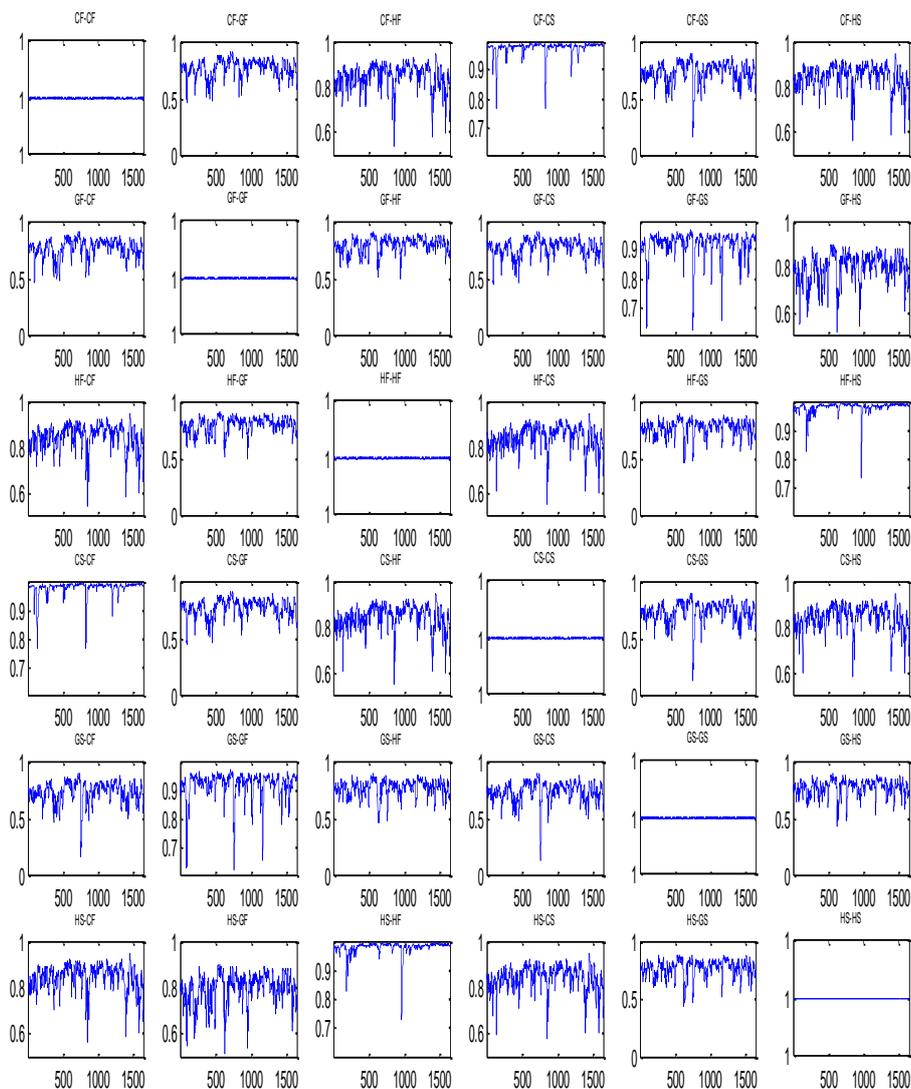
	CF	GF	HF	CS	GS	HS
CF	1	0.7625	0.8435	0.9773	0.7243	0.8444
GF	0.7625	1	0.7915	0.7641	0.912	0.7923
HF	0.8435	0.7915	1	0.8406	0.7622	0.9788
CS	0.9773	0.7641	0.8406	1	0.7259	0.8413
GS	0.7243	0.912	0.7622	0.7259	1	0.7625
HS	0.8444	0.7923	0.9788	0.8413	0.7625	1

#### 4.3.2. Stime DCC

Il punto di forza del modello è proprio questo: non dare per scontato che la correlazione sia costante nel tempo ma che vari in ogni periodo temporale; questo fatto, come vedremo nel seguito, gioverà sulla performance dell'hedging.

Come si nota dalla matrice di grafici di Figura 16 le correlazioni stimate, tra Futures e Spot, sembrano non essere sempre costanti, ma si possono notare dei momenti in cui la correlazione si riduce di molto, addirittura arrivando a dimezzarsi, soprattutto in momenti di mercato caratterizzati da alta volatilità, come negli ultimi mesi del 2008.

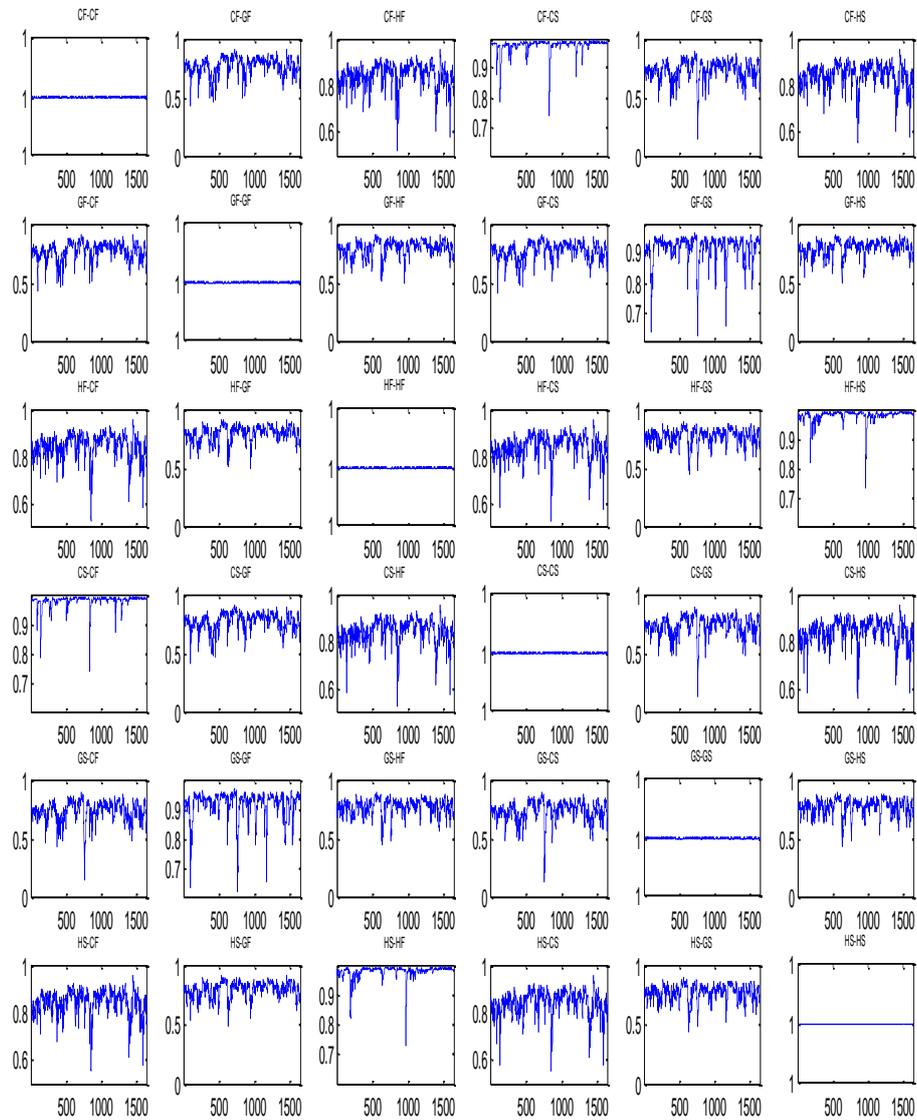
**Figura 16: Matrice di correlazioni dinamiche DCC**



### 4.3.2. Stime BEKK

Un altro modello a correlazione dinamica alternativo al DCC è il sistema BEKK. Con lo scopo di mettere a confronto i DCC, sono state eseguite delle stime con un diagonal BEKK, in modo da vedere quale dei due a parità di correlazione dinamica ci permette di ottenere il migliore hedging.

**Figura 17: Matrice di correlazioni dinamiche BEKK**



Se mettiamo a confronto questa matrice di correlazioni dinamiche (figura 17) con quella ottenuta con il sistema DCC si possono notare delle differenze, seppur minime, sui valori stimati; saranno proprio questi scarti che andranno ad inficiare sulla copertura, facendoci ottenere, come vedremo nel seguito, migliori risultati con il modello DCC.

#### 4.4. Copertura del crack spread (hedging)

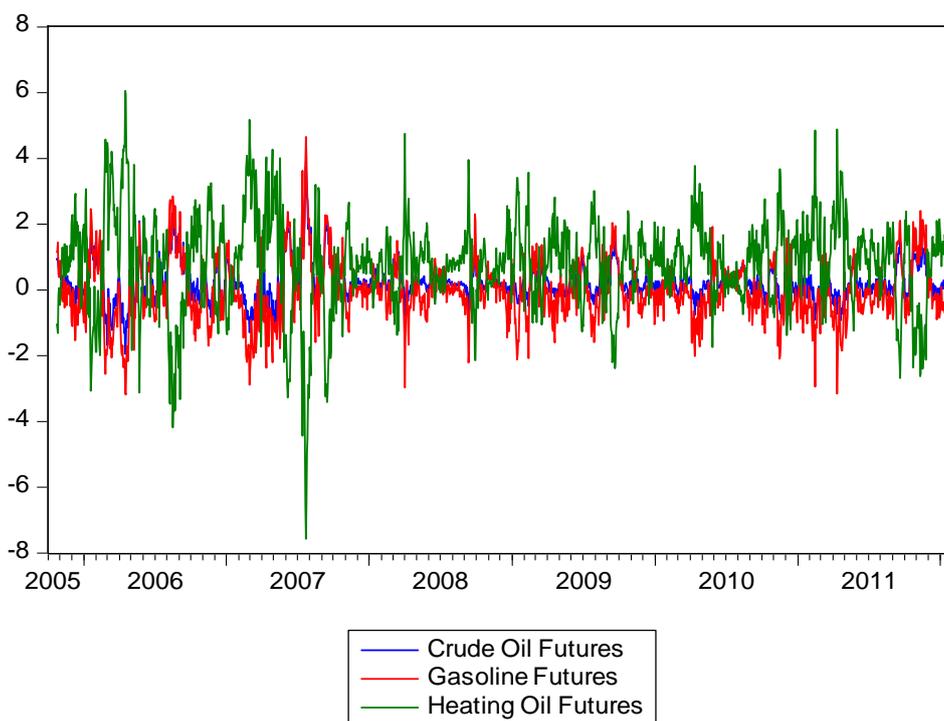
Una volta ottenute le stime di volatilità, il passo finale è stato quello di creare la copertura seguendo il modello di ottimizzazione descritto nel primo capitolo: ci siamo quindi calcolati i rapporti di copertura ottimali per ogni singolo modello con dei coefficienti lambda diversi (2, 4, 6,...100). L'obiettivo è stato quello di vedere con quale modello si ottenga la migliore copertura e con quale livello di rischio, in termini di performance media (CEI medio), applicando così alle singole quote una commissione di \$3 per le transazioni.

Infine un confronto finale è stato fatto creando altre due coperture: un full-hedge vendendo un Crude Oil futures e comprandone uno sul Gasoline e uno sul Heating Oil, e un un-hedge in cui non si entra sul mercato dei contratti futures.

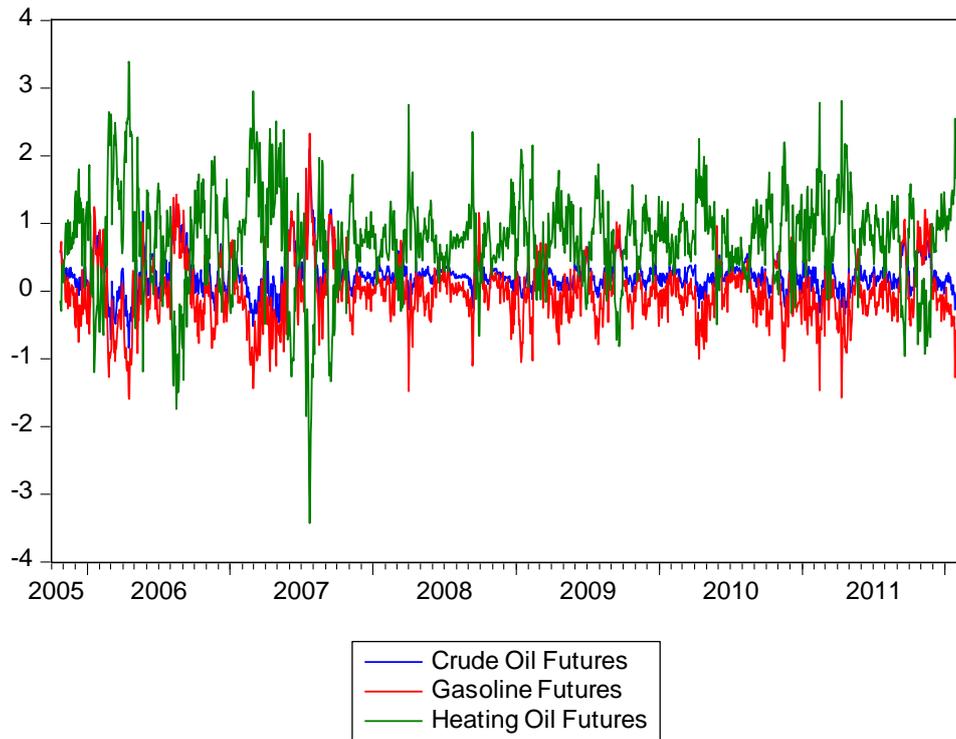
##### 4.4.1. Copertura con CCC

I rapporti di copertura, calcolati con il modello a correlazione costante, dimostrano che l'Heating oil è molto variabile, mentre si ottiene un buon hedging sul Crude oil, che risulta molto costante. La variabilità massima dei coefficienti di copertura si ottiene con grado di avversione al rischio pari a quattro ( $\lambda=4$ ).

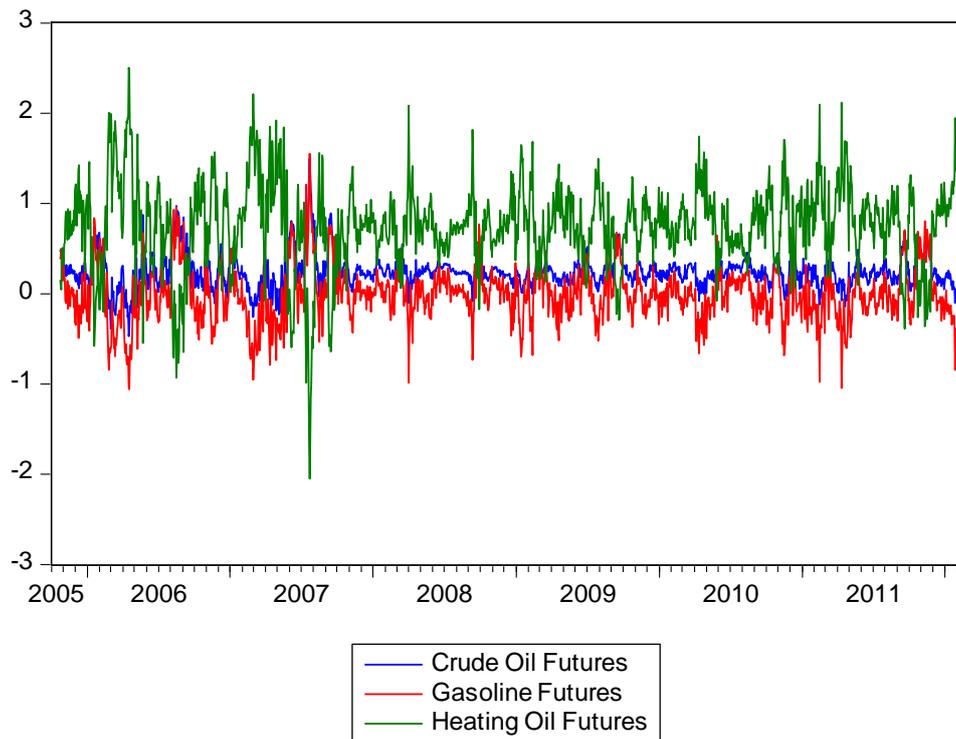
**Figura 18: Rapporti di copertura ottimi CCC ( $\lambda=2$ )**



**Figura 19: Rapporti di copertura ottimi CCC ( $\lambda=4$ )**

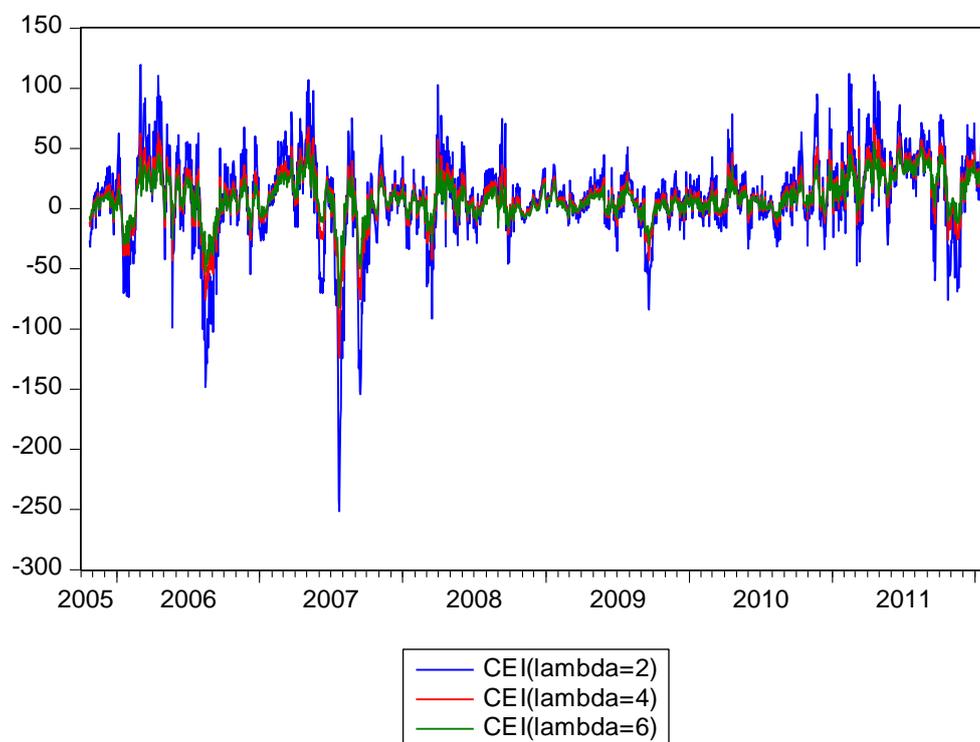


**Figura 20: Rapporti di copertura ottimi CCC ( $\lambda=6$ )**



In termini di performance la maggiore volatilità dei P&L si ottiene in corrispondenza di un lambda uguale a 2 ma in termini medi il punto in cui la performance media è maggiore si trova in corrispondenza di un grado di avversione al rischio pari a 4, come si può notare dalle figure 21 e 22.

**Figura 21: CEI (CCC)**



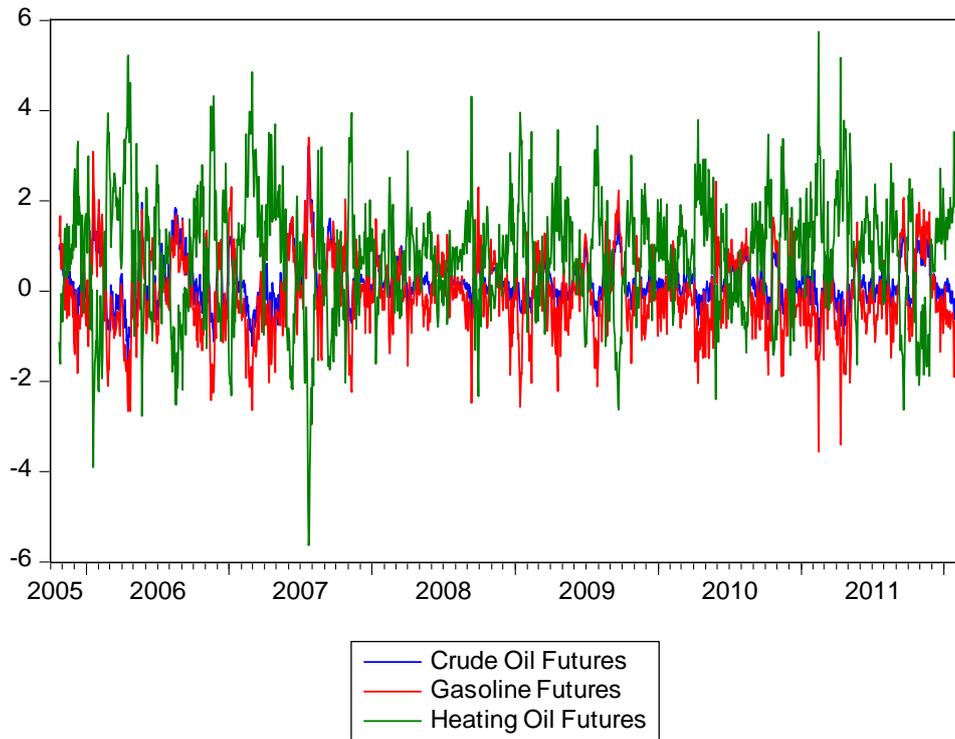
**Tabella 12: Confronto CCC**

	Crude Oil hedge ratio	Gasoline hedge ratio	Heating Oil hedge ratio	CEI Medio
$\lambda=2$	0.2465	-0.0338	0.7873	9.5367
$\lambda=4$	0.2492	-0.0107	0.7615	9.6582
$\lambda=6$	0.2501	-0.003	0.7529	8.6721

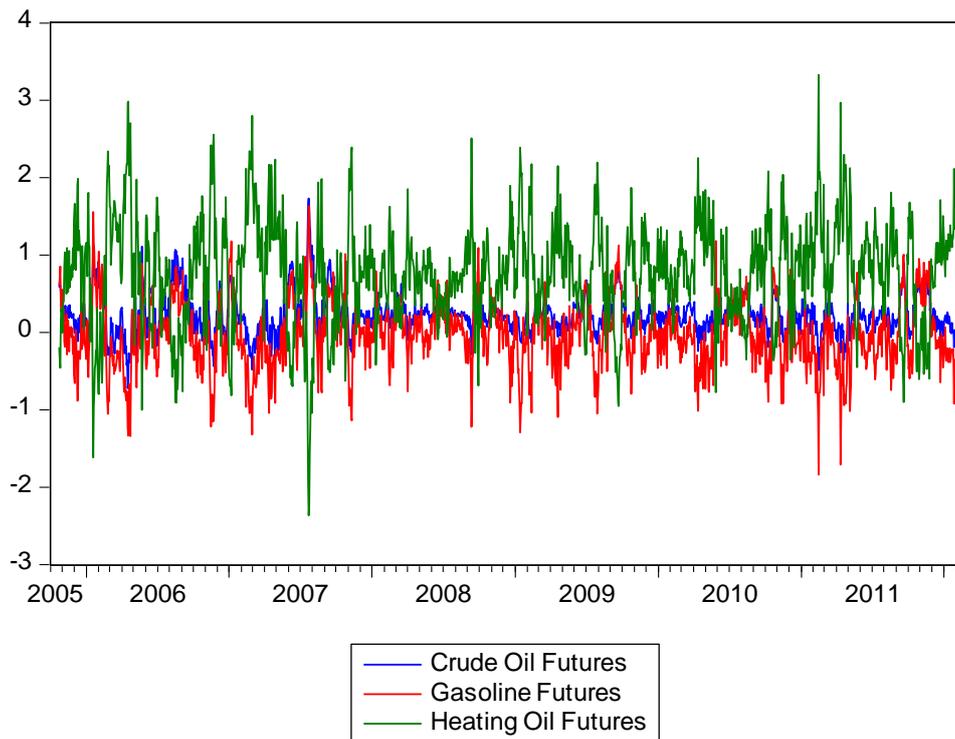
#### 4.4.2. Copertura con DCC

Dall'analisi dei grafici 22, 23 e 24, si evidenzia, anche in questo caso, come il rapporto di copertura dell'Heating Oil sia il più volatile dei tre; mentre la copertura sul Crude oil limita di molto le variazioni. Inoltre all'aumentare dell'avversione al rischio la variabilità dei rapporti di copertura diminuisce notevolmente; questo è dato dal fatto che, aumentando l'avversione al rischio, diminuisce la volatilità.

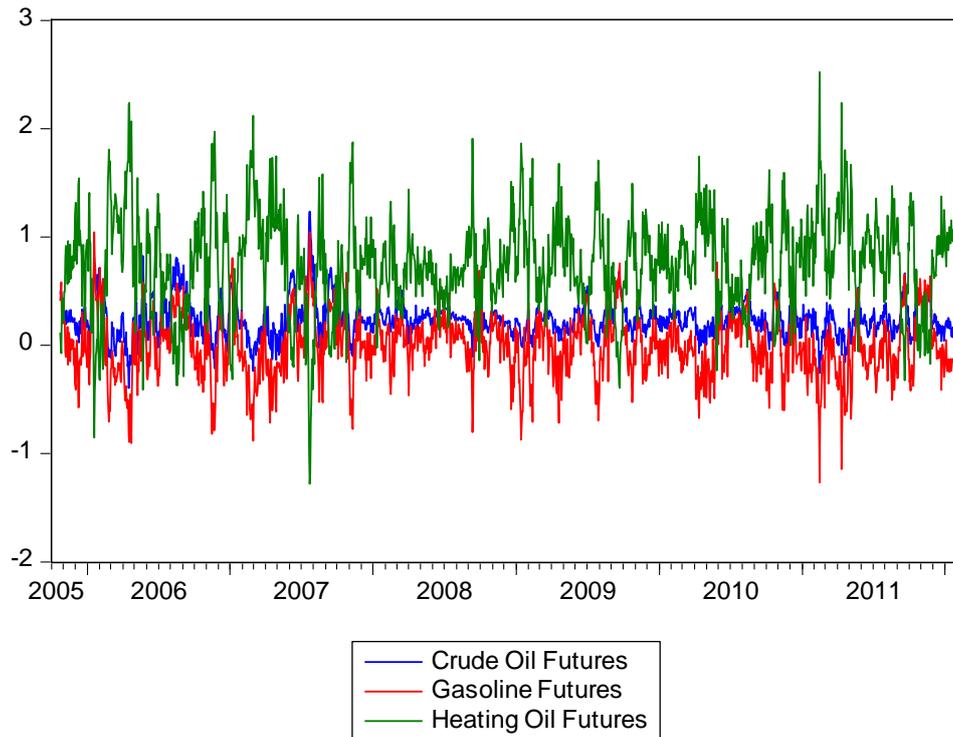
**Figura 22: Rapporti di copertura ottimi DCC ( $\lambda=2$ )**



**Figura 23: Rapporti di copertura ottimi DCC ( $\lambda=4$ )**

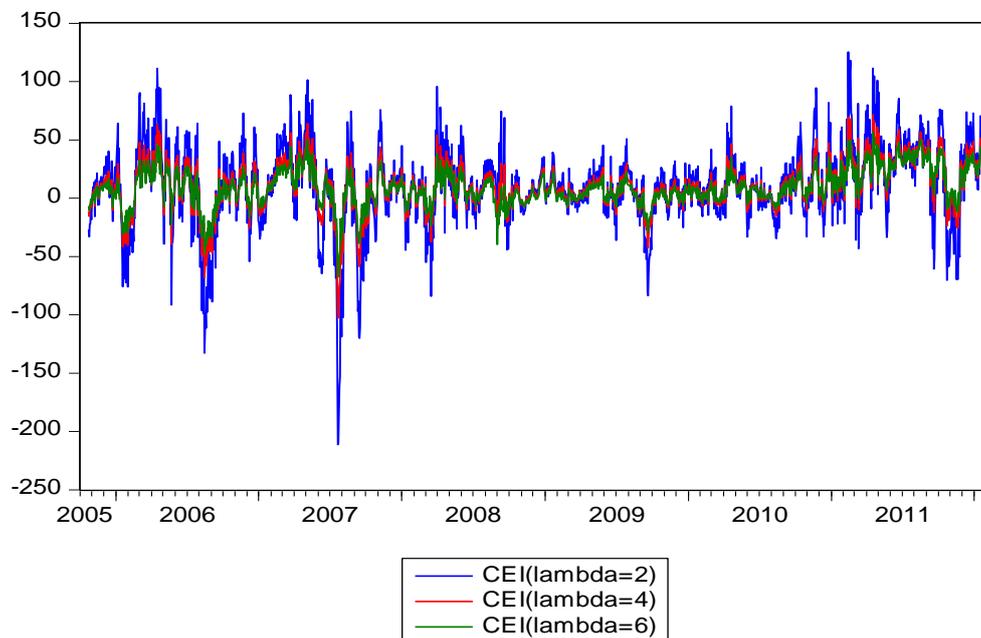


**Figura 24: Rapporti di copertura ottimi DCC ( $\lambda=6$ )**



Anche in termini di performance la variabilità è maggiore con una alta propensione al rischio piuttosto che con una alta avversione al rischio, come lo si può notare dalla figura 23.

**Figura 25: CEI (DCC)**



Completando l'analisi, in termini medi, si ottengono i risultati migliori in corrispondenza di un basso coefficiente di avversione al rischio ( $\lambda=2$ ), anche con le stime DCC (vedi tabella 13).

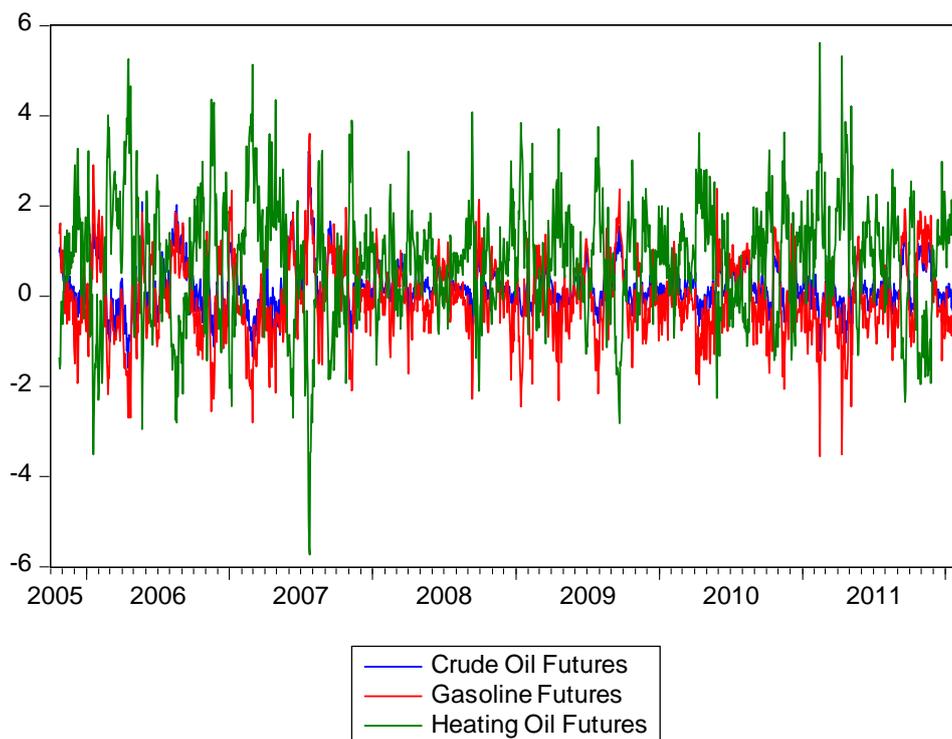
**Tabella 13: Confronto DCC**

	Crude Oil hedge ratio	Gasoline hedge ratio	Heating Oil hedge ratio	CEI Medio
$\lambda=2$	0.241	-0.0444	0.8034	10.1038
$\lambda=4$	0.2464	-0.0156	0.7691	9.9635
$\lambda=6$	0.2482	-0.006	0.7577	8.9006

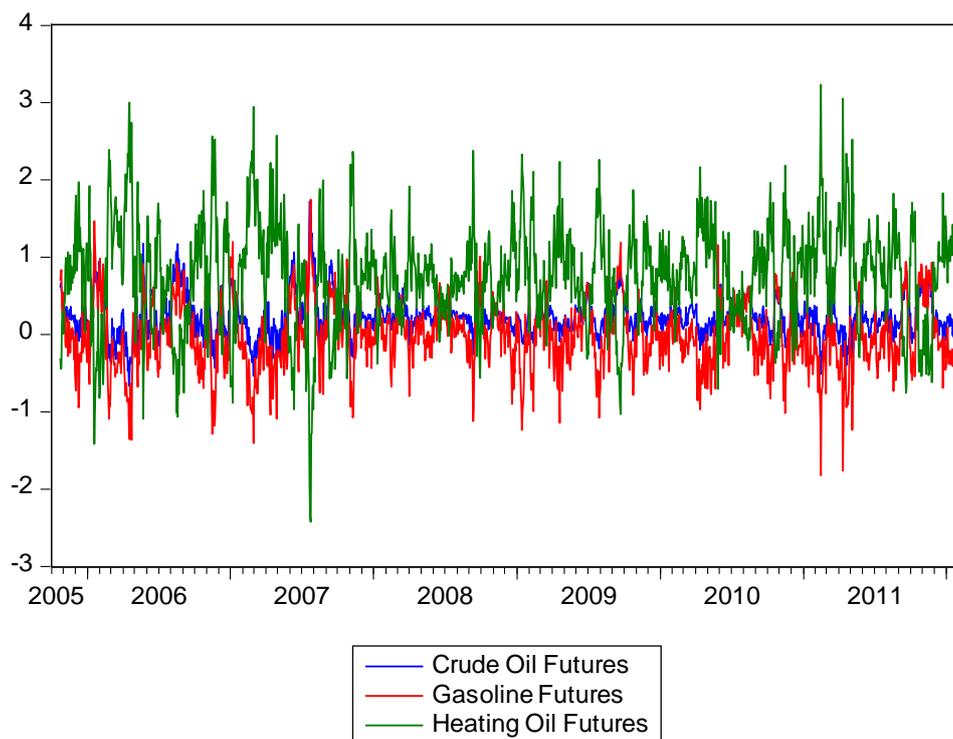
#### 4.4.3. Copertura con BEKK

Le stesse considerazioni fatte per la copertura con DCC valgono anche per il BEKK, infatti grafici (26, 27 e 28) dei rapporti di hedging sono molto simili tra questi due modelli a correlazione dinamica.

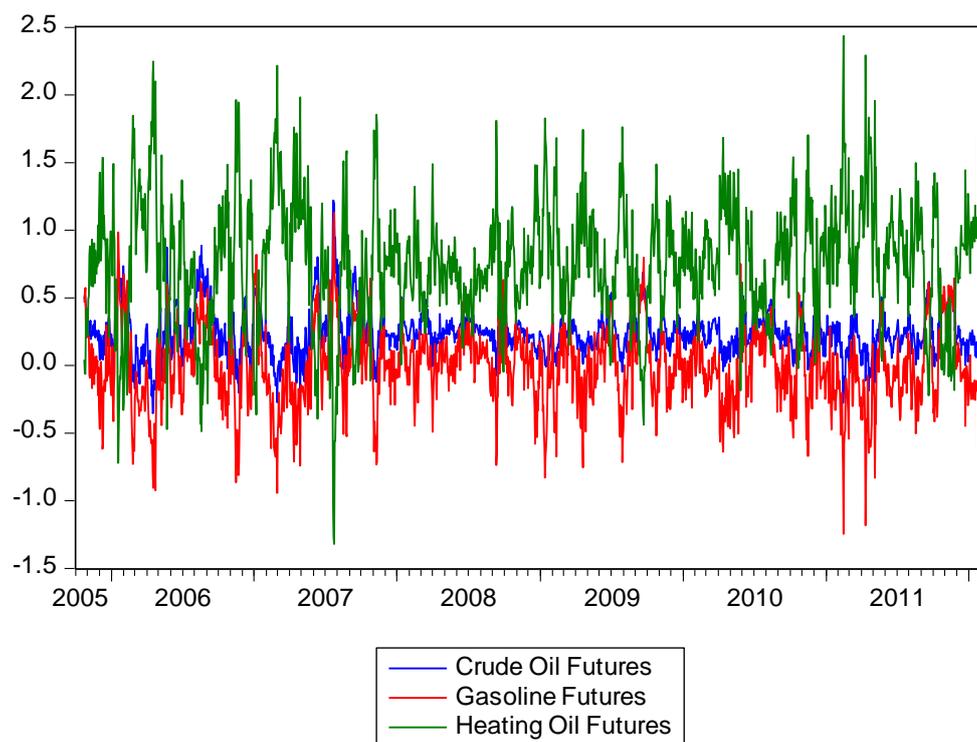
**Figura 26: Rapporti di copertura ottimi BEKK ( $\lambda=2$ )**



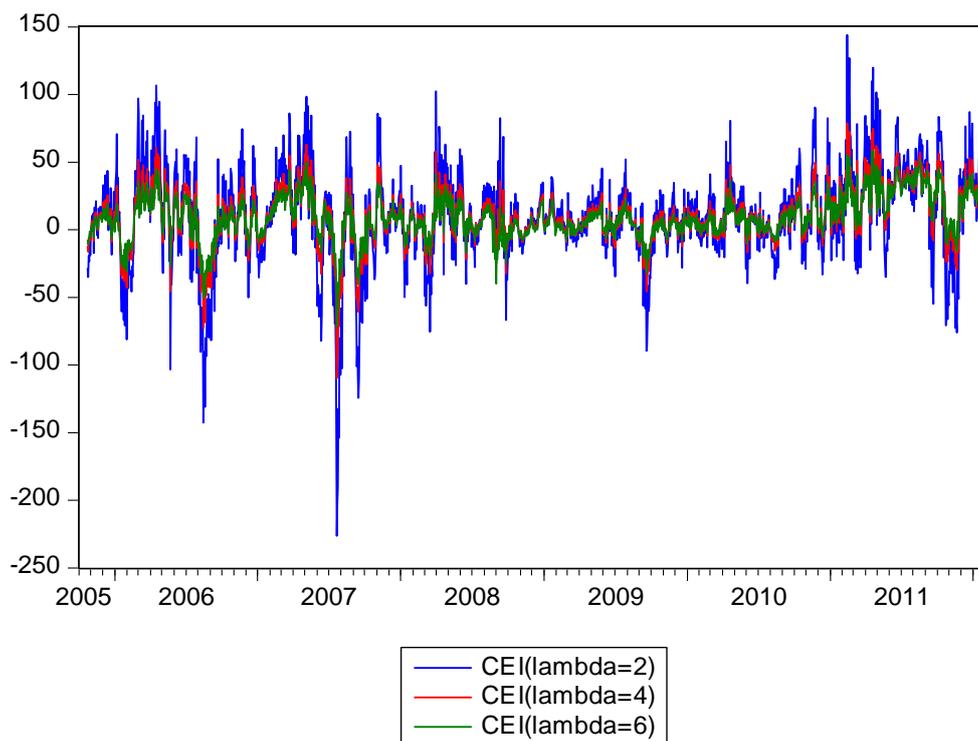
**Figura 27: Rapporti di copertura ottimi BEKK ( $\lambda=4$ )**



**Figura 28: Rapporti di copertura ottimi BEKK ( $\lambda=6$ )**



**Figura 29: CEI (BEKK)**



Anche in questo caso, come per il modello DCC, i risultati migliori in termini medi si ottengono in corrispondenza di un coefficiente di avversione al rischio pari a 2.

**Tabella 14: Confronto BEKK**

	Crude Oil hedge ratio	Gasoline hedge ratio	Heating Oil hedge ratio	CEI Medio
$\lambda=2$	0.241	-0.0434	0.8024	10.1019
$\lambda=4$	0.2464	-0.0151	0.7688	9.953
$\lambda=6$	0.2481	-0.0057	0.7575	8.883

#### 4.4.4. Coperture alternative: Full-hedge e Un-hedge ( $\lambda=6$ )

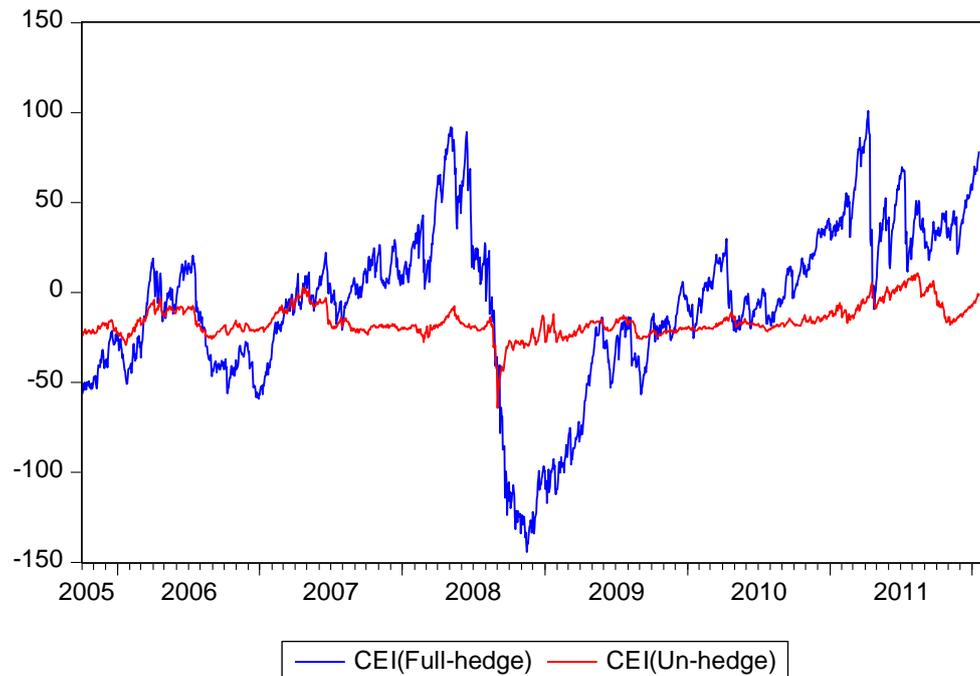
Come accennato precedentemente, sono state costruite altre due coperture:

- **full-hedge:** si concretizza nella vendita di un Crude Oil futures e nel contemporaneo acquisto di un Gasoline futures e di un Heating Oil futures,

- **un-hedge**: è invece una strategia di investimento senza copertura, cioè non si compra né vende nessun contratto futures, ma ci si posiziona solamente sul mercato spot.

Quello che è interessante vedere di queste due strategie, sono le performance graficamente (vedi figura 30): si può notare infatti come la performance della strategia full-hedge sia molto più volatile della strategia con crack spread 3:2:1 (ottenuta dai nostri GARCH-Multivariati); la strategia un-hedge anche se molto costante in termini di performance, non riesce praticamente quasi mai a creare profitto.

**Figura 30: CEI Full-hedge e Un-hedge**



#### 4.5. Confronto finale

Mettendo a confronto tutti i modelli e le strategie, si può notare come, in termini medi, con modelli a correlazione dinamica si ottengano (vedi tabella 15) i risultati migliori, infatti operando con una strategia un-hedge e full-hedge le performance sono negative, date dal fatto che le stime sono state fatte in un range temporale ad alta volatilità.

La copertura ottenuta con i GARCH-Multivariati con sistema DCC è quella che ci fornisce migliori risultati in termini di performance assolute: infatti, per qualsiasi

coefficiente di avversione al rischio, le stime DCC “performano” di più rispetto agli altri modelli. La superiorità del DCC-GARCH è verificata per ogni grado di avversione al rischio ( $\lambda=2,4,6$ ).

Concludendo, a parità di volatilità, il modello DCC, grazie alla stima dinamica della correlazione, ci permette di costruire una strategia di migliore copertura rispetto agli altri modelli analizzati; superiore addirittura al suo gemello BEKK.

**Tabella 15: Confronto risultati**

	Crude Oil hedge ratio	Gasoline hedge ratio	Heating Oil hedge ratio	CEI Medio
<b><math>\lambda=2</math></b>				
Un-hedge	0.0000	0.0000	0.0000	-13.1231
Full-hedge	-1.0000	1.0000	1.0000	1.0867
CCC	0.2465	-0.0338	0.7873	9.5367
BEKK	0.2410	-0.0434	0.8024	10.1019
DCC	0.2410	-0.0444	0.8034	10.1038
<b><math>\lambda=4</math></b>				
Un-hedge	0.0000	0.0000	0.0000	-14.5826
Full-hedge	-1.0000	1.0000	1.0000	-12.8434
CCC	0.2492	-0.0107	0.7615	9.6582
BEKK	0.2464	-0.0151	0.7688	9.9530
DCC	0.2464	-0.0156	0.7691	9.9635
<b><math>\lambda=6</math></b>				
Un-hedge	0.0000	0.0000	0.0000	-16.0421
Full-hedge	-1.0000	1.0000	1.0000	-26.7735
CCC	0.2501	-0.0030	0.7529	8.6721
BEKK	0.2481	-0.0057	0.7575	8.8830
DCC	0.2482	-0.0060	0.7577	8.9006

# CONCLUSIONI

In questo elaborato:

1. Abbiamo stimato le rispettive serie di residui, attraverso un modello VECM analizzando le serie storiche del Crude Oil, Gasoline, Heating Oil futures e spot;
2. I residui così ottenuti sono stati utilizzati per stimare le volatilità dei contratti (matrici di varianze e covarianze, correlazioni), con svariati modelli GARCH-Multivariati sia a correlazione costante che dinamica (CCC, BEKK, DCC).  
Nella letteratura corrente queste volatilità non erano ancora state stimate utilizzando contemporaneamente tutti questi sei tipi di prodotti;
3. Con le stime di volatilità, così ottenute, sono stati costruiti i relativi hedging (con rapporto di crack spread 3:2:1) e delle strategie alternative, con diversi gradi di avversione al rischio. Tutto questo ha permesso di individuare il modello o la strategia più idonea ad ottenere i risultati migliori.

Le considerazioni emerse:

- I migliori risultati in termini di relazioni intercorrenti tra variabili e di processi white noise si ottengono con le stime dei residui calcolate da un VECM (elemento di novità rispetto alla letteratura passata);
- I modelli a correlazione dinamica (BEKK, DCC) ci permettono di ottenere gli hedging migliori, in termini di performance medie, mentre con le strategie alternative (full-hedge, un-hedge) otteniamo quelli peggiori;
- Aumentando il grado di avversione al rischio diminuisce la performance, infatti con tutte le strategie il punto dove la performance media è più alta si trova in corrispondenza di un  $\lambda$  pari a 2, se non per il CCC che è posizionato ad un  $\lambda$  uguale a 4;
- In particolare l'hedging ottimale viene sempre ottenuto con il modello DCC, per ogni  $\lambda$ .

In momenti in cui la volatilità è molto alta, viste le ultime crisi finanziarie, il modello DCC sembra essere un ottimo stimatore della volatilità per gli Energy traders che vogliono coprirsi dal rischio di rialzo del prezzo del petrolio.

Idee per futuri aggiornamenti del presente lavoro possono essere:

- stimare la volatilità dividendo l'intero range temporale in sottoperiodi;

- creare la copertura ottimale utilizzando funzioni obiettivo diverse da quella quadratica;
- minimizzare misure di rischio.

## APPENDICE

Riportiamo parte dei codici MatLab utilizzati: per la stima della volatilità con modelli GARCH, sono stati utilizzati i codici presenti nella Toolbox di K. Sheppard ([http://www.kevinsheppard.com/wiki/UCSD\\_GARCH](http://www.kevinsheppard.com/wiki/UCSD_GARCH)), per la parte di hedging invece sono stati tutti creati personalmente.

Nel seguito vengono presentati i codici utilizzati per la stima dei DCC, a titolo d'esempio; per gli altri modelli sono stati usati gli stessi codici opportunamente modificati.

```
clear all
close all
clc;
data=xlsread('D:\Documenti\Università\Tesi Magistrale\resid VECM.xls',
'Foglio1');
dccP=1;
dccQ=1;
archP=1;
garchQ=1;
[parameters, loglikelihood, Ht, Qt, stdresid, likelihoods, stderrors,
A,B, jointscores]=dcc_mvlgarch(data,dccP,dccQ,archP,garchQ);

%% Correlation matrix
T=size(likelihoods,1);
Crr=zeros(6,6,T);
vrn=zeros(T,6);
for i=1:T

Crr(:, :, i)=inv(diag(sqrt(diag(Qt(:, :, i+1))))) * Qt(:, :, i+1) * inv(diag(sqrt
(diag(Qt(:, :, i+1))))) );
    vrn(i, :)=diag(Ht(:, :, i))';
end

%% Volatility&Correlation Graph/Matrix
figure(1)
tito={'CF', 'GF', 'HF', 'CS', 'GS', 'HS'};
for j=1:6;
    for k=1:6;
        subplot(6,6,6*(j-1)+k);
        plot(squeeze(Crr(j,k,:)));
        title(strcat(tito(j), '-', tito(k)), 'FontSize', 6)
        xlim([1 T])
    end
end

figure(2)
for j=1:3;
    subplot(3,1,j)
    plot(vrn(:,j))
```

```

        title(tito(j))
        xlim([1 T])
end

figure(3)
for j=4:6;
    subplot(3,1,j-3)
    plot(vrn(:,j))
    title(tito(j))
    xlim([1 T])
end

%% Upload Data
alfa=xlsread('D:\Documenti\Università\Tesi
Magistrale\vecm.xlsx','Foglio2','B95:G95');
A=xlsread('D:\Documenti\Università\Tesi
Magistrale\vecm.xlsx','Foglio2','B5:G94');
B=xlsread('D:\Documenti\Università\Tesi
Magistrale\vecm.xlsx','Foglio2','B2:G4');
coint=xlsread('D:\Documenti\Università\Tesi
Magistrale\cointeq.xlsx','Foglio1','B2:D1661');
historical=xlsread('D:\Documenti\Università\Tesi Magistrale\crack
spread evIEWS.xlsx','Foglio1','B2:G1661');
r=diff(historical);%deltadiff
r1=lagmatrix(r,1);
r2=lagmatrix(r,2);
r3=lagmatrix(r,3);
r4=lagmatrix(r,4);
r5=lagmatrix(r,5);
r6=lagmatrix(r,6);
r7=lagmatrix(r,7);
r8=lagmatrix(r,8);
r9=lagmatrix(r,9);
r10=lagmatrix(r,10);
r11=lagmatrix(r,11);
r12=lagmatrix(r,12);
r13=lagmatrix(r,13);
r14=lagmatrix(r,14);
r15=lagmatrix(r,15);

%% Optimal Hedge
lambda=2
y=historical(2:end,[2,3,5,1,4,6]);
T=size(y,1);
R=NaN(T,6);
yHat=NaN(T,6);
b=NaN(T,3);
for t=16:T;
    xt=[];
    for k=1:6;
        xt=[xt r(t,k) r1(t,k) r2(t,k) r3(t,k) r4(t,k) r5(t,k) r6(t,k)
r7(t,k) ...
r8(t,k) r9(t,k) r10(t,k) r11(t,k) r12(t,k) r13(t,k)
r14(t,k)];
    end
    R(t,:)=alfa+coint(t,:)*B+xt*A;
    yHat(t,:)=y(t,:)+R(t,:);
    b(t,:)=(hedge(-R(t,[1,2,3]),squeeze(Ht(1:3,1:3,t-15)),lambda))';
end

```

```

end
mean(b(16:end,:))
plot(b)

%% Graph: forecast and yield
for j=1:3;
    subplot(3,1,j)
    plot([y(16:end,j) yHat(16:end,j)])
    title(tito(j))
    xlim([16 T])
end

%% max CEI
w=[-1,2/3,1/3];
h=diag(w);
S=historical(16:end,[1,4,6]);
F=historical(16:end,[2,3,5]);
Ep=NaN(T-16,1);
Vp=NaN(T-16,1);
for t=1:(T-16);
    Ep(t,1)=w*S(t,:)'+b(t,:)*h*F(t,:)';

Vp(t,1)=w*squeeze(Ht([4,5,6],[4,5,6],t))*w'+b(t,:)*h*squeeze(Ht([1,2,3],
[1,2,3],t))*h*b(t,:)';
    CEI=Ep-lambda/2*Vp-3;
end
mean(CEI(16:end))

%% Unhedge
lambda=6
w=[-1,2/3,1/3];
h=diag(w);%diag([-1,2/3,1/3])
bUH=zeros(T-16,3);
S=historical(16:end,[1,4,6]);
F=historical(16:end,[2,3,5]);
Ep=NaN(T-16,1);
Vp=NaN(T-16,1);
for t=1:(T-16);
    Ep(t,1)=w*S(t,:)'+bUH(t,:)*h*F(t,:)';

Vp(t,1)=w*squeeze(Ht([4,5,6],[4,5,6],t))*w'+bUH(t,:)*h*squeeze(Ht([1,2,3],
[1,2,3],t))*h*bUH(t,:)';
    CEI=Ep-lambda/2*Vp-25;
end
mean(CEI(16:end,:))

%% Graph Spot&Fut
figure(1)
subplot
plot(S(16:end,:))
figure(2)
subplot
plot(F(16:end,:))

%% FullUnhedge
lambda=6
w=[-1,2/3,1/3];
h=diag(w);%diag([-1,2/3,1/3])

```

```

bFH= repmat([-1, 1, 1], T-16, 1);
S= historical(16:end, [1, 4, 6]);
F= historical(16:end, [2, 3, 5]);
Ep= NaN(T-16, 1);
Vp= NaN(T-16, 1);
for t=1:(T-16);
    Ep(t, 1)= w*S(t, :)' + bFH(t, :)*h*F(t, :)' ;

Vp(t, 1)= w*squeeze(Ht([4, 5, 6], [4, 5, 6], t))*w' + bFH(t, :)*h*squeeze(Ht([1, 2, 3], [1, 2, 3], t))*h*bFH(t, :)' ;
    CEI= (Ep-lambda/2*Vp)-170;
end
mean(CEI(16:end, :))

%% Function Hedge
function b=hedge(mu, sigma, lambda)
iota=[1, 1, 1]'; %ones(3,1)
H=[-1 0 0; 0 2/3 0; 0 0 1/3]; %diag([-1, 1/3, 2/3])
theta=-inv(iota'*inv(lambda*H*sigma*H')*iota)*(1-
mu'*H'*inv(lambda*H*sigma*H')*iota);
b=inv(lambda*H*sigma*H')*(H*mu-theta*iota);

```

Infine il programma creato, in Eviews, per stimare il test di Granger causality con la rolling window di 250 osservazioni (anche in questo caso, a titolo d'esempio, viene riportato solo quello per il Crude Oil).

```

'-----
' brute force rolling forecast
'-----

smpl @all
' set window size
!window = 250

' get size of workfile
!length = @obsrange
matrix(!length-!window, 2) caus

for !i = !window to !length-!window
    ' set sample to forecast period
    smpl @first+!i-!window+1 @first+!i
    var var1.ls 1 15 d(crud_oil_wti_fut) d(crude_oil_spot)
    var1.testexog(name=exog)
    caus(!i, 1)=exog(3, 1)
    caus(!i, 2)=exog(3, 2)
    delete exog
next
caus.line

```

## Riferimenti bibliografici

- [1] Ballie RT, Myers RJ, *Bivariate GARCH estimation of the optimal commodity futures hedge*, Journal of Applied Econometrics 6, 1991, pp. 109-124.
- [2] Bollerslev T., *Generalized autoregressive conditional heteroskedacity*, Journal of Applied Econometrics 31, 1986, pp. 301-327.
- [3] Caporin M., McAleer M., *Do We Really Need Both BEKK and DCC? A Tale of Two Covariance Models*, Working Paper, 2009.
- [4] Caporin M., McAleer M., *Scalr BEKK and indirect DCC*, Working Paper No. 05.03, 2005.
- [5] Cecchetti SG, Cumby RE, Figlewsky S., *Estimation of the optimal futures hedge*, The review of Economics and Statistics 70, 1988, pp. 623-630.
- [6] Chicago Mercantile Exchange, *Introduction to crack spreads*, Handbook CME Group.
- [7] Deaves R., Krinsky I., *Risk premiums and efficiency in the market for crude oil futures*, The Energy Journal 13, 1992, pp. 93-117.
- [8] Dempster M.A.H., Medova E., Ke Tang, *Long term spread option valuation and hedging*, Journal of Banking and Finance 32, 2008, pp. 2530-2540.
- [9] Ederington LH, *The hedging performance of the new futures markets*, Journal of Finance 34, 1979, pp. 157-170.
- [10] Engle RF, *Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation*, Econometrica 50, 1982, pp. 987-1008.

- [11] Engle RF, Granger CWJ, *Co-Integration and error correction: representation, estimation, and testing*, *Econometrica* 55, 1987, pp. 251-276.
- [12] Engle RF, Kroner KF, *Multivariate simultaneous generalized ARCH*, *Econometric Theory* 11, 1995, pp. 122-150.
- [13] Engle RF, Sheppard K., *Theoretical and Empirical properties of Dynamic Conditional Correlation Multivariate GARCH*, University of California at San Diego, Economics Working Paper, 2001.
- [14] Fama EF, *The behavior of stock market price*, *Journal of Business* 38, 1965, pp. 34-105.
- [15] Fleming J., Ostdiek B., *The impact of energy derivatives on the crude oil market*, *Energy Economics* 21, 1999, pp. 135-16.
- [16] Gnagnon L., Lypny GJ, McCurdy TH, *Hedging foreign currency portfolios*, *Journal of empirical Finance* 5, 1998, pp. 197-220.
- [17] Haigh M., Holt T., *Crack spread hedging: accounting for time-varying volatility spillovers in the energy futures markets*, *Journal of Applied Econometrics* 17, 2002, pp. 269-289.
- [18] Hamilton JD, *Time series analysis*, Princeton University Press, Princeton, New Jersey, 1994.
- [19] Herbert JH, *Trading volume, maturity and natural gas futures price volatility*, *Energy Economics* 17, 1995, pp. 293-299.
- [20] Kahl KH, *A reformulation of the portfolio model of hedging: comment*, *American Journal of Agricultural Economics* 68, 1986, pp. 1007-1009.

- [21] Kroner KF, Claessens S., *Optimal dynamic hedging portfolios and the currency composition of external debt*, Journal of International Money and Finance 10, 1991, pp. 131-148.
- [22] Kroner KF, Sultan J., *Time varying distributions and dynamic hedging with foreign currency futures*, Journal of Financial and Quantitative Analysis 28, 1993, pp. 535-551.
- [23] Lutkepohl H, *New introduction to multiple time series analysis* , Springer, 2005.
- [24] Moosa IA, Al-Loughani NE, *Unbiasedness and time-varying risk premia in the crude oil futures market*, Energy Economics 16, 1994, pp. 99-105.
- [25] Myers RJ, Thompson SR, *Generalized optimal hedge ratio estimation*, American Journal of Agricultural Economics 71, 1989, pp. 858-868.
- [26] Myers RJ, *Estimating time-varying optimal hedge ratios on futures market*, The Journal of Futures Market 11, 1991, pp. 39-53.
- [27] Nainar SMK, *Market information and price volatility in petroleum derivatives spot and futures markets*, Energy Economics, 1993, pp. 17-24.
- [28] New York Mercantile Exchange, *Crack spread handbook*, Handbook NYMEX Group.
- [29] Samuelson PA, *Proof of the properly anticipated prices fluctuate randomly*, Industrial Management Review 6, 1965, pp. 41-49.
- [30] Serlitis A., *Maturity effects in energy futures*, Energy Economics 12, 1991, pp.150-157.
- [31] Tansuchat R., Chang C.L., McAleer M., *Crude oil hedging strategies using dynamic multivariate GARCH*, Working paper No. 3/2010, College of Business and Economics, University of Canterbury, Christchurch, New Zealand, January 2010.

[32] Tong WHS, *An examination of dynamic hedging*, Journal of International Money and Finance 15, 1996, pp. 19-35.