



**Dipartimento di Statistica**  
**"Giuseppe Parenti"**

Dipartimento di Statistica "G. Parenti" - Viale Morgagni 59 - 50134 Firenze - [www.ds.unifi.it](http://www.ds.unifi.it)

S E R I E R I C E R C H E 2 0 0 5

Modelli Econometrici per  
Fenomeni Soggetti a Smooth  
Transition: Lo studio della  
volatilità sui mercati finanziari

Massimiliano Cecconi,  
Giampiero M. Gallo



Università degli Studi  
di Firenze

Modelli Econometrici  
per Fenomeni Soggetti a Smooth Transition:

Lo studio della volatilità sui mercati finanziari

Massimiliano Cecconi  
Giampiero M. Gallo



# Indice

<b>1</b>	<b>Introduzione</b>	<b>7</b>
1.1	Regolarità Empiriche . . . . .	9
<b>2</b>	<b>Breve rassegna sulla modellistica esistente</b>	<b>15</b>
2.1	Modelli per la media condizionata . . . . .	15
2.2	Modelli per la varianza condizionata . . . . .	19
2.2.1	ARCH . . . . .	19
2.2.2	GARCH . . . . .	20
2.2.3	ARCH esponenziale e ARCH in valore assoluto . . . . .	21
2.2.4	GARCH non-lineare . . . . .	21
2.2.5	ARCH moltiplicativo . . . . .	21
2.2.6	ARCH-M . . . . .	22
2.2.7	Modello con deviazione standard autoregressiva . . . . .	22
2.2.8	AGARCH . . . . .	23
2.2.9	EGARCH . . . . .	23
2.2.10	NARCH . . . . .	23
2.2.11	GJR . . . . .	24
2.2.12	A-PARCH . . . . .	24
2.2.13	TGARCH . . . . .	24
2.2.14	SWARCH . . . . .	25
2.2.15	GQARCH . . . . .	25
2.2.16	VS-GARCH . . . . .	26
2.3	Considerazioni conclusive . . . . .	26
<b>3</b>	<b>Modelli autoregressivi a transizione liscia</b>	<b>29</b>
3.1	I modelli . . . . .	30
3.1.1	Proprietà dei modelli STAR . . . . .	31
3.1.2	Alcune generalizzazioni . . . . .	34
3.2	Procedura empirica di specificazione . . . . .	37
3.3	Specificazione del modello lineare AR(p) . . . . .	39
3.4	Introduzione al test di linearità . . . . .	41
3.5	Test per STAR . . . . .	42
3.5.1	Test contro l'alternativa LSTAR . . . . .	44
3.5.2	Test contro l'alternativa ESTAR . . . . .	47

3.5.3	Test contro le alternative LSTAR-D ed ESTAR-D . . . . .	48
3.5.4	Selezione della variabile di transizione . . . . .	51
3.5.5	Scelta della funzione di transizione . . . . .	52
3.5.6	Applicazione del test al tasso sui Fondi Federali U.S.A. . . . .	54
3.6	Stima dei parametri . . . . .	60
3.6.1	Aspetti di calcolo nella fase di stima . . . . .	63
3.7	Valutazione del modello stimato . . . . .	66
3.7.1	Test per non-linearità rimanenti . . . . .	67
3.7.2	Verifica della costanza dei parametri . . . . .	69
3.7.3	Valutazione del modello mediante estrapolazione . . . . .	70
3.8	Utilizzo del modello . . . . .	71
3.9	Considerazioni conclusive . . . . .	74
<b>4</b>	<b>Sviluppi recenti</b>	<b>75</b>
4.1	Eteroschedasticità e test di linearità . . . . .	75
4.2	Outliers e test di linearità . . . . .	78
4.2.1	Una breve discussione sugli outliers . . . . .	78
4.2.2	Robustificazione del test di linearità . . . . .	80
4.3	Un cambiamento di notazione . . . . .	81
4.4	Modelli STAR a regimi multipli . . . . .	83
4.4.1	Specificazione dei modelli MR-STAR . . . . .	85
4.5	Modelli STAR con parametri variabili nel tempo . . . . .	87
4.5.1	Specificazione dei modelli TV-STAR . . . . .	88
4.5.2	Approccio “dallo specifico - al generale” . . . . .	89
4.5.3	Approccio “dallo specifico - al generale - allo specifico” . . . . .	90
4.6	Considerazioni conclusive . . . . .	91
<b>5</b>	<b>Modelli GARCH a transizione liscia</b>	<b>93</b>
5.1	I modelli STGARCH . . . . .	94
5.1.1	Proprietà dei modelli STGARCH . . . . .	96
5.2	Test per effetti STGARCH . . . . .	99
5.2.1	Scelta tra GARCH e STGARCH . . . . .	101
5.2.2	Una versione robusta del test . . . . .	107
5.3	Stima dei modelli STGARCH . . . . .	109
5.4	Valutazione del modello stimato . . . . .	112
5.4.1	Verifica della presenza di ulteriori regimi di volatilità . . . . .	115
5.4.2	Test per la non costanza dei parametri . . . . .	116
5.5	Possibili sviluppi dei modelli STGARCH . . . . .	117
5.5.1	Il modello STGARCH in media . . . . .	117
5.5.2	I modelli STGARCH a regimi multipli . . . . .	118
5.5.3	I modelli STGARCH con parametri variabili nel tempo . . . . .	119
5.6	Una strategia operativa . . . . .	120
5.7	Applicazione al Tasso sui Fondi Federali U.S.A. . . . .	122
5.8	Considerazioni conclusive . . . . .	124
<b>6</b>	<b>Conclusioni</b>	<b>127</b>

<i>INDICE</i>	5
<b>A Dati utilizzati nelle applicazioni</b>	<b>131</b>
<b>B Derivazione del test di linearità</b>	<b>133</b>
B.1 Caso LSTAR: approssimazione del primo ordine . . . . .	133
B.2 Caso LSTAR: approssimazione del terzo ordine . . . . .	133
B.3 Caso ESTAR: relazioni tra parametri . . . . .	134
B.4 Caso LSTAR-D: relazioni tra parametri . . . . .	135
<b>C Derivate analitiche per il test di dipendenza seriale</b>	<b>137</b>



# Capitolo 1

## Introduzione

Negli ultimi vent'anni, l'attenzione rivolta a modelli econometrici di tipo non-lineare è cresciuta molto rapidamente per due serie di motivi. Da un lato vi è stata la constatazione, oramai generalmente accettata, che il comportamento della maggior parte delle variabili economiche evolve in modo non-lineare. In secondo luogo, i miglioramenti apportati dalla tecnologia informatica hanno ampiamente ridotto i tempi di specificazione e stima di relazioni dinamiche complesse. Infatti, nonostante i modelli lineari possano essere considerati in molti casi una buona approssimazione della vera relazione non-lineare (almeno localmente), è soprattutto grazie ai moderni calcolatori che si è diffusa la pratica di ricercare formulazioni alternative a quelle tradizionali, abbandonando quella che per molti anni è stata la linea di condotta principale e che può essere definita come il "paradigma della linearità" (van Dijk, 1999). Con particolare riferimento ai mercati finanziari, l'osservazione empirica ha ampiamente dimostrato che l'evoluzione di molti fenomeni non può essere adeguatamente caratterizzata da formulazioni lineari. La presenza di un elevato grado di correlazione seriale nelle serie ricavate come trasformazioni non-lineari dei rendimenti (in particolare quadrati e valori assoluti), accompagnata dalla non-normalità della relativa distribuzione empirica, hanno rappresentato il fondamento dell'applicazione di metodologie non-lineari ai mercati finanziari. Inoltre, partendo dall'assunto teorico che l'evoluzione di molti fenomeni sia caratterizzata dallo stato, o regime, in cui si trova l'economia in quel particolare momento, numerosi sono stati i tentativi di formalizzare simili comportamenti. Lo scopo di questo lavoro è quello di presentare una classe di modelli parametrici non-lineari, i modelli a transizione liscia (*Smooth Transition Models*), i cui primi lavori risalgono a Luukkonen, Saikkonen e Teräsvirta (1988a e 1988b), Teräsvirta e Anderson (1992), Teräsvirta (1994) e Lin e Teräsvirta (1994), dove l'idea sottostante è che la dinamica futura delle variabili economiche dipende dallo stato, o regime, in corso. Così ad esempio, il rendimento di un attivo finanziario può essere caratterizzato da dinamiche differenti a seconda che l'intero sistema economico si trovi in una fase di espansione o di contrazione. In modo analogo, la volatilità corrente, intesa come misura del rischio di un determinato investimento, può dipendere del tutto o in parte dal regime di volatilità prevalente al tempo immediatamente precedente. Inoltre, a differenza dei tradizionali modelli per il cambiamento di regime, si ammette la possibilità



che il passaggio da uno stato all'altro avvenga in modo graduale, sulla base di una qualche relazione funzionale completamente specificata a partire dai dati. L'idea della transizione liscia rappresenta un'estensione dei modelli tradizionali in quanto, oltre ad ammettere un cambiamento istantaneo come caso limite, presuppone che l'economia consista di un insieme di agenti individuali o imprese, ciascuno dei quali passa da un regime all'altro istantaneamente, pur essendo difficile, se non improbabile, che tale cambiamento avvenga per tutti nello stesso momento. Così se il fenomeno osservato è di tipo aggregato, sarà molto probabile che il passaggio da uno stato all'altro avvenga nel continuo, ed in particolare sarà tanto più graduale quanto più smussato sarà il valore del termine che governa la transizione. La struttura del lavoro è la seguente: il resto del Capitolo introduce allo studio delle serie storiche finanziarie, presentandone peculiarità e caratteristiche distintive rispetto alle altre serie storiche economiche. Nel Capitolo 2 vengono introdotte le principali formulazioni parametriche non-lineari per la media e per la varianza condizionata di un generico fenomeno economico e si conclude con l'inquadramento dei modelli a transizione liscia all'interno della classe più generale dei modelli con cambiamento di regime. Nel Capitolo 3 vengono presentati i modelli autoregressivi a transizione liscia (STAR) utilizzati per l'analisi della media condizionata, soffermandosi in particolare sulla procedura di specificazione, stima, valutazione ed utilizzo dei modelli proposta da Teräsvirta (1994) ed Eitheim e Teräsvirta (1996). Ad una trattazione teorica dei singoli punti della procedura viene affiancata una esemplificazione su dati reali, dalla quale emerge chiaramente come l'impiego della parametrizzazione STAR sia in grado di migliorare considerevolmente le proprietà del modello in rapporto alla specificazione lineare. Sviluppi recenti del modello base sono considerati nel Capitolo 4 e consistono nella sua estensione al caso di regimi multipli e di parametri variabili nel tempo, all'interno di una notazione alternativa recentemente introdotta. Nello stesso Capitolo sono considerate versioni robuste del test di linearità che tengono conto della presenza di eteroschedasticità in forma generica ed *outliers*. Il Capitolo 5 si concentra invece sull'estensione dei modelli a transizione liscia al caso di eteroschedasticità condizionata di tipo autoregressivo (STGARCH), mettendo in evidenza come la possibilità di più regimi di volatilità possa costituire un aspetto tutt'altro che trascurabile nell'analisi dei fenomeni finanziari. Il Capitolo si concentra prevalentemente sulla ricerca di una coerente strategia operativa che sia in grado di selezionare adeguatamente il modello più idoneo a descrivere la dinamica della varianza condizionata delle serie storiche finanziarie, soffermandosi in particolare sulla possibilità di estendere a tale contesto la procedura utilizzata da Teräsvirta (1994) per la specificazione dei modelli STAR. La procedura descritta è affiancata da un esperimento di simulazione che ne conferma una buona operatività. Il Capitolo termina con una esemplificazione illustrativa su dati reali. Il Capitolo 6 presenta alcune considerazioni conclusive, mentre nelle Appendici sono riportate informazioni relative alle serie storiche utilizzate, nonché approfondimenti su temi specifici analizzati nel corso del testo. Tutte le applicazioni sono state programmate in GAUSS 3.2<sup>1</sup>.

---

<sup>1</sup> Copyright © 1995-1996 Aptech Systems, Inc.

## 1.1 Regolarità Empiriche

Le serie storiche finanziarie presentano alcune caratteristiche peculiari che le differenziano nettamente dalle altre serie storiche economiche. Alla base dell'evoluzione dei fenomeni finanziari, siano essi prezzi o tassi d'interesse, vi è l'ipotesi di passeggiata aleatoria (*random walk*). Secondo tale assunzione, la migliore previsione del prezzo di domani di un qualche attivo finanziario utilizzando l'informazione disponibile fino ad oggi è data dal prezzo odierno. In altri termini, si può dire che tra il tempo  $t - 1$  e il tempo  $t$  il cambiamento atteso nei prezzi è uguale a zero e non vi è correlazione tra la variazione dei prezzi nell'intervallo temporale considerato (Granger e Morgenstern, 1970 e Taylor, 1986).

Seguendo Gallo e Pacini (2002),<sup>2</sup> la traduzione dell'ipotesi di *random walk* in termini statistici risulta immediata. Per definizione, il prezzo di un attivo al tempo  $t$  è legato al prezzo dello stesso attivo al tempo  $t - 1$  attraverso il rendimento intercorso fra  $t - 1$  e  $t$ ,

$$P_t = P_{t-1}(1 + r_t) \quad (1.1)$$

che in termini logaritmici è esprimibile come<sup>3</sup>

$$p_t = p_{t-1} + \log(1 + r_t) \quad (1.2)$$

Per  $r_t$  sufficientemente piccolo, quale risulta senz'altro per variazioni di prezzo giornaliere, tramite l'espansione di Taylor del termine  $\log(1 + r_t)$  intorno ad 1, la (1.2) diventa

$$p_t \approx p_{t-1} + r_t \quad (1.3)$$

e da cui risulta che

$$r_t \approx \Delta p_t \equiv p_t - p_{t-1} \quad (1.4)$$

Dalla (1.3), eventualmente moltiplicata per 100 per lavorare con rendimenti percentuali così come sono usualmente espressi in pratica, si può notare che, al tempo  $t - 1$ , il valore di  $p_{t-1}$  risulta noto in quanto osservato, mentre la determinazione di  $p_t$  dipende in maniera cruciale dal valore assunto da  $r_t$  che, essendo incerto, lo caratterizza come variabile casuale. Caratterizzandosi come variabile aleatoria, diverse ipotesi possono essere formulate sui momenti (condizionati e non) e sulla distribuzione di  $r_t$ . In linea con l'idea di passeggiata casuale, l'ipotesi più semplice che si può fare su  $r_t$  è che

$$r_t \mid I_{t-1} = \epsilon_t \sim N(0, 1) \quad t = 1, \dots, T \quad (1.5)$$

dove  $I_{t-1}$  rappresenta tutta l'informazione disponibile fino al tempo  $t - 1$  ed è generalmente indicata in letteratura come *information set*. Questo significa che il processo che descrive i rendimenti,  $\{r_t\}_{t=1}^T$ , è rappresentato da una sequenza di variabili casuali indipendenti tra loro, con distribuzione identica di tipo Gaussiano, media zero e varianza  $\sigma^2$  costante nel tempo che, in altri termini, si definisce come *white noise* Gaussiano.

<sup>2</sup>Capitolo 5, relativo all'*Analisi dei prezzi*.

<sup>3</sup>Indicheremo con la lettera minuscola,  $p$ , il logaritmo naturale ( $\log$ ) dei prezzi.

Utilizzando la (1.5) e sostituendo ricorsivamente all'indietro l'espressione del prezzo nella (1.3), si ha che

$$p_t = p_0 + \sum_{i=0}^{t-1} \epsilon_{t-i} \quad (1.6)$$

ovvero, a partire da una condizione iniziale, il processo generatore dei prezzi si rappresenta come somma di variabili casuali indipendenti e identicamente distribuite. Un processo così definito prende il nome di *random walk* ed è caratterizzato da una totale persistenza delle innovazioni  $\epsilon_t$  (memoria lunga), la cui somma può anche essere definita trend stocastico.<sup>4</sup>

La verifica empirica dell'ipotesi di *random walk*, accompagnata da una quasi totale assenza di autocorrelazione nella serie dei rendimenti, è spesso citata come dimostrazione dell'ipotesi di efficienza dei mercati. Un mercato è detto perfettamente efficiente se i prezzi degli attivi al tempo  $t$  incorporano tutta l'informazione disponibile fino a quel momento. In tal modo, il livello dei prezzi si modifica istantaneamente al sopraggiungere di nuova informazione (Fama, 1976). In realtà, numerose evidenze empiriche note in letteratura come fatti stilizzati,<sup>5</sup> dimostrano che l'ipotesi di efficienza dei mercati non può essere completamente accettata. Valgono infatti le seguenti considerazioni:

- Un'autocorrelazione pressoché nulla non può essere interpretata come indipendenza dei rendimenti poiché la loro distribuzione empirica non è Gaussiana (Mandelbrot, 1963). Questo è essenzialmente dovuto alla presenza di un eccessivo numero di osservazioni sulle code della distribuzione,<sup>6</sup> in particolare su quella sinistra (asimmetria), e da un più alto addensamento di valori intorno alla media. Distribuzioni di questo tipo sono definite in letteratura leptocurtiche. Tali considerazioni sono facilmente verificabili calcolando il momento terzo (skewness,  $sk$ ) e il momento quarto (curtosi,  $k$ ) della distribuzione non condizionata dei rendimenti:

$$sk = \frac{1}{T} \sum_{t=1}^T \left( \frac{r_t - \bar{r}}{\sigma_r} \right)^3, \quad k = \frac{1}{T} \sum_{t=1}^T \left( \frac{r_t - \bar{r}}{\sigma_r} \right)^4. \quad (1.7)$$

Al calcolo di tali indicatori viene solitamente affiancato il test di normalità Jarque-Bera (1980),

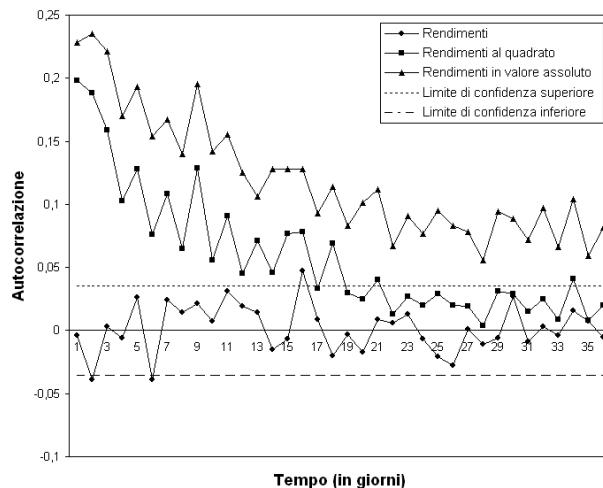
$$JB = \frac{T}{6} \left( sk^2 + \frac{1}{4}(k-3)^2 \right), \quad (1.8)$$

<sup>4</sup>Si dice inoltre che un processo così definito possiede una radice unitaria. Questo equivale ad affermare quanto segue: in un modello AR(1) (si veda la Sezione 2.1) abbiamo che  $(1 - \phi L)p_t = \epsilon_t$ , dove  $L$  rappresenta l'operatore ritardo definito come  $Lp_t \equiv p_{t-1}$ . Nel *random walk*,  $\phi = 1$  per cui, se la soluzione del polinomio caratteristico  $(1 - \phi L) = 0$  giace sul cerchio unitario nel piano complesso, si dice che vi è radice unitaria: il processo è non stazionario nei livelli poiché possiede media e varianza non costanti nel tempo. In un modello AR(1),  $|\phi| < 1$  per cui la radice del polinomio caratteristico si trova al di fuori del cerchio unitario, mentre l'effetto di persistenza si riduce gradualmente fino ad annullarsi. Per un'eccellente trattazione sull'argomento si veda Hamilton (1994) o, relativamente ad applicazioni ai mercati finanziari, Gallo e Pacini (2002).

<sup>5</sup>O, nella terminologia anglosassone, *stylized-facts*.

<sup>6</sup>Il cosiddetto fenomeno delle "code grasse", o *fat-tails*.

Figura 1.1: Autocorrelazioni dei rendimenti, dei quadrati e dei loro valori assoluti per l'indice DAX di Francoforte.



la cui statistica test, asintoticamente distribuita come un  $\chi^2$  con 2 gradi di libertà, è basata sul confronto tra i momenti terzo e quarto della distribuzione normale standardizzata, rispettivamente pari a 0 e 3, con i corrispondenti momenti della distribuzione empirica. A tal proposito si veda la Tabella 1.1 dove, assieme al calcolo di skewness, curtosi e test Jarque-Bera (JB), si riportano le principali statistiche descrittive dei rendimenti percentuali per gli indici S&P500 e CAC40, per i tassi di cambio JPY/USD e DM/USD e per le azioni INTC e T quotate rispettivamente al NASDAQ e al NYSE.<sup>7</sup> Come si vede, non solo siamo in presenza di una marcata asimmetria negativa della relativa distribuzione empirica, ma è un valore dell'indice di curtosi di molto superiore a 3, a far sì che anche il test di normalità porti a risultati ampiamente insoddisfacenti. Tali caratteristiche sono comuni alla maggioranza delle serie storiche finanziarie.

- Se si analizza la struttura temporale dei rendimenti, pur essendo in linea con l'ipotesi di incorrelazione seriale, esiste una forte autocorrelazione su  $|r_t|^d$ , con  $d > 0$  (Ding, Granger e Engle, 1993). In particolare, i rendimenti al quadrato e quelli in valore assoluto presentano una forte struttura dinamica. In Figura 1.1 è riportato il grafico delle prime 36 autocorrelazioni delle serie dei rendimenti, dei loro quadrati e dei valori assoluti per la serie dell'indice DAX quotato alla borsa di Francoforte. Le due linee orizzontali rappresentano i limiti di confidenza al 95% ( $\pm 2/\sqrt{T}$ , sotto l'ipotesi di normalità) per la verifica dell'assenza di autocorrelazione. Come si vede, quasi tutti i valori della serie dei rendimenti sono compresi tra le bande mentre, all'opposto, nessuno dei valori per i rendimenti in valore assoluto vi è incluso. Questo conferma ulteriormente la non normalità di  $r_t$  poiché, in tal caso, non vale che incorrelazione implica indipendenza. Il

<sup>7</sup>L'Appendice A riporta una dettagliata descrizione dei dati utilizzati nelle applicazioni.

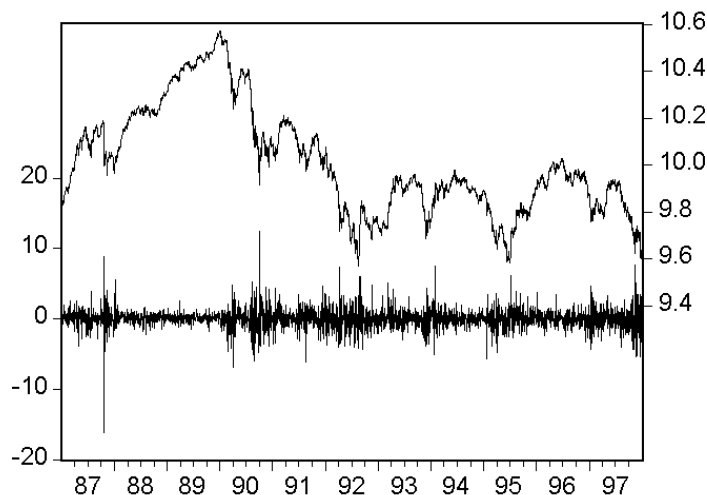
Tabella 1.1: Statistiche descrittive e test di normalità dei rendimenti percentuali di alcuni indici di borsa, tassi di cambio e azioni.

	S&P500	CAC40	JPY/USD	DEM/USD	INTC	T
Media	0.048	0.025	-0.015	-0.0007	0.114	-0.020
Mediana	0.038	0.000	0.008	0.019	0.106	0.000
Massimo	8.708	8.225	3.365	3.102	12.85	12.38
Minimo	-22.83	-10.13	-5.630	-4.140	-24.88	-29.54
Dev. Standard	0.993	1.198	0.671	0.667	2.618	1.936
Skewness	-4.300	-0.528	-0.563	-0.164	-0.539	-1.397
Curtosi	99.73	10.57	7.213	5.134	8.343	27.64
Jarque-Bera	1228919	6670	3929	963	3551	73513
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Osservazioni	3127	2734	4958	4958	2868	2868

primo ad evidenziare questo tipo di comportamento per le serie dei rendimenti giornalieri è stato Taylor (1986), mentre, tramite uno studio simile condotto su una serie storica molto lunga dello *Standard & Poor 500*, Ding *et al.* (1993) evidenziano una decadenza di tipo iperbolico delle autocorrelazioni campionarie dei rendimenti al quadrato e di quelli in valore assoluto, aprendo così un filone di ricerca nella modellazione di  $r_t^2$  che fa ampiamente uso dei concetti di memoria lunga e integrazione frazionaria introdotti da Granger (1980).

- La varianza dei rendimenti non è costante nel tempo, ma cambia con il variare dell'intensità con cui l'attività di trading si sviluppa sui mercati finanziari. Questo perché i prezzi hanno la tendenza a cambiare più spesso in alcuni giorni piuttosto che in altri, dando luogo a quello che in letteratura viene definito *volatility clustering* (raggruppamento di volatilità). In Figura 1.2 è riportato l'andamento del logaritmo dei prezzi e dei rendimenti percentuali per l'indice Nikkei di Tokyo. Come si vede, periodi di turbolenza sono caratterizzati da oscillazioni elevate nei rendimenti che tendono a persistere nel tempo; viceversa periodi più calmi danno luogo a variazioni più contenute. In tale regolarità empirica trova quindi fondamento l'idea dell'esistenza di diversi regimi di volatilità che, semplificando al massimo, si può ricondurre ai soli due stati di alta e bassa volatilità. Da questa considerazione ha preso avvio una vasta letteratura sull'argomento che, prevalentemente in ambito non-lineare, ha dato luogo a una molteplicità di formulazioni parametriche in grado di riprodurre comportamenti simili. Più in particolare, ciò che differenzia maggiormente tra loro le varie alternative è il diverso modo di interpretare, e quindi formalizzare, il passaggio da un regime all'altro di volatilità. Vedremo in seguito come ciò si sostanzia in pratica; basti qui accennare che il grosso della letteratura si è concentrato su tre diverse macro-alternative: quella del cambiamento istantaneo, quella del cambiamento graduale (o *smooth*) e quella che incorpora un cambiamento regolato da un processo stocastico sottostante non osservabile.

Figura 1.2: Andamento del logaritmo dei prezzi (scala di destra) e dei rendimenti percentuali (scala di sinistra) per l'indice Nikkei di Tokyo nel periodo 6 gennaio 1986 – 31 dicembre 1999.



Fra le spiegazioni della dipendenza seriale è stata avanzata l'ipotesi di non-linearità di  $r_t$  (Taylor, 1986). Tutta l'analisi econometrica si è quindi spostata sulla ricerca del processo generatore dei dati (DGP) che meglio di altri sia in grado di riprodurre le regolarità empiriche osservate. In particolare, l'attenzione è stata rivolta allo studio delle dinamiche non-lineari presenti nell'evoluzione di  $r_t$  e  $r_t^2$ . La scelta di lavorare sul quadrato dei rendimenti in luogo del loro valore assoluto è da attribuire a considerazioni di interpretabilità. Modellare  $r_t^2$  significa approssimare adeguatamente la varianza dei rendimenti, poiché vale che  $Var(r_t) = E(r_t^2)$  a meno di un errore casuale.

La varianza dei rendimenti è generalmente utilizzata come misura di volatilità e, di conseguenza, della rischiosità dell'asset sottostante. Tenere sotto controllo il rischio derivante dai propri investimenti costituisce un aspetto di indubbia utilità che qualsiasi operatore finanziario vorrebbe ridurre al minimo. Inoltre, la possibilità di disporre di previsioni accurate sulla volatilità futura del mercato permette di monitorare e controllare l'evoluzione temporale del rischio.

Ipotizzando di avere a disposizione tutta l'informazione rilevante al tempo  $t$  per la dinamica futura di  $r_t$ ,  $I_{t-1}$ , l'attenzione si sposta sulla ricerca del più adeguato modello statistico da adattare ai dati, operando in un contesto prevalentemente ateorico. In particolare saranno presi in considerazione i primi due momenti condizionati della distribuzione di  $r_t$ ,

$$m_t = E[r_t | I_{t-1}] \quad (1.9)$$

e

$$h_t = E[(r_t - m_t)^2 | I_{t-1}] \quad (1.10)$$

dove  $m_t$  e  $h_t$  indicano rispettivamente la media e la varianza condizionata di  $r_t$ . La regola generale che seguiremo nel corso della trattazione si basa sulla specificazione e

stima, prima del modello per la media condizionata e, solo successivamente, di quello per la varianza condizionata. La motivazione di questo modo di procedere è dovuta a una duplice considerazione. In primo luogo facciamo riferimento alla constatazione empirica che una qualche rappresentazione non-lineare per  $r_t$  può essere in grado di coglierne meglio la dinamica. In secondo luogo consideriamo che la teoria statistica sulla stima dimostra che non è possibile ottenere valori statisticamente accettabili per i parametri che caratterizzano  $h_t$  senza aver prima specificato adeguatamente la dinamica dei rendimenti (Lundbergh e Teräsvirta, 1998). Si è visto che la quasi completa assenza di autocorrelazione nella struttura temporale di  $r_t$  non è da attribuire all'assenza di dipendenza seriale poiché è accompagnata da una generalizzata non normalità della relativa distribuzione empirica. Questo implica che il miglior modello per la media condizionata non è da ricercare tra le specificazioni lineari, ma deve essere inevitabilmente non-lineare (Taylor, 1986). Solo successivamente l'attenzione si sposta sulla ricerca della miglior parametrizzazione per la varianza condizionata, al fine di catturare gli effetti di *clustering* tipici delle serie storiche finanziarie. Di questo ci occuperemo in ciò che segue.

## Capitolo 2

# Breve rassegna sulla modellistica esistente

### 2.1 Modelli per la media condizionata

Nella ricerca della miglior specificazione per la media condizionata di un generico processo  $\{y_t\}_{t=1}^T$ , l'ipotesi più semplice che si può formulare è che  $y_t$  dipenda da un certo insieme di valori passati ai quali si aggiunge una componente non sistematica  $\epsilon_t$  (innovazione) nella quale sono racchiusi tutti quei fattori non osservabili, o comunque non prevedibili per cause esterne. Il termine  $\epsilon_t$  è dunque una variabile casuale che rende la relazione tra  $y_t$  e i suoi valori passati non deterministica.

In ciò che segue e in generale nel corso di tutta la trattazione, assumeremo che  $\{y_t\}_{t=1}^T$  rappresenti un processo stocastico stazionario in covarianza, o debolmente stazionario. Ciò equivale a supporre la validità delle seguenti condizioni (Piccolo, 1990 e Hamilton, 1994):

- $E(y_t) = \mu$  costante  $\forall t$ ;
- $Var(y_t) = \sigma^2$  costante e finita (non dipende dal tempo);
- $Cov(y_t, y_{t-k}) = \gamma(k)$ , ovvero dipende da  $k$  (l'intervallo temporale tra  $t - k$  e  $t$ ), ma non da  $t$ .

Laddove queste condizioni non dovessero essere soddisfatte, assumeremo che  $y_t$  rappresenti una qualche trasformazione del processo originario tale per cui risulti debolmente stazionario. Ad esempio, nel contesto dei mercati finanziari, sappiamo che il logaritmo naturale del livello dei prezzi risulta essere non stazionario, mentre se ne prendiamo le differenze prime logaritmiche (rendimenti), eventualmente espresse in termini percentuali, quello che otteniamo è una serie storica che, almeno apparentemente, soddisfa le condizioni richieste per la stazionarietà in covarianza.<sup>1</sup>

---

<sup>1</sup>Per un'analisi dettagliata su quanto qui brevemente accennato si veda Gallo e Pacini (2002).



Se assumiamo che l'effetto di  $y_{t-1}, y_{t-2}, \dots, y_0$  si trasmetta in modo proporzionale su  $y_t$ , allora ipotizziamo un legame lineare tra la variabile che vogliamo spiegare e/o prevedere e l'*information set*,  $I_{t-1}$ , utilizzato. Questo equivale ad ipotizzare una relazione del tipo:

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \epsilon_t \quad (2.1)$$

dove i  $p$  valori più recenti di  $y_t$  ne determinano l'evoluzione futura tramite i coefficienti  $\phi_1, \dots, \phi_p$ , ai quali si aggiunge una costante e un termine di innovazione o sorpresa. Affinché  $\epsilon_t$  possa essere considerato tale, occorre che  $E(\epsilon_t | I_{t-1}) = 0$ . Ipotesi aggiuntive possono poi essere formulate sulla varianza (costante o meno) e sulla distribuzione di  $\epsilon_t$  (generalmente assunta di tipo Gaussiano). L'espressione (2.1) rappresenta quello che nella letteratura delle serie temporali viene definito modello autoregressivo di ordine  $p$ ,  $AR(p)$ . Alternativamente, si poteva ipotizzare per  $y_t$  una dipendenza dalle  $q$  innovazioni passate più recenti (che sono note al tempo  $t$ ) nel modo seguente:

$$y_t = \mu + \epsilon_t + \psi_1 \epsilon_{t-1} + \dots + \psi_q \epsilon_{t-q} \quad (2.2)$$

La (2.2) prende il nome di rappresentazione a media mobile di ordine  $q$ ,  $MA(q)$ , mentre unendo la (2.1) e la (2.2) si ricava il modello autoregressivo a media mobile,  $ARMA(p, q)$ ,

$$y_t - \phi_1 y_{t-1} - \dots - \phi_p y_{t-p} = \mu + \epsilon_t + \psi_1 \epsilon_{t-1} + \dots + \psi_q \epsilon_{t-q} \quad (2.3)$$

la cui caratteristica principale è quella di consentire una rappresentazione più parsimoniosa di  $y_t$ , riducendo il numero dei parametri da stimare. Si è evidenziato nel Capitolo precedente che, nell'ambito dei mercati finanziari, lavorare con specificazioni lineari per  $m_t$  non consente di ottenere risultati apprezzabili, specialmente in ambito previsivo, a causa del bassissimo grado di autocorrelazione della serie dei rendimenti che, in molti casi, può anche essere del tutto assente.

A partire dal lavoro di Taylor (1986), numerosi sono stati i tentativi di caratterizzare la dinamica dei fenomeni finanziari tramite formulazioni non-lineari: ad esempio, se consideriamo

$$y_t = f(y_{t-1}, \dots, y_{t-p}; \theta) + \epsilon_t \quad (2.4)$$

otteniamo il modello autoregressivo non-lineare di ordine  $p$ ,  $NLAR(p)$ , dove  $f(\cdot)$  indica una qualche forma funzionale. Inoltre, analogamente al caso lineare, si possono definire i modelli  $NLMA(q)$  e  $NLARMA(p, q)$ . Si noti, tuttavia, che l'enorme varietà di modelli che possono essere derivati in tale contesto porta ad un'inevitabile incremento delle complessità interpretative e di stima del modello proposto. Pertanto, la letteratura sulla modellistica non-lineare si è solitamente concentrata sui soli termini autoregressivi, eventualmente aumentati di qualche variabile esogena, tralasciando del tutto o in parte i termini a media mobile con conseguente riduzione delle complessità inerenti l'implementazione degli algoritmi di stima.

Due specificazioni, ricavabili dalla (2.4), che hanno incontrato grosso favore nelle applicazioni pratiche sono il modello autoregressivo di tipo *threshold* (TAR) sviluppato da Tong e Lim (1980) e da Tong (1983) e l'autoregressivo esponenziale (EAR) di Haggan e Ozaki (1981). Il modello TAR assume la forma

$$y_t = \begin{cases} \mu_1 + \phi_1^{(1)} y_{t-1} + \dots + \phi_p^{(1)} y_{t-p} + \epsilon_t^{(1)} & \text{se } y_{t-d} \leq c \\ \mu_2 + \phi_1^{(2)} y_{t-1} + \dots + \phi_p^{(2)} y_{t-p} + \epsilon_t^{(2)} & \text{se } y_{t-d} > c \end{cases} \quad (2.5)$$

mentre quello EAR è espresso come

$$y_t = \mu + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + (\theta_1 y_{t-1} + \dots + \theta_p y_{t-p}) \exp\{-\gamma y_{t-1}^2\} + \epsilon_t \quad (2.6)$$

Entrambe le formulazioni appartengono alla classe dei modelli con cambiamento di regime dove, in linea con la teoria economica dominante, si assume la presenza di più stati, o regimi, dell'economia. In ciascun istante sarà attivo solo uno dei regimi previsti, mentre il passaggio da uno stato all'altro avviene, nel primo caso in maniera istantanea in corrispondenza dell'attraversamento di una certa soglia  $c$  e, nel secondo caso, il cambiamento di regime avviene nel continuo ed è regolato dalla funzione esponenziale.

Una versione modificata della (2.5) è stata proposta da Ozaki (1981),

$$y_t = \begin{cases} (\phi_1 + \pi_1 |y_{t-1}|)y_{t-1} + \dots + (\phi_p + \pi_p |y_{t-1}|)y_{t-p} + \epsilon_t & \text{se } |y_{t-1}| \leq c \\ (\phi_1 + \pi_1 c)y_{t-1} + \dots + (\phi_p + \pi_p c)y_{t-p} + \epsilon_t & \text{se } |y_{t-1}| > c \end{cases} \quad (2.7)$$

e prende il nome di modello *threshold* non-lineare<sup>2</sup> (NLTAR).

Un altro modello che è stato a lungo utilizzato in pratica, specialmente in ambito finanziario, è il modello bilineare (BL), introdotto da Granger e Andersen (1978):

$$y_t - \sum_{i=1}^p \phi_i y_{t-i} = \mu + \epsilon_t + \sum_{j=1}^q \psi_j \epsilon_{t-j} + \sum_{i=1}^m \sum_{j=1}^k b_{ij} y_{t-i} \epsilon_{t-j} \quad (2.8)$$

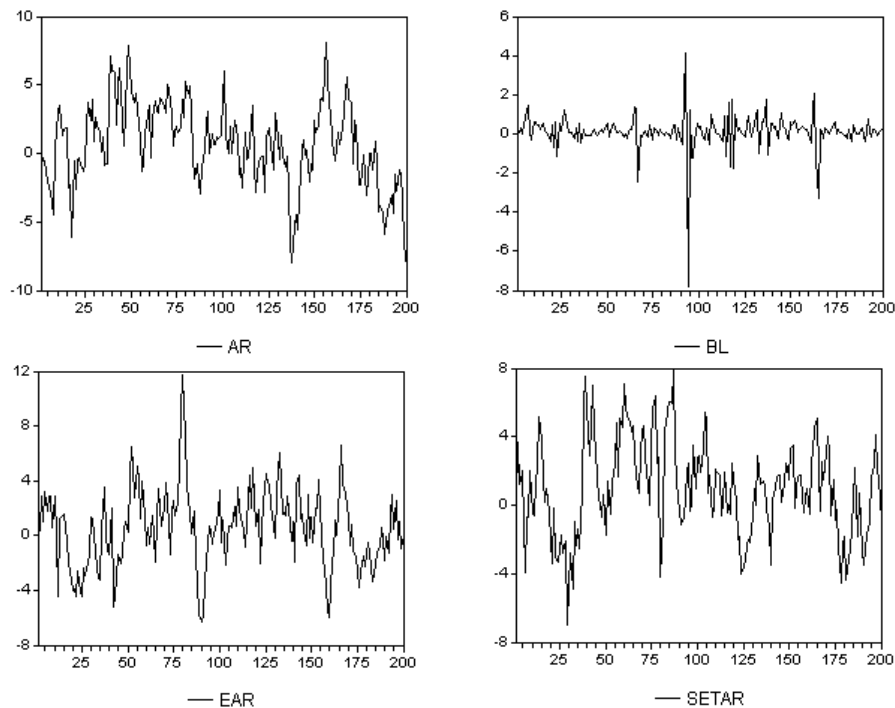
Si tratta di un'estensione del modello ARMA( $p, q$ ) lineare ottenuta semplicemente aggiungendo alla (2.3) la forma bilineare in  $\{y_{t-i}, \epsilon_{t-j}\}$ . Per ulteriori dettagli si rimanda a Priestley (1988) e Granger e Teräsvirta (1993).

La Figura 2.1 mostra delle simulazioni ricavate da alcuni dei modelli appena descritti. Sono stati considerati il modello autoregressivo lineare di ordine uno, AR(1), con parametro 0.8, il modello bilineare, BL,  $y_t = 0.8y_{t-1}\epsilon_{t-1} + \epsilon_t$ , il modello autoregressivo esponenziale, EAR, nella forma  $y_t = 0.8y_{t-1} - 0.8y_{t-1} \exp(-\gamma y_{t-1}^2) + \epsilon_t$ , con  $\gamma = 1$ , e il modello autoregressivo di tipo *threshold*, TAR nella parametrizzazione  $y_t = 0.8y_{t-1} + \epsilon_t$  se  $y_{t-1} \leq 0$  e  $y_t = -0.8y_{t-1} + \epsilon_t$  se  $y_{t-1} > 0$ . Tutti i dati, 200 osservazioni per ciascuna simulazione, sono stati generati ipotizzando una distribuzione Normale a media zero e varianza quattro del termine di disturbo  $\epsilon_t$ . Inoltre, l'intercetta di ciascuna delle parametrizzazioni considerate è stata posta pari a zero e, al fine di evitare distorsioni causate dalla scelta dei valori iniziali, fissati anch'essi convenzionalmente a zero, le prime cento simulazioni ricavate da ciascun modello sono state scartate.

Infine, una classe di modelli che negli ultimi anni ha ricevuto grossa attenzione è rappresentata dai modelli di tipo *Markov-Switching* (MS), sviluppati a lungo da Hamilton (1989). L'idea sottostante a questi modelli è l'assunzione della presenza di più stati dell'economia, ognuno dei quali caratterizzato da un processo stocastico a valori discreti. L'analogia con i modelli TAR ed EAR è abbastanza evidente: in entrambi i casi viene presa in considerazione la possibilità che il comportamento della maggior parte delle variabili economiche dipenda in maniera cruciale dallo stato in cui si trova l'economia in quel particolare momento. Diverso è invece l'approccio seguito nel modellare

<sup>2</sup>Per ulteriori dettagli sui modelli TAR, NLTAR e EAR si rimanda, fra gli altri, a Priestley (1988) e Tong (1990).

Figura 2.1: Simulazioni di modelli AR, BL, EAR e SETAR con estrazioni da una distribuzione Normale a media 0 e varianza 4.



queste eventualità. Mentre nei modelli TAR ed EAR il cambiamento di regime è determinato da variabili osservabili tali da permettere di conoscere con certezza il regime passato e quello presente, nel modello MS lo stato prevalente in un dato momento non può mai essere determinato con esattezza poiché dipende da un processo stocastico non osservabile. Si potranno allora assegnare soltanto delle probabilità ai diversi regimi e, pertanto, al passaggio da uno stato all'altro dell'economia. Tali probabilità dovranno essere stimate assieme agli altri parametri del modello. Formalmente avremo:

$$y_t = \mu_{s_t} + \tilde{y}_t \quad (2.9)$$

con

$$\tilde{y}_t = \phi_1 \tilde{y}_{t-1} + \dots + \phi_p \tilde{y}_{t-p} + \epsilon_t, \quad \epsilon_t \sim iid N(0, 1) \quad (2.10)$$

e dove  $\{s_t\}$  indica il processo stocastico discreto non osservabile, la cui dinamica è governata da una catena di Markov a  $N$  stati. In tal modo avremo che  $\mu_{s_t} = \mu_1$  se  $s_t = 1$ ,  $\mu_{s_t} = \mu_2$  quando  $s_t = 2$ , e così via. La caratteristica principale di una catena di Markov a  $N$  stati è che, ad ogni istante  $t$ , la probabilità che  $s_t$  assuma un certo valore intero  $j \in \{1, 2, \dots, N\}$  dipende dal passato solo attraverso il suo valore più recente  $s_{t-1}$ . Inoltre, la probabilità che lo stato  $i$  sia seguito dallo stato  $j$ , indicata

con  $p_{ij}$  e definita probabilità di transizione, è raccolta in una matrice  $\mathbf{P}$  di dimensioni  $N \times N$ , nota come matrice di transizione:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{bmatrix} \quad (2.11)$$

dove

$$p_{1j} + p_{2j} + \dots + p_{Nj} = 1 \quad \forall j = 1, \dots, N. \quad (2.12)$$

L'elemento della riga  $i$  e della colonna  $j$  di  $\mathbf{P}$  rappresenta la probabilità di transizione  $p_{ij}$ , ovvero la probabilità che lo stato  $i$  sia seguito dallo stato  $j$  (Hamilton, 1994). L'approccio appena descritto si generalizza facilmente al caso in cui la probabilità  $s_t = j$  dipende non solo dal valore di  $s_{t-1}$ , ma anche da un vettore di variabili esogene. Per ulteriori dettagli si rimanda comunque a Hamilton (1989 e 1994), Filardo (1992), Diebold, Lee e Weinbach (1992) ed Avesani e Gallo (1996).

## 2.2 Modelli per la varianza condizionata

Nella specificazione della varianza condizionata ci riferiamo all'approccio ad eteroschedasticità condizionata di tipo autoregressivo (ARCH) che, a partire dalla sua introduzione da parte di Engle (1982), ha riscosso grosso successo in ambito finanziario, sia tra gli accademici che tra i *practitioners*. In termini generali, considerando il generico fenomeno  $y_t$  (ad esempio il rendimento di un qualche attivo finanziario), la cui media condizionata indicheremo con  $\mu_t$ , avremo:

$$y_t = \mu_t + \epsilon_t, \quad \epsilon_t = \eta_t \sqrt{h_t}, \quad \eta_t \sim iid(0, 1). \quad (2.13)$$

Più in particolare, ciascun modello per la varianza condizionata si contraddistingue per la diversa specificazione funzionale assunta da  $h_t$ . Assumeremo inoltre che  $\eta_t$  sia normalmente distribuito<sup>3</sup>. In quel che segue presenteremo brevemente le principali formulazioni parametriche proposte per  $h_t$ , partendo dal lavoro di Engle (1982).

### 2.2.1 ARCH

Engle (1982) suggerisce di modellare la varianza condizionata  $h_t$  come funzione delle  $q$  innovazioni al quadrato più recenti. Il modello assume la forma

$$h_t = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 \quad (2.14)$$

Le restrizioni  $\omega > 0$  e  $\alpha_i \geq 0$  sono necessarie per garantire la positività della varianza condizionata, mentre il vincolo  $\alpha_1, \dots, \alpha_q < 1$  è necessario per non incorrere in una

<sup>3</sup>Per considerazioni sul caso in cui  $\eta_t$  segue una distribuzione di tipo  $t$  - *Student* si veda, fra gli altri, Hamilton (1994), Capitolo 21.

varianza esplosiva che, se da un lato non è certamente impossibile, dall'altro non è conforme con l'ipotesi sottostante di normalità. In questo modello, l'effetto di uno shock avvenuto  $i$  periodi precedenti ( $i \leq q$ ) sulla volatilità corrente è determinato dal parametro  $\alpha_i$  (Engle e Ng, 1993). Per contro, le notizie più vecchie di oltre  $q$  periodi non hanno alcun effetto su  $h_t$ . Quindi, risulta naturale aspettarsi che  $\alpha_i < \alpha_j$  per  $i > j$ , ovvero le notizie più vecchie hanno un impatto minore sulla volatilità corrente.

### 2.2.2 GARCH

Bollerslev (1986) generalizza il modello ARCH( $q$ ) fino ad includervi  $p$  ritardi della varianza condizionata. Empiricamente, la famiglia dei modelli GARCH ha avuto molto successo e, di questi, la specificazione GARCH(1,1) è risultata preferibile nella maggioranza dei casi:<sup>4</sup>

$$h_t = \omega + \alpha \epsilon_{t-1}^2 + \beta h_{t-1} \quad (2.15)$$

Il modello GARCH ha rappresentato un significativo miglioramento rispetto all'ARCH poiché ha permesso una più parsimoniosa struttura dei ritardi delle innovazioni al quadrato nell'equazione della varianza condizionata. In tal modo si è in grado di riprodurre strutture di elevata persistenza nella dinamica di  $h_t$  impiegando un limitato numero di parametri. Tra le caratteristiche principali del processo GARCH vi è senz'altro la sua struttura autoregressiva a media mobile,

$$\epsilon_t^2 = \omega + (\alpha + \beta)\epsilon_{t-1}^2 - \beta(\epsilon_{t-1}^2 - h_{t-1}) + (\epsilon_t^2 - h_t) \quad (2.16)$$

dove la differenza tra i residui al quadrato e la varianza condizionata ha valore atteso nullo. Così questi termini non sono altro che differenze di martingale e possono essere interpretati come sequenze di *white noise* (WN). Dunque, il modello GARCH(1,1) così riscritto può essere visto come un ARMA(1,1) sul processo delle innovazioni al quadrato. Tra le implicazioni della (2.16) vi è la possibilità di stabilire, almeno intuitivamente, le condizioni per la stazionarietà debole del processo GARCH (si veda Bollerslev, 1986).

Infine, come i processi ARCH, anche quelli GARCH sono in grado di generare un coefficiente di curtosi della distribuzione non condizionata di  $\epsilon_t$  maggiore di quello della distribuzione condizionata. Pertanto, risultano utili nella modellizzazione di quelle serie storiche caratterizzate dalla presenza, più o meno abbondante, di osservazioni nelle code della distribuzione.

Nonostante l'apparente successo di queste semplici parametrizzazioni della varianza condizionata, i modelli ARCH e GARCH non sono in grado di descrivere alcune importanti caratteristiche delle serie finanziarie. Tra le più note vi sono la non-linearità e la presenza di effetti *leverage* o di asimmetria riscontrati per la prima volta da Black (1976). Tali effetti consistono nel diverso impatto che le innovazioni positive e negative hanno sulla varianza condizionata: una diminuzione dei prezzi degli attivi finanziari tende ad aumentare la volatilità in misura molto maggiore di quanto faccia un incremento dei prezzi della stessa entità. Queste considerazioni sono alla base di una

<sup>4</sup>Bollerslev, Chou e Kroner (1992), Bera e Higgins (1993) e Bollerslev, Engle e Nelson (1994).

sconfinata letteratura sulla modellistica (G)ARCH che, allo stato attuale, può vantare oltre mille articoli pubblicati. Dei più importanti di questi modelli ci occuperemo brevemente in ciò che segue.

### 2.2.3 ARCH esponenziale e ARCH in valore assoluto

Sebbene lo stesso Engle (1982) si concentri su una specificazione ARCH di tipo lineare, nello stesso articolo ritiene probabile che altre formulazioni del modello possano trovare riscontro oggettivo in molte applicazioni. Pertanto, propone due semplici alternative al modello lineare, l'ARCH esponenziale e l'ARCH in valore assoluto, espressi rispettivamente come:

$$h_t = \exp(\omega + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_q \epsilon_{t-q}^2) \quad (2.17)$$

$$h_t = \omega + \alpha_1 |\epsilon_{t-1}| + \dots + \alpha_q |\epsilon_{t-q}| \quad (2.18)$$

Il modello esponenziale (2.17) ha il vantaggio di avere una varianza sempre positiva; tuttavia, come nota lo stesso Engle (1982), i dati generati da un simile processo possiedono varianza infinita, rendendo difficile sia la stima dei parametri che la successiva fase di inferenza. Il modello in valore assoluto (2.18), così come quello lineare, necessita dell'imposizione di restrizioni che assicurino la positività di  $h_t$ , anche se i dati generati da questo modello hanno varianza finita per tutti i valori dei parametri.

### 2.2.4 GARCH non-lineare

Engle e Bollerslev (1986) hanno stimato una varietà di modelli GARCH per la serie del tasso di cambio Dollaro U.S.A. – Franco Svizzero. I due autori mostrano risultati per i soli due modelli con performance migliore. Si tratta di due specificazioni non-lineari che, nel caso in cui  $q = p = 1$  assumono la forma

$$h_t = \omega + \alpha |\epsilon_{t-1}|^\gamma + \beta h_{t-1} \quad (2.19)$$

e

$$h_t = \omega + \alpha \left[ 2F \left( \frac{\epsilon_{t-1}^2}{\mu} - 1 \right) \right] + \beta h_{t-1} \quad (2.20)$$

dove  $F$  indica la funzione di distribuzione cumulata della normale. Il modello (2.19) rappresenta una semplice estensione di quello in valore assoluto di Engle (1982), mentre il modello (2.20) rende una parametrizzazione nella quale la varianza condizionata cambia a seconda dell'impatto di  $\epsilon_{t-1}^2$ . Il cambiamento di volatilità avviene nel continuo ed è determinato dalla funzione di distribuzione cumulata della variabile casuale normale.

### 2.2.5 ARCH moltiplicativo

Geweke (1986), Pantula (1986) e Mihoj (1987) suggeriscono la seguente forma funzionale per la varianza condizionata:

$$\ln(h_t) = \omega + \alpha_1 \ln(\epsilon_{t-1}^2) + \dots + \alpha_q \ln(\epsilon_{t-q}^2) \quad (2.21)$$

Questa formulazione assicura la positività di  $h_t$  per tutti i valori di  $\alpha$ . Geweke (1986) dimostra che la log-verosimiglianza di questo modello è globalmente concava, rendendone piuttosto semplice la stima di massima verosimiglianza. La principale critica a questa impostazione è stata sollevata da Engle e Bollerslev (1986), i quali notano che la funzione di verosimiglianza diventa indefinita se si incontra un valore dei residui pari a zero.

### 2.2.6 ARCH-M

Seguendo le indicazioni di Markowitz (1959), secondo cui una maggiore variabilità *ex-ante* di una qualsiasi attività finanziaria debba necessariamente ripercuotersi in un maggior rendimento atteso, Engle, Lilien e Robins (1987) propongono il modello ARCH-M (ARCH in Media), originariamente specificato nella forma

$$y_t = \mu_t + \delta\sigma_t + \epsilon_t \quad \text{con} \quad \sigma_t = \sqrt{h_t} \quad (2.22)$$

$$\epsilon_t = \eta_t\sqrt{h_t}, \quad h_t \sim \text{ARCH}(q), \quad \eta_t \sim iid N(0, 1) \quad (2.23)$$

In tal modo avremo che  $y_t|I_{t-1} \sim (\mu_t + \delta\sigma_t, h_t)$  per cui, se  $y_t$  rappresenta il rendimento di un attivo finanziario, la sua media condizionata dipende in modo cruciale dalla deviazione standard condizionata della componente non anticipabile al tempo  $t$  (innovazione). L'impatto che una maggiore variabilità di  $\epsilon_t$  ha sul livello di  $y_t$  è misurato dal parametro  $\delta$ . Un valore di  $\delta$  statisticamente non significativo indica assenza di effetti ARCH in media.

Tra le possibili estensioni del modello c'è la possibilità di considerare la varianza in luogo della deviazione standard condizionata nell'equazione che descrive il livello di  $y_t$ , mentre diverse specificazioni per la dinamica di  $h_t$  possono essere introdotte. Considerazioni simili valgono per la specificazione dell'equazione della media condizionata.

### 2.2.7 Modello con deviazione standard autoregressiva

Schwert (1989) utilizza un modello in valore assoluto per la deviazione standard condizionata  $\sigma_t = \sqrt{h_t}$  del tipo

$$\sigma_t = \omega + \alpha|\epsilon_{t-1}| + \beta\sigma_{t-1} \quad (2.24)$$

Il modello può anche essere ricavato come caso particolare della (2.19) imponendo il vincolo  $\gamma = 1$ . Inoltre, Schwert (1990) lo impiega anche per descrivere l'andamento della varianza condizionata nel modo seguente:

$$h_t = (\omega + \alpha|\epsilon_{t-1}|)^2 + \beta h_{t-1} \quad (2.25)$$

Entrambe le formulazioni permettono un'evoluzione della volatilità simile a quella di un modello GARCH standard.

### 2.2.8 AGARCH

Il modello AGARCH (GARCH asimmetrico) è stato proposto da Engle (1990) e, nel caso più frequente in pratica in cui  $q = p = 1$  assume la forma

$$h_t = \omega + \alpha (\epsilon_{t-1} + \gamma)^2 + \beta h_{t-1} \quad (2.26)$$

Nella (2.26), un valore negativo del parametro  $\gamma$  fa sì che shock negativi abbiano un impatto sulla volatilità maggiore di quello prodotto da shock della stessa taglia, ma di segno opposto. Il modello è stato studiato indipendentemente anche da Sentana (1991 e 1995), il quale si riferisce a questa specificazione come GQARCH (GARCH quadratico), della quale ci occuperemo tra breve.

### 2.2.9 EGARCH

Nelson (1991) propone il modello GARCH esponenziale o, più semplicemente, EGARCH, nel quale si ipotizza una dinamica asimmetrica della varianza condizionata. Nel caso EGARCH(1,1), si ipotizza che il logaritmo naturale di  $h_t$  segua il processo

$$\ln(h_t) = \omega + \beta \ln(h_{t-1}) + \alpha \left( \frac{|\epsilon_{t-1}|}{\sqrt{h_{t-1}}} - \sqrt{\frac{2}{\pi}} \right) + \gamma \frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} \quad (2.27)$$

Uno dei vantaggi della specificazione di Nelson (1991) è che la (2.27) descrive il logaritmo naturale di  $h_t$ , per cui la varianza condizionata sarà sempre positiva indipendentemente dalla positività dei parametri  $\alpha$  e  $\beta$ . Nelson (1991) indica altre due caratteristiche quali punti di forza della sua parametrizzazione rispetto a quella tradizionale. In primo luogo, il modello GARCH non è in grado di spiegare il comportamento asimmetrico di  $h_t$  e, in secondo luogo, nel modello GARCH è difficile valutare se vi sia persistenza o meno delle innovazioni sulla varianza condizionata. Tale persistenza è invece controllata dal parametro  $\beta$  nel modello EGARCH.

### 2.2.10 NARCH

Higgins e Bera (1992) introducono la classe dei modelli ARCH non-lineari (NARCH) espressa dalla formulazione

$$h_t = \left[ \omega^\delta + \alpha_1 (\epsilon_{t-1}^2)^\delta + \dots + \alpha_q (\epsilon_{t-q}^2)^\delta \right]^{1/\delta} \quad (2.28)$$

La caratteristica principale di questa classe di modelli è che, oltre a comprendere al suo interno i modelli (2.19) e (2.21), ha il pregio di permettere la verifica dell'ipotesi di linearità nell'equazione della varianza condizionata contro diverse specificazioni non-lineari. Gli autori derivano un test dei moltiplicatori di Lagrange, il quale è stato poi maggiormente sviluppato e analizzato in Bera e Higgins (1992), ai quali si rimanda.



### 2.2.11 GJR

Il modello GJR di Glosten, Jagannathan e Runkle (1993) rappresenta un'estensione del GARCH includendovi un termine che cattura l'evoluzione asimmetrica della varianza condizionata. Per il caso  $q = p = 1$  avremo

$$h_t = \omega + \alpha \epsilon_{t-1}^2 + \gamma S_{t-1}^- \epsilon_{t-1}^2 + \beta h_{t-1} \quad (2.29)$$

dove  $S_{t-1}^-$  è la variabile indicatrice che assume valore uno se la condizione  $\epsilon_{t-1} < 0$  è vera e valore zero altrimenti. Il coefficiente  $\gamma$  misura il diverso impatto che gli shock negativi hanno su  $h_t$ . Per essere in linea con le considerazioni di Black (1976), il segno atteso di  $\gamma$  è positivo: gli shock negativi amplificano la volatilità. Infine, un  $\gamma$  statisticamente non diverso da zero indica l'assenza di effetti *leverage*.

### 2.2.12 A-PARCH

Ding, Granger e Engle (1993) introducono la classe dei modelli ARCH a "potenza asimmetrica" (*Asymmetric Power ARCH*). Nel modello A-PARCH(1,1), la varianza condizionata è espressa dall'equazione

$$h_t^{\delta/2} = \omega + \alpha (|\epsilon_{t-1}| - \gamma \epsilon_{t-1})^\delta + \beta h_{t-1}^{\delta/2} \quad (2.30)$$

dove l'asimmetria è introdotta dal parametro  $\gamma$ , mentre  $\delta > 0$ . La caratteristica principale di questa classe di modelli è quella di comprendere molte delle tradizionali specificazioni GARCH come casi particolari. Così ad esempio,  $\delta = 2$  e  $\gamma = 0$  danno il modello GARCH(1,1); il solo  $\delta = 2$  rende il modello GJR(1,1), mentre  $\delta = 1$  implica dinamiche simili a quelle del modello TGARCH di cui sotto. Dal momento che tutte queste formulazioni sono incluse nella classe A-PARCH, un test del tipo rapporto di verosimiglianza (LR) può essere impiegato per verificare la significatività dei parametri. Di conseguenza, l'ipotesi nulla di uno specifico modello GARCH contro l'alternativa A-PARCH può essere verificata con relativa facilità.

### 2.2.13 TGARCH

Il modello *Threshold GARCH* è stato introdotto da Zakoïan (1994). La sua caratteristica principale è data dal fatto che ad essere modellata non è la varianza, ma la deviazione standard condizionata  $\sigma_t = \sqrt{h_t}$ . Nella versione TGARCH(1,1) il modello assume la forma

$$\sigma_t = \omega + \alpha^+ \epsilon_{t-1}^+ - \alpha^- \epsilon_{t-1}^- + \beta \sigma_{t-1} \quad (2.31)$$

dove  $\epsilon_t^+ = \max(\epsilon_t, 0)$  e  $\epsilon_t^- = \min(\epsilon_t, 0)$ . Questo approccio è strettamente connesso a quello utilizzato da Tong e Lim (1980) e da Tong (1983) per l'equazione della media condizionata. La deviazione standard condizionata è espressa come combinazione lineare dei valori passati di  $\epsilon_t$  e  $\sigma_t$ , mentre il regime in corso al tempo  $t$  dipende dal segno delle innovazioni passate. Si noti che la (2.31) può essere riparametrizzata come

$$\sigma_t = \omega + \alpha |\epsilon_{t-1}| + \gamma S_{t-1}^- \epsilon_{t-1} + \beta \sigma_{t-1} \quad (2.32)$$

In questo modo, nel modello TGARCH(1,1), la deviazione standard condizionata presenta la stessa forma funzionale del modello GJR(1,1). Inoltre, per  $\alpha^+ = \alpha^- = \alpha$  nella (2.31) e prendendo il valore assoluto delle innovazioni, si ottiene il modello di Schwert (1989) (equazione 2.24). Equivalentemente, basta porre  $\gamma = 0$  nella (2.32).

### 2.2.14 SWARCH

Hamilton e Susmel (1994) propongono un modello ARCH di tipo *Markov-Switching* (SWARCH) dove si assume l'esistenza di più regimi di volatilità (nel caso più semplice se ne considerano solo due). Il passaggio da uno stato di volatilità all'altro è determinato da un processo stocastico discreto  $\{s_t\}$  non osservabile, la cui struttura dinamica è regolata da una matrice di probabilità di transizione di Markov (si veda la Sezione 2.1). Si tratta di una naturale estensione dell'approccio di Hamilton (1989) per la media condizionata al caso di varianza non costante nel tempo. Formalmente avremo

$$y_t = \mu_t + \epsilon_t, \quad \epsilon_t = \tilde{\epsilon}_t \sqrt{g_{s_t}} \quad (2.33)$$

dove per  $\tilde{\epsilon}_t$  si assume un'evoluzione dinamica di tipo ARCH( $q$ ):

$$\tilde{\epsilon}_t = \eta_t \sqrt{h_t}, \quad \eta_t \sim iid N(0, 1) \quad (2.34)$$

Hamilton e Susmel (1994) considerano due diverse specificazioni per  $h_t$ ; una è rappresentata dal modello ARCH( $q$ ) di Engle (1982), l'altra considera un termine aggiuntivo,  $\gamma S_{t-1}^- \tilde{\epsilon}_{t-1}^2$ , allo scopo di catturare dinamiche asimmetriche nell'evoluzione di  $h_t$  secondo quanto descritto a proposito del modello GJR. Nel primo dei due casi parleremo semplicemente di modello SWARCH( $N, q$ ), dove  $N$  indica il numero di regimi considerati, mentre nel secondo si parlerà di modello SWARCH-L( $N, q$ ), con la "L" che indica un termine che tiene conto dell'effetto *leverage*.

Nella (2.33),  $\sqrt{g_{s_t}}$  rappresenta un fattore di scala moltiplicativo per la dinamica ARCH( $q$ ) di  $\tilde{\epsilon}_t$  che assume valore  $\sqrt{g_1}$  quando il processo si trova nello stato  $s_t = 1$ , valore  $\sqrt{g_2}$  quando  $s_t = 2$ , e così via. Il fattore per il primo stato,  $g_1$ , viene normalizzato ad uno tramite  $g_j \geq 1$  per  $j = 2, 3, \dots, N$ . L'idea di una simile standardizzazione è quella di descrivere il cambiamento di regime come cambiamento di scala del processo. Condizionatamente alla conoscenza del regime corrente e dei regimi passati, la varianza di  $\epsilon_t$  è data da

$$E(\epsilon_t^2 | s_t, s_{t-1}, \dots, s_{t-q}, \epsilon_{t-1}, \dots, \epsilon_{t-q}) = g_{s_t} \left[ \omega + \alpha_1 \left( \frac{\epsilon_{t-1}^2}{g_{s_{t-1}}} \right) + \dots + \alpha_q \left( \frac{\epsilon_{t-q}^2}{g_{s_{t-q}}} \right) \right] \equiv \sigma_t^2(s_t, s_{t-1}, \dots, s_{t-q}) \quad (2.35)$$

Il valore di  $\sigma_t^2$  è dunque determinato dallo stato di volatilità corrente e da quelli passati. Per ulteriori dettagli ed estensioni si rimanda a Hamilton e Susmel (1994), Cai (1994), Gray (1996) ed Avesani e Gallo (1996).

### 2.2.15 GQARCH

Sentana (1991 e 1995) introduce il modello ARCH quadratico (QARCH). Il termine quadratico è utilizzato in quanto il modello può essere interpretato come un'ap-

prossimazione di Taylor al secondo ordine di una qualche funzione incognita che descrive l'evoluzione della varianza condizionata. Nel caso generalizzato, il modello GQARCH(1,1) assume la forma

$$h_t = \omega + \alpha \epsilon_{t-1}^2 + \gamma \epsilon_{t-1} + \beta h_{t-1} \quad (2.36)$$

dove il parametro  $\gamma$  determina una dinamica asimmetrica per  $h_t$ . Si noti che questo modello è identico a quello AGARCH di Engle (1990), anche se diverse sono le motivazioni che introducono le due formulazioni. Engle (1990) motiva la sua specificazione per  $h_t$  come funzione di una parabola non centrata (si veda anche Engle e Ng, 1993), mentre Sentana (1991 e 1995) fa riferimento all'approssimazione di Taylor. La speranza è che la duplicazione di acronimo non crei confusione.

### 2.2.16 VS-GARCH

Il modello ARCH con cambiamento di volatilità (*Volatility-Switching*) è stato introdotto da Fornari e Mele (1996). In esso, la varianza condizionata segue il processo

$$h_t = \omega + \alpha \epsilon_{t-1}^2 + \beta h_{t-1} + \gamma S_{t-1} \frac{\epsilon_{t-1}^2}{h_{t-1}} \quad (2.37)$$

con  $S_{t-1} = 1$ , se  $\epsilon_t > 0$ ,  $S_{t-1} = 0$ , se  $\epsilon_t = 0$  e  $S_{t-1} = -1$ , se  $\epsilon_t < 0$ . Il termine  $\epsilon_{t-1}^2/h_{t-1}$  misura quanto un determinato residuo al quadrato devia dal suo valore atteso  $h_t$ , mentre  $S_t$  indica il segno del  $t$ -esimo residuo. Il modello (2.37), ad esempio, è in grado di generare dati dove grandi shock inattesi di segno negativo aumentano  $h_t$ , grandi shock di segno positivo diminuiscono  $h_t$ , piccoli shock negativi diminuiscono  $h_t$  e piccoli shock positivi aumentano  $h_t$ . Il livello dell'asimmetria dipende invece dal parametro  $\gamma$ , oltre che dalla forza relativa dei residui.

Fornari e Mele (1997) propongono una leggera estensione della (2.37), nella quale sono inclusi due parametri aggiuntivi:

$$h_t = \omega_1 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} + \left( \omega_2 + \alpha_2 \epsilon_{t-1}^2 + \beta_2 \frac{\epsilon_{t-1}^2}{h_{t-1}} \right) S_{t-1} \quad (2.38)$$

La principale questione sollevata da Fornari e Mele (1996 e 1997) è che la volatilità corrente è in parte determinata dagli shock inattesi di quella passata.

## 2.3 Considerazioni conclusive

In questo Capitolo sono stati presentati i principali modelli non-lineari per la media condizionata utilizzati in letteratura e ci siamo poi soffermati sulla descrizione dei modelli (G)ARCH che, a partire dal lavoro di Engle (1982) prima e da quello di Bollerslev (1986) poi, hanno riscosso un enorme successo nel descrivere la dinamica della varianza condizionata delle serie storiche finanziarie. La caratteristica comune alla gran parte delle formulazioni presentate, sia per la media che per la varianza condizionata, è quella di cercare di formalizzare l'idea dell'esistenza di un limitato numero di regimi, o stati, che caratterizzano l'economia. In questo ambito si collocano, ad esempio, i modelli

TAR, EAR, NLNAR, *Markov-Switching* (per la media condizionata) e quelli GARCH non-lineari, GJR, TGARCH, SWARCH e VS-GARCH (per la varianza condizionata). Concettualmente, i modelli con cambiamento di regime possono essere pensati come un insieme di formulazioni lineari, ciascuna delle quali corrispondente ad un particolare regime. Le diversità tra i vari modelli dipendono dal modo con cui lo stato prevalente in un determinato momento evolve nel tempo. Così è possibile distinguere due classi generali di modelli: quelli della prima classe assumono che il cambiamento di regime sia determinato da una variabile osservabile, mentre quelli della seconda classe ipotizzano che il passaggio da uno stato all'altro sia caratterizzato da un sottostante processo stocastico non osservabile. Si noti che, nel primo caso, il regime in corso e quello immediatamente precedente sono noti con certezza (nonostante siano stati individuati mediante tecniche statistiche), mentre, nel secondo caso, non sapremo mai con esattezza quale sia il regime in corso ad un certo istante  $t$ , ma potremo soltanto assegnare una certa probabilità alla sua manifestazione. È questo il caso dei modelli di tipo *Markov-Switching*, ampiamente discussi da Hamilton (1989) e la cui estensione all'equazione della varianza condizionata (modelli SWARCH) è avvenuta per opera di Hamilton e Susmel (1994). Al primo gruppo di rappresentazioni con cambiamento di regime appartengono invece i modelli TAR, EAR, TGARCH, VS-GARCH, ecc. Di questa classe fa parte anche la famiglia dei modelli a transizione liscia che presenteremo nei capitoli successivi. In tali modelli, si ammette la presenza di più stati dell'economia e la transizione da un regime all'altro avviene gradualmente sulla base di una qualche relazione funzionale. Tra le caratteristiche principali dei modelli vi è quella di comprendere molte delle formulazioni descritte in questo Capitolo come casi particolari, sia per la media (*Smooth Transition Autoregressive Models*, STAR) che per la varianza condizionata (*Smooth Transition GARCH*, STGARCH). Di questi modelli e delle loro proprietà ci occuperemo nei Capitoli successivi.



## Capitolo 3

# Modelli autoregressivi a transizione liscia

La moderna econometria sulle serie temporali ha proposto una vasta gamma di rappresentazioni non-lineari. L'abbondanza di soluzioni diverse pone problemi per l'utilizzatore poiché non è generalmente in grado di decidere quale tipo di modello adattare ai dati. Costruire una procedura utilizzabile per specificare un appropriato modello o famiglia di modelli, richiede inevitabilmente di restringere il campo di scelta.

Haggan, Heravi e Priestley (1984)<sup>1</sup> assumono che l'equazione da stimare faccia parte della famiglia dei modelli stato-dipendenti, dove le variabili economiche vengono modellate definendo differenti stati, o regimi, dell'economia, ed ammettendo la possibilità che il comportamento dinamico di tali variabili dipenda esclusivamente dallo stato attivo in quel particolare momento. Questa classe contiene molte formulazioni ben note in letteratura: si spazia dai modelli lineari di tipo ARMA, a quelli bilineari (BL), autoregressivi esponenziali (EAR), autoregressivi di tipo *threshold* (TAR), fino ad arrivare ai più recenti modelli autoregressivi a transizione liscia (STAR).

Scopo di questo Capitolo è introdurre i modelli autoregressivi a transizione liscia e discuterne le principali caratteristiche che saranno poi analizzate nei Capitoli successivi. Con riferimento agli articoli di Teräsvirta (1994) ed Eitrheim e Teräsvirta (1996), la discussione è basata sulla presentazione di una procedura empirica di specificazione dei modelli, arricchita da una esemplificazione su dati reali.

Il Capitolo si sviluppa come segue: nella Sezione 3.1, dopo aver derivato i modelli STAR a partire dalla rappresentazione stato-dipendente, ne vengono discusse le proprietà e possibili generalizzazioni, mentre la Sezione 3.2, oltre a presentare sinteticamente la procedura empirica di specificazione dei modelli, introduce la serie storica oggetto di analisi nel resto del Capitolo. Le Sezioni successive (3.3-3.8) si concentrano sui singoli punti della procedura di specificazione, mettendo di volta in volta in evidenza come questa operi in pratica, con riferimento alla serie del Tasso sui Fondi Federali U.S.A. Infine, la Sezione 3.9 presenta alcune considerazioni conclusive.

---

<sup>1</sup>Si veda anche Priestley (1980).

### 3.1 I modelli

Seguendo Haggan *et al.* (1984), la rappresentazione stato-dipendente per una serie temporale univariata  $y_t$ ,  $t = 1, \dots, T$ , assume la seguente forma non-lineare:

$$y_t - \phi_1(\mathbf{x}_{t-1})y_{t-1} - \dots - \phi_p(\mathbf{x}_{t-1})y_{t-p} = \mu(\mathbf{x}_{t-1}) + \epsilon_t + \psi_1(\mathbf{x}_{t-1})\epsilon_{t-1} + \dots \\ \dots + \psi_q(\mathbf{x}_{t-1})\epsilon_{t-q} \quad (3.1)$$

dove  $\{\epsilon_t\}$  è una sequenza di errori casuali indipendenti, normalmente e identicamente distribuiti, con media zero e varianza costante  $\sigma_\epsilon^2$ .  $\mathbf{x}_{t-1} = (y_{t-1}, \dots, y_{t-p}, \epsilon_{t-1}, \dots, \epsilon_{t-q})$  rappresenta il vettore di stato, o regime, del processo  $\{Y_t\}$ .  $\mathbf{x}_{t-1}$  indica che la sola informazione rilevante per la dinamica futura di  $y_t$  è contenuta nel vettore di stato, insieme con il valore futuro di  $\{\epsilon_t\}$ ;  $\phi_i(\mathbf{x}_{t-1})$ ,  $i = 1, \dots, p$  e  $\psi_j(\mathbf{x}_{t-1})$ ,  $j = 1, \dots, q$  esprimono una rappresentazione locale del modello non-lineare.

La modellistica stato-dipendente estende l'idea dei modelli di tipo ARMA al caso in cui un processo  $\{Y_t\}$  segue un'evoluzione non-lineare il cui comportamento può essere localmente rappresentato da un modello lineare. In questo contesto, il termine "locale" è riferito a piccole deviazioni del processo dal suo stato corrente. Alternativamente, si potrebbe dire che questi modelli permettono alla dinamica di  $y_t$  di variare tra un infinito numero di regimi, poiché i parametri sono funzioni delle variabili di stato<sup>2</sup>. Questa interpretazione dei modelli stato-dipendenti ci consente di considerarli parte della classe dei modelli con cambiamento di regime, i quali possono essere pensati come un insieme di parametrizzazioni lineari, ognuna delle quali corrispondente ad un particolare regime del processo. Ad ogni istante sarà attivo solo uno dei modelli (o una loro combinazione lineare), in corrispondenza del regime in corso.

Specificando opportunamente la forma dei coefficienti incogniti  $\mu(\mathbf{x}_{t-1})$ ,  $\phi_i(\mathbf{x}_{t-1})$  e  $\psi_j(\mathbf{x}_{t-1})$ , è agevole mostrare come dalla (3.1) si ricavano i modelli tradizionali descritti nella Sezione 2.1. Ad esempio, se assumiamo che  $\mu$ ,  $\{\phi_i\}$  e  $\{\psi_j\}$  siano costanti e indipendenti da  $\mathbf{x}_{t-1}$ , si deriva subito il modello ARMA(p,q), che ha come casi particolari i più semplici AR(p) e MA(q). In modo assolutamente analogo si ottengono i modelli bilineari, EAR, TAR, ecc<sup>3</sup>.

Quello a cui siamo interessati in questa sede è mostrare come dalla (3.1) si ricavano i modelli autoregressivi a transizione liscia (STAR). Assumiamo che  $\psi_j = 0 \forall j$ ,  $\mu(\mathbf{x}_{t-1}) = \pi_0 + \theta_0 F(s_t; \gamma, c)$  e  $\phi_i(\mathbf{x}_{t-1}) = \pi_i + \theta_i F(s_t; \gamma, c)$ ,  $i = 1, \dots, p$  e, dopo qualche passaggio algebrico, la (3.1) diventa:

$$y_t = \pi_0 + \pi' \mathbf{w}_t + (\theta_0 + \theta' \mathbf{w}_t) F(s_t; \gamma, c) + \epsilon_t, \quad t = 1, \dots, T \quad (3.2)$$

dove  $\pi = (\pi_1, \dots, \pi_p)'$ ,  $\theta = (\theta_1, \dots, \theta_p)'$ ,  $\mathbf{w}_t = (y_{t-1}, \dots, y_{t-p})'$  e  $\epsilon_t \sim iid N(0, \sigma_\epsilon^2)$ . In termini espliciti, la (3.2) può essere scritta come:

$$y_t = \pi_0 + \pi_1 y_{t-1} + \dots + \pi_p y_{t-p} + (\theta_0 + \theta_1 y_{t-1} + \dots + \theta_p y_{t-p}) F(s_t; \gamma, c) + \epsilon_t \quad (3.3)$$

<sup>2</sup>In questo contesto ci limitiamo a considerare il vettore di stato composto da sole variabili osservabili o da trend temporali; la presenza di variabili latenti esula dalla nostra trattazione. Più in dettaglio, se il vettore di stato è caratterizzato da un processo stocastico non osservabile la cui dinamica è governata da una catena di Markov, dalla (3.1) si ricava la classe dei modelli *Markov-Switching*, sviluppati da Hamilton (1989) e discussi nella Sezione 2.1.

<sup>3</sup>Per ulteriori dettagli si rimanda al lavoro di Haggan *et al.* (1984).

In generale, la (3.2) rappresenta la classe dei modelli autoregressivi a transizione liscia (STAR), dove la non-linearità è introdotta dalla funzione di transizione  $F(s_t; \gamma, c)$ . Si tratta di una funzione continua, assunta almeno due volte differenziabile, di solito limitata tra zero ed uno, della quale sono comunemente considerate due formulazioni: logistica ed esponenziale<sup>4</sup>. Se ipotizziamo che

$$F(s_t; \gamma, c) = \frac{1}{1 + \exp\{-\gamma(s_t - c)\}}, \quad \gamma > 0 \quad (3.4)$$

allora abbiamo il modello STAR logistico (LSTAR), mentre se consideriamo

$$F(s_t; \gamma, c) = 1 - \exp\{-\gamma(s_t - c)^2\}, \quad \gamma > 0 \quad (3.5)$$

otteniamo il modello STAR esponenziale (ESTAR). Nella (3.4) e nella (3.5),  $s_t$  indica la variabile di transizione e può essere una variabile endogena ritardata ( $s_t = y_{t-d}$  per un certo intero  $d > 0$ ), una variabile esogena ( $s_t = z_{kt}$ ), o una funzione di variabili endogene ritardate e/o esogene ( $s_t = h(\cdot)$  per alcune funzioni  $h$ ). Un'altra possibilità è quella di considerare  $s_t$  uguale a (o funzione di) un trend temporale ( $s_t = t$  oppure  $s_t = h(t)$ ), ottenendo in tal modo il modello autoregressivo con parametri che variano gradualmente (*smoothly*) nel tempo (TV-AR), introdotto da Lin e Teräsvirta (1994)<sup>5</sup>.

La funzione di transizione è caratterizzata anche dai parametri  $\gamma$  e  $c$ . Il primo, vincolato ad assumere valori strettamente positivi a causa di una restrizione di identificazione, esprime il grado di lisciamento della funzione, mentre il secondo indica la soglia di passaggio da un regime all'altro del modello<sup>6</sup>.

### 3.1.1 Proprietà dei modelli STAR

In questa sezione ci proponiamo di sottolineare più dettagliatamente le caratteristiche dei modelli STAR, mettendo anche in evidenza come essi contengano al loro interno altre famiglie di modelli come casi particolari.

Con riferimento al modello STAR logistico (LSTAR), ottenuto dalla (3.2) insieme con la (3.4),

$$y_t = \pi_0 + \pi' \mathbf{w}_t + (\theta_0 + \theta' \mathbf{w}_t) \frac{1}{1 + \exp\{-\gamma(s_t - c)\}} + \epsilon_t, \quad \gamma > 0 \quad (3.6)$$

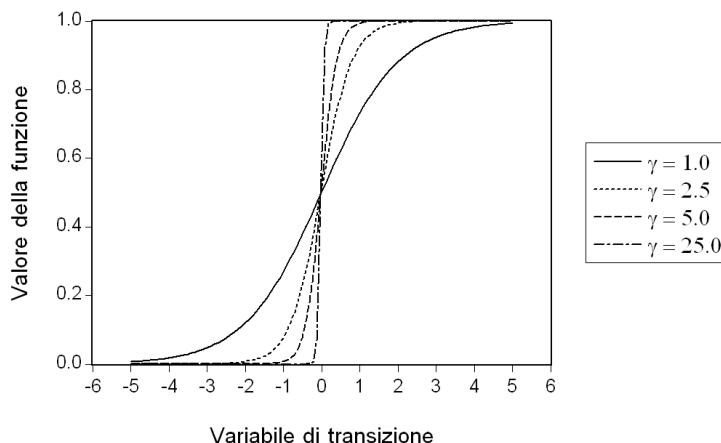
possiamo subito evidenziare che per  $\gamma = 0$  il modello diventa lineare, ed in particolare un modello autoregressivo di ordine  $p$ . Il modello diventa lineare anche quando  $\theta_0 = 0$  e  $\theta = 0$ .

<sup>4</sup>Chan e Tong (1986) hanno utilizzato la funzione di distribuzione cumulata della normale standard. Il loro modello non ha, tuttavia, riscosso grosso favore in pratica poiché la derivazione del test di linearità è, in tal caso, più complicata rispetto alle soluzioni logistica ed esponenziale.

<sup>5</sup>Un'ulteriore possibilità è quella di considerare soluzioni miste in cui  $s_t$  è funzione di variabili endogene ritardate, variabili esogene e trend temporali. In tal modo però si complica il modello, incontrando problemi in fase di stima dei parametri. È pertanto preferibile una rappresentazione più parsimoniosa rispetto a una più ricca, quando il miglioramento in termini di bontà di adattamento ai dati non è tale da giustificare l'aumento eccessivo dei tempi di stima.

<sup>6</sup>Ad esempio, il passaggio da una fase di recessione ad una di espansione dell'economia, il quale avviene più o meno gradualmente a seconda del valore di  $\gamma$ .



Figura 3.1: Funzione di transizione logistica per diversi valori di  $\gamma$  e  $c = 0$ .

Per  $\gamma$  che tende a zero, la funzione logistica tende alla costante 0.5; mentre se  $\gamma$  tende all'infinito, essa tende alla funzione indicatrice  $\mathfrak{I}(s_t > c)$ , che assume valore uno se la condizione è vera ( $s_t > c$ ) e valore zero altrimenti. Di conseguenza, il passaggio tra i due regimi, ottenuti in corrispondenza di  $F(s_t; \gamma, c) = 0$  e  $F(s_t; \gamma, c) = 1$ , diventa quasi istantaneo per  $s_t = c$ .<sup>7</sup> Al limite ( $\gamma \rightarrow \infty$ ) avremo che il modello LSTAR coincide con il modello autoregressivo di tipo *threshold* a due regimi (TAR). Inoltre, se  $s_t = y_{t-d}$ , il modello è detto *self-exciting* TAR (SETAR).<sup>8</sup>

La Figura 3.1 mostra l'andamento della funzione di transizione logistica per diversi valori del parametro  $\gamma$ . Come si può vedere, già per  $\gamma = 25$ , il passaggio da un regime all'altro diventa quasi istantaneo, mentre più  $\gamma$  è piccolo (ma pur sempre positivo), tanto più il cambiamento di stato risulta graduale.

La Figura 3.2 mostra alcune simulazioni ricavate dal modello LSTAR

$$y_t = \pi_0 + \pi_1 y_{t-1} + (\theta_0 + \theta_1 y_{t-1}) \frac{1}{1 + \exp\{-\gamma y_{t-1}\}} + \epsilon_t \quad (3.7)$$

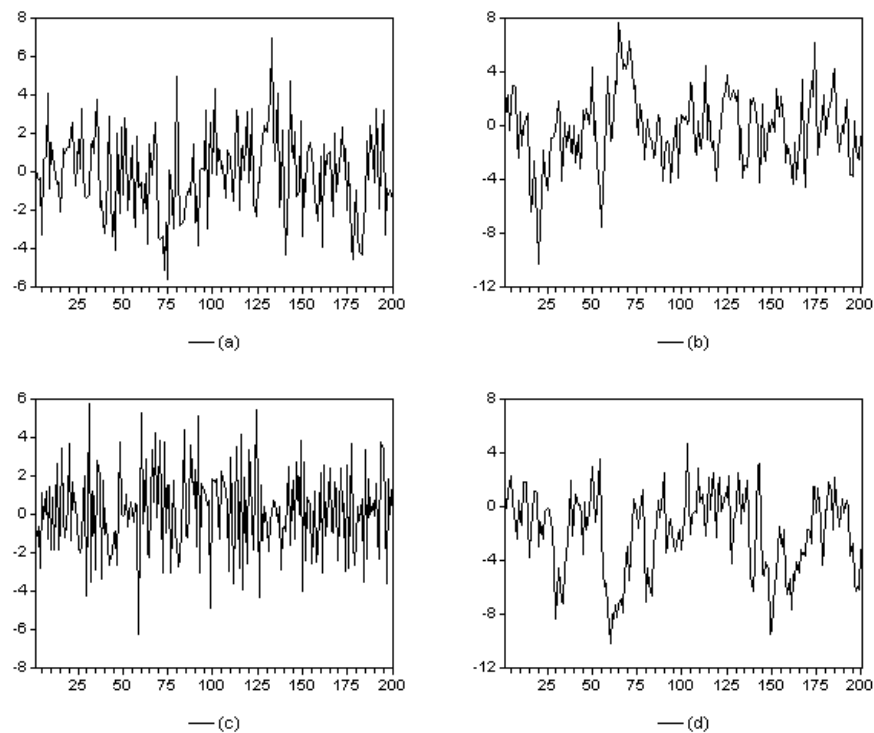
per diverse combinazioni dei parametri  $\pi_i$  e  $\theta_i$ ,  $i = 0, 1$  e con  $\gamma$  e  $c$  fissi rispettivamente pari ad uno e zero (si veda la Figura per tutti i dettagli). Il numero di simulazioni è pari a 200 e il termine di disturbo è ricavato estraendo dei numeri casuali a partire da una distribuzione Normale a media zero e varianza quattro.

I due regimi sono associati a piccoli o grandi valori, relativamente alla soglia  $c$ , della variabile di transizione  $s_t$ . Questa caratteristica ci suggerisce di utilizzare il modello LSTAR quando il comportamento di  $y_t$  è tale da mostrare la presenza di asimmetrie nella sua evoluzione. Ad esempio, nella letteratura sul ciclo economico si è soliti distinguere comportamenti asimmetrici delle variabili economiche a seconda che l'economia si trovi in una fase di espansione o in una di recessione.

<sup>7</sup>Si noti che  $F(c; \gamma, c) = 0.5$ .

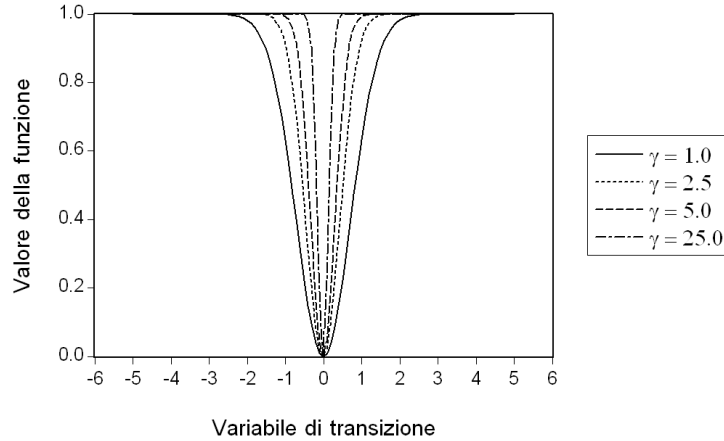
<sup>8</sup>Una esauriente trattazione dei modelli TAR e SETAR può essere trovata in Tong (1990).

Figura 3.2: Alcune simulazioni di processi STAR logistici con  $\gamma = 1$ ,  $c = 0$  ed  $s_t = y_{t-1}$ . (a):  $\pi_0 = 0.2$ ,  $\pi_1 = 0.5$ ,  $\theta_0 = 0.1$  e  $\theta_1 = -0.5$ . (b):  $\pi_0 = 0.1$ ,  $\pi_1 = 0.8$ ,  $\theta_0 = -0.1$  e  $\theta_1 = -0.3$ . (c):  $\pi_0 = -0.4$ ,  $\pi_1 = -0.6$ ,  $\theta_0 = 0.2$  e  $\theta_1 = 0.5$ . (d):  $\pi_0 = 0.1$ ,  $\pi_1 = 0.9$ ,  $\theta_0 = -0.3$  e  $\theta_1 = -0.9$ .



Altre volte può essere conveniente associare il cambiamento di regime a piccoli o grandi valori di  $s_t$  presi in valore assoluto. A titolo esemplificativo, si potrebbe pensare che il comportamento del tasso di cambio reale dipenda dall'entità delle deviazioni dalla parità del potere d'acquisto (PPP). In particolare, si ipotizza che la relazione di equilibrio descritta dalla PPP sia mantenuta dai profitti da arbitraggio su merci. La presenza di costi di transazione, come i costi di trasporto e stoccaggio, può dar luogo a diversi stati di evoluzione del tasso di cambio reale. Questo implica l'esistenza di una banda intorno al tasso di equilibrio all'interno della quale non c'è la tendenza del tasso di cambio reale a ritornare verso il suo valore di equilibrio. Al di fuori di questa banda, l'arbitraggio su merci diventa vantaggioso, spingendo di nuovo il tasso di cambio reale all'interno della banda.<sup>9</sup> Dinamiche simili a quella descritta possono essere ottenute

<sup>9</sup>Per alcune applicazioni si vedano Michael, Nobay e Peel (1997). Altre applicazioni dei modelli ESTAR relative all'aggiustamento di lungo periodo della domanda di moneta si possono trovare in Michael, Nobay e Peel (1999).

Figura 3.3: Funzione di transizione esponenziale per diversi valori di  $\gamma$  e  $c = 0$ .

utilizzando una funzione di transizione esponenziale del tipo della (3.5) nella (3.2):

$$y_t = \pi_0 + \pi' \mathbf{w}_t + (\theta_0 + \theta' \mathbf{w}_t)(1 - \exp\{-\gamma(s_t - c)^2\}) + \epsilon_t, \quad \gamma > 0 \quad (3.8)$$

Il risultante modello STAR esponenziale (ESTAR) diventa lineare per  $\gamma = 0$  o per  $\theta_0 = 0$  e  $\theta = 0$ . In Figura 3.3 è riportato l'andamento della funzione di transizione esponenziale per diversi valori di  $\gamma$ . Se facciamo tendere  $\gamma$  a zero o all'infinito, in entrambi i casi la funzione di transizione tende alla costante 1 e il modello diventa lineare. Inoltre, a differenza del modello logistico, quello esponenziale non ha il modello (SE)TAR come caso particolare. Tuttavia, la rappresentazione ESTAR costituisce una sottile generalizzazione di quella EAR proposta da Haggan e Ozaki (1981); quest'ultima si ricava dalla (3.8) assumendo  $\theta_0 = 0$  e  $c = 0$ .<sup>10</sup>

Alcune simulazioni di modelli ESTAR

$$y_t = \pi_0 + \pi_1 y_{t-1} + (\theta_0 + \theta_1 y_{t-1})(1 - \exp\{-\gamma y_{t-1}^2\}) + \epsilon_t \quad (3.9)$$

per diverse combinazioni dei parametri  $\pi_i$  e  $\theta_i$ ,  $i = 0, 1$  e con  $\gamma$  e  $c$  fissi rispettivamente pari ad uno e zero sono riportate in Figura 3.4.

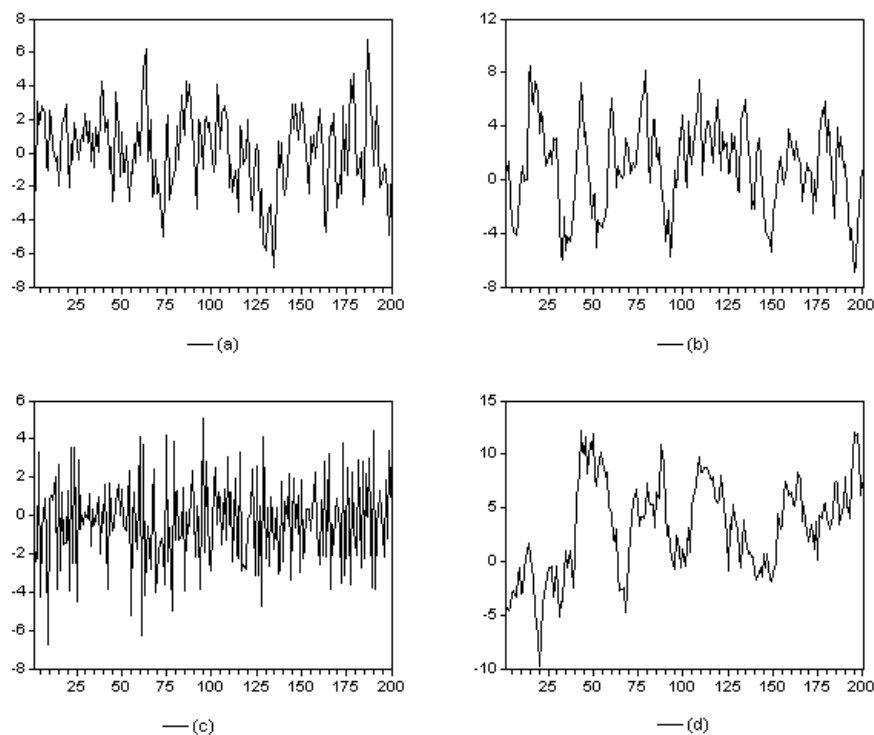
### 3.1.2 Alcune generalizzazioni

Lo scopo di questa sezione è quello di presentare sinteticamente delle estensioni dei modelli STAR, alcune delle quali saranno analizzate in modo più dettagliato nei capitoli successivi.

Le generalizzazioni che consideriamo possono essere ripartite in due gruppi: un primo gruppo riguarda la parte deterministica del modello, mentre un secondo gruppo si riferisce alla parte stocastica  $\epsilon_t$ .

<sup>10</sup>Nella formulazione originaria di Haggan e Ozaki (1981),  $s_t$  è pari ad una delle variabili endogene ritardate del vettore di stato (generalmente  $y_{t-1}$ ).

Figura 3.4: Alcune simulazioni di processi STAR esponenziali con estrazioni da una distribuzione Normale a media 0 e varianza 4 e con  $\gamma = 1$ ,  $c = 0$  ed  $s_t = y_{t-1}$ . (a):  $\pi_0 = 0.2$ ,  $\pi_1 = 0.5$ ,  $\theta_0 = 0.1$  e  $\theta_1 = -0.5$ . (b):  $\pi_0 = 0.1$ ,  $\pi_1 = 0.8$ ,  $\theta_0 = -0.1$  e  $\theta_1 = -0.3$ . (c):  $\pi_0 = -0.4$ ,  $\pi_1 = -0.6$ ,  $\theta_0 = 0.2$  e  $\theta_1 = 0.5$ . (d):  $\pi_0 = 0.1$ ,  $\pi_1 = 0.9$ ,  $\theta_0 = -0.3$  e  $\theta_1 = -0.9$ .



Riguardo al primo gruppo di condizioni, nella Sezione 3.1 è stata presentata la classe dei modelli autoregressivi a transizione liscia (STAR) espressa dalla (3.2). In quella sede si è assunto che il vettore  $\mathbf{w}_t$  fosse composto da sole variabili endogene ritardate. Una possibile estensione è quella di allargare la composizione del vettore  $\mathbf{w}_t$ , e quindi dell'insieme informativo a disposizione, fino a comprendervi  $k$  variabili esogene; chiameremo  $\mathbf{w}_t^*$  il vettore così risultante:

$$\mathbf{w}_t^* = (y_{t-1}, \dots, y_{t-p}, z_{1t}, \dots, z_{kt})'$$

In tal modo la (3.2) può essere così riscritta:

$$y_t = \pi_0 + \pi^* \mathbf{w}_t^* + (\theta_0 + \theta^* \mathbf{w}_t^*) F(s_t; \gamma, c) + \epsilon_t^*, \quad t = 1, \dots, T \quad (3.10)$$

e assume la denominazione di classe dei modelli a transizione liscia (STR),<sup>11</sup> dove i

<sup>11</sup>Al limite il vettore  $\mathbf{w}_t^*$  può essere composto da sole variabili esogene, in analogia al modello di regressione classico.

vettori  $\pi^*$  e  $\theta^*$  sono composti da  $m = p + k$  componenti.

Seguendo Granger e Teräsvirta (1993), un'altra possibilità si riferisce all'argomento della funzione di transizione  $F(s_t; \gamma, c)$ . Si è visto che sia la funzione logistica che quella esponenziale sono caratterizzate dalla variabile di transizione  $s_t$  e dai parametri  $\gamma$  e  $c$ . Se assumiamo che  $z_t = (s_t - c)$ , la (3.4) e la (3.5) possono essere così riscritte:

$$F(z_t) = \frac{1}{1 + \exp(-\gamma z_t)}, \quad \gamma > 0 \quad (3.11)$$

e

$$F(z_t) = 1 - \exp(-\gamma z_t^2), \quad \gamma > 0 \quad (3.12)$$

Un particolare modello STAR<sup>12</sup> può essere definito assumendo nella (3.11) e nella (3.12)  $z_t = (y_{t-d} - \pi_0 - \pi' \mathbf{w}_{t-d})$ ,  $d > 0$ . Si tratta del modello STAR con deviazione (STAR-D) nel quale la non-linearità è determinata dallo scarto tra i valori passati della variabile dipendente per un certo  $d > 0$  e un valore fisso, espresso dalla parte lineare del modello. Si può ulteriormente generalizzare il modello nel modo seguente:

$$z_t = \sum_{d=1}^D (y_{t-d} - \pi_0 - \pi' \mathbf{w}_{t-d}), \quad \gamma > 0 \quad (3.13)$$

In modo più compatto si può scrivere la (3.13):

$$z_t = \sum_{d=1}^D \tilde{\epsilon}_{t-d}, \quad \gamma > 0 \quad (3.14)$$

dove  $\tilde{\epsilon}_{t-d} = y_{t-d} - \pi_0 - \pi' \mathbf{w}_{t-d}$ ,  $d > 0$ . Inserendo  $\tilde{\epsilon}_{t-d}$  nella (3.11) e nella (3.12), si ottiene rispettivamente:

$$F(\tilde{\epsilon}_{t-d}; \gamma) = \frac{1}{1 + \exp\{-\gamma(\tilde{\epsilon}_{t-d})\}}, \quad \gamma > 0, d > 0 \quad (3.15)$$

e

$$F(\tilde{\epsilon}_{t-d}; \gamma) = 1 - \exp\{-\gamma(\tilde{\epsilon}_{t-d}^2)\}, \quad \gamma > 0, d > 0 \quad (3.16)$$

che, nella (3.2) al posto di  $F(s_t; \gamma, c)$  danno luogo ai modelli LSTAR-D ed ESTAR-D.

Il secondo gruppo di estensioni riguarda il termine di disturbo  $\epsilon_t$ . Si è ipotizzato che  $\epsilon_t$  fosse indipendente, normalmente e identicamente distribuito con media zero e varianza costante  $\sigma_\epsilon^2$ . L'ipotesi di normalità non è strettamente necessaria e può essere rilassata fino ad assumere che  $\epsilon_t$  sia una sequenza di differenze di martingale rispetto alla storia passata di  $y_t$ , indicata con  $I_{t-1} = \{y_{t-1}, y_{t-2}, \dots, y_1\}$ :

$$E[\epsilon_t | I_{t-1}] = 0 \quad t = 2, 3, \dots \quad (3.17)$$

L'ipotesi di normalità congiunta degli  $\epsilon_t$  è utilizzata in fase di stima in quanto, così facendo, le stime dei minimi quadrati non-lineari (NLS) sono equivalenti alle stime di massima verosimiglianza. La più semplice ipotesi di differenze di martingale è

<sup>12</sup>Quanto diremo vale sia che consideriamo  $\mathbf{w}_t$  (STAR), sia  $\mathbf{w}_t^*$  (STR).

sufficiente per interpretare le stime NLS come stime di quasi-massima verosimiglianza (Davidson e MacKinnon, 1993).

La (3.17) non basta ad escludere la possibilità che momenti di ordine superiore, ad esempio la varianza, dipendano dai valori passati di  $y_t$ . Nel seguito di questo Capitolo consideriamo valida l'ipotesi di non dipendenza di  $\epsilon_t$  dalla storia passata per i momenti di ordine superiore al primo, ma si tenga presente la possibilità che la varianza condizionata di  $\epsilon_t$  possa non essere costante:

$$E[\epsilon_t^2 | I_{t-1}] = h_t \quad (3.18)$$

dove  $h_t$  può seguire, ad esempio, una dinamica di tipo GARCH. Condizionatamente a  $I_{t-1}$ , avremo dunque che

$$\epsilon_t | I_{t-1} \sim (0, h_t). \quad (3.19)$$

Questa eventualità sarà analizzata in modo specifico nei Capitoli successivi.

## 3.2 Procedura empirica di specificazione

Il primo problema che il ricercatore deve affrontare quando si appresta ad analizzare un fenomeno economico dal punto di vista statistico-econometrico consiste nella specificazione del modello più idoneo da adattare ai dati.

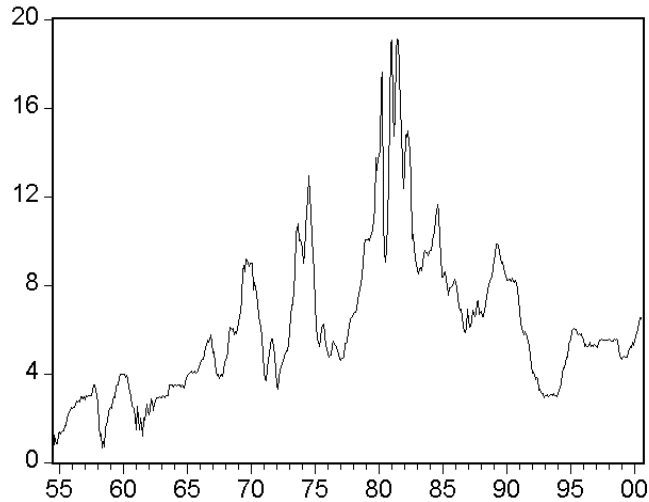
In certi casi è la teoria economica a suggerire un certo tipo di rappresentazione; il più delle volte, invece, la teoria economica non è in grado di fornire un contributo apprezzabile nella selezione del modello. È questo il caso dell'analisi statistica delle serie temporali nel cui ambito, a partire dai dati, si cerca di selezionare il modello che meglio degli altri li interpola, procedendo essenzialmente in modo del tutto ateorico. A differenza dell'approccio strutturale classico, sembra che questo secondo modo di procedere porti a miglioramenti significativi nei risultati ottenuti, prevalentemente in fase di previsione. Tuttavia, non essendo vincolato da alcuna teoria sottostante, il ricercatore è libero di scegliere il modello che ritiene più opportuno tra tutti quelli che superano una determinata batteria di test. Così facendo diventa difficile fornire un'accurata interpretazione economica dei risultati ottenuti.

Granger (1993) suggerisce di utilizzare una procedura che va dallo specifico al generale quando si vogliono impiegare modelli non-lineari per descrivere il comportamento e le caratteristiche di una particolare variabile economica.

La strategia proposta da Teräsvirta (1994) per la specificazione dei modelli autoregressivi a transizione liscia (STAR) segue le raccomandazioni di Granger (1993) e si basa sui seguenti passi successivi:

1. Specificare un appropriato modello lineare autoregressivo di ordine  $p$  [AR(p)] per la variabile oggetto di indagine.
2. Testare l'ipotesi nulla di linearità contro l'alternativa di non-linearità di tipo STAR, per diverse scelte della variabile di transizione  $s_t$ .
3. Se l'ipotesi di linearità è rifiutata, selezionare la forma della funzione di transizione  $F(s_t; \gamma, c)$ , un'appropriata variabile di transizione, e scegliere fra le parametrizzazioni LSTAR, ESTAR, LSTAR-D o ESTAR-D.

Figura 3.5: Andamento del Tasso sui Fondi Federali U.S.A. Dati mensili destagionalizzati, luglio 1954 – settembre 2000.

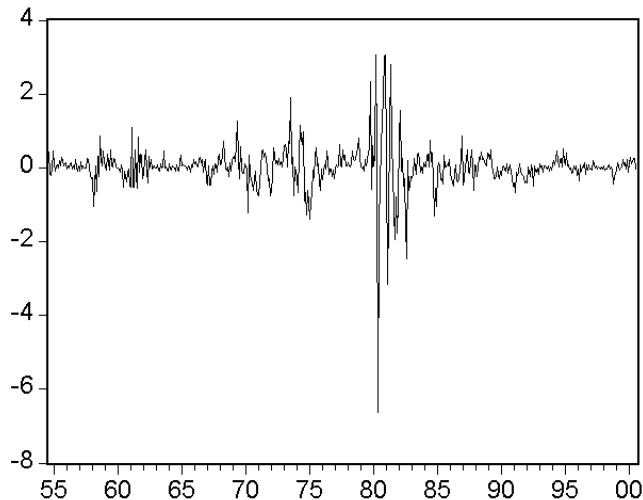


4. Stimare i parametri del modello scelto.
5. Valutare il modello stimato attraverso l'applicazione di diverse batterie di test.
6. Modificare, se necessario, il modello stimato (e ritornare al punto 4).
7. Utilizzare il modello per scopi descrittivi, di previsione o di simulazione.

Ciascuno di questi punti sarà dettagliatamente analizzato nei paragrafi successivi dove affiancheremo ad una trattazione teorica del singolo argomento, considerazioni empiriche su dati reali. A tale scopo, consideriamo la serie del Tasso sui Fondi Federali U.S.A., la cui ampiezza campionaria va dal luglio 1954 al settembre 2000, dati mensili destagionalizzati, per un totale di 555 osservazioni. In Figura 3.5 è riportata la dinamica della serie per il periodo considerato. L'osservazione del grafico lascia presagire la possibilità che la serie sia caratterizzata dalla presenza di una radice unitaria, eventualità che viene confermata dal test ADF<sup>13</sup>, la cui statistica test è pari a -2.5801 e si trova ampiamente nella regione di accettazione dell'ipotesi nulla, limitata dai valori critici tabulati di -3.4446 e -2.8671, rispettivamente per  $\alpha = 0.01$  e 0.05. Per questa ragione si è deciso di applicare ai dati l'operatore differenze prime, ricavando così la serie delle variazioni mensili, il cui andamento è riportato in Figura 3.6. La modellistica presentata nel seguito del Capitolo sarà quindi applicata alla serie delle differenze prime, mentre il campione utilizzato per la stima si riferirà al periodo agosto 1954 – dicembre 1994.<sup>14</sup> Il periodo rimanente (gennaio 1995 – settembre 2000) potrà essere impiegato per una valutazione del modello in un contesto di tipo *ex-post*.

<sup>13</sup>Per un'eccellente trattazione del test *Augmented Dickey-Fuller* per la presenza di radici unitarie su dati finanziari, si rimanda a Gallo e Pacini (2002). Inoltre, per un'estensione del test al caso *Smooth Transition*

Figura 3.6: Differenze prime del Tasso sui Fondi Federali U.S.A., agosto 1954 – settembre 2000.



### 3.3 Specificazione del modello lineare AR(p)

L'elemento principale da considerare nella specificazione del modello lineare AR(p) per la media condizionata è la scelta dell'ordine autoregressivo  $p$ . In tale contesto, si dovrebbe fare in modo che i residui del modello stimato corrispondano approssimativamente ad un *white noise* (WN); questo perché i test di non-linearità utilizzati nei passi successivi sono sensibili alla presenza di autocorrelazione residua.<sup>15</sup> Questa evenienza può essere sufficiente per rifiutare l'ipotesi nulla di linearità; in un tale contesto è pertanto preferibile una sovra-parametrizzazione piuttosto che una sotto-parametrizzazione del modello.

L'assenza di autocorrelazione congiunta nei residui fino ad un certo ritardo  $k$  è solitamente verificata mediante la statistica  $Q$  di Ljung-Box (1978) (*Portmanteau test*):

$$Q_{LB}(k) = T(T+2) \sum_{j=1}^k \frac{\hat{\rho}_j^2(\hat{\epsilon})}{T-j} \quad (3.20)$$

dove  $\hat{\rho}_j(\hat{\epsilon})$  rappresenta la  $j$ -esima autocorrelazione stimata dai residui del modello lineare ed è data da

$$\hat{\rho}_j(\hat{\epsilon}) = \frac{\sum_{t=j+1}^T \hat{\epsilon}_t \hat{\epsilon}_{t-j}}{\sum_{t=1}^T \hat{\epsilon}_t^2}, \quad j = 0, 1, 2, \dots \quad (3.21)$$

ADF si vedano Leybourne, Newbold e Vougas (1998) e Mills (1999).

<sup>14</sup>Per un totale di 485 osservazioni. Si noti che, anche nel sottocampione considerato, l'ipotesi nulla di radice unitaria non può essere rifiutata. Infatti, a fronte di valori critici tabulati di -3.4462 ( $\alpha = 0.01$ ) e -2.8678 ( $\alpha = 0.05$ ), la statistica ADF è pari a -2.4072.

<sup>15</sup>Si veda a tal proposito Granger e Teräsvirta (1993) o Teräsvirta (1994).



mentre  $T$  indica il numero di osservazioni. Sotto l'ipotesi nulla di assenza di autocorrelazione congiunta dal ritardo 1 al ritardo  $k$ , la statistica  $Q_{LB}(k)$  ha una distribuzione asintotica di tipo  $\chi^2$  con  $(k-r)$  gradi di libertà, dove  $r$  è il numero totale dei parametri stimati.<sup>16</sup>

Qualora i modelli lineari stimati candidati ad essere selezionati siano più di uno, si è soliti scegliere tra di essi impiegando delle tecniche di identificazione automatica (Piccolo, 1990). Tra tutte le varietà presenti in letteratura, è frequente utilizzare il criterio di informazione di Akaike (1974):

$$AIC = \frac{-2\ell}{T} + \frac{2r}{T} \quad (3.22)$$

oppure il criterio di Schwarz (1978):

$$SC = \frac{-2\ell}{T} + \frac{r \log T}{T} \quad (3.23)$$

dove, al solito,  $r$  è il numero dei parametri stimati,  $T$  è il numero di osservazioni, mentre  $\ell$  è il valore della funzione di log-verosimiglianza calcolato utilizzando gli  $r$  parametri stimati. Il valore di  $p$  che minimizza l'espressione di AIC o SC è scelto come ritardo appropriato per il modello autoregressivo.

Con riferimento alla serie delle differenze prime del Tasso sui Fondi Federali U.S.A. ( $\Delta y_t$ ), l'applicazione della procedura descritta porta al modello stimato mostrato in Tabella 3.1, mentre la Tabella 3.2 riporta le principali diagnostiche calcolate sui residui di stima.

Tabella 3.1: Stima del modello AR(p) sulla serie del Tasso sui Fondi Federali U.S.A. Variabile dipendente:  $\Delta y_t$ .

Regressori	Stima	Std. Error
$\Delta y_{t-1}$	0.4186	0.0448
$\Delta y_{t-2}$	-0.1126	0.0453
$\Delta y_{t-4}$	-0.1281	0.0421
$\Delta y_{t-7}$	-0.1612	0.0445
$\Delta y_{t-8}$	0.1392	0.0483
$\Delta y_{t-9}$	0.1088	0.0455
$\Delta y_{t-11}$	-0.1366	0.0420
$\Delta y_{t-13}$	0.1276	0.0421
AIC	1.6528	
SC	1.7233	
Log-Lik	-382.0731	

Come si può notare, a fronte di una curtosi molto elevata e di una significativa asimmetria a sinistra della distribuzione empirica dei residui, il test Jarque-Bera porta a rifiutare abbondantemente l'ipotesi di normalità. Inoltre, se da un lato il modello

<sup>16</sup>Nel caso in cui il modello stimato sia un AR(p) più un termine costante,  $r = p + 1$ . Se la costante non c'è, allora  $r = p$ .

Tabella 3.2: Diagnostiche sui residui di stima del modello AR(p) sulla serie del Tasso sui Fondi Federali U.S.A.

$\hat{\sigma}_\epsilon$	0.5436	
Skewness	-2.6731	
Curtosi	40.78	
Test	Statistica-test	p-value
Jarque-Bera	28632.9	0.0000
$\hat{Q}_{LB}(15)$	6.648	0.4664
$\hat{Q}_{LB}(24)$	21.02	0.1776
$\hat{Q}_{LB}(30)$	34.236	0.0465
ARCH(1)	1.1127	0.2915
ARCH(4)	19.2638	0.0007

stimato non lascia presenza di autocorrelazione seriale nei residui, dall'altro non è in grado di modellare adeguatamente la persistenza presente nella varianza condizionata della serie, come evidenziato dal test ARCH. Tali inefficienze non possono essere rimosse semplicemente aumentando l'ordine autoregressivo del modello o inserendo altre variabili esplicative. Una più accurata analisi di specificazione rivolta al passaggio da un modello lineare ad uno non-lineare si rivela essenziale, tenuto conto anche del fatto che il test ARCH potrebbe essere sensibile ad una errata specificazione del modello per la media condizionata (Lundbergh e Teräsvirta, 1998). Il modello scelto è il risultato della stima di una sequenza di modelli AR(p) più un termine costante, con un ritardo massimo di  $p = 15$ . Tra tutti i modelli che non lasciavano presenza di autocorrelazione nei residui, è stato scelto quello con i valori più bassi dei criteri informativi AIC e SC. Da tale specificazione sono stati poi eliminati i coefficienti non significativi, giungendo al modello finale della Tabella 3.1.

### 3.4 Introduzione al test di linearità

Dal momento che i modelli non-lineari sono molto più complicati da stimare e, successivamente, valutare e impiegare a fini previsivi rispetto a quelli lineari, la loro effettiva convenienza deve essere opportunamente investigata. Un modo naturale per verificare tale evenienza è quello di costruire un test di linearità da applicare alla serie dei dati.

A volte, il test di linearità può anche essere impiegato per specificare un particolare modello non-lineare, ma questo può essere fatto solamente se l'insieme delle alternative possibili appartiene ad una famiglia di rappresentazioni sufficientemente ristretta.

Un'altra occasione nella quale il test di linearità trova applicazione è nella valutazione delle proprietà dei residui di un modello stimato. Se quest'ultimo è rappresentato da un modello lineare, il test di linearità avrà l'obiettivo di valutare la presenza di possibili nonlinearità residue, mentre se il modello stimato è di tipo non-lineare, il test sarà utilizzato per verificare l'adeguatezza del modello nel descrivere completamente

la dinamica della serie considerata. In tal caso, se la linearità dovesse essere ancora una volta rifiutata, il modello dovrebbe essere nuovamente specificato e stimato.

I test di linearità possono essere divisi in due categorie. Da una parte si collocano quelli derivati senza un'alternativa specifica in mente, dall'altra quelli contro una particolare alternativa. In quest'ultimo caso, il test è di solito formulato come test dei moltiplicatori di Lagrange (LM) poiché, tale metodo, non necessita della stima del modello sotto l'ipotesi alternativa (ovvero, il modello non-lineare).<sup>17</sup>

In ciò che segue ci occuperemo di una classe particolare di modelli non-lineari, i modelli a transizione liscia (STR) e, pertanto, nel sottoporre a verifica la linearità faremo riferimento al test dei moltiplicatori di Lagrange. In generale, si veda ad esempio Greene (1997), sotto l'ipotesi di normalità del termine di errore  $\epsilon_t$ , la statistica test LM assume la forma

$$LM = \frac{1}{T} \left[ \frac{\partial \ell(\hat{\theta}_0)}{\partial \theta} \right]' \mathbf{I}(\hat{\theta}_0)^{-1} \left[ \frac{\partial \ell(\hat{\theta}_0)}{\partial \theta} \right] \quad (3.24)$$

dove  $\ell(\cdot)$  indica la funzione di log-verosimiglianza,  $\mathbf{I}(\cdot)$  la matrice di informazione e  $\hat{\theta}_0$  il vettore dei parametri stimati sotto l'ipotesi nulla. La distribuzione asintotica della (3.24) è, sotto l'ipotesi nulla, una  $\chi^2$  con gradi di libertà pari al numero di restrizioni imposte.

Come è stato sollevato da molti autori (Davies, 1977 e 1987, Andrews e Ploberger, 1994 e Hansen, 1996), la statistica (3.24) non è direttamente applicabile in tutti quei casi in cui uno o più parametri non sono identificati sotto l'ipotesi di linearità. Come conseguenza si ha che i test asintotici non possiedono distribuzione standard, ma necessitano di essere derivati ogni volta sulla base dello specifico modello considerato sotto l'ipotesi alternativa, ricorrendo spesso anche ad appropriati metodi di simulazione. Interessanti riferimenti di test costruiti avendo in mente una specifica formulazione non-lineare sono Saikkonen e Luukkonen (1988), Luukkonen, Saikkonen e Teräsvirta (1988a e 1988b), Tong (1990) e Granger e Teräsvirta (1993), dove i modelli considerati sono quello bilineare, autoregressivo esponenziale, *threshold* ed autoregressivo a transizione liscia (STAR), che costituisce l'oggetto principale di questo lavoro.<sup>18</sup> Gli stessi autori prendono in considerazione anche l'eventualità che il termine di errore sia caratterizzato da una qualche forma di eteroschedasticità, mostrando le correzioni da apportare al test in simili circostanze. Per ulteriori dettagli si rimanda alle opere citate.

### 3.5 Test per STAR

Il passo successivo alla specificazione del modello lineare consiste nel testare l'ipotesi di linearità contro l'alternativa di tipo STAR:

$$y_t = \pi_0 + \pi' \mathbf{w}_t + (\theta_0 + \theta' \mathbf{w}_t) F(s_t; \gamma, c) + \epsilon_t, \quad t = 1, \dots, T \quad (3.25)$$

<sup>17</sup>Questo lo differenzia sostanzialmente da altre famiglie di test, quali il rapporto di verosimiglianza (LR) ed il test di Wald (W). In entrambi i casi occorre stimare il modello sotto l'ipotesi alternativa e ciò può risultare molto dispendioso (si veda Greene, 1997).

<sup>18</sup>Si vedano, tra gli altri, anche Teräsvirta (1996), Rothman (1996) e Lütkepohl, Teräsvirta e Wolters (1999).

dove  $\pi = (\pi_1, \dots, \pi_p)'$ ,  $\theta = (\theta_1, \dots, \theta_p)'$ ,  $\mathbf{w}_t = (y_{t-1}, \dots, y_{t-p})'$ ,  $\epsilon_t \sim iid N(0, \sigma_\epsilon^2)$ .

L'ipotesi nulla può essere espressa come uguaglianza a zero dei parametri  $\theta_0$  e  $\theta$ :

$$H_{0,1} : \theta_0 = 0, \theta = 0 \quad (3.26)$$

contro l'alternativa che almeno uno dei  $\theta_i \neq 0$ ,  $i = 0, 1, \dots, p$ . Così facendo i problemi si complicano per il fatto che ci sono dei parametri che non sono identificati sotto  $H_{0,1}$ . In particolare, pur non essendo presenti quando  $H_{0,1}$  è vera,  $\gamma$  e  $c$  possono assumere qualsiasi valore. Alternativamente, si potrebbe pensare di sottoporre a test l'ipotesi:

$$H_{0,2} : \gamma = 0 \quad (3.27)$$

ma, anche in tal caso, sotto  $H_{0,2}$ , i parametri  $\theta_0$ ,  $\theta$  e  $c$  restano non identificati.

La presenza di parametri non identificati<sup>19</sup> sotto l'ipotesi nulla non consente di ottenere distribuzioni asintotiche standard per le statistiche test; quindi la teoria tradizionale dei test d'ipotesi non può essere applicata (si veda Davies, 1977 e 1987 e, per recenti sviluppi, Andrews e Ploberger, 1994 e Hansen, 1996). Le statistiche test tendono ad avere distribuzioni non-standard per le quali spesso non è disponibile un'espressione analitica. I loro valori critici devono quindi essere determinati attraverso appropriati metodi di simulazione.

Nel caso specifico del test di linearità contro l'alternativa di tipo STAR, il problema è stato risolto con successo da Luukkonen *et al.* (1988b). La loro proposta consiste nel rimpiazzare la funzione di transizione  $F(s_t; \gamma, c)$  mediante una sua adeguata approssimazione di Taylor. Nel modello così riparametrizzato, il problema della non identificazione di alcuni parametri viene eliminato, per cui la linearità può essere verificata attraverso un test dei moltiplicatori di Lagrange<sup>20</sup> ( $LM$ ), il cui vantaggio è quello di possedere distribuzione asintotica standard  $\chi^2$  sotto l'ipotesi nulla.

La derivazione del test di linearità necessita della distinzione tra diverse situazioni, ciascuna delle quali dipende dalla natura della variabile di transizione  $s_t$ . Comunemente, sono considerati i seguenti casi:

1.  $s_t$  variabile endogena ritardata  $y_{t-d}$ , con  $1 \leq d \leq p$ .
2.  $s_t$  variabile endogena ritardata  $y_{t-d}$ , con  $d > p$ , o variabile esogena  $z_{kt}$ .
3.  $s_t$  combinazione lineare di  $y_{t-1}, \dots, y_{t-p}$ , ovvero  $\alpha' \mathbf{w}_t$  con  $\alpha$  incognito.
4.  $s_t$  trend temporale  $t$ .

Rimandando l'analisi del quarto punto al capitolo successivo, qui di seguito ci soffermeremo dettagliatamente sul primo caso, mettendo in evidenza le complicazioni che insorgono quando ci riferiamo ai casi due e tre. Analizzeremo infine il caso dei modelli STAR-D.

<sup>19</sup>Nella teoria statistica dei test si parla di parametri di disturbo (*nuisance*).

<sup>20</sup>Sulle condizioni che dovrebbero essere verificate nella derivazione del test  $LM$ , si veda Tsay (1986), Saikkonen e Luukkonen (1988) e Granger e Teräsvirta (1993).

### 3.5.1 Test contro l'alternativa LSTAR

Si consideri il modello LSTAR espresso sinteticamente dalla (3.25) con  $s_t = y_{t-d}$  e  $1 \leq d \leq p$  noto. Luukkonen, Saikkonen e Teräsvirta (1988b) suggeriscono di sostituire la funzione  $F(y_{t-d}; \gamma, c)$  mediante la sua approssimazione di Taylor intorno a  $\gamma = 0$ , troncata al primo ordine (si veda l'Appendice B per i dettagli), pervenendo così al modello ausiliario

$$y_t = \beta_{0,0} + \beta'_0 \mathbf{w}_t + \beta'_1 \mathbf{w}_t y_{t-d} + e_t \quad (3.28)$$

dove  $e_t = \epsilon_t + \theta_0 R_1(\cdot) + \theta'_1 \mathbf{w}_t R_1(\cdot)$ . Si noti che, sotto l'ipotesi nulla  $R_1(\cdot) = 0$ , e quindi  $e_t = \epsilon_t$ . Di conseguenza, le proprietà dei residui e la distribuzione delle statistiche test non sono alterate dal resto  $R_1(\cdot)$  (Granger e Teräsvirta, 1993 e van Dijk, 1999). La linearità può dunque essere verificata sottoponendo a test l'ipotesi:

$$H_{0,3} : \beta_1 = 0 \quad (3.29)$$

In pratica, la procedura di test si basa sui seguenti passi:

1. Stimare il modello sotto l'ipotesi di linearità regredendo  $y_t$  su  $\mathbf{w}_t$  più una costante. Calcolare i residui  $\hat{\epsilon}_t$  e la relativa somma dei quadrati  $SSR_0 = \sum_{t=1}^T \hat{\epsilon}_t^2$ .
2. Stimare il modello ausiliario regredendo  $\hat{\epsilon}_t$  su una costante,  $\mathbf{w}_t$  e  $\mathbf{w}_t y_{t-d}$ . Calcolare i residui  $\hat{e}_t$  e la somma dei quadrati dei residui  $SSR_1 = \sum_{t=1}^T \hat{e}_t^2$ .
3. Calcolare la versione  $F$  o la versione  $\chi^2$  della statistica test  $LM_1$ :

$$LM_1 = \frac{(SSR_0 - SSR_1)/p}{SSR_1/(T - 2p - 1)} \sim F_{p, T-2p-1} \quad (3.30)$$

$$LM_1 = T \frac{(SSR_0 - SSR_1)}{SSR_0} \sim \chi_p^2 \quad (3.31)$$

Granger e Teräsvirta (1993) raccomandano di utilizzare la versione  $F$  della statistica test poiché possiede proprietà migliori in piccoli campioni.<sup>21</sup>

In Appendice B.1 è mostrata la relazione<sup>22</sup> che c'è tra i parametri del modello ausiliario e quelli del modello LSTAR originario espresso dalla (3.25). Dall'osservazione di queste equazioni si vede come testare le ipotesi (3.26) o (3.27) sia equivalente a testare l'ipotesi (3.29) contro l'alternativa che almeno uno dei  $\beta_{1,i} \neq 0$ ,  $i = 1, \dots, p$ , in quanto i parametri del modello ausiliario sono funzioni di quelli del modello originario.

Come mostrato da Luukkonen *et al.* (1988b), la statistica test  $LM_1$  non ha potenza nei casi in cui i due regimi differiscano soltanto per il valore dell'intercetta, ovvero  $\pi_0 \neq \theta_0$  ma  $\theta_i = 0$ ,  $i = 1, \dots, p$ . Ciò comporta che, nella (3.28),  $\beta_{1,i} = 0$ ,  $i = 1, \dots, p$  portando all'accettazione dell'ipotesi nulla quando in realtà dovrebbe essere rifiutata.

<sup>21</sup>Ciò varrà anche per le altre statistiche test derivate nel corso di questa Sezione (Granger e Teräsvirta, 1993). Tuttavia, il valore operativo di tale raccomandazione viene meno qualora ci si trovi a dover lavorare con serie finanziarie ad alta frequenza, data l'enorme mole di dati a disposizione (Lundbergh e Teräsvirta, 1998).

<sup>22</sup>Equazioni (B.2) - (B.5).

Luukkonen *et al.* (1988b) suggeriscono di rimediare a questo inconveniente sostituendo la funzione di transizione con la sua approssimazione di Taylor troncata al terzo ordine e valutata nell'intorno del punto  $\gamma = 0$ .<sup>23</sup> Il modello ausiliario che si ottiene è il seguente:

$$y_t = \beta_{0,0} + \beta_0' \mathbf{w}_t + \beta_1' \mathbf{w}_t y_{t-d} + \beta_2' \mathbf{w}_t y_{t-d}^2 + \beta_3' \mathbf{w}_t y_{t-d}^3 + u_t \quad (3.32)$$

dove  $u_t = \epsilon_t + \theta_0 R_3(\cdot) + \theta' \mathbf{w}_t R_3(\cdot)$  e, sotto l'ipotesi nulla,  $u_t = \epsilon_t$  poiché  $R_3(\cdot) = 0$ . In analogia a quanto visto in precedenza, l'ipotesi nulla di linearità è data da:

$$H_{0,4} : \beta_1 = \beta_2 = \beta_3 = 0, \quad (3.33)$$

mentre il calcolo della statistica test si basa sui seguenti punti:

1. Stimare il modello lineare e calcolare  $SSR_0$  come in precedenza.
2. Regredire  $\hat{\epsilon}_t$  su una costante, su  $\mathbf{w}_t$ ,  $\mathbf{w}_t y_{t-d}$ ,  $\mathbf{w}_t y_{t-d}^2$  e  $\mathbf{w}_t y_{t-d}^3$ . Calcolare i residui  $\hat{u}_t$  e la relativa somma dei quadrati dei residui  $SSR_1 = \sum_{t=1}^T \hat{u}_t^2$ .
3. Calcolare le versioni  $F$  o  $\chi^2$  della statistica test  $LM_3$ :

$$LM_3 = \frac{(SSR_0 - SSR_1)/3p}{SSR_1/(T - 4p - 1)} \sim F_{3p, T-4p-1} \quad (3.34)$$

$$LM_3 = T \frac{(SSR_0 - SSR_1)}{SSR_0} \sim \chi_{3p}^2 \quad (3.35)$$

Il collegamento tra il modello originario e quello ausiliario è mostrato nell'Appendice B.2. L'analisi delle equazioni (B.7) - (B.14) conferma la validità dell'ipotesi nulla  $H_{0,4}$  formulata sul modello ausiliario; essa corrisponde infatti all'ipotesi  $H_{0,2}$  sul modello originario.

Il numero dei parametri della regressione ausiliaria (3.32) può essere ridotto (Luukkonen *et al.*, 1988b) semplicemente aggiungendo alla (3.28) il regressore  $y_{t-d}^3$  poiché il parametro da cui deriva,  $\beta_{2,d}$ , è l'unico tra quelli aggiuntivi  $\beta_2$  e  $\beta_3$  ad essere funzione di  $\theta_0$ :

$$y_t = \beta_{0,0} + \beta_0' \mathbf{w}_t + \beta_1' \mathbf{w}_t y_{t-d} + \beta_{2,d} y_{t-d}^3 + e_t \quad (3.36)$$

La statistica test risultante viene indicata col simbolo  $LM_3^e$  e, sotto l'ipotesi nulla  $H_{0,4}^e : \beta_1 = 0, \beta_{2,d} = 0$ , possiede distribuzione asintotica  $\chi^2$  con  $p + 1$  gradi di libertà.<sup>24</sup>

Nel caso in cui la variabile di transizione  $s_t$  sia una variabile endogena ritardata  $y_{t-d}$ , con  $d > p$  noto, o una variabile esogena, la procedura di test del terzo ordine di cui sopra porta al seguente modello ausiliario:

$$y_t = \beta_{0,0} + \beta_0' \mathbf{w}_t + \beta_{1,0} s_t + \beta_1' \mathbf{w}_t s_t + \beta_{2,0} s_t^2 + \beta_2' \mathbf{w}_t s_t^2 + \beta_{3,0} s_t^3 + \beta_3' \mathbf{w}_t s_t^3 + u_t \quad (3.37)$$

<sup>23</sup>Si veda l'Appendice B.2.

<sup>24</sup>La versione  $F$  della statistica test  $LM_3^e$  possiede  $p + 1$  e  $T - 2p - 2$  gradi di libertà. Il principale vantaggio dell'impiego di  $LM_3^e$  è dovuto al fatto che richiede una quantità considerevolmente inferiore di gradi di libertà rispetto ad  $LM_3$ , soprattutto nel caso in cui  $s_t$  è incognita o rappresenta una non specificata combinazione lineare di  $y_{t-1}, \dots, y_{t-p}$ . Nel presente contesto, dove  $s_t = y_{t-d}$ , questa argomentazione può non essere rilevante (van Dijk, 1999) e lo sarà tanto di meno quanto maggiore sarà l'ammontare di dati a disposizione.

L'ipotesi sottoposta a test è  $H'_{0,4} : \beta_1 = \beta_2 = \beta_3 = 0, \beta_{1,0} = \beta_{2,0} = \beta_{3,0} = 0$ , mentre la statistica test  $LM'_3$  ha distribuzione asintotica  $\chi^2$  con  $3p + 3$  gradi di libertà. Anche in tal caso è presente una versione più parsimoniosa della (3.37),

$$y_t = \beta_{0,0} + \beta'_0 \mathbf{w}_t + \beta_{1,0} s_t + \beta'_1 \mathbf{w}_t s_t + \beta_{2,0} s_t^2 + \beta_{3,0} s_t^3 + e_t \quad (3.38)$$

la cui statistica test  $LM_3^{e'}$  possiede, sotto l'ipotesi  $H_{0,4}^{e'} : \beta_1 = 0, \beta_{1,0} = \beta_{2,0} = \beta_{3,0} = 0$ , distribuzione asintotica  $\chi^2$  con  $p + 3$  gradi di libertà. Tuttavia, le motivazioni che portavano a preferire la procedura del terzo ordine a quella del primo, in tal caso vengono meno. Pertanto è sufficiente adottare il modello ausiliario

$$y_t = \beta_{0,0} + \beta'_0 \mathbf{w}_t + \beta_{1,0} s_t + \beta'_1 \mathbf{w}_t s_t + e_t \quad (3.39)$$

costruito a partire dall'approssimazione di Taylor della funzione logistica nell'intorno di  $\gamma = 0$ , troncata al primo ordine. L'ipotesi di linearità implica l'uguaglianza a zero dei parametri  $\beta_{1,0}$  e  $\beta_1$ , mentre la distribuzione asintotica della statistica test  $LM_3^{e'}$  è una  $\chi^2$  con  $p + 1$  gradi di libertà. Essa ha potenza anche nel caso in cui i due regimi differiscano per il solo valore dell'intercetta.<sup>25</sup> Quando  $s_t$  è espresso come combinazione lineare di  $y_{t-1}, \dots, y_{t-p}$ ,  $s_t = \alpha' \mathbf{w}_t$  con  $\alpha$  incognito, il numero dei regressori nel modello riparametrizzato aumenta considerevolmente. Ad esempio, Teräsvirta, Lin e Granger (1993) hanno derivato un test contro l'alternativa LSTAR, nel quale l'unica differenza tra i due regimi è dovuta soltanto ai termini costanti. Il modello ausiliario è stato ottenuto rimpiazzando la funzione logistica  $F(\alpha' \mathbf{w}_t; \gamma, c)$  mediante la sua approssimazione di Taylor al terzo ordine, valutata sempre nell'intorno del punto  $\gamma = 0$ :

$$y_t = \beta_{0,0} + \beta'_0 \mathbf{w}_t + \sum_{i=1}^p \sum_{j=i}^p \lambda_{ij} y_{t-i} y_{t-j} + \sum_{i=1}^p \sum_{j=i}^p \sum_{k=j}^p \delta_{ijk} y_{t-i} y_{t-j} y_{t-k} + u_t \quad (3.40)$$

Anche il numero dei gradi di libertà aumenta rapidamente; nel caso trattato da Teräsvirta *et al.* (1993) tale numero è pari a  $p(p+1)/2 + p(p+1)(p+2)/6$  per la versione  $\chi^2$  (ad esempio, per  $p = 5$  occorrono 50 gradi di libertà).

Queste considerazioni auspicano una riduzione del numero dei regressori utilizzati nel modello ausiliario. Luukkonen *et al.* (1988b) assumono  $\alpha' = (0, \dots, 0, 1, 0, \dots, 0)$ , dove l'uno appare nella  $d$ -esima posizione. Il modello ausiliario a cui pervengono, utilizzando la procedura di approssimazione del terzo ordine,

$$\begin{aligned} y_t = & \beta_{0,0} + \beta'_0 \mathbf{w}_t + \sum_{i=1}^p \sum_{j=i}^p \lambda_{ij} y_{t-i} y_{t-j} + \sum_{i=1}^p \sum_{j=i}^p \varphi_{ij} y_{t-i} y_{t-j}^2 + \\ & + \sum_{i=1}^p \sum_{j=i}^p \vartheta_{ij} y_{t-i} y_{t-j}^3 + u_t \end{aligned} \quad (3.41)$$

può essere utilizzato per verificare l'ipotesi di linearità contro l'alternativa di tipo LSTAR, in cui  $s_t = y_{t-d}$  con  $d$  incognito. Il numero dei gradi di libertà richiesti

<sup>25</sup>Questo lo si vede bene analizzando la struttura dei parametri del modello ausiliario in funzione di quelli del modello originario; in particolare, nel caso in cui  $\pi_0 \neq \theta_0$  e  $\theta = 0, \beta_1 = 0$ , ma  $\beta_{1,0} \neq 0$  poiché  $\beta_{1,0} = \frac{1}{4} \gamma \theta_0$ .

dalla statistica test  $\chi^2$  così costruita ammonta a  $p(p+1)/2 + 2p^2$  (ad esempio, per  $p = 5$  occorrono 65 gradi di libertà). In modo analogo, il caso della statistica test  $LM_1$  con  $d$  incognito, richiede  $p(p+1)/2$  gradi di libertà.

Soluzioni più ricche sono possibili ma implicano un aumento indesiderato dei parametri da stimare e dei gradi di libertà necessari; difficoltà queste che, se da un lato possono essere superate grazie all'ausilio dei moderni calcolatori, dall'altro, la scarsa quantità di dati disponibili per certi fenomeni economici non le rende di fatto operative: ci riferiamo essenzialmente alle applicazioni di tipo macroeconomico, nel cui ambito tali modelli sono nati e dove trovano largo consenso tra gli utilizzatori.

### 3.5.2 Test contro l'alternativa ESTAR

Si consideri l'espressione (3.25) in cui  $F(\cdot)$  è una funzione esponenziale del tipo della (3.5), con  $s_t = y_{t-d}$  e  $1 \leq d \leq p$  noto. Sostituendo la funzione esponenziale con la sua approssimazione di Taylor al primo ordine nell'intorno di  $\gamma = 0$ , si ricava il modello ausiliario

$$y_t = \beta_{0,0} + \beta'_0 \mathbf{w}_t + \beta'_1 \mathbf{w}_t y_{t-d} + \beta'_2 \mathbf{w}_t y_{t-d}^2 + u_t \quad (3.42)$$

L'ipotesi sottoposta a test è  $H_{0,5} : \beta_1 = \beta_2 = 0$  e può essere verificata costruendo la statistica test  $LM_2$  nel modo seguente:

1. Stimare il modello lineare e calcolare  $SSR_0 = \sum_{t=1}^T \hat{\epsilon}_t^2$ .
2. Regredire  $\hat{\epsilon}_t$  su una costante, su  $\mathbf{w}_t$ ,  $\mathbf{w}_t y_{t-d}$  e  $\mathbf{w}_t y_{t-d}^2$ . Calcolare i residui  $\hat{u}_t$  e  $SSR_1 = \sum_{t=1}^T \hat{u}_t^2$ .
3. Calcolare le versioni  $F$  o  $\chi^2$  della statistica test:

$$LM_2 = \frac{(SSR_0 - SSR_1)/2p}{SSR_1/(T - 3p - 1)} \sim F_{2p, T-3p-1} \quad (3.43)$$

$$LM_2 = T \frac{(SSR_0 - SSR_1)}{SSR_0} \sim \chi_{2p}^2 \quad (3.44)$$

Dalle espressioni di  $\beta_{0,0}$  e  $\beta_i$ ,  $i = 0, 1, 2$ , in funzione dei parametri del modello originario (equazioni (B.15) - (B.20) nell'Appendice B.3) si vede che la restrizione  $\gamma = 0$  equivale a testare l'ipotesi  $H_{0,5}$  nel modello riparametrizzato.

Se  $s_t$  rappresenta una variabile endogena ritardata  $y_{t-d}$ , con  $d > p$  noto, o una variabile esogena, il modello ausiliario assume la forma

$$y_t = \beta_{0,0} + \beta'_0 \mathbf{w}_t + \beta_{1,0} s_t + \beta'_1 \mathbf{w}_t s_t + \beta_{2,0} s_t^2 + \beta'_2 \mathbf{w}_t s_t^2 + u_t \quad (3.45)$$

dalla quale deriva la statistica test  $LM_2'$  che richiede, nella versione  $\chi^2$ ,  $2p+2$  gradi di libertà.

Se invece consideriamo  $s_t = \alpha' \mathbf{w}_t$ , con  $\alpha' = (0, \dots, 0, 1, 0, \dots, 0)$  e  $d$  incognito, Granger e Teräsvirta (1993) hanno derivato il modello

$$y_t = \beta_{0,0} + \beta'_0 \mathbf{w}_t + \sum_{i=1}^p \sum_{j=i}^p \lambda_{ij} y_{t-i} y_{t-j} + \sum_{i=1}^p \sum_{j=i}^p \varphi_{ij} y_{t-i} y_{t-j}^2 + u_t, \quad (3.46)$$



la cui statistica test  $LM_2''$  necessita di  $p(p+1)/2 + p^2$  gradi di libertà (sempre nella versione asintotica  $\chi^2$ ).

Escribano e Jordá (1999) mettono in evidenza come l'approssimazione di Taylor al primo ordine della funzione esponenziale non sia sufficiente a catturare alcune sue caratteristiche rilevanti; in particolare non è in grado di catturarne i due punti di flesso.<sup>26</sup> Pertanto, suggeriscono di rimediare a questo inconveniente estendendo lo sviluppo di Taylor fino a comprendere un termine del secondo ordine, pervenendo così al modello ausiliario

$$y_t = \beta_{0,0} + \beta_0' \mathbf{w}_t + \beta_1' \mathbf{w}_t y_{t-d} + \beta_2' \mathbf{w}_t y_{t-d}^2 + \beta_3' \mathbf{w}_t y_{t-d}^3 + \beta_4' \mathbf{w}_t y_{t-d}^4 + u_t \quad (3.47)$$

L'ipotesi nulla da sottoporre a test è adesso  $H_{0,6} : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$  e la risultante statistica test, indicata con  $LM_4$ , ha distribuzione asintotica  $\chi^2$  con  $4p$  gradi di libertà.

Anche in tal caso sono facilmente ricavabili generalizzazioni della (3.47) che considerano diverse specificazioni per la variabile di transizione  $s_t$ . Il punto che merita sottolineare non è tanto la diversa espressione formale che tali modelli possono assumere, ma piuttosto il *trade-off* tra le variabili aggiuntive nella regressione ausiliaria e l'aumento delle dimensioni dell'ipotesi nulla e dei gradi di libertà nella risultante statistica test. È in tale contesto che Escribano e Jordá (1999) sottolineano l'importanza dell'utilizzo di modelli ausiliari basati sulla procedura di approssimazione del primo ordine, aumentati dei regressori aggiuntivi necessari per tener conto del fatto che i due regimi possono differire per il solo valore dell'intercetta. La versione parsimoniosa della (3.47) è pertanto espressa dal modello

$$y_t = \beta_{0,0} + \beta_0' \mathbf{w}_t + \beta_1' \mathbf{w}_t y_{t-d} + \beta_{2,d} y_{t-d}^3 + \beta_{3,d} y_{t-d}^4 + e_t \quad (3.48)$$

la cui statistica test  $LM_4^e$ , nella versione  $\chi^2$ , necessita di  $p+2$  gradi di libertà. Simulazioni effettuate da Escribano e Jordá (1999) mostrano che la statistica test  $LM_3$  ottiene risultati migliori quando il vero processo generatore dei dati (DGP) è un modello della classe logistica; viceversa, se il vero DGP appartiene alla famiglia esponenziale, è la statistica  $LM_4$  a fornire risultati migliori. Ciò risulta facilmente comprensibile poiché i  $p$  regressori aggiuntivi utilizzati nella (3.47), se da un lato sono di estrema importanza per un modello di tipo ESTAR, dall'altro risultano ridondanti per un modello LSTAR, provocando una perdita di potenza del test.

### 3.5.3 Test contro le alternative LSTAR-D ed ESTAR-D

Si considerino i modelli LSTAR-D e ESTAR-D definiti nella Sezione 3.1.2:

$$y_t = \pi_0 + \pi' \mathbf{w}_t + (\theta_0 + \theta' \mathbf{w}_t) F(\tilde{\epsilon}_{t-d}; \gamma) + \epsilon_t \quad (3.49)$$

dove

$$F(\tilde{\epsilon}_{t-d}; \gamma) = \frac{1}{1 + \exp\{-\gamma(\tilde{\epsilon}_{t-d})\}}, \quad \gamma > 0, \quad d > 0 \quad (3.50)$$

<sup>26</sup>Come si evince dalla Figura 3.1 a pagina 32, questo è particolarmente importante quando  $\gamma$  è piccolo in valore assoluto, oppure quando la varianza dei residui ricavati dal modello STAR è grande a tal punto da far sì che una grossa quantità di osservazioni si concentri sul regime, o stato, superiore (Escribano e Jordá, 1999).

nel caso logistico, mentre

$$F(\tilde{\epsilon}_{t-d}; \gamma) = 1 - \exp\{-\gamma(\tilde{\epsilon}_{t-d}^2)\}, \quad \gamma > 0, \quad d > 0 \quad (3.51)$$

in quello esponenziale.

La procedura di approssimazione di Luukkonen *et al.* (1988b) con  $d$  noto applicata in questi casi porta ai modelli ausiliari

$$y_t = \beta_{0,0} + \beta_0' \mathbf{w}_t + \beta_{1,0} \tilde{\epsilon}_{t-d} + \beta_1' \mathbf{w}_t \tilde{\epsilon}_{t-d} + \beta_{3,0} \tilde{\epsilon}_{t-d}^3 + \beta_3' \mathbf{w}_t \tilde{\epsilon}_{t-d}^3 + u_t \quad (3.52)$$

per la versione LSTAR-D, e

$$y_t = \beta_{0,0} + \beta_0' \mathbf{w}_t + \beta_{2,0} \tilde{\epsilon}_{t-d}^2 + \beta_2' \mathbf{w}_t \tilde{\epsilon}_{t-d}^2 + u_t \quad (3.53)$$

per il caso ESTAR-D.

L'analisi dei parametri del modello (3.52) in funzione di quelli del modello originario mostra che i termini  $\beta_{3,0} \tilde{\epsilon}_{t-d}^3$  e  $\beta_3' \mathbf{w}_t \tilde{\epsilon}_{t-d}^3$  non sono necessari nella regressione ausiliaria poiché vengono meno le cause che, nel caso LSTAR, giustificano la procedura di approssimazione del terzo ordine in luogo di quella del primo (si veda l'Appendice B.4). Nella (B.22),  $\beta_{1,0} \neq 0$  anche quando i due regimi differiscono soltanto per il valore dell'intercetta. Il test è dunque sensibile a questa possibilità e può essere semplicemente basato sulla procedura del primo ordine:

$$y_t = \beta_{0,0} + \beta_0' \mathbf{w}_t + \beta_{1,0} \tilde{\epsilon}_{t-d} + \beta_1' \mathbf{w}_t \tilde{\epsilon}_{t-d} + e_t \quad (3.54)$$

L'ipotesi sottoposta a test è  $\beta_{1,0} = 0, \beta_1 = 0$  contro l'alternativa che almeno uno dei  $\beta_{1,i} \neq 0, i = 0, 1, \dots, p$ . In modo del tutto analogo si formula l'ipotesi nulla per il caso esponenziale, mentre entrambe le statistiche test sono asintoticamente distribuite come una  $\chi^2$  con  $p + 1$  gradi di libertà.

Nel caso in cui  $d$  sia incognito, la (3.53) e la (3.54) sono facilmente generalizzabili, rispettivamente come (Granger e Teräsvirta, 1993):

$$y_t = \beta_{0,0} + \beta_0' \mathbf{w}_t + \sum_{d=1}^D \varphi_d \tilde{\epsilon}_{t-d}^2 + \sum_{d=1}^D \delta_d \mathbf{w}_t \tilde{\epsilon}_{t-d}^2 + u_t \quad (3.55)$$

e

$$y_t = \beta_{0,0} + \beta_0' \mathbf{w}_t + \sum_{d=1}^D \varphi_d \tilde{\epsilon}_{t-d} + \sum_{d=1}^D \delta_d \mathbf{w}_t \tilde{\epsilon}_{t-d} + e_t \quad (3.56)$$

e i gradi di libertà delle statistiche test<sup>27</sup>  $\chi^2$  così derivate sono pari a  $D + pD$ .

Granger e Teräsvirta (1993) mettono in evidenza come le espressioni (3.55) e (3.56) rappresentino dei modelli bilineari e concludono che i modelli STAR-D e quelli bilineari sono localmente equivalenti.

<sup>27</sup>Indicheremo con  $LM_1 - D$  e  $LM_2 - D$  rispettivamente le statistiche test ricavate dai modelli ausiliari (3.54) e (3.53). Chiameremo inoltre  $LM_2 - D(A)$  (dove la "A" sta per *augmented*) la statistica test risultante dal modello ausiliario ottenuto inserendo nella (3.53) i regressori  $\beta_{1,0} \tilde{\epsilon}_{t-d}$  e  $\beta_1' \mathbf{w}_t \tilde{\epsilon}_{t-d}$ .

Tabella 3.3: Risultati delle simulazioni per il DGP1 con diversi valori del parametro incognito  $b_{11} = -0.8, -0.5, -0.3, -0.1, 0.1, 0.3, 0.5, 0.8$ . La tabella riporta le % di accettazione dell'ipotesi nulla di bilinearità.

$\alpha$			-0.8	-0.5	-0.3	-0.1	0.1	0.3	0.5	0.8
5%	$\chi^2$	$LM_1 - D$	83.9	70.1	94.0	94.7	92.6	94.4	69.4	85.1
		$LM_2 - D$	90.6	82.0	98.3	94.7	92.2	99.0	82.4	91.0
		$LM_2 - D(A)$	96.3	90.1	99.6	89.9	87.6	99.6	91.9	97.6
5%	$F$	$LM_1 - D$	83.8	69.9	94.0	94.6	92.5	94.4	69.1	85.0
		$LM_2 - D$	90.6	82.0	98.3	94.7	92.2	99.0	82.4	90.9
		$LM_2 - D(A)$	96.3	90.1	99.6	89.8	87.6	99.6	91.9	97.6
1%	$\chi^2$	$LM_1 - D$	75.6	56.4	89.9	83.0	81.3	90.2	56.1	79.1
		$LM_2 - D$	85.5	71.3	96.4	83.1	80.6	97.4	73.3	86.0
		$LM_2 - D(A)$	94.3	82.6	98.4	71.9	70.7	99.3	83.5	96.0
1%	$F$	$LM_1 - D$	75.8	56.5	89.9	83.0	81.4	90.2	56.2	79.1
		$LM_2 - D$	85.5	71.6	96.6	83.2	80.6	97.5	73.3	86.0
		$LM_2 - D(A)$	94.3	83.0	98.4	72.0	71.6	99.3	83.9	96.1

A conferma di ciò è stato effettuato un esercizio di simulazione con lo scopo di valutare la potenza del test contro le alternative LSTAR-D ed ESTAR-D<sup>28</sup> nel riconoscere la presenza di un processo generatore dei dati (DGP) di tipo bilineare. Sono stati considerati tre diversi processi, ciascuno dei quali con diversi valori dei parametri incogniti: in totale sono stati quindi generati 24 diversi modelli bilineari<sup>29</sup> dai seguenti DGP:

- DGP1:  $y_t = b_{11}y_{t-1}\epsilon_{t-1} + \epsilon_t$
- DGP2:  $y_t = b_{12}y_{t-1}\epsilon_{t-2} + \epsilon_t$
- DGP3:  $y_t = b_{21}y_{t-2}\epsilon_{t-1} + \epsilon_t$

Per ciascuno dei DGP considerati e per una data scelta del parametro incognito  $b_{ij}$  sono state generate 1000 serie storiche di ampiezza  $T = 200$ , attraverso l'estrazione di numeri casuali da una distribuzione Normale a media zero e varianza quattro.<sup>30</sup> A ciascuna serie storica così generata sono stati applicati i test contro le alternative LSTAR-D ed ESTAR-D, sia nella forma riportata nelle equazioni (3.53) e (3.54), sia nella forma *augmented*, le cui statistiche test sono indicate rispettivamente come  $LM_1 - D$ ,  $LM_2 - D$  ed  $LM_2 - D(A)$ . Le Tabelle 3.3, 3.4 e 3.5 riportano i risultati delle simulazioni per due diversi valori di  $\alpha$ , rispettivamente del 5% e dell'1%, sia nella versione  $\chi^2$  che in quella  $F$  di Fisher valida per piccoli campioni. I valori all'interno di ciascuna Tabella indicano la percentuale di corretta accettazione dell'ipotesi nulla di una

<sup>28</sup>Cfr. le equazioni (3.53) e (3.54).

<sup>29</sup>Per un'introduzione ai modelli bilineari si vedano, oltre alla Sezione 2.1 della presente monografia, Granger e Andersen (1978) e Granger e Teräsvirta (1993).

<sup>30</sup>In realtà, per ciascuna serie sono stati generati 300 valori, di cui i primi 100 sono stati scartati per evitare problemi legati ai valori iniziali del DGP, convenzionalmente fissati a zero.

Tabella 3.4: Risultati delle simulazioni per il DGP2 con diversi valori del parametro incognito  $b_{11} = -0.8, -0.5, -0.3, -0.1, 0.1, 0.3, 0.5, 0.8$ . La tabella riporta le % di accettazione dell'ipotesi nulla di bilinearità.

$\alpha$			-0.8	-0.5	-0.3	-0.1	0.1	0.3	0.5	0.8
5%	$\chi^2$	$LM_1 - D$	86.4	70.6	39.0	10.7	12.0	40.3	68.0	85.6
		$LM_2 - D$	92.4	80.3	49.0	11.5	11.6	48.1	78.7	90.9
		$LM_2 - D(A)$	97.8	95.0	99.9	63.7	63.0	99.9	95.6	98.1
5%	F	$LM_1 - D$	86.4	70.2	38.8	10.4	11.7	40.2	67.9	85.3
		$LM_2 - D$	92.3	79.9	48.9	11.5	11.5	47.8	78.6	90.9
		$LM_2 - D(A)$	97.8	95.0	99.9	63.7	63.0	99.9	95.6	98.1
1%	$\chi^2$	$LM_1 - D$	79.9	60.3	23.7	4.3	3.0	24.3	54.5	78.4
		$LM_2 - D$	88.3	69.9	30.4	3.6	2.8	31.0	68.7	85.4
		$LM_2 - D(A)$	95.8	91.6	99.6	40.2	36.9	99.8	92.7	95.8
1%	F	$LM_1 - D$	80.0	60.4	23.8	4.4	3.2	24.4	54.5	78.4
		$LM_2 - D$	88.3	69.9	30.5	3.7	2.8	31.0	68.7	85.4
		$LM_2 - D(A)$	95.8	91.8	99.6	41.0	38.0	99.8	93.1	96.0

forma funzionale del DGP di tipo bilineare<sup>31</sup>. Da una prima analisi dei risultati emerge chiaramente come *performance* relativamente migliori si ottengono con la versione *augmented* del test, arrivando in molti casi vicino al 100% di accettazioni dell'ipotesi nulla. Per contro, la versione ESTAR-D del test ( $LM_2 - D$ ) sembra essere quella con *performance* peggiore. Ulteriori simulazioni, anche eventualmente impiegando processi bilineari più complessi di quelli qui considerati si rendono comunque necessarie prima di trarre qualsiasi tipo di conclusione.

### 3.5.4 Selezione della variabile di transizione

I diversi test di linearità sono stati costruiti prevalentemente assumendo che la variabile di transizione  $s_t$  fosse nota, mettendo di volta in volta in evidenza le complicazioni che sorgevano quando  $s_t$  era incognita. Nelle applicazioni concrete, è proprio questo il caso in cui si trova a dover operare il ricercatore, il quale deve essere in grado di selezionare adeguatamente un'opportuna variabile di transizione.

Come evidenziato da Teräsvirta (1994), sebbene la statistica test  $LM_3$  sia stata derivata contro l'alternativa di tipo LSTAR, essa dovrebbe avere potenza anche contro l'alternativa ESTAR. La motivazione di una tale affermazione si trova nei modelli ausiliari (3.32) e (3.42), utilizzati per costruire rispettivamente le statistiche test  $LM_3$  e  $LM_2$  nel caso in cui  $s_t = y_{t-d}$ , con  $d$  noto. Come si vede, tutti i regressori ausiliari del modello (3.42) sono inclusi nel modello (3.32), giustificando così l'osservazione di Teräsvirta (1994).<sup>32</sup> Una appropriata variabile di transizione può quindi essere selezionata senza specificare la forma della funzione di transizione, semplicemente calcolando la statistica  $LM_3$  per diversi valori di  $d$ , scegliendo poi quella per la quale il p-value

<sup>31</sup>Tutte le versioni del test STAR-D sono state calcolate fissando  $c = 0$ ,  $p = 1$  e  $d = 1$ .

<sup>32</sup>In modo analogo, se  $s_t$  rappresenta una variabile esogena  $z_{kt}$  o endogena  $y_{t-d}$ , con  $d > p$ , i modelli posti a confronto sono espressi dalle equazioni (3.37) e (3.45), ma le conclusioni restano immutate.

Tabella 3.5: Risultati delle simulazioni per il DGP3 con diversi valori del parametro incognito  $b_{11} = -0.8, -0.5, -0.3, -0.1, 0.1, 0.3, 0.5, 0.8$ . La tabella riporta le % di accettazione dell'ipotesi nulla di bilinearità.

$\alpha$			-0.8	-0.5	-0.3	-0.1	0.1	0.3	0.5	0.8
5%	$\chi^2$	$LM_1 - D$	59.8	44.5	24.1	13.7	12.9	25.7	45.7	64.0
		$LM_2 - D$	5.2	16.1	21.2	13.7	13.4	26.4	18.2	7.3
		$LM_2 - D(A)$	77.8	92.6	99.4	59.0	56.5	99.6	93.5	78.5
5%	$F$	$LM_1 - D$	59.5	44.1	23.8	13.7	12.7	25.3	45.7	63.9
		$LM_2 - D$	5.2	16.1	20.9	13.6	13.3	26.3	18.1	7.3
		$LM_2 - D(A)$	77.8	92.6	99.4	58.9	56.4	99.6	93.4	78.5
1%	$\chi^2$	$LM_1 - D$	50.1	28.8	10.5	3.6	2.8	11.7	30.7	56.4
		$LM_2 - D$	3.8	8.0	9.2	2.9	3.2	12.3	10.3	4.9
		$LM_2 - D(A)$	71.2	87.1	97.2	33.0	33.7	98.0	87.3	71.2
1%	$F$	$LM_1 - D$	50.2	28.8	10.7	3.6	2.9	11.7	30.7	54.4
		$LM_2 - D$	3.8	8.1	9.3	2.9	3.2	12.3	10.3	4.9
		$LM_2 - D(A)$	71.4	87.2	97.3	33.4	34.8	98.1	87.4	71.5

della statistica test è più piccolo. Una tale regola di decisione trova fondamento nel fatto che il test dovrebbe avere potenza massima nel caso in cui l'ipotesi alternativa sia correttamente specificata, ovvero se una appropriata variabile di transizione è stata scelta. Evidenze empiriche ottenute per simulazione che confermano la validità di questo modo di operare, almeno per il caso univariato, si possono trovare in Teräsvirta (1994).

van Dijk, Teräsvirta e Franses (2000) suggeriscono di far precedere la selezione della variabile  $s_t$  da un test generale di linearità contro l'alternativa di tipo STAR, assumendo solamente che la corretta variabile di transizione sia una tra le possibili candidate  $s_{1t}, \dots, s_{mt}$ . Il calcolo della statistica test  $LM_3$  è subordinato alla condizione che  $s_t = \sum_{i=1}^m \alpha_i s_{it}$ , con  $\alpha_d = 1$  per certi  $d \in \{1, \dots, m\}$  e  $\alpha_i = 0$  per  $i \neq d$ . Questo corrisponde ad una generalizzazione del modello ausiliario (3.41) in cui  $s_t$  può essere sia una variabile esogena  $z_{kt}$ , sia una variabile endogena  $y_{t-d}$ , con  $d$  qualsiasi.

### 3.5.5 Scelta della funzione di transizione

Quando la linearità del modello è stata rifiutata in favore della non-linearità di tipo STAR e una corretta variabile di transizione è stata selezionata, il passo successivo riguarda la scelta della forma della funzione di transizione  $F(s_t; \gamma, c)$ . Tale scelta consiste nella discriminazione tra una funzione logistica del tipo della (3.4) e una esponenziale del tipo della (3.5).

Nel caso in cui  $s_t = y_{t-d}$  con  $1 \leq d \leq p$  noto, Teräsvirta (1994) suggerisce di utilizzare una procedura basata su una sequenza di test  $F$ , ciascuno ricavato dal modello ausiliario (3.32):

$$H_{0,1} : \beta_3 = 0 \quad \text{contro} \quad H_{1,1} : \beta_3 \neq 0 \quad (3.57)$$

$$H_{0,2} : \beta_2 = 0 \mid \beta_3 = 0 \quad \text{contro} \quad H_{1,2} : \beta_2 \neq 0 \mid \beta_3 = 0 \quad (3.58)$$

$$H_{0,3} : \beta_1 = 0 \mid \beta_2 = \beta_3 = 0 \quad \text{contro} \quad H_{1,3} : \beta_1 \neq 0 \mid \beta_2 = \beta_3 = 0 \quad (3.59)$$

Dall'analisi delle ipotesi (3.57) - (3.59) e, tenuto conto delle espressioni dei parametri  $\beta_i, i = 1, 2, 3$ , in funzione dei parametri del modello originario, si ricavano le seguenti interpretazioni:

- Accettare  $H_{0,1}$  e rifiutare  $H_{0,2}$  implica la scelta di un modello della classe ESTAR.
- Accettare sia  $H_{0,1}$  che  $H_{0,2}$  e rifiutare  $H_{0,3}$  comporta che un modello della famiglia LSTAR è preferibile a uno della famiglia ESTAR.

Granger e Teräsvirta (1993) e Teräsvirta (1994) hanno sollevato delle critiche alle regole di scelta di cui sopra. Le loro argomentazioni si fondano sul fatto che il modello (3.32), utilizzato per verificare le ipotesi (3.57) - (3.59), è stato costruito facendo riferimento ad una procedura di approssimazione del terzo ordine. Approssimazioni di ordine più elevato potrebbero evidenziare che una funzione di transizione scelta sulla base delle suddette regole si riveli in realtà sbagliata. Ad esempio, se consideriamo un'approssimazione del secondo ordine del modello ESTAR,  $\beta_2$  sarà uguale a zero solo se  $c = \theta_0 = 0$ . Supponendo che il vero modello sia un ESTAR, almeno uno dei parametri  $c$  o  $\theta_0$  deve essere diverso da zero. Quindi  $H_{0,1}$  può essere rifiutata solo se il livello di significatività del test è sufficientemente elevato e il range temporale della variabile oggetto di analisi è abbastanza ampio. Nella maggior parte dei casi invece, le condizioni descritte portano alla scelta di un modello LSTAR in luogo di un più appropriato modello ESTAR. Questo inconveniente potrebbe essere trascurato, o quantomeno si potrebbe ridurre significativamente la distorsione del test, nell'ambito dell'analisi delle serie finanziarie ad alta frequenza.

Per ovviare a questo tipo di problemi, Teräsvirta (1994) propone di concentrare la scelta nella seguente regola di decisione, basata sul confronto tra le forze relative delle regioni di rifiuto dei tre diversi test  $F$ . Pertanto, sulla base di tale regola, se il modello vero è un LSTAR, allora  $H_{0,1}$  e  $H_{0,3}$  sono tipicamente rifiutate con una frequenza molto più marcata rispetto a  $H_{0,2}$ . Viceversa, vale l'opposto se il vero modello è un ESTAR. Il confronto è dunque posto in termini di p-value; se il p-value del test  $H_{0,2}$  è il più piccolo dei tre, allora la scelta dovrebbe ricadere su un modello esponenziale, altrimenti, se così non è, il modello scelto dovrebbe appartenere alla famiglia logistica.

Risultati ottenuti mediante simulazione, per diversi valori di  $c$  e  $\pi_0$  ad un livello di significatività del 95% confermano la validità di questa regola (Teräsvirta, 1994). I valori ottenuti, se da un lato migliorano all'aumentare dell'ampiezza campionaria, dall'altro peggiorano nei casi in cui i due modelli possono essere considerati l'uno una buona approssimazione dell'altro.

La procedura di cui sopra si semplifica se  $s_t$  rappresenta una variabile endogena ritardata  $y_{t-d}$  con  $d > p$ , o una variabile esogena  $z_{kt}$ . Dal confronto dei modelli ausiliari (3.39) e (3.45) si nota che, nel caso logistico, non sono presenti i termini  $\beta_{2,0}s_t^2$  e  $\beta_2'w_t s_t^2$ . Questo suggerisce (Granger e Teräsvirta, 1993) di stimare il modello (3.45) e sottoporre a test  $F$  le ipotesi

$$H'_{0,1} : \beta_2 = 0, \beta_{2,0} = 0 \quad (3.60)$$

$$H'_{0,2} : \beta_1 = 0, \beta_{1,0} = 0 \mid \beta_2 = 0, \beta_{2,0} = 0 \quad (3.61)$$

Se il p-value del test  $H'_{0,1}$  è più piccolo di quello del test  $H'_{0,2}$ , si dovrebbe optare per un modello della famiglia ESTAR; viceversa nel caso opposto.

Una simile regola di decisione, ricavabile dalle equazioni (3.53) e (3.54), può essere applicata per scegliere tra un modello della famiglia ESTAR-D e uno della famiglia LSTAR-D.

Escribano e Jordá (1999) presentano una procedura alternativa per la scelta della funzione di transizione. La loro proposta si basa sulla stima del modello ausiliario espresso dalla (3.47) e consiste nel sottoporre a test- $F$  le ipotesi

$$H_{0,L} : \beta_1 = \beta_3 = 0 \quad (3.62)$$

$$H_{0,E} : \beta_2 = \beta_4 = 0 \quad (3.63)$$

e scegliendo un modello della classe LSTAR se  $H_{0,L}$  possiede p-value più piccolo di  $H_{0,E}$  e viceversa. Mediante simulazioni di tipo Monte Carlo, Escribano e Jordá (1999) mostrano che la loro procedura (EJP) ottiene risultati migliori rispetto alla procedura di Teräsvirta (1994) (TP) nella scelta della funzione di transizione, ed è anche più semplice da applicare.

Nelle Tabelle 3.6, 3.7 e 3.8 sono riportati i risultati di alcune applicazioni alle serie dei rendimenti giornalieri degli indici di borsa, dei tassi di cambio e delle azioni descritte nell'Appendice A. Ad ogni serie è stato applicato il test di linearità contro l'alternativa di tipo STAR. Per fare ciò si è selezionato come modello lineare quello autoregressivo di ordine  $p$ , dove il  $p$  è stato scelto in modo tale da eliminare completamente, o quantomeno ridurre al minimo, la presenza di correlazione seriale nei residui. Per ragioni di semplicità di calcolo, sono state omesse analisi riguardanti la presenza di eventuali effetti giornalieri, effetti *week-end* ed effetti "vacanza" o festività. Dal momento che la linearità è quasi sempre rifiutata con un p-value della statistica test corrispondente praticamente nullo, riteniamo che una simile omissione non pregiudichi l'interpretabilità dei risultati derivanti dai test.

I risultati ottenuti mostrano in tutti i casi la forte presenza di non-linearità nell'evoluzione delle serie considerate. I valori riportati si riferiscono alla statistica test  $LM_3$ , di cui si è ampiamente parlato nella Sezione 3.5.1; risultati analoghi sono stati ottenuti con tutte le altre statistiche test della Sezione 3.5. Tutto questo mostra che modelli della classe STAR potrebbero costituire un'adeguata parametrizzazione della media condizionata di molte serie finanziarie.

La scelta tra un modello logistico ed uno esponenziale è operata sulla base delle procedure appena descritte di Teräsvirta (1994) (TP) e di Escribano e Jordá (1999) (EJP). I risultati indicano differenze sostanziali tra le due procedure, che si accentuano maggiormente quando il modello selezionato è di tipo esponenziale. Questo giustificerebbe quanto sollevato da Escribano e Jordá (1999) come critica alla regola di Teräsvirta (1994): quando il "vero" modello è di tipo esponenziale, la procedura di Teräsvirta (1994) non è in grado di selezionare un'adeguata funzione di transizione. Le evidenze empiriche qui riportate sembrerebbero confermare questa affermazione.

### 3.5.6 Applicazione del test al tasso sui Fondi Federali U.S.A.

Con riferimento alla serie relativa al Tasso sui Fondi Federali U.S.A., in questa Sezione si riportano i risultati delle diverse varianti del test di linearità secondo quanto descritto

Tabella 3.6: Risultati relativi agli indici di borsa.

Indice	$p$	Min p-val( $\chi^2$ )		Tipo di Modello	
		$1 < d < 15$	$d$	TP	EJP
EOE	11	$2.76 \times 10^{-73}$	2	ESTAR	ESTAR
DAX	6	$8.36 \times 10^{-24}$	7	LSTAR	ESTAR
CAC40	11	$6.54 \times 10^{-30}$	2	LSTAR	ESTAR
FTSE100	11	$2.04 \times 10^{-72}$	1	LSTAR	ESTAR
HANG SENG	15	$2.87 \times 10^{-128}$	5	ESTAR	ESTAR
NIKKEI	10	$7.16 \times 10^{-14}$	1	LSTAR	LSTAR
ALL SHARES	13	$2.83 \times 10^{-30}$	1	LSTAR	ESTAR
S&P500	5	$7.02 \times 10^{-57}$	1	LSTAR	LSTAR

Tabella 3.7: Risultati relativi ai tassi di cambio.

Tasso di Cambio contro \$ U.S.A.	$p$	Min p-val( $\chi^2$ )		Tipo di Modello	
		$1 < d < 15$	$d$	TP	EJP
AUD/USD	3	$1.65 \times 10^{-15}$	7	LSTAR	LSTAR
GPB/USD	6	$9.86 \times 10^{-6}$	7	LSTAR	LSTAR
CAD/USD	1	$2.93 \times 10^{-6}$	10	LSTAR	ESTAR
DEM/USD	1	$3.96 \times 10^{-6}$	4	LSTAR	LSTAR
NLG/USD	6	$4.63 \times 10^{-5}$	2	LSTAR	ESTAR
FRF/USD	6	$9.21 \times 10^{-6}$	10	LSTAR	LSTAR
JPY/USD	10	$2.58 \times 10^{-11}$	1	LSTAR	LSTAR
CHF/USD	1	$6.34 \times 10^{-5}$	11	LSTAR	LSTAR

nelle pagine precedenti. Il modello lineare preso come punto di riferimento del test è quello stimato nella Sezione 3.3, ovvero un AR(13) dove, per semplicità di calcolo, sono stati inclusi tutti i ritardi dal primo al tredicesimo più un termine costante. I risultati ottenuti portano in tutti i casi al rifiuto abbondante dell'ipotesi nulla di linearità e indicano che un modello della classe STAR potrebbe costituire una valida alternativa alla rappresentazione lineare.

In ciascuna Tabella il primo valore indica il p-value della versione per piccoli campioni della statistica test ( $F$ ), mentre il valore tra parentesi indica la sua controparte asintotica ( $\chi^2$ ). La Tabella 3.9 riporta i p-value delle statistiche test di Luukkonen *et al.* (1988b), mentre la Tabella 3.10 mostra i risultati del test secondo quanto indicato da Escribano e Jordá (1999). I test sono calcolati per diversi valori della variabile di transizione  $s_t$  in modo da poter selezionare quella più adeguata mediante la regola del minimo p-value (si veda la Sezione 3.5.4). La scelta della variabile di transizione si rivela abbastanza controversa in quanto le possibili candidate sono più di una. In particolare, i dubbi più grossi riguardano la discriminazione tra  $\Delta y_{t-2}$  e  $\Delta y_{t-7}$  poiché, se da un lato la statistica  $LM_3$  privilegia  $\Delta y_{t-2}$ , dall'altro la statistica  $LM_4$  indica in  $\Delta y_{t-7}$  la scelta migliore.

Il test di linearità è stato effettuato anche contro l'alternativa di tipo STAR-D, i cui risultati sono riportati in Tabella 3.11 e confermano l'ampio rifiuto dell'ipotesi nulla di



Tabella 3.8: Risultati relativi alle azioni.

Titolo	$p$	Min p-val( $\chi^2$ )		Tipo di Modello	
		$1 \leq d \leq 15$	$d$	TP	EJP
AAPL	15	$6.66 \times 10^{-18}$	1	LSTAR	LSTAR
AMD	2	$6.38 \times 10^{-5}$	3	LSTAR	ESTAR
INTC	12	$3.63 \times 10^{-16}$	3	ESTAR	ESTAR
KO	2	$8.41 \times 10^{-7}$	11	ESTAR	LSTAR
MCD	3	$2.80 \times 10^{-6}$	1	LSTAR	ESTAR
MHP	10	$1.69 \times 10^{-23}$	3	ESTAR	ESTAR
PEP	5	$3.45 \times 10^{-12}$	1	ESTAR	ESTAR
T	3	$7.40 \times 10^{-18}$	5	LSTAR	LSTAR

linearità. Il confronto tra i risultati delle statistiche  $LM_3$  ed  $LM_2 - D(A)$ , entrambe potenti sia contro l'alternativa logistica che esponenziale, ci consente di poter scegliere tra un modello della classe STAR e uno della classe STAR-D. Sebbene la differenza tra le due statistiche test non sia particolarmente accentuata, la scelta dovrebbe privilegiare un modello STAR poiché i p-value del test sono leggermente più piccoli di quelli della controparte STAR-D.

Avendo optato per un modello STAR, l'ultimo stadio delle fasi di test si concentra sulla scelta della funzione di transizione tra le rappresentazioni logistica ed esponenziale.

Tabella 3.9: p-value  $F$  e ( $\chi^2$ ) del test di Luukkonen, Saikkonen e Teräsvirta (1988b) sulla serie del Tasso sui Fondi Federali U.S.A.

$s_t$	$LM_1$	$LM_2$	$LM_3$	$LM_3^e$
$\Delta y_{t-1}$	0.00407 (0.00356)	$8.25 \times 10^{-10}$ ( $3.75 \times 10^{-9}$ )	0.00000 ( $9.76 \times 10^{-15}$ )	0.00265 (0.00234)
$\Delta y_{t-2}$	$5.51 \times 10^{-14}$ ( $1.43 \times 10^{-20}$ )	$2.86 \times 10^{-14}$ ( $5.77 \times 10^{-34}$ )	$1.69 \times 10^{-49}$ ( $1.57 \times 10^{-33}$ )	0.00000 ( $2.03 \times 10^{-24}$ )
$\Delta y_{t-3}$	$6.32 \times 10^{-8}$ ( $9.92 \times 10^{-8}$ )	$1.84 \times 10^{-9}$ ( $7.43 \times 10^{-9}$ )	$1.45 \times 10^{-10}$ ( $1.72 \times 10^{-9}$ )	$3.62 \times 10^{-8}$ ( $6.18 \times 10^{-8}$ )
$\Delta y_{t-4}$	$5.28 \times 10^{-5}$ ( $5.29 \times 10^{-5}$ )	$2.82 \times 10^{-11}$ ( $2.21 \times 10^{-10}$ )	$5.28 \times 10^{-12}$ ( $1.28 \times 10^{-10}$ )	$2.72 \times 10^{-5}$ ( $2.86 \times 10^{-5}$ )
$\Delta y_{t-5}$	$1.32 \times 10^{-11}$ ( $5.10 \times 10^{-11}$ )	$2.46 \times 10^{-13}$ ( $4.17 \times 10^{-12}$ )	0.00000 ( $3.29 \times 10^{-22}$ )	$1.43 \times 10^{-11}$ ( $5.90 \times 10^{-11}$ )
$\Delta y_{t-6}$	$4.67 \times 10^{-12}$ ( $2.08 \times 10^{-11}$ )	$3.66 \times 10^{-14}$ ( $2.20 \times 10^{-19}$ )	0.00000 ( $3.30 \times 10^{-18}$ )	0.00000 ( $9.90 \times 10^{-14}$ )
$\Delta y_{t-7}$	0.00000 ( $1.17 \times 10^{-16}$ )	$3.34 \times 10^{-14}$ ( $9.59 \times 10^{-33}$ )	$1.46 \times 10^{-46}$ ( $4.24 \times 10^{-32}$ )	0.00000 ( $2.31 \times 10^{-16}$ )
$\Delta y_{t-8}$	$1.91 \times 10^{-7}$ ( $2.74 \times 10^{-7}$ )	$7.79 \times 10^{-10}$ ( $3.58 \times 10^{-9}$ )	$1.28 \times 10^{-12}$ ( $4.38 \times 10^{-11}$ )	$5.94 \times 10^{-8}$ ( $9.71 \times 10^{-8}$ )
$\Delta y_{t-9}$	$2.37 \times 10^{-9}$ ( $5.07 \times 10^{-9}$ )	$1.03 \times 10^{-9}$ ( $4.53 \times 10^{-9}$ )	$1.11 \times 10^{-11}$ ( $2.26 \times 10^{-10}$ )	$4.88 \times 10^{-9}$ ( $1.01 \times 10^{-8}$ )
$\Delta y_{t-10}$	0.00059 (0.00054)	$1.63 \times 10^{-5}$ ( $2.18 \times 10^{-5}$ )	$1.47 \times 10^{-5}$ ( $2.52 \times 10^{-5}$ )	0.00078 (0.00071)
$\Delta y_{t-11}$	$2.06 \times 10^{-6}$ ( $2.48 \times 10^{-6}$ )	$9.38 \times 10^{-10}$ ( $4.18 \times 10^{-9}$ )	$3.09 \times 10^{-14}$ ( $4.62 \times 10^{-15}$ )	$2.84 \times 10^{-6}$ ( $3.43 \times 10^{-6}$ )
$\Delta y_{t-12}$	$7.41 \times 10^{-10}$ ( $1.79 \times 10^{-9}$ )	$5.67 \times 10^{-12}$ ( $5.91 \times 10^{-11}$ )	$1.12 \times 10^{-12}$ ( $3.91 \times 10^{-11}$ )	$1.73 \times 10^{-9}$ ( $4.01 \times 10^{-9}$ )
$\Delta y_{t-13}$	0.00420 (0.00367)	$1.03 \times 10^{-5}$ ( $1.43 \times 10^{-5}$ )	$8.21 \times 10^{-12}$ ( $1.78 \times 10^{-10}$ )	0.00687 (0.00597)
$\Delta y_{t-14}$	$7.81 \times 10^{-5}$ ( $7.78 \times 10^{-5}$ )	$1.79 \times 10^{-5}$ ( $2.46 \times 10^{-5}$ )	$9.54 \times 10^{-7}$ ( $2.59 \times 10^{-6}$ )	0.00016 (0.00015)
$\Delta y_{t-15}$	0.01287 (0.01112)	0.00141 (0.00137)	$2.10 \times 10^{-6}$ ( $5.05 \times 10^{-6}$ )	0.00098 (0.00090)

Tabella 3.10: p-value  $F$  e ( $\chi^2$ ) del test di Escribano e Jordá (1999) sulla serie del Tasso sui Fondi Federali U.S.A.

$s_t$	$LM_4$	$LM_4^e$
$\Delta y_{t-1}$	0.00000 ( $2.35 \times 10^{-13}$ )	0.00426 (0.00374)
$\Delta y_{t-2}$	$3.54 \times 10^{-47}$ ( $1.20 \times 10^{-30}$ )	$6.82 \times 10^{-14}$ ( $5.89 \times 10^{-26}$ )
$\Delta y_{t-3}$	$2.85 \times 10^{-13}$ ( $3.92 \times 10^{-11}$ )	$8.16 \times 10^{-8}$ ( $1.354 \times 10^{-7}$ )
$\Delta y_{t-4}$	0.00000 ( $1.14 \times 10^{-12}$ )	$3.57 \times 10^{-5}$ ( $3.76 \times 10^{-5}$ )
$\Delta y_{t-5}$	0.00000 ( $4.12 \times 10^{-25}$ )	$6.12 \times 10^{-13}$ ( $3.90 \times 10^{-12}$ )
$\Delta y_{t-6}$	0.00000 ( $1.07 \times 10^{-19}$ )	$4.62 \times 10^{-14}$ ( $6.03 \times 10^{-16}$ )
$\Delta y_{t-7}$	$8.16 \times 10^{-50}$ ( $7.99 \times 10^{-32}$ )	0.00000 ( $1.33 \times 10^{-19}$ )
$\Delta y_{t-8}$	0.00000 ( $8.61 \times 10^{-15}$ )	$2.76 \times 10^{-8}$ ( $5.03 \times 10^{-8}$ )
$\Delta y_{t-9}$	$5.35 \times 10^{-13}$ ( $1.57 \times 10^{-9}$ )	$9.42 \times 10^{-9}$ ( $1.91 \times 10^{-8}$ )
$\Delta y_{t-10}$	$6.84 \times 10^{-13}$ ( $1.70 \times 10^{-13}$ )	0.00095 (0.00087)
$\Delta y_{t-11}$	0.00000 ( $3.74 \times 10^{-14}$ )	$5.67 \times 10^{-6}$ ( $6.67 \times 10^{-6}$ )
$\Delta y_{t-12}$	$1.72 \times 10^{-11}$ ( $6.71 \times 10^{-10}$ )	$3.99 \times 10^{-10}$ ( $1.15 \times 10^{-9}$ )
$\Delta y_{t-13}$	$2.82 \times 10^{-13}$ ( $3.62 \times 10^{-11}$ )	0.00416 (0.00365)
$\Delta y_{t-14}$	$1.52 \times 10^{-6}$ ( $5.45 \times 10^{-6}$ )	0.00028 (0.00027)
$\Delta y_{t-15}$	$4.58 \times 10^{-8}$ ( $3.35 \times 10^{-7}$ )	0.00127 (0.00116)

Tabella 3.11: p-value  $F$  e ( $\chi^2$ ) del test contro l'alternativa STAR-D.

$s_t$	$\bar{LM}_1 - \bar{D}$	$\bar{LM}_2 - \bar{D}$	$\bar{LM}_2 - \bar{D}(A)$
$\Delta\tilde{\epsilon}_{t-1}$	$8.65 \times 10^{-12}$ ( $3.98 \times 10^{-11}$ )	$1.57 \times 10^{-12}$ ( $9.13 \times 10^{-12}$ )	$2.27 \times 10^{-14}$ ( $8.79 \times 10^{-17}$ )
$\Delta\tilde{\epsilon}_{t-2}$	<b>0.00000</b> ( $1.61 \times 10^{-21}$ )	<b>0.00000</b> ( $5.19 \times 10^{-28}$ )	$1.40 \times 10^{-14}$ ( $8.59 \times 10^{-27}$ )
$\Delta\tilde{\epsilon}_{t-3}$	$3.34 \times 10^{-9}$ ( $7.41 \times 10^{-9}$ )	$1.64 \times 10^{-6}$ ( $2.07 \times 10^{-6}$ )	$5.45 \times 10^{-14}$ ( $7.01 \times 10^{-13}$ )
$\Delta\tilde{\epsilon}_{t-4}$	$5.47 \times 10^{-5}$ ( $5.55 \times 10^{-5}$ )	$9.85 \times 10^{-6}$ ( $1.10 \times 10^{-5}$ )	$4.95 \times 10^{-9}$ ( $1.97 \times 10^{-8}$ )
$\Delta\tilde{\epsilon}_{t-5}$	$9.92 \times 10^{-12}$ ( $4.51 \times 10^{-11}$ )	$8.64 \times 10^{-11}$ ( $2.94 \times 10^{-10}$ )	<b>0.00000</b> ( $1.97 \times 10^{-26}$ )
$\Delta\tilde{\epsilon}_{t-6}$	<b>0.00000</b> ( $2.09 \times 10^{-23}$ )	$3.83 \times 10^{-15}$ ( $3.50 \times 10^{-29}$ )	<b>0.00000</b> ( $2.72 \times 10^{-30}$ )
$\Delta\tilde{\epsilon}_{t-7}$	$6.14 \times 10^{-15}$ ( $4.05 \times 10^{-20}$ )	<b>0.00000</b> ( $3.88 \times 10^{-32}$ )	<b>0.00000</b> ( $3.57 \times 10^{-29}$ )
$\Delta\tilde{\epsilon}_{t-8}$	<b>0.01742</b> (0.01497)	<b>0.00084</b> (0.00076)	$6.44 \times 10^{-5}$ ( $7.92 \times 10^{-5}$ )
$\Delta\tilde{\epsilon}_{t-9}$	$7.96 \times 10^{-7}$ ( $1.06 \times 10^{-6}$ )	$8.16 \times 10^{-8}$ ( $1.32 \times 10^{-7}$ )	$2.22 \times 10^{-9}$ ( $9.99 \times 10^{-9}$ )
$\Delta\tilde{\epsilon}_{t-10}$	$3.52 \times 10^{-6}$ ( $4.21 \times 10^{-6}$ )	$3.47 \times 10^{-8}$ ( $6.07 \times 10^{-8}$ )	$3.08 \times 10^{-8}$ ( $9.32 \times 10^{-8}$ )
$\Delta\tilde{\epsilon}_{t-11}$	$3.43 \times 10^{-15}$ ( $1.22 \times 10^{-13}$ )	$3.76 \times 10^{-12}$ ( $1.98 \times 10^{-11}$ )	$3.24 \times 10^{-14}$ ( $7.11 \times 10^{-15}$ )
$\Delta\tilde{\epsilon}_{t-12}$	$7.84 \times 10^{-9}$ ( $1.59 \times 10^{-8}$ )	$1.85 \times 10^{-6}$ ( $2.32 \times 10^{-6}$ )	$3.13 \times 10^{-7}$ ( $6.90 \times 10^{-7}$ )
$\Delta\tilde{\epsilon}_{t-13}$	$1.49 \times 10^{-6}$ ( $1.90 \times 10^{-6}$ )	<b>0.00072</b> (0.00066)	$8.90 \times 10^{-8}$ ( $2.32 \times 10^{-7}$ )
$\Delta\tilde{\epsilon}_{t-14}$	<b>0.00134</b> (0.00120)	<b>0.03011</b> (0.02587)	<b>0.00722</b> (0.00646)
$\Delta\tilde{\epsilon}_{t-15}$	<b>0.00033</b> (0.00031)	$6.12 \times 10^{-5}$ ( $6.18 \times 10^{-5}$ )	$2.60 \times 10^{-5}$ ( $3.48 \times 10^{-5}$ )

Nella Tabella 3.12 sono riportati i risultati del test, sia nella versione di Teräsvirta (TP) che in quella di Escribano-Jordá (EJP), per cinque diverse scelte della variabile di transizione. Nei casi che più ci interessano, quelli con  $s_t = \Delta y_{t-2}$  e  $s_t = \Delta y_{t-7}$ , sia la regola di Teräsvirta (1994) che quella di Escribano e Jordá (1999) (si veda la Sezione 3.5.5) indicano in un modello esponenziale la scelta migliore per  $s_t = \Delta y_{t-2}$ , mentre per  $s_t = \Delta y_{t-7}$  la regola di Teräsvirta risulta inconclusiva e quella di Escribano-Jordá seleziona una funzione logistica. Quanto fin qui esposto si basa sulle versioni per piccoli campioni dei diversi test. Auspicando un'accettabile potenza delle rispettive versioni asintotiche con le circa 470 osservazioni a disposizione, la regola di Teräsvirta (1994) porta a scegliere nettamente un modello esponenziale in entrambi i casi, mentre quella di Escribano e Jordá (1999) conferma questa scelta per  $s_t = \Delta y_{t-2}$  e privilegia

Tabella 3.12: p-value  $F$  e ( $\chi^2$ ) del test per la scelta tra LSTAR ed ESTAR.

$s_t$	Teräsvirta (TP)			Escribano-Jordá (EJP)	
	$H_{0,1}$	$H_{0,2}$	$H_{0,3}$	$H_{0,L}$	$H_{0,E}$
$\Delta y_{t-2}$	$3.8 \times 10^{-6}$ ( $1.16 \times 10^{-6}$ )	$2.1 \times 10^{-14}$ ( $1.94 \times 10^{-22}$ )	$0.00000$ ( $4.59 \times 10^{-20}$ )	$3.7 \times 10^{-5}$ ( $1.24 \times 10^{-5}$ )	$1.9 \times 10^{-5}$ ( $6.62 \times 10^{-6}$ )
$\Delta y_{t-5}$	$4.6 \times 10^{-14}$ ( $9.1 \times 10^{-17}$ )	$0.00068$ ( $0.00041$ )	$3.5 \times 10^{-11}$ ( $1.3 \times 10^{-10}$ )	$1.4 \times 10^{-14}$ ( $4.9 \times 10^{-19}$ )	$1.4 \times 10^{-14}$ ( $2.8 \times 10^{-16}$ )
$\Delta y_{t-6}$	$0.02194$ ( $0.01113$ )	$1.3 \times 10^{-12}$ ( $2.5 \times 10^{-12}$ )	$1.3 \times 10^{-11}$ ( $5.3 \times 10^{-11}$ )	$0.00000$ ( $1.1 \times 10^{-15}$ )	$3.2 \times 10^{-7}$ ( $1.3 \times 10^{-7}$ )
$\Delta y_{t-7}$	$1.6 \times 10^{-5}$ ( $5.6 \times 10^{-6}$ )	$0.00000$ ( $1.9 \times 10^{-24}$ )	$0.00000$ ( $3.4 \times 10^{-16}$ )	$7.8 \times 10^{-13}$ ( $1.1 \times 10^{-12}$ )	$9.3 \times 10^{-8}$ ( $3.9 \times 10^{-8}$ )
$\Delta y_{t-11}$	$7.4 \times 10^{-10}$ ( $3.2 \times 10^{-10}$ )	$6.4 \times 10^{-5}$ ( $3.9 \times 10^{-5}$ )	$4.3 \times 10^{-6}$ ( $5.1 \times 10^{-6}$ )	$4.6 \times 10^{-5}$ ( $1.6 \times 10^{-5}$ )	$1.3 \times 10^{-8}$ ( $6.3 \times 10^{-9}$ )

una funzione logistica nel caso  $s_t = \Delta y_{t-7}$ .

Le considerazioni che precedono non ci consentono di determinare univocamente il modello STAR più appropriato per la serie in esame. La strategia operativa da seguire in simili circostanze si basa sulla stima di una prima formulazione del modello e in sede di valutazione successiva si dovranno apportare i necessari aggiustamenti fino ad arrivare ad una rappresentazione finale. Il modello da cui potremmo partire in ciò che segue si potrebbe individuare in un ESTAR con variabili di transizione  $\Delta y_{t-2}$  o  $\Delta y_{t-7}$ , anche se formulazioni alternative non possono essere escluse a priori.

### 3.6 Stima dei parametri

Il momento della stima è uno dei più importanti e delicati di tutta la procedura empirica di specificazione. Le successive fasi di valutazione e utilizzo, seppur concettualmente distinte, sono strettamente connesse alla stima dei parametri. Disporre di una metodologia “efficiente” consente di ridurre al minimo i rischi derivanti da errate valutazioni inferenziali tratte dal modello che, come noto, rappresentano affermazioni probabilistiche. Scopo di questa Sezione è mostrare come vengono stimati i parametri dei modelli a transizione liscia, e come i risultati ottenuti possano poi essere correttamente impiegati per gli scopi preposti.

A tal fine, si supponga di aver osservato il fenomeno  $y_t$ ,  $t = -p+1, \dots, -1, 0, 1, \dots, T$ , rappresentabile tramite un modello della classe STAR. Sotto certe condizioni di regolarità, descritte fra gli altri in Gallant (1987), la teoria dei modelli dinamici mostra che lo stimatore dei minimi quadrati non-lineari (NLS) è consistente e asintoticamente normale e può pertanto essere impiegato per stimare i parametri di un modello STAR:

$$\hat{\delta}_{NLS} = \min_{\delta} Q_T(\delta) = \min_{\delta} \sum_{t=1}^T (y_t - f(\mathbf{w}_t; \delta))^2 \quad (3.64)$$

dove  $\delta = (\pi_0, \pi', \theta_0, \theta', \gamma, c)'$  rappresenta il vettore dei parametri incogniti e  $f(\mathbf{w}_t; \delta) = \pi_0 + \pi' \mathbf{w}_t + (\theta_0 + \theta' \mathbf{w}_t) F(s_t; \gamma, c)$ . Lo stimatore  $\hat{\delta}_{NLS}$  è condizionato alle prime  $p$

osservazioni campionarie<sup>33</sup> ma, poiché tali stime possiedono giustificazione teorica solamente in un contesto asintotico, in pratica, la perdita delle prime  $p$  informazioni costituisce un problema trascurabile, specialmente se il campione considerato è abbastanza ampio. Quindi, asintoticamente vale che:

$$\sqrt{T}(\hat{\delta}_{NLS} - \delta_0) \longrightarrow N(\mathbf{0}, \text{AsyVar}(\hat{\delta}_{NLS})) \quad (3.65)$$

dove  $\delta_0$  è il vero valore dei parametri, mentre  $\text{AsyVar}(\hat{\delta}_{NLS})$  indica la matrice di varianze e covarianze asintotica di  $\hat{\delta}_{NLS}$ , la cui stima consistente può essere calcolata come segue (White, 1982):

$$\frac{1}{T} \left[ \left( -\frac{\partial^2 Q_T(\delta)}{\partial \delta \partial \delta'} \right)_{\delta=\hat{\delta}_{NLS}}^{-1} \begin{pmatrix} \frac{\partial Q_T(\delta)}{\partial \delta} & \frac{\partial Q_T(\delta)}{\partial \delta'} \end{pmatrix}_{\delta=\hat{\delta}_{NLS}} \left( -\frac{\partial^2 Q_T(\delta)}{\partial \delta \partial \delta'} \right)_{\delta=\hat{\delta}_{NLS}}^{-1} \right] \quad (3.66)$$

Il calcolo della matrice di varianze e covarianze asintotica di  $\hat{\delta}_{NLS}$  secondo quanto indicato dalla (3.66) ci permette di interpretare le stime ottenute con i minimi quadrati non-lineari come stime di quasi-massima verosimiglianza (White, 1982 ed anche White e Domowitz, 1984). Nell'introdurre l'approccio di quasi-massima verosimiglianza (QML), White (1982) mostra che lo stimatore di massima verosimiglianza (ML) continua a fornire risultati accettabili anche quando il modello non è correttamente specificato. Ciò che invece deve essere aggiustata è la stima della matrice di varianze e covarianze asintotica dei parametri, così come espresso dalla (3.66).

Se aggiungiamo l'ipotesi di normalità dei disturbi casuali, lo stimatore  $\hat{\delta}_{NLS}$  è equivalente allo stimatore di massima verosimiglianza condizionata:

$$\begin{aligned} & \max_{\delta, \sigma_\epsilon^2} \log L_{y_T, y_{T-1}, \dots, y_1 | y_0, y_{-1}, \dots, y_{-p+1}}(y_T, y_{T-1}, \dots, y_1 | y_0, y_{-1}, \dots, y_{-p+1}; \delta, \sigma_\epsilon^2) \\ & = -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log(\sigma_\epsilon^2) - \frac{1}{2\sigma_\epsilon^2} \sum_{t=1}^T (y_t - f(\mathbf{w}_t; \delta))^2 \end{aligned} \quad (3.67)$$

dove  $\log(\cdot)$  indica il logaritmo naturale dell'argomento. Derivando la (3.67) rispetto a  $\sigma_\epsilon^2$  e ponendo la derivata uguale a zero si ottiene:

$$\sigma_\epsilon^2 = \frac{1}{T} \sum_{t=1}^T (y_t - f(\mathbf{w}_t; \delta))^2 \quad (3.68)$$

Questo risultato intermedio mostra che la stima di  $\sigma_\epsilon^2$  può essere univocamente determinata a partire dai residui del modello STAR stimato<sup>34</sup> per cui, in pratica, la procedura di massimizzazione è basata sulla log-verosimiglianza concentrata, ottenuta sostituendo la (3.67) nella (3.68):

$$\max_{\delta} \log \ell_{y_T, y_{T-1}, \dots, y_1 | y_0, y_{-1}, \dots, y_{-p+1}}(y_T, y_{T-1}, \dots, y_1 | y_0, y_{-1}, \dots, y_{-p+1}; \delta)$$

<sup>33</sup>Da  $-p+1$  fino a 0.

<sup>34</sup>Un simile risultato si rivela particolarmente utile in una eventuale estensione del modello STAR a un contesto multivariato. La possibilità di ricavare una stima consistente della matrice di varianze-covarianze del vettore dei disturbi casuali a partire dal vettore dei residui, consente un'apprezzabile riduzione del numero di parametri da stimare.

$$= -\frac{T}{2} (\log(2\pi) - \log(T) + 1) - \frac{T}{2} \log \left( \sum_{t=1}^T (y_t - f(\mathbf{w}_t; \delta))^2 \right) \quad (3.69)$$

La soluzione simultanea delle  $2p + 4$  equazioni ricavate dall'uguaglianza a zero delle derivate parziali della (3.69) rispetto a  $\delta$ , ci fornisce le stime di massima verosimiglianza condizionata dei parametri del modello STAR, indicate con  $\hat{\delta}_{ML}$ . Tali stime, oltre ad essere equivalenti alle stime NLS<sup>35</sup> e a possedere le proprietà di consistenza e normalità asintotica, sono anche asintoticamente efficienti:

$$\sqrt{T}(\hat{\delta}_{ML} - \delta_0) \longrightarrow N \left( \mathbf{0}, \frac{1}{T} [\hat{\mathfrak{S}}(\hat{\delta}_{ML})]^{-1} \right). \quad (3.70)$$

$[\hat{\mathfrak{S}}(\hat{\delta}_{ML})]^{-1}$  indica l'inversa della matrice di informazione, la cui stima consistente può essere ricavata indifferentemente in due modi distinti: il primo basato sul calcolo delle derivate seconde della log-verosimiglianza concentrata,

$$[\hat{\mathfrak{S}}(\hat{\delta}_{ML})]^{-1} = [-\mathbf{H}(\hat{\delta}_{ML})]^{-1} = \left[ - \left( \frac{\partial^2 \log \ell(\cdot)}{\partial \delta \partial \delta'} \right)_{\delta=\hat{\delta}_{ML}} \right]^{-1} \quad (3.71)$$

e il secondo limitato al solo calcolo delle derivate prime:

$$[\hat{\mathfrak{S}}(\hat{\delta}_{ML})]^{-1} = [\mathbf{OP}(\hat{\delta}_{ML})]^{-1} = \left[ \left( \frac{\partial \log \ell(\cdot)}{\partial \delta} \frac{\partial \log \ell(\cdot)}{\partial \delta'} \right)_{\delta=\hat{\delta}_{ML}} \right]^{-1} \quad (3.72)$$

dove  $\mathbf{H}(\hat{\delta}_{ML})$  e  $\mathbf{OP}(\hat{\delta}_{ML})$  indicano rispettivamente l'hessiano e il prodotto esterno del gradiente della log-verosimiglianza concentrata, valutati in corrispondenza di  $\hat{\delta}_{ML}$ . Se il modello è stato correttamente specificato non c'è ragione di ritenere che vi sia una differenza apprezzabile tra i due modi di procedere.

Qualora l'ipotesi di normalità delle innovazioni non risultasse valida, il metodo della massima verosimiglianza continua a fornire stime consistenti dei parametri. Tuttavia, lo stimatore della matrice di varianze e covarianze di  $\hat{\delta}_{ML}$  deve essere modificato per tener conto del fatto che non è valida l'uguaglianza, in termini di valore atteso, tra le espressioni (3.71) e (3.72). White (1982) fornisce un'approssimazione della matrice di varianze e covarianze asintotica calcolata come:

$$\widehat{AsyVar}(\hat{\delta}_{ML}) \simeq \frac{1}{T} \left( [-\mathbf{H}(\hat{\delta}_{ML})]^{-1} (\mathbf{OP}(\hat{\delta}_{ML})) [-\mathbf{H}(\hat{\delta}_{ML})]^{-1} \right) \quad (3.73)$$

In pratica, prevalentemente in ambito monetario-finanziario, è a questo secondo approccio a cui ci dovremmo riferire. La constatazione, oramai generalmente accettata, della non-normalità della distribuzione empirica dei rendimenti implica un inevitabile aggiustamento dei risultati inferenziali tratti dal modello. Ed è proprio in questo senso che trova giustificazione l'approccio di quasi-massima verosimiglianza.

<sup>35</sup>A parte il termine costante, la (3.69) esprime  $-\frac{T}{2}$  volte il logaritmo naturale della somma dei quadrati dei residui. Così vediamo che massimizzare la funzione di log-verosimiglianza concentrata è equivalente a minimizzare  $Q_T(\delta)$ . Da qui l'uguaglianza tra le stime NLS e quelle di massima verosimiglianza (Davidson e MacKinnon, 1993).

### 3.6.1 Aspetti di calcolo nella fase di stima

Nella Sezione precedente si è visto che la stima dei parametri di un modello STAR può essere ricavata indifferentemente dall'applicazione dei minimi quadrati non-lineari (NLS) o del metodo della massima verosimiglianza (ML). In entrambi i casi non è disponibile un'espressione dello stimatore in forma chiusa e si deve quindi ricorrere a tecniche di ottimizzazione numerica.

Scopo di questa Sezione non è tanto quello di descrivere i vari algoritmi che in tali circostanze possono essere impiegati, per i quali sono tra l'altro disponibili in letteratura numerosi riferimenti bibliografici ai quali si rimanda,<sup>36</sup> ma vorremmo discutere alcuni punti che possono essere di aiuto nell'implementazione dell'algoritmo.

Supponendo di aver preliminarmente risolto il problema della scelta del metodo di ottimizzazione numerica da impiegare, fatto che ha perso oggi parte della sua importanza grazie all'ausilio di calcolatori sempre più potenti e di software applicativi sui quali sono implementati pressoché tutti gli algoritmi disponibili,<sup>37</sup> il punto che in ogni caso merita considerare in queste circostanze è la scelta dei valori iniziali. Ciascuno di questi metodi necessita di input predeterminati la cui individuazione può influire pesantemente sulla convergenza dell'algoritmo stesso. Una accurata e attenta scelta di questi valori costituisce uno dei momenti più importanti e delicati di tutta la fase di stima.

Con riferimento specifico ai modelli STAR, indipendentemente dal tipo di algoritmo scelto, il ricercatore deve preliminarmente determinare un valore iniziale per ciascuno dei parametri incogniti, valore dal quale prende avvio la procedura iterativa che, con un pò di fortuna, dovrebbe convergere più o meno rapidamente al massimo (o minimo) globale.

Un modo per ridurre la complessità dei calcoli deriva dall'osservazione che, considerando fissi i valori di  $\gamma$  e  $c$  nella funzione di transizione, il modello STAR diventa lineare nei parametri autoregressivi  $\pi_0$ ,  $\pi$ ,  $\theta_0$  e  $\theta$ . Così condizionatamente a  $\gamma$  e  $c$ , la stima di  $\phi = (\pi_0, \pi', \theta_0, \theta)'$  può essere ricavata mediante minimi quadrati ordinari (OLS):

$$\hat{\phi}(\gamma, c) = \left( \sum_{t=1}^T \mathbf{x}_t(\gamma, c) \mathbf{x}_t(\gamma, c)' \right)^{-1} \left( \sum_{t=1}^T \mathbf{x}_t(\gamma, c) y_t \right) \quad (3.74)$$

dove  $\mathbf{x}_t = (1, \mathbf{w}_t')$ ,  $\mathbf{x}_t(\gamma, c) = (\mathbf{x}_t', \mathbf{x}_t' F(s_t; \gamma, c))'$  e la notazione  $\phi(\gamma, c)$  è utilizzata per indicare la stima di  $\phi$  condizionata a  $\gamma$  e  $c$ . I corrispondenti residui possono essere calcolati come  $\hat{\epsilon}_t(\gamma, c) = y_t - \hat{\phi}(\gamma, c)' \mathbf{x}_t(\gamma, c)$  e la varianza residua è data da  $\hat{\sigma}_\epsilon^2(\gamma, c) = T^{-1} \sum_{t=1}^T \hat{\epsilon}_t^2(\gamma, c)$ . Un metodo conveniente per ottenere dei "buoni" valori iniziali per  $\gamma$  e  $c$  è quello di effettuare una ricerca a griglia e selezionare quei valori  $\hat{\gamma}$  e  $\hat{c}$  che minimizzano il valore di  $\hat{\sigma}_\epsilon^2(\gamma, c)$ .

<sup>36</sup>Si veda, ad esempio, Quandt (1983), Thisted (1988), Hamilton (1994) Sezione 5.7 e Hendry (1995) Appendice A.5.

<sup>37</sup>In molti di questi software è anche possibile cambiare algoritmo all'interno dello stesso problema di ottimizzazione, beneficiando così delle potenzialità tipiche di ogni algoritmo. Si pensi, ad esempio, alla possibilità di poter effettuare le prime iterazioni con il metodo BHHH (più efficiente quando si parte da lontano rispetto al punto di massimo o minimo globale) e le successive iterazioni col metodo di Newton-Raphson (migliore quando si parte da una regione di valori vicini al massimo o al minimo) (cfr. Fiorentini, Calzolari e Panattoni, 1996).



Nella procedura descritta non si è tenuto conto del fatto che la stima di  $\gamma$  è particolarmente sensibile all'unità di misura utilizzata. Per rimediare a questo inconveniente, Granger e Teräsvirta (1993) propongono la seguente standardizzazione della funzione di transizione:

$$F(s_t; \gamma, c) = \frac{1}{1 + \exp\{-\gamma (s_t - c) / \hat{\sigma}(s_t)\}}, \quad \gamma > 0 \quad (3.75)$$

per il caso logistico e

$$F(s_t; \gamma, c) = 1 - \exp\left\{-\gamma (s_t - c)^2 / \hat{\sigma}^2(s_t)\right\}, \quad \gamma > 0 \quad (3.76)$$

in quello esponenziale, dove  $\hat{\sigma}(s_t)$  e  $\hat{\sigma}^2(s_t)$  indicano rispettivamente la deviazione standard e la varianza della variabile di transizione. Questo tipo di standardizzazione è particolarmente utile quando  $\gamma$  è elevato; ciò rende anche la scelta del corrispondente valore iniziale molto più semplice. I due autori indicano in  $\gamma = 1$  un ragionevole punto di inizio, prevalentemente nel caso esponenziale.

Un altro modo per semplificare il problema di stima viene suggerito da Leybourne, Newbold e Vougas (1998) e consiste nel concentrare la funzione di perdita quadratica  $Q_T(\delta)$  rispetto a  $\phi$ :

$$Q_T(\gamma, c) = \sum_{t=1}^T (y_t - \phi(\gamma, c)' \mathbf{x}_t(\gamma, c))^2. \quad (3.77)$$

Questo riduce considerevolmente le dimensioni del problema di stima poiché la (3.77) necessita di essere minimizzata rispetto ai soli due parametri incogniti  $\gamma$  e  $c$ . L'importanza di questo metodo è particolarmente evidente quando si considerano estensioni del modello STAR, in particolare nel caso dei modelli a regimi multipli (MR-STAR) che saranno analizzati nel Capitolo successivo.

La Tabella 3.13 riporta le stime dei parametri del modello ESTAR precedentemente selezionato per il Tasso sui Fondi Federali U.S.A.<sup>38</sup>. In tale contesto, si è tenuto conto delle difficoltà menzionate e si è deciso di utilizzare la seguente strategia operativa<sup>39</sup>:

1. Standardizzare la funzione di transizione secondo quanto indicato dalla (3.76).
2. Fissato  $\gamma = 1$ , stimare gli altri parametri del modello ESTAR tramite massima verosimiglianza.
3. Stimare  $\gamma$  mediante massima verosimiglianza utilizzando i coefficienti ricavati al punto 2 come valori fissi per gli altri parametri del modello.
4. Stimare nuovamente i parametri di cui al punto 2 utilizzando il  $\hat{\gamma}$  ricavato al punto 3.

<sup>38</sup>La specificazione finale del modello ha beneficiato degli utilissimi suggerimenti forniti dal Prof. Timo Teräsvirta della Stockholm School of Economics, incontrato in occasione di un suo seminario all'Istituto Universitario Europeo presso la Badia Fiesolana, Firenze, che ringrazio per l'enorme pazienza e la disponibilità concessami.

<sup>39</sup>Le stime sono state ottenute mediante l'impiego della libreria CML (*Constrained Maximum Likelihood*) di GAUSS 3.2.35 per Windows.

Tabella 3.13: Stima di diversi modelli ESTAR sulla serie del Tasso sui Fondi Federali U.S.A. Variabile dipendente:  $\Delta y_t$ . Variabili di transizione:  $\Delta y_{t-7}$  e  $\Delta y_{t-2}$ .

Regressori	$\Delta y_{t-7}$		$\Delta y_{t-2}$	
	Stima	Std. Error	Stima	Std. Error
Primo Regime				
Costante	0.1031	0.0498	0.0788	0.0313
$y_{t-1}$	-0.0099	0.0069		
$\Delta y_{t-1}$	0.3508	0.0387	0.3172	0.0417
$\Delta y_{t-2}$	0.1835	0.0572		
$\Delta y_{t-4}$	-0.1331	0.0402	-0.1323	0.0384
$\Delta y_{t-5}$	0.4166	0.0785		
$\Delta y_{t-6}$	-0.2491	0.0560		
$\Delta y_{t-7}$			0.2623	0.0846
$\Delta y_{t-9}$	0.1469	0.0375	-0.1264	0.0760
$\Delta y_{t-12}$			0.1389	0.0725
$\Delta y_{t-13}$	0.2685	0.0775		
Secondo Regime				
Costante	-0.2056	0.0845		
$y_{t-1}$			-0.0360	0.0074
$\Delta y_{t-2}$	-0.6762	0.1050	-0.1324	0.0449
$\Delta y_{t-5}$	-0.8613	0.1015		
$\Delta y_{t-6}$	0.6659	0.1106		
$\Delta y_{t-7}$	-0.3996	0.0512	-0.6523	0.1139
$\Delta y_{t-8}$			0.1770	0.0559
$\Delta y_{t-9}$			0.3170	0.1023
$\Delta y_{t-11}$	-0.2106	0.0599		
$\Delta y_{t-12}$	-0.2209	0.0595	-0.4910	0.0987
$\Delta y_{t-13}$	-0.1906	0.1062	0.4190	0.0649
Funzione di Transizione				
$\gamma$	0.7904	0.1233	0.7717	0.1368
$c$	0.3418	0.0502	0.4435	0.0633
AIC	1.3729		1.4809	
SC	1.5402		1.6130	
Log-Lik	-305.0006		-334.5015	

Questo modo di procedere è giustificato dalla constatazione che la stima dei parametri di un modello ESTAR non è molto sensibile alla determinazione del valore di  $\gamma$  (Granger e Teräsvirta, 1993), per cui si può scindere il momento della stima in due fasi: quella relativa a  $\gamma$  e quella relativa agli altri parametri, condizionatamente a  $\gamma$ .

Il modello è stato stimato per due diverse scelte della variabile di transizione,  $s_t = \Delta y_{t-2}$  e  $s_t = \Delta y_{t-7}$  e, partendo da una rappresentazione AR(13) in differenze prime con intercetta e un termine in livello di ritardo uno per entrambi i regimi, si sono via via eliminati i parametri non significativi fino ad arrivare alla rappresentazione finale riportata in Tabella 3.13.

### 3.7 Valutazione del modello stimato

Successivamente alla stima dei parametri e prima di un suo impiego per gli scopi preposti è necessario sottoporre il modello a una attenta e approfondita fase di verifica. È a questo stadio che si valutano le proprietà della funzione di transizione (Figure 3.7 e 3.8 per la serie relativa al Tasso sui Fondi Federali U.S.A.), la capacità del modello di descrivere le diversità incorporate nei differenti regimi, nonché l'attendibilità delle stime ottenute in relazione alle ipotesi formulate sul modello.

In merito a quest'ultimo punto si potrebbe pensare di sottoporre a una serie di test i residui del modello STAR, in analogia a quanto sviluppato nel contesto dei modelli lineari di tipo ARMA. Tuttavia, il passaggio da una modellizzazione lineare a una di tipo non-lineare non consente di utilizzare tutte le statistiche test usualmente impiegate in quell'ambito. Ad esempio, se da un lato il test di normalità Jarque-Bera e il test per la presenza di effetti di tipo ARCH sviluppato da Engle (1982) restano validi, dall'altro la statistica di Ljung-Box costituisce un caso di test non più utilizzabile in ambito di modellistica non-lineare (Eitrheim e Teräsvirta, 1996).

Posto  $f(\mathbf{w}_t; \delta) = \pi_0 + \pi' \mathbf{w}_t + (\theta_0 + \theta' \mathbf{w}_t) F(s_t; \gamma, c)$ , un test del tipo LM per verificare la presenza di dipendenza seriale di ordine  $q$  nei residui del modello STAR può essere ottenuto calcolando la statistica  $LM_{SI}(q) = TR^2$ .  $R^2$  rappresenta l'indice di determinazione lineare della regressione ausiliaria di  $\hat{\epsilon}_t$  su  $\nabla f(\mathbf{w}_t; \hat{\delta}) = \frac{\partial f(\mathbf{w}_t; \hat{\delta})}{\partial \delta}$ , con  $\hat{\delta} = (\hat{\pi}_0, \hat{\pi}', \hat{\theta}_0, \hat{\theta}', \hat{\gamma}, \hat{c})'$ , e  $q$  residui ritardati  $\hat{\epsilon}_{t-1}, \dots, \hat{\epsilon}_{t-q}$ , mentre  $T$  indica il numero di osservazioni della stessa regressione. Sotto l'ipotesi nulla di assenza di dipendenza seriale nei residui, la statistica  $LM_{SI}(q)$  si distribuisce asintoticamente come una  $\chi^2$  con  $q$  gradi di libertà.<sup>40</sup> Le proprietà del test sono state investigate da Eitrheim e Teräsvirta (1996) mediante simulazioni e mostrano che in piccoli campioni è preferibile l'impiego della versione  $F$  del test. Nell'appendice C sono riportate le derivate analitiche rispetto ai parametri  $\gamma$  e  $c$  della funzione di transizione, sia per il caso logistico che per quello esponenziale. Le derivate rispetto a ciascuno dei parametri  $\pi_0, \pi', \theta_0$  e  $\theta'$  non sono state esplicitate data la loro semplicità di calcolo.

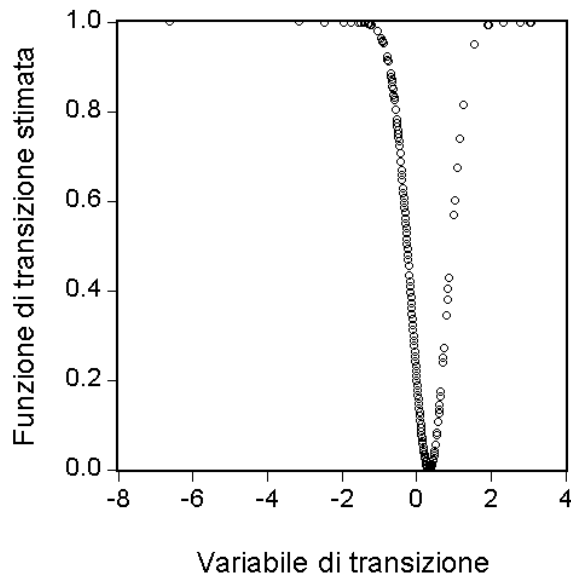
Con riferimento alla serie del Tasso sui Fondi Federali U.S.A., le principali diagnostiche sul modello ESTAR stimato (cfr. Tabella 3.14) mostrano significativi miglioramenti rispetto a quanto ottenuto attraverso il modello lineare nella Sezione 3.3:

La deviazione standard residua del modello ESTAR è del 15% inferiore a quella del modello AR. Ciò è abbastanza per compensare l'aumento del numero dei parametri (da 8 a 19) nel modello ESTAR. Si noti inoltre che le quantità AIC e SC sono inferiori rispetto alle relative controparti calcolate per il modello lineare, confermando così una migliore adattabilità ai dati del modello ESTAR.

La distribuzione empirica dei residui, pur continuando a rifiutare, anche se in misura molto inferiore, l'ipotesi di normalità come si evince dal test Jarque-Bera, presenta una significativa riduzione dell'asimmetria a sinistra (da -2.6731 a -0.8534) e della curtosi (da 40.78 a 12.49). Se da un lato questi risultati non ci consentono di trarre un'adeguata inferenza dal modello stimato, dall'altro, l'osservazione del test per ARCH e della Tabella 3.15 dove sono riportati i p-value del test di indipendenza seriale dei resi-

<sup>40</sup>Si noti che la statistica test  $LM_{SI}(q)$  non è altro che una generalizzazione del test di Godfrey (1979) dove  $f(\mathbf{w}_t; \delta) = \pi' \mathbf{w}_t$ , con  $\delta = \pi$  (van Dijk, Teräsvirta e Franses, 2000).

Figura 3.7: Funzione di transizione esponenziale stimata. Modello con variabile di transizione  $s_t = \Delta y_{t-7}$ .



dui, ci suggeriscono di concentrare l'attenzione su eventuali dinamiche evolutive della varianza condizionata.

Infine, occorre sottolineare che, se da un lato il modello ESTAR con  $s_t = \Delta y_{t-2}$  costituisce un significativo miglioramento rispetto a quello lineare, dall'altro presenta delle diagnostiche sui residui (cfr. Tabella 3.14) ampiamente peggiori di quelle del modello con  $s_t = \Delta y_{t-7}$  e per questa ragione è stato scartato. Tale considerazione è ulteriormente confermata dall'osservazione dei criteri di informazione AIC e SC, nonché dal valore della funzione di log-verosimiglianza (Tabella 3.13)

### 3.7.1 Test per non-linearità rimanenti

Quando si utilizzano modelli non-lineari per descrivere la dinamica di certi fenomeni economici, un punto che merita attenzione è senz'altro la verifica della presenza di non-linearità rimanenti nei residui del modello stimato. Nel caso particolare dei modelli STAR, che come si è visto possono essere interpretati come modelli con cambiamento di regime, un'ovvia soluzione per sottoporre a test la presenza di non-linearità residue è quella di specificare l'ipotesi nulla in modo tale da supporre la presenza di un regime, o stato, aggiuntivo.

Eitrheim e Teräsvirta (1996) derivano un test LM contro la presenza di un terzo regime di tipo additivo nel modello STAR, così espresso:

$$y_t = \pi_0 + \pi' \mathbf{w}_t + (\theta_0 + \theta' \mathbf{w}_t) F_1(s_{1t}; \gamma_1, c_1) + (\rho_0 + \rho' \mathbf{w}_t) F_2(s_{2t}; \gamma_2, c_2) + \epsilon_t \quad (3.78)$$

L'ipotesi nulla può essere formulata sia come  $H_{0,1} : \rho_0 = 0, \rho = 0$  che come  $H_{0,2} :$

Tabella 3.14: Diagnostiche sui residui di stima dei modelli ESTAR sulla serie del Tasso sui Fondi Federali U.S.A.

	$\Delta y_{t-7}$	$\Delta y_{t-2}$		
$\hat{\sigma}_\epsilon$	0.4617	0.4915		
$\hat{\sigma}_{ESTAR}/\hat{\sigma}_{AR}$	0.8493	0.9042		
Skewness	-0.8534	-1.4929		
Curtosi	12.49	27.64		
Test	Statistica-test	p-value	Statistica-test	p-value
Jarque-Bera	1827.54	0.0000	12117.13	0.0000
ARCH(1)	7.0951	0.0077	1.7479	0.1861
ARCH(4)	40.1821	$3.97 \times 10^{-8}$	44.4603	$5.15 \times 10^{-9}$

Tabella 3.15: Test di indipendenza seriale sui residui del modello ESTAR stimato con variabile di transizione  $s_t = \Delta y_{t-7}$ .

$q$	Autocorrelazione seriale di ordine $q$					
	2	4	6	8	10	12
p-value ( $F$ )	0.1475	0.0108	0.0047	$7.25 \times 10^{-5}$	0.0003	0.0002
p-value ( $\chi^2$ )	0.1361	0.0089	0.0037	$5.80 \times 10^{-5}$	0.0002	0.0002

$\gamma_2 = 0$ . In entrambi i casi, in analogia a quanto discusso nella Sezione 4.5, il test soffre di un problema di identificazione. La soluzione proposta è sempre la stessa:<sup>41</sup> rimpiazzare la funzione di transizione  $F_2(s_{2t}; \gamma_2, c_2)$  mediante la sua approssimazione di Taylor al terzo ordine, valutandola nell'intorno del punto  $\gamma_2 = 0$ . Dal modello ausiliario così risultante,<sup>42</sup>

$$y_t = \beta_{0,0} + \beta'_0 \mathbf{w}_t + (\theta_0 + \theta' \mathbf{w}_t) F_1(s_{1t}; \gamma_1, c_1) + \beta'_1 \mathbf{w}_t s_{2t} + \beta'_2 \mathbf{w}_t s_{2t}^2 + \beta'_3 \mathbf{w}_t s_{2t}^3 + u_t, \quad (3.79)$$

si evince che testare l'ipotesi nulla  $H_{0,2} : \gamma_2 = 0$  nella (3.78), equivale a testare l'ipotesi  $H_{0,3} : \beta_1 = \beta_2 = \beta_3 = 0$  nella (3.79).<sup>43</sup> La statistica test  $LM_{AMR}$  può essere ricavata come  $TR^2$  dalla regressione ausiliaria di  $\hat{e}_t$  (residui del modello STAR a due regimi stimato) sulle derivate parziali della funzione di regressione rispetto ai parametri del modello a due regimi  $(\pi_0, \pi', \theta_0, \theta', \gamma_1, c_1)$ , valutate sotto l'ipotesi nulla, e i regressori ausiliari  $\mathbf{w}_t s_{2t}^i$ ,  $i = 1, 2, 3$ . La risultante quantità<sup>44</sup>  $LM_{AMR}$  possiede distribuzione asintotica di tipo  $\chi^2$  con  $3p$  gradi di libertà.

Ovviamente, generalizzazioni del test fino ad ipotizzare la presenza di più regi-

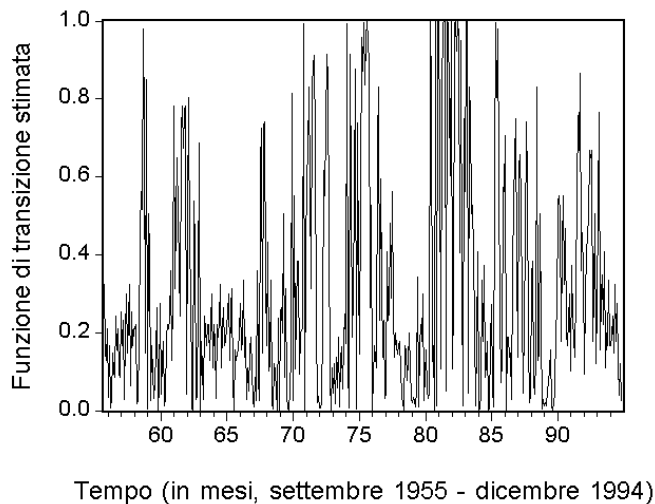
<sup>41</sup>Luukkonen *et al.* (1988b).

<sup>42</sup>Dove i parametri  $\beta_{0,0}$  e  $\beta_i$ ,  $i = 1, 2, 3$  sono funzioni di quelli del modello originario espresso dalla (3.78),  $\pi_0, \pi, \rho_0, \rho, \gamma_2, c_2$ , e nel quale si è ipotizzata una funzione di transizione logistica. Seguendo Teräsvirta (1994) il test così derivato ha potenza anche contro il caso di una funzione di transizione esponenziale (Eitrheim e Teräsvirta, 1996).

<sup>43</sup>Nella (3.79) si è implicitamente assunto che  $s_{2t}$  sia un elemento di  $\mathbf{w}_t$ . Se così non fosse, i regressori ausiliari  $\beta_{i,0} s_{2t}^i$ ,  $i = 1, 2, 3$ , dovrebbero essere aggiunti alla (3.79) (van Dijk *et al.*, 2000).

<sup>44</sup>Dove il pedice AMR indica che si sottopone a test un modello STAR a regimi multipli di tipo additivo.

Figura 3.8: Andamento della funzione di transizione esponenziale stimata contro il tempo. Modello con variabile di transizione  $s_t = \Delta y_{t-7}$ .



mi sono possibili e saranno dettagliatamente considerate nel capitolo successivo dove verranno presi in considerazione i modelli MR-STAR.

### 3.7.2 Verifica della costanza dei parametri

Fino ad ora si è implicitamente assunto che, sia i parametri del modello lineare, sia quelli del modello STAR fossero costanti nel tempo. In entrambi i casi, un'accurata specificazione del modello non può esimersi dalla verifica empirica di una tale ipotesi che, se rifiutata, implicherebbe di dover nuovamente specificare il modello, o comunque non accettare quello stimato.

La dinamica di molti fenomeni economici è tale da lasciar supporre un qualche tipo di fondatezza in questo modo di procedere. Ad esempio, se l'oggetto di studio è rappresentato da serie relative a tassi d'interesse, il cambiamento di politica monetaria da parte di una banca centrale, che passa da una politica restrittiva ad una espansiva o viceversa, è un elemento da tenere in grossa considerazione in quanto il modello stimato, pur potendo rimanere valido nella sua specificazione formale, può subire delle variazioni nel valore dei parametri. In queste circostanze è la stessa dinamica degli eventi a suggerire la possibilità di cambiamenti strutturali<sup>45</sup>. Altre volte invece, non si è in grado di ricondurre questa evenienza a momenti particolari dell'economia, e si deve pertanto ricorrere a metodologie puramente statistiche che ci permettano di individuare il cosiddetto "punto di rottura".

Eitrheim e Teräsvirta (1996) derivano un test per verificare l'ipotesi di costan-

<sup>45</sup>Ed è in questo ambito che si colloca il test di Chow (1960), il quale presuppone la conoscenza del momento in cui avviene il cambiamento strutturale ipotizzato (il momento del passaggio da una politica espansiva ad una restrittiva, o viceversa).

za dei parametri in un modello STAR, contro l'alternativa di parametri che variano gradualmente nel tempo. Il modello STAR oggetto di analisi può essere così espresso:

$$y_t = \pi_0(t) + \pi'(t)\mathbf{w}_t + (\theta_0(t) + \theta'(t)\mathbf{w}_t)F_1(s_t; \gamma_1, c_1) + \epsilon_t \quad (3.80)$$

con

$$\pi_0(t) = \pi_0 + \varphi_{1,0}F_2(t; \gamma_2, c_2) \quad (3.81)$$

$$\pi'(t) = \pi' + \varphi_1'F_2(t; \gamma_2, c_2) \quad (3.82)$$

$$\theta_0(t) = \theta_0 + \varphi_{2,0}F_2(t; \gamma_2, c_2) \quad (3.83)$$

$$\theta'(t) = \theta' + \varphi_2'F_2(t; \gamma_2, c_2) \quad (3.84)$$

e dove, al solito,  $F_i(\cdot)$ ,  $i = 1, 2$ , può assumere le specificazioni logistica o esponenziale. Sostituendo le espressioni (3.81) - (3.84) nella (3.80), il modello può essere riscritto come

$$y_t = \pi_0 + \pi'\mathbf{w}_t + (\theta_0 + \theta'\mathbf{w}_t)F_1(s_t; \gamma_1, c_1) + (\varphi_{1,0} + \varphi_1'\mathbf{w}_t)F_2(t; \gamma_2, c_2) + (\varphi_{2,0} + \varphi_2'\mathbf{w}_t)F_1(s_t; \gamma_1, c_1)F_2(t; \gamma_2, c_2) + \epsilon_t \quad (3.85)$$

e prende il nome di modello STAR con parametri che variano nel tempo (TV-STAR). La caratteristica principale di questo modello è quella di permettere l'evoluzione di dinamiche non-lineari di tipo STAR, all'interno delle quali il valore dei parametri cambia gradualmente nel tempo.

Senza entrare in modo specifico nel merito dei modelli TV-STAR, che costituiscono oggetto del Capitolo 4, per il momento ci basta sottolineare che il test per la costanza dei parametri in un modello STAR può essere facilmente ricondotto alla verifica dell'ipotesi  $\gamma_2 = 0$  nell'equazione (3.85), la cui statistica test si ottiene a partire da quanto già sviluppato nella Sezione 3.5. Per ulteriori dettagli e generalizzazioni si rimanda al Capitolo successivo.

### 3.7.3 Valutazione del modello mediante estrapolazione

Come per tutti i modelli non-lineari, la derivazione di requisiti per la stazionarietà di una rappresentazione STAR è molto complessa. Anche se fosse possibile trovare delle condizioni teoriche a cui il modello dovesse sottostare, la loro applicazione pratica si rivelerebbe piuttosto macchinosa. Tuttavia, se lo scopo per cui è stato costruito il modello è quello previsivo, la conoscenza della sua stabilità o meno si rivela particolarmente importante in quanto da un modello esplosivo si possono trarre soltanto previsioni a breve o brevissimo termine. Un aumento dell'orizzonte previsivo implica inevitabilmente la verifica della stabilità del modello.

Granger e Teräsvirta (1993) presentano una condizione necessaria basata sullo *skeleton* (Chan e Tong, 1985) del modello così definito:

$$f(\mathbf{w}_t; \delta) = \pi_0 + \pi'\mathbf{w}_t + (\theta_0 + \theta'\mathbf{w}_t)F(s_t; \gamma, c) \quad (3.86)$$

Estrapolando  $f(\mathbf{w}_t; \delta)$  per diversi valori iniziali, la condizione di cui sopra consiste nella verifica della convergenza ad un singolo punto, detto stazionario, della serie dei

valori ricavati dallo *skeleton*. Si noti che il punto stazionario stabile non deve essere necessariamente unico, ma ne possono esistere anche più di uno. Se il modello STAR stimato soddisfa questa condizione, lo stesso esercizio potrebbe essere ripetuto aggiungendo allo *skeleton* il termine di errore  $\epsilon_t$ . Per ulteriori dettagli si rimanda a Tong (1990) e Granger e Teräsvirta (1993).

### 3.8 Utilizzo del modello

L'ultima fase della procedura di specificazione consiste nell'impiego del modello per gli scopi preposti. Sebbene non siano da trascurare impieghi relativi ad analisi descrittive del fenomeno oggetto di studio o a simulazioni dello stesso, un ruolo predominante è senz'altro svolto dalla previsione. Ed è proprio a quest'ultima modalità di utilizzo che ci riferiremo in questo paragrafo.

La metodologia impiegata per prevedere dai modelli STAR è la stessa che viene adoperata per tutti gli altri modelli non-lineari. Consideriamo la generica formulazione

$$y_t = f(\mathbf{x}_t; \delta) + \epsilon_t \quad (3.87)$$

dove, in tal caso,  $f(\cdot)$  rappresenta lo *skeleton* del modello STAR,  $\mathbf{x}_t = (1, \mathbf{w}_t')'$  e immaginiamo di voler prevedere  $h$  valori futuri per  $y_t$ . Utilizzando il criterio dell'errore quadratico medio di previsione (MSPE), il previsore ottimale ad un periodo per  $y_t$  è (Hamilton, 1994):

$$y_{t,1}^f = E[y_{t+1} | I_t] = f(\mathbf{x}_{t+1}; \delta) \quad (3.88)$$

dove  $I_t$  rappresenta l'informazione disponibile al tempo  $t$ ,<sup>46</sup> mentre il vettore  $\mathbf{x}_{t+1} = (1, y_t, y_{t-1}, \dots, y_{t-(p-1)})'$ . Le cose si complicano quando allarghiamo l'orizzonte previsionale. Ad esempio, per  $h = 2$  avremo

$$y_{t,2}^f = E[y_{t+2} | I_t] = E[f(\mathbf{x}_{t+2}^f; \delta) + \epsilon_{t+2} | I_t] = E[f(\mathbf{x}_{t+2}^f; \delta) | I_t] \quad (3.89)$$

con  $\mathbf{x}_{t+2}^f = (1, y_{t,1}^f + \epsilon_{t+1}, y_t, \dots, y_{t-(p-2)})'$ . Per il calcolo del valore atteso condizionato espresso dalla (3.89) esistono in letteratura varie metodologie (Granger e Teräsvirta, 1993), ciascuna delle quali presenta un trade-off tra vantaggi e svantaggi nei confronti di tutte le altre:

- Calcolo esatto del valore atteso condizionato

$$y_{t,2}^{f(e)} = E[f(\mathbf{x}_{t+2}^f; \delta) | I_t] = \int_{-\infty}^{+\infty} f(\mathbf{x}_{t+2}^f; \delta) d\Phi(z) dz \quad (3.90)$$

dove  $\Phi(z)$  indica la funzione di distribuzione cumulata di  $\epsilon_{t+1}$ . Il calcolo dell'integrale avviene per via numerica e, se da un lato può non essere molto complicato

<sup>46</sup>Nel modello STAR,  $I_t$  è costituito dai valori passati di  $y_t$ . Si noti che la (3.88) coincide con il previsore ottimale ad un periodo per il caso in cui  $f(\cdot)$  è una funzione lineare. In particolare, in questo caso, coincide col previsore ottimale di un modello AR(p).



nel caso di modelli STAR, dall'altro le sue dimensioni crescono considerevolmente con l'orizzonte previsivo. Questo ci fa notare che sarebbe più flessibile, anche se meno preciso, ottenere le previsioni ricorsivamente senza dover calcolare l'integrale. Un modo semplice è quello di non considerare il termine  $\epsilon_{t+1}$  e basare la previsione sullo *skeleton* del modello.

- Metodo *naïve* (Granger e Teräsvirta, 1993) dove

$$y_{t,2}^{f(n)} = f(\mathbf{x}_{t+2}^f; \delta) \quad \text{con} \quad \mathbf{x}_{t+2}^f = \left(1, y_{t,1}^f, y_t, \dots, y_{t-(p-2)}\right)' \quad (3.91)$$

L'utilizzo di questo metodo si rivela piuttosto semplice, ma non si può evitare di ottenere estrapolazioni distorte. Altri metodi sono basati sulla simulazione<sup>47</sup> (Granger e Teräsvirta, 1993).

- Monte Carlo:

$$y_{t,2}^{f(mc)} = \frac{1}{M} \sum_{m=1}^M f(\mathbf{x}_{t+2,m}^f; \delta) \quad (3.92)$$

dove ognuno degli  $M$  valori di  $\epsilon_{t+1}$  all'interno di  $\mathbf{x}_{t+2,m}^f$  è generato indipendentemente dalla distribuzione ipotizzata per il termine di errore del modello STAR.<sup>48</sup> Per  $M \rightarrow +\infty$ ,  $y_{t,2}^{f(mc)}$  e  $y_{t,2}^{f(e)}$  sono uguali grazie alla legge debole dei grandi numeri (WLLN).

- *Bootstrap*:

$$y_{t,2}^{f(b)} = \frac{1}{B} \sum_{b=1}^B f(\mathbf{x}_{t+2,b}^f; \delta) \quad (3.93)$$

dove ciascuno dei  $B$  valori di  $\epsilon_{t+1}$  dentro  $\mathbf{x}_{t+2,b}^f$  è generato indipendentemente dalla funzione di distribuzione empirica (*empirical density function*, EDF) dei residui ricavati dal modello stimato. Il vantaggio derivante dall'utilizzo del metodo *bootstrap* in luogo di quello Monte Carlo è dovuto all'assenza di ipotesi aggiuntive concernenti la distribuzione di  $\epsilon_t$ .

Fino ad ora si è assunto che i parametri del vettore  $\delta$  fossero noti. In pratica ciò non accade, per cui il vettore  $\delta$  deve essere sostituito con una sua stima, aumentando così l'incertezza delle previsioni. Di questa incertezza campionaria si può tener conto estendendo i metodi Monte Carlo e *bootstrap* di cui sopra fino a comprendere diversi valori per i parametri stimati, ognuno dei quali ottenuto numericamente a partire dalla distribuzione empirica di  $\hat{\delta}$  (van Dijk *et al.*, 2000 e Lundbergh e Teräsvirta, 2000).

Alle previsioni puntuali ottenute secondo uno dei metodi visti si è soliti affiancare un intervallo di confidenza. Mentre per un modello lineare la regione di confidenza standard per le previsioni è simmetrica intorno al valore puntuale, questo non è generalmente vero quando si passa a lavorare con modelli non-lineari. La distribuzione condizionata  $g(y_{t,h} | I_t)$  non è necessariamente simmetrica intorno a  $E[y_{t,h} | I_t]$  (stimato

<sup>47</sup>Per una introduzione ai metodi di simulazione di tipo Monte Carlo e *bootstrap* si veda, ad esempio, Davidson e MacKinnon (1993), Capitolo 21.

<sup>48</sup>Nel nostro caso la distribuzione di probabilità della normale.

secondo uno dei metodi sopra descritti), e può anche essere plurimodale. Ammettendo queste possibilità, Hyndman (1995) propone tre metodi per costruire intervalli di confidenza di previsione ad un generico livello di confidenza del  $100(1 - \alpha)\%$ :

1. Un intervallo simmetrico intorno alla previsione puntuale

$$S_\alpha = \left( y_{t,h}^f - w, y_{t,h}^f + w \right) \quad (3.94)$$

dove  $w > 0$  è tale che  $\Pr(y_{t,h} \in S_\alpha | I_t) = 1 - \alpha$ .

2. Un intervallo definito dai quantili  $\alpha/2$  e  $(1 - \alpha/2)$  della distribuzione di previsione, indicati rispettivamente con  $q_{\alpha/2}$  e  $q_{1-\alpha/2}$ :

$$Q_\alpha = (q_{\alpha/2}, q_{1-\alpha/2}) \quad (3.95)$$

3. Costruire la regione con la più alta densità (HDR)

$$HDR_\alpha = \{y_{t,h} | g(y_{t,h} | I_t) \geq g_\alpha\} \quad (3.96)$$

dove  $g_\alpha$  è tale che  $\Pr(y_{t,h} \in HDR_\alpha | I_t) = 1 - \alpha$ , mentre  $g(\cdot)$  indica la densità del suo argomento.

Per distribuzioni simmetriche e unimodali, i tre metodi portano agli stessi identici risultati, mentre per distribuzioni asimmetriche o plurimodali, Hyndman (1995) indica in  $HDR_\alpha$  la scelta più appropriata poiché rappresenta la più piccola regione di previsione possibile al livello di confidenza specificato, ed ogni punto all'interno di HDR possiede densità condizionata  $g(y_{t,h} | I_t)$  grande almeno quanto ogni punto esterno a tale regione. Per ulteriori dettagli sul calcolo e sulla realizzazione del grafico di HDR si veda Hyndman (1996).

L'importanza di previsioni ottenute da modelli non-lineari è spesso valutata con riferimento a un modello lineare che costituisce il cosiddetto *benchmark*. Dal momento che previsioni multiperiodali con modelli non-lineari sono molto più complicate da ottenere, la loro effettiva convenienza va quindi confrontata con quelle ricavabili da modelli più semplici allo scopo di determinarne il valore aggiunto realizzabile. Un confronto tra simili previsioni può essere facilmente basato sul calcolo di alcuni dei tradizionali indicatori della bontà di adattamento, fra i quali merita segnalare l'errore medio quadratico di previsione (MSPE), l'errore medio assoluto di previsione (MAPE) e gli indici  $U_1$  e  $U_2$  di Theil.<sup>49</sup> Sfortunatamente, questo confronto può fornire risultati inaccettabili se le caratteristiche non-lineari della serie non sono presenti durante il periodo di previsione. Per questa ragione, se la performance ottenuta da una parametrizzazione non-lineare risulta peggiore, la specificazione del modello dovrebbe essere riconsiderata (Granger e Teräsvirta, 1993).

<sup>49</sup>Sull'impiego di tali indici nella valutazione della performance di previsione, si veda ad esempio Greene (1997).

### 3.9 Considerazioni conclusive

Lo scopo del Capitolo è stato quello di introdurre i modelli autoregressivi a transizione liscia (STAR), mettendone in evidenza le principali caratteristiche e le proprietà teoriche. Sulla falsariga di quanto proposto da Teräsvirta (1994) ed Eitrheim e Teräsvirta (1996), la struttura del Capitolo è stata suddivisa in modo da rispecchiare le varie fasi di specificazione, stima, valutazione e utilizzo del modello, evidenziandone di volta in volta gli aspetti peculiari, senza tuttavia tralasciare la possibilità di ulteriori generalizzazioni.

Ad una trattazione teorica dell'argomento, è stata affiancata un'applicazione reale sulla serie del Tasso sui Fondi Federali U.S.A., la cui analisi non può certo considerarsi completamente esaustiva. Come evidenziato in fase di valutazione, la forte presenza di eteroschedasticità sui residui del modello ESTAR stimato, non consente di utilizzarlo per scopi previsivi, senza aver prima specificato e stimato un'adeguata parametrizzazione della varianza condizionata. I risultati raggiunti sulla media condizionata sono comunque di buon auspicio, data la particolare difficoltà riscontrata in letteratura nel modellare la serie in oggetto e, in generale, tutte quelle caratterizzate da un andamento simile.<sup>50</sup> Numerosi sono stati i tentativi di descriverne il comportamento tramite modelli TAR, *Markov-Switching* e, di recente (Lanne e Saikkonen, 2000), autoregressivi di tipo *mixture* (MAR), senza però ottenere specificazioni finali completamente esenti da critiche.

Dato il grosso successo riscontrato dai modelli STAR in ambito macroeconomico, specialmente nell'ottenere previsioni più accurate di quelle ricavabili da formulazioni lineari, l'applicazione qui svolta è stata un tentativo di estensione dei modelli al contesto monetario/finanziario. La speranza è che nel futuro possano essere considerate ulteriori applicazioni in quest'ambito, allo scopo di valutare se, effettivamente, una tale specificazione per la media condizionata dei rendimenti degli attivi finanziari possa essere ritenuta in grado di coglierne completamente la struttura dinamica. Il tutto alla luce di una relativamente facile modalità d'impiego dei modelli STAR rispetto a formulazioni alternative, spesso in concorrenza, che si rivelano invece più complicate da utilizzare.

---

<sup>50</sup>In particolare, ci riferiamo alle serie relative ai tassi d'interesse a breve termine degli Stati Uniti.

## Capitolo 4

# Sviluppi recenti

Negli ultimi anni l'attenzione rivolta ai modelli di tipo STAR è aumentata considerevolmente. La generalità di questa classe di modelli, che comprende altre formulazioni già note oramai da più tempo in letteratura, e la sua relativamente facile operatività sono state sicuramente i precursori di nuovi sviluppi e nuovi miglioramenti. Se poi teniamo conto dei buoni risultati raggiunti da questi modelli in applicazioni di tipo macroeconomico, sia in ambito previsivo che in ambito di analisi del ciclo economico, si capisce subito il perchè di un tale successo e di un tale interesse.

Le principali linee di sviluppo del modello base riguardano, da una parte la ricerca di possibili robustificazioni del test di linearità quando sono presenti nella serie eteroschedasticità generica od *outliers*, e dall'altra l'estensione del modello ai casi di regimi multipli (MR-STAR) e di parametri che variano nel tempo (TV-STAR).

La struttura del Capitolo è la seguente: nelle Sezioni 4.1 e 4.2 vengono brevemente discusse le correzioni da apportare al test di linearità in presenza, rispettivamente di eteroschedasticità ed *outliers*. La Sezione 4.3 introduce un nuovo tipo di notazione che sarà utile per descrivere i modelli STAR a più regimi (MR-STAR) nella Sezione 4.4, e quelli con parametri variabili nel tempo (TV-STAR) nella 4.5. La Sezione 4.6 presenta alcune considerazioni conclusive.

### 4.1 Eteroschedasticità e test di linearità

I diversi test di linearità presentati nella Sezione 3.5 presuppongono che la varianza condizionata del termine di errore  $\epsilon_t$  sia costante. Effettuare un test di linearità contro l'alternativa di tipo STAR tralasciando la presenza di effetti eteroschedastici produce distorsioni simili a quelle che si manifestano quando il modello lineare preso come punto di riferimento del test non è stato specificato correttamente.<sup>1</sup>

Davidson e MacKinnon (1985) e Wooldridge (1990 e 1991) sviluppano un test che può essere utilizzato in presenza di eteroschedasticità, senza la necessità di specificarne

---

<sup>1</sup>Ad esempio, i residui del modello lineare non sono stati completamente ripuliti dalla presenza di correlazione seriale.

la forma (che spesso è incognita). La loro procedura (si veda anche Granger e Teräsvirta, 1993 e van Dijk *et al.*, 2000) può essere impiegata per costruire una versione robusta del test di linearità contro l'alternativa di tipo STAR. Con riferimento alla statistica test  $LM_3$  derivata nella Sezione 3.5.1 e basata sull'equazione (3.32), il calcolo della sua variante sensibile alla presenza di eteroschedasticità consiste nei seguenti passi:

1. Stimare il modello lineare e calcolare i residui  $\hat{\epsilon}_t$ .
2. Regredire ciascun regressore ausiliario  $\mathbf{w}_t y_{t-d}^i$ ,  $i = 1, 2, 3$ , su una costante e  $\mathbf{w}_t$  e calcolarne i residui  $\hat{r}_t$ .
3. Ponderare la matrice dei residui di cui al punto precedente con i residui  $\hat{\epsilon}_t$  e regredire il vettore unitario sulla matrice dei residui ponderata. La devianza spiegata (ESS) di quest'ultima regressione rappresenta la versione robusta della statistica test  $LM_3$ .

La procedura descritta si può estendere facilmente al calcolo delle altre statistiche test della Sezione 3.5. Luundbergh e Teräsvirta (1998) presentano simulazioni del test in presenza di eteroschedasticità di tipo GARCH standard. I risultati ottenuti mostrano che spesso il test di linearità perde molta della sua potenza portando all'accettazione dell'ipotesi nulla quando in realtà dovrebbe essere rifiutata. Se l'obiettivo della ricerca è quello di modellare adeguatamente la nonlinearità presente nella serie, certamente questo approccio non può essere raccomandato. La motivazione di ciò è dovuta al fatto che, se il modello è veramente lineare, in sede di valutazione ci dovremmo accorgere della presenza di eteroschedasticità condizionata che, come già accennato, può distorcere i risultati del test. Si tenga anche presente che il test per ARCH di Engle (1982) ha potenza contro la presenza di nonlinearità nella media condizionata; risultati ottenuti per simulazione si possono trovare, ad esempio, in Luukkonen, Saikkonen e Teräsvirta (1988a) e in Lee, White e Granger (1993).

Una procedura alternativa che tiene conto della presenza di eteroschedasticità condizionata nel test di linearità è stata proposta da Andersson, Eklund e Lyhagen (1999). Allo scopo di correggere la distorsione del test in simili circostanze propongono di utilizzare la tecnica del *bootstrap* parametrico. La metodologia di tipo *bootstrap* fa uso di un singolo data set per effettuare una sorta di esperimento Monte Carlo all'interno del quale i dati disponibili sono impiegati per approssimare la distribuzione del termine di errore. Non si tratta di una vera e propria simulazione Monte Carlo poiché i valori per il termine di errore non sono estratti da una densità ipotizzata, ma sono generati mediante l'impiego della loro funzione di densità empirica (*empirical density function*, EDF, Davidson e MacKinnon, 1993).

Nonostante siano presenti in letteratura molti modi d'impiego del *bootstrap*, quello utilizzato da Andersson *et al.* (1999) si basa sul calcolo del p-value. La teoria asintotica dei test d'ipotesi è esatta solamente se il p-value è indipendente dal processo generatore dei dati (DGP) e dall'ampiezza campionaria. In pratica, questa eventualità non è quasi mai verificata e la soluzione consiste nel rimpiazzare il p-value mediante la sua controparte di tipo *bootstrap*, che può essere stimata come

$$\hat{p}^*(\hat{\tau}) \equiv \frac{1}{B} \sum_{j=1}^B \mathbb{I}(\tau_j^* > \hat{\tau}) \quad (4.1)$$

dove  $\mathfrak{S}(\cdot)$  rappresenta la funzione indicatrice,  $\hat{\tau}$  indica la statistica test calcolata nel modo tradizionale, mentre  $\tau_j^*$  è la sua controparte bootstrap e  $B$  indica il numero delle replicazioni necessarie. Davidson e MacKinnon (1998) dimostrano che un  $B$  almeno pari a 399 è necessario per garantire una perdita di potenza del test inferiore all'1% per  $\alpha$  pari a 0.05. Inoltre, per  $B \rightarrow \infty$ , il p-value *bootstrap* stimato  $\hat{p}^*(\hat{\tau})$  tende al p-value *bootstrap* vero  $p^*(\hat{\tau})$ .

Se l'analisi svolta consiste in un test, ciascuna delle  $B$  replicazioni *bootstrap* deve obbedire all'ipotesi nulla. Nel caso STAR quindi, ogni campione è generato a partire dal modello lineare, selezionato mediante i criteri di informazione AIC e SC (si veda la Sezione 3.3).

La presenza di eteroschedasticità condizionata deve essere opportunamente specificata. La scelta più naturale, specialmente in un contesto monetario/finanziario, dovrebbe ricadere su una dinamica di tipo GARCH(1,1) per la varianza del termine di errore. Più in generale, una volta stimato il modello lineare, è dall'analisi dei suoi residui al quadrato e dal test per ARCH che ci dovremmo fare un'idea precisa sulla forma dell'eteroschedasticità presente. Sulla base di queste informazioni e con riferimento alla statistica  $LM_3$ , la procedura di test può essere sintetizzata nei seguenti punti (Skalin, 1998)<sup>2</sup>:

1. Calcolare la statistica  $LM_3$  nel modo tradizionale, ovvero secondo quanto specificato nella Sezione 3.5.1. Ciò equivale a calcolare  $\hat{\tau}$  nella (4.1).
2. Stimare il modello lineare AR(p) con residui di tipo GARCH (ad esempio GARCH(1,1)). La stima dei parametri di questo modello è ottenuta mediante il metodo della massima verosimiglianza.
3. Costruire i residui *bootstrap*  $b_t^*$  a partire dalla distribuzione  $N(0, 1)$  ipotizzata per  $\epsilon_t$  nel modello STAR; se  $\epsilon_t^*$  rappresenta un'estrazione indipendente da una  $N(0, 1)$ , i  $b_t^*$  sono così generati:

$$\hat{h}_t = \hat{\omega} + \hat{\alpha}_1 b_{t-1}^{*2} + \hat{\beta}_1 \hat{h}_{t-1} \quad (4.2)$$

$$b_t^* = \epsilon_t^* \sqrt{\hat{h}_t} \quad (4.3)$$

dove  $\hat{\omega}$ ,  $\hat{\alpha}_1$  e  $\hat{\beta}_1$  sono stati stimati al punto precedente.

4. Costruire i  $B$  campioni *bootstrap* ricorsivamente mediante l'equazione

$$y_{t,b}^* = \hat{\phi}_0 + \hat{\phi}_1 y_{t-1,b}^* + \dots + \hat{\phi}_p y_{t-p,b}^* + b_t^* \quad (4.4)$$

dove i parametri  $\hat{\phi}_0, \hat{\phi}_1, \dots, \hat{\phi}_p$  sono quelli stimati al punto 2.

5. Per ogni campione generato, calcolare la statistica  $LM_3$  secondo quanto specificato nella Sezione 3.5.1. Questo, nella (4.1) corrisponde al calcolo dei  $B$  valori di  $\tau_j^*$ ,  $j = 1, \dots, B$ .
6. Calcolare il p-value *bootstrap* utilizzando la (4.1).

<sup>2</sup>La procedura è facilmente generalizzabile al calcolo delle altre statistiche test della Sezione 3.5.

Naturalmente, complicazioni e specificazioni alternative per quanto concerne la forma dell'eteroschedasticità sono possibili. L'unica cosa da tenere presente in questi casi è che si tratta di una metodologia numerica che può richiedere anche molto tempo per essere impiegata. La sua effettiva convenienza va quindi valutata caso per caso e sarebbe anche utile uno studio di tipo Monte Carlo per valutare la potenza del test in contrapposizione alla procedura, più generica ma più semplice da calcolare, di Davidson e MacKinnon (1985) e Wooldridge (1990 e 1991).

Anderson *et al.* (1999) ipotizzano una dinamica della varianza condizionata di tipo ARCH(1) e mostrano, mediante simulazioni Monte Carlo, che in presenza di eteroschedasticità specifica il test *bootstrap* raggiunge performance nettamente migliori di quelle del test tradizionale di Luukkonen *et al.* (1988b). Tuttavia, se i disturbi casuali non sono caratterizzati da varianza condizionata variabile, i risultati ottenuti confermano la bontà del test tradizionale e la sua probabile maggior potenza.

## 4.2 Outliers e test di linearità

La gran parte delle serie economiche e finanziarie è caratterizzata dalla presenza di valori eccezionalmente grandi o eccezionalmente piccoli rispetto alla sua media. Molto spesso queste osservazioni sono associate ad altrettanti eventi straordinari che, il più delle volte, si esauriscono in quella sporadica apparizione. Altre volte invece, può accadere che simili circostanze si verificano più di frequente lungo il range temporale considerato, lasciando presagire la presenza di una qualche struttura nel DGP della serie, tale per cui questi valori non possono essere considerati casi isolati. In queste circostanze, la sostituzione dei valori anomali mediante le tradizionali tecniche di trattamento degli *outliers* non si rivela generalmente appropriata.<sup>3</sup> La drastica rimozione di queste osservazioni può provocare la perdita di alcune delle proprietà intrinseche alla serie, quali la non-linearità, che possono invece essere molto utili in fase di previsione o di semplice descrizione del fenomeno. Per contro, se il vero modello è effettivamente lineare ma risulta contaminato dalla presenza di *outliers*, anche il test di linearità può subire pesanti alterazioni. E' questo il caso del rifiuto dell'ipotesi di linearità dovuto alla presenza di valori anomali che modificano la struttura, apparentemente in senso non-lineare, della serie.

Le considerazioni che precedono indicano la necessità di disporre di una strategia che sia in grado di distinguere efficacemente tra non-linearità e *outliers* nella specificazione del modello. Prima però di presentare la metodologia proposta in letteratura in simili circostanze e prima di descrivere le modifiche da apportare al test di linearità derivato nella Sezione 3.5, occorre soffermarsi sul trattamento degli *outliers*. Di questo ci occuperemo nella Sezione che segue.

### 4.2.1 Una breve discussione sugli outliers

In questa Sezione verranno brevemente presentate alcune tra le più importanti tipologie di *outliers* che si possono considerare all'interno di un modello lineare di tipo AR(p).

<sup>3</sup>Per una rassegna delle principali tecniche di trattamento dei valori anomali si veda, ad esempio, Balke e Fonby (1994).

Tra le motivazioni che ci inducono a considerare modelli lineari in luogo di modelli più complicati (quali il modello STAR) vi è senz'altro la constatazione che il test di linearità contro l'alternativa di tipo STAR è basato sulla stima di un modello ausiliario che è lineare nei parametri.

A tal fine si consideri il modello (Martin e Yohai, 1986),

$$y_t = \phi(L)y_t(1 - \delta_t) + \zeta_t\delta_t + \epsilon_t, \quad t = 1, \dots, T \quad (4.5)$$

dove  $\phi(L) = 1 - \phi_1L - \dots - \phi_pL^p$  rappresenta il polinomio autoregressivo nell'operatore ritardo  $L$ ,  $\delta_t$  è una variabile binaria che assume valore 1 con probabilità  $\pi$  e valore 0 altrimenti, mentre differenti specificazioni del processo  $\zeta_t$  possono generare diverse tipologie di *outliers*; infine, al solito,  $\epsilon_t \sim iid(0, \sigma_\epsilon^2)$ .

Il processo descritto dalla (4.5) può essere considerato come composto da due parti: una parte che rappresenta la linea guida del processo  $\phi(L)y_t$  (il cosiddetto *core process*) e un processo contaminante  $\zeta_t$  (*contaminating process*). Queste definizioni sono giustificate dalla considerazione che, di solito,  $\pi$  è molto piccolo ( $\leq 0.10$ ), per cui il più delle volte  $y_t = \phi(L)y_t + \epsilon_t$ , mentre solo occasionalmente vale che  $y_t = \zeta_t$ .

Diverse specificazioni per  $\zeta_t$  sono possibili:

- $\zeta_t = \phi(L)y_t + \zeta_t$  porta ad una specificazione *outliers* di tipo additivo (AO, *Additive Outliers*),

$$y_t = \phi(L)y_t + \zeta_t\delta_t + \epsilon_t \quad (4.6)$$

dove solamente l'osservazione corrente è affetta da un'anomalia nell'andamento del suo livello.

- $\zeta_t = \phi(L)y_t + \frac{\zeta_t}{\phi(L)}$  specifica una relazione di tipo *outliers* innovativo (IO, *Innovative Outliers*),

$$y_t = \phi(L)y_t + \frac{\zeta_t}{\phi(L)}\delta_t + \epsilon_t \quad (4.7)$$

dove uno *shock* intercorso al tempo  $t$  si ripercuote anche sulle osservazioni future secondo la stessa dinamica del processo *core*. Questo significa che la serie ritorna al suo andamento naturale molto lentamente.

- $\zeta_t = \phi(L)y_t + \frac{\zeta_t}{(1-\xi L)}$  con  $|\xi| < 1$  porta ad un modello *outliers* con cambiamenti temporanei (TCO, *Temporary Changes Outliers*),

$$y_t = \phi(L)y_t + \frac{\zeta_t}{(1-\xi L)}\delta_t + \epsilon_t \quad (4.8)$$

il quale produce un effetto simile a quello precedente, nel senso che si riflette sulle osservazioni future.

- Se nel modello TCO assumiamo  $\xi = 1$ ,

$$y_t = \phi(L)y_t + \frac{\zeta_t}{1-L}\delta_t + \epsilon_t \quad (4.9)$$

otteniamo il modello *outliers* con cambiamento nei livelli (LSO, *Level Shifts Outliers*).



Specificazioni alternative per  $\zeta_t$  possono essere considerate; tuttavia, in un contesto di analisi di serie temporali, due specificazioni sono comunemente impiegate (van Dijk, Franses e Lucas, 1999a): l'*outlier* di tipo additivo (AO) e l'*outlier* di tipo innovativo (IO). Con riferimento a queste due tipologie, le stime dei parametri di un modello AR(p) ottenute mediante minimi quadrati ordinari (OLS) sono consistenti ma inefficienti, nel caso IO, mentre il caso AO è addirittura peggiore poiché non garantisce neppure la consistenza (van Dijk *et al.*, 1999a).

Per ovviare alle deficienze causate dallo stimatore OLS, i parametri del modello AR(p) possono essere stimati impiegando delle tecniche robuste di stima, le quali assegnano un peso inferiore alle osservazioni considerate anomale. Ad esempio, lo stimatore robusto per un modello AR(p),  $y_t = \phi' \mathbf{x}_t + \epsilon_t$  con  $\mathbf{x}_t = (1, y_{t-1}, \dots, y_{t-p})'$ , può essere definito (van Dijk *et al.*, 1999a) come soluzione della condizione del primo ordine

$$\sum_{t=1}^T w_r \left( \frac{y_t - \phi' \mathbf{x}_t}{\sigma_\epsilon w_x(\mathbf{x}_t)} \right) \mathbf{x}_t (y_t - \phi' \mathbf{x}_t) = 0 \quad (4.10)$$

dove  $\sigma_\epsilon$  rappresenta una misura di scala dei residui  $\epsilon_t \equiv y_t - \phi' \mathbf{x}_t$ , mentre  $w_r(\cdot)$  e  $w_x(\cdot)$  sono funzioni di ponderazione limitate tra 0 e 1. Dalla (4.10) si vede che lo stimatore robusto non è altro che un tipo particolare di stimatore dei minimi quadrati ponderati (WLS), dove il peso della  $t$ -esima osservazione è dato dal valore di  $w_r(\cdot)$ . Le funzioni  $w_x(\cdot)$  e  $w_r(\cdot)$  dovrebbero essere scelte in modo tale da assegnare un valore molto piccolo (o addirittura nullo) alle osservazioni contaminate.  $w_r(\cdot)$  è di solito specificata in termini di una funzione  $\psi(r_t)$  tale che si abbia  $w_r(r_t) = \frac{\psi(r_t)}{r_t}$  per  $r_t \neq 0$  e  $w_r(0) = 1$  e dove  $r_t \equiv \left( \frac{y_t - \phi' \mathbf{x}_t}{\sigma_\epsilon w_x(\mathbf{x}_t)} \right)$ . In particolare, se i regressori  $\mathbf{x}_t$  o la quantità  $\left( \frac{y_t - \phi' \mathbf{x}_t}{\sigma_\epsilon} \right)$  sono grandi in valore assoluto,<sup>4</sup> alle funzioni  $w_x(\cdot)$  e  $\psi(\cdot)$  dovrebbe essere assegnato un peso tale da far sì che l'osservazione contaminata riceva un valore relativamente piccolo.

#### 4.2.2 Robustificazione del test di linearità

van Dijk *et al.* (1999a) mostrano che i test tradizionali di linearità contro l'alternativa di tipo STAR tendono a rifiutare troppo spesso l'ipotesi nulla, anche asintoticamente. La ragione di un simile comportamento è dovuta alla distorsione che provoca lo stimatore OLS in presenza di *outliers*. Come soluzione, gli autori suggeriscono di utilizzare delle tecniche robuste di stima, analogamente a quanto presentato nella sezione precedente. Gli stimatori robusti consentono di ottenere stime più accurate dei parametri assegnando un peso inferiore alle osservazioni contaminate.

In analogia a quanto visto per il caso dell'eteroschedasticità, il calcolo della versione robusta alla presenza di *outliers* della statistica test  $LM_3$  è basato sui seguenti passi:

1. Stimare il modello lineare sotto l'ipotesi nulla utilizzando stimatori robusti. Ri-

<sup>4</sup>Per una discussione completa sulle diverse specificazioni che può assumere la funzione  $\psi(\cdot)$  si veda Hampel, Ronchetti, Rousseeuw e Stahel (1986).

cavare le quantità  $\hat{w}_r(\hat{r}_t)$  e  $\hat{w}_x(\mathbf{x}_t)$  di ponderazione<sup>5</sup> e calcolare i residui pesati  $\hat{\psi}(\hat{r}_t) = \hat{w}_r(\hat{r}_t) \hat{r}_t$ .

2. Regredire  $\hat{\psi}(\hat{r}_t)$  sui regressori ponderati  $\hat{w}_x(\mathbf{x}_t) \cdot (\mathbf{x}'_t, \mathbf{x}'_t y_{t-d}, \mathbf{x}'_t y_{t-d}^2, \mathbf{x}'_t y_{t-d}^3)'$ , dove il simbolo  $\cdot$  indica la moltiplicazione elemento per elemento.
3. La quantità  $TR^2$ , dove  $R^2$  rappresenta l'indice di determinazione lineare della regressione di cui al punto 2, rappresenta la statistica test  $LM_3$ . Sotto l'ipotesi nulla di linearità, la statistica  $LM_3$  si distribuisce asintoticamente come una  $\chi^2$  con  $3p$  gradi di libertà.<sup>6</sup>

Naturalmente, la procedura di cui sopra si adatta facilmente al calcolo delle versioni robuste delle altre statistiche test trattate nella Sezione 3.5.

Risultati ottenuti tramite simulazione da Escribano, Franses e van Dijk (1998) e da van Dijk *et al.* (1999a) mostrano che la statistica test così ricavata possiede una buona potenza in piccoli campioni. Tuttavia, come è naturale attendersi, la sua potenza risulta inferiore rispetto alla statistica test tradizionale nei casi in cui la serie considerata non sia contaminata dalla presenza di valori anomali. È comunque auspicabile l'impiego di questa versione della statistica test poiché presenta risultati soddisfacenti anche in assenza di *outliers* (van Dijk, 1999).

### 4.3 Un cambiamento di notazione

La rappresentazione formale del modello STAR che abbiamo utilizzato nelle pagine precedenti ha assunto la forma

$$y_t = \pi_0 + \pi' \mathbf{w}_t + (\theta_0 + \theta' \mathbf{w}_t) F(s_t; \gamma, c) + \epsilon_t \quad t = 1, \dots, T \quad (4.11)$$

dove  $\pi = (\pi_1, \dots, \pi_p)'$ ,  $\theta = (\theta_1, \dots, \theta_p)'$ ,  $\mathbf{w}_t = (y_{t-1}, \dots, y_{t-p})'$ , con  $\epsilon_t \sim iid N(0, \sigma_\epsilon^2)$  e dove la funzione  $F(\cdot)$  assumeva le tradizionali formulazioni logistica ed esponenziale (si veda la Sezione 4.1).

Nella parte restante di questo Capitolo ci riferiremo ad una diversa notazione che solo di recente è stata introdotta, ma che non comporta alcuna variazione nei risultati precedentemente descritti. La sua introduzione è giustificata da una migliore interpretabilità del modello, prevalentemente in un contesto di regimi multipli e di parametri che variano nel tempo poiché viene messo bene in evidenza l'effetto netto che il cambiamento di regime provoca sulla variabile oggetto di indagine. La (4.11) può dunque essere riscritta come

$$y_t = \phi'_1 \mathbf{x}_t (1 - F(s_t; \gamma, c)) + \phi'_2 \mathbf{x}_t F(s_t; \gamma, c) + \epsilon_t \quad t = 1, \dots, T \quad (4.12)$$

<sup>5</sup>La scelta del valore da attribuire alle quantità  $\hat{w}_r(\hat{r}_t)$  e  $\hat{w}_x(\mathbf{x}_t)$  avviene endogenamente a partire dalla stima robusta del modello sotto l'ipotesi nulla. Le osservazioni con valore entro l'intervallo  $(-c_1, c_1)$  riceveranno un peso pari a 1, mentre quelle più grandi di un certo valore  $c_2$  in valore assoluto assumeranno valore 0. Pesi intermedi verranno attribuiti alle osservazioni rimanenti. Per una descrizione dettagliata sulla determinazione di tali quantità e sul valore numerico che possono prendere le costanti  $c_1$  e  $c_2$ , si veda van Dijk *et al.* (1999a) e van Dijk (1999).

<sup>6</sup>La versione  $F$  della stessa statistica test può essere ricavata senza difficoltà.

dove  $\phi_i = (\phi_{i,0}, \phi_{i,1}, \dots, \phi_{i,p})'$ ,  $i = 1, 2$  e  $\mathbf{x}_t = (1, \mathbf{w}_t')' = (1, y_{t-1}, \dots, y_{t-p})'$ . Al solito,  $F(\cdot)$  rappresenta la funzione di transizione, mentre  $\epsilon_t \sim iid N(0, \sigma_\epsilon^2)$ .

L'interpretazione della (4.12) è che, ad ogni istante  $t$ , l'evoluzione di  $y_t$  è determinata da una media ponderata tra due diversi modelli lineari autoregressivi. I pesi assegnati ai due modelli dipendono dal valore assunto dalla funzione di transizione che, a sua volta, dipende dalla variabile di transizione  $s_t$ . La relazione che c'è tra le espressioni (4.12) e (4.11),

$$y_t = \phi_1' \mathbf{x}_t + (\phi_2 - \phi_1)' \mathbf{x}_t F(s_t; \gamma, c) + \epsilon_t \quad t = 1, \dots, T \quad (4.13)$$

descrive bene l'effetto prodotto da incrementi della variabile di transizione che si riflettono sul cambiamento graduale dei parametri autoregressivi da  $\phi_1$  a  $\phi_2$ . L'importanza di una simile notazione sarà particolarmente evidente nel prosieguo di questo Capitolo.

Seguendo Jansen e Teräsvirta (1996) è possibile anche ricavare una generalizzazione della funzione di transizione nel modo seguente:

$$F(s_t; \gamma, \mathbf{c}) = \left( 1 + \exp \left\{ -\gamma \prod_{k=1}^n (s_t - c_k) \right\} \right)^{-1}, \quad c_1 \leq c_2 \leq \dots \leq c_n, \quad \gamma > 0 \quad (4.14)$$

Per  $n = 1$ , la (4.14) porta al modello LSTAR tradizionale, mentre per  $n = 2$  e  $c_1 = c_2 = c$  si ottiene una buona approssimazione della funzione esponenziale, e quindi del relativo modello ESTAR. La letteratura recente (Jansen e Teräsvirta, 1996) considera anche la possibilità che  $n = 2$ , ma  $c_1 \neq c_2$ . Questo porta alla funzione logistica di tipo quadratico

$$F(s_t; \gamma, \mathbf{c}) = \frac{1}{1 + \exp \{ -\gamma (s_t - c_1)(s_t - c_2) \}}, \quad c_1 \leq c_2, \quad \gamma > 0 \quad (4.15)$$

dove  $\mathbf{c} = (c_1, c_2)'$ . La peculiarità di questa formulazione è dovuta al fatto che, per  $\gamma \rightarrow \infty$ , la funzione tende ad 1 per  $s_t < c_1$  e  $s_t > c_2$ , mentre per  $c_1 \leq s_t \leq c_2$  tende a 0. Quindi, il modello STAR con questa particolare funzione di transizione comprende il modello (SE)TAR a tre regimi come caso particolare.

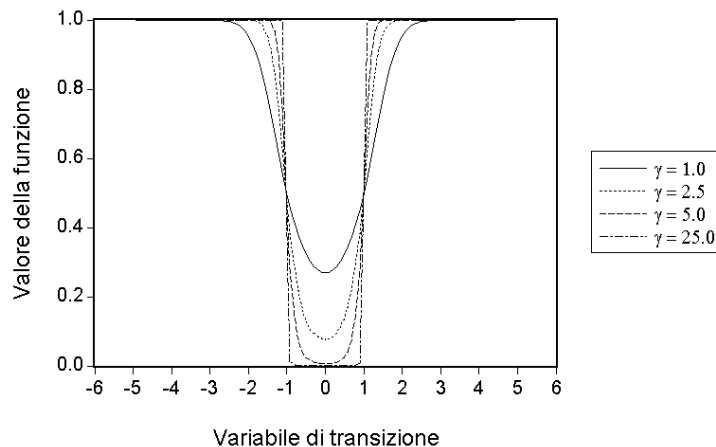
La Figura 4.1 sintetizza l'andamento della funzione di transizione logistica di tipo quadratico per diversi valori di  $\gamma$ . Si noti che per piccoli valori di  $\gamma$ , il punto di minimo della funzione, ottenuto in corrispondenza del valore di  $s_t = \frac{c_1 + c_2}{2}$ , non è lo zero. Questo dovrebbe essere tenuto bene in mente quando si interpretano i risultati delle stime ricavate da questo particolare modello STAR.

Sviluppi recenti (si veda Lundbergh e Teräsvirta, 2003) hanno proposto ulteriori generalizzazioni della funzione di transizione del tipo

$$F(s_t; \gamma, \delta, c) = \frac{\delta}{1 + \exp \{ -\gamma (s_t - c) \}}, \quad \gamma > 0, \delta > 0 \quad (4.16)$$

dove, al solito,  $s_t$  indica la variabile di transizione,  $\gamma$  il grado di liscio della funzione di transizione,  $c$  la soglia di passaggio da un regime all'altro, mentre l'ulteriore parametro  $\delta$  rappresenta un parametro di asimmetria. Le restrizioni  $\gamma > 0$  e  $\delta > 0$  sono condizioni di identificazione del modello. Questo tipo di funzione di transizione,

Figura 4.1: Funzione di transizione logistica quadratica per diversi valori di  $\gamma$  e  $c_1 = -1$  e  $c_2 = 1$ .



originariamente introdotta da Nelder (1961) e poi riutilizzata da Sollis, Leybourne e Newbold (1999), è stata definita dagli stessi autori “funzione logistica generalizzata”. Infatti, per  $\delta = 1$  si ottiene la funzione logistica solitamente impiegata nei modelli STAR (cfr. equazione (3.4) del Capitolo 3). Il parametro di asimmetria  $\delta$  si rivela essenziale nello studio delle *Target Zone* come in Lundbergh e Teräsvirta (2003), dove le oscillazioni del tasso di cambio sono vincolate a variare entro i limiti imposti dalla *Target Zone*.

#### 4.4 Modelli STAR a regimi multipli

Il modello a transizione liscia descritto nelle pagine precedenti distingue solamente tra due possibili stati dell’economia. Sebbene questa eventualità possa essere considerata ragionevole in molte applicazioni concrete, a volte può essere necessario considerare un modello più elaborato per rappresentare fenomeni con dinamiche più flessibili. In quest’ottica si collocano i modelli STAR a più regimi, i quali trovano una naturale applicazione nella descrizione del comportamento di variabili macroeconomiche all’interno delle varie fasi del ciclo economico (van Dijk, 1999).

La letteratura tradizionale sul ciclo economico è solita distinguere tra due soli stati: contrazione ed espansione. Tuttavia, recenti evidenze mostrano che un simile modo di operare si rivela il più delle volte troppo restrittivo. Ad esempio, Sichel (1993) e Ramsey e Rothman (1996) discutono concetti quali la “profondità” (*deepness*), la “ripidezza” (*steepness*) e la “affilatezza” (*sharpness*) in merito a differenti aspetti che determinano l’asimmetria tra le varie fasi del ciclo economico. Quindi, almeno intuitivamente, se una variabile macroeconomica esibisce simultaneamente diverse tipologie di asimmetria, la sola distinzione tra espansione e contrazione può non essere sufficiente a caratterizzare completamente il suo comportamento.

van Dijk e Franses (1999) e van Dijk (1999) modellano queste eventualità mediante l'introduzione di regimi aggiuntivi al modello STAR tradizionale. Così a partire dalla (4.13), un modello STAR a tre regimi si può ottenere semplicemente aggiungendo una seconda componente non lineare:

$$y_t = \phi'_1 \mathbf{x}_t + (\phi_2 - \phi_1)' \mathbf{x}_t F_1(s_t; \gamma_1, c_1) + (\phi_3 - \phi_2)' \mathbf{x}_t F_2(s_t; \gamma_2, c_2) + \epsilon_t \quad (4.17)$$

Se  $F_i(\cdot)$ ,  $i = 1, 2$  sono funzioni logistiche<sup>7</sup> e se, senza perdita di generalità, assumiamo che  $c_1 < c_2$ , all'aumentare di  $s_t$  i parametri autoregressivi del modello (4.17) variano gradualmente da  $\phi_1$  a  $\phi_3$ , passando per  $\phi_2$ . La (4.17) è facilmente generalizzabile al caso di  $m$  regimi:

$$y_t = \phi'_1 \mathbf{x}_t + (\phi_2 - \phi_1)' \mathbf{x}_t F_1(s_t; \gamma_1, c_1) + (\phi_3 - \phi_2)' \mathbf{x}_t F_2(s_t; \gamma_2, c_2) + \dots \\ \dots + (\phi_m - \phi_{m-1})' \mathbf{x}_t F_{m-1}(s_t; \gamma_{m-1}, c_{m-1}) + \epsilon_t \quad (4.18)$$

È interessante notare come il modello additivo espresso dalla (4.17) sia molto simile a quello utilizzato da Eitrheim e Teräsvirta (1996) in fase di verifica della presenza di non-linearità residue su un modello a due regimi, discusso nella Sezione 3.7.1. L'unica differenza, a parte il cambiamento di notazione, è che nel modello di Eitrheim e Teräsvirta (1996) le due funzioni  $F_i(\cdot)$ ,  $i = 1, 2$  non sono vincolate a dipendere dalla medesima variabile di transizione. Questo implica che  $F_1(s_{1t}; \gamma_1, c_1)$  e  $F_2(s_{2t}; \gamma_2, c_2)$  variano più o meno indipendentemente l'una dall'altra, dando luogo così a quattro regimi distinti, ognuno dei quali corrispondente a una particolare combinazione di valori estremi della funzione di transizione. In realtà, dalla (3.78) si vede anche che l'insieme delle parametrizzazioni lineari che compongono il modello STAR consistono di tre soli elementi, per cui le dinamiche tra i quattro regimi non sono completamente indipendenti tra loro (van Dijk, 1999).

Un modo alternativo per estendere il modello STAR ai casi di regimi multipli determinati da differenti variabili di transizione, è basato sulla notazione utilizzata nella (4.12). Ad esempio, un modello a quattro regimi si può ricavare annidando tra loro due diversi modelli STAR a due regimi:

$$y_t = [\phi'_1 \mathbf{x}_t (1 - F_1(s_{1t}; \gamma_1, c_1)) + \phi'_2 \mathbf{x}_t F_1(s_{1t}; \gamma_1, c_1)] [1 - F_2(s_{2t}; \gamma_2, c_2)] + \\ + [\phi'_3 \mathbf{x}_t (1 - F_1(s_{1t}; \gamma_1, c_1)) + \phi'_4 \mathbf{x}_t F_1(s_{1t}; \gamma_1, c_1)] F_2(s_{2t}; \gamma_2, c_2) + \epsilon_t \quad (4.19)$$

Si noti che adesso i quattro stati del modello, associati alle diverse combinazioni di  $F_1(\cdot) = 0, 1$  e  $F_2(\cdot) = 0, 1$ , sono indipendenti. Per valori di  $F_1(\cdot)$  e  $F_2(\cdot)$  compresi tra zero e uno, la relazione effettiva tra  $y_t$  e i suoi valori ritardati è espressa da una combinazione lineare dei quattro modelli. L'espressione (4.19), generalizzabile al caso di  $2^m$  regimi con  $m > 2$  grazie al principio dell'annidamento, prende il nome di modello STAR a regimi multipli (MR-STAR, van Dijk e Franses, 1999 e van Dijk, 1999). Un

<sup>7</sup>Come è noto dalla Sezione 3.1.1, le funzioni logistiche hanno la caratteristica di descrivere adeguatamente comportamenti asimmetrici della variabile oggetto di analisi. Pertanto, nell'ottica in cui sono stati qui introdotti i modelli STAR a più regimi, risulta naturale riferirsi a una tale parametrizzazione della funzione di transizione. Formulazioni alternative, quali quella esponenziale, possono essere trattate senza difficoltà (van Dijk, 1999).

modello a tre regimi può essere ottenuto dalla (4.19) imponendo appropriate restrizioni ai parametri dei modelli autoregressivi che prevalgono nei diversi stati.

Come già accennato, il modello (4.19) è appropriato solamente quando  $s_{1t} \neq s_{2t}$ . Se  $s_{1t} = s_{2t} = s_t$  e se  $c_1 < c_2$ ,  $F_1(\cdot)$  varia da zero a uno prima di  $F_2(\cdot)$  per valori crescenti di  $s_t$ . Di conseguenza, il prodotto  $(1 - F_1(\cdot)) F_2(\cdot)$  sarà zero per quasi tutti i valori di  $s_t$  e lo sarà a maggior ragione se  $\gamma_1$  e  $\gamma_2$  assumono valori elevati. Questa eventualità può essere esclusa imponendo la restrizione  $\phi_3 = 0$  poiché, in tal caso,  $F_1(\cdot) F_2(\cdot) \approx F_2(\cdot)$  e il modello risultante può essere scritto in forma additiva come espresso dalla (4.17). Considerazioni analoghe valgono per generalizzazioni della (4.19).

Per ulteriori considerazioni sul modello e per un'analisi della varietà di casi particolari ricavabili dal più generale MR-STAR, si rimanda alla letteratura sull'argomento.<sup>8</sup> In ciò che segue ci soffermeremo brevemente sulla specificazione empirica del modello.

#### 4.4.1 Specificazione dei modelli MR-STAR

Un approccio che va dallo specifico al generale, nel quale il numero dei regimi è incrementato gradualmente a seguito di test sulla desiderabilità o meno di stati aggiuntivi, si rivela la scelta più appropriata nella specificazione di un modello MR-STAR. In particolare, un modello a due regimi dovrebbe essere testato contro la presenza di regimi aggiuntivi. Come punto di partenza si potrebbe utilizzare il test per nonlinearità residue di Eitrheim e Teräsvirta (1996), sebbene si sia messo in evidenza come una tale formulazione ristretta sia giustificata solo quando  $s_{1t} = s_{2t} = s_t$ . Questo, infatti, assicura l'indipendenza fra i tre regimi previsti da modello (cfr. equazione (4.17)). Più in generale, van Dijk (1999) e van Dijk e Franses (1999) derivano un test per la verifica della presenza di regimi multipli in forma annidata come espresso dalla (4.19). A tale scopo la (4.19) viene riscritta in una forma più conveniente,

$$y_t = \phi_1^* \mathbf{x}_t + \phi_2^* \mathbf{x}_t F_1(s_{1t}; \gamma_1, c_1) + \phi_3^* \mathbf{x}_t F_2(s_{2t}; \gamma_2, c_2) + \phi_4^* \mathbf{x}_t F_1(s_{1t}; \gamma_1, c_1) F_2(s_{2t}; \gamma_2, c_2) + \varepsilon_t \quad (4.20)$$

dove  $\phi_1^* = \phi_1$ ,  $\phi_2^* = \phi_2 - \phi_1$ ,  $\phi_3^* = \phi_3 - \phi_1$  e  $\phi_4^* = \phi_1 - \phi_2 - \phi_3 + \phi_4$ . L'ipotesi nulla da sottoporre a test riguarda l'aggiunta dei regimi determinati dalla funzione di transizione  $F_2(\cdot)$ . Il modello (4.20) soffre degli stessi problemi di identificazione descritti nella Sezione 3.5, per cui si rivela essenziale l'ulteriore ricorso all'idea dell'approssimazione di Luukkonen, Saikkonen e Teräsvirta (1988b). Restringendo l'attenzione al caso logistico<sup>9</sup>, dall'approssimazione di Taylor al terzo ordine della funzione  $F_2(\cdot)$  nell'intorno del punto  $\gamma_2 = 0$  si ricava il modello ausiliario

$$y_t = \theta_1' \mathbf{x}_t + \theta_2' \mathbf{x}_t F_1(s_{1t}; \gamma_1, c_1) + \beta_1' \mathbf{w}_t s_{2t} + \beta_2' \mathbf{w}_t s_{2t}^2 + \beta_3' \mathbf{w}_t s_{2t}^3 + (\beta_4' \mathbf{w}_t s_{2t} + \beta_5' \mathbf{w}_t s_{2t}^2 + \beta_6' \mathbf{w}_t s_{2t}^3) F_1(s_{1t}; \gamma_1, c_1) + u_t \quad (4.21)$$

<sup>8</sup>Si veda van Dijk (1999), van Dijk e Franses (1999) e van Dijk *et al.* (2000).

<sup>9</sup>Il caso esponenziale può infatti essere facilmente ricavato adattando a questo contesto le considerazioni già svolte a suo tempo (si veda la Sezione 3.5.2).

dove  $u_t = \epsilon_t + R_3(s_{2t}; \gamma_2, c_2)$ ,  $\mathbf{x}_t = (1, \mathbf{w}'_t)'$  e  $\mathbf{w}_t = (y_{t-1}, \dots, y_{t-p})'$ . Nel passaggio dalla (4.20) alla (4.21) si è implicitamente assunto che  $s_{2t}$  fosse un elemento di  $\mathbf{w}_t$ . In caso contrario, i regressori  $\beta_{i,0}s_{2t}^i$ ,  $i = 1, 2, 3$  e  $\beta_{i,0}s_{2t}^{i-3}F_1(s_{1t}; \gamma_1, c_1)$ ,  $i = 4, 5, 6$  dovrebbero essere aggiunti alla (4.21).

L'ipotesi nulla da sottoporre a test, per il caso in cui  $s_{2t}$  è un elemento di  $\mathbf{w}_t$ , può essere formulata come

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0, \quad (4.22)$$

mentre la corrispondente statistica test può essere calcolata sulla base dei seguenti passi:

1. Stimare il modello STAR a due regimi, calcolare i residui  $\hat{\epsilon}_t$  e la relativa somma dei quadrati  $SSR_0 = \sum_{t=1}^T \hat{\epsilon}_t^2$ .
2. Regredire  $\hat{\epsilon}_t$  su  $\mathbf{x}_t$ ,  $\mathbf{x}_t \hat{F}_1(s_{1t})$ ,  $\hat{\theta}'_2 \mathbf{x}_t \hat{F}_{\gamma_1}(s_{1t})$ ,  $\hat{\theta}'_2 \mathbf{x}_t \hat{F}_{c_1}(s_{1t})$  e i regressori ausiliari  $\mathbf{w}_t s_{2t}^i$  e  $\mathbf{w}_t \hat{F}_1(s_{1t}) s_{2t}^i$ ,  $i = 1, 2, 3$ , e calcolare la quantità  $SSR_1 = \sum_{t=1}^T \hat{u}_t^2$ . Le espressioni  $\hat{F}_1(s_{1t})$ ,  $\hat{F}_{\gamma_1}(s_{1t})$  e  $\hat{F}_{c_1}(s_{1t})$  indicano rispettivamente le quantità  $F_1(s_{1t}; \hat{\gamma}_1, \hat{c}_1)$ ,  $\frac{\partial F_1(s_{1t}; \hat{\gamma}_1, \hat{c}_1)}{\partial \gamma_1}$  e  $\frac{\partial F_1(s_{1t}; \hat{\gamma}_1, \hat{c}_1)}{\partial c_1}$ , mentre  $\hat{\theta}_2$  può essere stimato consistentemente sotto l'ipotesi nulla come  $\hat{\theta}_2 = \hat{\phi}_2 - \hat{\phi}_1$ .
3. Calcolare la statistica test<sup>10</sup>

$$LM_{EMR,3} = \frac{(SSR_0 - SSR_1)/6p}{SSR_1/(T - 6p - 2(p+1))} \sim F_{6p, T-6p-2(p+1)} \quad (4.23)$$

$$LM_{EMR,3} = T \frac{(SSR_0 - SSR_1)}{SSR_0} \sim \chi_{6p}^2 \quad (4.24)$$

Al solito, la versione  $F$  della statistica test è preferibile in un contesto di piccoli campioni.

Un inconveniente che si può presentare è che i residui  $\hat{\epsilon}_t$  ricavati dal modello a due regimi non siano esattamente ortogonali alla matrice dei gradienti per il semplice motivo che le stime dei parametri sono state ottenute tramite algoritmi numerici di ottimizzazione. Eitrheim e Teräsvirta (1996), van Dijk e Franses (1999) e van Dijk (1999) suggeriscono di aggiungere alla procedura di test il passo seguente:

- Regredire  $\hat{\epsilon}_t$  su  $\mathbf{x}_t$  e  $\mathbf{x}_t \hat{F}_1(s_{1t})$  (ed anche  $\hat{\theta}'_2 \mathbf{x}_t \hat{F}_{\gamma_1}(s_{1t})$  e  $\hat{\theta}'_2 \mathbf{x}_t \hat{F}_{c_1}(s_{1t})$  se questi termini non sono stati esclusi), calcolare i residui  $\tilde{\epsilon}_t$  e la quantità  $SSR_0 = \sum_{t=1}^T \tilde{\epsilon}_t^2$ .

I residui  $\tilde{\epsilon}_t$  saranno poi utilizzati nei passi 2 e 3 in luogo di  $\hat{\epsilon}_t$ .

Se l'ipotesi di un modello a due regimi viene rifiutata in favore di un modello a quattro regimi del tipo di quello espresso dalla (4.19), il passo successivo è rappresentato dalla stima del modello alternativo tramite massima verosimiglianza (ML) o minimi

<sup>10</sup>Il pedice EMR indica che si sottopone a test un modello STAR a regimi multipli di tipo annidato (*encapsulated*), mentre il 3 indica che ci riferiamo alla statistica test  $LM_3$ , di cui alla Sezione 3.5.1.

quadrati non-lineari (NLS). Merita qui segnalare che la ricerca di adeguati valori iniziali si rivela in tal caso ancora più importante di quanto non lo fosse già, dal momento che le dimensioni del modello crescono molto rapidamente. Diventano quindi fondamentali le tecniche discusse a questo proposito nella Sezione 3.6.1, le quali possono essere facilmente estese ai casi di regimi multipli. La procedura di specificazione si conclude poi con le solite fasi di valutazione e utilizzo del modello stimato.

## 4.5 Modelli STAR con parametri variabili nel tempo

La non-linearità è soltanto una tra le caratteristiche che può esibire una serie temporale. Un altro importante aspetto, che contraddistingue principalmente serie di elevata ampiezza campionaria, è rappresentato dall'instabilità strutturale. Nonostante l'evidenza empirica confermi che sia la non-linearità che l'instabilità strutturale rappresentano dinamiche rilevanti dell'evoluzione di molte serie macroeconomiche, l'analisi dei due fenomeni è avvenuta prevalentemente in modo isolato.

Tra le motivazioni che hanno spinto in questa direzione vi è senz'altro la constatazione che le due caratteristiche possono essere considerate come alternative in competizione quando si vuole sottoporre a test la linearità di un qualche fenomeno economico/finanziario. Per contro, è possibile che una stessa serie presenti simultaneamente caratteristiche di non-linearità e di instabilità strutturale.

Lundbergh, Teräsvirta e van Dijk (2000) analizzano questa evenienza mediante l'introduzione dei modelli STAR con parametri che variano nel tempo (TV-STAR). Come già accennato nella Sezione 3.7.2, si tratta di una classe di modelli che permette di descrivere simultaneamente sia la non-linearità (di tipo *smooth transition*), sia la variabilità dei parametri nel corso del tempo. Un simile modello si ricava dall'espressione (4.19) assumendo  $s_{1t} = y_{t-d}$ , per  $d > 0$  e  $s_{2t} = t$ . Il risultante TV-STAR rappresenta dunque un caso particolare del più generale MR-STAR:

$$y_t = [\phi'_1 \mathbf{x}_t (1 - F_1(y_{t-d}; \gamma_1, c_1)) + \phi'_2 \mathbf{x}_t F_1(y_{t-d}; \gamma_1, c_1)] [1 - F_2(t; \gamma_2, c_2)] + [\phi'_3 \mathbf{x}_t (1 - F_1(y_{t-d}; \gamma_1, c_1)) + \phi'_4 \mathbf{x}_t F_1(y_{t-d}; \gamma_1, c_1)] F_2(t; \gamma_2, c_2) + \epsilon_t \quad (4.25)$$

Si vede dalla (4.25) che  $y_t$  segue un'evoluzione di tipo STAR ad ogni istante  $t$ , accompagnata da un graduale cambiamento nel valore dei parametri autoregressivi in entrambi i regimi, da  $\phi_1$  a  $\phi_3$  per  $F_1(y_{t-d}; \gamma_1, c_1) = 0$  e da  $\phi_2$  a  $\phi_4$  per  $F_1(y_{t-d}; \gamma_1, c_1) = 1$ . Ovviamente, imponendo opportune restrizioni al valore dei parametri autoregressivi nei differenti regimi si possono ottenere casi particolari, come i modelli lineari che diventano modelli STAR in corrispondenza di certi istanti temporali  $t$ , o i modelli STAR a due regimi con parametri variabili in uno solo dei due stati. Anche il modello autoregressivo con soli parametri che variano gradualmente nel tempo (TV-AR) di Lin e Teräsvirta (1994) può essere ricavato come caso particolare della (4.25); è infatti sufficiente imporre le restrizioni  $\phi_1 = \phi_2$  e  $\phi_3 = \phi_4$  in modo tale che risulti<sup>11</sup>

$$y_t = \phi'_1 \mathbf{x}_t (1 - F_2(t; \gamma_2, c_2)) + \phi'_3 \mathbf{x}_t F_2(t; \gamma_2, c_2) + \epsilon_t \quad (4.26)$$

<sup>11</sup>Fissando  $\phi_1 = \phi_3$  e  $\phi_2 = \phi_4$  si ottiene invece il modello STAR a due regimi, ampiamente discusso nel Capitolo precedente.



Come per il caso MR-STAR ci limiteremo qui a descrivere brevemente la specificazione empirica dei modelli TV-STAR. Per una rassegna completa sui casi particolari che si possono ottenere a partire dalla (4.25) si rimanda invece a van Dijk (1999) e Lundbergh *et al.* (2000).

#### 4.5.1 Specificazione dei modelli TV-STAR

L'applicazione del modello TV-STAR ai dati richiede l'impiego di una coerente strategia operativa. In primo luogo dovrebbe essere deciso se una tale specificazione si rivela appropriata o se invece sono da privilegiare modelli più semplici, ricavabili come casi particolari del modello TV-STAR.

Supponendo di aver preliminarmente risolto il problema della scelta del valore di  $d$  e dell'ordine autoregressivo  $p$ , Lundbergh *et al.* (2000) individuano due possibili strategie alternative. La prima, già impiegata da Teräsvirta (1994) nella specificazione dei modelli STAR e da van Dijk e Franses (1999) per quelli MR-STAR, si basa sulle raccomandazioni di Granger (1993): partire da un modello specifico per arrivare a un modello più generale se quello stimato non supera una determinata batteria di test. La soluzione alternativa che propongono si concentra su un approccio che va dallo specifico al generale e poi di nuovo allo specifico. Si tratta di una procedura nella quale la linearità è verificata direttamente contro l'alternativa di tipo TV-STAR (*specific-to-general*). Se la linearità viene rifiutata, dall'applicazione di un sottoinsieme di test si valuta se specificazioni più parsimoniose, come quella STAR o quella TV-AR, possano fornire una adeguata caratterizzazione della serie (*general-to-pecific*).

Al fine di analizzare dettagliatamente le due procedure, si riscriva la (4.25) nella forma

$$y_t = \alpha' \mathbf{x}_t + \beta' \mathbf{x}_t F_1(y_{t-d}; \gamma_1, c_1) + \pi' \mathbf{x}_t F_2(t; \gamma_2, c_2) + \theta' \mathbf{x}_t F_1(y_{t-d}; \gamma_1, c_1) F_2(t; \gamma_2, c_2) + \epsilon_t \quad (4.27)$$

dove  $\alpha = \phi_1$ ,  $\beta = \phi_2 - \phi_1$ ,  $\pi = \phi_3 - \phi_1$  e  $\theta = \phi_4 - \phi_2 - \phi_3 + \phi_1$ . Nella (4.27) la linearità può essere testata verificando l'ipotesi  $H_0 : \gamma_1 = \gamma_2 = 0$ . Le considerazioni di non identificabilità di certi parametri sotto l'ipotesi nulla rendono necessaria l'applicazione della procedura di approssimazione di Luukkonen *et al.* (1988b). Se assumiamo che  $F_1(\cdot)$  e  $F_2(\cdot)$  siano funzioni logistiche dalle quali si sottrae la quantità  $\frac{1}{2}$ , l'approssimazione del primo ordine porta al modello ausiliario<sup>12</sup>

$$y_t = \alpha'^* \mathbf{x}_t + \beta'^* \mathbf{w}_t y_{t-d} + \pi'^* \mathbf{x}_t t + \theta'^* \mathbf{x}_t y_{t-d} t + e_t \quad (4.28)$$

L'ipotesi di linearità si può adesso verificare sottoponendo a test  $H_0^* : \beta^* = \pi^* = \theta^* = 0$  nella (4.28) e si basa sui seguenti passi:

1. Regredire  $y_t$  su  $\mathbf{x}_t$  e calcolare la somma dei quadrati dei residui  $SSR_0$ .
2. Regredire  $y_t$  su  $\mathbf{x}_t$ ,  $\mathbf{w}_t y_{t-d}$ ,  $\mathbf{x}_t t$  e  $\mathbf{x}_t y_{t-d} t$  e calcolare la quantità  $SSR_1$ .

<sup>12</sup>Assumiamo inoltre che  $d \leq p$  in modo tale che  $y_{t-d}$  sia un elemento di  $\mathbf{w}_t = (y_{t-1}, \dots, y_{t-p})'$ .

3. Calcolare la statistica test

$$LM_{TV-STAR} = \frac{(SSR_0 - SSR_1) / (3p + 2)}{SSR_1 / (T - 4p - 3)} \sim F_{3p+2, T-4p-3} \quad (4.29)$$

$$LM_{TV-STAR} = T \frac{(SSR_0 - SSR_1)}{SSR_0} \sim \chi_{3p+2}^2 \quad (4.30)$$

I regressori ausiliari del modello (4.28) possono essere impiegati anche per testare la linearità contro specifici sotto-modelli del TV-STAR. Ad esempio, la linearità può essere accertata contro l'alternativa di tipo STAR sottoponendo a test l'ipotesi  $H_0^{**} : \beta^* = 0 \mid \pi^* = \theta^* = 0$  nella (4.28).

Dall'equazione (4.27) è anche possibile verificare l'ipotesi di costanza dei parametri sui residui di un modello STAR a due regimi (si veda la Sezione 3.7.2). La procedura di test è uguale a quella descritta a proposito della presenza di regimi multipli analizzata nella Sezione 4.4.1. L'unica differenza è rappresentata dall'espressione di  $s_{2t}$  nel modello ausiliario (4.21) che, in tal caso, è  $s_{2t} = t$ . Per il resto, con l'accortezza che  $s_{2t} = t$  non è un elemento di  $w_t$ , la procedura di test rimane inalterata.<sup>13</sup> Si noti che la teoria sulla distribuzione asintotica della statistica test non cambia anche se la variabile di transizione è rappresentata da un trend deterministico non stazionario (Lin e Teräsvirta, 1994).

## 4.5.2 Approccio “dallo specifico - al generale”

Questo tipo di strategia operativa può essere riassunto nei seguenti passi:

1. Specificare un modello lineare di tipo AR(p) per la serie  $y_t$ .
2. Testare l'ipotesi di linearità separatamente contro le alternative TV-AR ( $H_0 : \pi^* = 0 \mid \beta^* = \theta^* = 0$ ) e STAR ( $H_0 : \beta^* = 0 \mid \pi^* = \theta^* = 0$ ) a partire dal modello (4.28). Il valore appropriato di  $d$  può essere selezionato ripetendo ciascun test per diverse scelte di  $y_{t-d}$ ,  $d = 1, \dots, d_{\max}$  come variabile di transizione, e scegliendo quella per cui il p-value della statistica test risulta più piccolo (Teräsvirta, 1994).
3. Se la linearità non può essere accettata, stimare il modello sotto l'ipotesi alternativa per il quale l'ipotesi nulla è rifiutata in modo più marcato. Testare successivamente il modello stimato per la presenza di non-linearità di tipo STAR aggiuntive o per la costanza dei parametri.
4. Stimare il modello TV-STAR se i test di cui al punto 3 forniscono adeguate conferme in questa direzione. Valutare infine il modello stimato contro ulteriori errate specificazioni.

<sup>13</sup>Indicheremo con  $LM_{C,3}$  la statistica test così risultante, dove il pedice C indica che si sottopone a test la costanza (*constancy*) dei parametri in un modello STAR a due regimi, mentre il 3 si riferisce, al solito, alla statistica test  $LM_3$  di cui alla Sezione 3.5.1.

### 4.5.3 Approccio “dallo specifico - al generale - allo specifico”

La procedura appena descritta si rivela abbastanza dispendiosa in termini di tempo e può portare alla specificazione di un modello finale crucialmente alterato dalle decisioni assunte nei diversi passi che la compongono. Per ovviare a questi inconvenienti, van Dijk (1999) e Lundbergh *et al.* (2000) propongono questo secondo approccio (*specific to general to specific*):

1. Specificare un appropriato modello lineare AR(p) per la serie  $y_t$ .
2. Utilizzare la statistica  $LM_{TV-STAR}$  per testare direttamente l'ipotesi di linearità contro l'alternativa di tipo TV-STAR ( $\beta^* = \pi^* = \theta^* = 0$  nella (4.28)). Ripetendo il test per diverse scelte di  $d$  è possibile selezionare un'adeguata variabile di transizione.
3. Se la linearità viene rifiutata, verificare l'effettiva esigenza di un modello TV-STAR sottoponendo a test le ipotesi:

$$H_0^{STAR} : \beta^* = \theta^* = 0$$

$$H_0^{TV-AR} : \pi^* = \theta^* = 0$$

nel modello ausiliario (4.28). Le corrispondenti statistiche test, indicate con  $LM_{STAR}$  e  $LM_{TV-AR}$ , possiedono distribuzione asintotica  $\chi^2$  con  $2p + 1$  e  $2p + 2$  gradi di libertà. Dall'espressione dei parametri  $\beta^*$ ,  $\pi^*$  e  $\theta^*$  in funzione di quelli del modello originario (si veda van Dijk, 1999 o Lundbergh *et al.*, 2000) otteniamo che, sotto  $H_0^{STAR}$  il modello si riduce a un TV-AR, mentre sotto  $H_0^{TV-AR}$  diventa un modello STAR a due regimi. Da queste osservazioni si ricavano le seguenti regole di decisione:

- se sia  $H_0^{STAR}$  che  $H_0^{TV-AR}$  sono rifiutate, mantenere il modello TV-STAR;
- se  $H_0^{STAR}$  è rifiutata mentre  $H_0^{TV-AR}$  non lo è, selezionare un modello STAR;
- se  $H_0^{TV-AR}$  è rifiutata e  $H_0^{STAR}$  no, scegliere un modello TV-AR.

Una possibile complicazione, che non consente di scegliere nettamente tra una delle alternative, sorge quando né  $H_0^{STAR}$  né  $H_0^{TV-AR}$  sono rifiutate, mentre  $H_0^{TV-STAR}$  lo era.

4. Stimare il modello selezionato al punto 3 e valutarne le proprietà.

A differenza di quello precedente, il vantaggio principale dell'approccio *specific to general to specific* è che non richiede la stima di alcun modello non-lineare intermedio. Sono infatti sufficienti le regressioni OLS dei vari test e la stima del modello selezionato al punto 3 tramite massima verosimiglianza o minimi quadrati non-lineari. A questo proposito, relativamente alle tecniche di stima, valgono le medesime considerazioni accennate per i modelli a regimi multipli.

## **4.6 Considerazioni conclusive**

In questo capitolo sono state presentate alcune delle principali estensioni del modello STAR tradizionale. Oltre alla robustificazione del test di linearità in presenza di eteroschedasticità generica ed outliers, si sono introdotti i modelli a più regimi (MR-STAR) e quelli con parametri variabili nel tempo (TV-STAR) all'interno di una notazione alternativa. In nessun caso si tratta di una completa ed esaustiva trattazione dell'argomento, ma si è voluto semplicemente accennare a possibili estensioni del modello STAR di base, sulle quali si sta concentrando la letteratura più recente. Ulteriori generalizzazioni sono possibili ed auspicabili. Infine, alcune delle tecniche presentate saranno estese all'analisi della volatilità sui mercati finanziari nel capitolo successivo.



## Capitolo 5

# Modelli GARCH a transizione liscia

Nei Capitoli precedenti si è ipotizzato che la varianza condizionata del termine di errore sia costante nel tempo e ci siamo concentrati sulla ricerca di una adeguata rappresentazione della dinamica del livello di una generica serie  $y_t$ . Tuttavia, il più delle volte, una semplificazione del genere non trova riscontri oggettivi nelle analisi empiriche, limitandone di fatto l'attendibilità dei risultati.

La verifica dell'ipotesi di omoschedasticità diventa quindi importante per due motivi: uno economico e uno più prettamente statistico-econometrico. Dal punto di vista statistico, una varianza che cambia nel tempo ha implicazioni sulla validità e sull'efficienza dei risultati inferenziali tratti dai parametri stimati nell'equazione della media condizionata. L'interpretazione economica fa invece riferimento al fatto che molte serie economiche, ed in particolare finanziarie, reagiscono più o meno rapidamente all'insorgere di nuova informazione. La concentrazione di tali innovazioni in momenti particolari provoca reazioni a catena nel comportamento degli agenti economici, che si riflettono in un aumento della variabilità<sup>1</sup> (e quindi della varianza) del singolo fenomeno.

Il modello econometrico di riferimento utilizzato per descrivere la dinamica della varianza di una serie temporale è quello ad eteroschedasticità condizionata autoregressiva (ARCH) di Engle (1982), dove la varianza condizionata è rappresentata come funzione lineare di  $q$  ritardi delle innovazioni al quadrato. Da allora, numerose sono state le variazioni e le applicazioni del modello base (si veda la Sezione 2.2). Ad esempio, Bollerslev (1986) introduce il modello ARCH generalizzato (GARCH) estendendo quello di Engle fino ad includervi  $p$  ritardi della varianza condizionata. Con questa nuova rappresentazione il numero dei parametri da stimare può essere considerevolmente ridotto ed è sempre possibile esprimere un modello GARCH stazionario come un ARCH con un infinito numero di parametri (Bollerslev, 1986).

Engle e Bollerslev (1986) sono stati tra i primi a introdurre il concetto di non-linearità nel contesto (G)ARCH. Essi propongono un modello nel quale la dinamica

---

<sup>1</sup>O, nella terminologia dei mercati finanziari, volatilità.

della varianza condizionata cambia con l'impatto dei residui al quadrato e dove la transizione tra differenti stati di volatilità è determinata dalla funzione di distribuzione cumulata della normale (cfr. Sezione 2.2.4).

L'osservazione empirica di una correlazione negativa tra volatilità e innovazioni, la cui prima documentazione ufficiale risale a Black (1976) sotto il nome di effetto *leverage*, è alla base di una classe di modelli (G)ARCH volti proprio a catturare questo tipo di asimmetria.

I modelli che analizzeremo in questo Capitolo sono ispirati alla classe dei modelli autoregressivi a transizione liscia (STAR) di Luukkonen, Saikkonen e Teräsvirta (1988b) e riuniscono in un'unica formalizzazione, sia le caratteristiche di non-linearità, sia l'eventuale presenza di effetti *leverage*.

La struttura del Capitolo è la seguente: la Sezione 5.1 introduce i modelli e ne discute le principali caratteristiche e la Sezione 5.2 si concentra sulla verifica empirica degli effetti di tipo GARCH a transizione liscia (STGARCH). Nelle Sezioni 5.3 e 5.4 sono analizzate le fasi di stima e valutazione del modello, mentre nella Sezione 5.5 ne sono discusse brevemente alcune generalizzazioni. Il Capitolo si conclude con la presentazione di una strategia operativa e con una esemplificazione su dati reali relativi ai *Fed-Funds Rates* statunitensi, già impiegati nel Capitolo 3 per l'analisi della media condizionata tramite modelli della classe STAR.

## 5.1 I modelli STGARCH

Seguendo Engle (1982) e definendo con  $\mu_t$  la media condizionata del generico processo  $y_t$ , assumiamo che

$$y_t = \mu_t + \epsilon_t, \quad \epsilon_t = \eta_t \sqrt{h_t} \quad (5.1)$$

con

$$\eta_t \sim iid N(0, 1) \quad (5.2)$$

per cui

$$\epsilon_t | I_{t-1} \sim N(0, h_t) \quad (5.3)$$

e dove  $I_{t-1}$  indica l'informazione disponibile fino al tempo  $t-1$ , mentre  $h_t$  rappresenta una qualche dinamica della varianza condizionata di  $\epsilon_t$ . Per contro, assumeremo che la varianza non condizionata di  $\epsilon_t$  continui ad essere costante.

Specificazioni diverse per  $h_t$  sono possibili e danno luogo a differenti modelli (G)ARCH. Nel modello originario di Engle (1982),  $h_t$  dipende da  $q$  valori ritardati delle innovazioni al quadrato:

$$h_t = \omega + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_q \epsilon_{t-q}^2 \quad (5.4)$$

e prende il nome di modello ARCH( $q$ ).

Bollerslev (1986) aggiunge all'equazione (5.4)  $p$  ritardi della varianza condizionata,

$$h_t = \omega + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_q \epsilon_{t-q}^2 + \beta_1 h_{t-1} + \dots + \beta_p h_{t-p} \quad (5.5)$$

e chiama il suo modello ARCH generalizzato di ordine  $p, q$ , indicato con l'acronimo di GARCH( $p, q$ ).

Nei modelli (5.4) e (5.5),  $h_t$  è espressa come funzione lineare delle innovazioni al quadrato e dei ritardi della varianza condizionata. Estendendo l'idea dei modelli STAR all'equazione della varianza condizionata, Hagerud (1996) e González-Rivera (1998) propongono, in due lavori indipendenti, un modello (G)ARCH di tipo non-lineare dove, al solito, la non-linearità è introdotta da una funzione di transizione. Questo dà luogo a due diverse classi di modelli STGARCH: logistico (LSTGARCH) ed esponenziale (ESTGARCH). Hagerud (1996) definisce il modello STARCH come:

$$h_t = \omega + \sum_{i=1}^q \alpha_{1i} \epsilon_{t-i}^2 + \sum_{i=1}^q \alpha_{2i} F(\epsilon_{t-i}) \epsilon_{t-i}^2 \quad (5.6)$$

dove  $F(\epsilon_{t-i})$  può assumere le specificazioni logistica,

$$F(\epsilon_{t-i}) = \frac{1}{1 + \exp\{-\gamma \epsilon_{t-i}\}} - \frac{1}{2}, \quad \gamma > 0 \quad (5.7)$$

ed esponenziale,

$$F(\epsilon_{t-i}) = 1 - \exp\{-\gamma \epsilon_{t-i}^2\}, \quad \gamma > 0 \quad (5.8)$$

Lo stesso Hagerud (1996) estende poi il modello fino a comprendervi  $p$  ritardi della varianza condizionata nel modo seguente:

$$h_t = \omega + \sum_{i=1}^q \alpha_{1i} \epsilon_{t-i}^2 + \sum_{i=1}^q \alpha_{2i} F(\epsilon_{t-i}) \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \quad (5.9)$$

Un modello simile a quello STGARCH logistico di Hagerud (1996) è stato indipendentemente proposto da González-Rivera (1998):

$$h_t = \omega + \sum_{i=1}^q \alpha_{1i} \epsilon_{t-i}^2 + \left( \sum_{i=1}^q \alpha_{2i} \epsilon_{t-i}^2 \right) F(\epsilon_{t-d}; \gamma) + \sum_{j=1}^p \beta_j h_{t-j} \quad (5.10)$$

dove

$$F(\epsilon_{t-d}; \gamma) = \frac{1}{1 + \exp\{-\gamma \epsilon_{t-d}\}} - \frac{1}{2}, \quad \gamma > 0 \quad (5.11)$$

Si noti che per  $q = d = 1$ , il modello di Hagerud (1996) e quello di González-Rivera (1998) diventano uguali.<sup>2</sup>

Una generalizzazione del modello di Hagerud (1996) è stata proposta da Lundbergh e Teräsvirta (1998) e si basa sulla seguente struttura dinamica per la varianza condizionata:

$$h_t = \omega + \sum_{i=1}^q \alpha_{1i} \epsilon_{t-i}^2 + \sum_{i=1}^q \alpha_{2i} F(\epsilon_{t-i}) + \sum_{i=1}^q \alpha_{3i} F(\epsilon_{t-i}) \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \quad (5.12)$$

Il principale vantaggio di questa rappresentazione è quello di essere molto più generale di quelle precedenti, tanto da comprenderle come casi particolari.<sup>3</sup> Tuttavia, in nessuno

<sup>2</sup>Sebbene restrizioni su  $p$  non siano necessarie per assicurare l'uguaglianza tra i due modelli, è pratica comune scegliere  $p = 1$ .

<sup>3</sup>Lundbergh e Teräsvirta (1998) considerano una funzione di transizione più generale del tipo della (4.14) con  $s_t = \epsilon_{t-i}$  e dalla quale non viene sottratta la quantità  $\frac{1}{2}$ .



dei modelli che precedono è stata imposta una struttura non-lineare per  $h_{t-j}$ ,  $j = 1, \dots, p$ , poiché il modello è già molto flessibile senza una tale estensione (Lundbergh e Teräsvirta, 1998).

Anderson, Nam e Vahid (1999) estendono il modello di González-Rivera (1998) fino a comprendervi l'assunzione che la risposta della volatilità futura alle innovazioni del mercato dipenda in parte dallo stato di volatilità corrente:

$$h_t = \omega_1 + \sum_{i=1}^q \alpha_{1i} \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_{1j} h_{t-j} + \left( \omega_2 + \sum_{i=1}^q \alpha_{2i} \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_{2j} h_{t-j} \right) F(\epsilon_{t-d}; \gamma) \quad (5.13)$$

Con questa rappresentazione, che gli autori definiscono *Asymmetric Nonlinear Smooth Transition GARCH* (ANST-GARCH), è possibile determinare l'importanza del grado di lisciamento della funzione di transizione, nonché di valutare se il cambiamento di regime sia da attribuire maggiormente ai coefficienti di  $h_{t-j}$ ,  $j = 1, \dots, p$ , o alle sole innovazioni al quadrato  $\epsilon_{t-i}^2$ ,  $i, \dots, q$ . Si noti che, adesso, gli effetti della volatilità passata si riflettono su quella corrente attraverso la somma dei parametri  $\beta_{1j}$  e  $\beta_{2j}$ ,  $j = 1, \dots, p$ . Questo risulta potenzialmente importante poiché, in pratica, il valore di  $\beta$  rappresenta il più grande ed il più significativo parametro di una qualche specificazione GARCH (Anderson *et al.*, 1999).

In ciò che segue ci occuperemo delle proprietà dei diversi modelli STGARCH e faremo vedere come, per il caso più comune nella pratica,  $q = p = d = 1$ , le diverse specificazioni siano tutte comprese in quella di Anderson *et al.* (1999).

### 5.1.1 Proprietà dei modelli STGARCH

Il modello STGARCH di Hagerud (1996) presenta, nella versione logistica, dinamiche simili a quelle del modello GJR di Glosten, Jagannathan e Runkle (1993). Infatti, facendo tendere  $\gamma$  all'infinito, la funzione logistica tende alla funzione indicatrice  $\mathfrak{S}(\epsilon_{t-i} < 0)$ , che assume valore uno se la codizione è vera ( $\epsilon_{t-i} < 0$ ) e valore zero altrimenti. In entrambi i modelli, rispecchiando le considerazioni di Black (1976), la varianza condizionata segue un processo quando le innovazioni sono positive e un altro processo quando le innovazioni sono negative (effetto *leverage*). Tuttavia, il cambiamento da un processo all'altro avviene in modo istantaneo nel modello di Glosten *et al.* (1993), mentre in quello STGARCH logistico, il passaggio da un processo all'altro, ovvero da un regime di volatilità ad un altro, è governato dal grado di lisciamento della funzione di transizione.

Nel modello STGARCH esponenziale, se da un lato il passaggio da uno stato di volatilità all'altro avviene più o meno gradualmente a seconda del valore di  $\gamma$ , dall'altro, la dinamica di  $h_t$  è indipendente dal segno di  $\epsilon_{t-i}$ , ed è invece strettamente connessa all'impatto delle innovazioni al quadrato. Ciò rende un modello GARCH simmetrico (assenza di effetti *leverage*) nel quale la dinamica della varianza condizionata evolve in modo non-lineare.

Inoltre, sia il modello logistico che quello esponenziale, per  $\gamma = 0$  o per  $\alpha_{2i} = 0$ ,  $i = 1, \dots, q$ , collassano a quello lineare di Bollerslev (1986), se ci riferiamo all'equazione (5.9), o a quello di Engle (1982), se consideriamo l'equazione (5.6).

Tabella 5.1: Condizioni per la positività di  $h_t$  nei diversi modelli.

Modello	Restrizioni sui parametri
LSTGARCH Hagerud (1996) González-Rivera (1998)	$\omega > 0$ $\alpha_{1i} \geq 0 \quad i = 1, \dots, q$ $\alpha_{1i} \geq \frac{1}{2}  \alpha_{2i}  \quad i = 1, \dots, q$ $\beta_j \geq 0 \quad j = 1, \dots, p$
ESTGARCH Hagerud (1996)	$\omega > 0$ $\alpha_{1i} \geq 0 \quad i = 1, \dots, q$ $\alpha_{1i} + \alpha_{2i} \geq 0 \quad i = 1, \dots, q$ $\beta_j \geq 0 \quad j = 1, \dots, p$
LSTGARCH Lundbergh e Teräsvirta (1998)	$\omega > 0$ $\omega + \sum_{i=1}^q \alpha_{2i} > 0$ $\alpha_{1i} \geq 0 \quad i = 1, \dots, q$ $\alpha_{1i} + \alpha_{3i} \geq 0 \quad i = 1, \dots, q$ $\beta_j \geq 0 \quad j = 1, \dots, p$
LSTGARCH Anderson, Nam e Vahid (1999)	$\omega_1 > 0$ $\alpha_{1i} \geq 0 \quad i = 1, \dots, q$ $\beta_{1j} \geq 0 \quad j = 1, \dots, p$ $\omega_1 + \omega_2 > 0$ $\alpha_{1i} + \alpha_{2i} \geq 0 \quad i = 1, \dots, q$ $\beta_{1j} + \beta_{2j} \geq 0 \quad j = 1, \dots, p$

Considerazioni analoghe valgono per il modello logistico di González-Rivera (1998), il quale, come già osservato, per  $q = d = 1$ , è uguale al modello di Hagerud (1996).

L'estensione di Lundbergh e Teräsvirta (1998) costituisce un'ulteriore generalizzazione del modello. Infatti, non solo comprende il modello GJR come caso particolare, ma incorpora anche una versione non-lineare del modello GARCH quadratico (GQARCH) di Sentana (1991 e 1995) attraverso il termine  $\sum_{i=1}^q \alpha_{2i} F(\epsilon_{t-i})$ .

Infine, il modello ANST-GARCH di Anderson *et al.* (1999) può essere visto come una generalizzazione di quello di Fornari e Mele (1997), dove la volatilità futura dipende in parte dagli shock inattesi di quella corrente. L'innovazione introdotta da Anderson *et al.* (1999) riguarda il modo con cui si determina il cambiamento di volatilità fra i diversi regimi; questo corrisponde all'idea della transizione liscia come in González-Rivera (1998).<sup>4</sup>

<sup>4</sup>Per ulteriori dettagli si rimanda ad Anderson *et al.* (1999). Si noti che, sia Anderson *et al.* (1999)

Tabella 5.2: Condizioni per la stazionarietà di  $h_t$  nei diversi modelli.

Modello	Restrizioni sui parametri
LSTGARCH Hagerud (1996)	$\sum_{i=1}^q [\alpha_{1i} - \frac{1}{2}  \alpha_{2i}  + \max(\alpha_{2i}, 0)] + \sum_{j=1}^p \beta_j < 1$
ESTGARCH Hagerud (1996)	$\sum_{i=1}^q [\alpha_{1i} + \max(\alpha_{2i}, 0)]$
LSTGARCH González-Rivera (1998)	$\sum_{i=1}^q \alpha_{1i} \pm \frac{1}{2} \sum_{i=1}^q \alpha_{2i} + \sum_{j=1}^p \beta_j < 1, F(\cdot) = \pm \frac{1}{2}$ $\sum_{i=1}^q \alpha_{1i} + \sum_{j=1}^p \beta_j < 1, F(\cdot) = 0$
LSTGARCH Lundbergh e Teräsvirta (1998)	$\sum_{i=1}^q \alpha_{1i} + \sum_{i=1}^q \alpha_{2i} + \sum_{i=1}^q \alpha_{3i} + \sum_{j=1}^p \beta_j < 1$
LSTGARCH Anderson, Nam e Vahid(1999)	$\sum_{i=1}^q \alpha_{1i} + \sum_{j=1}^p \beta_{1j} + \frac{1}{2} (\sum_{i=1}^q \alpha_{2i} + \sum_{i=1}^q \beta_{2j}) < 1$

In analogia ai modelli tradizionali di Engle (1982) e Bollerslev (1986), restrizioni sui parametri dei diversi modelli STGARCH sono imposte per assicurare la positività della varianza condizionata. Tali restrizioni sono riassunte nella Tabella 5.1. La Tabella 5.2 riporta invece le condizioni, solo sufficienti, che assicurano la stazionarietà in covarianza di  $h_t$  nelle diverse formulazioni.

Consideriamo adesso il caso particolare in cui  $q = p = d = 1$  e facciamo vedere che tutte le altre rappresentazioni STGARCH si possono ottenere come casi particolari del modello ANST-GARCH di Anderson *et al.* (1999):

$$h_t = \omega_1 + \alpha_{11}\epsilon_{t-1}^2 + \beta_{11}h_{t-1} + \omega_2 F(\epsilon_{t-1}) + \alpha_{21}\epsilon_{t-1}^2 F(\epsilon_{t-1}) + \beta_{21}h_{t-1} F(\epsilon_{t-1}) \quad (5.14)$$

Imponendo la restrizione  $\omega_2 = \beta_{21} = 0$ , si ottengono i modelli di Hagerud (1996) e González-Rivera (1998), mentre se solo  $\beta_{21} = 0$ , si ricava il modello di Lundbergh e Teräsvirta (1998).<sup>5</sup> Se aumentiamo l'ordine di  $q$ ,  $p$  e  $d$  (o, al limite, del solo  $q$ ), quanto sopra non risulta più verificato con esattezza.<sup>6</sup>

che González-Rivera (1998) considerano esclusivamente una funzione di transizione logistica. In tutti quei casi in cui la presenza di effetti *leverage* non risulta verificata può essere comunque conveniente modellare la varianza condizionata tramite una parametrizzazione della classe STGARCH. In simili circostanze, una funzione di transizione esponenziale potrebbe essere più appropriata, per cui, in quel che segue, prenderemo in esame anche questa eventualità.

<sup>5</sup>Si osservi che dalla (5.14) si possono ricavare anche i modelli ARCH ( $\beta_{11} = \omega_2 = \alpha_{21} = \beta_{21} = 0$ ) e GARCH ( $\omega_2 = \alpha_{21} = \beta_{21} = 0$ ).

<sup>6</sup>Ne fa eccezione il solo caso in cui  $q = d = 1$ , con  $p$  qualsiasi.

## 5.2 Test per effetti STGARCH

Dal momento che la stima di un modello GARCH a transizione liscia risulta più dispendiosa di quella di un modello GARCH più semplice, la prima operazione da compiere nella fase di specificazione riguarda la verifica empirica di una tale ipotesi. Diventa allora fondamentale derivare una procedura che sia in grado di discriminare tra una parametrizzazione lineare della varianza condizionata contro l'alternativa di tipo non-lineare, qual'è quella a transizione liscia. Tuttavia, prima di sottoporre a test una simile ipotesi è naturale verificare la presenza o meno di eteroschedasticità condizionata di tipo ST(G)ARCH. A tale scopo, la procedura che deriveremo è molto simile a quella di Engle (1982) per l'assenza di effetti ARCH, ed è basata sull'idea dell'approssimazione di Luukkonen *et al.* (1988b). La motivazione di un tale modo di operare è sempre la stessa: il modello STARCH, in una qualunque delle formulazioni descritte, è caratterizzato dalla presenza di parametri non identificati sotto l'ipotesi nulla di omoschedasticità.

Nel caso logistico, l'approssimazione di Taylor al primo ordine della funzione di transizione valutata nell'intorno di  $\gamma = 0$ , è pari a

$$T_1(\epsilon_{t-i}; \gamma) = \frac{\gamma}{4} \epsilon_{t-i} + R_1(\epsilon_{t-i}; \gamma), \quad i = 1, \dots, q \text{ oppure } i = d \text{ fisso}, \quad (5.15)$$

mentre per il caso esponenziale avremo<sup>7</sup>

$$T_1(\epsilon_{t-i}; \gamma) = \gamma \epsilon_{t-i}^2 + R_1(\epsilon_{t-i}; \gamma), \quad i = 1, \dots, q \text{ oppure } i = d \text{ fisso}, \quad (5.16)$$

Sostituendo  $T_1(\cdot)$  al posto della funzione di transizione nei diversi modelli STARCH e riordinando i termini, si ottengono i modelli ausiliari su cui basare il test per l'assenza di eteroschedasticità condizionata di tipo *smooth transition*:

$$h_t = \alpha' \mathbf{v}_t \quad (5.17)$$

dove  $\alpha$  rappresenta il vettore dei coefficienti, mentre  $\mathbf{v}_t$  indica il vettore dei regressori da includere nel modello ausiliario. La composizione del vettore  $\mathbf{v}_t$  cambia a seconda del tipo di modello STARCH considerato e dalla forma della funzione di transizione. Nella Tabella 5.3 sono riassunte tutte le possibilità che si possono presentare.

In analogia a quanto proposto da Engle (1982), il test per l'assenza di effetti STARCH può essere derivato a partire dal principio dei moltiplicatori di Lagrange (LM) e si basa sul calcolo della statistica test  $TR^2$ , dove  $R^2$  indica l'indice di determinazione lineare della regressione

$$\hat{\epsilon}_t^2 = \alpha' \hat{\mathbf{v}}_t + \xi_t \quad (5.18)$$

<sup>7</sup>L'unica eccezione si ha nel modello di Lundbergh e Teräsvirta (1998), i quali considerano una funzione di transizione più generica che comprende quella esponenziale solo come approssimazione (si veda la Sezione 4.3 tenendo presente che, in tal caso,  $s_t = \epsilon_{t-i}$ ,  $c_k = 0 \forall k$  e viene tolta dalla funzione la quantità  $\frac{1}{2}$ ). L'approssimazione di Taylor al primo ordine sarà dunque  $T_1(\epsilon_{t-i}; \gamma) = \frac{\gamma}{4} \epsilon_{t-i}^2 + R_1(\epsilon_{t-i}; \gamma)$ . Ciò che cambia è l'espressione dei coefficienti in funzione di quelli del modello originario, ma non i regressori da includere in quello ausiliario.

mentre  $T$  rappresenta il numero di osservazioni della stessa regressione.<sup>8</sup> La risultante statistica test possiede distribuzione asintotica di tipo  $\chi^2$  con gradi di libertà pari al numero dei regressori del vettore  $\mathbf{v}_t$ , diminuito di una unità.

Tabella 5.3: Composizione del vettore  $\mathbf{v}_t$  nel test per STARCH( $q$ ).

Engle (1982)		
$\mathbf{v}_t = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2)'$		ARCH( $q$ )
Hagerud (1996)		
$\mathbf{v}_t = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, \epsilon_{t-1}^3, \dots, \epsilon_{t-q}^3)'$		Logistico
$\mathbf{v}_t = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, \epsilon_{t-1}^4, \dots, \epsilon_{t-q}^4)'$		Esponenziale
González-Rivera (1998)		
$\mathbf{v}_t = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, \epsilon_{t-1}^2 \epsilon_{t-d}, \dots, \epsilon_{t-d}^3, \dots, \epsilon_{t-q}^2 \epsilon_{t-d})'$	$q \geq d$	Logistico
$\mathbf{v}_t = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, \epsilon_{t-1}^2 \epsilon_{t-d}^2, \dots, \epsilon_{t-d}^4, \dots, \epsilon_{t-q}^2 \epsilon_{t-d}^2)'$	$q \geq d$	Esponenziale
$\mathbf{v}_t = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, \epsilon_{t-1}^2 \epsilon_{t-d}, \dots, \epsilon_{t-q}^2 \epsilon_{t-d})'$	$q < d$	Logistico
$\mathbf{v}_t = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, \epsilon_{t-1}^2 \epsilon_{t-d}^2, \dots, \epsilon_{t-q}^2 \epsilon_{t-d}^2)'$	$q < d$	Esponenziale
Lundbergh e Teräsvirta (1998)		
$\mathbf{v}_t = (1, \epsilon_{t-1}, \dots, \epsilon_{t-q}, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, \epsilon_{t-1}^3, \dots, \epsilon_{t-q}^3)'$		Logistico
$\mathbf{v}_t = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, \epsilon_{t-1}^4, \dots, \epsilon_{t-q}^4)'$		Esponenziale
Anderson, Nam e Vahid (1999)		
$\mathbf{v}_t = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, \epsilon_{t-d}, \epsilon_{t-1}^2 \epsilon_{t-d}, \dots, \epsilon_{t-d}^3, \dots, \epsilon_{t-q}^2 \epsilon_{t-d})'$	$q \geq d$	Logistico
$\mathbf{v}_t = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, \epsilon_{t-1}^2 \epsilon_{t-d}^2, \dots, \epsilon_{t-d}^4, \dots, \epsilon_{t-q}^2 \epsilon_{t-d}^2)'$	$q \geq d$	Esponenziale
$\mathbf{v}_t = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, \epsilon_{t-d}, \epsilon_{t-1}^2 \epsilon_{t-d}, \dots, \epsilon_{t-q}^2 \epsilon_{t-d})'$	$q < d$	Logistico
$\mathbf{v}_t = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, \epsilon_{t-d}^2, \epsilon_{t-1}^2 \epsilon_{t-d}^2, \dots, \epsilon_{t-q}^2 \epsilon_{t-d}^2)'$	$q < d$	Esponenziale

La composizione del vettore  $\mathbf{v}_t$  in Tabella 5.3 è stata deliberatamente separata a seconda che la funzione di transizione fosse di tipo logistico o esponenziale. Tuttavia, è anche possibile sottoporre a test l'ipotesi nulla di omoschedasticità contro l'alternativa di tipo *smooth transition* ARCH generica, includendo semplicemente nel vettore  $\mathbf{v}_t$  i regressori ausiliari di entrambe le parametrizzazioni. Così ad esempio, nel modello di Hagerud (1996), il calcolo della statistica test  $\chi^2$  è basato sull'indice di determinazione lineare della regressione ausiliaria

$$\begin{aligned} \hat{\epsilon}_t^2 = & \alpha_0 + \alpha_{11} \hat{\epsilon}_{t-1}^2 + \dots + \alpha_{1q} \hat{\epsilon}_{t-q}^2 + \alpha_{21} \hat{\epsilon}_{t-1}^3 + \dots + \alpha_{2q} \hat{\epsilon}_{t-q}^3 + \\ & + \alpha_{31} \hat{\epsilon}_{t-1}^4 + \dots + \alpha_{3q} \hat{\epsilon}_{t-q}^4 + \xi_t \end{aligned} \quad (5.19)$$

A cambiare sarà anche il numero dei gradi di libertà che, nel caso in esame, è pari a  $3q$

<sup>8</sup>La dimostrazione formale dell'appartenenza del test alla famiglia dei test dei moltiplicatori di Lagrange è riportata in Engle (1982). Hagerud (1996) la estende al caso STARCH.

contro i  $2q$  per i singoli casi logistico o esponenziale. In modo assolutamente analogo si procede per gli altri modelli STARCH.

Fino ad ora si è sottoposta a test l'ipotesi di varianza condizionata costante contro l'alternativa di tipo *smooth transition* ARCH. Nel caso in cui l'alternativa sia di tipo ARCH generalizzato a transizione liscia (STGARCH), valgono le stesse motivazioni fornite da Bollerslev (1986) per il caso GARCH: sotto l'ipotesi nulla di omoschedasticità non esiste una forma generale del test dei moltiplicatori di Lagrange per il modello GARCH( $p,q$ ). Questo è dovuto al fatto che la matrice d'informazione è singolare quando sia  $p$  che  $q$  sono maggiori di zero. Il test contro l'alternativa STGARCH può essere pertanto basato sulla stessa procedura utilizzata nel caso STARCH, includendovi un numero di ritardi abbastanza lungo. La motivazione di ciò risiede nel fatto che il modello STGARCH può essere riscritto come un modello STARCH con un infinito numero di ritardi (Hagerud, 1996).

### 5.2.1 Scelta tra GARCH e STGARCH

La possibilità di disporre di una metodologia in grado di selezionare adeguatamente il modello GARCH più appropriato<sup>9</sup> per modellare l'evoluzione della varianza condizionata di un qualche fenomeno economico o finanziario, costituisce un aspetto di non secondaria importanza. Infatti, se da un lato si è visto che modelli più complicati sono in grado di descrivere meglio alcune delle caratteristiche intrinseche alla dinamica della volatilità, dall'altro, l'utilizzo di questi modelli richiede tempi di stima e valutazione successiva molto più dispendiosi di quelli di un modello (G)ARCH più semplice. Quindi, la loro effettiva convenienza dovrebbe essere opportunamente valutata e confrontata con un modello di riferimento, quale il GARCH lineare.

Sfortunatamente, non esiste una metodologia efficiente che sia in grado di selezionare *ex-ante* il modello più appropriato. Una volta verificata la presenza di effetti ARCH, la soluzione migliore sembra essere quella di procedere alla stima di un modello GARCH(1,1). Questo costituisce in molti casi una buona parametrizzazione della varianza condizionata per la maggior parte delle serie finanziarie,<sup>10</sup> ed è poi sulla base di questo modello che si dovrebbe procedere alla specificazione, se ritenuta necessaria, di modelli più complicati.

Tuttavia, si è anche visto che nel caso dei modelli (G)ARCH a transizione liscia c'è la possibilità di verificare direttamente questa eventualità. Ciò suggerisce di confrontare il test per ARCH di Engle (1982) con quello per STARCH precedentemente derivato e scegliere come parametrizzazione più adeguata quella con il minimo p-value. Questo perché il test dovrebbe avere potenza massima nel caso in cui l'ipotesi alternativa sia stata correttamente specificata. Si tratta della stessa regola e della stessa motivazione utilizzate da Teräsvirta (1994) per selezionare un'adeguata variabile di transizione nel modello STAR e che può essere estesa anche alla scelta di  $d$  nei modelli di González-Rivera (1998) e Anderson *et al.* (1999). Inoltre, una simile strategia ci consente di selezionare la funzione di transizione più idonea a caratterizzare la dinamica non-lineare della varianza condizionata. In particolare, se i risultati del test vanno nella

---

<sup>9</sup>Sia esso lineare o non-lineare, simmetrico o asimmetrico.

<sup>10</sup>Bollerslev, Chou e Kroner (1992), Bera e Higgins (1993) e Bollerslev, Engle e Nelson (1994).

Tabella 5.4: Risultati della simulazione (dati percentuali).

		H	LT	GR	ANV
$T = 500$	$\chi^2$	53.6	60.3	46.0	48.7
	$F$	52.6	57.1	44.5	47.6
$T = 1500$	$\chi^2$	55.3	65.1	49.9	53.1
	$F$	53.7	62.8	47.8	50.6

direzione di un modello ST(G)ARCH logistico come parametrizzazione più adeguata, la conclusione che se ne può trarre è quella della presenza di effetti *leverage* all'interno dell'evoluzione di  $h_t$ . Le fasi successive di specificazione, stima e verifica del modello non possono dunque esimersi da queste considerazioni.

La potenza di questa regola di decisione è stata investigata mediante il seguente esperimento di simulazione. Si è ritenuto che sarebbe pericoloso accettare una specificazione del modello di tipo non-lineare (date le complicazioni in fase di stima e verifica successiva) quando in realtà il vero processo generatore dei dati (DGP) è lineare. A tal fine si è assunto come DGP il GARCH(1,1),

$$h_t = 0.005 + 0.1\epsilon_{t-1}^2 + 0.8h_{t-1} \quad (5.20)$$

con estrazioni per  $\eta_t$  da una distribuzione  $N(0, 1)$ . L'esperimento è stato condotto per due diverse lunghezze campionarie, rispettivamente di 500 e 1500 osservazioni, e si basa su 1000 replicazioni. Inoltre, le prime 100 osservazioni di ogni serie generata sono state scartate al fine di eliminare ogni possibile influenza della scelta di porre uguali a zero i loro valori iniziali. Ad ogni serie è stato applicato il test per ARCH di Engle (1982) e quello per STARARCH congiunto nelle diverse formulazioni presentate. Secondo la regola descritta, il p-value del test di Engle (1982) dovrebbe essere più piccolo del corrispondente p-value del test per STARARCH. La frequenza percentuale con cui si verifica una simile ipotesi è riportata nella Tabella 5.4, dove i diversi modelli STARARCH sono indicati, rispettivamente con H (Hagerud, 1996), LT (Lundbergh e Teräsvirta, 1998), GR (González-Rivera, 1998 e ANV (Anderson *et al.*, 1999). Nei modelli GR e ANV si è assunto come variabile di transizione  $\epsilon_{t-1}$ , mentre in tutti i casi il test per (ST)ARCH è stato calcolato scegliendo  $q = 4$ .

Dalla Tabella 5.4 si nota che, quando il modello considerato è quello di Lundbergh e Teräsvirta (1998), la regola di decisione ha potenza maggiore, mentre in tutti gli altri casi i risultati sono sensibilmente inferiori. Tra le motivazioni che possono portare all'accettazione di un modello non-lineare quando in realtà il DGP della serie è lineare vi è senz'altro la constatazione che il numero dei regressori inclusi nel test per STARARCH è notevolmente maggiore di quello del test di Engle (1982) (si veda la Tabella 5.3). Un modo per ovviare a questo inconveniente potrebbe essere quello di porre il confronto tra i p-value sulla base di una diversa scelta di  $q$ . Selezionare un valore di  $q$  più lungo per il test di Engle (1982) (ad esempio, 4 o 8) ed uno più corto (ad esempio, 2 o 4) in quello per STARARCH. Questo dovrebbe, almeno in parte, ridurre la distorsione della regola proposta. Ulteriori verifiche e simulazioni in tal senso sono comunque auspicabili.

La procedura descritta è stata applicata alle serie dei rendimenti giornalieri degli

Tabella 5.5: Risultati relativi agli indici di borsa.

Indice	ARCH(8)	STARCH(8)			
		H	LT	GR ( $d = 1$ )	ANV ( $d = 1$ )
EOE	$5.1 \times 10^{-140}$	Logistico	Logistico	Espon.zle	Espon.zle
DAX	$3.6 \times 10^{-48}$	Espon.zle	Logistico	Espon.zle	Logistico
CAC40	$1.4 \times 10^{-94}$	Espon.zle	Espon.zle	Espon.zle	Espon.zle
FTSE100	$4.0 \times 10^{-273}$	Espon.zle	Espon.zle	Espon.zle	Espon.zle
HANG SENG	$2.7 \times 10^{-5}$	Espon.zle	Espon.zle	Espon.zle	Espon.zle
NIKKEI	$3.0 \times 10^{-36}$	Logistico	Logistico	Espon.zle	Logistico
ALL SHARES	$1.0 \times 10^{-21}$	Logistico	Logistico	Logistico	Logistico
S&P500	$9.9 \times 10^{-20}$	Espon.zle	Espon.zle	Espon.zle	Espon.zle

Tabella 5.6: Risultati relativi ai tassi di cambio.

Tasso di Cambio contro \$	ARCH(8)	STARCH(8)			
		H	LT	GR ( $d = 1$ )	ANV ( $d = 1$ )
AUD/USD	$2.6 \times 10^{-14}$	Espon.zle	Espon.zle	Espon.zle	Espon.zle
GPB/USD	$1.2 \times 10^{-54}$	Espon.zle	Espon.zle	Logistico	Logistico
CAD/USD	$4.5 \times 10^{-62}$	Logistico	Logistico	Logistico	Logistico
DEM/USD	$5.9 \times 10^{-44}$	Espon.zle	GARCH	GARCH	GARCH
NLG/USD	$6.5 \times 10^{-46}$	Espon.zle	GARCH	GARCH	GARCH
FRF/USD	$1.2 \times 10^{-24}$	Logistico	Logistico	Logistico	Logistico
JPY/USD	$1.4 \times 10^{-54}$	Espon.zle	Espon.zle	Espon.zle	Espon.zle
CHF/USD	$8.7 \times 10^{-33}$	GARCH	GARCH	GARCH	GARCH

indici di borsa, dei tassi di cambio e delle azioni descritte nell'Appendice A. I test per ARCH e STARCH sono stati applicati sui residui al quadrato ricavati dalla stima del modello lineare AR(p) selezionato nella Sezione 3.5.5. Dal confronto tra i p-value del test per ARCH con quelli del corrispondente test congiunto per STARCH è stato possibile decidere se l'evoluzione della varianza condizionata sia di tipo lineare o meno. Laddove il valore di  $d$  nella funzione di transizione deve essere specificato dai dati (modelli di González-Rivera, 1998 ed Anderson *et al.*, 1999), si è assunto un  $d = 1$  poiché, in pratica, la specificazione più comune è rappresentata da un modello STGARCH(1,1) con variabile di transizione  $\epsilon_{t-1}$ . I risultati per i tre gruppi di serie analizzate sono riportati nelle Tabelle 5.5, 5.6 e 5.7. Per ciascun test si è scelto un  $q = 8$ . Ogni Tabella riporta il p-value del test per ARCH e, confrontando quest'ultimo con quelli dei corrispondenti test per STARCH, la scelta del particolare modello selezionato (logistico, esponenziale o GARCH lineare). Come si vede, quasi sempre emergono forti evidenze a favore della nonlinearietà di tipo STGARCH, mentre solo occasionalmente la scelta ricade su un modello GARCH lineare.

Hagerud (1996) e González-Rivera (1998) derivano un test dei moltiplicatori di Lagrange (LM) per la verifica della linearità contro l'alternativa di tipo *smooth transi-*



Tabella 5.7: Risultati relativi alle azioni.

Titolo	ARCH(8)	STARCH(8)			
		H	LT	GR ( $d = 1$ )	ANV ( $d = 1$ )
AAPL	0.945	Var. Cost.	Logistico	Logistico	Logistico
AMD	0.033	Espon.zle	GARCH	Espon.zle	Espon.zle
INTC	$7.5 \times 10^{-15}$	Espon.zle	Espon.zle	Logistico	Logistico
KO	$3.9 \times 10^{-41}$	Logistico	Logistico	GARCH	GARCH
MCD	$5.7 \times 10^{-24}$	GARCH	GARCH	Espon.zle	Espon.zle
MHP	$1.7 \times 10^{-34}$	Logistico	Logistico	GARCH	GARCH
PEP	$4.8 \times 10^{-51}$	Logistico	Logistico	Espon.zle	Espon.zle
T	$1.0 \times 10^{-8}$	Espon.zle	Espon.zle	Logistico	Logistico

tion nell'equazione della varianza condizionata. Com'è noto, il test dei moltiplicatori di Lagrange è la forma tipica con cui si sottopone a verifica una qualche forma di non-linearità specifica (Granger e Teräsvirta, 1993). Questo perché la sua applicazione richiede la stima del solo modello sotto l'ipotesi nulla, ed è pertanto molto più operativo di altre famiglie di test. Se nell'equazione della media condizionata, il modello sotto l'ipotesi nulla è rappresentato da un modello lineare autoregressivo, per il quale non sorge alcuna complicazione di stima, altrettanto non si può dire per quanto riguarda l'equazione della varianza condizionata. L'ipotesi che si sottopone a test riguarda un'evoluzione dinamica di  $h_t$  che, seppur di tipo lineare, rende quest'ultima non costante nel tempo. Di conseguenza, non è possibile sottoporre a test la non-linearità della varianza condizionata senza aver prima stimato un modello (G)ARCH lineare. Pertanto, la statistica test del tipo LM per la verifica dell'ipotesi di linearità di tipo ARCH contro l'alternativa di tipo STARCH assume la forma

$$\frac{1}{2} \left\{ \sum_{t=1}^T \frac{1}{h_{0t}} \left[ \frac{\hat{\epsilon}_t^2}{h_{0t}} - 1 \right] \hat{\mathbf{v}}_t \right\}' \left\{ \sum_{t=1}^T \frac{1}{h_{0t}^2} \hat{\mathbf{v}}_t \hat{\mathbf{v}}_t' \right\}^{-1} \left\{ \sum_{t=1}^T \frac{1}{h_{0t}} \left[ \frac{\hat{\epsilon}_t^2}{h_{0t}} - 1 \right] \hat{\mathbf{v}}_t \right\} \quad (5.21)$$

dove  $h_{0t}$  indica la varianza condizionata sotto l'ipotesi nulla.  $h_{0t}$  può essere ricavata dalla stima di un modello ARCH(q) sui dati, mentre  $\mathbf{v}_t$  indica il vettore dei regressori di cui alla Tabella 5.3. La statistica test (5.21), sotto l'ipotesi nulla, si distribuisce asintoticamente come una  $\chi^2$  con  $q$  gradi di libertà. L'inclusione nella (5.21) del vettore  $\mathbf{v}_t$  in luogo del vettore dei regressori originari del modello STARCH si rivela essenziale a causa dei soliti problemi di identificazione di alcuni parametri sotto l'ipotesi nulla. Ed è proprio grazie all'idea dell'approssimazione di Luukkonen *et al.* (1988b) che si può rendere operativo il test dei moltiplicatori di Lagrange in simili circostanze.<sup>11</sup>

Alternativamente, e in analogia con quanto visto in precedenza, è possibile derivare una forma della statistica test basata su una regressione ausiliaria, la cui distribuzione

<sup>11</sup>Per una derivazione formale della statistica test (5.21) si vedano Hagerud (1996) e González-Rivera (1998).

asintotica è equivalente a quella della (5.21). Il modello ausiliario di riferimento è

$$\left[ \frac{\hat{\epsilon}_t^2}{h_{0t}} - 1 \right] = \alpha' \left( \frac{\hat{\mathbf{v}}_t}{h_{0t}} \right) + \xi_t \quad (5.22)$$

e la statistica test può essere calcolata come  $TR^2$ , ricavati dalla stessa regressione.

Infine, Hagerud (1997a) sviluppa un test per la verifica dell'ipotesi nulla di tipo GARCH(1,1) contro l'alternativa STGARCH(1,1) logistica. Con riferimento all'equazione (5.9) dove  $q = p = 1$ , il test può essere formulato come

$$H_0 : \alpha_{21} = 0 \quad (5.23)$$

$$H_1 : \alpha_{21} \neq 0 \quad (5.24)$$

La risultante statistica test non è però calcolabile a causa della non identificabilità di  $\gamma$  sotto  $H_0$ . Per renderla operativa si considera ancora una volta l'approssimazione di Taylor al primo ordine della funzione logistica (Luukkonen *et al.*, 1988b), pervenendo così al modello ausiliario

$$\tilde{h}_t = \omega + \alpha_{11}\epsilon_{t-1}^2 + \alpha_{21}\frac{\gamma}{4}\epsilon_{t-1}^3 + \beta_1 h_{t-1} \quad (5.25)$$

dove il test può essere espresso attraverso la verifica dell'ipotesi nulla

$$H_0 : \alpha_{21}\frac{\gamma}{4} = 0 \quad (5.26)$$

contro l'alternativa

$$H_1 : \alpha_{21}\frac{\gamma}{4} \neq 0 \quad (5.27)$$

Sotto l'ipotesi di normalità condizionata degli  $\epsilon_t$ , la statistica test LM assume la forma

$$\begin{aligned} & \frac{1}{2} \left\{ \sum_{t=1}^T \frac{1}{2h_{0t}} \left[ \frac{\hat{\epsilon}_t^2}{h_{0t}} - 1 \right] \frac{\partial \tilde{h}_t}{\partial \tilde{\alpha}} \right\}' \left\{ \sum_{t=1}^T \left[ \frac{1}{h_{0t}} \frac{\partial \tilde{h}_t}{\partial \tilde{\alpha}} \right] \left[ \frac{1}{h_{0t}} \frac{\partial \tilde{h}_t}{\partial \tilde{\alpha}} \right]' \right\}^{-1} \\ & \times \left\{ \sum_{t=1}^T \frac{1}{2h_{0t}} \left[ \frac{\hat{\epsilon}_t^2}{h_{0t}} - 1 \right] \frac{\partial \tilde{h}_t}{\partial \tilde{\alpha}} \right\} \end{aligned} \quad (5.28)$$

dove

$$\frac{\partial \tilde{h}_t}{\partial \tilde{\alpha}'} = \left[ \sum_{i=1}^{t-1} \hat{\beta}_1^{i-1}, \sum_{i=1}^{t-1} \hat{\beta}_1^{i-1} \hat{\epsilon}_{t-i}^2, \sum_{i=1}^{t-1} \hat{\beta}_1^{i-1} \hat{\epsilon}_{t-i}^3, \sum_{i=1}^{t-1} \hat{\beta}_1^{i-1} h_{0t-i} \right], \quad (5.29)$$

$h_{0t}$  indica la varianza condizionata sotto l'ipotesi nulla di GARCH(1,1),  $\tilde{\alpha}'$  rappresenta il vettore dei parametri nel modello (5.25),  $(\omega, \alpha_{11}, \alpha_{21}\frac{\gamma}{4}, \beta_1)$  e  $\hat{\beta}_1$  è il parametro  $\beta_1$  stimato nel modello GARCH(1,1). La (5.28) si distribuisce asintoticamente come una  $\chi^2$  con un grado di libertà, mentre la sua versione asintoticamente equivalente può essere calcolata come  $TR^2$  dalla regressione

$$\begin{aligned} \frac{\hat{\epsilon}_t^2}{h_{0t}} - 1 &= \left( \frac{\sum_{i=1}^{t-1} \hat{\beta}_1^{i-1}}{h_{0t}} \right) + \theta_1 \left( \frac{\sum_{i=1}^{t-1} \hat{\beta}_1^{i-1} \hat{\epsilon}_{t-i}^2}{h_{0t}} \right) + \\ &+ \theta_2 \left( \frac{\sum_{i=1}^{t-1} \hat{\beta}_1^{i-1} h_{0t-i}}{h_{0t}} \right) + \theta_3 \left( \frac{\sum_{i=1}^{t-1} \hat{\beta}_1^{i-1} \hat{\epsilon}_{t-i}^3}{h_{0t}} \right) + \xi_t \end{aligned} \quad (5.30)$$

In modo analogo, per il caso esponenziale, la regressione ausiliaria è identica alla (5.30) con l'eccezione che il termine

$$\theta_3 \left( \frac{\sum_{i=1}^{t-1} \hat{\beta}^{i-1} \hat{\epsilon}_{t-i}^3}{h_{0t}} \right)$$

è sostituito da

$$\theta_4 \left( \frac{\sum_{i=1}^{t-1} \hat{\beta}^{i-1} \hat{\epsilon}_{t-i}^4}{h_{0t}} \right)$$

Lo stesso Hagerud (1997a) mostra che una versione più semplificata della (5.30) è in grado di ottenere quasi gli stessi risultati. Pertanto, suggerisce di basare il test sui seguenti passi:

1. Stimare il modello GARCH(1,1) e formare i vettori

$$\mathbf{d}_t = \left( 1, \frac{\sum_{i=1}^{t-1} \hat{\beta}^{i-1} \hat{\epsilon}_{t-i}^3}{h_{0t}} \right)$$

2. Regredire  $\left\{ \frac{\hat{\epsilon}_t^2}{h_{0t}} - 1 \right\}$  su  $\mathbf{d}_t$  e calcolare la quantità  $TR^2$  che, sotto l'ipotesi nulla, si distribuisce asintoticamente come una  $\chi^2$  con un grado di libertà.

Un test analogo a quello di Hagerud (1997a), sviluppato seguendo le considerazioni di Davies (1977 e 1987), è stato indipendentemente proposto da González-Rivera (1998). Entrambi gli autori presentano esaurienti risultati ottenuti per simulazione che confermano la buona potenza del test nel discriminare tra un modello GARCH(1,1) ed uno STGARCH(1,1) di tipo logistico. La scelta di un modello logistico sotto l'ipotesi alternativa in luogo di un modello esponenziale è dovuta alla considerazione empirica che le innovazioni hanno un diverso impatto sull'evoluzione della varianza condizionata a seconda del loro segno (Black, 1976). Una tale caratteristica, come precedentemente mostrato (si veda la Sezione 5.1.1), è ben catturata da un modello STGARCH logistico. Il test così derivato può essere utilizzato per verificare la presenza di effetti *leverage* sui residui standardizzati di un modello (G)ARCH. Infatti, Hagerud (1997a), mediante simulazioni, mostra come tale test abbia una potenza ampiamente superiore ai test tradizionali di asimmetria proposti da Engle e Ng (1993), anche nei casi in cui il DGP della serie non sia un modello LSTGARCH(1,1), ma bensì uno qualsiasi dei modelli asimmetrici per la varianza condizionata presenti in letteratura.<sup>12</sup> Applicazioni empiriche su serie finanziarie si possono trovare in Hagerud (1997b).

<sup>12</sup>In particolare, i modelli considerati sono il GQARCH(1,1), l'EGARCH(1,1), il modello GJR(1,1), il TGARCH(1,1), l'A-PARCH(1,1) e quello VS-GARCH (si veda la Sezione 2.2). Inoltre, Hagerud (1997a) deriva un test per l'ipotesi nulla di GARCH(1,1) contro l'alternativa GQARCH(1,1) in modo simile a quanto visto per il caso LSTGARCH(1,1). Le stesse simulazioni di cui sopra confermano, anche in tal caso, una maggiore potenza del test rispetto ai tradizionali strumenti per la verifica dell'asimmetria nell'evoluzione della varianza condizionata proposti da Engle e Ng (1993): si tratta del *sign bias test*, del *negative* e *positive size bias test* e del *joint test* (si veda anche Gallo e Pacini, 2002). In aggiunta, il test contro l'alternativa GQARCH(1,1) presenta potenza maggiore anche nei confronti del test per LSTGARCH(1,1).

### 5.2.2 Una versione robusta del test

Il test per la presenza di effetti ARCH derivato da Engle (1982), non solo richiede che il momento quarto condizionato di  $\epsilon_t$  esista, ma anche che sia costante e finito. Partendo da queste considerazioni, Wooldridge (1990) sviluppa una versione robusta del test che ne assicura una soddisfacente performance asintotica. Così il test può essere calcolato utilizzando la procedura seguente:

1. Stimare la varianza campionaria sotto l'ipotesi nulla di omoschedasticità:

$$\hat{\sigma}_\epsilon^2 = \frac{1}{T} \sum_{t=1}^T \hat{\epsilon}_t^2$$

dove  $\hat{\epsilon}_t$  indica i residui ricavati dal modello per la media condizionata.

2. Regredire 1 su  $(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2)$ ,  $(\hat{\epsilon}_{t-1}^2 - \hat{\sigma}_\epsilon^2)$ , ...,  $(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2)$ ,  $(\hat{\epsilon}_{t-q}^2 - \hat{\sigma}_\epsilon^2)$  e calcolare la statistica test  $TR^2 = T - RSS$  dalla stessa regressione, dove  $R^2$  indica l'indice di determinazione lineare,  $RSS$  la somma dei quadrati dei residui e  $T$  il numero di osservazioni. La statistica test converge in distribuzione ad una  $\chi^2$  con  $q$  gradi di libertà.

La strategia proposta è immediatamente generalizzabile al test per ARCH a transizione liscia precedentemente derivato, le cui diverse specificazioni sono riassunte nella Tabella 5.3.

Hagerud (1997b) presenta una robustificazione della procedura di test per la scelta tra un modello GARCH(1,1) e uno STGARCH(1,1) logistico espresso dalla (5.9). La correzione del test si rivela necessaria poiché la statistica (5.28) è stata derivata sotto l'ipotesi di normalità condizionata degli  $\epsilon_t$ . Una tale assunzione si rivela però inadeguata per la quasi totalità delle serie storiche finanziarie, implicando così una probabile distorsione del test.

Utilizzando la metodologia proposta da Wooldridge (1991) in simili circostanze, Hagerud (1997b) corregge la procedura di test nel modo seguente:

1. Stimare il modello GARCH(1,1) e formare la serie della varianza condizionata sotto l'ipotesi nulla  $\{h_{0t}\}_{t=1}^T$ . Sia inoltre  $\hat{\beta}$  la stima del parametro  $\beta$  nel modello GARCH(1,1).
2. Regredire

$$\frac{\sum_{i=1}^{t-1} \hat{\beta}^{i-1} \hat{\epsilon}_{t-i}^3}{h_{0t}}$$

su

$$\mathbf{y}_t = \left\{ \frac{\sum_{i=1}^{t-1} \hat{\beta}^{i-1}}{h_{0t}}, \frac{\sum_{i=1}^{t-1} \hat{\beta}^{i-1} \hat{\epsilon}_{t-i}^2}{h_{0t}}, \frac{\sum_{i=1}^{t-1} \hat{\beta}^{i-1} h_{0t-i}}{h_{0t}} \right\}$$

e sia  $\hat{\mathbf{b}}$  il vettore dei parametri stimati. Formare quindi la serie dei residui

$$s_t = \frac{\sum_{i=1}^{t-1} \hat{\beta}^{i-1} \hat{\epsilon}_{t-i}^3}{h_{0t}} - \hat{\mathbf{b}} \mathbf{y}_t'$$

3. Regredire 1 su

$$\left( \frac{\hat{\epsilon}_t^2}{h_{0t}} - 1 \right) .* s_t$$

dove il simbolo  $.*$  indica la moltiplicazione elemento per elemento e calcolare la statistica test  $TR^2 = T - RSS$  dalla stessa regressione. Tale quantità converge in distribuzione a una  $\chi^2$  con un grado di libertà.

Seguendo Wooldridge (1990 e 1991) e Hagerud (1997b), è agevole derivare una versione robusta della statistica (5.21), ovvero della sua versione asintoticamente equivalente ricavata dal modello ausiliario (5.22). Premessa la seguente notazione,  $\hat{\mathbf{w}}_t = (1, \hat{\epsilon}_{t-1}^2, \dots, \hat{\epsilon}_{t-q}^2)'$  e indicato con  $\hat{\mathbf{z}}_t$  il vettore dei regressori ausiliari in ciascuna versione del test per STARCH (quelli aggiuntivi a  $\hat{\mathbf{w}}_t$ , secondo la Tabella 5.3), la procedura si basa sui seguenti passi:

1. Stimare un modello ARCH(q) e formare le serie  $\{h_{0t}\}_{t=1}^T$  e  $\left( \frac{\hat{\epsilon}_t^2}{h_{0t}} - 1 \right)$ .
2. Standardizzare i vettori  $\hat{\mathbf{w}}_t$  e  $\hat{\mathbf{z}}_t$  tramite  $h_{0t}$  e regredire ciascun elemento di  $\left( \frac{\hat{\mathbf{z}}_t}{h_{0t}} \right)$  su  $\left( \frac{\hat{\mathbf{w}}_t}{h_{0t}} \right)$  e calcolare i residui  $\hat{u}_t$ .
3. Ponderare la matrice dei residui calcolata al punto precedente tramite il vettore  $\left( \frac{\hat{\epsilon}_t^2}{h_{0t}} - 1 \right)$  e regredire 1 sulla matrice dei residui ponderati. La quantità  $TR^2 = T - RSS$  rappresenta la statistica test che, in distribuzione, converge a una  $\chi^2$  con gradi di libertà pari al numero dei regressori inclusi nel vettore  $\hat{\mathbf{z}}_t$ .

Alla luce delle robustificazioni del test qui presentate, si potrebbe riconsiderare l'esperimento di simulazione precedentemente effettuato, in modo da valutarne la potenza in situazioni più idonee a descrivere il comportamento della maggior parte delle serie storiche finanziarie (in particolare la non normalità degli  $\epsilon_t$ ).

Infine, secondo quanto descritto nella Sezione 4.2, tecniche robuste di stima possono essere utilizzate per costruire dei test per ARCH e STARCH sensibili alla presenza di *outliers*. Il raggruppamento (*clustering*) di grandi innovazioni tipico dei modelli ARCH potrebbe essere ben mimetizzato da una sequenza di *outliers* additivi (AO) (Balke e Fomby, 1994 e Franses e van Dijk, 2000). Il comportamento del test per ARCH in presenza di *outliers* è stato analizzato da Franses, van Dijk e Lucas (1998) e da van Dijk, Franses e Lucas (1999b). È stato mostrato che la presenza di *outliers* produce effetti distorsivi sul test simili a quelli che si riscontrano nel test di linearità applicato all'equazione della media condizionata di un generico processo  $y_t$  (si veda la Sezione 4.2.2). I valori anomali possono portare ad accettare l'ipotesi di effetti ARCH quando in realtà ciò non è vero, mentre altre volte possono anche mascherarne la presenza. Statistiche test robuste a questa evenienza sono in grado di discriminare correttamente fra omoschedasticità e eteroschedasticità di tipo ARCH. Intuitivamente (cfr. Franses e van Dijk, 1997), il test per ARCH modificato per la presenza di *outliers* additivi (AO) può essere calcolato formando la serie dei residui ponderati dal modello per la media condizionata stimato utilizzando tecniche robuste, e regredendo poi tali residui ponderati al quadrato su una costante e  $q$  valori ritardati. La risultante statistica test è ancora

asintoticamente distribuita come un  $\chi^2$  con  $q$  gradi di libertà. Simulazioni effettuate da van Dijk *et al.* (1999b) mostrano che la versione robusta della statistica test possiede buona potenza già per campioni di 100 osservazioni, anche in assenza di *outliers*. Estensioni al caso STARCH possono essere previste senza difficoltà.

### 5.3 Stima dei modelli STGARCH

Avendo ipotizzato che le innovazioni  $\eta_t$  siano Gaussiane, i parametri di un modello STGARCH possono essere stimati massimizzando la funzione di log-verosimiglianza condizionata.<sup>13</sup> Posto  $\theta$  il vettore dei parametri incogniti in uno dei modelli STGARCH di cui alla Sezione 6.1, dalla (5.3) si ricava la log-verosimiglianza condizionata della  $t$ -esima osservazione come

$$\log l_t(\theta) = -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(h_t) - \frac{1}{2} \frac{\epsilon_t^2}{h_t} \quad (5.31)$$

dalla quale è possibile costruire la funzione di log-verosimiglianza

$$\log L(\theta) = \sum_{t=1}^T \log l_t(\theta) \quad (5.32)$$

dove si è supposto di aver osservato un campione di  $T + \max\{q, p\}$  osservazioni. In forma meno compatta avremo

$$\log L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(h_t) - \frac{1}{2} \sum_{t=1}^T \frac{\epsilon_t^2}{h_t} \quad (5.33)$$

e la dipendenza della  $\log L(\cdot)$  dal vettore  $\theta$  diventa più chiara nel momento in cui si sostituisce al posto di  $h_t$  una delle rappresentazioni STGARCH descritte nella Sezione 5.1. Così ad esempio, se consideriamo il modello ANST-GARCH di Anderson *et al.* (1999) dove  $q = p = d = 1$ , il vettore  $\theta$  sarà composto dai parametri incogniti  $(\omega_1, \alpha_{11}, \beta_{11}, \omega_2, \alpha_{21}, \beta_{21}, \gamma)'$ . In modo simile si procede negli altri casi.

Poiché non esiste una soluzione analitica della (5.33), la procedura di massimizzazione sarà vincolata all'utilizzo di uno dei metodi di ottimizzazione numerica presenti in letteratura (si veda, ad esempio, Hamilton, 1994, Sezione 5.7). Ciascuno di questi algoritmi richiede la fissazione di valori iniziali per i parametri incogniti. Seguendo Bollerslev (1986) e Fiorentini, Calzolari e Panattoni (1996), una scelta ragionevole è rappresentata dai valori delle autocorrelazioni dei residui al quadrato.

Un altro problema è legato al calcolo della sequenza delle varianze condizionate  $\{h_t\}_{t=1}^T$  da inserire nella (5.33) in quanto è vincolata alla presenza di valori precampionari per  $h_t$ ,  $(h_{-p+1}, \dots, h_0)$  e per  $\epsilon_t^2$ ,  $(\epsilon_{-q+1}^2, \dots, \epsilon_0^2)$ . Bollerslev (1986) ha suggerito di porre

$$h_j = \epsilon_j^2 = \hat{\sigma}_\epsilon^2 \quad \text{per } j = -\max\{q+1, p+1\}, \dots, 0$$

<sup>13</sup>Anche qui, come nella Sezione 3.6,  $\log$  indica il logaritmo naturale.

e dove

$$\hat{\sigma}_\epsilon^2 = \sum_{t=1}^T \hat{\epsilon}_t^2 \quad (5.34)$$

La sequenza  $\{h_t\}_{t=1}^T$  così calcolata può essere impiegata per valutare la funzione di log-verosimiglianza condizionata espressa dalla (5.33). A questo punto può prendere avvio il procedimento iterativo di massimizzazione.

Se l'evoluzione di  $\mu_t$  è governata da una qualche specificazione dinamica (ad esempio un modello della classe STAR), i parametri di  $h_t$  dovrebbero essere stimati simultaneamente a quelli del modello per la media condizionata. Questo risulta particolarmente importante in tutti quei casi in cui la matrice d'informazione non è diagonale a blocchi e/o la varianza condizionata non è simmetrica rispetto al termine di errore  $\epsilon_t$ . Engle (1982) dimostra che quando  $h_t$  è simmetrica rispetto alle innovazioni, così come accade nei modelli ARCH, GARCH e STGARCH esponenziale, le due equazioni possono essere stimate separatamente senza perdita di efficienza asintotica. Tali considerazioni non valgono invece nel modello logistico che, per costruzione, ha la caratteristica di descrivere dinamiche asimmetriche nell'evoluzione di  $h_t$  e che, a parità di entità, dipendono dal segno delle innovazioni. In simili circostanze,<sup>14</sup> può comunque essere conveniente stimare separatamente le due equazioni e solo successivamente, dopo una prima fase di verifica del modello, procedere ad una stima congiunta per recuperare l'efficienza asintotica. Le stime dei parametri ottenute separatamente potranno essere convenientemente utilizzate come valori iniziali della procedura iterativa di massimizzazione sul modello congiunto.

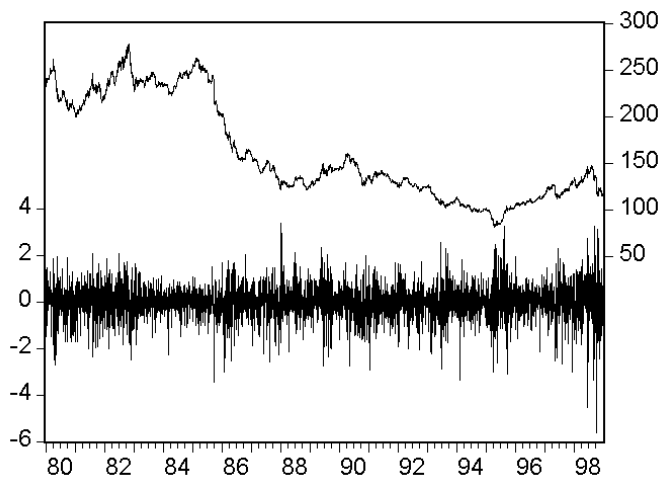
Al fine di inferire correttamente sui parametri del modello stimato, occorre disporre di una stima consistente della matrice di varianze e covarianze di  $\theta$ . Il problema maggiore che si incontra in questo caso è legato all'utilizzo di una funzione di verosimiglianza Gaussiana. Si è già visto che la maggior parte delle serie finanziarie tende a rifiutare ampiamente l'ipotesi che  $\eta_t \sim iid N(0, 1)$ . Bollerslev e Wooldridge (1992) mostrano che la massimizzazione della verosimiglianza Gaussiana continua a fornire stime consistenti di  $\theta$  anche quando la distribuzione condizionata di  $\epsilon_t$  non è Gaussiana. Tuttavia, sulla falsariga di quanto evidenziato da White (1982) nella stima del modello per la media condizionata, i due autori mettono in evidenza che gli *standard error* delle stime devono essere corretti. In particolare, propongono di ricavare una stima consistente della matrice di varianze e covarianze di  $\theta$  calcolando la quantità

$$Asy\widehat{Var}(\hat{\theta}_{ML}) \simeq \frac{1}{T} \left[ \mathbf{I}(\hat{\theta}_{ML})^{-1} \mathbf{OP}(\hat{\theta}_{ML}) \mathbf{I}(\hat{\theta}_{ML})^{-1} \right] \quad (5.35)$$

dove  $\mathbf{I}(\cdot)$  e  $\mathbf{OP}(\cdot)$  indicano rispettivamente la matrice di informazione e il prodotto esterno del gradiente della funzione di log-verosimiglianza, valutati in corrispondenza di  $\hat{\theta}_{ML}$ . Per una derivazione formale delle espressioni assunte dalle quantità  $\mathbf{I}(\cdot)$  e  $\mathbf{OP}(\cdot)$  nel caso GARCH si veda, fra gli altri, Bollerslev e Wooldridge (1992), Hamilton (1994) e Fiorentini *et al.* (1996). Nel caso specifico dei modelli GARCH a transizione liscia, simili espressioni si possono trovare in Lundbergh e Teräsvirta (1998 e 1999).

<sup>14</sup>La discussione non è infatti limitata al modello STGARCH logistico, ma vale anche per tutti i modelli della classe GARCH dove si prevede esplicitamente la presenza di effetti *leverage* (si veda la Sezione 2.2).

Figura 5.1: Andamento del tasso di cambio JPY/USD (scala di destra) e dei corrispondenti rendimenti percentuali (scala di sinistra) nel periodo 31 dicembre 1979 – 31 dicembre 1998.



A scopo illustrativo, consideriamo la stima di un modello AR-STGARCH sulla serie del tasso di cambio JPY/USD il cui andamento è riportato, assieme a quello dei rendimenti, in Figura 5.1. La Tabella 5.8 riporta invece alcune caratteristiche descrittive della serie dei rendimenti percentuali.

Tabella 5.8: Principali statistiche descrittive dei rendimenti percentuali del tasso di cambio JPY/USD.

Statistica	Valore	p-value
Media	-0.0152	
Mediana	0.0083	
Min	-5.6302	
Max	3.3657	
Deviazione standard	0.6717	
Skewness	-0.5627	
Curtosi	7.2061	
Jarque-Bera	3916.35	0.0000

L'equazione della media condizionata è stata specificata stimando un modello autoregressivo lineare con 10 ritardi più un termine costante, così come risulta dalla Tabella 5.9, al quale sono state aggiunte delle variabili *dummy* per verificare la presenza o meno di effetti legati al giorno della settimana. Per evitare problemi di perfetta collinearità, si è deciso di mantenere la costante e di omettere la variabile *dummy* relativa al venerdì. In tal modo, i coefficienti stimati per gli altri giorni indicano la variazione intercorsa in quel giorno rispetto al venerdì, ovvero rispetto al termine costante. Per l'equazione del-



la varianza condizionata si è invece deciso di stimare un modello ESTGARCH(1,1) con variabile di transizione  $\epsilon_{t-1}$  nella formulazione di Anderson *et al.* (1999) sui residui del modello per la media.<sup>15</sup> I risultati delle stime sono riportati nella Tabella 5.9.

Tabella 5.9: Risultati della stima di un modello AR-ESTAGRCH con *dummy* giornaliera e variabile di transizione  $\epsilon_{t-1}$  sulla serie del tasso di cambio JPY/USD.

Parametro	Stima	Std. Error
$\pi_0$	0.0294	0.0213
$\pi_1$	0.0551	0.0142
$\pi_2$	0.0273	0.0143
$\pi_{10}$	0.0549	0.0141
Lunedì	-0.0385	0.0302
Martedì	-0.0655	0.0301
Mercoledì	-0.0439	0.0300
Giovedì	-0.0661	0.0301
Primo Regime		
$\omega_1$	0.0879	0.0134
$\alpha_{11}$	0.0000	–
$\beta_{11}$	0.7843	0.0230
Secondo Regime		
$\omega_2$	-0.1842	0.0305
$\alpha_{21}$	0.3348	0.0440
$\beta_{21}$	0.2358	0.0488
Funzione di Transizione		
$\gamma$	1.0249	0.0717

## 5.4 Valutazione del modello stimato

Una volta che il modello STGARCH è stato stimato, occorre valutarne le proprietà al fine di poterlo impiegare correttamente per descrivere o prevedere la dinamica della varianza condizionata del generico fenomeno  $y_t$ . In analogia a quanto visto per il modello STAR (cfr. Sezione 3.7), è necessario ispezionare l'andamento della funzione di transizione stimata per verificare che, in linea di massima, rispecchi quello della sua controparte teorica.

Di estrema importanza si rivela poi la valutazione delle proprietà dei residui standardizzati, in quanto una loro non adeguatezza alle ipotesi formulate implicherebbe inevitabilmente una riconsiderazione del modello stimato. Così diventa importante verificare che la loro distribuzione empirica sia prossima a quella della normale standard poiché si è assunto

$$\epsilon_t = \eta_t \sqrt{h_t}, \quad \eta_t \sim iid N(0, 1), \quad (5.36)$$

<sup>15</sup>Si veda la Tabella 5.6.

da cui dovrebbe risultare che

$$\eta_t = \frac{\epsilon_t}{\sqrt{h_t}} \sim N(0, 1). \quad (5.37)$$

La verifica empirica di una simile ipotesi può essere effettuata tramite il test di normalità Jarque-Bera sui residui standardizzati.

Inoltre, è necessario valutare se il modello stimato è in grado di ripulire completamente  $\epsilon_t$  dalla presenza di eteroschedasticità condizionata. Pertanto, i test per ARCH e STARARCH discussi nella Sezione 5.2 dovrebbero essere applicati alla serie dei residui standardizzati al quadrato.

Se i risultati ottenuti da queste prime valutazioni sono soddisfacenti, allora una prima fase di valutazione può dirsi conclusa. Infatti, come accadeva per la media condizionata, occorre verificare la presenza o meno di non-linearità rimanenti, nonché la costanza dei parametri stimati. Di questo ci occuperemo nelle Sezioni successive.

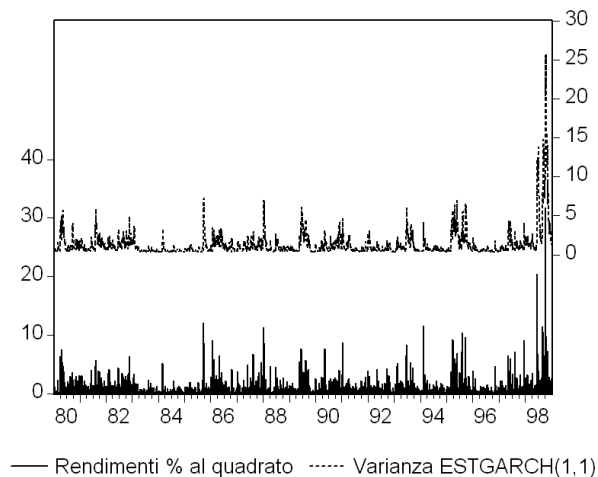
Tabella 5.10: Diagnostiche sui residui standardizzati del modello AR-ESTGARCH con *dummy* giornaliero.

Statistica	Valore	p-value
Media	0.0066	
Mediana	0.0262	
Min	-4.6939	
Max	3.3680	
Deviazione standard	0.6727	
Skewness	-0.4767	
Curtosi	6.5039	
Jarque-Bera	2717.99	0.0000
ARCH(4)	7.9826	0.0922

Da una prima analisi del modello stimato sul tasso di cambio JPY/USD (cfr. Tabella 5.10) emerge come la distribuzione empirica dei residui standardizzati presenti asimmetria a sinistra accompagnata da una forte curtosi, portando a rifiutare abbondantemente l'ipotesi di normalità. Il problema è comune alla maggior parte delle stime su serie finanziarie e difficilmente può essere risolto impiegando altri modelli della classe GARCH, nonostante che una delle loro caratteristiche principali sia quella di generare un coefficiente di curtosi maggiore di quello della normale. Apprezzabile è invece la stima della varianza condizionata in quanto, oltre a caratterizzare adeguatamente periodi di alta e bassa volatilità, riduce ampiamente il grado di persistenza tipico dei modelli GARCH lineari. Le Figure 5.2 e 5.3 mostrano l'andamento dei rendimenti al quadrato confrontato con quello della stima della varianza condizionata, rispettivamente nei modelli ESTGARCH(1,1) e GARCH(1,1).<sup>16</sup>

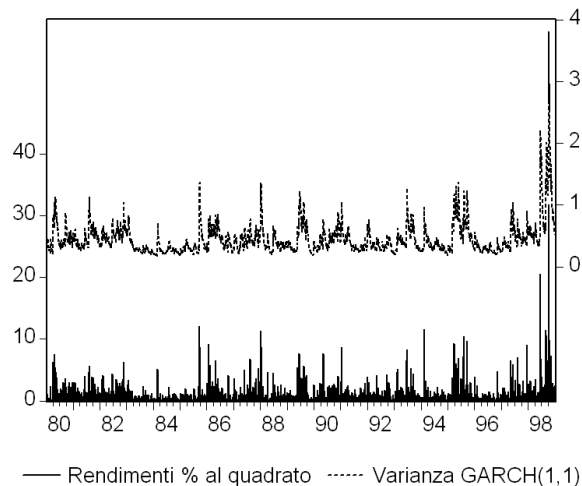
<sup>16</sup>Il modello GARCH(1,1) sui residui di quello autoregressivo con *dummy* giornaliero e costante non è stato riportato poiché i risultati dei test nella Tabella 5.6 spingono nella direzione di una parametrizzazione non-lineare della varianza condizionata, indicando in un modello ESTGARCH la scelta preferibile. Pertanto il riferimento al GARCH(1,1) è solo quello di modello *benchmark* su cui basare il confronto. Per completezza, le stime dei parametri del GARCH(1,1) sono risultate essere (tra parentesi gli *standard error*):  $\omega = 0.0129$  (0.0013),  $\alpha = 0.0605$  (0.0040) e  $\beta = 0.9105$  (0.0061).

Figura 5.2: Confronto tra andamento dei residui al quadrato e varianza condizionata stimata col modello ESTGARCH(1,1) sulla serie del tasso di cambio JPY/USD.



Come si vede, l'adattamento della varianza stimata alla dinamica dei rendimenti al quadrato è migliore nel caso ESTGARCH in quanto, oltre ad assorbire più velocemente le innovazioni (minore persistenza della volatilità), descrive meglio la dinamica di  $h_t$  in periodi di alta volatilità. In particolare, la stima della varianza col modello GARCH(1,1) risulta pesantemente sottostimata rispetto all'andamento dei rendimenti al quadrato. Nel modello ESTGARCH(1,1) invece, un tale inconveniente si riduce ampiamente confermando empiricamente quelle che sono le premesse del modello, ovvero la presenza di più regimi di volatilità che si alternano nel corso del tempo sulla base di una qualche relazione funzionale completamente specificata dai dati (in tal caso la funzione di transizione esponenziale, la cui lisciatura è determinata dalla stima del parametro  $\gamma$ ); caratteristica questa che non può essere catturata dai modelli GARCH tradizionali. La Figura 5.4 mostra l'andamento della funzione di transizione esponenziale ricavata dalla stima del modello ESTGARCH(1,1) con variabile di transizione  $\epsilon_{t-1}$  e un valore stimato di  $\gamma$  pari a 1.0249. A conferma del buon adattamento ai dati del modello GARCH a transizione liscia, si riporta in Figura 5.5 lo *scatterplot* tra la varianza stimata col modello ESTGARCH(1,1) e quella stimata col GARCH(1,1). Se quest'ultimo fosse in grado di caratterizzare completamente la dinamica di  $h_t$ , non dovrebbero esserci apprezzabili differenze tra le due stime e i punti dello *scatter* dovrebbero disporsi approssimativamente intorno ad una retta positivamente inclinata. In realtà, a parte la diversità di scala delle due stime, lo scostamento dalla retta appare piuttosto marcato, specialmente nella parte bassa del grafico dove l'andamento sembra essere quello di un polinomio di grado superiore con concavità rivolta verso il basso e nella parte alta dove la concavità sembra cambiare segno.

Figura 5.3: Confronto tra andamento dei residui al quadrato e varianza condizionata stimata col modello GARCH(1,1) sulla serie del tasso di cambio JPY/USD.



#### 5.4.1 Verifica della presenza di ulteriori regimi di volatilità

A tal fine si riscrive il modello STGARCH nel modo seguente (Lundbergh e Teräsvirta, 1998 e 1999):

$$\epsilon_t = \eta_t \sqrt{h(\mathbf{z}_t; \delta, \theta, \pi, \gamma)} \quad \eta_t \sim iid N(0, 1) \quad (5.38)$$

In forma meno compatta,

$$\epsilon_t = \eta_t \sqrt{\delta' \mathbf{z}_t + f(\mathbf{z}_t; \theta, \gamma_1) + g(\mathbf{z}_t; \pi, \gamma_2)} \quad (5.39)$$

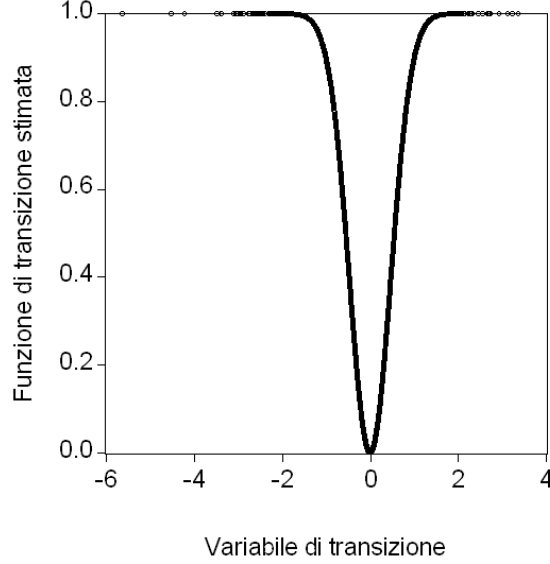
dove  $\mathbf{z}_t = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, h_{t-1}, \dots, h_{t-p})'$ ,  $f(\cdot) = \theta' \mathbf{z}_t F_1(\epsilon_{t-d}; \gamma_1)$  e  $g(\cdot) = \pi' \mathbf{z}_t F_2(\epsilon_{t-d}; \gamma_2)$ . In analogia a quanto proposto da Eitrheim e Teräsvirta (1996) per la media condizionata, la presenza di ulteriori non-linearità viene verificata aggiungendo al modello un terzo regime di volatilità, così come espresso dalla (5.39). L'ipotesi nulla può quindi essere formulata come  $H_{0,1} : \pi = 0$  oppure  $H_{0,2} : \gamma_2 = 0$ . Approssimando la funzione di transizione tramite lo sviluppo di Taylor al primo ordine nell'intorno di  $\gamma_2 = 0$ , la (5.39) si può riscrivere come

$$\epsilon_t \simeq \eta_t \sqrt{\delta'^* \mathbf{z}_t + \theta'^* \mathbf{z}_t F_1(\epsilon_{t-d}; \gamma_1) + \beta'^* \mathbf{x}_t} \quad (5.40)$$

dove  $\mathbf{x}_t$  indica i regressori ausiliari ricavati dall'approssimazione di Taylor del termine  $\pi' \mathbf{z}_t F_2(\epsilon_{t-d}; \gamma_2)$  e i parametri  $\delta$  e  $\theta$  sono stati indicati con una “\*” poiché vi confluiscono alcuni termini derivanti dall'approssimazione. In tal modo, l'ipotesi nulla di assenza di regimi aggiuntivi di volatilità può essere verificata utilizzando la statistica

$$LM = T \left( \frac{1}{T} \sum_{t=1}^T \frac{\partial L_t}{\partial \pi} \Big|_{H_0} \right)' \hat{\mathbf{I}}(\hat{\delta}, \hat{\theta}, \hat{\gamma}_1, \pi, \gamma_2)^{-1} \Big|_{H_0} \left( \frac{1}{T} \sum_{t=1}^T \frac{\partial L_t}{\partial \pi} \Big|_{H_0} \right) \quad (5.41)$$

Figura 5.4: Funzione di transizione esponenziale stimata nel modello ESTGARCH(1,1); variabile di transizione  $\epsilon_{t-1}$ .



che, sotto l'ipotesi nulla, si distribuisce asintoticamente come una  $\chi^2$  con gradi di libertà pari alla dimensione del vettore  $\pi$ . Per ulteriori dettagli e per una derivazione formale della (5.41) si rimanda a Lundbergh e Teräsvirta (1998 e 1999).

### 5.4.2 Test per la non costanza dei parametri

La verifica di una simile ipotesi si rivela particolarmente importante poiché la non costanza dei parametri nel tempo può essere causa di una apparente, seppur debole, non stazionarietà (Lamoureux e Lastrapes, 1990).

Seguendo Lin e Teräsvirta (1994) ed Eitrheim e Teräsvirta (1996), Lundbergh e Teräsvirta (1998 e 1999) sottopongono a test la costanza dei parametri contro l'alternativa di una loro variazione nel tempo di tipo *smooth transition* nel modo seguente:

$$\epsilon_t = \eta_t \sqrt{\delta'(t) \mathbf{z}_t + \theta'(t) \mathbf{z}_t F_1(\epsilon_{t-d}; \gamma_1)} \quad (5.42)$$

dove

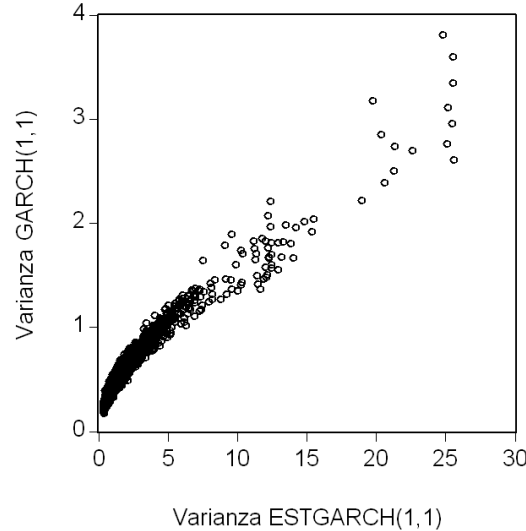
$$\delta'(t) = \delta' + \varphi_1' F_2(t; \gamma_2) \quad (5.43)$$

$$\theta'(t) = \theta' + \varphi_2' F_2(t; \gamma_2) \quad (5.44)$$

e nel quale l'ipotesi nulla può essere espressa come  $H_0 : \gamma_2 = 0$ . Anche qui, i soliti problemi di non identificazione di alcuni parametri sotto  $H_0$  vengono risolti mediante l'idea dell'approssimazione di Luukkonen *et al.* (1988b). Riscrivendo la (5.42) come

$$\epsilon_t = \eta_t \sqrt{\delta' \mathbf{z}_t + \theta' \mathbf{z}_t F_1(\epsilon_{t-d}; \gamma_1) + \varphi_1' \mathbf{z}_t F_2(t; \gamma_2) + \varphi_2' \mathbf{z}_t F_1(\epsilon_{t-d}; \gamma_1) F_2(t; \gamma_2)} \quad (5.45)$$

Figura 5.5: Scatter tra varianza condizionata stimata con i modelli ESTGARCH(1,1) e GARCH(1,1) sulla serie del tasso di cambio JPY/USD.



e applicando lo sviluppo di Taylor di  $F_2(\cdot)$  nell'intorno di  $\gamma_2 = 0$ , si ricava il modello ausiliario

$$\epsilon_t = \eta_t \sqrt{\delta' \mathbf{z}_t + \theta' \mathbf{z}_t F_1(\epsilon_{t-d}; \gamma_1) + \beta_1' \mathbf{x}_t^* + \beta_2' \mathbf{x}_t^* F_1(\epsilon_{t-d}; \gamma_1)} \quad (5.46)$$

dove  $\mathbf{x}_t^*$  indica i regressori ausiliari ricavati dall'approssimazione di Taylor dei termini  $\varphi_1' \mathbf{z}_t F_2(t; \gamma_2)$  e  $\varphi_2' \mathbf{z}_t F_2(t; \gamma_2)$ . Ancora una volta, l'ipotesi di interesse può essere verificata tramite la statistica (5.41) dove, in tal caso,  $\pi = (\varphi_1, \varphi_2)'$ . Asintoticamente, tale quantità converge ad una distribuzione  $\chi^2$  con gradi di libertà pari alla dimensione di  $\pi$  (ovvero,  $\dim\{\varphi_1\} + \dim\{\varphi_2\}$ ). Ulteriori dettagli e generalizzazioni si trovano in Lundbergh e Teräsvirta (1998 e 1999).

## 5.5 Possibili sviluppi dei modelli STGARCH

In questa Sezione presenteremo brevemente alcune semplici estensioni del modello STGARCH di base, così come era stato fatto nel Capitolo 4 per il modello STAR. In particolare saranno descritti il modello STGARCH in media (STGARCH-M), il modello a regimi multipli e quello con parametri che variano nel tempo. La discussione è finalizzata alla presentazione di quelle che possono essere le direttrici di sviluppo dei modelli STGARCH, senza tuttavia fornirne un'analisi dettagliata ed esaustiva.

### 5.5.1 Il modello STGARCH in media

Nella Sezione 2.2.6 è stato analizzato il modello ARCH-M di Engle, Lilien e Robins (1987), la cui motivazione economica risiede nella relazione rendimento-rischio

di Markowitz (1959): un investimento più rischioso deve essere necessariamente accompagnato da un più alto rendimento e viceversa, mentre a parità di rendimento è preferibile una rischiosità inferiore.

Nel corso di tutto il lavoro si è ampiamente mostrato che la struttura dinamica della maggior parte delle serie finanziarie evolve in modo non-lineare. Pertanto, è ragionevole supporre che anche la relazione rendimento-rischio sia caratterizzata da dinamiche simili. In tal modo, la specificazione STGARCH in media (STGARCH-M) assume la forma

$$y_t = \mu_t + \delta\sigma_t + \epsilon_t \quad \text{con } \sigma_t = \sqrt{h_t} \quad (5.47)$$

$$\epsilon_t = \eta_t \sqrt{h_t}, \quad \eta_t \sim iid \ N(0,1) \quad (5.48)$$

e dove  $h_t \sim \text{STGARCH}$ , ovvero rappresenta una delle formulazioni parametriche per la varianza condizionata descritte nella Sezione 5.1. Un valore di  $\delta$  statisticamente nullo indica l'assenza di effetti STGARCH-M. Tra i principali inconvenienti di questa formulazione vi è la constatazione che la verifica di effetti STGARCH in media è subordinata alla stima di una relazione dinamica non-lineare che, in quanto tale, può risultare molto dispendiosa. Questo perché, allo stato attuale, non siamo in grado di disporre di una efficiente metodologia operativa che ci consenta di verificare sui dati l'eventuale presenza di effetti STGARCH in media prima di un'eventuale stima del modello stesso.

### 5.5.2 I modelli STGARCH a regimi multipli

Nella Sezione 5.4.1 si è considerata la possibilità di un regime aggiuntivo al modello STGARCH di base. Generalizzeremo adesso questa eventualità fino ad ammettere la possibilità di  $m$  regimi di volatilità. Per semplicità, ci riferiremo al modello ANSTGARCH di Anderson *et al.* (1999), mettendo di volta in volta in evidenza le differenze che insorgono quando si considerano le altre formulazioni STGARCH. A tal fine, riscriviamo il modello in notazione matriciale,

$$h_t = \theta'_1 \mathbf{z}_t + \theta'_2 \mathbf{z}_t F(\epsilon_{t-d}, \gamma) \quad (5.49)$$

dove  $\theta_1 = (\omega_1, \alpha_{11}, \dots, \alpha_{1q}, \beta_{11}, \dots, \beta_{1p})'$ ,  $\theta_2 = (\omega_2, \alpha_{21}, \dots, \alpha_{2q}, \beta_{21}, \dots, \beta_{2p})'$  e  $\mathbf{z}_t = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2, h_{t-1}, \dots, h_{t-p})'$ . In tal modo, il modello STGARCH a tre regimi di volatilità è espresso come

$$h_t = \theta'_1 \mathbf{z}_t + \theta'_2 \mathbf{z}_t F_1(\epsilon_{t-d}, \gamma_1) + \theta'_3 \mathbf{z}_t F_2(\epsilon_{t-d}, \gamma_2) \quad (5.50)$$

con  $\theta_3 = (\omega_3, \alpha_{31}, \dots, \alpha_{3q}, \beta_{31}, \dots, \beta_{3p})'$ . I tre diversi regimi risultano indipendenti tra loro poiché si è assunta la medesima variabile di transizione nelle funzioni  $F_1(\cdot)$  e  $F_2(\cdot)$ . Se così non fosse,  $F_1(\cdot)$  e  $F_2(\cdot)$  varierebbero più o meno indipendentemente l'una dall'altra dando luogo a quattro regimi distinti. In realtà, come accadeva per il modello STAR a tre regimi proposto da Eitrheim e Teräsvirta (1996), la (5.50) è composta da tre soli modelli GARCH lineari, per cui le dinamiche tra i quattro regimi non possono essere completamente indipendenti tra loro.<sup>17</sup> Pertanto, questo modo di

<sup>17</sup>Si veda la Sezione 4.4 per considerazioni analoghe sul modello STAR.

procedere è giustificato solamente se  $F_1(\cdot)$  e  $F_2(\cdot)$  dipendono dalla stessa variabile di transizione.<sup>18</sup> La (5.50) si generalizza facilmente al caso di  $m$  regimi, mentre un modo per ottenere l'indipendenza tra i diversi stati con variabili di transizione diverse si può ricavare dalla riparametrizzazione della (5.49) nel modo seguente:

$$h_t = \phi'_1 \mathbf{z}_t (1 - F(\epsilon_{t-1}; \gamma)) + \phi'_2 \mathbf{z}_t F(\epsilon_{t-1}; \gamma) \quad (5.51)$$

e dove  $\theta_1 = \phi_1$  e  $\theta_2 = (\phi_2 - \phi_1)$ . In analogia a quanto proposto da van Dijk e Franses (1999) e van Dijk (1999) per il modello STAR, la formulazione STGARCH a quattro regimi si ricava annidando tra loro due modelli STGARCH di base:

$$h_t = [\phi'_1 \mathbf{z}_t (1 - F_1(\epsilon_{t-d_1}; \gamma_1)) + \phi'_2 \mathbf{z}_t F_1(\epsilon_{t-d_1}; \gamma_1)] [1 - F_2(\epsilon_{t-d_2}; \gamma_2)] + [\phi'_3 \mathbf{z}_t (1 - F_1(\epsilon_{t-d_1}; \gamma_1)) + \phi'_4 \mathbf{z}_t F_1(\epsilon_{t-d_1}; \gamma_1)] F_2(\epsilon_{t-d_2}; \gamma_2) \quad (5.52)$$

con  $d_1 \neq d_2$ . In tal modo, i quattro regimi risultano completamente indipendenti tra loro. Si noti che questa rappresentazione è giustificata solo per i modelli di González-Rivera (1998) e Anderson *et al.* (1999) in quanto la variabile di transizione è considerata fissa e specificata dai dati tramite la regola del minimo p-value. Per i modelli di Hagerud (1996) e Lundbergh e Teräsvirta (1998) è valida solo se si assume che la variabile di transizione sia fissa. Alla specificazione dei modelli MR-STGARCH si può applicare la stessa procedura descritta per il caso MR-STAR con gli opportuni adattamenti poiché si tratta di lavorare sui residui al quadrato ricavati dal modello per la media condizionata. Tramite l'idea dell'approssimazione (Luukkonen *et al.*, 1988b), la statistica test per la verifica della presenza di regimi multipli può essere ricavata senza difficoltà eccessive.

### 5.5.3 I modelli STGARCH con parametri variabili nel tempo

A partire dalla (5.52), il modello STGARCH con parametri che variano nel tempo (TV-STGARCH) si ricava come

$$h_t = [\phi'_1 \mathbf{z}_t (1 - F_1(\epsilon_{t-d_1}; \gamma_1)) + \phi'_2 \mathbf{z}_t F_1(\epsilon_{t-d_1}; \gamma_1)] [1 - F_2(t; \gamma_2)] + [\phi'_3 \mathbf{z}_t (1 - F_1(\epsilon_{t-d_1}; \gamma_1)) + \phi'_4 \mathbf{z}_t F_1(\epsilon_{t-d_1}; \gamma_1)] F_2(t; \gamma_2) \quad (5.53)$$

L'interpretazione del modello è quella di permettere all'evoluzione della varianza condizionata sia dinamiche non-lineari, sia la possibilità di cambiamenti strutturali nel valore dei parametri. Come caso particolare della (5.53) si può ricavare il modello TV-GARCH imponendo le restrizioni  $\phi_1 = \phi_2$  e  $\phi_3 = \phi_4$ :

$$h_t = \phi'_1 \mathbf{z}_t (1 - F_2(t; \gamma_2)) + \phi'_3 \mathbf{z}_t F_2(t; \gamma_2) \quad (5.54)$$

il quale non è altro che un GARCH lineare con parametri variabili gradualmente nel tempo. Si noti la forte analogia col modello TV-AR per la media condizionata introdotto da Lin e Teräsvirta (1994). Le fasi di specificazione, stima e verifica dei modelli TV-STGARCH e dei relativi casi particolari, possono essere ricavate applicando in

<sup>18</sup>Il problema non si pone nei modelli di Hagerud (1996) e Lundbergh e Teräsvirta (1998) (si veda la Sezione 5.1 per dettagli).



questo contesto le considerazioni della Sezione 4.5 e seguenti. Anche in tal caso, l'approccio "dallo specifico - al generale - allo specifico" sembra risultare la scelta migliore poiché non è richiesta alcuna stima non-lineare intermedia. Ulteriori analisi e verifiche in tal senso sono comunque necessarie.

## 5.6 Una strategia operativa

In questo paragrafo presenteremo una strategia operativa che può essere utilizzata per descrivere l'evoluzione delle serie storiche finanziarie mediante l'impiego dei modelli a transizione liscia. La letteratura prevalente sull'argomento si è soffermata in particolare sulla ricerca della parametrizzazione migliore per la varianza condizionata, tralasciando del tutto o in parte la specificazione della media condizionata. Tuttavia, si è visto (cfr. Sezione 3.5.5) che ci sono forti evidenze empiriche a sostegno di una più o meno accentuata dipendenza non-lineare nell'equazione che descrive il livello dei rendimenti, nonostante la quasi totale assenza di autocorrelazione. Questo dovrebbe essere esplicitamente considerato per evitare di ottenere formulazioni errate per la dinamica di  $h_t$ . Infatti, è possibile stimare consistentemente i parametri della media condizionata anche se la varianza condizionata è stata assunta costante (ipotesi di omoschedasticità). D'altra parte, non è invece possibile stimare consistentemente i parametri del modello per la varianza condizionata se  $m_t$  è mal specificata (Lundbergh e Teräsvirta, 1998).

Il nostro obiettivo è quello di costruire una coerente strategia di specificazione, stima e valutazione dei modelli STAR-STGARCH, in analogia a quanto proposto da Lundbergh e Teräsvirta (1998). I singoli punti della procedura possono essere così sintetizzati:

1. Testare la linearità della media condizionata contro l'alternativa a transizione liscia.
2. Se la linearità viene rifiutata, specificare il particolare modello STAR tramite la regola del minimo p-value.
3. Stimare i parametri del modello specificato al punto precedente sotto l'ipotesi di omoschedasticità e valutare le proprietà dei residui e della funzione di transizione stimata. In particolare, verificare la presenza di non-linearità rimanenti e di variabilità nel tempo dei parametri.
4. Verificare la presenza di eteroschedasticità condizionata di tipo ARCH e STAR-CH sulla serie dei residui al quadrato.
5. Se l'ipotesi di omoschedasticità non può essere accettata, dal confronto tra i p-value dei diversi test, valutare se un modello della classe STGARCH può risultare una valida alternativa al GARCH lineare.
6. Una volta verificata la presenza di non-linearità di tipo STGARCH nell'equazione della varianza condizionata, selezionare la funzione di transizione e, nei modelli di González-Rivera (1998) ed Anderson *et al.* (1999), il valore di  $d$ . Si

noti che a questo stadio della specificazione non è stata ancora stimata alcuna parametrizzazione per  $h_t$ ; tuttavia, l'applicazione dei test descritti, ciascuno basato su una semplice regressione lineare, ci consente di trarre significative informazioni sulla corretta specificazione del modello. In particolare, oltre a valutare l'assenza o meno di non linearità, siamo in grado di farci un'idea abbastanza precisa sull'eventuale presenza di effetti *leverage*: ciò equivale infatti alla scelta di un modello STGARCH logistico.

7. Stimare il modello STGARCH selezionato al punto precedente e valutarne le proprietà.
8. Ricavare una stima congiunta del modello STAR-STGARCH, utilizzando come valori iniziali dell'algoritmo di ottimizzazione le stime ottenute singolarmente sui due modelli.
9. Valutare le proprietà dei residui standardizzati e, se non risultano ulteriori "mis-specificazioni", utilizzare il modello per gli scopi preposti.

Naturalmente, può essere che un sottomodulo della famiglia STAR-STGARCH risulti come miglior specificazione finale. Così ad esempio, la media condizionata può essere lineare, mentre è la varianza condizionata a seguire una dinamica di tipo STGARCH, o viceversa. Al limite, sia la media che la varianza condizionata possono essere lineari.

Nella specificazione del modello per la varianza condizionata, la strategia descritta si differenzia abbastanza da quella di Lundbergh e Teräsvirta (1998). I due autori infatti, propongono di stimare un modello GARCH lineare sui residui del modello per la media condizionata e, solo successivamente in fase di valutazione, verificare la presenza di effetti STGARCH. Un simile modo di procedere può essere definito "dallo specifico - al generale" poiché si basa sulla stima di un modello più semplice per poi verificarne l'adeguatezza. Il test per ST(G)ARCH è dunque applicato alla serie dei residui standardizzati tratti dal modello STAR-GARCH (o da uno più semplice). Per contro, la procedura qui presentata può essere definita "dal generale - allo specifico" ed è giustificata sulla base delle seguenti considerazioni:

- la quasi totalità delle serie storiche finanziarie presenta forti evidenze empiriche a favore della non-linearità, tanto da giustificare l'affermazione che deve essere ritenuta come caso limite un'eventuale linearità e non il viceversa;
- la strategia proposta è basata sull'applicazione di diverse batterie di test sulla serie dei residui al quadrato ricavati dal modello per la media condizionata. Pertanto, non è necessario stimare alcuna formulazione intermedia per la varianza condizionata, ma utili informazioni sul modello più appropriato possono essere ricavate mediante semplici regressioni lineari.

Tra le principali critiche che possono essere mosse a questa procedura vi è senz'altro la considerazione che i test per ARCH e STARCH potrebbero essere sensibili a problemi di errata specificazione. In particolare, il loro fondamento teorico risiede nell'ipotesi di normalità che, come è stato più volte evidenziato, ben difficilmente

si adatta alle serie storiche finanziarie. A questo si aggiunga l'eventuale presenza di osservazioni aberranti che possono alterare, apparentemente in senso non-lineare, la struttura dinamica della varianza condizionata. Simili problematiche potrebbero essere parzialmente risolte impiegando le versioni robuste delle statistiche test (si veda la Sezione 5.2.2), anche se ulteriori verifiche e simulazioni in tal senso sono auspicabili.

## 5.7 Applicazione al Tasso sui Fondi Federali U.S.A.

In questa Sezione riprendiamo l'esempio analizzato nel Capitolo 3 al fine di ricavarne un'adeguata rappresentazione per la varianza condizionata. Dalle diagnostiche sul modello ESTAR stimato (cfr. Sezione 3.7) si è visto che sia la normalità della distribuzione empirica dei residui sia l'ipotesi di varianza condizionata costante erano ampiamente rifiutate. Pertanto, abbiamo deciso di applicare alla serie in oggetto la strategia di specificazione, stima e verifica della varianza condizionata, così come discusso nel presente Capitolo.

Tabella 5.11: p-values dei test per STARARCH (versione  $\chi^2$ ) con  $q = 4$  e 8.

Test	H	LT	GR ( $d = 1$ )	ANV ( $d = 1$ )
$q = 4$				
Congiunto	$1.58 \times 10^{-25}$	$4.46 \times 10^{-26}$	$1.87 \times 10^{-21}$	$3.16 \times 10^{-23}$
Logistico	$1.12 \times 10^{-23}$	$3.27 \times 10^{-23}$	$1.88 \times 10^{-7}$	$9.65 \times 10^{-9}$
Esponenziale	$1.23 \times 10^{-11}$	$1.23 \times 10^{-11}$	$4.78 \times 10^{-15}$	$4.78 \times 10^{-15}$
$q = 8$				
Congiunto	$5.94 \times 10^{-48}$	$8.09 \times 10^{-50}$	$2.75 \times 10^{-23}$	$2.84 \times 10^{-24}$
Logistico	$1.75 \times 10^{-41}$	$6.92 \times 10^{-39}$	$1.54 \times 10^{-12}$	$1.69 \times 10^{-13}$
Esponenziale	$5.12 \times 10^{-18}$	$5.12 \times 10^{-18}$	$5.37 \times 10^{-19}$	$5.37 \times 10^{-19}$

In primo luogo, sono stati impiegati i tests per STARARCH alla serie dei residui al quadrato, i cui risultati sono riportati nella Tabella 5.11, dove le sigle H, LT, GR ed ANV indicano rispettivamente i diversi modelli STGARCH.<sup>19</sup> I risultati dei test sono riportati per due diverse scelte di  $q$  (rispettivamente 4 e 8) e mostrano che, in entrambi i casi, l'assenza di effetti STARARCH non può essere accettata. Tuttavia, l'effettiva convenienza di un modello della classe *smooth transition* GARCH deve essere valutata confrontando i risultati ottenuti con quelli del test per ARCH di Engle (1982), riportati in Tabella 5.12.

Come si vede, in tutti i casi emerge la presenza di forti nonlinearità nell'equazione della varianza condizionata<sup>20</sup>, tali da lasciar supporre che il vero processo generatore dei dati sia non-lineare e, in particolare, un modello GARCH a transizione liscia. Dal confronto tra i p-value dei test per STARARCH è anche possibile selezionare la funzione di transizione. Poiché si è deciso di stimare un modello STGARCH(1,1), ed avendo

<sup>19</sup>H, Hagerud (1996); LT, Lundbergh e Teräsvirta (1998); GR, González-Rivera (1998); ANV, Anderson *et al.* (1999).

<sup>20</sup>I risultati non cambiano se il confronto è basato su scelte diverse di  $q$  nei test per ARCH ( $q = 8$ ) e STARARCH ( $q = 4$ ).

Tabella 5.12: Risultati del test ARCH (Engle, 1982).

$q$	4	8
p-value ( $\chi^2$ )	$3.97 \times 10^{-8}$	$2.56 \times 10^{-12}$

precedentemente dimostrato che tutti gli altri modelli STGARCH sono compresi nella formulazione di Anderson *et al.* (1999) come casi particolari, faremo riferimento al test ANV. Così sia per  $q = 4$  che per  $q = 8$ , il p-value del test nel caso esponenziale è più piccolo del corrispondente valore nel caso logistico. Pertanto, stimeremo un modello ESTGARCH(1,1) sui residui ricavati dal modello per la media condizionata. In Tabella 5.13 si riportano i risultati delle stime (tra parentesi gli *standard error* non robusti). Nel caso esponenziale,  $h_t$  risulta simmetrica rispetto alle innovazioni, per cui la procedura di stima a due passi non provoca perdita di efficienza asintotica (Engle, 1982 e Hagerud, 1996).

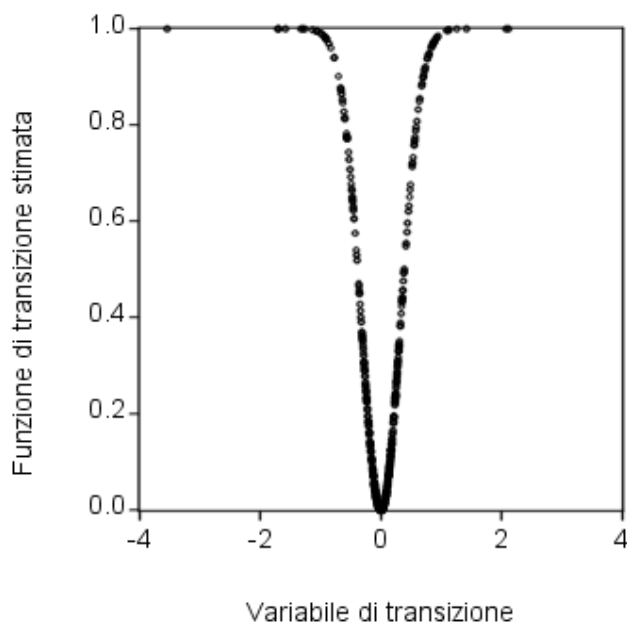
Tabella 5.13: Risultati di stima del modello ESTAGRCH(1,1) con variabile di transizione  $\epsilon_{t-1}$ .

Parametro	Stima	Std. Error
Primo Regime		
$\omega_1$	0.0002	0.0006
$\alpha_{11}$	0.9120	0.9799
$\beta_{11}$	0.6426	0.0307
Secondo Regime		
$\omega_2$	-0.1974	0.2039
$\alpha_{21}$	1.2467	0.7120
$\beta_{21}$	-0.3505	0.0974
Funzione di Transizione		
$\gamma$	1.0048	0.0326

Il grafico della funzione di transizione stimata è riportato in Figura 5.6, mentre la Figura 5.7 mostra l'andamento dei residui al quadrato (ricavati dal modello ESTAR stimato) e della varianza condizionata stimata col modello ESTGARCH(1,1). A parte una sovrastima della varianza condizionata nei periodi di alta volatilità (primi anni '80, in corrispondenza del cambiamento di politica monetaria da parte della Federal Reserve), il modello sembra descrivere bene l'evoluzione di  $h_t$  nel corso dell'arco di tempo considerato. In particolare non sembra mostrare quel grado di eccessiva persistenza tipico invece del modello GARCH tradizionale dove la stima della volatilità non si concilia bene con l'andamento delle innovazioni al quadrato.

L'analisi delle diagnostiche sulla serie dei residui standardizzati (Tabella 5.14) indica che la parametrizzazione di tipo ESTAR-ESTGARCH ottenuta, non può essere ancora accettata come modello finale. L'ampio rifiuto dell'ipotesi di normalità, dovuta ad una accentuata asimmetria a destra della distribuzione empirica, alla quale si aggiunge un indice di curtosi che va ben oltre il corrispondente valore della normale standard, rappresentano elementi importanti per concludere che il modello stimato necessita di

Figura 5.6: Funzione di transizione esponenziale stimata nel modello ESTGARCH(1,1).

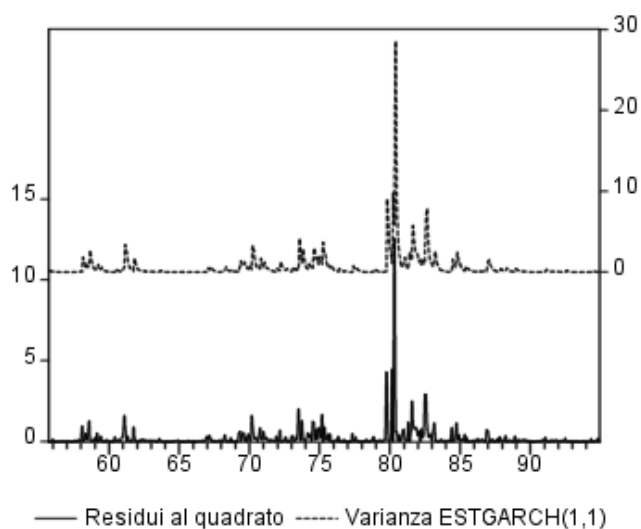


ulteriori specificazioni e verifiche. Tuttavia, non procederemo oltre nell'analisi dato il carattere meramente illustrativo con cui è stato affrontato lo studio della serie relativa al Tasso sui Fondi Federali U.S.A. I risultati presentati hanno ampiamente raggiunto lo scopo che si erano prefissi, ovvero semplicemente quello di aiuto nella comprensione delle metodologie descritte, cercando di mettere in evidenza come queste potessero essere impiegate nelle applicazioni pratiche.

## 5.8 Considerazioni conclusive

In questo Capitolo sono stati presentati i modelli GARCH a transizione liscia (STGARCH), analizzandone gli aspetti peculiari e le differenze con gli altri modelli della famiglia GARCH. In particolare, ci siamo soffermati sulla specificazione empirica dei modelli, cercando di individuare una coerente strategia operativa tale da permetterci di selezionare correttamente un modello della classe STGARCH quando realmente se ne presenti la necessità. Si è cercato di estendere la regola del minimo p-value di Teräsvirta (1994) sulla selezione del modello STAR alla scelta tra una rappresentazione GARCH lineare contro l'alternativa a transizione liscia. La regola proposta è stata investigata mediante un piccolo esperimento di simulazione i cui risultati sono abbastanza confortanti. Gli eventuali insuccessi della regola sono da ricercare tra le ipotesi teoriche (in particolare la normalità) su cui si fondano i diversi test. Per ovviare a que-

Figura 5.7: Confronto tra andamento dei residui al quadrato e varianza condizionata stimata col modello ESTGARCH(1,1).



sti inconvenienti sono state proposte delle robustificazioni da apportare ai diversi test, rimandando ad analisi successive la loro investigazione su serie simulate e reali.

La maggior parte del Capitolo è stata suddivisa secondo il tradizionale schema di specificazione, stima e valutazione dei modelli, mentre gli ultimi due paragrafi si concentrano sulla sintesi della procedura di specificazione proposta (Sezione 5.6) e su una esemplificazione reale applicata alla serie del Tasso sui Fondi Federali U.S.A. (Sezione 5.7), già utilizzata nel Capitolo 3 per mostrare le implicazioni empiriche derivanti dalla strategia di specificazione, stima e verifica dei modelli STAR sull'equazione della media condizionata. L'esempio è a scopo puramente illustrativo e non può certo essere considerato come una esauriente e completa analisi del fenomeno considerato.

Tabella 5.14: Diagnostiche sui residui standardizzati.

Statistica	Valore	p-value
Media	0.0459	
Mediana	0.0195	
Min	-5.0981	
Max	8.1921	
Deviazione standard	1.2712	
Skewness	1.0773	
Curtosi	10.954	
Jarque-Bera	1332.69	0.0000
ARCH(4)	1.683	0.7938

Interessanti sviluppi sull'argomento hanno riguardato la valutazione di previsioni da modelli STGARCH, sia in un contesto *in-sample* che *out-of-sample*, in rapporto a parametrizzazioni alternative della varianza condizionata. A tal proposito si vedano Franses e van Dijk (1996 e 1998b), Franses, Neele e van Dijk (1998) e Cecconi, Gallo e Lombardi (2002).

## Capitolo 6

# Conclusioni

Lo scopo di questa monografia è stato quello di presentare la classe dei modelli a transizione liscia (*Smooth Transition Models*) e di descrivere le principali caratteristiche che li contraddistinguono dagli altri modelli per il cambiamento di regime. Dopo una breve introduzione allo studio delle serie storiche finanziarie ci siamo soffermati sulle principali formulazioni dinamiche utilizzate in letteratura per descrivere il comportamento di molti fenomeni economici, sia per quanto riguarda la media che la varianza condizionata. Il termine “condizionata” è riferito a tutta l’informazione rilevante per la dinamica futura del fenomeno considerato. Si è operato poi un inquadramento generale dei modelli all’interno del contesto più ampio dei modelli con cambiamento di regime, evidenziandone le peculiarità e i caratteri distintivi.

L’esposizione è proseguita con l’analisi dei modelli autoregressivi a transizione liscia (STAR), strutturata secondo il tradizionale schema di specificazione, stima, valutazione ed utilizzo dei modelli proposto da Teräsvirta (1994) ed Eitrheim e Teräsvirta (1996). Nel corso della trattazione è stata presentata una esemplificazione su dati reali applicata alla serie mensile del tasso sui Fondi Federali U.S.A. Sebbene l’analisi condotta non possa considerarsi completamente esaustiva, i risultati ottenuti confermano una buona adattabilità ai dati del modello ESTAR stimato. In particolare, migliorano considerevolmente le proprietà dei residui di stima in rapporto a quelli di un modello lineare autoregressivo. Ampiamente significative risultano la riduzione della variabilità non spiegata dal modello (del 15% rispetto al modello lineare) e quelle dell’asimmetria e della curtosi della distribuzione empirica dei residui. La presenza di diversi regimi nell’evoluzione del fenomeno è confermata dai grafici della funzione di transizione stimata e della funzione di transizione contro il tempo. Il modello stimato, se da un lato rappresenta un cospicuo miglioramento rispetto alla parametrizzazione lineare, dall’altro non può essere ancora impiegato per scopi previsivi. La presenza di eteroschedasticità condizionata e, come si evince dal grafico a pagina 38, un’eventuale cambiamento strutturale in corrispondenza dei primi anni ’80 (cambiamento di politica monetaria da parte della *Federal Reserve*), dovrebbero essere maggiormente investigati prima di utilizzare il modello per gli scopi preposti.

La trattazione si sofferma poi sugli sviluppi recenti dei modelli STAR, introducendo i modelli a regimi multipli (MR-STAR) e quelli con parametri che variano nel tempo



(TV-STAR). La notazione utilizzata è leggermente diversa da quella tradizionale, ma risulta di maggiore utilità nell'interpretazione dei risultati. Si nota infatti che i modelli a transizione liscia, siano essi a due o più regimi o con parametri variabili nel tempo, possono essere visti come una media ponderata tra diverse parametrizzazioni lineari, dove il passaggio da una all'altra avviene per opera della funzione di transizione. Nello stesso Capitolo sono state anche presentate delle robustificazioni da apportare al test di linearità nei casi di presenza di eteroschedasticità generica ed *outliers*.

Nel Capitolo 5 si è analizzato il caso di eteroschedasticità condizionata autoregressiva a transizione liscia (STGARCH), mettendo in evidenza come la presenza di diversi regimi di volatilità nella dinamica dei fenomeni finanziari costituisca un aspetto di indubbia importanza. Dopo un'introduzione delle diverse formulazioni STGARCH presentate in letteratura, si è cercato di fornire una trattazione univoca dell'argomento evidenziando di volta in volta le diversità e i punti comuni a ciascun modello. In quest'ottica ne sono state discusse le proprietà e si è derivato un test di specificazione del modello basato su una semplice regressione ausiliaria. Anche l'effettiva convenienza di un modello più complicato, qual'è quello a transizione liscia, rispetto al GARCH lineare è stata opportunamente valutata presentando una regola di decisione empirica, la cui fondatezza si ritrova nella regola proposta da Teräsvirta (1994) per la specificazione dei modelli STAR. La procedura presentata è stata valutata mediante un esperimento di simulazione, i cui risultati sono abbastanza confortanti. Tra i vantaggi principali di questa strategia operativa vi è senz'altro quello di una valutazione *ex-ante* della reale convenienza all'utilizzo di un modello non-lineare, quando in realtà il vero processo generatore dei dati può essere lineare. Robustificazioni dei diversi test sono state proposte per correggere le distorsioni della regola decisionale in presenza della violazione delle ipotesi generali su cui si fondano: in particolare la non-normalità della distribuzione empirica dei rendimenti e la presenza di osservazioni eccezionalmente grandi o piccole rispetto alla media (*outliers*). Ulteriori simulazioni ed esemplificazioni su serie reali, per le quali si rimanda a ricerche successive, si rendono necessarie per valutare la potenza della regola in simili circostanze. Tutto il Capitolo è organizzato sulla base della stessa procedura di specificazione, stima e valutazione dei modelli proposta da Teräsvirta (1994) per il caso STAR. A conclusione dello stesso è presentata una esemplificazione a scopo illustrativo su dati reali, riprendendo la serie del Tasso sui Fondi Federali U.S.A. precedentemente utilizzata. Dai risultati emerge chiaramente la presenza di forti nonlinearità di tipo STGARCH, sebbene il modello finale stimato non possa ancora essere considerato definitivo.

Nel corso della trattazione si sono anche analizzate numerose serie finanziarie rappresentate da indici di borsa, tassi di cambio ed azioni quotate alla borsa di New York, alle quali sono stati applicati i test di non-linearità di tipo STAR (nella media condizionata, Sezione 3.5.5) e STGARCH (nella varianza condizionata, Sezione 5.2.1 e seguenti). I risultati mostrano una quantomeno generalizzata nonlinearità di tali fenomeni, dove la linearità può considerarsi solamente come una manifestazione isolata. Più in particolare si è analizzata la serie del tasso di cambio Yen Giapponese-Dollaro USA mediante un modello della classe STGARCH. I risultati ottenuti confermano l'ampio miglioramento nella caratterizzazione della varianza condizionata rispetto al modello GARCH lineare. Il modello riesce bene a descrivere periodi di alta e bassa volatilità, senza per questo lasciare una eccessiva persistenza nella varianza condizionata, tipica

invece del GARCH lineare.

A conclusione del lavoro possiamo affermare che i modelli a transizione liscia potrebbero costituire un'adeguata caratterizzazione dei fenomeni finanziari, soprattutto per il fatto che riescono a catturare le caratteristiche di non-linearità intrinseche a ciascuna serie ed ammettono la presenza di più regimi, o stati, dell'economia. La loro utilità è supportata anche dalla considerazione che comprendono al loro interno molte delle altre formulazioni per il cambiamento di regime come casi particolari. Pertanto, se il vero processo generatore dei dati è meglio caratterizzato da uno degli altri modelli, dalla stima di una parametrizzazione a transizione liscia ce ne dovremmo rendere conto. Un valore del parametro che governa la transizione particolarmente alto o la non significatività di alcuni parametri ci dovrebbero suggerire di restringere il modello ad una formulazione alternativa che, al limite, può essere anche quella lineare.

Tra i possibili sviluppi dei modelli non si è considerata l'ipotesi che il termine di disturbo si distribuisca come una variabile casuale  $t$ -*Student*, laddove invece estensioni in tal senso hanno costituito un miglioramento, anche apprezzabile, dei modelli stimati in ambito finanziario. Desideriamo infine segnalare i recenti sviluppi dei modelli STAR nell'ambito dell'integrazione frazionaria, *Fractionally Integrated STAR* (FISTAR), introdotti da van Dijk, Franses e Paap (2002) con applicazioni al tasso di disoccupazione statunitense, e i modelli *STARTZ*, *Smooth Transition Autoregressive Target Zone*, introdotti da Lundbergh e Teräsvirta (2003) con l'idea di caratterizzare le dinamiche dei tassi di cambio all'interno delle cosiddette *Target Zone* dove la banda di oscillazione è fissata a priori dalle autorità monetarie. Uno studio recente di Sensier, Osborn e Öcal (2002) applica i modelli della classe STAR agli effetti del tasso di interesse sul GDP inglese, mentre altri studi di Potter (1999) e Chan e McAleer (2001) forniscono rispettivamente una valutazione comparativa dei principali modelli non-lineari applicati allo studio delle serie temporali (dove i modelli STAR costituiscono una delle formulazioni principali e maggiormente impiegate), e un'analisi di tipo Monte Carlo delle stime di massima verosimiglianza ricavate da modelli STAR e STAR-GARCH. È pertanto probabile, ed anche auspicabile, che, dato il grosso successo di tali parametrizzazioni e data la loro facilità di utilizzo, numerosi saranno gli studi sull'argomento che, sia dal punto di vista teorico che applicato, si succederanno nei prossimi anni.



## Appendice A

# Dati utilizzati nelle applicazioni

Le serie storiche impiegate negli esempi sono costituite da otto indici rappresentativi delle maggiori borse internazionali, otto tassi di cambio espressi come numero di unità di moneta estera contro il dollaro americano e otto azioni quotate nei mercati americani del NYSE (*New York Stock Exchange*) e del NASDAQ (*National Association of Securities Dealers Automated Quotation*). I campioni utilizzati vanno dal 6 gennaio 1986 al 31 dicembre 1999 per le serie relative agli indici di borsa (con l'eccezione del CAC40 che parte dal 9 luglio 1987), dal 31 dicembre 1979 al 31 dicembre 1998 per i tassi di cambio e dal 2 gennaio 1990 al 29 dicembre 2000 per le azioni. I dati raccolti si riferiscono alle quotazioni di chiusura della giornata borsistica, rispettivamente 3128 osservazioni per gli indici (2735 per il CAC40), 4959 per i tassi di cambio e 2869 per le azioni, mentre le osservazioni mancanti sono state rimpiazzate dal valore medio tra il dato del giorno immediatamente precedente e quello successivo. Le serie storiche relative a indici e tassi di cambio sono le stesse utilizzate da Franses e van Dijk (2000b) e possono essere prelevate dal sito web <http://www.few.eur.nl/few/people/djvandijk/nltsmef/nltsmef.htm>. Inoltre, abbiamo fatto riferimento a Yahoo! Finance, <http://finance.yahoo.com>, per quanto riguarda i corsi azionari americani e all'*Economic and Financial Database* sul sito web della *Federal Reserve Bank* di St. Louis, <http://www.stls.frb.org/fred>, relativamente alla serie del Tasso sui Fondi Federali U.S.A. utilizzata negli esempi del Capitolo 3 e in parte del Capitolo 5.

Tabella A.1: Indici di borsa.

Piazza Finanziaria	Indice
New York	EOE
Francoforte	DAX
Parigi	CAC40
Londra	FTSE100
Hong Kong	HANG SENG
Giappone	NIKKEI
Singapore	ALL SHARES
New York	S&P500

Tabella A.2: Tassi di cambio (contro dollaro USA).

Valuta	Simbolo
Dollaro Australiano	AUD/USD
Sterlina Inglese	GPB/USD
Dollaro Canadese	CAD/USD
Marco Tedesco	DEM/USD
Fiorino Olandese	NLG/USD
Franco Francese	FRF/USD
Yen giapponese	JPY/USD
Franco Svizzero	CHF/USD

Tabella A.3: Titoli azionari.

Societ	Simbolo	Mercato
Intel Corporation	INTC	NASDAQ
AT&T Corporation	T	NYSE
Apple Computer Inc	AAPL	NASDAQ
Advanced Micro Devices (AMD) Inc	AMD	NYSE
Coca Cola Co (Coke)	KO	NYSE
McDonald's Corporation	MCD	NYSE
McGraw-Hill Inc	MHP	NYSE
PepsiCo Inc	PEP	NYSE

## Appendice B

# Derivazione del test di linearità

### B.1 Caso LSTAR: approssimazione del primo ordine

Secondo quanto indicato da Luukkonen, Saikkonen e Teräsvirta (1988b), l'approssimazione di Taylor della funzione di transizione logistica espressa dalla (3.4), nell'intorno di  $\gamma = 0$  e troncata al primo ordine è pari a:

$$T_1(y_{t-d}; \gamma, c) = F(y_{t-d}; 0, c) + g_1\gamma + R_1(y_{t-d}; \gamma, c) \quad (\text{B.1})$$

dove  $F(y_{t-d}; 0, c) = \frac{1}{2}$  e  $g_1 = \frac{\partial F(y_{t-d}; \gamma, c)}{\partial \gamma} \Big|_{\gamma=0} = \frac{1}{4}(y_{t-d} - c)$ . Sostituendo  $T_1(\cdot)$  al posto di  $F(\cdot)$  nella (3.25) e riordinando i termini, si ricava il modello ausiliario (3.28). La relazione che c'è tra i parametri del modello ausiliario e i parametri del modello LSTAR è la seguente:

$$\beta_{0,0} = \pi_0 + \theta_0 \left( \frac{1}{2} - \frac{1}{4}\gamma c \right) \quad (\text{B.2})$$

$$\beta_{0,i} = \pi_i + \theta_i \left( \frac{1}{2} - \frac{1}{4}\gamma c \right) + \frac{1}{4}\gamma\theta_0 \quad \text{per } i = d \quad (\text{B.3})$$

$$\beta_{0,i} = \pi_i + \theta_i \left( \frac{1}{2} - \frac{1}{4}\gamma c \right) \quad i = 1, \dots, p \quad \text{con } i \neq d \quad (\text{B.4})$$

$$\beta_{1,i} = \frac{1}{4}\gamma\theta_i \quad i = 1, \dots, p \quad (\text{B.5})$$

### B.2 Caso LSTAR: approssimazione del terzo ordine

In tal caso, l'approssimazione di Taylor della funzione logistica valutata nell'intorno di  $\gamma = 0$  è

$$T_3(y_{t-d}; \gamma, c) = F(y_{t-d}; 0, c) + g_1\gamma + g_3\gamma^3 + R_3(y_{t-d}; \gamma, c) \quad (\text{B.6})$$

dove  $F(y_{t-d}; 0, c) = \frac{1}{2}$ ,  $g_1 = \frac{\partial F(y_{t-d}; \gamma, c)}{\partial \gamma} \Big|_{\gamma=0} = \frac{1}{4}(y_{t-d} - c)$  e, in ultimo,  $g_3 = \left(\frac{1}{6}\right) \frac{\partial^3 F(y_{t-d}; \gamma, c)}{\partial \gamma^3} \Big|_{\gamma=0} = -\frac{1}{48}(y_{t-d} - c)^3$ . Dalla (B.6) si ricava il modello ausiliario (3.32), mentre il collegamento tra parametri del modello originario e parametri del modello ausiliario è sintetizzato dalle relazioni:

$$\beta_{0,0} = \pi_0 + \frac{1}{2}\theta_0 - \frac{1}{4}\gamma c\theta_0 \left(1 - \frac{1}{12}\gamma^2 c^2\right) \quad (\text{B.7})$$

$$\beta_{0,i} = \pi_i + \frac{1}{2}\theta_i - \frac{1}{4}\gamma c\theta_i \left(1 - \frac{1}{12}\gamma^2 c^2\right) + \frac{1}{4}\gamma\theta_0 \left(1 - \frac{1}{4}\gamma^2 c^2\right) \quad \text{per } i = d \quad (\text{B.8})$$

$$\beta_{0,i} = \pi_i + \frac{1}{2}\theta_i - \frac{1}{4}\gamma c\theta_i \left(1 - \frac{1}{12}\gamma^2 c^2\right) \quad i = 1, \dots, p \quad \text{con } i \neq d \quad (\text{B.9})$$

$$\beta_{1,i} = \frac{1}{4}\gamma\theta_i - \frac{1}{16}\gamma^3 c(c\theta_i - \theta_0) \quad \text{per } i = d \quad (\text{B.10})$$

$$\beta_{1,i} = \frac{1}{4}\gamma\theta_i \left(1 - \frac{1}{4}\gamma^2 c^2\right) \quad i = 1, \dots, p \quad \text{con } i \neq d \quad (\text{B.11})$$

$$\beta_{2,i} = \frac{1}{16}\gamma^3 \left(c\theta_i - \frac{1}{3}\theta_0\right) \quad \text{per } i = d \quad (\text{B.12})$$

$$\beta_{2,i} = \frac{1}{16}\gamma^3 c\theta_i \quad i = 1, \dots, p \quad \text{con } i \neq d \quad (\text{B.13})$$

$$\beta_{3,i} = -\frac{1}{48}\gamma^3\theta_i \quad i = 1, \dots, p \quad (\text{B.14})$$

### B.3 Caso ESTAR: relazioni tra parametri

Quando consideriamo la funzione di transizione esponenziale, è sufficiente far riferimento alla procedura di approssimazione di Taylor troncata al primo ordine, dalla quale si ricava il modello ausiliario (3.42), i cui parametri in funzione di quelli del modello originario sono espressi come:

$$\beta_{0,0} = \pi_0 + \gamma c^2\theta_0 \quad (\text{B.15})$$

$$\beta_{0,i} = \pi_i + \gamma c(c\theta_i - 2\theta_0) \quad \text{per } i = d \quad (\text{B.16})$$

$$\beta_{0,i} = \pi_i + \gamma c^2\theta_i \quad i = 1, \dots, p \quad \text{con } i \neq d \quad (\text{B.17})$$

$$\beta_{1,i} = \gamma(\theta_0 - 2c\theta_i) \quad \text{per } i = d \quad (\text{B.18})$$

$$\beta_{1,i} = -2\gamma c\theta_i \quad i = 1, \dots, p \quad \text{con } i \neq d \quad (\text{B.19})$$

$$\beta_{2,i} = \gamma\theta_i \quad i = 1, \dots, p \quad (\text{B.20})$$

**B.4 Caso LSTAR-D: relazioni tra parametri**

Relazioni tra i parametri del modello ausiliario (3.52) e modello LSTAR-D originario:

$$\beta_{0,i} = \pi_i + \frac{1}{2}\theta_i \quad i = 0, 1, \dots, p \quad (\text{B.21})$$

$$\beta_{1,i} = \frac{1}{4}\gamma\theta_i \quad i = 0, 1, \dots, p \quad (\text{B.22})$$

$$\beta_{3,i} = -\frac{1}{48}\gamma^3\theta_i \quad i = 0, 1, \dots, p \quad (\text{B.23})$$





## Appendice C

# Derivate analitiche per il test di dipendenza seriale

Presentiamo di seguito le espressioni delle derivate analitiche del test per la verifica di ulteriore dipendenza seriale nei residui di un modello STAR stimato. Il test, derivato da Eitrheim e Teräsvirta (1996), è stato presentato nella Sezione 3.7 a cui si rimanda per tutti i dettagli. Posto  $f(\mathbf{w}_t; \delta) = \pi_0 + \pi' \mathbf{w}_t + (\theta_0 + \theta' \mathbf{w}_t) F(s_t; \gamma, c)$ , per il caso LSTAR avremo:

$$\frac{\partial f(\mathbf{w}_t; \delta)}{\partial \gamma} = (1 + \exp \{-\gamma (s_t - c)\})^{-2} \exp \{-\gamma (s_t - c)\} (s_t - c) (\theta_0 + \theta' \mathbf{w}_t) \quad (\text{C.1})$$

$$\frac{\partial f(\mathbf{w}_t; \delta)}{\partial c} = \gamma (1 + \exp \{-\gamma (s_t - c)\})^{-2} \exp \{-\gamma (s_t - c)\} (\theta_0 + \theta' \mathbf{w}_t), \quad (\text{C.2})$$

mentre per il caso ESTAR le corrispondenti derivate saranno:

$$\frac{\partial f(\mathbf{w}_t; \delta)}{\partial \gamma} = \exp \{-\gamma (s_t - c)^2\} (s_t - c)^2 (\theta_0 + \theta' \mathbf{w}_t) \quad (\text{C.3})$$

$$\frac{\partial f(\mathbf{w}_t; \delta)}{\partial c} = 2\gamma \exp \{-\gamma (s_t - c)^2\} (s_t - c) (\theta_0 + \theta' \mathbf{w}_t) \quad (\text{C.4})$$

I valori di  $\gamma$ ,  $c$ ,  $\theta_0$  e  $\theta'$  riportati nelle equazioni (C.1) – (C.4), per il calcolo della regressione ausiliaria su cui si basa il test per la presenza di dipendenza seriale nei residui di un modello STAR stimato, dovranno essere opportunamente sostituiti con le rispettive stime campionarie.



# Bibliografia

- AKAIKE, H. (1974): "A New Look at Statistical Model Identification," *IEEE Transactions on Automatic Control*, AC-19, 716–723.
- ANDERSON, H. M., K. NAM, AND F. VAHID (1999): *Asymmetric Nonlinear Smooth Transition GARCH Models*, chap. 10, pp. 191–207. In Rothman, P. (editor), *Nonlinear Time Series Analysis of Economic and Financial Data*. Boston, Kluwer Academic Press.
- ANDERSON, M. K., B. EKLUND, AND J. LYHAGEN (1999): "An ARCH Robust STAR Test," Working Paper Series in Economics and Finance No. 317, Stockholm School of Economics.
- ANDREWS, D. W. K., AND W. PLOBERGER (1994): "Optimal Test when a Nuisance Parameter is Present Only Under the Alternative," *Econometrica*, 62, 1383–1414.
- AVESANI, R. G., AND G. M. GALLO (1996): "The Hamilton's Treatment of Shifts in Regimes," Background Paper for a Series of Lectures given at MAD, Université de Paris–I Panthéon Sorbonne.
- BALKE, N. S., AND T. B. FONBY (1994): "Large Shocks, Small Shocks, and Economic Fluctuations: Outliers in Macroeconomic Time Series," *Journal of Applied Econometrics*, 9, 181–200.
- BERA, A. K., AND M. L. HIGGINS (1992): "A Test for Conditional Heteroskedasticity in Time Series Models," *Journal of Time Series Analysis*, 13, 305–366.
- (1993): "A Survey of ARCH Models: Properties, Estimation and Testing," *Journal of Economic Surveys*, 7, 305–366.
- BLACK, F. (1976): "Studies in Stock Price Volatility Changes," Proceedings of the 1976 Business Meeting of the Business Economics Section, *American Statistical Association*, pp. 177–181.
- BOLLERSLEV, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327.
- BOLLERSLEV, T., R. CHOU, AND K. KRONER (1992): "ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 52, 5–59.

- BOLLERSLEV, T., R. F. ENGLE, AND D. B. NELSON (1994): *ARCH Models*, chap. 49, pp. 2959–3038. In McFadden, D. L. and Engle, R. F. (editors), *Handbook of Econometrics* Vol. 4.
- BOLLERSLEV, T., AND J. M. WOOLDRIDGE (1992): “Quasi-Maximum Likelihood Estimation and Inference in Dynamics Models with Time-Varying Covariances,” *Econometric Reviews*, 11, 143–172.
- CAI, J. (1994): “A Markov Model of Switching-Regime ARCH,” *Journal of Business and Economic Statistics*, 12, 309–316.
- CECCONI, M., G. M. GALLO, AND M. J. LOMBARDI (2002): “GARCH-based Volatility Forecasts for Market Volatility Index,” Working Paper del Dipartimento di Statistica “G. Parenti” No. 6, Università di Firenze.
- CHAN, F., AND M. MCALEER (2001): “Maximum Likelihood Estimation of STAR and STAR-GARCH Models: A Monte Carlo Analysis,” Discussion paper, Department of Economics, University of Western Australia, Working Paper, presentato a ESEM 2002, Venezia.
- CHAN, K. S., AND H. TONG (1985): “On the Use of the Deterministic Lyapunov Function for the Ergodicity of Stochastic Difference Equations,” *Advances in Applied Probability*, 17, 666–678.
- (1986): “On Estimating Thresholds in Autoregressive Models,” *Journal of Time Series Analysis*, 7, 179–190.
- CHOW, G. C. (1960): “Testing for Equality between Sets of Coefficients in Two Linear Regressions,” *Econometrica*, 28, 591–605.
- DAVIDSON, R., AND J. G. MACKINNON (1985): “Heteroschedasticity-Robust Tests in Regression Directions,” *Annales de l'INSEE*, 59/60, 183–218.
- (1993): *Estimation and Inference in Econometrics*. Oxford University Press, New York.
- (1998): “Bootstrap Test: How Many Bootstrap?,” Working Paper, Department of Economics, Queen’s University.
- DAVIES, R. B. (1977): “Hypothesis Testing when a Nuisance Parameter is Present Only Under the Alternative,” *Biometrika*, 64, 247–254.
- (1987): “Hypothesis Testing when a Nuisance Parameter is Present Only Under the Alternative,” *Biometrika*, 74, 33–43.
- DIEBOLD, F. X., J. H. LEE, AND G. C. WEINBACH (1992): *Regime Switching with Time-Varying Transition Probabilities*. In Hargreaves, C. (editor), *Nonstationary Time Series Analysis and Cointegration*, Oxford University Press, Oxford.

- DING, Z., C. W. J. GRANGER, AND R. F. ENGLE (1993): "A Long Memory Property of Stock Market Returns and A New Model," *Journal of Empirical Finance*, 1, 83–106.
- EITRHEIM, O., AND T. TERÄSVIRTA (1996): "Testing the Adequacy of Smooth Transition Autoregressive Models," *Journal of Econometrics*, 74, 59–75.
- ENGLE, R. F. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of The Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1007.
- (1990): "Discussion: Stock Market Volatility and The Crash of 87," *Review of Financial Studies*, 3, 103–106.
- ENGLE, R. F., AND T. BOLLERSLEV (1986): "Modelling the Persistence of Conditional Variances," *Econometric Review*, 5, 1–50.
- ENGLE, R. F., D. M. LILIEN, AND R. P. ROBINS (1987): "Estimating Time Varying Risk Premia in The Term Structure: The ARCH–M Model," *Econometrica*, 55, 391–407.
- ENGLE, R. F., AND V. K. NG (1993): "Measuring and Testing the Impact of News on Volatility," *Journal of Finance*, 48, 1749–1778.
- ESCRIBANO, A., P. H. FRANSES, AND D. VAN DIJK (1998): "Nonlinearities and Outliers: Robust Specification of STAR Models," Working Paper, Erasmus University, Rotterdam.
- ESCRIBANO, A., AND O. JORDÁ (1999): *Improved Testing and Specification of Smooth Transition Regression Models*, chap. 14, pp. 289–319. In Rothman, P. (editor), *Nonlinear Time Series Analysis of Economic and Financial Data*. Boston, Kluwer Academic Press.
- FAMA, E. F. (1976): *Foundations of Finance*. Basil Blackwell, Oxford.
- FILARDO, A. J. (1992): "Business Cycle Phases and Their Transitional Dynamics," Mimeo, Federal Reserve Bank of Kansas City.
- FIorentini, G., G. CALZOLARI, AND L. PANATTONI (1996): "Analytic Derivatives and The Computation of GARCH Estimates," *Journal of Applied Econometrics*, 11, 399–417.
- FORNARI, F., AND A. MELE (1996): "Modelling the Changing Asymmetry of Conditional Variances," *Economics Letters*, 50, 197–203.
- (1997): "Sign and Volatility–Switching ARCH Models: Theory and Applications to International Stock Markets," *Journal of Applied Econometrics*, 12, 49–66.
- FRANSES, P. H., J. NEELE, AND D. VAN DIJK (1998): "Forecasting Volatility with Switching Persistence GARCH Models," Working Paper, Erasmus University, Rotterdam.

- FRANSES, P. H., AND D. VAN DIJK (1996): "Forecasting Stock Market Volatility Using Nonlinear GARCH Models," *Journal of Forecasting*, 15, 229–235.
- (1997): "Do We Often Find ARCH because of Neglected Outliers?," Working Paper, Erasmus University, Econometric Institute, Rotterdam.
- (2000a): "Outlier Detection in GARCH Models," Research Report EI-9926/RV, Erasmus University, Econometric Institute, Rotterdam.
- (2000b): *Non-linear Time Series Models in Empirical Finance*. Cambridge University Press, Cambridge.
- FRANSES, P. H., D. VAN DIJK, AND A. LUCAS (1998): "Short Patches of Outliers, ARCH, and Volatility Modelling," Discussion Paper 98-057/4, Tinbergen Institute, di prossima pubblicazione su *Applied Financial Economics*, 2003.
- GALLANT, A. R. (1987): *Nonlinear Statistical Models*. John Wiley, New York.
- GALLO, G. M., AND B. PACINI (2002): *Metodi Quantitativi per i Mercati Finanziari*. Carocci Editore, Roma.
- GEWEKE, J. (1986): "Modelling the Persistence of Conditional Variances: A Comment," *Econometric Review*, 5, 57–61.
- GLOSTEN, L. R., R. JAGANNATHAN, AND D. E. RUNKLE (1993): "On the Relation between Expected Value and The Nominal Excess Return on Stocks," *Journal of Finance*, 48, 1779–1801.
- GODFREY, L. G. (1979): "Testing the Adequacy of A Time Series Model," *Biometrika*, 66, 67–72.
- GONZÁLEZ-RIVERA, G. (1998): "Smooth Transition GARCH Models," *Studies in Nonlinear Dynamics and Econometrics*, 3, 61–78.
- GRANGER, C. W. J. (1980): "Long Memory Relationships and The Aggregation of Dynamic Models," *Journal of Econometrics*, 14, 227–238.
- (1993): "Strategies for Modelling Nonlinear Time-Series Relationships," *The Economic Record*, 69, 233–238.
- GRANGER, C. W. J., AND A. P. ANDERSEN (1978): *An Introduction to Bilinear Time Series Models*. Vandenhoeck and Ruprecht, Gottingen.
- GRANGER, C. W. J., AND O. MORGENSTERN (1970): *Predictability of Stock Market Prices*. Heath, Lexington, Massachusetts.
- GRANGER, C. W. J., AND T. TERÄSVIRTA (1993): *Modelling Nonlinear Economic Relationships*. Oxford University Press, Oxford.
- GRAY, S. F. (1996): "Modeling the Conditional Distribution of Interest Rates as A Regime-Switching Process," *Journal of Financial Economics*, 42, 27–62.

- GREENE, H. (1997): *Econometric Analysis*. Prentice Hall, New Jersey, third edn.
- HAGERUD, G. E. (1996): "A Smooth Transition ARCH Models for Asset Returns," Working Paper Series in Economics and Finance No. 162, Stockholm School of Economics.
- (1997a): "Specification Tests for Asymmetric GARCH," Working Paper Series in Economics and Finance No. 163, Stockholm School of Economics.
- (1997b): "Modelling Nordic Stock Returns with Asymmetric GARCH Models," Working Paper Series in Economics and Finance No. 164, Stockholm School of Economics.
- HAGGAN, V., S. M. HERAVI, AND M. B. PRIESTLEY (1984): "A Study of The Application of State-Dependent Models in Non-Linear Time Series Analysis," *Journal of Time Series Analysis*, 5, 69–102.
- HAGGAN, V., AND T. OZAKI (1981): "Modelling Nonlinear Random Vibrations Using an Amplitude-Dependent Autoregressive Time Series Model," *Biometrika*, 68, 189–196.
- HAMILTON, J. D. (1989): "A New Approach to The Economic Analysis of Nonstationary Time Series and The Business Cycle," *Econometrica*, 57, 357–384.
- (1994): *Time Series Analysis*. Princeton University Press, Princeton.
- HAMILTON, J. D., AND R. SUSMEL (1994): "Autoregressive Conditional Heteroskedasticity and Changes in Regime," *Journal of Econometrics*, 64, 307–333.
- HAMPEL, H. R., E. M. RONCHETTI, P. J. ROUSSEEUW, AND W. A. STAHEL (1986): *Robust Statistics – The Approach Based on Influence Functions*. John Wiley, New York.
- HANSEN, B. E. (1996): "Inference When a Nuisance Parameter is Not Identified Under the Null Hypothesis," *Econometrica*, 64, 413–430.
- HENDRY, D. F. (1995): *Dynamics Econometrics*. Oxford University Press, Oxford.
- HIGGINS, M. L., AND A. K. BERA (1992): "A Class of Nonlinear ARCH Models," *International Economic Review*, 33, 137–158.
- HYNDMAN, R. J. (1995): "Highest-Density Forecast Regions for Nonlinear and Nonnormal Time Series," *Journal of Forecasting*, 14, 431–441.
- (1996): "Computing and Graphing Highest-Density Regions," *American Statistician*, 50, 120–126.
- JANSEN, E. S., AND T. TERÄSVIRTA (1996): "Testing Parameter Constancy and Super Exogeneity in Econometric Equations," *Oxford Bulletin of Economics and Statistics*, 58, 735–763.



- JARQUE, C. M., AND A. K. BERA (1980): "Efficient Test for Normality, Homoskedasticity and Serial Dependence of Regression Residuals," *Economic Letters*, 6, 255–259.
- LAMOUREUX, C. G., AND W. D. LASTRAPES (1990): "Persistence in Variance, Structural Change, and The GARCH Model," *Journal of Business and Economic Statistics*, 8, 225–234.
- LANNE, M., AND P. SAIKKONEN (2000): "Modelling the U.S. Short-Term Interest Rate by Mixture Autoregressive Processes," Working Paper, University of Helsinki, pubblicato su *Journal of Financial Econometrics*, 2003.
- LEE, T.-H., H. WHITE, AND C. W. J. GRANGER (1993): "Testing for Neglected Nonlinearity in Time Series Models," *Journal of Econometrics*, 56, 269–290.
- LEYBOURNE, S., P. NEWBOLD, AND D. VOUGAS (1998): "Unit Roots and Smooth Transitions," *Journal of Time Series Analysis*, 19, 83–97.
- LIN, C.-F., AND T. TERÄSVIRTA (1994): "Testing the Constancy of Regression Parameters against Continuous Structural Change," *Journal of Econometrics*, 62, 211–228.
- LJUNG, G. M., AND G. E. P. BOX (1978): "On a Measure of Lack of Fit in Time Series Models," *Biometrika*, 65, 297–303.
- LUNDBERGH, S., AND T. TERÄSVIRTA (1998): "Modelling Economic High-Frequency Time Series with STAR-STGARCH Models," Working Paper Series in Economics and Finance No. 291, Stockholm School of Economics.
- (1999): "Evaluating GARCH Models," Working Paper Series in Economics and Finance No. 292, Stockholm School of Economics, pubblicato su *Journal of Econometrics*, 110, pp. 417-435, 2002.
- (2000): "Forecasting with Smooth Transition Autoregressive Models," Working Paper Series in Economics and Finance No. 390, Stockholm School of Economics, pubblicato su Clements Michael P. e David F. Hendry (eds.) *A companion to Economic Forecasting*, Chapter 21, pp. 485-509.
- (2003): "A Time Series Model for an Exchange Rate in a Target Zone with Applications," Discussion paper, Stockholm School of Economics, unpublished manuscript, presentato al Symposium *New Frontiers in Financial Volatility Modelling*, Firenze, 25-27 maggio 2003.
- LUNDBERGH, S., T. TERÄSVIRTA, AND D. VAN DIJK (2000): "Time-Varying Smooth Transition Autoregressive Models," Working Paper Series in Economics and Finance No. 376, Stockholm School of Economics, pubblicato su *Journal of Business & Economic Statistics*, 21, pp. 104-121, 2003.
- LÜTKEPOHL, H., T. TERÄSVIRTA, AND J. WOLTERS (1999): "Investigating Stability and Linearity of a German M1 Money Demand Function," *Journal of Applied Econometrics*, 14, 511–525.

- LUUKKONEN, R., P. SAIKKONEN, AND T. TERÄSVIRTA (1988a): "Testing Linearity in Univariate Time Series Models," *Scandinavian Journal of Statistics*, 16, 161–175.
- (1988b): "Testing Linearity against Smooth Transition Autoregressive Models," *Biometrika*, 75, 491–499.
- MANDELBROT, B. (1963): "The Variation of Certain Speculative Prices," *Journal of Business*, 36, 394–419.
- MARKOWITZ, H. M. (1959): *Portfolio Selection: Efficient Diversification of Investments*. Wiley, New York.
- MARTIN, R. D., AND V. J. YOHAI (1986): "Influence Functionals for Time Series," *Annals of Statistics*, 14, 781–818.
- MICHAEL, P., A. R. NOBAY, AND D. A. PEEL (1997): "Transactions Costs and Non-linear Adjustment in Real Exchange Rates: An Empirical Investigation," *Journal of Political Economy*, 105, 862–879.
- (1999): *Nonlinear Adjustment Towards Long-Run Money Demand*, chap. 9, pp. 179–190. In Rothman, P. (editor), *Nonlinear Time Series Analysis of Economic and Financial Data*. Boston, Kluwer Academic Press.
- MIHOJ, A. (1987): "A Multiplicative Parametrization of ARCH Models," Working Paper, Department of Statistics, University of Copenhagen.
- MILLS, T. C. (1999): *The Econometric Modelling of Financial Time Series*. Cambridge University Press, Cambridge, seconda edn.
- NELDER, J. A. (1961): "The Fitting of A Generalisation of The Logistic Curve," *Biometrics*, 17, 89–110.
- NELSON, D. B. (1991): "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347–370.
- OZAKI, T. (1981): "Non-Linear Threshold Autoregressive Models for Non-Linear Random Vibrations," *Journal of Applied Probability*, 18, 443–451.
- PANTULA, S. G. (1986): "Modelling the Persistence of Conditional Variances: A Comment," *Econometric Review*, 5, 71–74.
- PICCOLO, D. (1990): *Introduzione all'Analisi delle Serie Storiche*. La Nuova Italia Scientifica, Roma.
- POTTER, S. M. (1999): "Nonlinear Time Series Modelling: An Introduction," *Journal of Economic Surveys*, 13, 505–528.
- PRIESTLEY, M. B. (1980): "State-Dependent Models: A General Approach to Non-Linear Time Series Analysis," *Journal of Time Series Analysis*, 1, 47–71.

- (1988): *Non-Linear and Non-Stationary Time Series Analysis*. Academic Press, London and San Diego.
- QUANDT, R. (1983): *Computational Problems and Methods*, chap. 12, pp. 699–764. In Griliches, Z. and M.D. Intriligator (editors), *Handbook of Econometrics I*, Elsevier Science, Amsterdam.
- RAMSEY, J. B., AND P. ROTHMAN (1996): “Time Irreversibility and Business Cycle Asymmetry,” *Journal of Money, Credit and Banking*, 28, 1–21.
- ROTHMAN, P. (1996): *International Evidence of Business-Cycle Nonlinearity*, chap. 14, pp. 333–341. In Barnett, W. A., Kirman, A. P. and Salmon, M. (editors), *Nonlinear Dynamics and Economics*.
- SAIKKONEN, P., AND R. LUUKKONEN (1988): “Lagrange Multiplier Tests for Non-Linearities in Time Series Models,” *Scandinavian Journal of Statistics*, 15, 55–68.
- SCHWARZ, G. (1978): “Estimating the Dimension of A Model,” *Annals of Statistics*, 6, 461–464.
- SCHWERT, G. W. (1989): “Why does Stock Market Volatility Change over Time?,” *Journal of Finance*, 44, 1115–1153.
- (1990): “Stock Volatility and The Crash of 87,” *Review of Financial Studies*, 3, 77–102.
- SENSIER, M., D. OSBORN, AND N. ÖCAL (2002): “Asymmetric Interest Rate Effects for the UK Real Economy,” Discussion Paper Series 010, Centre for Growth & Business Cycle Research, University of Manchester.
- SENTANA, E. (1991): “Quadratic ARCH Models: A Potential Reinterpretation of ARCH Models,” Discussion Paper 122, LSE Financial Market Group.
- (1995): “Quadratic ARCH Models,” *Review of Economic Studies*, 62, 639–661.
- SICHEL, D. E. (1993): “Business Cycles Asymmetry: A Deeper Look,” *Economic Inquiry*, 31, 224–236.
- SKALIN, J. (1998): “Testing Linearity against Smooth Transition Autoregression Using a Parametric Bootstrap,” Working Paper Series in Economics and Finance No. 276, Stockholm School of Economics.
- SOLLIS, R., S. LEYBOURNE, AND P. NEWBOLD (1999): “Unit Roots and Asymmetric Smooth Transitions,” *Journal of Time Series Analysis*, 20, 671–677.
- TAYLOR, S. (1986): *Modelling Financial Time Series*. Wiley, New York.
- TERÄSVIRTA, T. (1994): “Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models,” *Journal of the American Statistical Association*, 89, 208–218.

- (1996): *Linearity Testing and Nonlinear Modeling of Economic Time Series*, chap. 11, pp. 281–293. In Barnett, W. A., Kirman, A. P. and Salmon, M. (editors), *Nonlinear Dynamics and Economics*.
- TERÄSVIRTA, T., AND H. M. ANDERSON (1992): “Characterizing Nonlinearities in Business Cycles Using Smooth Transition Autoregressive Models,” *Journal of Applied Econometrics*, 7, S119–S136.
- TERÄSVIRTA, T., C.-F. LIN, AND C. W. J. GRANGER (1993): “Power of The Neural Network Linearity Test,” *Journal of Time Series Analysis*, 14, 209–220.
- THISTED, R. A. (1988): *Elements of Statistical Computing*. Chapman and Hall, New York.
- TONG, H. (1983): *Threshold Models in Non-Linear Time Series Analysis*. Springer-Verlag, Berlin.
- (1990): *Non-Linear Time Series: A Dynamical Systems Approach*. Oxford University Press, Oxford.
- TONG, H., AND K. S. LIM (1980): “Threshold Autoregression, Limit Cycles and Cyclical Data,” *Journal of Royal Statistical Society*, (B) 42, 245–292.
- TSAY, R. S. (1986): “Nonlinearity Tests for Time Series,” *Biometrika*, 73, 461–466.
- VAN DIJK, D. (1999): “Smooth Transition Models: Extensions and Outlier Robust Inference,” Ph.D. thesis, Tinbergen Institute, Erasmus University, Rotterdam.
- VAN DIJK, D., AND P. H. FRANSES (1999): “Modelling Multiple Regimes in The Business Cycle,” *Macroeconomic Dynamics*, 3, 311–340.
- VAN DIJK, D., P. H. FRANSES, AND A. LUCAS (1999a): “Testing for Smooth Transition Nonlinearity in The Presence of Outliers,” *Journal of Business and Economic Statistics*, 17, 217–235.
- (1999b): “Testing for ARCH in The Presence of Additive Outliers,” *Journal of Applied Econometrics*, 14, 539–562.
- VAN DIJK, D., P. H. FRANSES, AND R. PAAP (2002): “A Nonlinear Long Memory Model for US Unemployment,” *Journal of Econometrics*, 102, 135–165.
- VAN DIJK, D., T. TERÄSVIRTA, AND P. H. FRANSES (2000): “Smooth Transition Autoregressive Models – A Survey of Recent Developments,” Working Paper Series in Economics and Finance No. 380, Stockholm School of Economics, publicato su *Econometric Reviews*, 21, pp. 1-47, 2002.
- WHITE, H. (1982): “Maximum Likelihood Estimation of Misspecified Models,” *Econometrica*, 50, 1–25.
- WHITE, H., AND I. DOMOWITZ (1984): “Nonlinear Regression with Dependent Observations,” *Econometrica*, 52, 143–161.

WOOLDRIDGE, J. M. (1990): "A Unified Approach to Robust, Regression-Based Specification Tests," *Econometric Theory*, 6, 17–43.

——— (1991): "On the Application of Robust, Regression-Based Diagnostics to Models of Conditional Means and Conditional Variances," *Journal of Econometrics*, 47, 5–46.

ZAKOÏAN, J. M. (1994): "Threshold Heteroskedastic Models," *Journal of Economic Dynamic and Control*, 18, 931–955.

Copyright © 2005  
Massimiliano Cecconi,  
Giampiero M. Gallo