



UNIVERSITÀ POLITECNICA DELLE MARCHE

---

Dottorato di Ricerca in Economia Politica  
V Ciclo Nuova Serie  
Corso di Econometria

*Giulio Palomba*

MODELLI ARCH

Giugno 2004



---

# Indice

---

<b>1</b>	<b>Introduzione</b>	<b>5</b>
1.1	La volatilità . . . . .	6
1.2	Il modello con innovazioni i.i.d. . . . .	7
<b>2</b>	<b>Regolarità empiriche</b>	<b>9</b>
2.1	Leptocurtosi . . . . .	9
2.2	Volatility clustering . . . . .	9
2.3	Persistenza della volatilità . . . . .	11
2.4	Effetto leverage . . . . .	11
2.5	Informazione e volatilità . . . . .	11
2.5.1	Chiusura dei mercati . . . . .	11
2.5.2	Eventi prevedibili . . . . .	11
2.5.3	Volume degli scambi . . . . .	12
2.5.4	Movimenti comuni . . . . .	12
2.5.5	Volatilità e variabili macroeconomiche . . . . .	12
<b>3</b>	<b>Modelli univariati di tipo ARCH</b>	<b>13</b>
3.1	Modello ARCH . . . . .	15
3.2	Modello GARCH . . . . .	16
3.3	Modello GARCH-M . . . . .	17
3.4	Modelli con asimmetria . . . . .	18
3.4.1	Modello EGARCH . . . . .	18
3.4.2	Modello AGARCH . . . . .	19
3.4.3	Modello QGARCH . . . . .	21
3.5	Altri modelli univariati di tipo ARCH . . . . .	21
3.5.1	Modello IGARCH . . . . .	21
3.5.2	Modello FIGARCH . . . . .	22
3.5.3	Modello SWARCH . . . . .	23

<b>4</b>	<b>Modelli multivariati di tipo ARCH</b>	<b>25</b>
4.1	Modello Vech . . . . .	26
4.2	Modello GARCH diagonale . . . . .	27
4.3	Modello BEKK . . . . .	28
4.4	Modello ARCH fattoriale . . . . .	30
4.5	Modello con correlazioni condizionali costanti . . . . .	31
4.6	Modello con correlazioni condizionali dinamiche . . . . .	33
4.7	Modello O-GARCH . . . . .	35
<b>5</b>	<b>Inferenza nei modelli di tipo ARCH</b>	<b>37</b>
5.1	Stima ML . . . . .	37
5.2	Stima QML . . . . .	39
5.3	Test statistici . . . . .	41

Nell'analisi dei dati macroeconomici spesso si trovano valori in alcuni fenomeni per i quali le varianze dei termini di errore nei modelli temporali sono meno stabili rispetto a quelle che vengono generalmente assunte. Inoltre le serie storiche e soprattutto quelle finanziarie mostrano spesso legami non lineari tra loro e ciò ha causato non pochi problemi, dato che i modelli a disposizione non contemplavano affatto questo aspetto. I risultati ottenuti suggeriscono che nell'analisi di modelli su tale tipologia di dati, errori grandi o piccoli sembrano ricorrere, generando una forma di eteroschedasticità nella quale la varianza del termine di errore dipende da essi.

La prima caratteristica che è emersa sin dai primi contributi pionieristici di Mandelbrot (1963) e Fama (1965) è quella che le serie storiche dei rendimenti sono caratterizzate da distribuzioni leptocurtiche e “volatility clustering”; la loro presenza determina perciò il crollo dell'ipotesi della normalità in quanto le stesse serie storiche mostrano alternanza di periodi con ampie oscillazioni intorno al valore medio e periodi caratterizzati da variazioni di piccola entità. Successivamente altri autori hanno mostrato che la volatilità tende ad aumentare quando si verificano alcuni eventi che determinano un aumento dell'incertezza sui mercati<sup>1</sup>.

L'evidenza empirica ha inoltre fornito anche i seguenti risultati:

- le serie storiche riguardanti i prezzi delle attività finanziarie generalmente sono processi integrati, mentre quelle dei rendimenti sono stazionarie;
- spesso le serie dei rendimenti sono processi frazionalmente integrati;
- i rendimenti solitamente non sono autocorrelati;
- i quadrati dei rendimenti mostrano autocorrelazioni rilevanti avvalorando l'assunzione dell'esistenza di relazioni non lineari tra i rendimenti stessi con i loro valori assunti nel passato.

In base a tali affermazioni crolla l'ipotesi di normalità i.i.d. dei rendimenti, condizione essenziale soprattutto nell'ambito dei modelli di asset allocation<sup>2</sup>, che determina tre condizioni:

---

<sup>1</sup>Si veda in proposito il capitolo 2.

<sup>2</sup>Si veda ad esempio Lee (2000).

1. volatilità dei rendimenti costante per ciascun titolo,
2. correlazioni tra i rendimenti di diversi titoli costante nel tempo,
3. circa il 99% dei dati disponibili hanno valore compreso nell'intervallo  $[\mu - 3\sigma, \mu + 3\sigma]$  dove  $\mu$  e  $\sigma$  rappresentano rispettivamente media e deviazione standard della distribuzione dei rendimenti.

Oltre all'analisi media-varianza la teoria della finanza si è avvalsa di altri contributi che puntano la loro attenzione su aspetti differenti: in questo quadro si inseriscono i numerosi modelli di volatilità variabile che rappresentano utili strumenti per l'interpretazione di tutte le caratteristiche empiriche delle attività finanziarie sin qui elencate. La loro principale caratteristica è quella di riuscire a catturare la non linearità di alcuni fenomeni economici, finalità non raggiungibile mediante i modelli precedenti sull'argomento.

Le principali applicazioni empiriche riguardanti i mercati finanziari erano incentrate soprattutto sull'analisi delle serie storiche dal punto di vista del momento primo condizionale, assumendo invece i momenti successivi sostanzialmente come vincoli. Il sempre maggiore ruolo giocato dal rischio e dall'incertezza nei modelli di decisione ed il fatto che sono stati trovati risultati per i quali il rischio stesso, quindi la volatilità, sono variabili nel tempo, hanno condotto la moderna teoria economica ad elaborare nuove tecniche sulle serie storiche che segnano la nascita di metodi di indagine incentrati principalmente sullo studio dei momenti condizionali successivi al primo.

All'interno di questa categoria si inserisce la classe dei modelli di tipo Autoregressive Conditional Heteroskedasticity (ARCH) che assume una particolare importanza soprattutto per il fatto che per la prima volta viene posta l'attenzione sulla distinzione tra momento secondo condizionale e momento secondo non condizionale: l'elemento innovativo è perciò rappresentato dal fatto che, mentre la matrice delle varianze e covarianze non condizionale di una generica variabile di interesse (rendimenti di titoli, tassi di cambio, tassi di inflazione, ecc.) può non variare nel tempo, quella condizionale spesso mostra un andamento dinamico.

## 1.1 La volatilità

La volatilità è la variabilità di un valore o di un indice finanziario calcolata in un determinato intervallo di tempo: essa rappresenta un nuovo oggetto di studio nell'analisi delle serie storiche da quando ci si è resi conto del fatto che componenti come rischio, incertezza o mutamenti strutturali svolgono un ruolo importante e spesso decisivo all'interno del sistema economico.

La volatilità costituisce l'espressione osservabile dell'incertezza presente sui mercati finanziari ed analiticamente si identifica nel concetto di varianza condizionale (qualora esista) ad un dato set informativo disponibile al tempo  $t$ : nel caso dei modelli univariati la volatilità è uno scalare ( $h_t$ ) poiché si lavora con un'unica variabile di interesse, mentre nei modelli multivariati essa è rappresentata da una matrice

quadrata, simmetrica e almeno semidefinita positiva ( $\Omega_t$ ) con dimensioni pari al numero delle variabili considerate. La volatilità è un fenomeno caratterizzato da una memoria che può risalire molto lontano nel tempo, quindi, in tutte le sue espressioni, contempla anche i suoi valori passati.

Ad esempio, nell'ottica delle strategie di asset allocation, risulta di fondamentale importanza la possibilità di poter esaminare contemporaneamente i movimenti nella volatilità dei rendimenti di più attività in quanto il rischio di portafoglio potrebbe essere ridotto tramite la diversificazione.

Per spiegare il fenomeno della volatilità variabile Engle (1982) ha elaborato il primo modello di tipo ARCH, un nuovo metodo di analisi delle serie storiche basato sull'intuizione che la varianza condizionale sia in relazione con i valori da essa assunti nel passato; dal punto di vista econometrico questo discorso si traduce nel fatto che la volatilità mostra dinamiche autoregressive nel tempo.

In riferimento al modello di regressione lineare, il contributo di Engle (1982) ha perciò fornito le basi per la formalizzazione successiva di numerosi modelli di specificazione dell'errore non lineari né in media né in varianza e soprattutto in grado di spiegare diverse caratteristiche connesse al fenomeno della volatilità variabile.

## 1.2 Il modello con innovazioni i.i.d.

Il teorema di rappresentazione di Wold afferma che ogni processo stocastico  $y_t$  stazionario in covarianza può essere scritto come somma di una componente deterministica ( $d_t$ ) e di una idiosincratca che segue un processo MA( $\infty$ ). Analiticamente si ha perciò:

$$y_t = d_t + u_t \quad [1.1]$$

dove  $u_t = C(L)\varepsilon_t$ ,  $\sum_{i=1}^{\infty} c_i^2 < \infty$  e  $c_0 = 1$ . I momenti dell'innovazione i.i.d. sono:

- $E(\varepsilon_t) = 0$
- $E(\varepsilon_t \varepsilon_\tau) = \begin{cases} \sigma_\varepsilon^2 & \text{se } t = \tau \\ 0 & \text{se } t \neq \tau \end{cases}$

La condizione di incorrelazione è sufficiente per concludere che in un modello lineare di questo tipo le innovazioni non devono essere necessariamente indipendenti o avere distribuzione normale. Quando le innovazioni non sono indipendenti generalmente si è in presenza di serie storiche non lineari, categoria alla quale appartengono quelle caratterizzate dalla presenza di eteroschedasticità condizionale.

Si consideri la rappresentazione MA( $\infty$ ) del processo stocastico  $y_t = C(L)\varepsilon_t$  stazionario in covarianza, dove  $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$ ; i momenti non condizionali non dipendono dal tempo, infatti risulta:

1.  $E(y_t) = 0$

$$2. \text{Var}(y_t) = E(y_t^2) = \sigma_\varepsilon^2 \sum_{i=1}^{\infty} c_i^2$$

Dato il set informativo  $I_{t-1}$ , che consiste in una  $\sigma$ -algebra generata dalle osservazioni passate  $\{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots\}$ , i momenti condizionali<sup>3</sup> sono:

$$1. E(y_t | I_{t-1}) = \sum_{i=1}^{\infty} c_i \varepsilon_{t-i}$$

$$2. \text{Var}(y_t | I_{t-1}) = E\{[y_t - E(y_t | I_{t-1})]^2 | I_{t-1}\} = \sigma_\varepsilon^2$$

Il valore atteso condizionale ha espressione variabile in base al tempo, mentre la varianza condizionale è costante: questa caratteristica rende il modello con innovazioni i.i.d. inadeguato per l'analisi della volatilità variabile. Per sottolineare tale aspetto si consideri la previsione condizionale di  $k$  periodi:

$$E(y_{t+k} | I_t) = \sum_{i=1}^{\infty} c_{k+i} \varepsilon_{t-i} \quad [1.2]$$

L'errore di previsione associato è:

$$y_{t+k} - E(y_{t+k} | I_t) = \sum_{i=1}^{\infty} c_i \varepsilon_{t+k-i} \quad [1.3]$$

La varianza condizionale della [1.3] è la seguente:

$$E\{[y_{t+k} - E(y_{t+k} | I_t)]^2 | I_t\} = \sigma_\varepsilon^2 \sum_{i=1}^{k-1} c_i^2 \quad [1.4]$$

dove

$$\lim_{k \rightarrow \infty} E\{[y_{t+k} - E(y_{t+k} | I_t)]^2 | I_t\} = \sigma_\varepsilon^2 \sum_{i=1}^{\infty} c_i^2 = \text{Var}(y_t)$$

L'espressione della varianza dell'errore di previsione dipende solo da  $k$  e non dal set informativo disponibile  $I_t$ , quindi il modello con innovazioni i.i.d. non è in grado di tenere conto della nuova informazione rilevante al tempo  $t$ .

---

<sup>3</sup>In molti testi per indicare il momento condizionale rispetto al set informativo  $I_{t-h}$  viene utilizzata un'altra notazione rispetto a queste pagine: ad esempio nel caso del valore atteso condizionale le notazioni  $E(y_t | I_{t-h})$  e  $E_{t-h}(y_t)$  sono del tutto equivalenti.



Questo capitolo consiste in una breve rassegna di alcuni aspetti ricorrenti comuni delle serie storiche finanziarie. I modelli di tipo ARCH sono stati sviluppati per rendere conto di tali caratteristiche empiriche.

### 2.1 Leptocurtosi

Mandelbrot (1963) e Fama (1965, 1970) sono stati i primi a documentare il fatto che le serie storiche sui rendimenti sono caratterizzate da distribuzioni leptocurtiche, cioè da distribuzioni all'interno delle quali la massa di probabilità che si addensa sulle code è maggiore di quella che si registra nella funzione di densità della variabile casuale normale: dal punto di vista statistico ciò si traduce con il fatto che dai campioni osservati emerge un eccesso di curtosi. La Figura 2.1 mostra questa evidenza dal punto di vista grafico<sup>1</sup>.

### 2.2 Volatility clustering

La volatilità è un fenomeno persistente cioè è in stretta relazione con il valore da essa assunto nel periodo precedente, ma anche con quello successivo al momento dell'analisi. Per mettere in evidenza tale fenomeno Mandelbrot (1963) scrive:

«grandi cambiamenti tendono ad essere seguiti da grandi cambiamenti, mentre piccoli cambiamenti tendono ad essere seguiti da piccoli cambiamenti»

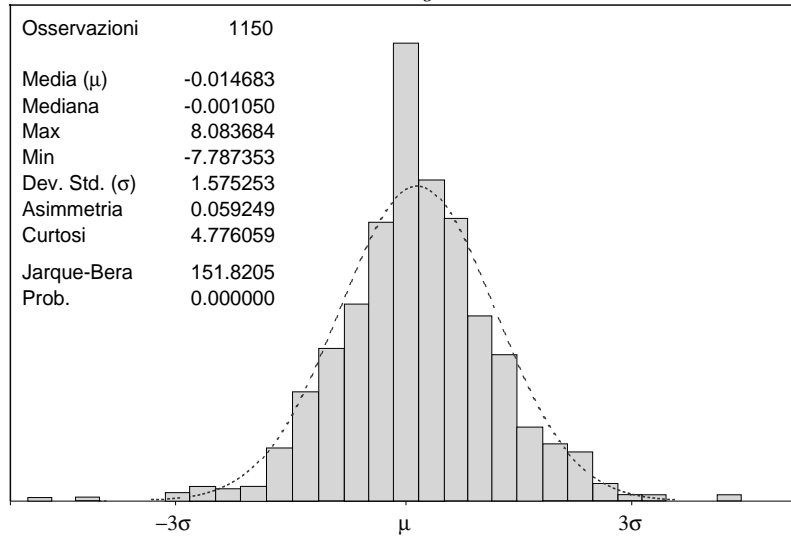
In queste parole risiede il concetto di volatility clustering (volatilità a grappoli) che si identifica nell'andamento delle serie storiche che mostrano un continuo alternarsi di ampie oscillazioni e di piccole oscillazioni intorno al loro valore medio (Figura 2.2). Volatility clustering e leptocurtosi sono due componenti in stretta

---

<sup>1</sup>Le Figure 2.1 e 2.2 si riferiscono ad un campione di 1150 osservazioni relative ai rendimenti giornalieri dell'indice azionario MIB30 registrati dal 3 novembre 1998 al 31 Marzo 2003. La linea tratteggiata in Figura 2.1 rappresenta la funzione di densità di una variabile casuale normale avente media e deviazione standard pari rispettivamente ai  $\mu$  e  $\sigma$  del grafico.

## 2. REGOLARITÀ EMPIRICHE

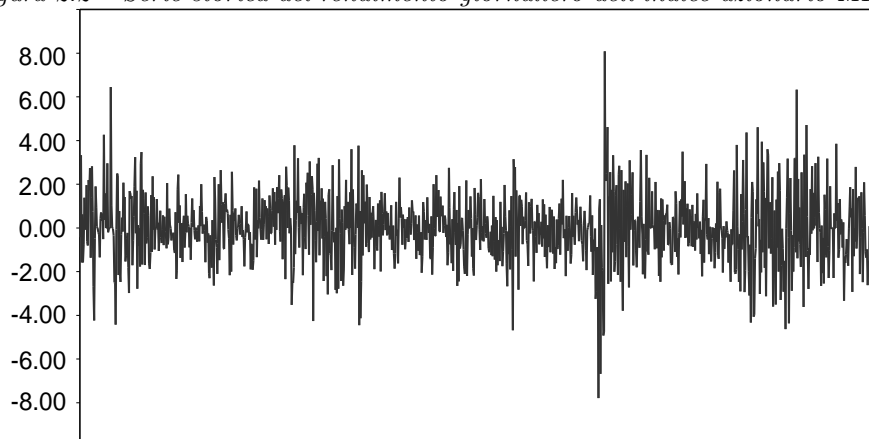
Figura 2.1 – Distribuzione del rendimento giornaliero dell'indice azionario MIB30



relazione tra loro. I modelli di eteroschedasticità condizionale come quelli di tipo ARCH o di volatilità stocastica<sup>2</sup> (modelli SV) sono stati introdotti sostanzialmente per modellare questo fenomeno.

Diebold (1988) e Drost e Nijman (1993) documentano inoltre il fatto che attraverso l'aggregazione temporale dei dati il fenomeno del volatility clustering tende a svanire.

Figura 2.2 – Serie storica del rendimento giornaliero dell'indice azionario MIB30



<sup>2</sup>Per una rassegna sui modelli SV si veda Ghysels, Harvey e Renault (1996).

### 2.3 Persistenza della volatilità

La volatilità dei rendimenti è un fenomeno dotato di memoria lunga, cioè accade frequentemente che il processo stocastico che la governa non è stazionario in covarianza poiché presenta radici unitarie nei suoi ritardi. Ciò significa che qualsiasi valore assunto dalla varianza condizionale è in relazione coi valori assunti da essa nei precedenti periodi di osservazione. In termini analitici questa caratteristica si traduce nel fatto che le serie storiche dei quadrati dei rendimenti non sono serialmente indipendenti, quindi si è in presenza di relazioni di tipo non lineare tra le osservazioni campionarie.

### 2.4 Effetto leverage

Black (1976) scopre che esiste una tendenza delle variazioni dello stock di prezzi ad essere negativamente correlate con quelle che si verificano per la volatilità. Ciò accade perché una diminuzione dei prezzi determina un aumento del grado di leverage delle imprese, quindi dell'incertezza e della volatilità. Lo stesso autore aggiunge che l'effetto leverage da solo non è però in grado di spiegare in modo completo le asimmetrie nei prezzi.

### 2.5 Informazione e volatilità

I rendimenti vengono solitamente misurati ed analizzati tramite osservazioni campionarie che vengono ottenute con cadenze regolari (dati giornalieri, mensili o annuali); tuttavia le informazioni non giungono sui mercati in modo uniforme nel tempo e spesso non sono neppure osservabili poiché possono verificarsi alcuni eventi che generano una maggiore incertezza.

#### 2.5.1 Chiusura dei mercati

I periodi di chiusura dei mercati rappresentano uno dei momenti principali in cui le notizie giungono in modo frammentato: alla riapertura esse si riflettono inevitabilmente sui prezzi dei titoli e ciò determina un aumento dell'incertezza, quindi della volatilità. Fama (1965) e French e Roll (1986) hanno notato che le informazioni si accumulano molto più lentamente quando i mercati sono chiusi rispetto al caso in cui sono aperti, quindi le varianze dei rendimenti sono maggiori a seguito dei weekends o di periodi legati a festività o vacanze.

#### 2.5.2 Eventi prevedibili

La volatilità è maggiore quando si è in presenza di eventi prevedibili e ciò è stato riscontrato in diverse indagini empiriche<sup>3</sup>: in presenza di annunci, operazioni tra

<sup>3</sup>Si vedano i contributi di Cornell (1978) oppure quelli di Patell e Wolfson (1979, 1981).

Banche Centrali, realizzazioni di avvenimenti attesi nel sistema economico, aumenta il volume delle informazioni, quindi la volatilità assume valori più alti.

### 2.5.3 Volume degli scambi

Numerosi studi empirici<sup>4</sup> hanno mostrato un legame diretto tra la volatilità ed il numero di scambi che avvengono all'interno del mercato: un aumento delle transazioni provoca inevitabilmente un flusso aggiuntivo in entrata di nuove informazioni che va ad incrementare il valore delle varianze dei rendimenti.

### 2.5.4 Movimenti comuni

Black (1976) evidenzia il fatto che se nel mercato è stata riscontrata una certa variazione nella volatilità, la stessa variazione si ottiene per ciascun titolo presente in quel mercato: ciò significa che se le volatilità cambiano in un dato periodo, esse tendono a muoversi tutte nella stessa direzione.

La volatilità che in un dato istante investe un determinato titolo nello stesso tempo interessa in modo analogo molti altri titoli. In questo modo l'intero mercato nel tempo è caratterizzato a livello aggregato da valori di volatilità sostanzialmente coincidenti con quelli particolari registrati nei mercati di ogni singola attività finanziaria.

Diversi autori<sup>5</sup> hanno mostrato che tale caratteristica può riguardare non solo un singolo mercato, ma anche una pluralità di mercati contemporaneamente.

### 2.5.5 Volatilità e variabili macroeconomiche

Schwert (1989) mette in evidenza che la volatilità cresce nel corso di periodi di recessione e di crisi finanziarie, mentre diminuisce quando si è in presenza di fasi di espansione.

Glosten, Jagannathan e Runkle (1993) scoprono una relazione diretta tra volatilità dei rendimenti e tassi di interesse: l'incertezza determina una crescita della volatilità, ma il suo legame con essa risulta debole.

---

<sup>4</sup>Il legame tra volatilità e volume degli scambi è stato rilevato in numerosi contributi: a titolo di esempio si vedano Karpoff (1987) oppure Gallant, Rossi e Tauchen (1992).

<sup>5</sup>Si vedano i seguenti contributi: Schwert (1989), Engle e Susmel (1993), King, Sentana e Wadhvani (1994).

L'introduzione dei modelli di tipo ARCH si giunge alla formalizzazione di una classe specifica di processi non lineari né in media né in varianza in grado di spiegare diverse caratteristiche connesse al fenomeno della volatilità variabile.

In questo capitolo saranno richiamati i concetti basilari relativi ai principali modelli univariati di tipo ARCH e le loro proprietà. Dato il set informativo  $I_{t-1} = \{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}\}$ , il termine di disturbo di un modello di regressione lineare segue un processo di tipo ARCH se sono verificate le seguenti condizioni:

1. La media di  $\varepsilon_t$  condizionale al set informativo  $I_{t-1}$  è nulla per ogni  $t$ .

$$E(\varepsilon_t | I_{t-1}) = 0 \quad [3.1]$$

Ciò implica anche che  $E(\varepsilon_t) = 0$  e soprattutto che tale processo risulti serialmente incorrelato condizionatamente allo stesso set informativo, infatti:

$$Cov(\varepsilon_t \varepsilon_{t+k} | I_{t-h}) = E(\varepsilon_t \varepsilon_{t+k} | I_{t-h}) - E(\varepsilon_t | I_{t-h}) E(\varepsilon_{t+k} | I_{t-h})$$

Poiché per la legge dei valori attesi iterati<sup>1</sup>  $E(\varepsilon_t | I_{t-h}) = E[E(\varepsilon_t | I_{t-1}) | I_{t-h}] = 0$ , risulta:

$$\begin{aligned} Cov(\varepsilon_t \varepsilon_{t+k} | I_{t-h}) &= E(\varepsilon_t \varepsilon_{t+k} | I_{t-h}) \\ Cov(\varepsilon_t \varepsilon_{t+k} | I_{t-h}) &= E[E(\varepsilon_t \varepsilon_{t+k} | I_{t+k-1}) | I_{t-h}] \\ Cov(\varepsilon_t \varepsilon_{t+k} | I_{t-h}) &= E[\varepsilon_t \cdot E(\varepsilon_{t+k} | I_{t+k-1}) | I_{t-h}] \end{aligned}$$

Dato che  $E(\varepsilon_{t+k} | I_{t+k-1}) = 0$  si ha:

$$Cov(\varepsilon_t \varepsilon_{t+k} | I_{t-h}) = 0$$

<sup>1</sup>Date le variabili casuali  $X$  ed  $Y$ , la legge dei valori attesi iterati afferma che:

$$E[E(Y|X)] = E(Y)$$

2. La componente idiosincratca o innovazione  $\varepsilon_t$  è data dalla relazione:

$$\varepsilon_t = u_t h_t^{1/2} \quad [3.2]$$

dove  $u_t \sim i.i.d. (0, 1)$  è detto processo standardizzato<sup>2</sup>. La [3.2] rappresenta l'equazione di definizione dell'innovazione universalmente valida per qualsiasi modello di tipo ARCH: in base a questa espressione si assume che  $\varepsilon_t^2$  è un previsore corretto per la volatilità in quanto per definizione si ha:

$$h_t = E(\varepsilon_t^2 | I_{t-1}) = Var(\varepsilon_t | I_{t-1}) \quad [3.3]$$

Date queste premesse la distribuzione condizionale dell'innovazione risulta essere:

$$\varepsilon_t | I_{t-1} \sim N(0, h_t) \quad [3.4]$$

Da questa equazione emerge chiaramente che la varianza condizionale è variabile nel tempo. In sintesi un modello di regressione lineare sui rendimenti  $y_t$  con innovazioni che seguono un modello di tipo ARCH si configura come segue:

$$\begin{cases} y_t = x_t' b + \varepsilon_t \\ \varepsilon_t = u_t h_t^{1/2} \\ h_t = E(\varepsilon_t^2 | I_{t-1}) \\ y_t | I_{t-1} \sim N(x_t' b, h_t) \end{cases} \quad [3.5]$$

I modelli di tipo ARCH rappresentano quindi le espressioni analitiche per la varianza condizionale  $h_t$ . Un importante elemento da evidenziare è il fatto che questi processi rappresentano lo strumento formale che mette in relazione le dinamiche (condizionali) della volatilità con il concetto di leptocurticità (non condizionale). In questo contesto infatti l'innovazione rispetta pienamente la relazione:

$$\frac{E(\varepsilon_t^4)}{[E(\varepsilon_t^2)]^2} \geq 3 \quad [3.6]$$

La dimostrazione di questo enunciato parte dalla distribuzione condizionale di cui alla [3.4] per la quale deve risultare:

$$E(\varepsilon_t^4 | I_{t-1}) = 3[E(\varepsilon_t^2 | I_{t-1})]^2$$

Sostituendo la [3.3], per la legge dei valori attesi iterati si ha:

$$E(\varepsilon_t^4) = 3E(h_t^2)$$

Poiché per la disuguaglianza di Jensen<sup>3</sup> si ha  $E(h_t^2) \geq [E(h_t)]^2 = [E(\varepsilon_t^2)]^2$ , si ottiene:

$$E(\varepsilon_t^4) - 3[E(\varepsilon_t^2)]^2 = 3E(h_t^2) - 3[E(h_t)]^2$$

<sup>2</sup>Per semplicità molto spesso si ipotizza  $u_t \sim N(0, 1)$ .

<sup>3</sup>Data la variabile casuale  $X$  con media  $E(X)$  e la funzione convessa  $g(X)$  risulta:

$$E[g(X)] \geq g[E(X)]$$

Il valore dell'indice di curtosi sarà perciò:

$$\frac{E(\varepsilon_t^4)}{[E(\varepsilon_t^2)]^2} = 3 + 3 \frac{E(h_t^2) - 3[E(h_t)]^2}{[E(\varepsilon_t^2)]^2} = 3 + 3 \frac{Var(h_t)}{[E(\varepsilon_t^2)]^2} \quad [3.7]$$

### 3.1 Modello ARCH

Introdotta da Engle (1982), il modello ARCH specifica la varianza condizionale come una funzione lineare dei quadrati dei valori passati delle innovazioni, cioè:

$$h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \quad [3.8]$$

dove  $\omega \geq 0$  e tutti gli  $\alpha_i \geq 0$  per  $i = 1, 2, \dots, q$  rappresentano i parametri da stimare. La [3.8] è l'equazione relativa all'innovazione  $\varepsilon_t$  che segue un generico processo ARCH(q) dove  $q$  è il numero di ritardi di  $\varepsilon_t^2$ .

L'ARCH è quindi un processo con media nulla, varianza costante e varianza condizionale linearmente dipendente dai quadrati delle innovazioni. Esso riesce a catturare il fenomeno delle oscillazioni delle serie storiche relative ai rendimenti dei titoli, quindi interpreta il volatility clustering.

L'ARCH(q) di cui alla [3.8] può essere riscritto mediante un processo MA(q) per i quadrati delle innovazioni<sup>4</sup>, quindi:

$$h_t = \omega + A(L)\varepsilon_t^2 \quad [3.9]$$

dove  $A(L) = \alpha_1 L + \alpha_2 L^2 + \dots + \alpha_q L^q$  è il polinomio nell'operatore ritardo. Ponendo inoltre  $v_t = \varepsilon_t^2 - h_t$ , l'ARCH(q) può essere riscritto anche come un modello AR(q) per i quadrati dei disturbi, cioè:

$$\varepsilon_t^2 = \omega + A(L)\varepsilon_t^2 + v_t \quad [3.10]$$

dove  $v_t$  è un termine non normale, asimmetrico ed eteroschedastico con media condizionale nulla. L'ARCH(q) è stazionario in covarianza quando le radici del polinomio  $1 - A(L)$  cadono al di fuori del cerchio unitario<sup>5</sup>. Ciò si traduce nella condizione:

$$\sum_{i=1}^q \alpha_i < 1 \quad [3.11]$$

Se il processo è stazionario la varianza non condizionale dell'innovazione  $\varepsilon_t$  assume il seguente valore:

$$Var(\varepsilon_t) = \frac{\omega}{1 - \sum_{i=1}^q \alpha_i} \quad [3.12]$$

---

<sup>4</sup>In questo caso occorre tenere presente che tutti gli incrementi condizionali e non condizionali di  $h_t$  non hanno media nulla.

<sup>5</sup>Poiché le radici del polinomio ritardo possono essere date sia da numeri reali, sia da numeri complessi, tale caratteristica si identifica nella condizione nella quale il loro modulo risulta essere maggiore di 1.

### 3.2 Modello GARCH

Elaborato da Bollerslev (1986) il modello Generalized ARCH (GARCH) rappresenta un utile strumento per analizzare la persistenza dei movimenti della volatilità senza dover stimare l'alto numero di parametri presenti nel polinomio  $A(L)$ . Dato che anch'esso si basa sul set informativo  $I_{t-1}$ , l'equazione di un generico modello GARCH(p,q) specifica la varianza condizionale come segue:

$$h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad [3.13]$$

dove  $\omega \geq 0$ ,  $\alpha_i \geq 0$  per  $i = 1, 2, \dots, q$  e  $\beta_j \geq 0$  per  $j = 1, 2, \dots, p$ . La varianza condizionale dipende da  $(1+p+q)$  parametri dove  $p$  si riferisce all'ordine del ritardo della parte autoregressiva rappresentata dai valori della varianza condizionale stessa nel passato, mentre  $q$  mostra il numero dei ritardi della componente  $\varepsilon_t^2$ . Riscrivendo l'espressione utilizzando l'operatore ritardo si ha:

$$h_t = \omega + A(L)\varepsilon_t^2 + B(L)h_t \quad [3.14]$$

Definendo  $m = \max\{p, q\}$ , il GARCH(p,q) è facilmente riconducibile ad un ARMA(m,p) per i quadrati delle innovazioni passate semplicemente imponendo la solita condizione  $v_t = \varepsilon_t^2 - h_t$ . Dopo alcuni passaggi si ricava la seguente espressione:

$$\varepsilon_t^2 = \omega + [A(L) + B(L)]\varepsilon_t^2 - B(L)v_t + v_t \quad [3.15]$$

La [3.15] è stata utilizzata per la previsione della volatilità nel tempo  $t+k$  da da Engle e Bollerslev (1986) i quali ottengono:

$$\begin{aligned} h_{t+k} &= \omega + \sum_{i=1}^N (\alpha_i \varepsilon_{t+k-i}^2 + \beta_i h_{t+k-i}) + \sum_{i=k}^m (\alpha_i \varepsilon_{t+k-i}^2 + \beta_i h_{t+k-i}) \\ E(h_{t+k}|I_t) &= \omega + \sum_{i=1}^N (\alpha_i + \beta_i) E(h_{t+k-i}|I_t) + \sum_{i=k}^m (\alpha_i \varepsilon_{t+k-i}^2 + \beta_i h_{t+k-i}) \end{aligned} \quad [3.16]$$

dove  $N = \min\{m, k-1\}$ . La formula [3.16] è applicata interamente solo nel caso in cui i  $k$  periodi di previsione considerati sono in numerosità maggiore rispetto al valore massimo dei ritardi del modello: se  $k > m$  infatti il secondo addendo è nullo in quanto non potrebbe esistere una somma da  $k$  ad  $m$ , se  $k = m$  il primo addendo si annulla perché costituirebbe una somma da 1 a 0.

Bollerslev (1986) afferma che il processo GARCH(p,q) è stazionario in covarianza quando le radici del polinomio  $1 - A(L) - B(L)$  cadono al di fuori del cerchio unitario. Ciò si traduce nella seguente relazione:

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1 \quad [3.17]$$



con tutti i parametri  $\omega$ ,  $\alpha_i$  e  $\beta_j$  non negativi<sup>6</sup>. Quando vale questa condizione, la varianza non condizionale dell'innovazione risulta essere:

$$Var(\varepsilon_t) = \frac{\omega}{1 - \sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j} \quad [3.18]$$

Bollerslev (1986) inoltre fornisce una utile discussione che mostra le condizioni necessarie per assicurare la stabilità dei momenti di ordine superiore al dodicesimo per un modello GARCH(1,1).

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \quad [3.19]$$

$$h_t = \omega + (\alpha + \beta)h_{t-1} + \alpha(\varepsilon_{t-1}^2 - h_{t-1}) \quad [3.20]$$

$$h_t = \omega + (\alpha + \beta)h_{t-1} + \alpha h_{t-1}(u_{t-1}^2 - 1) \quad [3.21]$$

Nell'equazione [3.20] il termine  $(\varepsilon_{t-1}^2 - h_{t-1})$ , che ha media condizionale al tempo  $t-1$  pari a zero, può essere considerato come uno shock sulla volatilità. Il coefficiente  $\alpha$  misura l'incidenza dell'effetto che lo shock sulla volatilità di un periodo esercita su quella del periodo successivo, mentre  $(\alpha + \beta)$  è il tasso al quale questo effetto va a ridursi nel tempo. L'equazione [3.21] riscrive gli shock con la formula  $h_{t-1}(u_{t-1}^2 - 1)$  cioè come quadrato di una variabile casuale normale standardizzata meno la sua media e moltiplicata per il valore passato della volatilità<sup>7</sup>.

### 3.3 Modello GARCH-M

In tutti i modelli GARCH si determina il momento secondo condizionale variabile nel tempo, mentre il momento primo condizionale è assunto costante uguale a zero. Nella realtà dei mercati finanziari media e varianza dei rendimenti dei titoli mutano continuamente, quindi occorre elaborare un nuovo modello che tenga conto di tale assunzione.

Engle, Lilien e Robins (1987) hanno modificato la parte strutturale del modello di regressione inserendo al suo interno una funzione della varianza condizionale  $g(h_t)$ , ottenendo la forma:

$$\begin{cases} y_t = X_t b + \delta g(h_t) + \varepsilon_t \\ \varepsilon_t | I_{t-1} \sim N(0, h_t) \end{cases} \quad [3.22]$$

dove  $y_t$  e  $X_t$  sono rispettivamente la variabile dipendente ed i regressori del modello lineare, mentre la volatilità  $h_t$  segue un qualsiasi processo di tipo ARCH. Quando l'innovazione segue un processo ARCH(q) si parla di modello ARCH-M. La definizione di tale modello dipende espressamente dalla forma funzionale per la varianza

<sup>6</sup>Nelson e Cao (1992) forniscono un set di condizioni sufficienti per la stazionarietà in covarianza di un GARCH(p,q) che risultano essere meno stringenti della [3.17] in quanto non occorre che tutti i parametri siano necessariamente non negativi.

<sup>7</sup>È quindi evidente che risulta  $u_{t-1}^2 \sim \chi_1^2$ .

condizionale: generalmente risulta  $g(h_t) = h_t$ , ma in alcuni casi sono state utilizzate anche  $g(h_t) = h_t^{1/2}$  e  $g(h_t) = \ln(h_t)$ .

Il modello GARCH-M esprime la media condizionale dei rendimenti  $\mu_t$  come funzione lineare della varianza condizionale, cioè:

$$\mu_t = X_t b + \delta g(h_t) \quad [3.23]$$

La [3.23] consente l'interpretazione del trade-off tra rischio e rendimento proprio delle teorie di portafoglio in quanto mette in evidenza un legame diretto tra  $\mu_t$  e  $h_t$ ; una relazione di questo tipo determina il fatto che l'aumento di uno dei due parametri implica l'aumento anche dell'altro.

Il GARCH-M consente anche un'interpretazione del premio per il rischio variabile nel tempo poiché, per ciascun valore della varianza condizionale, è possibile ottenere il corrispondente valore di  $\mu_t$ . Ponendo  $y_t \equiv R_t - R_f$  infatti risulta:

$$y_t | I_{t-1} = N(\mu_t, h_t) \quad [3.24]$$

Stime consistenti dei parametri di tali modelli richiedono che le funzioni relative alla media e alla varianza condizionale siano correttamente specificate e ottenute simultaneamente. Mentre nei modelli GARCH(p,q) è possibile una valutazione dei parametri separata e consistente (ma inefficiente), nei modelli GARCH-M ciò non accade poiché la matrice di informazione di Fisher non è diagonale a blocchi.

Bollerslev, Engle e Wooldridge (1988) hanno generalizzato questo modello in un contesto multivariato, ma il numero dei parametri incogniti aumenta sensibilmente al crescere del numero dei rendimenti. Per questa ragione esso viene applicato solo in presenza di poche attività.

### 3.4 Modelli con asimmetria

I modelli di asimmetria nascono dall'intuizione di Black (1976) il quale si era accorto che i movimenti dei rendimenti e quelli della volatilità sono negativamente correlati; quando si verifica uno shock negativo sui rendimenti la volatilità registra valori più elevati rispetto a quelli che si ottengono a seguito di uno shock positivo della stessa ampiezza. L'asimmetria si configura perciò nel fatto che la volatilità tende ad aumentare in corrispondenza delle "bad news", ovvero in quei periodi in cui il livello dei rendimenti è minore di quello atteso, e tende a diminuire quando si è in presenza di "good news". I modelli presentati in questo sottoparagrafo tengono conto di questo aspetto.

#### 3.4.1 Modello EGARCH

Per meglio rendere conto del fatto che gli shock sulla volatilità spesso non sono simmetrici, Nelson (1991) elabora il modello GARCH Esponenziale (EGARCH) che ha la seguente espressione analitica per la volatilità:

$$\log(h_t) = \omega + \sum_{j=1}^p \beta_j \log(h_{t-j}) + \sum_{i=1}^q \alpha_i g(u_t) \quad [3.25]$$

dove  $g(u_t) = \{\phi u_{t-i} + \gamma[|u_{t-i}| - E|u_{t-i}|]\}$  è i.i.d. con media nulla ed i parametri  $\omega$ ,  $\alpha_i$  e  $\beta_j$  non necessariamente devono essere non negativi. Tale modello non ha bisogno di restrizioni per assicurare la positività della varianza condizionale poiché il logaritmo rimuove tutti i vincoli di non negatività sui parametri. L'EGARCH è strettamente stazionario, ma non stazionario in covarianza quando risulta  $A(L) + B(L) = 1$ .

Se  $\phi \neq 0$  l'EGARCH coglie l'asimmetria degli effetti sulla volatilità provocati da shock con segni differenti infatti se  $u_t > 0$ ,  $g(u_t)$  è lineare con gradiente pari a  $(\phi + \gamma)$ , mentre se  $u_t < 0$ , la stessa equazione è sempre lineare, ma con gradiente pari a  $(\phi - \gamma)$ . Un altro elemento da tenere presente è il segno del prodotto  $\alpha_i \phi$ : nel caso in cui questo risultasse minore di zero si avrebbe una volatilità che cresce maggiormente quando si verificano shock di segno negativo rispetto all'eventualità in cui ci siano per essi valori positivi.

L'impatto dell'innovazione  $u_{t-i}$  sul valore di  $\log(h_t)$  è in accordo con ciò che accade sui mercati finanziari: la volatilità tende ad aumentare in presenza di notizie cattive che evidenziano valori dei rendimenti dei titoli inferiori a quelli attesi dagli operatori, mentre diminuisce quando le notizie sono buone.

Poiché  $E[\ln(h_t)] \neq \ln E(h_t)$  il principale limite dell'EGARCH risiede nella previsione della varianza condizionale in quanto esso è strutturato per analizzare il logaritmo di  $h_t$  e non  $h_t$  stessa.

### 3.4.2 Modello AGARCH

Nel modello GARCH le previsioni sui valori futuri della varianza sono in relazione lineare con quelli relativi al presente e al passato, mentre i quadrati delle innovazioni fungono da correttivo per le previsioni stesse. Ricorrendo alla trasformazione di Box e Cox (1964) per il modello Absolute GARCH (AGARCH), Hentschel (1994) generalizza un modello per la famiglia dei GARCH(1,1) dove compare una funzione  $f(u_t)$  elevata ad una generica potenza  $v$ : in questo modo la volatilità viene individuata in base ad una potenza di  $h_t^{1/2}$  piuttosto che da  $h_t^{1/2}$  stessa. Tale famiglia è definita dalla seguente funzione dove la deviazione standard condizionale è elevata alla potenza  $\lambda$ :

$$\frac{\sigma_t^\lambda - 1}{\lambda} = \omega + \beta \left( \frac{\sigma_{t-1}^\lambda - 1}{\lambda} \right) + \alpha \sigma_{t-1}^\lambda [f(u_{t-1})]^v \quad [3.26]$$

dove  $\sigma_t = h_t^{1/2}$ . Quando  $\lambda < 1$  la [3.26] è concava, mentre se  $\lambda > 1$  è convessa. La funzione  $f(u_t)$  contenuta al suo interno è la curva che mostra l'impatto che ogni nuovo evento ha sulla volatilità: essa ha sempre avere valore positivo per evitare che la volatilità sia minore di zero. Pagan e Schwert (1990) introducono la seguente espressione per tale funzione:

$$f(u_t) = |u_t - b| + c(u_t - b) \quad [3.27]$$

Sia il parametro di spostamento  $b$  che il coefficiente angolare  $c$  mostrano due differenti tipi di asimmetria e non devono essere nulli contemporaneamente<sup>8</sup>.

Dalla [3.26], facendo le opportune ipotesi sui parametri in essa contenuti, si possono ricavare molti altri modelli di tipo ARCH che contemplano un ritardo sia per  $\varepsilon_t^2$ , sia per  $h_t$ , infatti:

- Per  $\lambda = v = 2$ ,  $b = c = 0$  e  $\beta = 0$  si ha il modello ARCH di cui all'equazione [3.8]
- Per  $\lambda = v = 2$ ,  $b = c = 0$  si ottiene il modello GARCH di cui alla [3.19]
- Per  $v = 1$ ,  $b = 0$  e  $\lambda \rightarrow 0$  si ottiene l'equazione [3.25] relativa all'EGARCH
- Per  $\lambda = v = 1$  e  $|c| \leq 1$  si ha l'equazione dell'AGARCH
- Per  $\lambda = v = 1$  e  $b = c = 0$  si ha l'Absolute Value GARCH<sup>9</sup> (AVGARCH). Questo modello, introdotto da Taylor (1986) e Schwert (1989), rappresenta un'alternativa ai modelli GARCH esaminati in precedenza, poiché giunge ad effettuare previsioni non sulla varianza, bensì sulla deviazione standard; la previsione sulla deviazione standard è funzione lineare dei valori da essa assunti nel presente e nel passato e il termine  $u_t$  funge da correttivo per tale valore.
- Per  $\lambda = v = 2$  e  $c = 0$  si ottiene il Nonlinear Asymmetric GARCH (NGARCH) di Engle e Ng (1993) dato dalla seguente espressione analitica per la varianza condizionale:

$$h_t = \omega + \beta h_{t-1} + \alpha \left( \varepsilon_{t-1} + b \cdot h_{t-1}^{1/2} \right)^2 \quad [3.28]$$

- Per  $\lambda = v = 2$  e  $b = 0$  si ha il modello GJR di Glosten, Jagannathan e Runkle (1993) di equazione:

$$h_t = \omega + \beta h_{t-1} + \alpha \varepsilon_{t-1}^2 + \gamma S_t^- \varepsilon_{t-1}^2 \quad [3.29]$$

dove  $S_t^-$  è una variabile dummy che assume valore unitario quando  $\varepsilon_{t-1} < 0$ .

- Per  $\lambda = v = 1$ ,  $b = 0$  e  $|c| \leq 1$  si ha il Threshold GARCH (TGARCH) di Zakoian (1994) relativamente alla deviazione standard. L'equazione per la varianza condizionale è invece:

$$h_t = \omega + \beta h_{t-1} + \alpha^+ \varepsilon_{t-1}^+ + \alpha^- \varepsilon_{t-1}^- \quad [3.30]$$

dove  $\varepsilon_{t-1}^+ = \max\{\varepsilon_{t-1}, 0\}$  e  $\varepsilon_{t-1}^- = \min\{\varepsilon_{t-1}, 0\}$ .

- Asymmetric Power GARCH (APARCH) di Ding, Engle e Granger (1993) quando  $\lambda = v = 0$  e  $|c| \leq 1$

---

<sup>8</sup>In quest'ambito è importante sottolineare che mentre  $b \in [-\infty, +\infty]$ , deve risultare  $|c| \leq 1$ , per fare in modo che  $f(u_t)$  abbia sempre valore positivo

<sup>9</sup>Si veda in proposito Campbell, Lo e MacKinlay (1997).

### 3.4.3 Modello QGARCH

Esistono diversi altri contributi riguardo a modelli univariati di tipo ARCH che tengono conto delle asimmetrie negli shock sulla volatilità che esulano dalla [3.26]. Il più noto in letteratura è il modello Quadratic GARCH (QGARCH) introdotto da Sentana (1995), con equazione:

$$h_t = \sigma^2 + \phi'x_{t-q} + x'_{t-q}Ax_{t-q} + \sum_{j=1}^p \beta_j h_{t-1} \quad [3.31]$$

dove  $x_{t-q} = [\varepsilon_{t-1} \ \varepsilon_{t-2} \ \dots \ \varepsilon_{t-q}]'$ . Da questa formula, attraverso l'imposizione di opportuni vincoli sui parametri, è possibile ricavare:

- ARCH quando  $\phi = 0$ ,  $\beta_j = 0$  e  $A$  diagonale
- GARCH quando  $\phi = 0$  e  $A$  diagonale
- NGARCH quando  $A$  è diagonale

## 3.5 Altri modelli univariati di tipo ARCH

La letteratura sui modelli univariati di tipo ARCH ha avuto un notevole sviluppo negli anni grazie talvolta anche all'introduzione di modelli "esotici" che si occupano di aspetti particolari di tali processi<sup>10</sup>. In questo paragrafo sono presentati i processi ideati nell'ambito delle problematiche della non stazionarietà, dei processi frazionalmente integrati e dei modelli con cambio di regime.

### 3.5.1 Modello IGARCH

Si consideri per semplicità un GARCH(1,1): supponendo di voler prevedere il valore atteso della volatilità per il periodo  $t+s$ , sostituendo ricorsivamente all'interno della [3.19] si ha:

$$E(h_{t+k}|I_t) = \omega \left[ \sum_{i=0}^{k-1} (\alpha + \beta)^i \right] + (\alpha + \beta)^k h_t \quad [3.32]$$

Quando risulta che  $(\alpha + \beta) < 1$ , si nota che per  $s \rightarrow \infty$  essa converge al valore non condizionale della varianza dato da  $(1 - \alpha - \beta)^{-1}\omega$  ed il processo è stazionario in covarianza. Se invece  $(\alpha + \beta) = 1$  come limite si ottiene la quantità:

$$E(h_{t+s}|I_t) = h_t + k \cdot \omega \quad [3.33]$$

<sup>10</sup>Dal lavoro pionieristico di Engle (1982) fino ad oggi la letteratura ha fornito diverse varianti sul tema a testimonianza dell'enorme successo di tali strumenti nell'ambito della ricerca scientifica culminato nel 2003 con l'assegnazione del Premio Nobel per l'Economia a C.W.J. Granger e allo stesso R.F. Engle. Un esempio recente di contributo alla letteratura è quello relativo all'Hyperbolic GARCH di Davidson (2003).

Generalizzando, quando  $A(L) + B(L) < 1$ , uno shock futuro sulla varianza condizionale tende a generare un effetto che tende ad annullarsi man mano che aumenta l'orizzonte previsivo, mentre se  $A(L) + B(L) = 1$  gli shock futuri sulla volatilità sono altamente persistenti e si riflettono nei suoi valori stimati.

Il modello Integrated GARCH (IGARCH) introdotto da Engle e Bollerslev (1986) si ottiene in quest'ultimo caso, cioè quando il polinomio autoregressivo della [3.15] ammette radice unitaria<sup>11</sup>.

Nel modello IGARCH(1,1) ad esempio risulta  $(\alpha + \beta) = 1$ : in questo modo la volatilità di un periodo influenza tutte le previsioni relative a qualsiasi periodo successivo. Un dato shock sulla varianza condizionale risulta persistente per qualsiasi orizzonte temporale futuro, diventando una componente rilevante nel lungo periodo.

Il processo di cui alla [3.33] è simile a quello di un random walk lineare con drift  $\omega$ , ma Nelson (1990) mostra che tale analogia va esaminata tenuto conto di alcune sostanziali differenze: in primo luogo l'IGARCH(1,1) è un processo stazionario in senso stretto per  $h_t$ , mentre il random walk non gode di questa proprietà. In secondo luogo quest'ultimo non è stazionario nelle covarianze poiché non possiede né momento primo né momento secondo non condizionali; nell'IGARCH(1,1) questa condizione circa  $h_t$  è generalmente, ma non necessariamente rispettata.

Quando  $\omega = 0$  è molto semplice dimostrare che la volatilità è una martingala con valori vincolati dal fatto che, essendo un quadrato, essa non può essere negativa, infatti si ha  $E(h_{t+k}|I_{t-1}) = h_t$ . Da questa espressione si deduce che l'IGARCH(1,1) è un processo stazionario in senso stretto.

### 3.5.2 Modello FIGARCH

Il modello Fractionally Integrated GARCH (FIGARCH) è stato introdotto da Baillie, Bollerslev e Mikkelsen (1996) per rispondere alla domanda se la varianza condizionale è caratterizzata da memoria lunga<sup>12</sup>. Ding, Engle e Granger (1993) hanno mostrato che nelle serie storiche finanziarie le autocorrelazioni dei quadrati dei residui decrescono nel tempo in maniera più lenta rispetto al caso esponenziale tipico dei modelli GARCH(p,q) ed IGARCH(p,q)<sup>13</sup>, quindi nasce l'esigenza di elaborare uno strumento che sia in grado di analizzare l'alto grado di persistenza talvolta riscontrato nella volatilità.

Dato che una classe di modelli in grado di interpretare questo requisito è data dai processi frazionalmente integrati<sup>14</sup> che costituiscono l'insieme di tutti i casi intermedi

<sup>11</sup>Esiste anche la possibilità nel quale  $A(L) + B(L) > 1$ : in questo caso il limite della [3.32] assume valore infinito e di conseguenza lo shock sulla volatilità diviene esplosivo.

<sup>12</sup>Il concetto di memoria lunga nelle serie storiche è stato introdotto inizialmente da Granger e Joyeux (1980) e Hosking (1981).

<sup>13</sup>Bollerslev (1988) ha infatti mostrato che la funzione di autocorrelazione dei quadrati dei residui decresce esponenzialmente per un GARCH(1,1), mentre Ding e Granger (1996) sono giunti alle stesse conclusioni per il processo IGARCH.

<sup>14</sup>Nell'ambito dei processi ARMA(p,q), i processi frazionalmente integrati di ordine  $d$  o ARFIMA(p,d,q) sono quei processi  $\{y_t\}$  per i quali vale la relazione:

$$A(L)(1 - L)^d y_t = B(L)\varepsilon_t$$

tra quelli a radice unitaria e quelli integrati di ordine 0, il FIGARCH(p,d,q) è ottenuto estendendo ai modelli di tipo ARCH le caratteristiche proprie di tali processi. Analiticamente esso si ottiene semplicemente definendo il polinomio nell'operatore ritardo come segue:

$$1 - A(L) - B(L) = f(L)(1 - L)^d \quad [3.34]$$

dove le radici che annullano  $f(L)$  sono maggiori di uno. Analogamente ai modelli ARFIMA  $(1 - L)^d$  è l'operatore differenza frazionale con  $0 < d < 1$ . In riferimento alla [3.15], il processo FIGARCH(p,d,q) si configura perciò come segue:

$$f(L)(1 - L)^d \varepsilon_t^2 = \omega + [1 - B(L)]v_t \quad [3.35]$$

Come tutti i processi integrati il FIGARCH è mean-reverting e si pone perciò a metà strada tra due casi estremi: quando  $d = 0$  si ha un modello GARCH(p,q) il quale prevede una caduta esponenziale della volatilità nel tempo, caratteristica che lo rende maggiormente efficace per previsioni di breve periodo. Nel caso in cui  $d = 1$  il grado di persistenza della volatilità risulta massimo e si ottiene un IGARCH(p,q). Quando invece  $0 < d < 1$  si ottiene uno strumento flessibile soprattutto per quanto riguarda le sue implicazioni nel lungo periodo in quanto le previsioni per la varianza condizionale decrescono nel tempo ad un tasso iperbolico.

Rispetto ai modelli ARFIMA il FIGARCH si distingue per due ragioni: innanzi tutto il polinomio nell'operatore ritardo non è applicato alle innovazioni, bensì ai loro quadrati e in secondo luogo, anche quando  $0 < d < 0.5$ , il processo non risulta essere stazionario in covarianza<sup>15</sup>.

Bollerslev e Mikkelsen (1996) introducono il modello FIEGARCH determinando così un processo a memoria lunga che tenga conto dell'asimmetria generata da shock sui rendimenti di segno opposto.

### 3.5.3 Modello SWARCH

I modelli di tipo ARCH introdotti finora non sono in grado di catturare l'effetto generato da break strutturali<sup>16</sup> nei dati causati da eventi di natura eccezionale come le guerre, le recessioni o variazioni nella politica monetaria di un paese.

Hamilton (1990) ha introdotto un modello basato sul fatto che lo stato o regime in cui si trova l'economia varia al trascorrere del tempo e che tutte le osservazioni campionarie sono influenzate da tali variazioni. Le ipotesi su cui si basa sono:

1. Si introduce nell'equazione del modello di regressione una variabile casuale discreta non osservabile ( $s_t$ ) detta variabile di "stato" o di "regime" del processo

dove  $0 < d < 1$ . Negli ARFIMA la funzione di autocorrelazione decresce iperbolicamente quindi più lentamente rispetto al caso dei modelli ARMA dove questa decresce esponenzialmente. Si veda in proposito Hamilton (1994).

<sup>15</sup>In effetti una delle più importanti proprietà dei modelli ARFIMA evidenziate in Hosking (1981) è il fatto che per  $0 < d < 0.5$  essi sono stazionari in covarianza.

<sup>16</sup>Con il termine "break strutturale" si intende un salto registrato all'interno della serie storica in esame nel passaggio da un'osservazione alla successiva. In corrispondenza di tale salto è possibile dividere le osservazioni in due sottocampioni che mostrano sensibili differenze nei loro momenti.

al tempo  $t$ . A seconda del valore assunto da la serie storica in esame mostra un differente andamento nei suoi sottocampioni, quindi ogni eventuale break può essere agevolmente interpretato.

2. L'economia passa da un regime all'altro nel corso del tempo, quindi accade spesso di incontrare delle fratture all'interno delle serie storiche.
3. Il regime in cui si trova l'economia può essere descritto come il risultato di una Catena di Markov a  $N$  stati.

Alla luce di queste considerazioni Cai (1994) e Hamilton e Susmel (1994) hanno introdotto lo Switching Regime ARCH (SWARCH) generalizzando il modello con cambio di regime permettendo alle innovazioni di seguire un processo di tipo ARCH.

L'equazione della varianza condizionale nello SWARCH(N,q) è la seguente:

$$\frac{h_t}{\gamma_{s_t}} = \omega + \sum_{i=1}^q \alpha_i \frac{\varepsilon_{t-i}}{\gamma_{s_{t-i}}} \quad [3.36]$$

dove  $\gamma_{s_{t-i}}$  sono i parametri di scala che catturano i cambi di regime. Il parametro  $\gamma_{s_t}$  è posto al valore 1 per l'identificazione del modello.



La volatilità analizzata nell'ambito dei modelli univariati è quella relativa ad un determinato titolo, mentre i suoi movimenti in relazione ad altri titoli non è considerata. Lavorare con più titoli contemporaneamente significa apportare alcune modifiche sostanziali a tali modelli in modo che siano in grado di interpretare una realtà costituita da più di un valore che varia nel tempo.

In un contesto multivariato bisogna quindi tenere conto dei seguenti aspetti:

1. Esiste un vettore di  $n$  componenti ciascuna associabile ad un titolo presente sul mercato.
2. Le innovazioni inerenti a ciascun titolo sono contenute nel vettore  $\varepsilon_t$  con media nulla.
3. Le covarianze condizionali del vettore  $\varepsilon_t$  sono contenute all'interno di una matrice  $n \times n$  simmetrica e semidefinita positiva ( $\Omega_t$ ).
4. Analogamente a quanto avveniva nei modelli univariati di tipo ARCH, il processo  $\varepsilon_t$  è definito come:

$$\varepsilon_t = \Omega_t^{1/2} u_t \quad [4.1]$$

dove  $u_t$  è un vettore  $n$ -dimensionale generalmente con distribuzione multinormale standardizzata. Da questa affermazione segue che la distribuzione dell'innovazione condizionale ad un dato set informativo al tempo  $t - 1$  è data da  $\varepsilon_t | I_{t-1} \sim MN(0, \Omega_t)$ .

5. Gli elementi della matrice  $\Omega_t$  dipendono dai valori assunti da essa nel passato e da quelli assunti dal prodotto incrociato  $\varepsilon_t \varepsilon_t'$ .

Con i modelli multivariati entrano in gioco anche le covarianze condizionali delle innovazioni che coi modelli univariati non erano state mai analizzate; attraverso lo studio della loro struttura si riesce a raggiungere una maggiore efficienza nelle stime ed una migliore interpretazione dei movimenti comuni che caratterizzano i rendimenti.

L'analisi condotta sulle covarianze della matrice  $\Omega_t$  risulta di fondamentale importanza soprattutto nell'ottica dei modelli di portafoglio: un soggetto intenzionato a minimizzare la componente rischio all'interno del proprio portafoglio ovviamente investirà su alcune attività  $X$  ed  $Y$  per le quali il valore della covarianza condizionale è negativo.

#### 4.1 Modello Vech

Bollerslev, Engle e Wooldridge (1988), partendo dal modello GARCH(p,q) di cui alla [3.13] sostengono che la più logica generalizzazione in un contesto multivariato per lo studio della volatilità sembra essere la seguente:

$$\text{vech}(\Omega_t) = W + A(L)\text{vech}(\varepsilon_t\varepsilon_t') + B(L)\text{vech}(\Omega_t) \quad [4.2]$$

quindi

$$\text{vech}(\Omega_t) = W + \sum_{i=1}^q A_i \text{vech}(\varepsilon_{t-i}\varepsilon_{t-i}') + \sum_{j=1}^p B_j \text{vech}(\Omega_{t-j}) \quad [4.3]$$

dove l'operatore *vech* considera solo la parte triangolare bassa di una matrice quadrata simmetrica disponendola all'interno di un vettore colonna. Il vettore così costruito contiene quindi tutte le componenti non ripetute della matrice stessa.

Si consideri ad esempio il caso in cui  $p = q = 1$  che determina un GARCH(1,1) multivariato così definito:

$$\text{vech}(\Omega_t) = W + A\text{vech}(\varepsilon_{t-1}\varepsilon_{t-1}') + B\text{vech}(\Omega_{t-1}) \quad [4.4]$$

dove il vettore  $W$  e le matrici  $A$  e  $B$  sono definite come nella [4.3]. Gli inconvenienti di questo modello sono il rapido aumento del numero dei parametri da stimare al crescere della numerosità dei titoli considerati e soprattutto il fatto che non è possibile vincolare le matrici  $A_i$  e  $B_j$  in modo da ottenere una matrice delle varianze e delle covarianze sicuramente definita positiva. Si nota subito infatti che in questo modello se  $n = 3$  si devono stimare 78 parametri, se  $n = 4$  il loro numero sale a 210, se  $n = 5$  diventano 465 e così via. Nella [4.3] i parametri da stimare sono contenuti nel vettore  $W$  di dimensione  $n(n+1)/2$  e da tutte le matrici del tipo  $A_i$  e  $B_j$  ognuna con  $n(n+1)/2 \times n(n+1)/2$  elementi da stimare. Il numero totale dei parametri ammonta perciò a  $n(n+1)/2[1 + n(p+q)(n+1)/2]$ , numero che aumenta alla quarta potenza di  $n$  e che rende ben presto tutta la struttura incontrollabile anche in presenza di valori bassi per  $n$ ,  $p$  e  $q$ . Ovviamente, se all'interno del modello è contemplata anche la parte "in mean", occorre aggiungere al numero totale dei parametri anche quelli ad essa relativi.

Per fare in modo che il totale dei parametri non esploda al crescere della numerosità del campione osservato, la letteratura ha dovuto porre dei vincoli di diversa natura alla [4.3]. Il modello GARCH diagonale costituisce un valido esempio in questo senso<sup>1</sup>.

<sup>1</sup>Si veda in proposito il sottoparagrafo 4.2.

Supponendo che ciascuna matrice del tipo  $B_j$  sia composta solo da zeri in ogni sua componente, si ottiene il modello ARCH Multivariato.

Il modello Vech non garantisce che la matrice  $\Omega_t$  sia semidefinita positiva, ma è stazionario in covarianza quando la persistenza in varianza ha distribuzione limite pari a zero, cioè, per  $\forall s > 0$ , deve risultare:

$$\hat{\Omega}_t = E[\text{vech}(\Omega_t)|I_s] - E[\text{vech}(\Omega_t)|I_0] \xrightarrow{d} 0 \quad [4.5]$$

Tale relazione è soddisfatta solo quando la matrice  $\Omega_t$  tende verso i valori della matrice non condizionata delle varianze e delle covarianze, cioè quando gli autovalori della matrice  $\sum_{i=1}^q A_i + \sum_{j=1}^p B_j$  sono minori dell'unità.

Ovviamente, se dovesse risultare il caso in cui la disuguaglianza non è stretta, si ricadrebbe nell'IGARCH(p,q) multivariato. Se invece gli autovalori di tale matrice eccedono l'unità si avrebbero shock esplosivi sulla volatilità che non attenuano i loro effetti nel tempo.

## 4.2 Modello GARCH diagonale

Introdotta da Bollerslev, Engle e Wooldridge (1988) semplicemente imponendo diagonali le matrici  $A_i$  e  $B_j$  del modello Vech, il GARCH(p,q) diagonale è costruito in modo che ciascun elemento della matrice condizionale delle varianze e covarianze segua un processo GARCH univariato. Il generico elemento  $\sigma_{mn}$  della matrice  $\Omega_t$  dipende perciò dagli elementi posti sulla  $m$ -esima riga e  $n$ -esima colonna di ciascuna delle  $q$  matrici relative ai valori passati del prodotto  $\varepsilon_t \varepsilon_t'$  e delle  $p$  matrici relative ai valori assunti in passato dalla stessa matrice delle varianze e delle covarianze.

Analiticamente si ha perciò:

$$\sigma_{mn,t} = w_{mn} + \sum_{i=1}^q \alpha_{i,mn} \varepsilon_{m,t-i} \varepsilon'_{n,t-i} + \sum_{j=1}^p \beta_{j,mn} \sigma_{mn,t-1} \quad [4.6]$$

Le condizioni necessarie e sufficienti affinché la matrice condizionale delle varianze e delle covarianze sia definita positiva derivano dalle proprietà del prodotto di Hadamard<sup>2</sup>. Nel caso in cui nella [4.3] le matrici dei parametri siano definite come  $A_i = \text{diag}[\text{vech}(\bar{A}_i)]$  e  $B_j = \text{diag}[\text{vech}(\bar{B}_j)]$ , il modello GARCH(p,q) diagonale può

<sup>2</sup>Il prodotto di Hadamard tra due matrici  $A$  e  $B$  aventi le stesse dimensioni è definito come segue:

$$A \odot B = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \dots & a_{1n}b_{1n} \\ a_{21}b_{21} & a_{22}b_{22} & \dots & a_{2n}b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}b_{m1} & a_{m2}b_{m2} & \dots & a_{mn}b_{mn} \end{bmatrix}$$

essere espresso come<sup>3</sup>:

$$\Omega_t = \bar{W} + \sum_{i=1}^q \bar{A}_i \odot (\varepsilon_{t-i} \varepsilon'_{t-i}) + \sum_{j=1}^p \bar{B}_j \odot \Omega_{t-j} \quad [4.7]$$

dove le matrici  $\bar{W}$ ,  $\bar{A}_i$  e  $\bar{B}_j$  sono simmetriche di dimensione  $n \times n$  e  $W = \text{vech}(\bar{W})$ . Per le proprietà del prodotto di Hadamard dalla [4.7] segue che la matrice  $\Omega_t$  è definita positiva per  $\forall t$  se lo è anche  $\bar{W}$  e le matrici  $\bar{A}_i$  e  $\bar{B}_j$  sono almeno semidefinite positive per  $i = 1, 2, \dots, q$  e  $j = 1, 2, \dots, p$ .

Il modello GARCH(p,q) diagonale presenta  $(1 + p + q)$  parametri da stimare per ciascun elemento di  $\Omega_t$ , quindi  $[n(n+1)/2](1 + p + q)$  parametri in tutto. Nonostante il numero dei parametri incogniti sia ridotto rispetto al modello Vech, tuttavia la struttura di questo modello risulta essere piuttosto rigida e non garantisce una matrice condizionale delle varianze e delle covarianze definita positiva a meno che non siano verificate le condizioni su cui la [4.7] si basa. Le condizioni di stazionarietà in covarianza sono le stesse viste per il modello Vech.

### 4.3 Modello BEKK

Formalizzato in Engle e Kroner (1995), il BEKK supera il problema posto dai precedenti modelli in quanto gli elementi di  $\Omega_t$  scaturiscono dalla somma di forme quadratiche, quindi tale matrice risulta sempre definita positiva sotto blande condizioni sui parametri. Tale modello, che prende il proprio nome dalle iniziali di coloro che hanno partecipato alla sua stesura<sup>4</sup>, assume la seguente forma analitica:

$$\Omega_t = C' C + \sum_{k=1}^K \sum_{i=1}^q A'_{ki} \varepsilon_{t-i} \varepsilon'_{t-i} A_{ki} + \sum_{k=1}^K \sum_{j=1}^p B'_{kj} \Omega_{t-j} B_{kj} \quad [4.8]$$

dove  $C$  è una matrice triangolare alta e  $K$  indica il grado di generalità del modello, cioè una componente utile per poter ottenere una qualsiasi matrice definita positiva nell'ambito dello spazio parametrico<sup>5</sup>.

Dato che la matrice  $C$  è composta di  $n(n+1)/2$  parametri, mentre le matrici  $A_{ki}$  e  $B_{kj}$  ne hanno  $n^2$  al loro interno, il totale dei parametri da stimare ammonta a  $n(n+1)/2 + K(p+q)n^2$ .

Engle e Kroner (1995) hanno mostrato alcune importanti proprietà del BEKK:

- a) Per le regole del prodotto di Kronecker<sup>6</sup> esso può essere facilmente ricondotto nella forma della [4.3] utilizzando l'operatore *vec* al posto del *vech* per tutte le matrici relative ai prodotti incrociati delle innovazioni e alle matrici

<sup>3</sup>cfr. Bollerslev, Engle e Nelson (1994).

<sup>4</sup>La sigla BEKK infatti deriva dalle iniziali dei suoi ideatori Baba, Engle, Kraft e Kroner.

<sup>5</sup>Nella pratica  $K = 1$  soprattutto per evitare che il numero dei parametri del BEKK sia troppo elevato.

<sup>6</sup>In particolare  $\text{vec}(ABC) = (C' \otimes A) \text{vec}(B)$ .

condizionali delle varianze e delle covarianze passate, ottenendo la seguente espressione:

$$\begin{aligned} \text{vec}(\Omega_t) = & W + \sum_{k=1}^K \sum_{i=1}^q (A'_{ki} \otimes A'_{ki}) \text{vec}(\varepsilon_{t-i} \varepsilon'_{t-i}) + \\ & + \sum_{k=1}^K \sum_{j=1}^p (B'_{kj} \otimes B'_{kj}) \text{vec}(\Omega_{t-1}) \end{aligned} \quad [4.9]$$

dove  $W = (C' \otimes C') \text{vec}(I)$  è un vettore di dimensione  $n^2$ . Da un BEKK si può ricavare un unico modello  $\text{vec}$ , mentre il passaggio inverso non conduce alla stessa conclusione<sup>7</sup>.

- b) Ogni modello GARCH(p,q) diagonale ha una sua rappresentazione di tipo BEKK semplicemente ponendo le matrici dei parametri di quest'ultimo in forma diagonale<sup>8</sup>. L'espressione che si ottiene è del tutto simile alla [4.7], infatti:

$$\Omega_t = \bar{W} + \sum_{k=1}^K \sum_{i=1}^q \alpha_{ki} \alpha'_{ki} \odot (\varepsilon_{t-i} \varepsilon'_{t-i}) + \sum_{k=1}^K \sum_{i=1}^p \beta_{kj} \beta'_{kj} \odot \Omega_{t-j} \quad [4.10]$$

- c) Deboli vincoli sulle matrici dei parametri garantiscono a  $\Omega_t$  di essere definita positiva per ogni  $t$ : ciò accade perché ogni suo elemento scaturisce dalla sommatoria di forme quadratiche.
- d) Dalla [4.9] si evince che il BEKK è stazionario in covarianza quando sono strettamente minori di 1 tutti gli autovalori della matrice

$$\sum_{k=1}^K \sum_{i=1}^q (A'_{ki} \otimes A'_{ki}) + \sum_{k=1}^K \sum_{j=1}^p (B'_{kj} \otimes B'_{kj}) \quad [4.11]$$

Da questa condizione deriva che la matrice non condizionata delle varianze e delle covarianze calcolata attraverso il BEKK è data da:

$$\Omega = \left[ I - \sum_{k=1}^K \sum_{i=1}^q (A'_{ki} \otimes A'_{ki}) + \sum_{k=1}^K \sum_{j=1}^p (B'_{kj} \otimes B'_{kj}) \right]^{-1} \text{vec}(C'C) \quad [4.12]$$

<sup>7</sup>Tale passaggio al BEKK è possibile solo nel caso in cui per ciascuna matrice dei parametri del modello  $\text{vec}$  è scomponibile nella forma  $\tilde{A} = A' \otimes A'$ ; tuttavia il rispetto di questa condizione è sempre garantito solo imponendo alcuni vincoli sui parametri del BEKK ed è per questo motivo in questo modello il loro numero risulta essere inferiore rispetto al modello  $\text{vec}$ .

<sup>8</sup>McCurdy e Stegans (1992) ad esempio hanno utilizzato questa proprietà ponendo  $K = p = q = 1$ .

#### 4.4 Modello ARCH fattoriale

Il modello ARCH fattoriale introdotto da Engle, Ng e Rothschild (1990) rappresenta un metodo alternativo per la parametrizzazione della matrice condizionale delle varianze e delle covarianze in quanto ha una struttura simile all'APT di Ross (1976); dato un vettore  $y_t$  contenente i rendimenti di  $n$  asset al tempo  $t$  si ha:

$$y_t = \mu_t + Bf_t + \varepsilon_t \quad [4.13]$$

dove  $f_t$  è il vettore con  $k$ -dimensionale che racchiude al suo interno i fattori,  $B$  è una matrice data di dimensione  $n \times k$  relativa ai pesi dei singoli fattori e  $\varepsilon_t$  è il vettore degli  $n$  shock idiosincratici.

Le ipotesi alla base di questo modello sono:

1. Gli shock idiosincratici hanno varianza condizionale costante pari a  $E(\varepsilon_t \varepsilon_t') = \Psi$ .
2. La matrice condizionale delle varianze e delle covarianze dei fattori invece è indicizzata rispetto al tempo ed è posta pari a  $\Lambda_t$ .
3. Vale la relazione di incorrelazione  $E(f_t \varepsilon_t) = 0$ .

Applicando il modello ARCH fattoriale si applica una scomposizione della matrice  $\Omega_t$  in due componenti distinte, la prima riconducibile ad un numero  $k$  di fattori e la seconda idiosincratica, imputabile al rischio proprio di ciascun titolo<sup>9</sup>. Analiticamente si ha:

$$\Omega_t = B\Lambda_t B' + \Psi \quad [4.14]$$

Se  $\Lambda_t$  è diagonale risulta:

$$\Omega_t = \sum_{i=1}^k \beta_i \beta_i' \lambda_{it} + \Psi \quad [4.15]$$

dove  $\beta_i$  rappresenta l' $i$ -esima colonna della matrice  $B$ , mentre  $\lambda_{it}$  è l' $i$ -esimo elemento sulla diagonale principale di  $\Lambda_t$ . Ci sono quindi  $k$  vettori di dimensione  $n$  che definiscono l'intera matrice condizionale delle varianze e delle covarianze dei rendimenti.

Si considerino i portafogli  $z_{it} = \Phi' y_t$  dove  $\Phi$  ha dimensione  $n \times n$  e valgano le condizioni:

$$\begin{cases} \phi_j' \beta_i = 1 & \text{se } i = j \\ \phi_j' \beta_i = 0 & \text{se } i \neq j \end{cases}$$

dove  $\phi_j$  è la  $j$ -esima colonna di  $\Phi$ . La varianza condizionale di  $z_{it}$  è perciò:

$$\Phi' \Omega_t \Phi = \tilde{B} \Lambda_t \tilde{B}' + \tilde{\Psi} \quad [4.16]$$

---

<sup>9</sup>cfr. Palm (1996).

dove  $\bar{\Psi} = \Phi' \Psi \Phi$  è costante per definizione e la matrice  $\tilde{B} = \Phi' B$  ha dimensione  $n \times k$  e risulta così partizionata:

$$\tilde{B} = \begin{bmatrix} I_{k \times k} \\ - - - \\ 0_{(n-k) \times k} \end{bmatrix}$$

Alla luce di questa riformulazione i portafogli  $z_{it}$  hanno la stessa variazione temporale che subiscono i fattori<sup>10</sup> e soprattutto la [4.14] dà luogo a due implicazioni: in primo luogo se  $k < n$  si possono costruire  $n - k$  portafogli come combinazione lineare del vettore  $y_t$  che hanno varianza costante, mentre i rimanenti  $k$  portafogli hanno varianza condizionale pari a  $\Lambda_t$  con l'aggiunta di una costante. La seconda implicazione è che esistono portafogli dove il vettore dei pesi dei titoli è ortogonale a tutte le colonne di  $B$ , tranne una. In termini pratici tutto ciò si riflette nella capacità dell'ARCH fattoriale di interpretare il fatto che un soggetto, mediante le sue scelte, ha la possibilità di investire in portafogli con varianza costante oppure in portafogli dove il rischio dipende dalla variabilità di alcuni fattori.

Per stimare  $\Omega_t$  Engle, Ng e Rothschild (1990) hanno imposto che ciascun elemento della diagonale di  $\Lambda_t$  segua un modello univariato GARCH(1,1) del tipo:

$$\lambda_{it} = \omega + \alpha_i (\phi_i' \varepsilon_{t-1})^2 + \gamma_i \lambda_{it-1} \quad [4.17]$$

Sostituendo nella [4.15] si ottiene un'espressione per l'ARCH fattoriale che si configura come caso speciale del BEKK di cui all'equazione [4.8], infatti:

$$\Omega_t = \bar{\Psi} + \sum_{i=1}^k \alpha_i (\beta_i \phi_i' \varepsilon_{t-1} \varepsilon_{t-1}' \phi_i \beta_i') + \sum_{i=1}^k \gamma_i \beta_i \phi_i' \Omega_{t-1} \phi_i \beta_i' \quad [4.18]$$

Anche se attraverso l'applicazione di questo modello è possibile esprimere la matrice condizionale delle varianze e delle covarianze attraverso le covarianze di un numero ristretto di fattori, tuttavia il problema dell'individuazione dei fattori costituisce la causa di un suo scarso utilizzo nelle analisi empiriche.

#### 4.5 Modello con correlazioni condizionali costanti

All'interno della letteratura sui modelli multivariati di tipo ARCH uno dei problemi più rilevanti è l'aumento del numero di parametri al crescere del numero di variabili che determinano ulteriori complicazioni durante la fase della stima. Una soluzione in questo senso viene dal modello con correlazioni condizionali costanti (CCC) elaborato da Bollerslev (1990) all'interno del quale la matrice condizionale delle varianze e delle covarianze variabile nel tempo è parametrizzata in modo da risultare proporzionale al prodotto delle corrispondenti deviazioni standard; in questo modo, oltre all'ottenimento di una procedura di stima meno complessa, si ottiene una matrice  $\Omega_t$  definita positiva sotto blande condizioni.

Le ipotesi formulate da Bollerslev in proposito sono:

<sup>10</sup>Per questo motivo  $z_{it}$  è il vettore contenente i portafogli dei fattori rappresentativi.

1. Si considera la matrice diagonale  $D_t$  di dimensione  $n \times n$ ; sulla sua diagonale vengono poste tutte le varianze condizionali  $h_{iit}$ .
2. Si considera inoltre la matrice  $R_t$  delle correlazioni condizionate, anch'essa di dimensione  $n \times n$ , tale che ogni coefficiente di correlazione soddisfi la relazione  $\rho_{ijt} = h_{ijt}(h_{iit}h_{jjt})^{-1/2}$ .
3. Poiché il generico elemento di  $\Omega_t$  risulta essere  $h_{ijt} = \rho_{ijt}(h_{iit}h_{jjt})^{1/2}$  la sua variazione nel tempo può dipendere da quella del coefficiente di correlazione oppure da quella delle singole varianze condizionali. L'imposizione della matrice  $R$  costante significa perciò ipotizzare che nel corso del tempo la struttura delle covarianze condizionali dipende esclusivamente dalle varianze condizionali.

Date tali ipotesi, in termini generali risulta:

$$\Omega_t = D_t^{1/2} R D_t^{1/2} \quad [4.19]$$

Dato che le varianze condizionali poste sulla diagonale principale di  $D_t$  sono tutte positive per definizione, se la matrice  $R$  è definita positiva anche  $\Omega_t$  lo sarà per ogni  $t$ . In questo caso è agevole calcolare:

$$\Omega_t^{-1} = D_t^{-1/2} R^{-1} D_t^{-1/2} \quad [4.20]$$

Ipotizzando che la varianza condizionale del rendimento di ciascun titolo segua un processo univariato GARCH(p,q) e che la covarianza dei rendimenti di una qualsiasi coppia di titoli sia data dalle deviazioni standard dei rendimenti stessi moltiplicate per coefficienti costanti di correlazione<sup>11</sup>, si ricava il seguente modello:

$$\begin{cases} h_{iit} = \omega_i + \sum_{r=1}^q \alpha_{ir} \varepsilon_{it-r}^2 + \sum_{r=i}^p \beta_{ir} h_{iit-r} \\ h_{ijt} = \rho_{ij} (h_{iit} h_{jjt})^{1/2} \end{cases} \quad [4.21]$$

dove, per le ipotesi del modello, i coefficienti di correlazione condizionale non sono indicizzati rispetto al tempo. Il numero totale dei parametri incogniti ammonta a  $n(1 + p + q) + n(n + 1)/2$ .

I vantaggi che si hanno mediante l'utilizzo del modello CCC sono la matrice  $\Omega_t$  definita positiva, la riduzione del numero dei parametri incogniti rispetto agli altri modelli multivariati di tipo ARCH e la semplificazione nel processo di stima. Se le varianze condizionali poste lungo la diagonale di  $D_t$  sono tutte positive e la matrice  $R$  è definita positiva il primo risultato è verificato senza la necessità di imporre particolari vincoli. L'imposizione di una matrice  $R$  costante nel tempo consente di combinare più modelli univariati di tipo ARCH all'interno di un contesto

<sup>11</sup>Nel lavoro originario di Bollerslev (1990) si ipotizza che ciascuna varianza condizionale segua un processo GARCH(1,1).



multivariato e ciò permette la stima di un numero contenuto di parametri anche in presenza valori elevati di  $n$ . Infine, considerando che all'interno della funzione log-verosimiglianza è richiesta l'inversione ed il calcolo del determinante della matrice  $\Omega_t$  per ogni periodo  $t$ , l'applicazione della [4.19] semplifica notevolmente la fase del calcolo, infatti:

- poiché  $R$  è costante, ponendo  $\tilde{\varepsilon}_t = D_t^{-1/2}\varepsilon_t$ , il numero di inversioni di matrici si riduce sensibilmente,
- il calcolo di  $|\Omega_t|$  per  $\forall t$  si riduce ad un'unica determinazione di  $|R|$ , mentre è noto che:

$$|D_t| = \prod_{i=1}^n h_{iit}$$

Il modello CCC è stato oggetto di discussione in letteratura perché la condizione di correlazioni condizionali costanti costituisce una forte imposizione all'interno della struttura delle covarianze condizionali; il modello oggetto del prossimo sottoparagrafo si propone di superare il problema consentendo anche alla matrice  $R$  di essere variabile nel tempo.

#### 4.6 Modello con correlazioni condizionali dinamiche

Il modello con correlazioni condizionali dinamiche (DCC) proposto in Engle e Shepard (2001) e successivamente in Engle (2002), aggiunge al modello CCC una dinamica seppur limitata alla matrice delle correlazioni introducendo la stessa struttura GARCH univariata per ogni suo elemento; in questo modo viene proposta una nuova classe di stimatori che mantengono le stesse proprietà del modello CCC, ma nel contempo parametrizzano la correlazione direttamente permettendole di essere variabile nel tempo.

Dato un vettore  $y_t$   $n$ -dimensionale, nel modello DCC la matrice condizionale delle varianze e delle covarianze è data dalla relazione:

$$\Omega_t = D_t^{1/2} R_t D_t^{1/2} \tag{4.22}$$

La differenza rispetto all'equazione [4.19] del modello di Bollerslev (1990) risiede nel fatto che la matrice delle correlazioni  $R_t$  non è costante nel tempo e questo si traduce in una stima a due stadi dell'intero processo. Nel primo stadio della stima del DCC si ipotizza che gli elementi della matrice diagonale  $D_t$  seguano un processo GARCH del tipo:

$$h_{iit} = \alpha_{0i} + \sum_{s=1}^{Q_i} \alpha_s y_{it-s}^2 + \sum_{s=1}^{P_i} \beta_s h_{iit-s} \tag{4.23}$$

dove  $Q_i$  e  $P_i$  rappresentano il numero di ritardi del GARCH propri di ciascuna serie dei rendimenti<sup>12</sup>.

<sup>12</sup>In questa fase la specificazione del modello per le varianze condizionali potrebbe essere un qualsiasi modello univariato di tipo ARCH.

Una volta stimata la matrice  $D_t$  si determinano i residui standardizzati  $u_{it} = y_{it}/h_{iit}^{1/2}$  e si procede al secondo stadio della stima partendo dall'ipotesi che la struttura di correlazioni dinamiche è data da:

$$R_t = \tilde{Q}_t^{-1} Q_t \tilde{Q}_t^{-1} \quad [4.24]$$

dove  $\tilde{Q}_t$ , introdotta per garantire che  $R_t$  sia effettivamente una matrice delle correlazioni<sup>13</sup>, è una matrice diagonale con le radici quadrate degli elementi posti lungo la diagonale di  $Q_t$ . La matrice  $Q_t$  è parametrizzata attraverso un modello GARCH(p,q) del tipo:

$$Q_t = \left(1 - \sum_{r=1}^q a_r - \sum_{r=1}^p b_r\right) \bar{Q} + \sum_{r=1}^q a_r u_{t-r} u'_{t-r} + \sum_{r=1}^p b_r Q_{t-r} \quad [4.25]$$

dove  $\bar{Q} = E(u_t u'_t)$  è la matrice non condizionale delle correlazioni di  $y_t$ . Questa relazione deriva direttamente dalla definizione del processo standardizzato<sup>14</sup>. Il numero totale dei parametri da stimare all'interno del modello DCC è pari perciò a

$$p + q + n + \sum_{i=1}^n (P_i + Q_i)$$

La costante del modello di cui alla [4.25] è esplicitata in base all'equazione della varianza non condizionale calcolata mediante un generico GARCH univariato. Nonostante il procedimento di stima a due stadi conduca a stime meno efficienti, tuttavia in Engle e Sheppard (2001) si mostra che esse sono consistenti ed asintoticamente normali.

La specificazione della matrice  $R_t$  come matrice dei residui standardizzati garantisce per costruzione che la matrice condizionale delle varianze e delle covarianze sia definita positiva. Un limite del modello DCC risiede soprattutto nel fatto che tutte le correlazioni condizionali  $\rho_{ij}$  seguono lo stesso processo quindi le possibilità di loro evoluzioni temporali risultano ristrette. Esaminando questo vincolo dal punto di vista dei parametri emerge invece che la sua imposizione consente al modello di essere facilmente stimabile anche in presenza di una dimensione elevata per il vettore  $y_t$ .

<sup>13</sup>cfr. Billio, Caporini e Gobbo (2003).

<sup>14</sup>Dato che le correlazioni di  $y_t$  condizionali al set informativo  $I_{t-1}$  si definiscono come segue:

$$\rho_{ijt} = \frac{E_{t-1}(y_{it} y_{jt})}{\sqrt{E_{t-1}(y_{it}^2) E_{t-1}(y_{jt}^2)}}$$

Poiché  $u_{it} = y_{it}/h_{iit}^{1/2}$ , da ciò segue che le correlazioni condizionali coincidono con le covarianze condizionali dei processi standardizzati:

$$\rho_{ijt} = E_{t-1}(u_{it} u_{jt})$$

## 4.7 Modello O-GARCH

Nell'ambito del risk management riuscire a modellare una struttura di varianze e di covarianze quando la numerosità  $n$  dei titoli in portafoglio è elevata costituisce un problema di non facile soluzione. Il modello GARCH ortogonale (O-GARCH) proposto da Alexander e Chibumba (1996) si inserisce perciò nell'ambito dei modelli multivariati di tipo ARCH come una via alternativa ai modelli CCC e DCC per l'analisi della volatilità variabile anche in presenza di un numero elevato di asset<sup>15</sup>.

L'idea centrale consiste non nell'applicazione del modello alle serie storiche originali bensì alle sue componenti principali che per costruzione hanno matrice non condizionale delle varianze e delle covarianze pari alla matrice identità. In questo modo stime e previsioni condotte sul vettore dei dati possono essere condotte semplicemente invertendo la scomposizione in base alle componenti principali stesse, data della relazione:

$$y_t = Wz_t \quad [4.26]$$

dove  $y_t$  è il vettore di dimensione  $n$  contenente le osservazioni che, in base a questa equazione, è generato dalla combinazione lineare di alcune variabili non osservabili  $z_t$ . Nell'ipotesi che la matrice ortogonale<sup>16</sup>  $W$  sia costante nel tempo ed invertibile, le variabili in  $z_t$  così determinate si identificano sostanzialmente in un set di fattori incorrelati che guidano le evoluzioni di un dato mercato quindi il modello O-GARCH diviene simile al modello fattoriale di cui al sottoparagrafo 4.4. Dalla [4.26] segue che la matrice condizionale delle varianze e delle covarianze di  $y_t$  è data da:

$$\Omega_t = WD_tW' \quad [4.27]$$

dove  $D_t = E_t(z_t z_t')$ . Partendo dall'assunto  $\varepsilon_t | I_{t-1} \sim MN(0, \Omega_t)$  proprio dei modelli multivariati di tipo ARCH, la distribuzione condizionale delle componenti principali è perciò:

$$z_t | I_{t-1} \sim MN(0, D_t) \quad [4.28]$$

L'ipotesi forte del modello è la condizione che pone  $D_t$  diagonale all'interno della quale sono disposte tutte le varianze condizionali  $d_{it}$ . Analogamente ai modelli CCC e DCC, questa condizione presenta il vantaggio di garantire una matrice  $\Omega_t$  definita positiva. Nell'O-GARCH ciascun elemento di  $D_t$  è descritto da un GARCH(p,q) univariato del tipo:

$$d_{it} = \left( 1 - \sum_{r=1}^q \alpha_{ir} - \sum_{r=1}^p \beta_{ir} \right) + \sum_{r=1}^q \alpha_{ir} y_{it-r}^2 + \sum_{r=1}^p \beta_{ir} d_{it-r} \quad [4.29]$$

<sup>15</sup>In letteratura sono stati introdotti altri modelli che utilizzano la tecnica delle componenti principali nell'ambito dei modelli multivariati di tipo ARCH: il più noto è il GARCH a componenti principali di Ding e Engle (2001), mentre una possibile alternativa è costituita dal modello di Giannini e Rossi (1999).

<sup>16</sup>Data la matrice  $Y_t$  di dimensione  $T \times n$  contenente le serie storiche disponibili, questa proprietà deriva dal fatto che le colonne della matrice  $W$  sono gli autovettori di  $Y_t' Y_t$  ordinati in base ai valori dei corrispondenti autovalori.

dove  $H_0 = I$  per costruzione. Il modello O-GARCH soffre di alcuni problemi legati essenzialmente all'identificazione dei parametri in quanto la condizione di ortogonalità risulta essere molto restrittiva<sup>17</sup>.

Van der Weide (2002) ha proposto il modello Generalized Orthogonal GARCH (GO-GARCH) che supera il problema dell'identificazione dei parametri inclusi in  $W$ . Le matrici ortogonali costituiscono un sottoinsieme di tutte le possibili matrici invertibili; il GO-GARCH consente di instaurare un legame attraverso una matrice qualsiasi matrice  $W$  invertibile.

---

<sup>17</sup>Dato che la stima dell'O-GARCH è basata sulla matrice non condizionale delle varianze e delle covarianze, quando si è in presenza di deboli correlazioni tra le variabili, il modello incontra notevoli difficoltà nell'identificazione di una matrice che sia realmente ortogonale (Alexander, 2002).

---

## 5 Inferenza nei modelli di tipo ARCH

---

Questo paragrafo si occupa dei problemi relativi alla stima di un modello lineare nel quale l'errore segue un processo di tipo ARCH e alle procedure di test per verificare la presenza di eteroschedasticità condizionale all'interno delle serie storiche.

### 5.1 Stima ML

Partendo dalla distribuzione delle innovazioni standardizzate  $u_t$ , per la stima dei parametri relativi alla media condizionale e alla varianza condizionale generalmente si ricorre al metodo della massima verosimiglianza (ML). Dato il vettore  $\theta$  contenente tutti i parametri incogniti relativi sia alla media condizionale sia alla varianza condizionale delle osservazioni  $y_t$ , le ipotesi di partenza sono:

1.  $u_t(\theta) \sim i.i.d. (0, 1)$
2.  $u_t(\theta) = \varepsilon_t(\theta)/h_t(\theta)^{1/2}$ ,
3. La funzione di densità dell'innovazione standardizzata condizionale al set informativo  $I_{t-1}$  è data da  $f(u_t(\theta), \eta)$  dove  $\eta \in H \subseteq R^k$  rappresenta il vettore dei parametri di disturbo<sup>1</sup>.

La funzione di verosimiglianza è costruita mediante la produttoria di funzioni di densità condizionali<sup>2</sup>, quindi può essere interpretata come il risultato di una fattorizzazione sequenziale in cui si assume una distribuzione degenere per  $\varepsilon_1$ . Ponendo  $y = (y_T, y_{T-1}, \dots, y_1)$ , essa risulta così definita:

$$L_T(y, \psi) = \prod_{t=1}^T f(\varepsilon_t(\theta), \eta | I_{t-1}) \quad [5.1]$$

---

<sup>1</sup>Parametri propri di una distribuzione statistica come ad esempio i g.d.l.

<sup>2</sup>Nel caso di variabili casuali i.i.d. la funzione di verosimiglianza è calcolata attraverso la produttoria delle funzioni di densità marginali.

dove  $f(\varepsilon_t(\theta), \eta | I_{t-1})$  è la funzione di densità dell'innovazione e  $\psi' = [\theta' \quad \eta']$ . In base alla distribuzione di  $u_t$  la [5.1] diventa:

$$L_T(y, \psi) = \prod_{t=1}^T f(u_t(\theta), \eta) h_t(\theta)^{-1/2} \quad [5.2]$$

dove  $h_t(\theta)^{-1/2} = |\partial u_t / \partial \varepsilon_t|$  è il determinante della Jacobiana ottenuta nel passaggio dall'innovazione  $\varepsilon_t$  a quella standardizzata  $u_t$ . La funzione log-verosimiglianza è perciò:

$$\ell_T(y, \psi) = \sum_{t=1}^T \left\{ \ln[f(u_t(\theta), \eta)] - \frac{1}{2} \ln[h_t(\theta)] \right\} \quad [5.3]$$

Lo stimatore ML consiste nella statistica  $\hat{\psi}$  per la quale risulta:

$$S_T(y, \psi) = \frac{\partial \ell_T(y, \psi)}{\partial \theta} = 0 \quad [5.4]$$

dove  $S_T$  è lo score relativo all'intero campione per i parametri della media condizionale e della varianza condizionale. Dopo alcuni passaggi si ottiene:

$$S_T(y, \psi) = \sum_{t=1}^T \left[ \frac{f'(u_t(\theta), \eta)}{f(u_t(\theta), \eta)} \frac{\partial u_t(\theta)}{\partial \theta} - \frac{1}{2h_t(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} \right] \quad [5.5]$$

con

$$\begin{cases} \frac{\partial u_t(\theta)}{\partial \theta} = \frac{\partial \mu_t(\theta)}{\partial \theta} h_t(\theta)^{-1/2} - \frac{1}{2h_t(\theta)^{3/2}} \frac{\partial h_t(\theta)}{\partial \theta} \varepsilon_t(\theta) \\ \varepsilon_t(\theta) = y_t - \mu_t(\theta) \end{cases}$$

Poiché non esiste una soluzione analitica alla [5.4], per ottenere lo stimatore ML bisogna ricorrere alle tecniche di ottimizzazione basate sul calcolo numerico. Un'altra ragione per la quale si deve ricorrere all'utilizzo di algoritmi numerici risiede nel fatto che la derivata di  $h_t$  rispetto al vettore  $\theta$  contiene al suo interno  $\partial h_{t-1} / \partial \theta$ , quindi tutti gli addendi di cui alla [5.5] mostrano una componente ricorsiva.

Nella fase di stima risulta di fondamentale la specificazione della funzione di densità delle innovazioni standardizzate. Nell'ipotesi in cui  $u_t \sim N(0, 1)$ , la log-verosimiglianza diventa:

$$\ell_T(y, \theta) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \ln[h_t(\theta)] - \frac{1}{2} \sum_{t=1}^T u_t(\theta)^2 \quad [5.6]$$

Sotto questa condizione la [5.5] assume perciò la seguente forma:

$$S_T(y, \theta) = \sum_{t=1}^T \left\{ \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\varepsilon_t^2(\theta)}{h_t(\theta)} + \frac{1}{2h_t} \frac{\partial h_t(\theta)}{\partial \theta} \left[ \frac{\varepsilon_t^2(\theta)}{h_t(\theta)} - 1 \right] \right\} \quad [5.7]$$

In molte applicazioni empiriche giustificare l'ipotesi di normalità del termine  $u_t$  non è facilmente giustificabile soprattutto per la presenza di leptocurtosi nella distribuzione delle serie storiche: per poter continuare a lavorare con la funzione di verosimiglianza, Bollerslev (1987) ha utilizzato la variabile casuale  $t$  di Student standardizzata con  $\eta > 2$  g.d.l. poiché essa è simmetrica e converge verso una variabile casuale normale quando  $\eta \rightarrow \infty$ . La curtosi della  $t$  di Student vale  $3(\eta - 2)/(\eta - 4) > 3$  per qualsiasi valore finito di  $\eta$  purché sia  $\eta > 4$ .

Nelson (1991) invece ha utilizzato la Generalized Error Distribution (GED), una particolare distribuzione con la seguente funzione di densità per l'innovazione standardizzata:

$$f[u_t(\theta), \eta] = \frac{\eta}{2^{(1+1/\eta)} \kappa \Gamma(\eta^{-1})} \exp \left\{ -\frac{1}{2} \left| \frac{u_t(\theta)}{\kappa} \right|^\eta \right\} \quad [5.8]$$

con  $\kappa = [2^{(-2/\eta)} \Gamma(\eta^{-1}) \Gamma(3\eta^{-1})^{-1}]^{1/2}$ . A seconda dal valore del parametro  $\eta$  la GED assume una diversa forma soprattutto per quanto riguarda la massa di probabilità presente alle code: quando  $\eta = 2$  la [5.8] diventa la funzione di densità di una variabile casuale normale standardizzata, quando  $\eta < 2$  essa mostra un valore di curtosi superiore a 3.

Crowder (1976) ha mostrato che sotto opportune condizioni di regolarità lo stimatore ML è consistente, asintoticamente efficiente ed ha la seguente distribuzione asintotica normale:

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{d} N(0, A_0^{-1}) \quad [5.9]$$

dove  $A_0$  è la matrice di informazione di Fisher calcolata in corrispondenza del vero vettore dei parametri  $\psi_0$ .

## 5.2 Stima QML

La stima di quasi-massima verosimiglianza (QML) rappresenta una possibilità alternativa alla stima ML quando viene a cadere l'ipotesi di normalità dell'innovazione standardizzata  $u_t$ : in questo si utilizza lo score di cui alla [5.7] il vettore stimato  $\hat{\theta}_T$  è lo stesso che si otterrebbe per la stima ML. La differenza tra i due metodi consiste essenzialmente nella conformazione della matrice delle varianze e delle covarianze, infatti sotto opportune condizioni di normalità Weiss (1986) ha dimostrato che:

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{d} N(0, A_0^{-1} B_0 A_0^{-1}) \quad [5.10]$$

dove  $B_0$  è la matrice contenente i valori attesi dei prodotti incrociati degli score calcolati per il vero parametro  $\theta_0$ . Lo stimatore QML fornisce quindi stime consistenti ed asintoticamente normali per il vettore  $\theta_0$ : Lee e Hansen (1994) hanno condotto degli studi in particolare sul GARCH(1,1) trovando che tali conclusioni non dipendono dal fatto che i termini  $u_t$  siano normali o indipendenti.

Analiticamente le matrici  $A_0$  e  $B_0$  hanno la seguente forma:

$$\left\{ \begin{array}{l} A_0 = -\frac{1}{T} \sum_{t=1}^T \left[ \frac{\partial^2 \ell_t(\theta)}{\partial \theta \partial \theta'} \right] \\ B_0 = \frac{1}{T} \sum_{t=1}^T \left[ \frac{\partial \ell_t(\theta)}{\partial \theta} \frac{\partial \ell_t(\theta)}{\partial \theta'} \right] \end{array} \right. \quad [5.11]$$

Svolgendo le derivate si ottiene:

$$\left\{ \begin{array}{l} A_0 = \sum_{t=1}^T \left[ \frac{1}{2h_t^2(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} \frac{\partial h_t(\theta)}{\partial \theta'} + \frac{1}{h_t(\theta)} \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\partial \mu_t(\theta)}{\partial \theta'} \right] \\ B_0 = \sum_{t=1}^T \left[ \frac{1}{4h_t^2(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} \frac{\partial h_t(\theta)}{\partial \theta'} (K_t(\theta) - 1) + \frac{1}{h_t(\theta)} \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\partial \mu_t(\theta)}{\partial \theta'} \right] + \\ + \sum_{t=1}^T \frac{1}{2} \left[ \frac{\partial h_t(\theta)}{\partial \theta} \frac{\partial \mu_t(\theta)}{\partial \theta'} + \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\partial h_t(\theta)}{\partial \theta'} \right] SK_t(\theta) h_t^{3/2}(\theta) \end{array} \right. \quad [5.12]$$

dove  $SK_t(\theta) = E_{t-1}[\varepsilon^3(\theta)/h_t^{3/2}(\theta)]$  e  $K_t(\theta) = E_{t-1}[\varepsilon^4(\theta)/h_t^2(\theta)]$  sono rispettivamente i valori condizionali per l'indice di asimmetria e per la curtosi della distribuzione delle innovazioni standardizzate. Si nota immediatamente che quando  $u_t \sim N(0, 1)$  risulta che  $SK_t(\theta) = 0$  e  $K_t(\theta) = 3$  e le due matrici  $A_0$  e  $B_0$  di cui alla [5.12] coincidono quindi anche le matrici delle varianze e delle covarianze asintotiche degli stimatori ML e QML assumono la stessa espressione.

Lo stimatore QML generalmente non è efficiente poiché la varianza asintotica non rispetta il limite di Cramer-Rao, tuttavia Bollerslev e Wooldridge (1992) hanno mostrato che la perdita di efficienza che si ha rispetto allo stimatore ML è rilevante solo quando si utilizzano distribuzioni con scarti non simmetrici dalla variabile casuale normale<sup>3</sup>.

In diverse applicazioni dei modelli di tipo ARCH il vettore dei parametri da stimare può essere partizionato in  $\theta' = [\theta'_1 \ \theta'_2]$ ; in questo modo si effettua una separazione tra i parametri relativi alla media condizionale da quelli inerenti alla varianza condizionale; sostituendo i vettori  $\theta_1$  e  $\theta_2$  all'interno della [5.12] si dimostra che la matrice di informazione di Fisher è diagonale a blocchi, condizione che semplifica notevolmente la determinazione della sua inversa e consente la stima separata dei parametri della media condizionale e quelli della volatilità.

<sup>3</sup>In particolare i due autori mostrano che la differenza tra la varianza asintotica e l'inversa della matrice di Informazione di Fisher è una combinazione lineare dei momenti centrali della distribuzione dal terzo in poi con pesi decrescenti man mano che aumenta il loro ordine. Poiché per le distribuzioni simmetriche il momento centrale di ordine 3 è nullo, è quindi evidente che in questo caso lo stimatore QML registra una minore perdita di efficienza.



Engle (1982) fornisce alcune condizioni formalizzando un ARCH( $q$ ) lineare nel caso  $u_t$  abbia distribuzione normale standard. Il risultato sono stime asintoticamente efficienti dei parametri relativi alla varianza condizionale ottenute sulla base di quelle consistenti per la media condizionale e viceversa: in questo modo, quando si ha un modello di regressione lineare con innovazioni ARCH stazionarie in covarianza, il metodo OLS fornisce stime consistenti per i coefficienti e stime asintoticamente efficienti per i parametri ARCH, calcolate sulla base dei residui. Per arrivare a tale conclusione occorre tenere conto della presenza di eteroschedasticità condizionale quindi si deve ridefinire in questo senso la componente idiosincratca del modello OLS. Per modelli come l'EGARCH o il GARCH-M la matrice di informazione di Fisher non è diagonale a blocchi, quindi non è possibile raggiungere gli stessi risultati soprattutto ed i parametri della media condizionale e quelli della varianza condizionale devono essere stimati simultaneamente.

### 5.3 Test statistici

L'utilizzo di appropriati test statistici costituisce la base sulla quale si può decidere se ad una o più serie storiche è applicabile un modello di tipo ARCH, cioè se effettivamente ci sia presenza di eteroschedasticità condizionale.

Il test più utilizzato è il Test ARCH introdotto da Engle (1982) che parte dalle seguenti ipotesi:

1. Si ha un modello di regressione lineare  $y_t = x_t b + \varepsilon_t$  nel quale  $x_t$  sono variabili esogene o ritardate,
2. Le serie storiche sono composte di  $T$  osservazioni.
3. La struttura delle ipotesi del test ARCH è la seguente:

$$\begin{cases} H_0 : \varepsilon_t | I_{t-1} \sim N(0, h) \\ H_1 : \varepsilon_t | I_{t-1} \sim N(0, h_t) \end{cases}$$

L'ipotesi nulla è costituita dalla presenza di omoschedasticità nelle serie storiche in esame quindi i quadrati delle innovazioni hanno variazioni casuali. Nell'ipotesi alternativa, secondo la quale c'è presenza di eteroschedasticità condizionale, il termine  $\varepsilon_t^2$  segue una dinamica legata ai valori da esso assunto nei periodi precedenti.

Il test ARCH si basa sulla regressione ausiliaria data dal prodotto  $\varepsilon_t^2 = z_t' \alpha$  dove  $\alpha' = [\alpha_0 \ \alpha_1 \ \dots \ \alpha_q]$  e  $z_t' = [1 \ \hat{\varepsilon}_{t-1}^2 \ \dots \ \hat{\varepsilon}_{t-q}^2]$  dove le componenti del vettore  $z_t$  sono i residui calcolati attraverso il modello OLS di cui all'ipotesi 1; analiticamente esso si configura come un test LM costruito sullo score nell'ipotesi che  $H_0$  sia vera cioè che valga la relazione  $\alpha_1 = \alpha_2 = \dots = \alpha_q = 0$ .

Nell'ipotesi di normalità della distribuzione delle innovazioni standardizzate è possibile costruire una funzione di log-verosimiglianza per  $\hat{\varepsilon}_t$  del tutto analoga alla

[5.6]:

$$\ell_T(\alpha) = -\frac{T}{2} \ln(2\pi) + \sum_{t=1}^T \left[ -\frac{1}{2} \ln(h_t) - \frac{1}{2} \frac{\hat{\varepsilon}_t^2}{h_t} \right] \quad [5.13]$$

Derivando rispetto ad  $\alpha$  si ottiene lo score:

$$S_T(\alpha) = \frac{\partial \ell_T(\alpha)}{\partial \alpha} = \sum_{t=1}^T \frac{G' z'_t}{2h_t} \left[ \frac{\hat{\varepsilon}_t^2}{h_t} - 1 \right] \quad [5.14]$$

dove  $G' z'_t = \partial h_t / \partial \alpha$  mentre  $G'$  è uno scalare che deriva dalla funzione  $h_t = G(z'_t \alpha)$ . Nell'ipotesi in cui  $H_0$  è vera la [5.14] diventa:

$$S_T(\alpha_0) = \frac{G'}{2h_0} Z' f_0 \quad [5.15]$$

dove  $h_0 = G(\alpha_0)$ ,  $f_0 = [ (\hat{\varepsilon}_1/h_0 - 1) \quad (\hat{\varepsilon}_2/h_0 - 1) \quad \dots \quad (\hat{\varepsilon}_T/h_0 - 1) ]'$  e  $Z' = [ z'_1 \quad z'_2 \quad \dots \quad z'_T ]$ . La matrice delle derivate seconde è:

$$\frac{\partial^2 \ell(\alpha)}{\partial \alpha \partial \alpha'} = \sum_{t=1}^T \left( \frac{G'}{h_t} \right)^2 \left( \frac{\hat{\varepsilon}_t^2}{h_t} z'_t z_t + \frac{1}{2} z'_t z_t \right) \quad [5.16]$$

Sotto  $H_0$  la matrice di informazione di Fisher è perciò<sup>4</sup>:

$$\mathcal{I}_T(\alpha_0) = -\frac{1}{T} \left[ \frac{\partial^2 \ell_T(\alpha_0)}{\partial \alpha_0 \partial \alpha_0'} \right] = \frac{1}{2} \left( \frac{G'}{h_t} \right)^2 \frac{1}{T} \sum_{t=1}^T E(z'_t z_t) \quad [5.17]$$

Derivando rispetto ad  $\alpha$  si ottiene l'equazione dello score che, valutata sotto  $H_0$ , entra nella statistica test LM data dalla seguente espressione:

$$LM = S_T(\alpha_0)' [\mathcal{I}_T(\alpha_0)]^{-1} S_T(\alpha_0) = \frac{1}{2} f_0' Z (Z' Z)^{-1} Z' f_0 \quad [5.18]$$

Nell'ipotesi che le innovazioni abbiano distribuzione normale il momento quarto è un multiplo della varianza quindi si può dimostrare agevolmente che  $f_0' f_0 / T \xrightarrow{\text{Pr}} 2$ ; un'espressione asintoticamente equivalente alla [5.18] è:

$$LM = T \frac{f_0' Z (Z' Z)^{-1} Z' f_0}{f_0' f_0} = TR^2 \quad [5.19]$$

Formalmente il test ARCH consiste perciò in una statistica test LM data dal prodotto  $TR^2$  dove il coefficiente  $R^2$  è quello di una regressione ausiliaria dei quadrati

<sup>4</sup>L'equazione [5.17] si ottiene utilizzando la legge dei valori attesi iterati e sfruttando la relazione

$$E(z'_t z_t) = E_{t-1} \left( \frac{\hat{\varepsilon}_t^2}{\alpha_0} z'_t z_t \right)$$

dei residui OLS  $\hat{\varepsilon}_t^2$  su una costante e su  $q$  suoi ritardi<sup>5</sup>. La distribuzione limite per la statistica test è quella di una variabile casuale chi quadrato con  $q$  g.d.l., dove  $q$  è il numero di vincoli imposti in  $H_0$ .

Quando il modello di regressione non è correttamente specificato, cioè nel caso in cui alcuni regressori rilevanti siano stati omessi oppure non riescano a catturare alcuni aspetti legati alla presenza di correlazione seriale o non linearità delle innovazioni, il test ARCH potrebbe rifiutare l'ipotesi nulla; in questo caso non necessariamente è implicata la presenza di eteroschedasticità condizionale nei dati.

Con disturbi del tipo GARCH(p,q) il test non sarebbe applicabile in quanto la matrice di informazione di Fisher è singolare (Bollerslev, 1986): ciò è dovuto al fatto che i parametri relativi ai ritardi di  $h_t$  non sono identificati quando  $H_0$  è vera. Lee (1991) ha mostrato che i risultati ottenibili con l'ARCH(q) si possono estendere anche ai GARCH(p,q).

Il test ARCH, anche se rappresenta il contributo più importante proposto dalla letteratura, tuttavia non è l'unico test diagnostico per l'eteroschedasticità condizionale. A questo proposito Tse (2001) distingue tre categorie nei test portmanteau, LM e infine in quelli sui residui.

Alla prima categoria appartengono tutti i procedimenti come ad esempio quello di Ljung e Box (1978) calcolato sui quadrati delle innovazioni.

Un altro contributo importante appartiene a Brock, Deschert e Sheinkman (1987) che hanno formulato l'omonimo test BDS per valutare a livello più generale la dipendenza temporale di tipo non lineare presente all'interno delle serie storiche sui rendimenti: la distribuzione asintotica della statistica BDS sotto l'ipotesi nulla per la quale le componenti del vettore  $y_t$  non sono i.i.d., è quella della variabile casuale normale<sup>6</sup>.

---

<sup>5</sup>Il passaggio da  $\hat{\varepsilon}_t^2$  a  $f_0$  (aggiunta di una costante e moltiplicazione per uno scalare) infatti non modifica il coefficiente  $R^2$ .

<sup>6</sup>Per una trattazione più dettagliata del test BDS si veda Bollerslev, Engle e Nelson (1994).

## BIBLIOGRAFIA

---

## Bibliografia

---

- ALEXANDER, C. (2002). *Principal components models for generating large GARCH covariance matrices*. Review of Banking, Finance and Monetary Economics.
- ALEXANDER, C. E CHIBUMBA, A. (1996). *Multivariate orthogonal factor GARCH*. Discussion Paper in mathematics, University of Sussex.
- BAILLIE, R. T., BOLLERSLEV, T. E MIKKELSEN, H. O. (1996). *Fractionally integrated generalized autoregressive conditional heteroskedasticity*. Journal of Econometrics, 74(1), pp. 3–30.
- BILLIO, M., CAPORINI, M. E GOBBO, M. (2003). *Block dynamic conditional correlation multivariate GARCH models*. GRETA working paper n. 03.03.
- BLACK, F. (1976). *Studies of stock price volatility changes*. in Proceedings of the 1976 Meeting of the Business and Economic Statistic Section, American Statistical Association, pp. 79–121.
- BOLLERSLEV, T. (1986). *Generalized autoregressive conditional heteroskedasticity*. Journal of Econometrics, 31, pp. 307–327.
- BOLLERSLEV, T. (1987). *A conditional heteroskedastic time series model for speculative prices and rates of return*. Review of Economics and Statistics, 69, pp. 542–547.
- BOLLERSLEV, T. (1988). *On the correlation structure for the GARCH process*. Journal of Time Series Analysis, 9, pp. 121–131.
- BOLLERSLEV, T. (1990). *Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH approach*. Review of Economics and Statistics, 72, pp. 498–505.
- BOLLERSLEV, T., ENGLE, R. F. E NELSON, D. B. (1994). *ARCH models*. in R.F. Engle & D.L. McFadden (eds.), *Handbook of Econometrics*, vol. IV, Elsevier, North Holland.

## BIBLIOGRAFIA

- BOLLERSLEV, T., ENGLE, R. F. E WOOLDRIDGE, J. M. (1988). *A capital asset pricing model with time varying covariances*. Journal of Political Economy, 96, pp. 116–131.
- BOLLERSLEV, T. E MIKKELSEN, H. O. (1996). *Modelling and pricing long memory in stock market volatility*. Journal of Econometrics, 73, pp. 151–184.
- BOLLERSLEV, T. E WOOLDRIDGE, J. M. (1992). *Quasi maximum likelihood estimation and inference in dynamic models with time varying covariances*. Economic Review, 11, pp. 143–172.
- BOX, G. E. P. E COX, D. R. (1964). *An analysis of transformations*. Journal of the Royal Statistical Society, serie B, 26, pp. 211–246.
- BROCK, A. W., DESCHERT, W. D. E SHEINKMAN, J. A. (1987). *A test for independence based on correlation dimension*. manoscritto, Department of Economics, University of Wisconsin, Madison.
- CAI, J. (1994). *A markov model of unconditional variance in ARCH*. Journal of Business and Economic Statistics, 12, pp. 309–316.
- CAMPBELL, J. Y., LO, A. W. E MACKINLAY, A. C. (1997). *The econometrics of financial markets*. Princeton University Press.
- CORNELL, B. (1978). *Using the options pricing model to measure the uncertainty producing effect of major announcements*. Financial Management, 7, pp. 54–59.
- CROWDER, M. J. (1976). *Maximum likelihood estimation with dependent observation*. Journal of the Royal Statistical Society (series B), 38, pp. 45–53.
- DAVIDSON, J. (2003). *Moment and memory properties of linear conditional heteroskedasticity models, and a new model*. working paper, Cardiff University.
- DIEBOLD, F. X. (1988). *Empirical modelling of exchange rate dynamics*. Springer-Verlag, New York.
- DING, Z. X. E ENGLE, R. F. (2001). *Large scale conditional covariance matrix estimation and testing*. working paper.
- DING, Z. X., ENGLE, R. F. E GRANGER, C. W. J. (1993). *A long memory property of stock markets returns and a new model*. Journal of Empirical Finance, 1, pp. 83–106.
- DING, Z. X. E GRANGER, C. W. J. (1996). *Modelling volatility persistence of speculative returns: a new approach*. Journal of Econometrics, 73, pp. 185–215.
- DROST, F. C. E NIJMAN, T. N. (1993). *Temporal aggregation of GARCH processes*. Econometrica, 61, pp. 909–927.

- ENGLE, R. F. (1982). *Autoregressive conditional heteroskedasticity with estimates of the U.K. inflation*. *Econometrica*, 50, pp. 987–1008.
- ENGLE, R. F. (2002). *Dynamic conditional correlation - A simple class of multivariate GARCH models*. *Journal of Business and Economic Statistics*, 20, pp. 339–350.
- ENGLE, R. F. E BOLLERSLEV, T. (1986). *Modelling the persistence of conditional variances*. *Econometric Review*, 5, pp. 1–50.
- ENGLE, R. F. E KRONER, K. F. (1995). *Multivariate simultaneous generalized ARCH*. *Econometric Theory*, 11, pp. 122–150.
- ENGLE, R. F., LILIEN, D. M. E ROBINS, R. P. (1987). *Estimating time varying risk premia in the term structure: The ARCH-M model*. *Econometrica*, 55, pp. 391–407.
- ENGLE, R. F. E NG, V. K. (1993). *Measuring and testing the the impact of news on volatility*. *Journal of Finance*, 48, pp. 1749–1778.
- ENGLE, R. F., NG, V. K. E ROTHSCILD, M. (1990). *Asset pricing with a factor ARCH covariance structure: Empirical estimates for Treasury Bills*. *Journal of Econometrics*, 45, pp. 213–238.
- ENGLE, R. F. E SHEPPARD, K. (2001). *Theoretical and empirical properties of dynamical conditional correlation model multivariate GARCH*. UCSD Discussion paper 2001-15.
- ENGLE, R. F. E SUSMEL, R. (1993). *Common volatility and international equity markets*. *Journal of Business and Economic Statistics*, 11, pp. 167–176.
- FAMA, E. F. (1965). *The behavior of stock market prices*. *Journal of Business*, 38, pp. 34–105.
- FAMA, E. F. (1970). *Efficient capital markets: a review of theory and empirical work*. *Journal of Finance*, 25, pp. 383–417.
- FRENCH, K. R. E ROLL, R. (1986). *Stock return variances: the arrival of information and the reaction of traders*. *Journal of Financial Economics*, 38, pp. 5–26.
- GALLANT, A. R., ROSSI, P. E. E TAUCHEN, G. (1992). *Stock prices and volume*. *Review of Financial Studies*, 5, pp. 1091–1120.
- GHYSELS, E., HARVEY, A. C. E RENAULT, E. (1996). *Stochastic volatility*. in G.S. Maddala & C.R. Rao (eds.), *Handbook of Statistics*, vol. 14, Statistical Methods in Finance, Elsevier, North Holland.

## BIBLIOGRAFIA

- GIANNINI, C. E ROSSI, E. (1999). *A principal components multivariate GARCH technique for medium size portfolio management*. Collana studi del Credito Italiano, supplemento a *Strategie d'investimento*.
- GLOSTEN, L. R., JAGANNATHAN, R. E RUNKLE, D. (1993). *On the reaction between the expected value and the volatility of the nominal excess return on stocks*. *Journal of Finance*, 48, pp. 1779–1801.
- GRANGER, C. W. J. E JOYEUX, R. (1980). *An introduction to long memory time series analysis and fractional differencing*. *Journal of Time Series Analysis*, 1, pp. 15–29.
- HAMILTON, J. D. (1990). *Analysis of time series subject to changes in regime*. *Journal of Econometrics*, 45, pp. 39–70.
- HAMILTON, J. D. (1994). *Time series analysis*. Princeton University Press.
- HAMILTON, J. D. E SUSMEL, R. (1994). *Autoregressive conditional heteroskedasticity and changes in regime*. *Journal of Econometrics*, 64, pp. 307–333.
- HENTSCHHEL, L. (1994). *All in the family: nesting symmetric and asymmetric GARCH models*. *Journal of Econometrics*, 39, pp. 71–104.
- HOSKING, J. R. (1981). *Fractional differencing*. *Biometrika*, 68, pp. 165–176.
- KARPOFF, J. (1987). *The relation between price changes and trading volume: a survey*. *Journal of Financial Quantitative Analysis*, 22, pp. 109–126.
- KING, M., SENTANA, E. E WADHWANI, S. (1994). *Volatility links between national stock markets*. *Econometrica*, 62, pp. 901–933.
- LEE, J. H. H. (1991). *A Lagrange multiplier test for GARCH models*. *Economic Letters*, 37, pp. 265–271.
- LEE, S. W. E HANSEN, B. E. (1994). *Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator*. *Econometric Theory*, 10, pp. 29–52.
- LEE, W. (2000). *Advanced theory and methodology of tactical asset allocation*. John Wiley, New York.
- LJUNG, G. M. E BOX, G. E. P. (1978). *On a measure of lag of fit in time series models*. *Biometrika*, 67, pp. 297–303.
- MANDELBROT, B. B. (1963). *The variation of certain speculative prices*. *Journal of Business*, 36, pp. 394–419.
- MCCURDY, T. H. E STEGONS, T. (1992). *A comparison risk premium forecast implied by parametric and non parametric conditional mean estimators*. *Journal of Econometrics*, 52, pp. 225–244.



- NELSON, D. E CAO, C. (1992). *Inequality constraints in the univariate GARCH model*. Journal of Business and Economic Statistic, 10, pp. 229–235.
- NELSON, D. B. (1990). *Stationarity and persistence in the GARCH(1,1) model*. Econometric Theory, 6, pp. 318–334.
- NELSON, D. B. (1991). *Conditional heteroskedasticity in asset returns: a new approach*. Econometrica, 59, pp. 347–370.
- PAGAN, A. R. E SCHWERT, G. W. (1990). *Alternative models for conditional stock volatility*. Journal of Econometrics, 45, pp. 267–290.
- PALM, F. C. (1996). *GARCH models of volatility*. in G.S. Maddala & C.R. Rao (eds.), *Handbook of Statistics*, vol. 14, Statistical Methods in Finance, Elsevier, North Holland.
- PATELL, J. M. E WOLFSON, M. A. (1979). *Anticipated information releases reflected in call options prices*. Journal of Accounting and Economics, 1, pp. 117–140.
- PATELL, J. M. E WOLFSON, M. A. (1981). *The ex-ante and ex-post price effects of quarterly earnings announcement reflected in option and stock price*. Journal of Accounting Research, 19, pp. 434–458.
- ROSS, S. (1976). *The arbitrage theory of capital asset pricing*. Journal of Economic Theory, 4, pp. 341–360.
- SCHWERT, G. W. (1989). *Why does stock market volatility change over time?* Journal of Finance, 44, pp. 1115–1153.
- SENTANA, E. (1995). *Quadratic ARCH models: a potential re-interpretation of ARCH models*. Review of Economic Studies, 62, pp. 639–661.
- TAYLOR, S. (1986). *Modelling financial time series*. John Wiley & Sons, New York.
- TSE, Y. K. (2001). *Residual-based diagnostics for conditional heteroskedasticity models*. working paper, Department of Economics, national University of Singapore.
- VAN DER WEIDE, R. (2002). *GO-GARCH: a multivariate generalized orthogonal GARCH model*. Journal of Applied Econometrics, 17, pp. 549–564.
- WEISS, A. A. (1986). *Asimptotic theory for ARCH models: estimation and testing*. Econometric Theory, 2, pp. 107–131.
- ZAKOIAN, J. M. (1994). *Threshold heteroskedastic models*. Journal of Economic Dynamic and Control, 18, pp. 931–955.