



Università Politecnica delle Marche

Dottorato in Economia Politica

XV ciclo

**Un modello con variabili esogene per la matrice delle
covarianze realizzate**

Andrea Bucci

Supervisor:

Prof. Giulio Palomba

Prof. Eduardo Rossi

Coordinatore del corso:

Prof. Riccardo Lucchetti

Indice

Introduzione	7
1 Misure di volatilità	11
1.1 Fondamenti di teoria	11
1.2 Volatilità: definizione e aspetti teorici	13
1.3 Modelli GARCH univariati	16
1.3.1 Modello GARCH	18
1.4 Modelli GARCH multivariati	18
1.4.1 VECH	19
1.4.2 BEKK	19
1.4.3 GARCH fattoriali e ortogonali	20
1.4.4 CCC e DCC	21
1.4.5 Stima dei modelli GARCH multivariati	22
1.5 Modelli di volatilità stocastica	23
1.5.1 Modelli multivariati a volatilità stocastica (MSV)	24
1.6 Varianza realizzata	25
1.6.1 Costruzione della varianza realizzata	26
1.6.2 Distribuzione della varianza realizzata	28
1.6.3 Problemi con la varianza realizzata	29
1.6.4 Covarianza realizzata	32
1.7 Modelli di varianza realizzata e previsioni	33
1.7.1 Modelli di serie storiche	33
1.7.2 Modelli MIDAS e GARCH-MIDAS	38
1.7.3 Realized GARCH e modello HEAVY	46
2 Modello di previsione della volatilità con variabili esogene	49
2.1 Modello Cholesky-VAR(p) con esogene	49
2.1.1 Modello Cholesky-VAR(p)	49
2.1.2 Stima del modello VAR e previsioni	51
2.2 Modello Cholesky-VLSTAR con esogene	53
2.2.1 Presenza di break strutturali e test di linearità	57
2.3 Analisi empirica	59

2.3.1	Determinanti macroeconomiche e finanziarie	59
2.3.2	Dati cross-country	60
2.3.3	Dati mercato americano	66
2.3.4	Risultati della stima del modello lineare cross-country	72
2.3.5	Risultati della stima del modello lineare americano	75
2.3.6	Risultati della stima del modello non lineare cross-country	77
2.3.7	Risultati della stima del modello non lineare americano	83
2.4	Previsioni	87
2.4.1	Valutazione delle previsioni	88
2.4.2	Risultati previsioni	95
3	Analisi empirica dei modelli	115
3.1	Ottimizzazione di portafoglio	115
3.1.1	Valutazione delle performance di portafoglio	115
3.1.2	Risultati	117
3.2	Gestione del rischio	123
3.2.1	Valutazione delle previsioni di VaR e CVaR	125
3.2.2	Risultati empirici	127
4	Conclusioni	133
	Bibliografia	137
A		
	Parametrizzazioni della matrice delle covarianze	151
A.1	Parametrizzazione di Cholesky	151
A.2	Parametrizzazione logaritmica di una matrice	152
B		
	Test linearità congiunto	153

Elenco delle tabelle

2.1	Statistiche descrittive Varianze e Covarianze cross-country	61
2.2	Statistiche descrittive fattori di Cholesky cross-country	62
2.3	Descrizione Variabili Esogene dataset cross-country	65
2.4	Statistiche descrittive determinanti macroeconomiche e finanziarie cross-country	65
2.5	Statistiche descrittive Varianze e Covarianze mercato americano .	67
2.6	Statistiche descrittive fattori di Cholesky mercato americano . . .	67
2.7	Descrizione Variabili Esogene mercato americano	70
2.8	Statistiche descrittive determinanti macroeconomiche e finanziarie mercato americano	71
2.9	Stima VAR dataset cross-country	74
2.10	Stima VAR dataset mercato americano	76
2.11	Test di break strutturali multipli nelle serie dei fattori di Cholesky con trimming pari a 0.1 e $M = 8$	79
2.12	Test non linearità cross-country	80
2.13	Test non linearità cross-country congiunto	81
2.14	Stima QML modello non lineare cross-country	82
2.15	Test di break strutturali multipli nelle serie dei fattori di Cholesky con trimming pari a 0.1 e $M = 8$	84
2.16	Test non linearità mercato americano	85
2.17	Stima QML modello non lineare mercato americano	86
2.18	RMSE cross-country	96
2.19	RMSE mercato americano	96
2.20	RMSE basato sulla norma di Frobenius dell'errore di previsione .	96
2.21	Mincer-Zarnowitz cross-country	98
2.22	Mincer-Zarnowitz US	99
2.23	Valutazione previsioni out-of-sample cross-country	101
2.24	Test in ambito multivariato, dataset cross-country	102
2.25	Valutazione previsioni out-of-sample mercato americano	103
2.26	Test in ambito multivariato, dataset mercato americano	104
2.27	Valutazione previsioni non lineari out-of-sample cross-country . . .	106

2.28	Test in ambito multivariato (modello non lineare), dataset cross-country	107
2.29	Valutazione previsioni non lineari out-of-sample mercato americano	108
2.30	Test in ambito multivariato (modello non lineare), dataset mercato americano	109
2.31	MCS cross-country univariato	111
2.32	MCS cross-country multivariato	112
2.33	MCS mercato americano univariato	113
2.34	MCS mercato americano multivariato	113
3.1	Performance Portafogli cross-country	120
3.2	Performance Portafogli mercato americano	122
3.3	Valutazione modelli VaR, cross-country	129
3.4	Valutazione modelli CVaR, cross-country	129
3.5	Valutazione modelli VaR, US	131
3.6	Valutazione modelli CVaR, US	131

Elenco delle figure

2.1	Autocorrelogramma delle serie dei fattori di Cholesky della matrice varianze e covarianze realizzate del dataset cross-country, RCOV.	63
2.2	Autocorrelogramma delle serie dei fattori di Cholesky della matrice varianze e covarianze realizzate del dataset sul mercato americano, RCOVUS.	68
3.1	Conditional Value at Risk	125

Introduzione

Negli ultimi decenni, la volatilità finanziaria ha assunto un ruolo importante in diversi ambiti applicativi come la gestione del rischio, la formazione del prezzo delle opzioni e l'allocazione di portafoglio. Alcuni studi si sono concentrati sull'individuazione delle determinanti della volatilità. Questa tesi si pone l'obiettivo di analizzare la relazione tra la volatilità e le determinanti macroeconomiche e finanziarie in ambito multivariato. L'analisi si basa su una procedura che prevede l'utilizzo di una misura di volatilità come la varianza realizzata, su una trasformazione della matrice delle covarianze realizzate, come la scomposizione di Cholesky, e sull'utilizzo di un modello lineare e di un modello non lineare con l'inclusione di determinanti macroeconomiche e finanziarie.

Schwert (1989), per primo, ha analizzato il collegamento tra la volatilità e lo stato dell'economia, dimostrando i movimenti contro-ciclici della volatilità rispetto al livello dell'attività economica. Più recentemente, una parte crescente della letteratura si è occupata del legame tra volatilità e variabili macroeconomiche e finanziarie. Mele (2007), in uno studio sul premio per il rischio, suggerisce di utilizzare le stesse variabili previsionali dei rendimenti per prevedere la volatilità. Paye (2012) mostra come la volatilità sia contro-ciclica e persistente, mentre la capacità previsionale di un modello con esogene risulta modesta. Christiansen *et al.* (2012) estendono il lavoro di Paye (2012), sottolineando l'utilità dell'inclusione di alcune variabili finanziarie in un modello previsionale.

In questa tesi, si utilizza una misura realizzata della volatilità. La crescente disponibilità dei dati ad alta frequenza ha permesso di porre le basi teoriche all'utilizzo di una misura non parametrica delle volatilità giornaliera, nota come varianza realizzata (si veda Andersen e Bollerslev (1998), Andersen *et al.* (2001b) e Barndorff-Nielsen e Shephard (2002)), pari alla somma dei quadrati dei rendimenti intragiornalieri. Tale misura rende osservabile la volatilità e permette l'uso di modelli di serie storiche per descriverne le dinamiche. In aggiunta, utilizzando i quadrati dei rendimenti giornalieri, è possibile una sua estensione per orizzonti temporali più lunghi.

L'uso della varianza realizzata come proxy della volatilità mensile o trimestrale permette di verificare quali variabili macroeconomiche e finanziarie determinino i movimenti della volatilità. Questo lavoro si pone l'obiettivo, sulla base delle indicazioni di Schwert (1989), Mele (2007), Paye (2012) e Christiansen *et al.* (2012), di testare la capacità previsiva di

tali variabili esogene. In particolare, si testa la capacità previsionale di alcune variabili macroeconomiche, come il tasso di crescita dell'inflazione e il tasso di crescita della produzione industriale già analizzati da Schwert (1989), e di variabili finanziarie, principalmente derivanti dal set di previsori del premio per il rischio individuato da Mele (2007).

Quando si costruisce un modello per la volatilità in ambito multivariato, per assicurare che la matrice delle covarianze prevista sia almeno semi-definita positiva e simmetrica, sono necessarie restrizioni sui parametri o una trasformazione della matrice. Partendo dal modello di Halbleib-Chiriac e Voev (2011), si fa uso della scomposizione di Cholesky della matrice delle covarianze realizzate per assicurare che le matrici previste siano definite positive.

L'effetto delle determinanti viene analizzato mediante un VAR e un modello *Vector Logistic Smooth Transition Autoregressive* (VLSTAR) sui fattori di Cholesky della matrice delle covarianze realizzate. L'uso di modelli non lineari è finalizzato ad analizzare le dinamiche asimmetriche della volatilità. Il modello *Generalized Autoregressive Conditional Heteroskedasticity* (GARCH) esponenziale di Nelson (1990) è il primo modello volto ad introdurre una componente asimmetrica nelle dinamiche della volatilità condizionale. Di recente, McAleer e Medeiros (2008) e Ghalanos e Rossi (2016) hanno utilizzato un modello regressivo *smooth transition* per modellare la volatilità in ambito univariato. Questo lavoro estende questi due ultimi approcci in ambito multivariato, facendo uso di un modello VLSTAR. Il nuovo modello non lineare ha il vantaggio di cogliere la componente asimmetrica e la persistenza di lungo periodo della volatilità, in ambito multivariato. L'analisi per la specificazione del modello prevede l'utilizzo di test sulla presenza di break strutturali, per la determinazione del numero di regimi del modello, e di test sulla linearità del modello, per determinare le variabili di transizione (o la variabile di transizione) del modello. Il modello prevede l'utilizzo di variabili esogene ritardate sia nella componente lineare che nella componente non lineare. Per confermare l'utilità dei modelli introdotti in questo lavoro, si analizzano due distinti dataset: nel primo dataset si analizza la volatilità dei rendimenti dei principali indici internazionali, mentre nel secondo dataset si analizza la volatilità dei rendimenti di diverse attività finanziarie nel mercato americano.

I modelli introdotti nel presente lavoro hanno una finalità esclusivamente previsionale. Pertanto, si analizzano nel dettaglio le performance out-of-sample delle previsioni dei modelli introdotti in questa tesi. I risultati sostengono l'ipotesi che la scomposizione di Cholesky della matrice delle covarianze realizzate e l'uso delle determinanti migliorino la capacità previsionale del modello. In particolare, in maniera contrastante con quanto

emerge dalla letteratura, è evidente la superiorità dei modelli non lineari per quanto concerne il dataset cross-country. Per una valutazione indiretta delle previsioni out-of-sample, si utilizzano le previsioni in un'ottimizzazione di portafoglio e l'uso delle previsioni della volatilità nella gestione del rischio. Dalla valutazione delle performance di portafoglio emergono risultati discordanti sulla accuratezza delle previsioni in termini di allocazione di portafoglio. Questa tesi confronta, poi, le previsioni dei *Value-at-Risk* (condizionali e non) derivanti dalle diverse specificazioni.

La tesi è organizzata come segue. Nel primo capitolo viene presentata la letteratura sui modelli per la stima della volatilità, con particolare attenzione posta sull'uso della volatilità realizzata e i suoi sviluppi. Nel secondo capitolo vengono analizzati i modelli lineare e non lineare della matrice delle covarianze e si verifica l'utilità dell'uso di variabili macroeconomiche e finanziarie esogene in termini previsionali. Si presenta, quindi, uno studio out-of-sample delle stesse e un confronto con altri modelli di riferimento. Nel terzo e ultimo capitolo è presente un'applicazione empirica dei modelli analizzati nel corso del secondo capitolo. L'ultima sezione è dedicata alle conclusioni.

1 Misure di volatilità

I modelli di volatilità rappresentano, da ormai 30 anni a questa parte, uno degli ambiti più attivi della ricerca in econometria per la finanza. Lo sviluppo di tale filone di letteratura è stato favorito anche dal massiccio utilizzo dei modelli sviluppati in ambito accademico nell'industria finanziaria.

Esistono diversi metodi per quantificare empiricamente la volatilità, alcuni parametrici, come i modelli di eteroschedasticità condizionale e quelli di volatilità stocastica, altri non parametrici, come la varianza realizzata. Questo capitolo presenta una rassegna della letteratura sui modelli di stima della volatilità ed è così diviso: nella prima parte sono definiti i concetti di volatilità nozionale e attesa; la seconda parte è dedicata alla letteratura sui modelli parametrici, basati su una forma funzionale della volatilità attesa e di quella istantanea. Nei modelli ad eteroschedasticità condizionale di tipo GARCH, e nei loro sviluppi nel multivariato, la volatilità viene definita in termini di variabili direttamente osservabili mentre i modelli di volatilità stocastica coinvolgono variabili di stato latenti; la terza parte riguarda l'approccio non-parametrico per modellare la volatilità, in particolare riguarda l'uso della volatilità realizzata e le sue applicazioni empiriche.

1.1 Fondamenti di teoria

In questa sezione vengono introdotte le ipotesi e alcuni aspetti teorici di base nella definizione delle misure di volatilità.

È noto che in un mercato aperto gli scambi di attività finanziarie avvengono ad intervalli di tempo estremamente brevi e che il processo del prezzo di tali attività possa essere modellato in tempo continuo. Pertanto, si consideri un processo univariato¹ del logaritmo del prezzo, $p(t)$, definito in uno spazio di probabilità (Ω, I, P) , che evolve in tempo continuo su un intervallo $[0, T]$, dove T è un intero; si assuma inoltre che il set informativo, $(I_t)_{t \in [0, T]} \subseteq I$, contenga l'intera storia, fino a t , dei valori del prezzo e di altre rilevanti variabili.

Sotto le assunzioni di assenza di arbitraggio e di momento primo finito, il processo del prezzo appartiene alla classe delle semimartingale speciali², come definito in Back (1991) e Shiryaev (1999). La classe delle semimartingale assume una notevole importanza in ambito

¹L'estensione nel multivariato risulta relativamente semplice. Si veda Andersen *et al.* (2003).

²Un processo X si definisce come una semimartingala speciale se può essere scritto come $X = X_0 + A + M$, dove $A_0 = M_0 = 0$, M è una martingala locale e A è un processo prevedibile e a variazione finita.

econometrico, in quanto include processi come la martingala e i processi di Lévy (per una discussione sugli aspetti probabilistici inerenti la classe di semimartingala si veda Protter (1992)). Il processo del logaritmo del prezzo $p(t)$, con media finita, è una semimartingala se può essere scomposto come somma di una componente di drift e di una martingala locale che a loro volta possono essere scomposti in una realizzazione di un processo continuo e una componente di salto

$$p(t) = p(0) + A(t) + M(t) = p(0) + A^C(t) + \Delta A(t) + M^C(t) + \Delta M(t), \quad (1.1.1)$$

dove $A(0) \equiv M(0) \equiv 0$, $A^C(t)$ e $M^C(t)$ sono le realizzazioni di un processo continuo, $\Delta A(t)$ e $\Delta M(t)$ sono le rispettive componenti di salto.

Si consideri, per $0 \leq h \leq t \leq T$, il rendimento composto continuo nell'intervallo $[t-h, t]$, dato dalla differenza

$$r(t, h) = p(t) - p(t-h); \quad (1.1.2)$$

poiché il rendimento nell'intervallo di tempo, $[0, t]$, è pari a

$$r(t) \equiv r(t, t) = p(t) - p(0), \quad (1.1.3)$$

si ha che

$$\begin{aligned} r(t, h) &= p(t) - p(0) + p(0) - p(t-h) \\ &= r(t) - (p(t-h) - p(0)) \\ &= r(t) - r(t-h). \end{aligned} \quad (1.1.4)$$

Si assume ulteriormente che il processo del prezzo sia strettamente positivo e finito, così che $p(t)$ e $r(t)$ siano ben definiti nell'intervallo $[0, T]$. Questo consente che $r(t)$ abbia solo un numero definito di salti tra $[0, T]$. Si assume inoltre che i quadrati dei processi del prezzo e quello dei rendimenti siano integrabili.

Senza perdita di generalità, si può assumere la versione continua a destra del processo, per cui

$$r(t-) \equiv \lim_{\tau \rightarrow t, \tau < t} r(\tau) \quad (1.1.5)$$

$$r(t+) \equiv \lim_{\tau \rightarrow t, \tau > t} r(\tau) \quad (1.1.6)$$

$$r(t) = r(t+) \quad a.s. \quad (1.1.7)$$

I salti nei processi cumulativi del prezzo e dei rendimenti sono dati da

$$\Delta r(t) \equiv r(t) - r(t-), \quad 0 \leq t \leq T. \quad (1.1.8)$$

Nei punti continui si ha $\Delta r(t) = 0$ e, in generale,

$$P(\Delta r(t) \neq 0) = 0 \quad t \in [0, T]. \quad (1.1.9)$$

Questo non implica necessariamente che i salti siano rari, esiste la possibilità che si presenti un numero infinito (numerabile) di salti in un intervallo discreto secondo un fenomeno definito *esplosione*.

Una conseguenza della scomposizione di una semimartingala è che il processo del rendimento può essere scritto come

$$r(t) \equiv p(t) - p(0) = \mu(t) + M(t) = \mu(t) + M^C(t) + M^J(t). \quad (1.1.10)$$

Il rendimento istantaneo viene scomposto in un processo prevedibile e a variazione finita, $\mu(t)$, e in una martingala locale, $M(t)$, la quale può essere ulteriormente scomposta in una componente locale a variazione infinita, $M^C(t)$, e in una componente di salto, $M^J(t)$. Il rendimento viene così scomposto nella componente di rendimento atteso e nella componente imprevedibile o innovazione.

1.2 Volatilità: definizione e aspetti teorici

In questa sezione si analizzano le diverse definizioni di volatilità e le relazioni che intercorrono tra di esse.

Per ogni semimartingala $X(t)$ e per una coppia di semimartingale, $X(t)$ e $Y(t)$, la variazione quadratica e la covarianza dei processi, rispettivamente $[X, X]_t$ e $[X, Y]_t$, per $t \in [0, T]$ sono definite come

$$[X, X]_t = X(t)^2 - 2 \int_0^t X(s-) dX(s) \quad (1.2.1)$$

$$[X, Y]_t = X(t)Y(t) - \int_0^t X(s-) dX(s) - \int_0^t Y(s-) dY(s), \quad (1.2.2)$$

dove l'integrale dei processi càdlàg³, $X(s-)$ e $Y(s-)$, è ben definito. Ne segue direttamente che la variazione quadratica, $[X, X]_t$, è un processo stocastico crescente.

La variazione quadratica di una semimartingala gode delle seguenti proprietà:

- i. se τ_m è una partizione di $[0, T]$, per cui $0 = \tau_{m,0} \leq \tau_{m,1} \leq \dots \leq \tau_{m,m} = T$, tale per cui $\sup_{j \geq 0} (\tau_{m,j+1} - \tau_{m,j}) \rightarrow 0$ per $m \rightarrow \infty$, si ha

$$\lim_{m \rightarrow \infty} \left\{ \sum_{j \geq 1} (X(t \wedge \tau_{m,j}) - X(t \wedge \tau_{m,j-1}))^2 \right\} \rightarrow [X, X]_t, \quad (1.2.3)$$

³Processi continui da sinistra con un limite a destra.

dove $t \wedge \tau \equiv \min(t, \tau)$ e la convergenza è uniforme in probabilità. Il processo di variazione quadratica rappresenta il sentiero di variabilità realizzato di $X(t)$ nell'intervallo $[0, t]$.

- ii. se $X(t)$ e $Y(t)$ sono semimartingale integrabili al quadrato, la covarianza tra X e Y in $[t - h, t]$ è data da

$$\text{Cov}(X(t), Y(t) \mid I_{t-h}) = E([X, Y]_t \mid I_{t-h}) - [X, Y]_{t-h}; \quad (1.2.4)$$

- iii. se la componente di variazione finita nella (1.1.1), A , è continua, allora

$$[X_i, X_j]_t = [M_i, M_j] = [M_i^C, M_j^C] + \sum_{0 \leq s \leq t} \Delta M_i(s) \Delta M_j(s). \quad (1.2.5)$$

La proprietà (iii) mostra che la variazione quadratica di processi continui a variazione finita è zero, questo implica che la componente di media sia irrilevante per la variazione quadratica. Se si assume, senza perdita di generalità, che il logaritmo del prezzo segua un processo di diffusione del tipo,

$$dp(t) = \mu(t)dt + \sigma(t)dW(t), \quad (1.2.6)$$

dove W è un processo di Wiener, $\mu(t)$ è un processo prevedibile⁴ a variazione finita e $\sigma(t)$ è un processo strettamente positivo e integrabile al quadrato, per cui

$$P\left[\int_{t-h}^t \sigma^2(s)ds < \infty\right] = 1, \quad (1.2.7)$$

allora, il rendimento composto nell'intervallo $[t - h, t]$ è dato da

$$r(t, h) = \mu(t, h) + M(t, h) = \int_{t-h}^t \mu(s)ds + \int_{t-h}^t \sigma(s)dW(s). \quad (1.2.8)$$

Ne segue che la variazione quadratica del processo è definita come

$$QV_t = [p, p]_t - [p, p]_{t-h} = \int_{t-h}^t \sigma^2(s)ds. \quad (1.2.9)$$

Ci si riferisce a questa quantità, necessaria per la definizione di varianza realizzata (si veda il paragrafo 1.6.1), con il nome di *varianza integrata*.

⁴Un processo è prevedibile al tempo t se il valore del processo è noto un istante prima di t . Esempi di processi prevedibili sono i trend deterministici e tutti i processi càdlàg.

La variazione quadratica è dunque il parametro di interesse che entra nella definizione di *volatilità nozionale*, quantificata dalla varianza realizzata. La volatilità nozionale è un'espressione naturale *ex-post* della variabilità dei rendimenti, si veda Andersen *et al.* (2000). Tale misura equivale alla volatilità integrata e dunque alla variazione quadratica che, sotto le assunzioni precedenti, nell'intervallo $[t-h, t]$, è data da

$$v^2(t, h) \equiv [r, r]_t - [r, r]_{t-h} = \int_{t-h}^t \sigma^2(s) ds. \quad (1.2.10)$$

La volatilità (o *volatilità attesa*), sia I_t il set informativo di r_t fino al tempo t , nell'intervallo $[t-h, t]$ si definisce come

$$\zeta^2(t, h) = \text{Var}(r_t | I_t) \equiv \mathbb{E} \left[\{r_t - E(r_t | I_t)\}^2 \middle| I_t \right] \quad (1.2.11a)$$

$$= \mathbb{E} \left[\left\{ \int_{t-h}^t \mu(s) ds - E \left(\int_{t-h}^t \mu(s) ds \middle| I_t \right) + \int_{t-h}^t \sigma(s) dW(s) \right\}^2 \middle| I_t \right] \quad (1.2.11b)$$

$$= \mathbb{E} \left[\left\{ \int_{t-h}^t \{\mu(s) - E(\mu(s) | I_t)\} ds \right\}^2 \middle| I_t \right] \quad (1.2.11c)$$

$$+ \mathbb{E} \left[\left\{ \int_{t-h}^t \sigma(s) dW(s) \right\}^2 \middle| I_t \right] \quad (1.2.11d)$$

$$+ 2 \mathbb{E} \left[\int_{t-h}^t \{\mu(s) - E(\mu(s) | I_t)\} ds \int_{t-h}^t \sigma(s) dW(s) \middle| I_t \right]. \quad (1.2.11e)$$

Si ha che $A_h = O_{a.s.}(B_h)$ se il rapporto A_h/B_h converge quasi sicuramente (*almost surely*) ad una costante finita, per $h \rightarrow 0$. Si ha che l'equazione (1.2.11c) = $O_{a.s.}(h^2)$, l'equazione (1.2.11d) = $\int_{t-h}^t \sigma^2(s) ds = O_{a.s.}(h)$ e che (1.2.11e) = $O_{a.s.}(h^{3/2})$. Così che

$$\text{Var}(r_t | I_{t-h}) \simeq \mathbb{E}[v^2(t, h) | I_{t-h}] = \mathbb{E}[QV(t, h) | I_{t-h}]. \quad (1.2.12)$$

Questo implica che la varianza condizionale dei rendimenti sia uguale al valore atteso condizionale della variazione quadratica, se $\mu(s) = 0$ o se $\mu(s)$ è misurabile rispetto a I_{t-h} . Tale risultato, approfondito nel paragrafo 1.6.1, garantisce che la varianza realizzata sia uno stimatore corretto della varianza condizionale.

La volatilità nozionale, così come quella attesa, non sono osservabili e possono essere quantificate empiricamente. La misurazione della volatilità può avvenire mediante la stima di modelli parametrici e mediante misure non parametriche. Le misure parametriche sono volte a modellare la volatilità attesa, $\zeta(t, h)$, attraverso diverse forme funzionali. Nei modelli ARCH, il set informativo I_{t-h} dipende dai valori dei rendimenti passati e altre variabili direttamente osservabili. Nei modelli di volatilità stocastica il set informativo,

I_{t-h} , incorpora sia i valori passati dei rendimenti che quelli di variabili di stato latenti. Le misure non parametriche di volatilità, invece, quantificano la volatilità nozionale, $v^2(t, h)$, direttamente. Rispetto alle misure parametriche, le misure non parametriche di volatilità non necessitano di una forma funzionale per il processo stocastico della martingala locale, $M(t)$, e per il processo a variazione finita, $\mu(t)$, nella scomposizione dei rendimenti.

1.3 Modelli GARCH univariati

In questo lavoro vengono trattati i modelli parametrici in tempo discreto, come i modelli della famiglia ARCH (AutoRegressive Conditional Heteroskedasticity) e i modelli di volatilità stocastica. Nella sezione corrente vengono presentati i modelli della classe ARCH.

Il successo dei modelli ARCH è dovuto all'ampia applicazione degli stessi in finanza, nello specifico in problemi di asset allocation e risk management. Questo è stato possibile grazie alla capacità di questi modelli di cogliere alcuni fatti stilizzati circa i rendimenti, come la loro non prevedibilità, la persistenza dei loro quadrati, la presenza di code spesse e i clustering di volatilità nella loro distribuzione. Questo aspetto ha fatto in modo che l'attenzione dei ricercatori si sia spostata dal momento primo condizionale dei rendimenti ai momenti condizionali di ordine superiore. In particolare, la classe di modelli ARCH si concentra sulla variabilità del momento secondo condizionale dei rendimenti, definito come

$$\zeta^2(t, h) = E \left[\left(r(t, h) - E[\mu(t, h) | I_{t-h}] \right)^2 \middle| I_{t-h} \right], \quad (1.3.1)$$

equivalente alla volatilità attesa del paragrafo 1.2.

Per spiegare la variabilità del momento secondo condizionale, Engle (1982) ha introdotto il modello Autoregressive Conditional Heteroskedasticity (ARCH). Tale modello specifica l'errore in una regressione lineare sui rendimenti y_t di una qualche variabile finanziaria, per cui

$$y_t = x_t' b + \varepsilon_t.$$

Si assume che valgano le assunzioni di Gauss-Markov, in particolare

$$E[\varepsilon_t | I_{t-1}] = 0,$$

dove ε_t è l'innovazione al tempo t , x_t è un insieme di variabili esplicative esogene e b è il vettore dei coefficienti. Inoltre, per la legge dei valori attesi iterati, risulta:

$$E \left[E(\varepsilon_t | I_{t-1}) \right] = E(\varepsilon_t) = 0.$$

L'innovazione ε_t è definita come

$$\varepsilon_t = u_t h_t^{1/2} \quad (1.3.2)$$

dove $u_t \sim i.i.d.(0, 1)$ è un processo standardizzato con media nulla e varianza unitaria e h_t è la varianza condizionale dell'innovazione. Si ipotizza che

$$Cov(\varepsilon_t \varepsilon_{t+k}) = 0.$$

Mentre la varianza non condizionale di tale processo è costante, è possibile che ci sia una variabilità nella varianza condizionale di ε_t :

$$h_t = E[\varepsilon_t^2 | I_{t-1}] = Var(\varepsilon_t | I_{t-1}). \quad (1.3.3)$$

Per cui la distribuzione condizionale dell'errore risulta essere

$$\varepsilon_t | I_{t-1} \sim N(0, h_t).$$

Engle (1982) specifica la varianza condizionale come una funzione lineare dei quadrati dei valori passati di ε_t , per cui un modello per h_t diventa

$$h_t = w + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \quad (1.3.4)$$

dove w è il termine costante e α_i è il coefficiente associato a ε_{t-i}^2 .

Per assicurare la positività della varianza è necessario che ogni $\alpha_i \geq 0$ per ogni ritardo $i = 1, 2, \dots, q$ e che $w \geq 0$.

Il processo ARCH(q) può essere riscritto secondo una rappresentazione MA(q) per i quadrati delle innovazioni, per cui

$$h_t = w + A(L)\varepsilon_t^2$$

è un processo debolmente stazionario se le radici del polinomio $1 - A(L)$ sono esterne al cerchio unitario, dove $A(L) = \alpha_1 L + \alpha_2 L^2 + \dots + \alpha_q L^q$ è il polinomio nell'operatore ritardo. La condizione necessaria e sufficiente affinché ci sia stazionarietà del processo è quindi che

$$\sum_{i=1}^q \alpha_i < 1.$$

La varianza non condizionale diventa per cui:

$$Var(\varepsilon_t) = E(\varepsilon_t^2) = \frac{w}{1 - \sum_{i=1}^q \alpha_i} = \frac{w}{1 - A(1)}. \quad (1.3.5)$$

1.3.1 Modello GARCH

Il modello Generalized Autoregressive Conditional Heteroskedasticity (GARCH), proposto da Bollerslev (1986), generalizza il modello ARCH introducendo una componente autoregressiva della varianza condizionale. La specificazione della varianza condizionale, al set informativo I_{t-1} sotto le medesime ipotesi per il modello ARCH, diventa

$$h_t = w + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \quad (1.3.6)$$

con $w \geq 0$, $\alpha_i \geq 0$ per $i = 1, 2, \dots, q$ ritardi e $\beta_j \geq 0$ per $j = 1, 2, \dots, p$ ritardi. La varianza condizionale diventa funzione di p ritardi della varianza condizionale stessa e q ritardi della componente ε_t^2 , cogliendo così gli effetti di breve termine legati all'evoluzione della variabile considerata e quelli di lungo legati alla persistenza della volatilità. Se si riscrive il modello utilizzando l'operatore ritardo si ha $h_t = A(L)\varepsilon_t^2 + B(L)h_t$ dove $A(\cdot)$ e $B(\cdot)$ sono i polinomi dei ritardi rispettivamente di ε_t^2 e di h_t di ordine $q - i$ e $p - j$, affinché il processo GARCH(p, q) risulti stazionario in covarianza è necessario che le radici del polinomio $1 - A(L) - B(L)$ cadano al di fuori del cerchio unitario, per cui

$$\sum_i^q \alpha_i + \sum_j^p \beta_j < 1.$$

La varianza non condizionale diventa così

$$E(\varepsilon_t^2) = \frac{w}{1 - \sum_i^q \alpha_i - \sum_j^p \beta_j} = \frac{w}{1 - A(1) - B(1)}.$$

Infine, imponendo la condizione $\beta_1 = \dots = \beta_p = 0$ si torna al modello ARCH di Engle.

Dai lavori di Engle (1982) e Bollerslev (1986) sono state prodotte numerose varianti ed estensioni dei modelli ARCH e GARCH, si veda Bollerslev (2009) per una rassegna dettagliata di tali modelli.

1.4 Modelli GARCH multivariati

I modelli ARCH e GARCH univariati rappresentano la base teorica per lo studio delle dinamiche della volatilità, tuttavia, in molteplici contesti, risulta più utile analizzare la volatilità in ambito multivariato.

In particolare, se y_t è un vettore di n componenti e ε_t è un vettore contenente n innovazioni con media nulla, dato il set informativo I_{t-1} si assume che

$$\varepsilon_t = H_t^{1/2} u_t \quad (1.4.1)$$

dove H_t è la matrice $n \times n$ delle covarianze condizionali e u_t è un vettore tale per cui $E(u_t u_t') = I_n$. Quello che cambia nelle varie specificazioni dei modelli multivariati è la definizione della matrice H_t . La prima classe di modelli analizzati specifica la matrice H_t direttamente; di questa classe fanno parte i modelli VECHE e BEKK.

1.4.1 VECHE

Il VECHE, dal lavoro di Bollerslev, Engle e Wooldridge (1988), è una generalizzazione nel multivariato del modello GARCH (1.3.6), infatti

$$vech(H_t) = c + \sum_{i=1}^q A_i vech(\varepsilon_{t-i} \varepsilon_{t-i}') + \sum_{j=1}^p B_j vech(H_{t-j}) \quad (1.4.2)$$

dove $vech(\cdot)$ è un operatore che trasforma in vettore una matrice simmetrica considerando solo la sua parte triangolare bassa. Poiché c è un vettore $n(n+1)/2 \times 1$ e A_i e B_j sono matrici $n(n+1)/2 \times n(n+1)/2$, il numero di parametri da stimare ammonta a $(p+q)[n(n+1)/2]^2 + n(n+1)/2$, se il numero di asset non è particolarmente limitato possono emergere alcuni problemi numerici nella stima. In aggiunta, non è possibile assicurare che la matrice delle covarianze sia almeno semi-definita positiva senza l'imposizione di vincoli.

Una versione ristretta del modello VECHE è stata introdotta da Bollerslev *et al.* (1988) assumendo che A_i e B_j siano matrici diagonali. Come dimostrato da Bollerslev *et al.* (1994), si può assicurare che H_t sia definita positiva per ogni t . Il modello GARCH(p,q) diagonale ha il vantaggio di ridurre il numero di parametri da stimare a $(p+q+1)n(n+1)/2$, ma la sua struttura è troppo restrittiva, in quanto non sono permesse interazioni tra le differenti varianze e covarianze condizionali.

1.4.2 BEKK

Il modello Baba-Engle-Kraft-Kroner (BEKK), formalizzato da Engle e Kroner (1995), garantisce matrici delle covarianze condizionali definite positive grazie alla forma quadratica. Il modello diventa

$$H_t = CC' + \sum_{i=1}^q \sum_{k=1}^K A'_{ki} \varepsilon_{t-i} \varepsilon_{t-i}' A_{ki} + \sum_{j=1}^p \sum_{k=1}^K B'_{kj} H_{t-j} B_{kj} \quad (1.4.3)$$

dove A_{ki} , B_{kj} sono matrici $n \times n$ simmetriche e non negative e C è una matrice triangolare bassa $n \times n$, K sono i gradi di generalità. Il BEKK risulta stazionario in covarianza se e

solo se gli autovalori della matrice

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

sono minori di 1 in modulo, dove \otimes rappresenta il prodotto di Kronecker. Nel caso in cui $K > 1$, si presenta un problema di identificazione poiché esistono diverse parametrizzazioni che portano alla stessa rappresentazione del modello (cfr. Silvennoinen e Teräsvirta (2008)).

1.4.3 GARCH fattoriali e ortogonali

Esiste, poi, una classe di modelli che cerca di minimizzare il numero di parametri da stimare che è quella dei modelli fattoriali. Questi modelli si basano sull'idea, ispirata dal modello CAPM (Capital Asset Pricing Model) di Sharpe (1964), che poche forze comuni guidino le dinamiche della volatilità degli asset. Il primo modello GARCH fattoriale è stato introdotto da Engle, Ng e Rothschild (1990), il quale prende spunto dall'Arbitrage Price Theory di Ross (1976). Il modello GARCH fattoriale presuppone che la matrice H_t sia generata da un numero k di fattori f_t in modo tale che

$$H_t = \Omega + \sum_{k=1}^K w_k w_k' f_{k,t}, \quad (1.4.4)$$

in cui Ω è una matrice $n \times n$ semidefinita positiva, $f_{k,t}$ sono i fattori e w_k , con $k = 1, \dots, K$, sono i vettori linearmente indipendenti dei pesi. Il modello assume che i fattori siano correlati. Risulta, però, difficile distinguere gli effetti dei singoli fattori. In aggiunta, risulta difficoltoso individuare i fattori stessi, motivo per cui è scarsa la diffusione di analisi empiriche con modelli fattoriali.

Per ovviare a questi problemi si è ricorso all'uso delle componenti principali. Le stime vengono condotte partendo dalla scomposizione invertibile W , tale per cui

$$y_t = W z_t. \quad (1.4.5)$$

Le osservazioni originali risultano essere, dunque, una combinazione lineare dei fattori non osservabili z_t . I modelli derivanti da questa specificazione prendono il nome di Orthogonal (O-) GARCH di Alexander e Chibumba (1996) e Generalized Orthogonal (GO-) GARCH di Van der Weide (2002). L'ipotesi alla base dei modelli è che la matrice ortogonale W sia costante nel tempo e invertibile e che i fattori siano condizionatamente eteroschedastici e seguano dei processi del tipo GARCH.

1.4.4 CCC e DCC

La terza e ultima classe di specificazioni multivariate si basa sulla scomposizione della matrice delle covarianze condizionali nelle correlazioni e nelle deviazioni standard condizionali. Il primo modello di questa letteratura è rappresentato dal Constant Conditional Correlation (CCC) di Bollerslev (1990)

$$H_t = D_t^{1/2} R D_t^{1/2} \quad (1.4.6)$$

dove D_t è una matrice diagonale $n \times n$ con le varianze condizionali h_{iit} poste sulla diagonale e R rappresenta la matrice delle correlazioni condizionali che ha come elementi i coefficienti di correlazione $\rho_{ijt} = h_{ijt}(h_{iit}h_{jtt})^{-1/2}$; nel caso del modello a correlazioni costanti si presuppone che tali elementi non varino nel tempo⁵.

Tenendo costanti le correlazioni, gli elementi di H_t , $h_{ijt} = \rho_{ij}(h_{iit}h_{jtt})^{1/2}$, variano nel tempo solo per effetto delle variazioni delle singole varianze condizionali. Inoltre, la matrice H_t è definita positiva per ogni t poichè R è definita positiva e costante e le varianze sulla diagonale di D_t sono positive per costruzione.

Si suppone, infine, che le varianze condizionali poste sulla diagonale di D_t seguano un processo GARCH univariato, per cui

$$\begin{cases} h_{iit} = w_i + \sum_{r=1}^q \alpha_{ir} \varepsilon_{it-r}^2 + \sum_{r=1}^p \beta_{ir} h_{iit-r} \\ h_{ijt} = \rho_{ij} (h_{iit} h_{jtt})^{1/2}. \end{cases} \quad (1.4.7)$$

Questo modello porta il vantaggio di avere una matrice H_t definita positiva e di ridurre il numero di parametri da stimare a $n(1+p+q) + n(n+1)/2$. L'imposizione delle correlazioni costanti nel tempo è, tuttavia, poco adacente dal punto di vista empirico. Si è reso necessario introdurre una componente dinamica nelle correlazioni, dando forma al modello a due stadi Dynamic Conditional Correlation (DCC) di Engle e Sheppard (2001) e di Engle (2002). Il modello diventa

$$H_t = D_t^{1/2} R_t D_t^{1/2}, \quad (1.4.8)$$

con R_t variabile nel tempo. Nel primo stadio si stimano le varianze condizionali che seguono un processo GARCH(p,q)

$$h_{iit} = \alpha_{0i} + \sum_{s=1}^{Q_i} \alpha_s y_{it-s}^2 + \sum_{s=1}^{P_i} \beta_s h_{iit-s}, \quad (1.4.9)$$

⁵Si noti infatti che la matrice R non è indicizzata nel tempo.

dove Q_i e P_i rappresentano il numero di ritardi del GARCH di ogni serie dei rendimenti. Il passo successivo alla stima della matrice D_t è la stima delle correlazioni dinamiche per cui

$$Q_t = \left(1 - \sum_{r=1}^q a_r - \sum_{r=1}^p b_r\right) \bar{Q} + \sum_{r=1}^q a_r u_{t-r} u'_{t-r} + \sum_{r=1}^p b_r Q_{t-r} \quad (1.4.10)$$

$$R_t = \tilde{Q}_t^{-1} Q_t \tilde{Q}_t^{-1} \quad (1.4.11)$$

dove \tilde{Q}_t è una matrice la cui diagonale contiene le radici quadrate degli elementi della diagonale di Q_t , $\bar{Q} = E(u_t u'_t)$ è la matrice non condizionale delle correlazioni di y_t e $u_{it} = y_{it}/h_{it}^{1/2}$ sono i residui standardizzati.

Il modello DCC presenta, dunque, il notevole vantaggio della riduzione del numero di parametri a $p + q + n + \sum_{i=1}^n (P_i + Q_i)$ e di tenere conto della variabilità nel tempo delle correlazioni. Lo svantaggio principale è quello di avere un unico processo per tutte le correlazioni condizionali.

1.4.5 Stima dei modelli GARCH multivariati

La procedura per la stima dei parametri dei modelli GARCH si basa sulla funzione di verosimiglianza. Sia $\{y_t : t = 1, 2, \dots\}$ una sequenza di vettori $n \times 1$ di variabili casuali y_t , si suppone che i primi due momenti condizionali siano

$$\begin{aligned} E(y_t | I_{t-1}) &= \mu_t(\theta) \\ \text{Var}(y_t | I_{t-1}) &= H_t(\theta) \end{aligned}$$

dove $\theta \in \Theta$ e Θ è un sottoinsieme di R^p . La log-verosimiglianza corrispondente è data da

$$\ell_T = \sum_{t=1}^T \ell_t(\theta)$$

dove T è il numero di osservazioni. Si ha dunque

$$\begin{aligned} \log \ell_t(\theta) &= -\frac{1}{2} \log \det H_t(\theta) - \frac{1}{2} \left[y_t - \mu_t(\theta) \right]' H_t(\theta) \left[y_t - \mu_t(\theta) \right] \\ &= -\frac{1}{2} \log \det H_t(\theta) - \frac{1}{2} \varepsilon_t(\theta)' H_t^{-1} \varepsilon_t(\theta). \end{aligned}$$

Lo stimatore di quasi-massima verosimiglianza si ottiene massimizzando $\ell_T(\theta)$, esaminando le condizioni del primo ordine per una massimizzazione:

$$\frac{\partial \ell_T(\hat{\theta}_T)}{\partial \theta} = \sum_{t=1}^T \frac{\partial \log \ell_t(\hat{\theta}_t)}{\partial \theta} = 0.$$

Se i momenti condizionali $\mu_t(\theta)$ e $H_t(\theta)$ sono differenziabili rispetto a θ e se $H_t(\theta)$ è una matrice non-singolare con probabilità uno per ogni $\theta \in \Theta$, allora il gradiente della t -esima osservazione è pari a:

$$s_t(\theta)' = \frac{\partial \mu_t(\theta)}{\partial \theta} H_t^{-1}(\theta) \varepsilon_t(\theta) + \frac{1}{2} \frac{\partial H_t(\theta)'}{\partial \theta} \left[H_t^{-1}(\theta) \otimes H_t^{-1}(\theta) \right] \text{vec} \left[\varepsilon_t(\theta) \varepsilon_t(\theta)' - H_t(\theta) \right].$$

Sotto alcune condizioni di regolarità sulle varianze condizionali del gradiente e sulla stazionarietà del processo osservato, si può stabilire che lo stimatore di quasi-massima verosimiglianza esiste asintoticamente, è consistente e asintoticamente normale (Bollerslev e Wooldridge (1992)).

1.5 Modelli di volatilità stocastica

Un'alternativa ai modelli di tipo GARCH per la stima della volatilità è rappresentata dai modelli di volatilità stocastica. Quello che differenzia i modelli di volatilità stocastica dai modelli ARCH è che il set informativo, I_{t-h} , non è direttamente osservabile rispetto al tempo $t - h$. Come risultato si ha che la volatilità latente sia guidata da un proprio processo stocastico sottostante. In questa sezione si presenta una rapida rassegna dei principali modelli a volatilità stocastica, per una analisi più completa della letteratura collegata si rimanda ai lavori di Taylor (1994), Shephard (1996) e Ghysels *et al.* (1996).

Si presuppone, dunque, che la volatilità segua un processo stocastico v_t e che il prezzo di un titolo rischioso segua un processo del tipo

$$dS_t = \mu S_t dt + \sqrt{v_t} S_t dW_t \quad (1.5.1)$$

dove W_t è un moto Browniano con media zero e varianza unitaria, μ è la componente costante di drift del prezzo S_t . L'equazione sottostante la volatilità stocastica varia da modello a modello, il caso generale è rappresentato da

$$dv_t = \alpha_{s,t} dt + \beta_{s,t} dB_t \quad (1.5.2)$$

dove $\alpha_{s,t}$ e $\beta_{s,t}$ sono funzioni di v_t . La maggiore differenza con i modelli GARCH è che, condizionatamente al set informativo I_{t-1} , la volatilità v_t non è nota ma è una variabile casuale non osservabile (cfr. Bauwens *et al.* (2012)).

Heston (1993), per primo, ha proposto un modello in cui la volatilità del prezzo di un titolo non sia indipendente dalle dinamiche del prezzo. Il processo stocastico alla base della volatilità diventa

$$dv_t = k(\theta - v_t)dt + \sigma \sqrt{v_t} dB_t \quad (1.5.3)$$

dove θ è la volatilità media di lungo periodo, k è il tasso al quale la volatilità torna verso il suo valore medio di lungo periodo, σ è la volatilità del processo della volatilità e dB_t è un processo gaussiano a media zero. I due processi, dW_t e dB_t , sono correlati ad un tasso costante ρ .

Quello di Heston (1993) rappresenta il modello a volatilità stocastica più importante, esistono, tuttavia, diversi modelli alternativi:

- modello CEV (constant elasticity of variance), la relazione tra volatilità e prezzo è sintetizzata dal modello

$$dS_t = \mu S_t dt + \sigma S_t^{\beta/2} dW_t \quad (1.5.4)$$

dove dW è un processo di Wiener, σ è una costante positiva e β è definito come il parametro CEV. La grandezza β influenza la direzione e l'entità dell'impatto del prezzo dell'attività sulla volatilità. Tale modello non presenta un processo separato per la volatilità, si parla di modello di volatilità locale;

- modello Chen (1996), le dinamiche che guidano i tassi di interesse, v_t , derivano dal sistema

$$dr_t = (\theta_t - \alpha_t)dt + \sqrt{r_t}\sigma_t dW_t \quad (1.5.5)$$

$$d\alpha_t = (\zeta_t - \alpha_t)dt + \sqrt{\alpha_t}\sigma_t dW_t \quad (1.5.6)$$

$$d\sigma_t = (\beta_t - \sigma_t)dt + \sqrt{\sigma_t}\eta_t dW_t; \quad (1.5.7)$$

- modello SABR (Stochastic Alpha, Beta, Rho), ha lo scopo di riprodurre le dinamiche della volatilità sul mercato dei derivati. Le equazioni del modello sono:

$$dS_t = \sigma_t S_t^\beta dW_t \quad (1.5.8)$$

$$d\sigma_t = \alpha\sigma_t dz_t, \quad (1.5.9)$$

con W_t e z_t che rappresentano due processi di Wiener correlati.

1.5.1 Modelli multivariati a volatilità stocastica (MSV)

Esistono modelli di volatilità stocastica anche in ambito multivariato. Considerando un vettore dei logaritmi dei prezzi $S = (S_1, \dots, S_n)$ di n asset con $y = (y_1, \dots, y_n)$ vettore dei rendimenti, il modello per il vettore S può essere espresso come

$$dS_t = H_t^{1/2} dW_t \quad (1.5.10)$$

$$df[\text{vech}(H_t)] = a[\text{vech}(H_t)]dt + b[\text{vech}(H_t)]dB_t \quad (1.5.11)$$

dove W_t e B_t sono due vettori di moti Browniani, H_t è la matrice delle covarianze istantanea e f , a e b sono funzioni note.

Da questa specificazione si ricava il modello MSV generico in tempo discreto

$$y_t = H_t^{1/2} \varepsilon_t \quad (1.5.12)$$

$$\varepsilon_t \sim N(0, I_n) \quad (1.5.13)$$

$$u_{t-1} \sim N(0, \Sigma_u) \quad (1.5.14)$$

$$f[\text{vech}(H_t)] = a[\text{vech}(H_{t-1})] + f[\text{vech}(H_{t-1})] + b[\text{vech}(H_{t-1})]u_{t-1}, \quad (1.5.15)$$

con $y_t = S_t - S_{t-1}$.

Nell'equazione generica l'assenza di vincoli non garantisce che H_t sia definita positiva; diversi lavori hanno cercato di ovviare a questo problema.

Il primo lavoro della letteratura MSV è quello di Harvey *et al.* (1994)

$$\varepsilon_t = H_t^{1/2} z_t \quad z_t \sim N(0, \Sigma_t) \quad (1.5.16)$$

$$H_t = \text{diag}(\exp(h_{1t}), \dots, \exp(h_{Nt})) \quad (1.5.17)$$

$$h_{t+1} = w + \beta \odot h_t + u_t \quad u_t \sim N(0, \Sigma_u) \quad (1.5.18)$$

dove $h_t = (h_{1t}, \dots, h_{Nt})$ è il vettore di volatilità al tempo t , Σ_t è la matrice delle correlazioni, \odot è l'operatore prodotto di Hadamard, w e β sono vettori di parametri. Risulta un modello stringente per il fatto di avere delle correlazioni costanti nel tempo, infatti presenta delle analogie con il CCC di Bollerslev (1990).

Al modello di Harvey sono seguite diverse formulazioni per tenere conto delle dinamiche delle correlazioni, degli effetti leverage, delle distribuzioni con code spesse delle innovazioni, si rimanda alla revisione della letteratura multivariata condotta da Asai, McAleer e Yu (2006) e al lavoro di Andersen (2009) per un approfondimento sui modelli in tempo continuo.

1.6 Varianza realizzata

I modelli di eteroschedasticità condizionale e quelli di volatilità stocastica rappresentano gli approcci più comuni per misurare la volatilità. Tali approcci dipendono fortemente dalla specificazione del modello sottostante alla volatilità e necessitano di forti assunzioni per la stima dei parametri.

La ricerca di un modello adeguato per la volatilità ha spostato l'attenzione dei ricercatori sui dati ad alta frequenza. Merton (1980), per primo, ha evidenziato come la varianza

possa essere espressa come la somma dei rendimenti al quadrato se i dati sono disponibili a frequenze molto alte. Recentemente, Andersen, Bollerslev, Diebold e Ebens (2001a) e Andersen, Bollerslev, Diebold e Labys (2001b) hanno mostrato che sommando i quadrati dei rendimenti intragiornalieri si ottiene una misura osservabile della volatilità giornaliera definita varianza realizzata (*realized variance*, RV). La volatilità *ex post*, al netto di errori di misura e salti, diventa "osservabile" e può essere modellata direttamente, piuttosto che essere trattata come una variabile latente. In aggiunta, sebbene le varianze realizzate giornaliere siano distribuite in maniera non-normale, i loro logaritmi sono approssimativamente normali (Andersen *et al.* (2001a)). La meccanica per il calcolo della varianza realizzata è piuttosto semplice, ma le implicazioni teoriche alla base sono profonde e necessitano di essere analizzate nel dettaglio.

1.6.1 Costruzione della varianza realizzata

L'idea alla base della teoria della varianza realizzata è che la volatilità realizzata diventa la migliore approssimazione della volatilità non osservabile quando i rendimenti giornalieri sono campionati con frequenze sufficientemente alte. Si suppone che il logaritmo del prezzo di uno specifico asset, $p(t)$, segua un processo di diffusione

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) \quad \text{con } t=1, 2, 3 \dots \quad (1.6.1)$$

che descrive le traiettorie di una *semimartingala* in tempo continuo nell'intervallo $[0, T]$, con $0 \leq s \leq t \leq T$. Dove $dp(t)$ è il differenziale del logaritmo del prezzo, μ_t è la componente di drift, σ_t è la volatilità istantanea del processo o deviazione standard, strettamente positiva e con il proprio quadrato integrabile (per cui $E(\int_0^t \sigma_s^2 ds) < \infty$), e W_t è un moto Browniano standard.

Se si definisce il rendimento composto continuo tra il tempo $t-h$ e il tempo t del suddetto asset, con $0 < h \leq t$, come

$$r_t = p(t) - p(t-h) = \int_{t-h}^t \mu_s ds + \int_{t-h}^t \sigma_s dW_s. \quad (1.6.2)$$

La variazione quadratica del rendimento, misura naturale della variabilità del sentiero di diffusione delle martingala secondo la teoria della integrazione stocastica, è definita come

$$[p, p]_t = QV_t = \int_{t-h}^t \sigma^2(s) ds. \quad (1.6.3)$$

L'equazione (1.6.3) mostra che le innovazioni del drift $\mu(t)$ non influenzano la variazione del sentiero di diffusione dei rendimenti. Quindi, in questo contesto, poiché la *variazione*

quadratica è totalmente indotta dalle innovazioni della martingala locale, essa coincide con la varianza integrata⁶, indice della variabilità cumulata *ex post* dal processo del rendimento e data da

$$IV_t = \int_{t-h}^t \sigma^2(s) ds = QV_t. \quad (1.6.4)$$

Dalla proprietà della variazione quadratica (definita dall'equazione (1.2.3) nella sezione 1.2) deriva che, in assenza di errori di microstruttura⁷ e di errori di misura (Andersen *et al.* (2000) e Barndorff-Nielsen e Shephard (2002)), la variazione quadratica dei rendimenti può essere approssimata come

$$[p, p]_t = \text{plim}_{n \rightarrow \infty} \sum_{j=1}^n \left[p(s_j) - p(s_{j-1}) \right]^2, \quad (1.6.5)$$

per ogni sequenza di partizioni $0 = s_0 < s_1 < \dots < s_n = t$ con $|s_j - s_{j-1}| \rightarrow 0$, quando il numero di partizioni n è tale per cui $n \rightarrow \infty$.

Poiché il rendimento intragiornaliero è definito come

$$r_{t,i} = p_{t,i} - p_{t,i-1} \quad \forall i = 1, \dots, n$$

e quello giornaliero come

$$r_t = \sum_{i=1}^n r_{t,i}.$$

La teoria della semimartingala assicura che la varianza realizzata, definita come la somma dei quadrati dei rendimenti intragiornalieri, converga in probabilità alla variazione quadratica, quindi alla varianza integrata del giorno t , per $n \rightarrow \infty$. Sia

$$RV_t = \sum_{i=1}^n r_{t,i}^2. \quad (1.6.6)$$

Si ha che

$$RV_t \xrightarrow{p} [r, r]_t - [r, r]_{t-h} \equiv QV_t \quad (1.6.7)$$

$$RV_t \xrightarrow{p} IV_t. \quad (1.6.8)$$

⁶La variazione quadratica e la varianza integrata non coincidono in un processo per i rendimenti più generale (e.g. modello di diffusione con salti).

⁷L'errore di microstruttura emerge per la presenza di scambi non sincronizzati, per l'assenza di scambi, per le proprietà del meccanismo di trading (Black (1976) e Amihud e Mendelson (1987)) e per la presenza di prezzi discreti (Harris (1990) e Harris (1991)).

La volatilità realizzata fornisce dunque uno stimatore consistente della volatilità nozionale, per cui

$$RV_t \xrightarrow{p} v^2(t, h). \quad (1.6.9)$$

Questo implica che la volatilità realizzata attesa sia uno stimatore consistente della volatilità nozionale attesa

$$E[RV_t | I_{t-h}] \xrightarrow{p} E[v^2(t, h) | I_{t-h}]. \quad (1.6.10)$$

Infine, dall'equazione (1.2.12) segue che, se il processo dei rendimenti è integrabile al quadrato e se $\mu(t) \equiv 0$, allora la volatilità realizzata è uno stimatore corretto della varianza condizionale dei rendimenti:

$$E[RV_t | I_{t-h}] = E[QV_t | I_{t-h}] = Var[r(t, h) | I_{t-h}]. \quad (1.6.11)$$

Questa uguaglianza lega il concetto di volatilità realizzata a quello della varianza condizionale dei modelli ARCH. In particolare, è possibile costruire un modello di serie storiche per la varianza realizzata mediante il quale si può ottenere una buona approssimazione della varianza condizionale dei rendimenti. Se $\mu(t) \neq 0$, la convergenza di RV a QV non implica automaticamente la convergenza in media degli stessi oggetti.

Primi esempi di varianza realizzata si trovano in Merton (1980), Poterba e Summers (1986), Schwert (1989), Richardson e Stock (1989), Schwert (1990), Taylor e Xu (1997) e Christensen e Prabhala (1998). Tuttavia, l'uso della misura realizzata si è diffuso solo in seguito alla formalizzazione della stessa da parte di Andersen e Bollerslev (1998) i quali, insieme al lavoro di Andersen *et al.* (2003), hanno dimostrato che la varianza realizzata è uno stimatore consistente della volatilità giornaliera solo se la distanza tra le osservazioni intragiornaliere tende a zero o se la frequenza di campionamento tende all'infinito.

1.6.2 Distribuzione della varianza realizzata

La distribuzione asintotica della varianza realizzata è stata analizzata in due diversi lavori, Jacod e Protter (1998) e Barndorff-Nielsen e Shephard (2002).

Dai loro risultati emerge che la varianza realizzata si distribuisce come

$$n^{1/2} \cdot \frac{1}{\sqrt{2IQ_t}} (RV_t - IV_t) \xrightarrow{d} N(0, 1) \quad (1.6.12)$$

dove n è la frequenza di campionamento intragiornaliero e IQ_t è definita come la quarticità integrata (*integrated quarticity*)

$$IQ_t = \int_{t-h}^t \sigma^4(s) ds.$$

Per poter fare inferenza è necessario uno stimatore consistente di IQ_t . Barndorff-Nielsen e Shephard (2002) hanno mostrato che uno stimatore consistente di IQ_t è dato dalla *Realized Quarticity*, RQ , definita come

$$RQ_t = \frac{1}{3} \sum_{i=1}^n r_i^4, \quad (1.6.13)$$

per cui

$$\frac{\delta^{-1/2}(RV_t - IV_t)}{\sqrt{\frac{2}{3}RQ_t}} \xrightarrow{d} N(0, 1). \quad (1.6.14)$$

Barndorff-Nielsen e Shephard (2002) dimostrano che un'approssimazione migliore si ha mediante una log-linearizzazione che ha come risultato

$$\frac{\delta^{-1/2}(\log RV_t - \log IV_t)}{\sqrt{\frac{2}{3} \frac{RQ_t}{RV_t^2}}} \xrightarrow{d} N(0, 1). \quad (1.6.15)$$

1.6.3 Problemi con la varianza realizzata

I risultati precedenti mostrano come sia preferibile scegliere la frequenza di campionamento più elevata possibile per la costruzione della varianza realizzata, tuttavia i diversi metodi di campionamento e la presenza dell'errore di microstruttura fanno sorgere dei problemi rilevanti.

Il logaritmo dei prezzi nella pratica non si diffonde in tempo continuo, come suggerito dalla teoria finanziaria, ma viene osservato ad intervalli discreti e non regolari nel tempo. I diversi metodi di campionamento influenzano la discretizzazione del prezzo, tra i vari metodi si ha:

- *Calendar time sampling*, si considerano intervalli equidistanti nel tempo, ad esempio ogni 5 minuti o ogni ora; in un lavoro di De Pooter *et al.* (2008) si cerca di trovare l'ottima frequenza di campionamento che massimizza l'utilità di un investitore avverso al rischio. Secondo il loro lavoro, la frequenza ottimale di campionamento per la stima della varianza realizzata giornaliera è compresa tra 30 e 65 minuti.
- *Transaction time sampling*, il campionamento avviene ogni n transazioni.
- *Business time sampling*, il campionamento è tale per cui $IV_{t,i} = \frac{IV_t}{n}$
- *Tick time sampling*, si campiona per ogni cambiamento di prezzo.

Diversi lavori si sono occupati di scegliere il miglior metodo di campionamento da utilizzare, in particolare Oomen (2005a, 2005b) conclude che il metodo *transaction time sampling* è superiore agli altri, in particolar modo nei periodi di maggiore volatilità; tuttavia, il metodo di campionamento più diffuso rimane il *calendar time sampling*.

Oltre alla scelta della metodologia di campionamento, la procedura per la costruzione della varianza realizzata deve far fronte alla presenza dell'errore di microstruttura. Esiste, infatti, un *trade-off* tra errore di microstruttura e accuratezza della stima che può portare la varianza realizzata ad essere uno stimatore non robusto della varianza integrata giornaliera.

L'effetto dell'errore di microstruttura si accentua quando viene usata la metodologia di campionamento *calendar time sampling*, quindi con serie misurate ad intervalli temporali equidistanti. Andersen *et al.* (2000, 2001, 2003) suggeriscono una soluzione al problema della presenza dell'errore di microstruttura che passa mediante il sottocampionamento. Una soluzione pratica è quella di non campionare troppo frequentemente, per esempio ogni 5 o 30 minuti. Questa procedura è chiamata *sparse sampling*.

Partendo da questa procedura Bandi e Russell (2005), Bandi e Russell (2007) e Zhang *et al.* (2005) hanno proposto un metodo per determinare l'ottima frequenza di sottocampionamento passando per la minimizzazione dell'MSE (errore quadratico medio):

$$MSE(n_t^{(sparse)}) = 2n_t^{(sparse)} E(\varepsilon_{t,i}^2) + 4n_t^{(sparse)} E(\varepsilon_{t,i}^4) + \left[8RV_t^{(sparse)} E(\varepsilon_{t,i}^2) - 2V(\varepsilon_{t,i}^2) \right] + \frac{2}{n_t^{(sparse)}} IQ_t^{(sparse)}$$

dove n_t è il numero di sottocampioni, RV_t è la varianza realizzata, IQ_t è la quarticità integrata, ε_t è l'errore di microstruttura con media zero. Da ciò segue che la frequenza di campionamento ottimale viene approssimata come

$$n_t^* = \left\{ \frac{IQ_t}{4[E(\varepsilon_{t,i}^2)]^2} \right\}^{1/3}.$$

Con $E(\varepsilon_{t,i}^2)$ che può essere stimato in maniera consistente da $\frac{1}{2n_t} RV_t$ e con la quarticità integrata che viene stimata mediante la quarticità realizzata.

Tuttavia, Zhang (2006) mostra che l'uso del sottocampionamento da un lato riduce la grandezza dell'errore di microstruttura, nello specifico di $2nE(\varepsilon_{t,i}^2)$, dall'altro lato aumenta la varianza dovuta alla maggiore discretizzazione per via dell'intervallo di campionamento maggiore, creando un *trade-off* tra distorsione e varianza.

Sono stati proposti, nel tempo, ulteriori stimatori alternativi per la varianza integrata che

riescano ad ovviare a questo problema e a tener conto dell'errore di microstruttura. Uno stimatore, proposto da Zhang *et al.* (2005) e definito *two time scales Estimator (TTSE)*, è dato da

$$RV_t^{(TTSE)} = \frac{1}{K} \sum_{k=1}^K RV_t^{(k)} - \frac{\bar{n}_t}{n_t} RV_t.$$

Dove l'insieme delle osservazioni giornaliere viene diviso in K sottoinsiemi non sovrapposti, con $k=1, \dots, K$, n_t è il numero delle osservazioni dell'intera griglia, $RV_t^{(k)}$ è la varianza realizzata del sottoinsieme k , RV_t è la varianza realizzata del giorno t e \bar{n}_t è definita come

$$\bar{n}_t = \frac{1}{K} \sum_{k=1}^K n_t^{(k)} = \frac{n_t - K - 1}{K}.$$

Una sua rivisitazione è stata proposta da Aït-Sahalia *et al.* (2006), con lo stimatore che diventa

$$RV_t^{(TTSE,adj)} = \left(1 - \frac{\bar{n}_t}{n_t}\right)^{-1} RV_t^{(TTSE)}.$$

Lo stimatore più diffuso sotto la presenza di errore di microstruttura è lo stimatore *kernel*. Zhou (1996) ha proposto per primo tale stimatore per i dati ad alta frequenza, dimostrando anche che l'errore di microstruttura è dipendente dal tempo e ha proprietà che variano nel tempo. Dallo stimatore proposto da Zhou (1996), Hansen e Lunde (2004) e Hansen e Lunde (2006) hanno elaborato il seguente stimatore

$$RV_t = RV_t + 2 \sum_{h=1}^H \frac{n_t}{n_t - h} \hat{\gamma}_h, \quad (1.6.16)$$

con

$$\hat{\gamma}_h = \frac{n_t}{n_t - h} \sum_{j=1}^{n_t-h} r_{t,j} r_{t,j+h}.$$

Tuttavia, gli stimatori proposti da Zhou (1996), con $H = 1$ per la (1.6.16), e da Hansen e Lunde (2004) risultano non consistenti.

Barndorff-Nielsen *et al.* (2008) cercano di ovviare a questo problema con lo stimatore *flat-top kernel-based*

$$RV_t^{(BHLS)} = RV_t + \sum_{h=1}^H k\left(\frac{h-1}{H}\right) (\hat{\gamma}_h + \hat{\gamma}_{-h}) \quad (1.6.17)$$

dove $k(x)$ per $x \in [0, 1]$ è una funzione dei pesi non stocastica tale per cui $k(0) = 1$ e $k(1) = 0$.

Alternativa agli stimatori proposti nel corso di questo paragrafo, in presenza di errori di

microstruttura, è il pre-filtraggio dei rendimenti intragiornalieri. Bollen e Inder (2002) ricorrono ad un filtro autoregressivo (AR), Ebens (1999) uno a media mobile (MA), così come Maheu e McCurdy (2002) e Andersen *et al.* (2001a); si rimanda ai loro articoli per un'analisi approfondita.

1.6.4 Covarianza realizzata

Come si è visto per i modelli ad eteroschedasticità condizionale e di volatilità stocastica, il passaggio dall'univariato al multivariato è una condizione necessaria per l'evoluzione della letteratura specializzata. Considerando n attività finanziarie, r_t diventa un vettore di dimensione n , le variabili p , μ e W diventano tutti vettori di processi n -dimensionali e il processo di diffusione del logaritmo dei prezzi è tale per cui:

$$dp(t) = \mu(t) + \Omega(t)dW(t) \quad t = 1, 2, \dots$$

dove $\Omega(t)$ è una matrice $n \times n$ definita come covolatilità istantanea tale per cui, supponendo che $\Omega(t)$ sia ortogonale a $W(t)$, $\Sigma_t = \Omega(t)\Omega(t)'$ è la matrice delle covarianze istantanee.

La variazione quadratica di $p(t)$ è data da

$$[p, p]_t = \int_{t-h}^t \Sigma_s ds = \int_{t-h}^t \Omega(s)\Omega(s)' ds.$$

Poiché in questo contesto la variazione quadratica coincide con la varianza integrata, la matrice di varianze integrate equivale a:

$$ICov_t = \int_{t-h}^t \Sigma(s) ds. \quad (1.6.18)$$

Andersen *et al.* (2003) hanno dimostrato che la covarianza realizzata

$$RCov_t = \sum_{i=1}^n r_{i,t} r'_{i,t} \quad (1.6.19)$$

risulta uno stimatore consistente per la covarianza integrata, dove n è il numero di partizioni intragiornaliere.

Nell'ambito multivariato emerge il problema della sincronicità degli scambi, poichè le contrattazioni avvengono in tempi differenti tra i vari asset, quindi, può sorgere il cosiddetto "effetto Epps" (Epps, 1979) per cui la correlazione tra gli asset tende ad essere sottostimata per effetto della non sincronicità. Anche in questo caso sono stati condotti alcuni studi per attenuare tale effetto, si veda Bandi e Russell (2005).

1.7 Modelli di varianza realizzata e previsioni

Il vantaggio di stime non parametriche per la volatilità, aggiunto al fatto che il logaritmo delle volatilità realizzate sia approssimativamente gaussiano, hanno favorito l'utilizzo di questo approccio nelle applicazioni empiriche. La letteratura si è divisa in diversi filoni, sia nell'univariato che nel multivariato, i prossimi paragrafi si occupano di analizzarli nel dettaglio.

1.7.1 Modelli di serie storiche

Per tenere conto della persistenza del logaritmo della varianza, il primo modello di serie storiche utilizzato è stato il fractionally integrated ARMA, Andersen *et al.* (2003), Oomen (2001). Il modello originale di Andersen *et al.* (2003) era:

$$\Phi(L)(1-L)^d(y_t - \mu) = \Theta(L)\varepsilon_t \quad (1.7.1)$$

dove y_t è il logaritmo della volatilità realizzata, L è l'operatore ritardo, $\Phi(L) = 1 - \Phi_1 L - \dots - \Phi_p L^p$, $\Theta(L) = 1 + \Theta_1 L + \dots + \Theta_q L^q$ e $(1-L)^d$ è l'operatore di differenza frazionale definito come

$$(1-L)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k-d)L^k}{\Gamma(-d)\Gamma(d+1)} \quad (1.7.2)$$

dove $\Gamma(\cdot)$ è la funzione gamma. Il parametro d può assumere qualsiasi valore tra 0 e 1 compresi, con $d = 0$ si ritorna al modello $I(0)$, con $d = 1$ si ha un modello integrato di ordine 1. La versione multivariata di questo modello è stata proposta da Halbleib-Chiriac e Voev (2011) e sarà analizzata nel corso del paragrafo.

In alternativa ai modelli ARFIMA, la memoria di lungo termine può essere catturata anche mediante il modello Heterogeneous Autoregressive (HAR) di Corsi (2009), con una struttura gerarchica delle volatilità realizzate su diversi orizzonti temporali. La volatilità realizzata giornaliera diventa funzione di quella giornaliera, settimanale e mensile ritardate

$$RV_{t+1g}^{(g)} = c^{(g)} + \beta^{(g)} RV_t^{(g)} + \beta^{(w)} RV_t^{(w)} + \beta^{(m)} RV_t^{(m)} + \varepsilon_{t+1,g} \quad (1.7.3)$$

dove g sta per frequenza giornaliera, w sta per quella settimanale (5 giorni) e m per quella mensile (20 giorni). I regressori $RV_t^{(w)}$, $RV_t^{(m)}$ sono la media dei valori passati di RV_t scalati per l'ampiezza della frequenza, ad esempio $RV_t^{(m)} = \frac{1}{20} \sum_{i=0}^{19} RV_{t-i}$. La struttura relativamente semplice e la possibilità di stimare il modello mediante il metodo dei minimi quadrati ha reso questa formulazione piuttosto popolare. Corsi (2009) ha mostrato come

questo modello generi previsioni out-of-sample migliori rispetto ai modelli con memoria breve e sia in linea con i risultati ottenuti dai modelli ARFIMA.

Il passaggio dai modelli univariati a quelli multivariati è reso relativamente semplice dalla possibilità di usare l'analisi delle serie storiche. Così, in Halbleib-Chiriac e Voev (2011) si propone un modello VARFIMA(p, d, q) sui fattori di Cholesky della matrice delle covarianze realizzate, in modo tale da cogliere la persistenza della volatilità e garantire che la matrice varianze e covarianze prevista sia almeno semi-definita positiva.

Se Y_t è la matrice delle covarianze realizzate al tempo t di dimensione $n \times n$, dove n è il numero degli asset, la scomposizione di Cholesky della matrice Y_t è data dalla matrice triangolare bassa P_t per cui $P_t P_t' = Y_t$. Se $X_t = vech(P_t)$ è il vettore di dimensione $n(n+1)/2$, ottenuto dalla trasformazione delle componenti della matrice P_t in un vettore colonna, il VARFIMA proposto diventa:

$$\Phi(L)D(L)[X_t - BZ_t] = \Theta(L)\varepsilon_t \quad \varepsilon_t \sim N(0, \Sigma_t) \quad (1.7.4)$$

dove Z_t è un vettore di variabili esogene di dimensione $k \times 1$, B è la matrice dei coefficienti di dimensione $n \times k$, $\Phi(L) = I_n - \Phi_1 L - \Phi_2 L^2 - \dots - \Phi_p L^p$ e $\Theta(L) = I_n - \Theta_1 L - \Theta_2 L^2 - \dots - \Theta_q L^q$ sono le matrici dei ritardi dove Φ_i , per $i = 1, \dots, p$ e Θ_j , per $j = 1, \dots, q$, sono i polinomi dei coefficienti AR- e MA- e $D(L) = diag\{(1-L)^{d_1}, \dots, (1-L)^{d_n}\}$, dove d_1, \dots, d_n sono i gradi di integrazione frazionale di ogni elemento del vettore X_t , Σ_t è la matrice delle covarianze di ε_t . Si assume che le radici di $\Phi(L)$ e $\Theta(L)$ cadano fuori dal cerchio unitario e che X_t sia stazionario se $d_i < 0,5$, come mostrato da Sowell (1992).

La stima viene condotta nella forma dell'equazione finale⁸ in modo tale da ridurre il numero di parametri da stimare e garantire l'unicità della rappresentazione. Il modello finale stimato mediante quasi-massima verosimiglianza è dunque un VARFIMA(1,d,1)

$$(1 - \Phi L)D(L)[X_t - c] = (1 - \Theta L)\varepsilon_t \quad \varepsilon_t \sim N(0, \Sigma) \quad (1.7.5)$$

dove c è un vettore di costanti $n \times 1$. Il numero di parametri da stimare rispetto ad un generico VARFIMA(p, d, q) passa da $qn^2 + (k+1)n + p$ a $2n + 2$, se $D(L) = diag\{(1-L)^{d_1}, \dots, (1-L)^{d_n}\}$, e a $n + 3$, se $D(L) = (1-L)^d I_n$.

Uno studio recente, da parte di Baruník e Čech (2016), prevede di modellare i fattori di Cholesky secondo un modello HAR generalizzato (GHAR) e di confrontare le performance

⁸Vedi Theil e Boot (1962). In questo caso si dice che l'equazione è nella forma finale la rappresentazione del VARMA(p, q) per cui $\Phi(L)Y_t = \Theta(L)\varepsilon_t$ se $\Theta_0 = I_n$ e $\Phi(L) = 1 - \Phi_1 L - \dots - \Phi_p L^p$ è un operatore scalare con $\Phi_p \neq 0$.

di un portafoglio ottenuto mediante tale modello con quelle di modelli esistenti in letteratura. Il modello proposto da Baruník e Čech (2016) è un'estensione del modello HAR nel multivariato, presente nel lavoro di Halbleib-Chiriac e Voev (2011). Gli autori propongono un sistema di regressioni HAR non correlate per tutti gli elementi del vettore di fattori di Cholesky, X_t , utilizzando il metodo di stima di regressioni non correlate introdotto da Zellner (1962). I loro risultati evidenziano l'utilità del loro modello in un'ottimizzazione di portafoglio se confrontato con un modello VARFIMA sui fattori di Cholesky, con il modello Riskmetrics e con un DCC-GARCH.

Un'alternativa alla scomposizione di Cholesky per garantire che la matrice delle covarianze realizzate sia almeno semi-definita positiva, introdotta da Bauer e Vorkink (2011), è la trasformazione logaritmica della stessa. La trasformazione produce le log-volatilità, definite come $a_t = \text{vech}(A_t)$, dove A_t è data da

$$A_t = B_t \log(G_t) B_t' \quad (1.7.6)$$

dove B_t e G_t sono delle matrici risultanti dalla scomposizione spettrale $Y_t = B_t G_t B_t'$ (si veda l'Appendice A).

Le dinamiche log-volatilità, a_t , prodotte dalla trasformazione logaritmica della matrice delle covarianze, vengono modellate mediante un VAR(1):

$$a_t = \gamma_0 + \gamma_1 a_{t-1} + \varepsilon_t \quad (1.7.7)$$

dove γ_0 è un vettore di intercette $n \times 1$, γ_1 è una matrice di coefficienti $n \times n$ e ε_t è un vettore di residui.

Viene, infine, ricostruita la matrice varianze e covarianze mediante i valori fittati $\hat{a}_t \equiv E_t(a_t) \equiv \hat{\gamma}_0 + \hat{\gamma}_1 a_{t-1}$; l'operatore inverso del vech permette la ricostruzione della matrice \hat{A}_t , la funzione esponenziale, poi, restituisce la stima della matrice delle covarianze \hat{V}_t :

$$\hat{V}_t = \exp(\hat{A}_t). \quad (1.7.8)$$

La matrice \hat{V}_t è definita positiva per costruzione.

L'ulteriore innovazione che accompagna il lavoro di Bauer e Vorkink è l'utilizzo di variabili esogene esplicative per la previsione della matrice delle covarianze. I due autori utilizzano come regressori, oltre ai valori ritardati delle variabili dipendenti, un set di variabili esogene che si sono mostrate utili nella previsione dei rendimenti di attività finanziarie nella letteratura di riferimento. Il modello diventa dunque:

$$a_t = \gamma_0 + \gamma_1 a_{t-1} + \gamma_2 a_{t-2} + \dots + \gamma_k a_{t-M} + \gamma_X X_{t-1} + \varepsilon_t. \quad (1.7.9)$$

Poiché il numero di parametri da stimare è relativamente alto, i due autori propongono tre metodologie alternative per ridurre il numero di parametri. Il primo approccio deriva dal modello Heterogeneous Autoregressive di Corsi, già descritto in precedenza. Considerando la trasformazione logaritmica della matrice *bi-power*⁹ delle covarianze e gli elementi del vettore ottenuto mediante il *vech*, $a^{BP}(d)_t$, si ha:

$$a_t = \gamma_0 + \gamma_1 a^{BP}(1)_{t-1} + \gamma_5 a^{BP}(5)_{t-1} + \gamma_{20} a^{BP}(20)_{t-1} + \gamma_X X_{t-1} + \varepsilon_t \quad (1.7.10)$$

dove $a^{BP}(1)_t$ è la log-covolatilità giornaliera, $a^{BP}(5)_t$ quella settimanale e $a^{BP}(20)_t$ quella mensile. In questo modo il numero di parametri da stimare è ridotto di $(M - 3)p^2$.

Un'ulteriore riduzione del numero di parametri si ha se si suppone che le serie $a_t^{BP}(d)_t$ siano guidate da un numero ridotto di fattori, in particolare si fa ricorso all'analisi delle componenti principali di $a^{BP}(d)_t$, dove $a^{BP}(d, i)$ è la i -esima componente principale per $i = 1, \dots, pc$, dove pc è il numero di componenti principali tale per cui $pc < p$.

Il terzo e ultimo approccio fa uso delle variabili sottostanti non osservabili (fattori latenti).

Se le variabili esplicative

$$Z_t = (a^{BP}(1, 1)_t, \dots, a^{BP}(5, 1)_t, \dots, a^{BP}(20, 1)_t, \dots, X_t) \quad (1.7.11)$$

hanno un collegamento con i fattori non noti della volatilità, allora il k -esimo fattore della volatilità, $v_{k,t}$, è dato da

$$v_{k,t} = \theta_k Z_{t-1}. \quad (1.7.12)$$

Il fattore è, dunque, una combinazione lineare delle N variabili Z_t , $\theta_k = \{\theta_{k,(1)}, \dots, \theta_{k,(N)}\}$ sono i coefficienti che combinano le variabili esplicative. Si ha così che ogni log-volatilità è funzione dei fattori:

$$a_t^i = \gamma_0^i + \beta^i \theta Z_{t-1} + \varepsilon_t^i, \quad (1.7.13)$$

con $i = 1, \dots, p$, dove γ_0^i è l' i -esimo elemento del vettore γ_0 , β^i è il vettore dei pesi delle log-volatilità e θ è una matrice $K \times N$ contenente i coefficienti sulle variabili Z_{t-1} per i k fattori. Aggregando per le p log-volatilità si ha

$$a_t = \gamma_0 + \beta \theta Z_{t-1} + \varepsilon_t. \quad (1.7.14)$$

⁹Introducendo una componente di salto nel processo di diffusione del logaritmo del prezzo, p_t , uno stimatore robusto della IV, introdotto da Barndorff-Nielsen e Shephard (2004), è dato da:

$$BV = \sum_{i=1}^{n-1} |r_i| \cdot |r_{(i+1)}|$$

dove n è il numero di partizioni intragiornaliere.

Questo approccio ha il vantaggio di ridurre sensibilmente il numero di parametri da stimare, in aggiunta permette di combinare le volatilità ritardate con variabili esogene esplicative.

Un nuovo modello per le dinamiche delle volatilità multivariate, chiamato Wishart Autoregressive (WAR), è stato proposto da *Gourieroux et al. (2009)*. Il modello è basato sulla distribuzione Wishart della matrice varianze e covarianze; se $X_{k,t}$ con $k = 1, \dots, K$ sono processi VAR(1) Gaussiani indipendenti di dimensione n si ha:

$$X_{k,t} = MX_{k,t-1} + \varepsilon_{k,t} \quad \varepsilon_{k,t} \stackrel{\text{i.i.d.}}{\sim} N(0, \Sigma) \quad (1.7.15)$$

Il processo, definito da

$$Y_t = \sum_{k=1}^K X_{k,t} X_{k,t}' \quad Y_t \sim W_n(K, M, \Sigma) \quad (1.7.16)$$

è un processo Wishart Autoregressivo di ordine 1 di dimensione n , dove K sono i gradi di libertà. Unendo le due espressioni, Y_t può essere scritta come

$$Y_t = MY_{t-1}M' + K\Sigma + \eta_t \quad (1.7.17)$$

dove η_t è il termine d'errore eteroschedastico con media condizionale pari a zero e M è una matrice $n \times n$. Affinché Y_t abbia una distribuzione Wishart e che Y_t sia definita positiva è necessario che K sia maggiore di n anche se, come mostrato da Halbleib-Chiriac (2007), non sempre accade empiricamente. Il valore K deve essere stimato e *Gourieroux et al. (2009)* ne presentano una stima mediante il metodo dei momenti:

$$\hat{k} = \frac{2 \left[\alpha' \hat{\Sigma}^*(\infty) \alpha \right]^2}{\hat{V} \left[\alpha' Y_t \alpha \right]}$$

dove α è un vettore $n \times 1$ definito come allocazione di portafoglio e $\hat{\Sigma}^*(\infty)$ è dato da $\hat{\Sigma}^*(\infty) = \hat{M} \hat{\Sigma}^*(\infty) \hat{M}' + \hat{\Sigma}^*$, \hat{M} e $\hat{\Sigma}^*$ sono le stime prodotte dal metodo dei momenti, $\hat{V}(\alpha' Y_t \alpha)$ è la varianza campionaria della volatilità di portafoglio.

Il modello presenta diverse estensioni, si veda *Bonato (2009)*, *Bonato et al. (2009)*, *Jin e Maheu (2012)* e *Halbleib-Chiriac e Voev (2011)*. Il problema di questo modello è che le assunzioni parametriche sono particolarmente restrittive e che l'unico metodo di stima possibile, una stima Bayesiana legata all'utilizzo di una simulazione di Monte Carlo basata su una Catena di Markov (MCMC), è dispendioso dal punto di vista computazionale.

Nell'ambito dell'uso dei modelli di serie storiche rientrano anche due modelli recenti che puntano ad introdurre l'uso di molteplici variabili esogene per spiegare le dinamiche

del logaritmo della volatilità e sono quelli di Paye (2012) e Christiansen *et al.* (2012). Un'analisi dettagliata dei due modelli si trova nel corso del secondo capitolo, il modello di base è:

$$LVOL_t = \alpha + \sum_{k=1}^K \rho_k LVOL_{t-k} + \beta' X_{t-1} + \varepsilon_t \quad (1.7.18)$$

dove X_{t-1} è il vettore delle variabili esogene al tempo $t - 1$, $LVOL_t$ è il logaritmo della volatilità realizzata e α è una costante. Interessanti risultano le motivazioni per l'uso delle variabili esplicative e le analisi *out-of-sample* dei due modelli.

1.7.2 Modelli MIDAS e GARCH-MIDAS

Accanto all'analisi dei modelli di serie storiche si è sviluppato un altro filone di ricerca che utilizza dati ad alta frequenza e misure realizzate, detto approccio Mixed Data Sampling (MIDAS). Il MIDAS, introdotto da Ghysels *et al.* (2004) e Ghysels *et al.* (2006), fa ricorso all'uso dei dati osservati a frequenze differenti in maniera efficiente per prevedere la volatilità, anche per più passi in avanti.

Se V_{t+1} è una misura di volatilità, come la volatilità realizzata, nel periodo successivo $t + h$ la regressione MIDAS può essere scritta come:

$$\hat{V}_{t+h} = \mu + \varphi \sum_{j=0}^{jmax} b(j, \theta) X_{t-j} + \varepsilon_t \quad (1.7.19)$$

dove $jmax$ è il numero massimo di ritardi considerati, X_{t-j} è un set di variabili esplicative e $b(j, \theta)$ è una funzione dei pesi dei ritardi dei regressori. I regressori vengono campionati a frequenze più alte rispetto alla variabile dipendente. La stima dei parametri del modello MIDAS viene condotta mediante massima verosimiglianza.

Diversi lavori hanno utilizzato l'approccio MIDAS per un'analisi empirica. In particolare, Becker *et al.* (2010) associano l'utilizzo del campionamento a frequenze miste alla scomposizione di Cholesky della matrice varianze e covarianze realizzate. I primi due step della procedura risultano identici a quelli posti in essere da Halbleib-Chiriac e Voev (2011), mentre la matrice varianze e covarianze realizzate giornaliera è costruita secondo l'approccio di Hansen e Lunde (2005). Tale approccio considera la differenza tra il periodo di chiusura e di apertura un rendimento separato rispetto a quello delle fasi di contrattazione, si ha così una matrice per l'intera giornata di 24 ore data dalla somma della componente chiusura-apertura, data da $r_{co,t} r'_{co,t}$, e della covarianza realizzata definita

come nella (1.6.19), per cui

$$V_t = r_{co,t}r'_{co,t} + \sum_{i=1}^n r_{i,t}r'_{i,t} \quad (1.7.20)$$

dove $r_{co,t}$ è il vettore dei rendimenti dalla chiusura al giorno $t - 1$ all'apertura del giorno t . Sommando le matrici le matrici varianze e covarianze giornaliere, V_t , si può ricavare la matrice per m giorni, $V_t^{(m)}$. La sua scomposizione di Cholesky è data da $V_t^{(m)} = C_t^{(m)}C_t^{(m)'$, mentre la trasformazione in vettore della triangolare alta C_t è data da $P_t^{(m)} = \text{vech}(C_t^{(m)})$.

Gli $n(n + 1)/2$ elementi di $P_t^{(m)}$ vengono poi modellati secondo l'approccio MIDAS, mediante una media ponderata dei loro valori passati:

$$P_{i,t+m}^{(m)} = \beta_{i0} + \beta_{i1} \sum_{k=1}^K B(k, 1, \theta_i) P_{i,t-k+1} + v_t \quad (1.7.21)$$

dove B è una funzione di pesi e in questo caso è una funzione beta:

$$B(k, i, \theta_i) = \frac{f(\frac{k}{K}, 1, \theta_i)}{\sum_{k=1}^K f(\frac{k}{K}, 1, \theta_i)} \quad (1.7.22)$$

$$f(z, a, b) = \frac{z^{a-1}(1-z)^{b-1}\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \quad (1.7.23)$$

dove K è il numero massimo di ritardi e β_{i0} , β_{i1} e θ_i sono parametri da stimare. Inserendo delle variabili esogene esplicative nel modello, si ha

$$P_{i,t+m}^{(m)} = \beta_{i0} + \beta_{i1} B(k, 1, \theta_i) P_{i,t-k+1} + \beta_{ix} B(k, 1, \theta_{ix}) X_{t-k+1} + v_t. \quad (1.7.24)$$

Becker *et al.* (2010) precisano che non è ancora chiara l'utilità delle variabili esogene nel prevedere la matrice delle varianze e covarianze realizzate; questa tesi, in particolare nel secondo capitolo, cerca di analizzare nel dettaglio l'argomento.

Per poter prevedere $P_{i,t+m}^{(m)}$ è necessario stimare i parametri del modello Cholesky-MIDAS. La stima viene condotta mediante il metodo dei minimi quadrati non lineari (Non-Linear Least Squares, NLS). Una volta ottenuta la stima dei parametri del modello, si producono le previsioni per m periodi in avanti di $P_{i,t+m}^{(m)}$ e si ricostituisce la matrice varianze e covarianze prevista.

Un'interessante analisi delle diverse specificazioni della funzione dei pesi, B , viene condotta da Ghysels *et al.* (2009). Nel loro lavoro confrontano le previsioni ottenute con un modello MIDAS con quelle di un GARCH e di un modello AR sulla varianza realizzata.

Le previsioni fuori campione derivano dalla regressione

$$\tilde{V}_{t+1}^k = \mu_k + \varphi_k \sum_{j=0}^{j^{max}} b_k(j, \theta) r_{t-j}^2 + \varepsilon_{k,t} \quad (1.7.25)$$

dove \tilde{V}_{t+1}^k è una qualche misura della volatilità come la volatilità realizzata, per cui $\tilde{V}_{t+1}^k = RV_{t+1}^k = \sum_{j=1}^k r_{t+j}^2$, $b_k(j, \theta)$ è una funzione di pesi. La costante μ_k , il coefficiente φ_k e il parametro θ della funzione di peso devono essere stimati mediante quasi-massima verosimiglianza. Essi considerano diverse specificazioni dei pesi e dimostrano come le previsioni ottenute mediante l'approccio MIDAS siano migliori, in termini di errore quadratico medio (MSE) e secondo i test di West (1996) e Giacomini e White (2006), rispetto a quelle ottenute con un GARCH o con un AR sulla varianza realizzata.

L'approccio MIDAS viene molto spesso abbinato all'approccio GARCH per una specificazione della volatilità a due componenti, come suggerito nel modello Spline-GARCH di Engle e Rangel (2008). La componente di breve periodo viene modellata secondo un processo GARCH che si muove intorno ad un trend di lungo periodo, la componente di lungo periodo viene modellata con una funzione *Spline*¹⁰. Engle *et al.* (2013) hanno introdotto il modello GARCH-MIDAS, modellando la componente di lungo periodo secondo l'approccio MIDAS. Il maggior vantaggio di questo modello è che permette di collegare le osservazioni giornaliere dei rendimenti azionari con variabili macroeconomiche osservate a frequenze più basse, in modo tale da esaminare l'impatto di queste variabili sulla volatilità dei rendimenti.

Il modello Spline-GARCH di Engle e Rangel (2008) prevede che i rendimenti seguano un processo del tipo:

$$r_{i,t} - E[r_{i,t} | I_{i-1,t}] = \sqrt{g_{i,t}\tau_t} Z_{i,t} \quad (1.7.26)$$

dove $r_{i,t}$ sono i logaritmi dei rendimenti giornalieri, $I_{i-1,t}$ è l'informazione disponibile al giorno i , $Z_{i,t} \stackrel{iid}{\sim} (0, 1)$ sono le innovazioni, $g_{i,t}$ è un processo GARCH e τ_t è una funzione spline esponenziale. La volatilità presenta, dunque, due componenti, quella di breve periodo per analizzare le fluttuazioni giornaliere, $g_{i,t}$, e quella di lungo periodo, τ_t . Nel modello GARCH-MIDAS la funzione spline viene sostituita con una specificazione MIDAS.

Engle *et al.* (2013), dunque, hanno introdotto il modello GARCH-MIDAS con l'uso di variabili macroeconomiche come variabili esogene, in particolare inflazione e tasso di

¹⁰Una funzione spline ha lo scopo di interpolare in un intervallo un insieme di punti, mediante un insieme di polinomi raccordati tra loro. Si veda Wold (1976) per ulteriori approfondimenti sulle funzioni spline.

crescita della produzione industriale. Partendo dalla (1.7.26), l'equazione dei rendimenti nel giorno i e nel mese¹¹ t diventa

$$r_{i,t} = \mu + \sqrt{\tau_t} g_{i,t} Z_{i,t} \quad \forall i = 1, \dots, N_t \quad (1.7.27)$$

dove N_t è il numero di giorni che compongono t e μ è la media non condizionale di $r_{i,t}$. $g_{i,t}$ segue un processo GARCH del tipo:

$$g_{i,t} = (1 - \alpha - \beta) + \alpha \frac{(r_{i-1,t} - \mu)^2}{\tau_t} + \beta g_{i-1,t}. \quad (1.7.28)$$

Prendendo spunto dalla letteratura sulla varianza realizzata e dal paper di Schwert (1989), modellano τ_t come funzione della varianza realizzata mensile, RV_t . Secondo lo schema MIDAS, la componente di lungo periodo diventa

$$\tau_t = m + \theta \sum_{k=1}^K \varphi_k(w_1, w_2) RV_{t-k} \quad (1.7.29)$$

dove m è una costante, θ misura l'impatto dei ritardi di r_t^2 e $RV_t = \sum_{i=1}^{N_t} r_{i,t}^2$. La funzione dei pesi, $\varphi_k(w_1, w_2)$, viene specificata in maniera duplice:

$$\left\{ \begin{array}{ll} \frac{(k/K)^{w_1-1} (1-k/K)^{w_2-1}}{\sum_{j=1}^K (j/K)^{w_1-1} (1-j/K)^{w_2-1}} & \text{Funzione Beta} \\ w^k / (\sum_{j=1}^K w^j) & \text{Pesi Esponenziali} \end{array} \right. \quad (1.7.30)$$

I pesi sommano a uno. L'indice t può rappresentare diverse basse frequenze, ad esempio mensile, trimestrale o semestrale, e il numero di ritardi della componente MIDAS può variare sensibilmente. La stima dei parametri del modello viene condotta mediante quasi-massima verosimiglianza.

Infine, questa specificazione permette di incorporare variabili macroeconomiche nel modello. Engle *et al.* (2013) presentano un modello ad un filtro e uno a due filtri. Il primo prevede l'utilizzo di variabili macroeconomiche ritardate come regressori nella componente di lungo periodo, per cui

$$\log \tau_t = m_l + \theta_l \sum_{k=1}^{K_l} \varphi_k(w_{1,l}, w_{2,l}) X_{l,t-k}^{mv} \quad (1.7.31)$$

¹¹La frequenza non deve necessariamente essere un mese, può essere anche più bassa come un trimestre o un semestre.

dove $X_{l,t-k}^{mv}$ rappresenta il livello di una variabile macroeconomica, nello specifico il tasso di inflazione o il tasso di crescita della produzione industriale, m_l è una costante e θ_l misura l'impatto dei ritardi della variabile esogena sul logaritmo della componente di lungo periodo.

Il modello definito "two-sided filter" utilizza osservazioni, passate e future, della variabile macroeconomica per modellare la componente di lungo periodo. La specificazione univariata è per cui

$$\log \tau_t = m_2 + \sum_{k=-K_l^{(l)}}^{K_f^{(l)}} \varphi_k(w_1, w_2) \theta_l^{(k)} X_{l,t+k}^{mv} \quad (1.7.32)$$

dove gli impatti delle realizzazioni delle variabili macro sono liberi di variare, infatti

$$\theta_l^{(k)} = \begin{cases} \theta_l^f & k \geq 0 \\ \theta_l^b & k < 0 \end{cases}$$

Nel loro lavoro il modello GARCH-MIDAS fornisce previsioni migliori di quelle dei modelli con esso confrontati, in particolare nei sotto-campioni come il periodo della Grande Depressione o nel periodo successivo alla Seconda Guerra Mondiale.

Conrad e Loch (2014) estendono il modello di Engle *et al.* (2013) includendo due variabili esogene contemporaneamente nella componente di lungo periodo, si ha così

$$\log(\tau_t) = m + \theta^X \sum_{k=1}^K \varphi_k(w_1^X, w_2^X) X_{t-k} + \theta^Y \sum_{k=1}^K \varphi_k(w_1^Y, w_2^Y) Y_{t-k} \quad (1.7.33)$$

dove Y_{t-k} rappresenta una seconda variabile macroeconomica esplicativa. Sia Y_{t-k} la varianza realizzata mensile, per costruire una componente di lungo periodo annuale, si ha

$$\log(\tau_t) = m + \theta^{RV} \sum_{k=1}^{K=12} \varphi_k(w_1^{RV}, w_2^{RV}) RV_{t-k} + \theta^{MV} \sum_{k=1}^{K=12} \varphi_k(w_1^{MV}, w_2^{MV}) X_{t-k}. \quad (1.7.34)$$

Con questo modello producono previsioni un passo avanti. All'inizio del periodo t la componente di lungo periodo τ_t è predeterminata rispetto al set informativo I_{t-1} , di conseguenza la previsione dell'intera volatilità per il giorno i nel periodo t , è data da

$$E \left[g_{i,t} \tau_t Z_{i,t}^2 \mid I_{t-1} \right] = \tau_t E \left[g_{i,t} \mid I_{t-1} \right].$$

Poiché $E[g_{i,t} \mid I_{t-1}] = 1 + (\alpha + \beta + \gamma/2)^{i-1} (g_{1,t} - 1)$ converge alla varianza non condizionale di $g_{i,t}$, le previsioni tendono alla componente di lungo periodo per i sufficientemente ampio.

La previsione per il periodo t è data da

$$E \left[\sum_{i=1}^{N_t} g_{i,t} \tau_t Z_{i,t}^2 \mid I_{t-1} \right] = \tau_t \left(N_t + (g_{1,t} - 1) \frac{1 - (\alpha + \beta + \gamma/2)^{N_t}}{1 - \alpha - \beta - \gamma/2} \right).$$

Se $g_{1,t}$ è uguale alla sua varianza non condizionale, si ha che la previsione per il periodo t è data da $\tau_t N_t$.

Essi utilizzano un ampio set di variabili macroeconomiche come variabili esogene, in particolare il tasso di crescita del PIL, della produzione industriale, del tasso di disoccupazione, il *term spread*¹², il deflatore del PIL, il CPI e altre. Anche Asgharian *et al.* (2013) analizzano l'importanza delle variabili macroeconomiche per prevedere la volatilità mediante l'approccio GARCH-MIDAS.

L'approccio GARCH-MIDAS viene esteso anche nel multivariato. Considerando un vettore di rendimenti di n asset, nella specificazione DCC di Engle (2002) esso segue un processo del tipo $r_t \sim N(\mu, H_t)$, dove la matrice delle covarianze condizionali, H_t , può essere scritta come specificato nella (1.4.8). Nel caso bivariato le volatilità condizionali per il titolo i e il titolo j , definite come $q_{i,t}$ e $q_{j,t}$, seguono dei modelli GARCH univariati che vengono stimati a parte nel primo stadio. La stima delle loro covarianze condizionali rappresenta il secondo passaggio. La covarianza condizionale è specificata come nella (1.4.11) e Q_t è data da

$$q_{ij,t} = \bar{\rho}_{ij,t}(1 - a - b) + a(u_{i,t-1}u_{j,t-1}) + b(q_{ij,t-1}) \quad (1.7.35)$$

dove $u_{i,t}$ e $u_{j,t}$ rappresentano i residui standardizzati dal modello univariato e la correlazione condizionale è

$$\rho_{ij,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t}q_{jj,t}}}. \quad (1.7.36)$$

$q_{ij,t}$ è la covarianza di breve periodo.

Il primo lavoro a proporre una combinazione del modello DCC con l'approccio MIDAS è quello di Colacito *et al.* (2011). Nel modello DCC-MIDAS la covarianza condizionale è definita come nella (1.7.35), la correlazione di lungo periodo viene specificata secondo l'approccio MIDAS

$$\bar{\rho}_{ij,t} = \sum_{k=1}^{K_c} \varphi_k (w_k^{ij}) C_{ij,t-k} \quad (1.7.37)$$

¹²Il *term spread* rappresenta la differenza tra il tasso di interesse di un'obbligazione a breve termine e una a lungo termine.

dove K_c è il numero di ritardi delle correlazioni storiche, $C_{ij,t-k}$, specificate come

$$C_{ij,t} = \frac{\sum_{k=t-N}^t u_{i,k} u_{j,k}}{\sqrt{\sum_{k=t-N}^t u_{i,k}^2} \sqrt{\sum_{k=t-N}^t u_{j,k}^2}}. \quad (1.7.38)$$

$\bar{\rho}_{ij,t}$ è la correlazione di lungo periodo che si muove lentamente e u_t è l'innovazione standardizzata.

La specificazione generale del modello in forma matriciale è data da

$$R_t = (Q_t^*)^{-1/2} Q_t (Q_t^*)^{-1/2} \quad (1.7.39)$$

$$Q_t^* = \text{diag}(Q_t) \quad (1.7.40)$$

$$Q_t = (1 - a - b) \bar{R}_t(\underline{w}_r) + a u_t u_t' + b Q_{t-1} \quad (1.7.41)$$

dove

$$\bar{R}_t(\underline{w}_r) = \sum_{k=1}^{K_c} \phi_k(\underline{w}_r) \odot C_{t-k} \quad (1.7.42)$$

$$C_t = \begin{pmatrix} v_{1,t} & 0 & 0 \\ \vdots & \ddots & 0 \\ 0 & \cdots & v_{n,t} \end{pmatrix}^{-\frac{1}{2}} \left(\sum_{k=t-N_c}^t u_k u_k' \right) \begin{pmatrix} v_{1,t} & 0 & 0 \\ \vdots & \ddots & 0 \\ 0 & \cdots & v_{n,t} \end{pmatrix}^{-\frac{1}{2}} \quad (1.7.43)$$

$$v_{i,t} = \sum_{k=t-N_c}^t u_{i,k}^2 \quad \forall i = 1, \dots, n \quad (1.7.44)$$

dove $\phi_k(\underline{w}_r) = \varphi_k(\underline{w}_r) u u'$. Con questa specificazione a e b sono parametri comuni a tutte le combinazioni di titoli. Senza questa imposizione, la componente di breve periodo diventa

$$Q_t = G \odot \bar{R}_t(\underline{w}_r) + A \odot u_{t-1} u_{t-1}' + B \odot Q_{t-1} \quad (1.7.45)$$

dove G , A e B sono matrici $n \times n$ di parametri.

Quando c'è un unico parametro w_r , la matrice delle covarianze è definita positiva sotto alcune assunzioni non stringenti. Infatti, si può notare che la matrice Q_t sia una media ponderata di tre matrici. Dove \bar{R}_t , essendo una media ponderata delle matrici delle correlazioni, è semi-definita positiva e $u_{t-1} u_{t-1}'$ è semi-definita positiva per costruzione. Con una matrice iniziale Q_0 semi-definita positiva, Q_t è dunque semi-definita positiva in ogni punto.

Quando sono presenti due o più funzioni di peso, la matrice \bar{R}_t non è semi-definita positiva per ogni specificazione MIDAS. Sono necessarie delle restrizioni affinché la sequenza

di matrici $\left\{ \phi_k \right\}_{k=1}^K$ sia semi-definita positiva. Un'analisi delle restrizioni necessarie per assicurare una matrice semi-definita positiva viene condotta nel lavoro di Colacito *et al.* (2011). La stima dei parametri del modello utilizza una quasi-massima verosimiglianza a due stadi, come in Engle (2002).

Un'estensione del lavoro di Colacito *et al.* (2011) viene data da Asgharian *et al.* (2014) con l'inclusione di variabili macroeconomiche e di correlazioni realizzate ritardate nel modello di lungo periodo. Il modello DCC-MIDAS-XC diventa

$$q_{ij,t} = \bar{\rho}_{ij,t}(1 - a - b) + au_{i,t-1}u_{j,t-1} + bq_{ij,t-1} \quad (1.7.46)$$

$$\bar{\rho}_{ij,t} = \frac{\exp(2\bar{z}_{ij,\tau}) - 1}{\exp(2\bar{z}_{ij,\tau}) + 1} \quad (1.7.47)$$

$$\bar{z}_{ij,\tau} = m_{ij} + \theta_{RC} \sum_{k=1}^K \varphi_k(w_1, w_2) RC_{ij,t-k} + \theta_X \sum_{k=1}^K \varphi_k(w_1, w_2) X_{t-k}^Q \quad (1.7.48)$$

$$RC_{ij,t} = \frac{\sum_{k=1}^{N_t} u_{i,k} u_{j,k}}{\sqrt{\sum_{k=1}^{N_t} u_{i,k}^2} \sqrt{\sum_{k=1}^{N_t} u_{j,k}^2}} \quad (1.7.49)$$

dove $RC_{ij,t}$ è la correlazione realizzata, misurata a frequenze trimestrali, X_t^Q è una variabile macroeconomica misurata alla stessa frequenza e K è il numero della componente MIDAS.

Come nel modello di Engle *et al.* (2009), gli autori utilizzano anche osservazioni future delle variabili macroeconomiche, trasformando il modello in un DCC-MIDAS-XCF. Se θ_{RC} è imposto uguale a zero, si ha:

$$\bar{z}_{ij,t} = m + \theta_X \sum_{k=1}^{K_{lag}} \varphi_k(w_1, w_2) X_{t-k}^Q + \theta_X \sum_{k=-K_{lead}}^0 \varphi_k(w_1, w_2) X_{t-k|t}^{SPF}. \quad (1.7.50)$$

Le osservazioni future, X^{SPF} , vengono sostituite dalle previsioni delle variabili macroeconomiche fornite dalla *Survey of Professional Forecasters*. Il problema maggiore è combinare i dati storici con le previsioni, così una soluzione suggerita è quella di considerare le previsioni come una variabile individuale con un proprio parametro, θ_{FX} , per cui

$$\bar{z}_{ij,t} = m + \theta_X \sum_{k=1}^{K_{lag}} \varphi_k(w_1, w_2) X_{t-k}^Q + \theta_{FX} \sum_{k=-K_{lead}}^0 \varphi_k(w_1, w_2) X_{t-k|t}^{SPF}. \quad (1.7.51)$$

Per la stima dei parametri del modello si usa la quasi-massima verosimiglianza a due stadi. La funzione di quasi-massima verosimiglianza da massimizzare è data da

$$L = - \sum_{t=1}^T (T \log(2\phi) + 2 \log |D_t| + u_t' D_t^{-2} u_t) - \sum_{t=1}^T (\log |R_t| + u_t' R_t^{-1} u_t - u_t' u_t)$$

dove D_t è una matrice diagonale con le deviazioni standard dei rendimenti sulla diagonale e R_t è la matrice delle correlazioni condizionali dei residui standardizzati.

Le variabili di loro interesse sono i rendimenti azionari e obbligazionari, in particolare dei future sullo S&P 500 e sul Treasury note a 10 anni, e usano un largo insieme di variabili macroeconomiche. I loro risultati empirici evidenziano l'utilità del modello DCC-MIDAS-XC.

1.7.3 Realized GARCH e modello HEAVY

Un recente metodo di stima della volatilità prevede l'utilizzo di un GARCH che include una misura realizzata (come la varianza realizzata, la variazione bi-quadratica, il kernel realizzato) nell'equazione GARCH. Alcune specificazioni di questa metodologia comportano l'uso di molteplici processi latenti per la volatilità, in particolare il Multiplicative Error Model (MEM) di Engle e Gallo (2006) e il modello HEAVY (High-frEQUENCY-bAsed VolatilitY) di Shephard e Sheppard (2010). Un modello che presenta un unico processo latente per la volatilità è il Realized GARCH di Hansen *et al.* (2011a).

La variabile di interesse è la varianza condizionale, $h_t = \text{Var}(r_{t-1} | I_{t-1})$, dove r_t è la serie di rendimenti di un qualche asset finanziario. Nella classica specificazione GARCH h_t è funzione solo di h_{t-1} e r_{t-1}^2 , nel modello di Hansen *et al.* (2011a) h_t è funzione anche di una qualche misura realizzata di volatilità, x_{t-1} . Un'equazione, definita *measurement equation*, completa il modello specificando le determinanti della misura realizzata. Il modello si presenta perciò come segue:

$$r_t = \sqrt{h_t} z_t \quad (1.7.52)$$

$$h_t = w + \beta h_{t-1} + \gamma x_{t-1} \quad (1.7.53)$$

$$x_t = \xi + \varphi h_t + \tau(z_t) + u_t \quad (1.7.54)$$

dove $z_t \sim i.i.d.(0, 1)$, $u_t \sim i.i.d.(0, \sigma_u^2)$. In questo modo h_t diventa un processo autoregressivo di ordine 1 del tipo $h_t = \mu + \phi h_{t-1} + w_{t-1}$, dove $\mu = w + \gamma \xi$, $\phi = \beta + \varphi \gamma$ e $w_t = \gamma \tau(z_t) + \gamma u_t$. Gli autori, per tenere conto di eventuali risposte asimmetriche a shock di volatilità, suggeriscono di specificare $\tau(z)$ come

$$\tau(z) = \tau_1 z + \tau_2 (z^2 - 1).$$

Il loro modello è relativamente semplice da stimare mediante quasi-massima verosimiglianza ed estendibile nel contesto multivariato, così come proposto da Hansen *et al.* (2014).

Il modello high-frequency-based volatility (HEAVY) di Shephard e Sheppard (2010) fa, invece, ricorso a due o più processi latenti. La specificazione di base di questo modello è data da

$$\text{var}(r_t | I_{t-1}^{HF}) = h_t = w + \alpha RM_{t-1} + \beta h_{t-1}, \quad w, \alpha \geq 0, \beta \in [0, 1) \quad (1.7.55)$$

$$E(RM_t | I_{t-1}^{HF}) = \mu_t = w_R + \alpha_R RM_{t-1} + \beta_R \mu_{t-1}, \quad w_R, \alpha_R, \beta_R \geq 0, \alpha_R + \beta_R \in [0, 1) \quad (1.7.56)$$

dove I_{t-1}^{HF} è l'informazione dei dati ad alta frequenza, RM_{t-1} è la misura realizzata, i pesi w e w_R sono costanti e μ_t è la variabile latente per la media condizionale della misura realizzata. Questo modello semi-parametrico può essere anche esteso a dinamiche più complesse nell'equazione di μ_t e all'uso di r_{t-1}^2 , come nel tradizionale modello GARCH, in h_{t-1} . La stima del modello viene condotta mediante quasi-massima verosimiglianza.

Il modello è esteso nel multivariato da Noureldin *et al.* (2011). Sia $r_{i,t}$ il vettore dei rendimenti intragiornalieri di dimensione $n \times 1$, la misura realizzata, V_t , sarà una matrice $n \times n$, un esempio è dato dalla covarianza realizzata RC_t :

$$RC_t = \sum_{i=1}^n r_{i,t} r'_{i,t}$$

dove n sono le partizioni intragiornaliere per una covarianza realizzata giornaliera. Il prodotto incrociato dei rendimenti giornalieri è dato da $P_t = r_t r'_t$, dunque il modello HEAVY diventa

$$\begin{aligned} E \left[P_t | I_{t-1}^{HF} \right] &= E \left[r_t r'_t | I_{t-1}^{HF} \right] = H_t \\ E \left[V_t | I_{t-1}^{HF} \right] &= M_t, \end{aligned}$$

con $E[r_t | I_{t-1}^{HF}] = 0$, in modo tale che H_t sia la matrice delle covarianze condizionali. Si possono scegliere diverse specificazioni per garantire che H_t e M_t siano semi-definite positive, gli autori scelgono una parametrizzazione di tipo BEKK (Engle e Kroner (1995)):

$$\begin{aligned} H_t &= \bar{C}_H \bar{C}'_H + \bar{B}_H H_{t-1} \bar{B}'_H + \bar{A}_H V_{t-1} \bar{A}'_H \\ M_t &= \bar{C}_M \bar{C}'_M + \bar{B}_M M_{t-1} \bar{B}'_M + \bar{A}_M V_{t-1} \bar{A}'_M \end{aligned}$$

dove $\bar{B}_H, \bar{A}_H, \bar{B}_M$ e \bar{A}_M sono matrici $n \times n$ con n^2 parametri e \bar{C}_H e \bar{C}_M sono triangolari basse con $n^* = n(n+1)/2$ parametri. Le matrici H_t e M_t sono semi-definite positive per

ogni t se H_0 e M_0 sono semi-definite positive. Anche questo modello viene stimato con la quasi-massima verosimiglianza, anche se il numero di parametri da stimare è elevato e può comportare problemi di convergenza. Una soluzione a questo problema è imporre che $\bar{B}_H, \bar{A}_H, \bar{B}_M$ e \bar{A}_M siano scalari o matrici diagonali.

L'uso di questi modelli è adatto alle previsioni della volatilità un passo avanti anche se specificazioni più complicate permettono di catturare al meglio le dipendenze nel dataset.

2 Modello di previsione della volatilità con variabili esogene

Lo scopo di questo capitolo è quello di presentare due nuovi modelli econometrici per la stima della volatilità in ambito multivariato, caratterizzati dall'uso di variabili esogene macroeconomiche e finanziarie. In particolare, vengono analizzati un modello vettoriale autoregressivo nei fattori di Cholesky della matrice varianze e covarianze realizzate con l'uso di variabili esogene e un modello vettoriale non lineare sui medesimi fattori. Le previsioni out-of-sample ottenute con questi modelli vengono successivamente confrontate con quelle ottenute mediante altri modelli multivariati di volatilità come un VAR(1) sugli elementi della matrice varianze e covarianze realizzate, un DCC, un BEKK, un VAR sulle log-volatilità definite nel paragrafo 1.7.1, un modello senza variabili esogene e un modello HEAVY, come quello introdotto da Noureldin *et al.* (2011).

2.1 Modello Cholesky-VAR(p) con esogene

Il modello presentato in questa sezione viene utilizzato per prevedere la volatilità dei rendimenti di attività finanziarie in un portafoglio di n titoli. Sia $r_{i,t}$ un vettore di n rendimenti per il giorno i del t -esimo mese, si suppone che la media dei rendimenti mensili, data l'informazione I_{t-1} , non sia prevedibile, cioè $E(r_t | I_{t-1}) = 0$, e che la matrice delle varianze e covarianze sia variabile nel tempo, prevedibile e che possa essere consistentemente stimata dalla matrice delle covarianze realizzate, $RCOV_t$, data da

$$RCOV_t = \sum_{i=1}^m r_{i,t} r'_{i,t} \quad (2.1.1)$$

dove $i = 1, \dots, m$ è il numero di giorni che compone il mese t .

Una volta ottenute le serie delle matrici varianze e covarianze realizzate, si segue una procedura simile a quella presente nel lavoro di Halbleib-Chiriac e Voev (2011) per ottenere le previsioni della matrice delle covarianze.

2.1.1 Modello Cholesky-VAR(p)

In generale sono disponibili due approcci per garantire che la matrice delle covarianze prevista sia almeno semi-definita positiva: da un lato si utilizza una stima vincolata, affinché le matrici siano semi-definite positive; dall'altro si utilizza una riparametrizzazione della matrice varianze e covarianze. Poiché la stima vincolata presenta alcune difficoltà dal punto

di vista computazionale, questo lavoro fa ricorso all'uso di una parametrizzazione¹³ della matrice varianze e covarianze realizzate, in particolare della scomposizione di Cholesky. Diversi lavori ricorrono a questa tecnica, in particolare Tsay (2005a) discute dell'uso di tale scomposizione in un approccio GARCH-multivariato, mentre Pourahmadi (1999) suggerisce di modellare i fattori di Cholesky dell'inverso della matrice varianze e covarianze. Più recentemente, l'idea di modellare i fattori di Cholesky della matrice delle covarianze realizzate è alla base della procedura posta in essere da Halbleib-Chiriac e Voev (2011), Becker *et al.* (2010) e Heiden (2015).

L'uso di tale parametrizzazione permette di avere stime e previsioni della matrice varianze e covarianze realizzate definite positive senza imporre restrizioni sui parametri del modello. Infatti, la matrice triangolare bassa, costruita degli elementi di P_t , genera una matrice definita positiva grazie all'inverso della scomposizione di Cholesky.

La scomposizione di Cholesky non è l'unica parametrizzazione della matrice delle covarianze che garantisce che essa sia garantita almeno semi-definita positiva. Pinheiro e Batels (1996) analizzano diverse parametrizzazioni della matrice varianze e covarianze, tra cui la trasformazione logaritmo di una matrice, utilizzata in Bauer e Vorkink (2011), come definita nella (1.7.6). In questo lavoro si confrontano le previsioni mediante un modello VAR sulle serie delle log-volatilità a_t definite nel paragrafo 1.7.1 con quelle derivanti dalla procedura Cholesky-VAR descritta in precedenza.

La procedura utilizzata in questo lavoro è simile a quella presente nel lavoro di Halbleib-Chiriac e Voev (2011) e prevede i seguenti passaggi:

1. Si costruisce la matrice varianze e covarianze realizzate, $RCOV_t$, come descritto nell'equazione (2.1.1);
2. La matrice $RCOV_t$ viene parametrizzata mediante la scomposizione di Cholesky, per cui

$$RCOV_t = C_t C_t' \quad (2.1.2)$$

dove C_t è una matrice triangolare bassa. Gli elementi di C_t sono tutti reali se il numero di giorni m è superiore al numero delle serie n (si veda Golub e Van Loan (1996)). Il vettore, P_t , di $\tilde{n} = n(n + 1)/2$ elementi si ottiene dalla trasformazione in vettore della matrice triangolare bassa C_t :

$$P_t = vech(C_t); \quad (2.1.3)$$

¹³Un'analisi approfondita delle tecniche di parametrizzazione della matrice varianze e covarianze viene condotta nell'Appendice A.

3. Le serie degli \tilde{n} fattori di Cholesky sono poi modellate mediante un VARX(p):

$$P_t = c + A_1 P_{t-1} + \cdots + A_p P_{t-p} + D_1 X_{t-1} + \cdots + D_q X_{t-q} + \varepsilon_t \quad (2.1.4)$$

dove X_{t-q} è la matrice delle k variabili esogene per il ritardo q -esimo (l'uso delle variabili esogene viene motivato nei prossimi paragrafi).

4. Il VAR stimato nel punto 3 viene utilizzato per prevedere k periodi avanti gli elementi di P , \hat{P}_{t+k} . In seguito viene ricostruita la triangolare bassa prevista \hat{C}_{t+k} .

5. Si produce, infine, una previsione della matrice varianze e covarianze realizzate, data da

$$RCOV_{t+k} = \hat{C}_{t+k} \hat{C}'_{t+k}. \quad (2.1.5)$$

La costruzione della matrice varianze e covarianze realizzate nella (2.1.1) garantisce che le t matrici siano simmetriche e, per $n < m$, definite positive.

La procedura descritta in precedenza comporta una serie di vantaggi nei confronti dei modelli GARCH, specialmente a fini previsivi. In primo luogo, come mostrato da Andersen *et al.* (2003), un vantaggio di questo approccio è quello di produrre previsioni migliori di quelle ottenute con metodi meno diretti, quali i modelli GARCH e quelli di volatilità stocastica. In secondo luogo, i modelli esistenti necessitano di alcune restrizioni ai parametri per poter ottenere una matrice delle covarianze definita positiva, questo può portare a distorsioni nella valutazione dell'effetto delle variabili nella previsione delle varianze e covarianze. La scomposizione di Cholesky permette di evitare qualsiasi tipo di vincolo. In aggiunta, storicamente i modelli GARCH multivariati presentano problemi di stima per matrici di grandi dimensioni, mentre l'uso delle misure realizzate permette di considerare un numero elevato di titoli (si veda Laurent *et al.* (2014)). Infine, i modelli di volatilità realizzata presentano un costo computazionale decisamente minore rispetto ai modelli esistenti, basti pensare che il modello VAR può essere stimato mediante una batteria di OLS, senza la necessità di trovare il minimo locale di una funzione di verosimiglianza.

2.1.2 Stima del modello VAR e previsioni

Come evidenziato nel paragrafo precedente, per poter produrre previsioni k passi in avanti della matrice delle covarianze realizzate, si modellano i fattori di Cholesky mediante un modello multivariato. All'interno di questo paragrafo si analizza il modello lineare autoregressivo vettoriale con l'uso di variabili esogene (VARX).

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Il modello vettoriale da stimare, per p ritardi delle dipendenti e q ritardi delle esogene, è dato da:

$$P_t = c + A_1 P_{t-1} + \cdots + A_p P_{t-p} + D_1 X_{t-1} + \cdots + D_q X_{t-q} + \varepsilon_t \quad (2.1.6)$$

dove c è un vettore $\tilde{n} \times 1$ di costanti, per $\tilde{n} = n(n+1)/2$, P_t è un vettore di $\tilde{n} \times 1$ fattori di Cholesky della matrice varianze e covarianze realizzate, $A_i, \forall i = 1, \dots, p$, è la matrice $\tilde{n} \times \tilde{n}$ di coefficienti autoregressivi per l'ordine di ritardo i -esimo, $D_j, \forall j = 1, \dots, q$, è la matrice $\tilde{n} \times k$ di coefficienti delle variabili esogene per il ritardo q -esimo, il vettore, $\tilde{n} \times 1$, ε è un white noise vettoriale.

Se si considera un solo ritardo per le esogene, il vettore $(k+1) \times 1$ delle esogene più la costante può essere ridefinito come $\delta_t = (x_{t-1}, 1)$, al quale è associata la matrice dei coefficienti Φ , di dimensione $\tilde{n} \times (k+1)$. Se si assume che F_t sia la matrice $\tilde{n} \times (p \cdot \tilde{n} + k + 1)$ tale per cui $F = (A_1, \dots, A_p, \Phi)$ e z_t il vettore $(p \cdot \tilde{n} + q \cdot k + 1) \times 1$, definito come $z_t = (y_{t-1}, \dots, y_{t-p}, \delta_t)$, allora la (2.1.6) diventa

$$y_t = F z_t + \varepsilon_t. \quad (2.1.7)$$

La stima del modello VAR di ordine p viene effettuata mediante il metodo dei minimi quadrati, attraverso il quale le \tilde{n} equazioni vengono stimate singolarmente. In maniera alternativa, la stima dei coefficienti del modello VAR può essere condotta secondo la massimizzazione della funzione di verosimiglianza condizionale.

Il numero dei ritardi in un VAR può essere determinato usando il test F oppure mediante i criteri di informazione. Una volta determinato il numero di ritardi, p , e stimato il modello VAR mediante il metodo OLS, il modello può essere utilizzato per effettuare previsioni. Le previsioni un passo avanti sono date da:

$$\begin{aligned} P_{T+1|T} &= E \left\{ P_{T+1} \mid P_T, P_{T-1}, \dots, X_T, X_{T-1}, \dots \right\} \\ &= \hat{c} + \hat{A}_1 P_T + \cdots + \hat{A}_p P_{T-p+1} + \hat{D}_1 X_T + \cdots + \hat{D}_q X_{T-q+1}, \end{aligned}$$

dove i coefficienti dell'equazione sono le stime OLS del modello VAR. Le previsioni per più periodi in avanti possono essere ottenute in maniera iterativa.

In questo lavoro vengono utilizzate le previsioni un periodo avanti mediante una finestra rolling di osservazioni per produrre previsioni out-of-sample, il metodo viene analizzato nel dettaglio nel paragrafo 2.4.

2.2 Modello Cholesky-VLSTAR con esogene

Alle stime e previsioni del modello lineare si affiancano quelle derivanti da un modello non lineare. La letteratura sulle dinamiche della volatilità presenta diversi modelli non lineari. Nelson (1990), per primo, ha introdotto una componente asimmetrica nel modello GARCH esponenziale, o E-GARCH. Accanto a questo modello sono stati sviluppati diversi modelli GARCH finalizzati a cogliere la non linearità della varianza condizionale, come il Threshold GARCH (TGARCH) di Rabemananjara e Zakoian (1993), il Volatility Switching GARCH (VSGARCH) di Fornari e Mele (1997) o lo Smooth Transition GARCH (STGARCH), basato sul modello autoregressivo Smooth Transition (STAR), di Hagerud (1997) e di Gonzalo-Rivera (1998). I modelli asimmetrici e non lineari sono presenti anche nella classe dei modelli di volatilità stocastica, uno dei primi lavori è quello di Harvey *et al.* (1996). Nella letteratura della volatilità realizzata, invece, raramente sono stati considerati modelli non lineari. Il primo lavoro finalizzato all'introduzione di asimmetrie e non linearità in un modello di varianza realizzata è quello di Martens *et al.* (2004). In seguito, McAleer e Medeiros (2008) riprendono questo lavoro proponendo un modello di lunga memoria che tiene conto della non linearità della serie della varianza realizzata, presentando un test per la non linearità e per il numero di regimi. Più recentemente, Ghalanos e Rossi (2016) analizzano l'uso di un modello Smooth Transition dinamico con l'uso di variabili esogene sul logaritmo della varianza realizzata dell'indice S&P 500.

In questo lavoro viene estesa la letteratura attuale in ambito multivariato, si utilizza un modello non lineare multivariato per il vettore dei fattori di Cholesky della matrice delle covarianze realizzate. Una volta ottenute le previsioni dei fattori di Cholesky, si segue la procedura descritta nella sezione 2.1.1 dal punto 4.

I modelli non lineari multivariati più utilizzati in letteratura sono il modello VAR a soglia (*Vector Threshold Autoregressive*, TVAR), che utilizza una variabile di switch, e il modello vettoriale autoregressivo Smooth Transition (*Smooth Transition VAR*, STVAR), che utilizza una variabile di transizione (per una breve rassegna dei modelli non lineari multivariati si veda Hubrich e Teräsvirta (2013)). Nel presente lavoro si è scelto di utilizzare un modello STVAR.

Il modello TVAR è la generalizzazione vettoriale della regressione switching univariata, definita come

$$y_t = \sum_{i=1}^r (\phi_i' z_t + \varepsilon_{it}) I(c_{i-1} \leq s_t \leq c_i) \quad (2.2.1)$$

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

dove $z_t = (w_t', x_t')$ è un vettore che contiene variabili esplicative, $w_t = (1, y_{t-1}, \dots, y_{t-p})'$, per p ritardi della variabile dipendente, e $x_t = (x_{1t}, \dots, x_{kt})$ è l'insieme delle variabili esogene, s_t è una variabile osservabile di switch, normalmente stazionaria, e c_0, c_1, \dots, c_r sono i valori soglia. La stima dell'equazione (2.2.1) può essere condotta mediante *conditional least squares* (CLS). Nel multivariato Tsay (2005b) specifica il modello come segue:

$$y_t = \sum_{i=1}^r \left\{ \sum_{j=1}^p (\mu_i + \Phi_{ij} y_{t-j} + \Gamma_{ij} x_{t-j}) + \varepsilon_{it} \right\} I(c_{i-1} < s_t \leq c_i) \quad (2.2.2)$$

dove y_t e ε_{it} sono vettori $\tilde{n} \times 1$ e μ_i è un vettore $\tilde{n} \times 1$ di intercette, per $i = 1, \dots, r$ numero di soglie, ϕ_{ij} sono le matrici $\tilde{n} \times \tilde{n}$ di coefficienti dei ritardi del vettore di variabili dipendenti, e Γ_{ij} sono le matrici $\tilde{n} \times k$ di coefficienti delle esogene, per $i = 1, \dots, r$ e $j = 1, \dots, p$. Gli errori ε_{it} sono serialmente incorrelati con media zero e matrice varianze e covarianze definita positiva. Una singola variabile di switch, s_t , stazionaria e continua, determina il regime dell'intero sistema.

Il modello STVAR è, invece, la generalizzazione della regressione smooth transition, proposta da Bacon e Watts (1971), definita come

$$\begin{aligned} y_t = & \phi_0 + \sum_{j=1}^p \phi_j y_{t-j} + \gamma' x_t \\ & + (\psi_0 + \sum_{j=1}^p \psi_j y_{t-j} + \eta' x_t) G(\gamma, c; s_t) + \varepsilon_t \end{aligned} \quad (2.2.3)$$

dove ϕ_j e ψ_j , per $j = 0, 1, \dots, p$ sono parametri, γ e η sono vettori di parametri $k \times 1$, dove k è il numero di esogene, $G(\gamma, c; s_t)$, detta funzione di transizione, è una funzione continua di s_t e ε_t è l'errore i.i.d con media zero e varianza costante σ^2 . In letteratura sono state suggerite diverse specificazioni per la funzione di transizione come una funzione tangente iperbolica, logistica e esponenziale. I modelli STAR possono essere stimati tramite massima verosimiglianza e, sotto alcune condizioni di regolarità, le stime sono consistenti e asintoticamente normali.

Anderson e Vahid (1998) hanno generalizzato il modello STAR nel multivariato come

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

segue:

$$\begin{aligned}
 y_t &= \mu_0 + \sum_{j=1}^p \Phi_j y_{t-j} + \Gamma x_t \\
 &+ G(\gamma, c; s_t) \cdot \left\{ \mu_1 + \sum_{j=1}^p (\Psi_j y_{t-j} + \Theta x_t) \right\} + \varepsilon_t \\
 &= \mu_0 + G(\gamma, c; s_t) \mu_1 + \sum_{j=1}^p \{ \Phi_j + G(\gamma, c; s_t) \Psi_j \} y_{t-j} \\
 &+ (\Gamma + G(\gamma, c; s_t) \Theta) x_t + \varepsilon_t
 \end{aligned} \tag{2.2.4}$$

dove y_t è il vettore di \tilde{n} variabili dipendenti, μ_0 e μ_1 sono i vettori $\tilde{n} \times 1$ di intercette, Φ_j e Ψ_j sono matrici di parametri, per $j = 1, \dots, p$, Γ e Θ sono matrici di parametri $\tilde{n} \times n$ e $G(\gamma, c; s_t)$ è una matrice diagonale $\tilde{n} \times \tilde{n}$ di funzioni di transizione, per cui

$$G(\gamma, c; s_t) = \text{diag}\{G_1(\gamma_1, c_1; s_{1t}), \dots, G_{\tilde{n}}(\gamma_{\tilde{n}}, c_{\tilde{n}}; s_{\tilde{n}t})\}. \tag{2.2.5}$$

Spesso il modello (2.2.4) viene semplificato assumendo che le varie serie condividano la medesima variabile di transizione, s_t , e che i parametri delle funzioni di transizione siano gli stessi per tutte le equazioni.

In questo lavoro si utilizza il modello *Vector Logistic Smooth Transition AutoRegressive* (VLSTAR). Se si considera una generalizzazione non lineare della (2.1.7), il modello può essere riscritto come

$$y_t = \left\{ \sum_{i=1}^m (G_t^{i-1} - G_t^i) F_i \right\} z_t + \varepsilon \tag{2.2.6}$$

dove m è il numero di regimi del modello e gli elementi della matrice G_t^i sono delle funzioni logistiche, tali per cui

$$G(s_{jt}; \gamma_{ij}, c_{ij}) = (1 + \exp\{-\gamma_{ij}(s_{jt} - c_{ij})\})^{-1}, \tag{2.2.7}$$

per $i = 1, \dots, m-1$ e $j = 1, \dots, \tilde{n}$, $G_t^0 = I_{\tilde{n}}$ e $G_t^m = 0$. Un'ulteriore riparametrizzazione del modello VLSTAR è data da

$$y_t = (B_1 + G_t^1 + \dots + G_t^{m-1} B_m) z_t + \varepsilon_t = \Psi_t B z_t + \varepsilon_t \tag{2.2.8}$$

dove $\Psi_t = (I_{\tilde{n}}, G_t^1, \dots, G_t^{m-1})$ è una matrice $\tilde{n} \times m\tilde{n}$, $B = (B_1, B_2, \dots, B_m)$ è una matrice $m\tilde{n} \times (p\tilde{n} + k + 1)$, $B_1 = F_1$ e $B_i = F_i - F_{i-1}$, per $i = 2, \dots, m$. Camacho (2004) ha utilizzato questo modello per due serie con un numero di regimi pari a 2, per cui

$$y_t = (B_1 + G_t^1 B_2) + \varepsilon_t. \tag{2.2.9}$$

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Se si assume che $\varepsilon_t \sim N(0, \Omega)$, dove Ω è la matrice delle covarianze dei termini di errore, il modello può essere rappresentato dalla funzione di densità condizionale multivariata

$$f(y_t | I_{t-1}) = (2\pi)^{-\frac{\tilde{n}}{2}} |\Omega|^{-\frac{1}{2}} \exp \left\{ - (1/2)(y_t - \Psi_t B z_t) \Omega^{-1} (y_t - \Psi_t B z_t) \right\} \quad (2.2.10)$$

dove I_{t-2} è il set informativo al tempo $t-1$. La funzione di log-verosimiglianza condizionale è data da

$$\begin{aligned} \log L(\theta) = & - (T\tilde{n}/2) \log 2\pi - (T/2) \log |\Omega| \\ & - (1/2) \sum_{t=1}^T (y_t - \Psi_t B z_t) \Omega^{-1} (y_t - \Psi_t B z_t). \end{aligned} \quad (2.2.11)$$

Il set di parametri da stimare è dato da $\theta = \{B, \Omega, \Gamma, C\}$, dove $\Gamma = [\gamma_{ij}]$ e $C = [c_{ij}]$.

La stima del modello (2.2.8) può essere condotta sia mediante minimi quadrati non lineari, sia mediante massima verosimiglianza. Una volta scelta la variabile di transizione unica per il modello, s_t , o il set di variabili di transizione, $S = \text{diag}(s_{1t}, \dots, s_{\tilde{n}t})$, lo stimatore dei minimi quadrati non lineare è il risultato del seguente problema di ottimo

$$\hat{\theta}_{NLS} = \arg \min_{\theta} \sum_{t=1}^T (y_t - \Psi_t B z_t)' (y_t - \Psi_t B z_t) \quad (2.2.12)$$

L'algoritmo di ottimizzazione può convergere ad un minimo locale differente da quello globale, di conseguenza assume notevole importanza l'insieme dei valori di partenza di θ , per un approfondimento sulla griglia di ricerca dei valori iniziali si veda Teräsvirta e Yang (2014).

Lo stimatore di massima verosimiglianza è la soluzione al problema di ottimo

$$\hat{\theta}_{ML} = \arg \max_{\theta} \log L(\theta) \quad (2.2.13)$$

dove la funzione $\log L(\theta)$ equivale alla (2.2.11), se $\varepsilon_t \sim N(0, \Omega)$.

Nel caso in cui le funzioni di transizione cambino da equazione a equazione, la stima può essere condotta per ogni singola equazione come nel caso univariato.

Le previsioni dei modelli non lineari per più passi in avanti possono essere condotte mediante tecniche numeriche. Sia dato un modello non lineare

$$y_t = g(z_t, \theta) + \varepsilon_t \quad (2.2.14)$$

dove θ è un vettore di parametri da stimare, z_t può rappresentare una combinazione dei valori ritardati di y_t e di variabili esogene x_{1t}, \dots, x_{kt} e ε_t è un white noise con media nulla

e varianza costante pari a σ^2 . Sia I_T il set informativo al tempo T e sia ε_t indipendente da I_{T-1} , la previsione al tempo T equivale alla media condizionale

$$y_{T+h|T} = E\{y_{T+h} | I_T\} = E\{g(z_{T+h-1}) | I_T\} \quad (2.2.15)$$

quando $h = 1$ nella (2.2.15), $y_{T+1|T} = g(z_T)$. Se $h \geq 2$, il valore atteso condizionale (2.2.15) può essere solo calcolato in maniera ricorsiva usando tecniche numeriche. Per una analisi approfondita sulle previsioni dei modelli non lineari si vedano Hubrich e Teräsvirta (2013), Kock e Teräsvirta (2011) e Teräsvirta *et al.* (2010).

2.2.1 Presenza di break strutturali e test di linearità

La procedura che porta alla stima di un modello VLSTAR e alle previsioni prodotte dallo stesso prevede diversi passaggi. In primo luogo, si indaga sulla presenza di break strutturali nelle serie dei fattori di Cholesky per determinare il numero di regimi, m , del modello. In seguito si passa alla specificazione del modello non lineare. La scelta della struttura del modello non lineare implica la selezione della variabile di transizione (o delle variabili di transizioni nel caso siano diverse per ogni fattore di Cholesky). La scelta della variabile di transizione può essere basata sulla base della teoria economica che accompagna il modello o può essere guidata dai dati. Se si ipotizza che ci siano molteplici variabili di transizione potenziali, una scelta comune è quella di testare la linearità del modello per ogni variabile di transizione e scegliere quella che produce il p -value minore nel test di linearità. In ambito multivariato, nel caso in cui si assuma che la variabile di transizione sia la stessa per l'intero modello, è possibile utilizzare un test congiunto per la linearità del modello, in caso contrario tale test viene effettuato equazione per equazione, come in Camacho (2004), Luukkonen *et al.* (1988) e Teräsvirta e Yang (2014).

Per verificare la presenza di cambiamenti nel valore medio dei fattori di Cholesky viene utilizzato il metodo introdotto da Bai e Perron (1998, 2003a), B-P. Il test B-P verifica la presenza di molteplici cambiamenti strutturali in un modello di regressione lineare. Per le finalità di questo lavoro, viene considerato il modello con m break strutturali e $m + 1$ regimi, dato da

$$y_{it} = \mu_{ij} + u_{it} \quad (2.2.16)$$

dove $j = 1, \dots, m + 1$, $i = 1, \dots, \tilde{n}$, μ_{ij} è la media di y_{it} . Il numero e i punti di break non sono considerati noti. Per determinare il numero di break viene usata la statistica test $\sup F_T(l)$, con l'ipotesi nulla di assenza di break contro quella alternativa di un numero di

break arbitrario fino ad un numero massimo di break M . B-P hanno introdotto due test aventi come ipotesi nulla l'assenza di break strutturali, per testare la presenza di break senza specificare un'ipotesi alternativa (si veda Bai e Perron (2003b)). Il primo test è una versione equipesata delle statistiche test F_T , definito come

$$UD_{max} = \max_{1 \leq m \leq M} F_T(\hat{\lambda}_1, \dots, \hat{\lambda}_m)$$

dove $\hat{\lambda}_j = T_j/T$, $j = 1, \dots, m$, mentre l'altro test è definito come

$$WD_{max} = \max_{1 \leq m \leq M} w_m F_T(\hat{\lambda}_1, \dots, \hat{\lambda}_m)$$

dove i pesi w_m sono tali per cui i p -value marginali sono uguali per ogni m . L'ipotesi nulla di entrambi i test è l'assenza di break strutturali contro l'alternativa di un numero non noto di break, dato un limite massimo M . Infine, il test $\sup F_T(l+1 | l)$ consente di selezionare il numero di break ottimale. Pertanto, si procede prima alla determinazione del numero di break, equazione per equazione, mediante i test UD_{max} e WD_{max} , si seleziona poi il numero di break ottimali mediante il test $\sup F_T(l+1 | l)$. I risultati dei test sui break strutturali sono riportati nelle sezioni 2.3.6 e 2.3.7.

Una volta testata la presenza di break strutturali, si procede alla specificazione del modello non lineare. La scelta tra linearità e non linearità assume una notevole importanza in quanto è possibile che un modello lineare descriva in maniera adeguata le relazioni considerate, facilitando il processo di stima dello stesso. In aggiunta, un gran numero di modelli non lineari non è identificato quando il processo generatore dei dati è lineare. In letteratura esiste un grande numero di test di linearità, per cui questo lavoro cerca di presentare solo alcuni di essi.

La linearità può essere testata mediante un test Lagrange multiplier (LM), equazione per equazione. Luukkonen *et al.* (1988) hanno suggerito di approssimare la funzione di transizione mediante un'espansione di Taylor e di testare l'ipotesi nulla di linearità, data da $\gamma_j = 0$ nella funzione di transizione (2.2.7).

Se le equazioni condividono la medesima variabile di transizione, è possibile applicare anche un test congiunto. Sia dato un modello VLSTAR a 2 regimi del tipo (2.2.9), l'ipotesi nulla di non linearità equivale a $\gamma_j = 0$, $j = 1, \dots, \tilde{n}$, mentre l'ipotesi alternativa è che almeno un parametro di forma, γ_j , sia maggiore di 0. Il test, come quello nell'univariato sviluppato da Luukkonen *et al.* (1988), si basa sull'espansione di Taylor della funzione di transizione ed è analizzato nel dettaglio nell'appendice B.

2.3 Analisi empirica

In questa sezione vengono presentate le stime e, in particolare, le previsioni dei modelli introdotti nella sezione precedente: il VAR sui fattori di Cholesky della matrice varianze e covarianze e lo STVAR sugli stessi fattori. La peculiarità di questo lavoro è quella di verificare la capacità previsiva di alcune variabili macroeconomiche e finanziarie esogene.

Per ottenere le previsioni della matrice varianze e covarianze realizzate si utilizzano metodi tradizionali basati su una funzione di perdita come la radice quadrata dell'errore quadratico medio (*Root mean squared error*, RMSE) e la regressione Mincer-Zarnowitz. Per valutare diversi modelli contemporaneamente si ricorre ai test di confronto tra due modelli, come quello di Diebold-Mariano (1995) o quello di Giacomini-White (2006). Infine si utilizza una metodologia recente, sviluppata in Hansen *et al.* (2011b) col nome di *Model Confidence Set*, per determinare l'ordinamento di un insieme di modelli in termini previsionali.

2.3.1 Determinanti macroeconomiche e finanziarie

L'obiettivo di questa sezione è individuare le determinanti della volatilità e la loro utilità nel prevedere la matrice varianze e covarianze. Schwert (1989), per primo, ha analizzato il comportamento della volatilità rispetto al ciclo economico. Dal suo studio emerge che la relazione tra le variazioni delle variabili macroeconomiche e la volatilità di un titolo non è particolarmente forte, ma mostra comunque un andamento anti-ciclico da parte della volatilità. Lo stesso risultato è stato confermato dai lavori di Fama e French (1989), per quanto riguarda il premio per il rischio, e da Conrad e Loch (2014), in un lavoro sulla previsione della volatilità mediante l'approccio MIDAS.

Recentemente, grazie alla diffusione della varianza realizzata come misura di volatilità e dell'approccio MIDAS, una larga parte di letteratura si è occupata del legame tra la volatilità dei titoli e i fondamentali macroeconomici del Paese di riferimento. In particolare, Engle e Rangel (2008) fanno uso di uno Spline-GARCH per analizzare il collegamento tra alcune variabili macroeconomiche e la volatilità a bassa frequenza. I loro risultati evidenziano che alcune variabili macroeconomiche, come la crescita del PIL, l'inflazione e il tasso di interesse a breve, aiutano a spiegare le dinamiche della volatilità. Engle *et al.* (2009) estendono l'uso di determinanti macroeconomiche all'approccio MIDAS. Tra le variabili macroeconomiche utilizzate anche da Schwert (1989), essi prendono in considerazione la crescita della produzione industriale e l'inflazione, concentrandosi sul potere previsionale

di tali variabili. Essi concludono che l'utilizzo delle determinanti macroeconomiche aiuta a prevedere le dinamiche della componente di lungo periodo della volatilità.

Nel suo studio del premio per il rischio, Mele (2007, 2008) suggerisce che le variabili utilizzate per prevedere i rendimenti possono essere usate per prevedere anche la volatilità. Più recentemente, Paye (2012) e Christiansen *et al.* (2012) si sono occupati della prevedibilità della volatilità di un titolo mediante variabili macroeconomiche e finanziarie. Paye (2012) evidenzia che esiste un legame tra diverse variabili esogene e la volatilità di mercato, mentre la capacità previsiva out-of-sample risulta modesta. Christiansen *et al.* (2012) estendono l'analisi di Paye (2012) considerando un maggior numero di variabili esogene e analizzando, oltre alla volatilità di mercato, la volatilità di ulteriori attività finanziarie come il tasso di cambio, il mercato obbligazionario e delle commodity. Dal loro lavoro emerge che alcune variabili aiutano a migliorare le previsioni della volatilità. Le variabili particolarmente utili nella previsione sono quelle legate alla variabilità del premio per il rischio e al rischio di credito, mentre le variabili macroeconomiche "pure" non mostrano un miglioramento significativo delle previsioni della volatilità finanziaria.

La letteratura non ha chiarito in pieno l'utilità delle variabili finanziarie e macroeconomiche nella previsione della volatilità di mercato. Il motivo di questo risultato potrebbe essere dettato dalla natura lineare dei modelli utilizzati, questo implica che la relazione tra variabili esogene e volatilità potrebbe essere di natura non lineare. A sostegno di tale ipotesi, Mele (2007) mostra come la volatilità dei rendimenti sia contro-ciclica in seguito alle risposte asimmetriche del premio per il rischio in risposta alle variazioni nelle condizioni economiche. Per questo motivo, in questo lavoro si analizza il ruolo delle determinanti macroeconomiche e finanziarie in un contesto non lineare.

Questo lavoro fa ricorso alle variabili esogene ritenute maggiormente significative in letteratura per la previsione out-of-sample della volatilità e analizza, come suggerito da Diebold e Yilmaz (2009), gli effetti spillover tra diversi Paesi.

2.3.2 Dati cross-country

Per la costruzione della matrice varianze e covarianze realizzate nel modello cross-country sono utilizzate le osservazioni dei prezzi di chiusura giornalieri di 4 indici azionari¹⁴. Le serie disponibili consistono nelle osservazioni giornaliere degli indici di mercato S&P 500

¹⁴L'ordine degli indici nel vettore dei rendimenti, r_i , è SP, JPN, UK, GER. Sono state utilizzate anche combinazioni degli indici diverse, ma le caratteristiche delle serie dei fattori di Cholesky derivanti non cambiano.

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

(SP), Nikkei (JPN), FTSE 100 (UK) e DAX (GER). La scelta delle serie è motivata dalla digitalizzazione avanzata dei mercati che ha reso possibile, per gli investitori, detenere nel proprio portafoglio di investimento attività quotate in diversi mercati e dalla necessità di verificare gli effetti spillover tra le quotazioni degli indici di Paesi differenti. I dati originali giornalieri coprono il periodo compreso tra l'1 Gennaio 1969 e il 31 Ottobre 2015 per un totale di 11186 osservazioni giornaliere.

Tutte le osservazioni giornaliere dei rendimenti delle serie sono utilizzate per costruire $T = 562$ matrici varianze e covarianze realizzate mensili, secondo l'equazione (2.1.1), per cui

$$RCOV_t = \sum_{i=1}^{m_t} r_{i,t} r'_{i,t}$$

dove m_t è il numero di giorni nel mese t e $r_{i,t}$ è un vettore $n \times 1$ di rendimenti giornalieri dati da

$$r_{i,t} = p_{i,t} - p_{i-1,t}, \quad i = 1, \dots, m_t.$$

Per costruzione, le matrici delle varianze e covarianze realizzate sono simmetriche e, per $n < m_t$, sono definite positive.

Tabella 2.1: Statistiche descrittive Varianze e Covarianze cross-country

Varianze realizzate dataset cross-country									
Simbolo	Media	Dev.St	Asimm.	Ex. Curt.	ρ_1	ρ_2	ADF		KPSS
							c	c-t	
SP	0.0024	0.0048	9.85	2.02	0.45	0.30	0.0000	0.0000	0.0407
JPN	0.0035	0.0058	9.67	1.68	0.36	0.20	0.0000	0.0000	0.0716
UK	0.0024	0.0041	6.58	1.64	0.56	0.33	0.0000	0.0000	0.1131
DAX	0.0034	0.0049	4.98	37.42	0.59	0.40	0.0000	0.0000	0.0754
Covarianze realizzate dataset cross-country									
Simbolo	Media	Dev.St	Asimm.	Ex. Curt.	ρ_1	ρ_2	ADF		KPSS
							c	c-t	
SP-JPN	0.0004	0.0016	14.51	276.84	0.24	0.04	0.0000	0.0000	0.0312
SP-UK	0.0009	0.0023	7.95	81.24	0.56	0.32	0.0000	0.0000	0.0669
SP-DAX	0.0012	0.0031	7.34	75.79	0.56	0.36	0.0000	0.0000	0.0729
JPN-UK	0.0008	0.0028	14.06	242.58	0.34	0.12	0.0000	0.0000	0.0369
JPN-DAX	0.0010	0.0028	13.90	256.10	0.37	0.13	0.0000	0.0000	0.0394
UK-DAX	0.0016	0.0033	6.64	67.60	0.64	0.38	0.0000	0.0000	0.0633

a , b e c indicano un KPSS significativo rispettivamente all'1%, 5% e 10%

La tabella 2.1 contiene una serie di statistiche descrittive della varianze e covarianze realizzate degli indici considerati. Dalla tabella emerge che le distribuzioni delle varianze e delle covarianze realizzate sono estremamente asimmetriche a destra e leptocurtiche, co-

Tabella 2.2: Statistiche descrittive fattori di Cholesky cross-country

Fattori di Cholesky dataset cross-country									
Simbolo	Media	Dev.St	Asimm.	Ex. Curt.	ρ_1	ρ_2	ADF		KPSS
							c	c-t	
Cholesky 1a	0.0421	0.0243	3.88	25.32	0.63	0.51	0.0000	0.0000	0.0522
Cholesky 2a	0.0069	0.0145	1.88	13.25	0.13	0.12	0.0000	0.0000	0.0428
Cholesky 3a	0.0155	0.0173	1.92	6.98	0.53	0.39	0.0308	0.0000	0.1315 ^c
Cholesky 4a	0.0192	0.0259	2.02	6.04	0.65	0.56	0.0021	0.0000	0.1551 ^b
Cholesky 5a	0.0489	0.0280	2.32	11.61	0.58	0.43	0.0001	0.0000	0.1614 ^b
Cholesky 6a	0.0079	0.0143	2.79	19.12	0.27	0.17	0.0000	0.0000	0.0576
Cholesky 7a	0.0099	0.0149	1.99	8.61	0.28	0.16	0.0000	0.0000	0.2297 ^a
Cholesky 8a	0.0355	0.0190	2.64	11.40	0.62	0.52	0.0000	0.0000	0.1330 ^c
Cholesky 9a	0.0172	0.0194	1.21	2.53	0.56	0.47	0.0543	0.0000	0.1786 ^b
Cholesky 10a	0.0334	0.0151	1.44	3.11	0.54	0.47	0.0001	0.0018	0.2365 ^a

a, *b* e *c* indicano un KPSS significativo rispettivamente all'1%, 5% e 10%

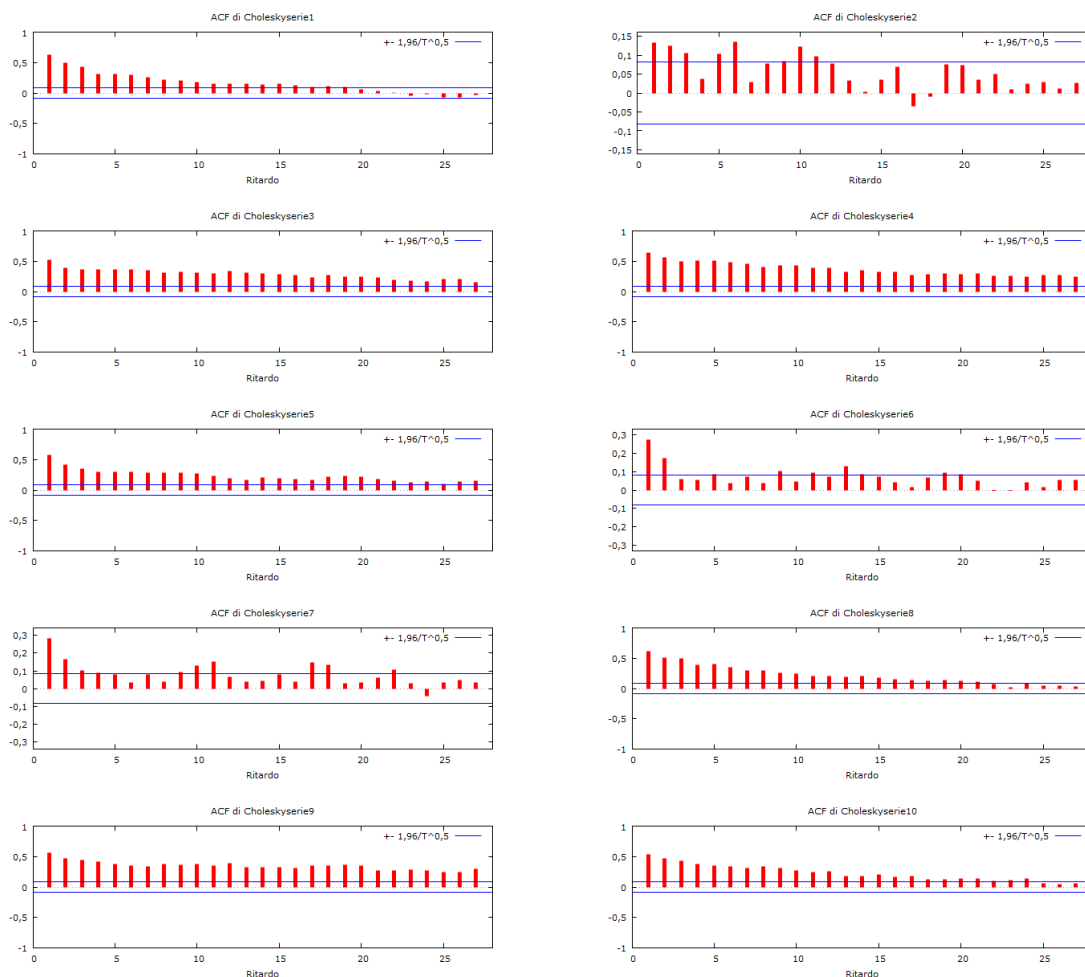
me documentato in Andersen *et al.* (2001a). Per verificare la persistenza delle serie delle varianze e covarianze sono riportate anche le autocorrelazioni campionarie di primo e di secondo ordine. Esse sono positive per tutte le serie delle covarianze realizzate, evidenziando la presenza di una persistenza relativamente forte. In aggiunta le ultime colonne della tabella riportano i *p*-value del test Augmented Dickey-Fuller (ADF) e i valori della statistica test KPSS. Il test ADF rifiuta l'ipotesi nulla di non stazionarietà per tutte le serie delle varianze e delle covarianze, mentre il KPSS accetta, per tutte le serie, l'ipotesi nulla di serie stazionarie.

Le caratteristiche delle serie dei fattori di Cholesky, derivanti dalla scomposizione (2.1.2), sono analizzate nella tabella 2.2. Si può osservare che le serie conservano la proprietà di lunga memoria delle serie delle covarianze realizzate, come evidenziato da Andersen e Bollerslev (1997) e Andersen *et al.* (2001a). Le autocorrelazioni delle serie dei fattori di Cholesky del primo dataset sono riportate nella figura 2.1. La figura mostra come le autocorrelazioni decadono lentamente, come per le serie delle covarianze realizzate. Il test ADF di radice unitaria con l'uso della sola costante mostra come tutte le serie rifiutino l'ipotesi nulla ad un livello del 10%, mentre 9 serie su 10 la rifiutano al 5%. In diverse serie dei fattori, il KPSS rifiuta l'ipotesi nulla di assenza di radici unitarie. Questo suggerisce la presenza di una certa persistenza nelle serie e la necessità di applicare un modello più complesso rispetto al VAR lineare.

Per il dataset cross-country è stato utilizzato un insieme di 4 variabili esplicative. Le variabili macroeconomiche utilizzate in questo lavoro, come l'inflazione (INF) e il tasso

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Figura 2.1: Autocorrelogramma delle serie dei fattori di Cholesky della matrice varianze e covarianze realizzate del dataset cross-country, RCOV.



di crescita della produzione industriale mensili (IP), sono le stesse utilizzate da Schwert (1989) e mirano a spiegare il movimento anti-ciclico della volatilità. Sono considerati i tassi di crescita dell'inflazione per ogni singolo Paese, mentre vengono considerati i tassi di crescita della produzione industriale per gli Stati Uniti, per il Giappone e per l'intera Unione Europea, per tenere conto di ulteriori effetti spillover con altri Paesi europei.

L'insieme delle variabili esplicative comprende anche due misure di valutazione dei titoli azionari come il logaritmo del rapporto tra i dividendi e il prezzo (D-P) e il logaritmo del rapporto prezzo-utili (E-P) dell'indice di borsa americano S&P 500, tipicamente incluse nell'insieme di predittori del rendimento extra, come evidenziato in Welch e Goyal (2008).

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

La tabella 2.3 presenta una descrizione delle variabili esogene utilizzate e indica la provenienza dei dati, mentre la 2.4 presenta le statistiche descrittive delle esogene.

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.3: Descrizione Variabili Esogene dataset cross-country

Simbolo	Variabile	Origine Dati		
		Descrizione		Fonte
ΔInflUS	Crescita Inflazione Stati Uniti	Differenza prima del tasso di inflazione americano		Datastream
ΔInflUK	Crescita Inflazione Gran Bretagna	Differenza prima del tasso di inflazione inglese		Datastream
$\Delta\text{InflGER}$	Crescita Inflazione Germania	Differenza prima del tasso di inflazione tedesco		Datastream
$\Delta\text{InflJPN}$	Crescita Inflazione Giappone	Differenza prima del tasso di inflazione giapponese		Datastream
IPUS	Crescita Produzione Industriale Stati Uniti	Crescita della produzione industriale americana		Database OECD
IPJPN	Crescita Produzione Industriale Giappone	Crescita della produzione industriale giapponese		Database OECD
IPEU	Crescita Produzione Industriale Unione Europea	Crescita della produzione industriale dell'Unione Europea		Database OECD
d_DY	Crescita Dividend Yield Ratio S&P 500	Crescita dividendi degli ultimi 12 mesi relativi ai prezzi correnti di mercato; indice S&P500		Sito di Robert Shiller
d_EP	Crescita Earning Price Ratio S&P 500	Crescita guadagni degli ultimi 12 mesi relativi ai prezzi correnti di mercato; indice S&P500		Sito di Robert Shiller

Tabella 2.4: Statistiche descrittive determinanti macroeconomiche e finanziarie cross-country

Simbolo	Media	Dev.St	Asimm.	Curt.	ρ_1	ρ_2	ADF		KPSS
							c	c-t	
ΔInflUS	-0.0078	0.3863	0.02	5.06	-0.27	-0.16	0.0000	0.0000	0.0130
ΔInflUK	-0.0104	0.5719	-0.31	15.11	0.40	0.32	0.0001	0.0011	0.0431
$\Delta\text{InflGER}$	-0.0034	0.3252	-0.27	2.46	0.02	0.04	0.0000	0.0000	0.0464
$\Delta\text{InflJPN}$	-0.0066	0.6518	-0.35	8.75	0.22	0.09	0.0000	0.0000	0.0328
IPUS	0.0018	0.0074	-1.18	5.35	0.36	0.29	0.0000	0.0000	0.0461
IPJPN	0.1852	1.6852	-2.16	16.97	0.03	0.16	0.0000	0.0000	0.0257
IPEU	0.1578	0.9226	-0.56	2.24	-0.10	0.13	0.0000	0.0000	0.0266
d_DY	-0.0006	0.0375	0.78	3.51	0.27	0.03	0.0000	0.0000	0.0981
d_EP	-0.0003	0.0615	1.77	39.53	0.55	0.35	0.0000	0.0000	0.0330

a, b e c indicano un KPSS significativo rispettivamente all'1%, 5% e 10%

2.3.3 Dati mercato americano

Per la costruzione delle serie delle matrici varianze e covarianze realizzate nel modello sul mercato americano sono state utilizzate le osservazioni dei prezzi¹⁵ di chiusura giornalieri dell'indice S&P 500 (SP), del future sui Treasury Note a 10 anni (10Y) e di quello sui Fed Funds a 30 giorni (30d), negoziati sul Chicago Board of Trade¹⁶ (CBOT). I dati coprono il periodo compreso tra il 01 Ottobre 1988 e il 30 Giugno 2015 per un totale di 6739 osservazioni giornaliere. La costruzione delle matrici delle covarianze realizzate avviene in maniera analoga al modello cross-country, per un totale di $T = 321$ osservazioni mensili. Nella tabella 2.5 sono riportate alcune statistiche descrittive delle varianze e covarianze realizzate. Le distribuzioni delle varianze realizzate sono altamente asimmetriche a destra e leptocurtiche, mentre le covarianze realizzate tra i rendimenti dell'indice di borsa con quelli dei future sui tassi a breve e lungo termine sono altamente asimmetriche a sinistra. Tutte le serie mostrano autocorrelazioni che decadono lentamente, mentre la serie della varianza realizzata dei future sui Treasury a 10 anni non risulta persistente. Il test ADF di radice unitaria rifiuta l'ipotesi nulla di non stazionarietà, ad un livello di significatività del 5%, per tutte le serie delle varianze e covarianze realizzate, mentre il KPSS rifiuta l'ipotesi nulla di assenza di radici unitarie per tutte le serie delle covarianze al 10%. Questo risultato può essere legato alla differente natura dei dati che portano alla definizione della matrice delle covarianze realizzate.

La tabella 2.6 riporta le statistiche descrittive delle serie dei fattori di Cholesky, anche in questo contesto le serie sembrano mantenere le proprietà mostrate dalle serie delle covarianze realizzate. La figura 2.2 mostra come le autocorrelazioni delle serie dei fattori di Cholesky del secondo dataset decadano lentamente come nel caso del primo dataset. Il test di radice unitaria sulle serie dei fattori di Cholesky rifiuta l'ipotesi nulla di non stazionarietà per la maggioranza delle serie ad un livello di significatività del 5%, il test sul secondo fattore di Cholesky presenta un p-value maggiore di 0.05 quando si considera un modello con una costante e il test sul terzo fattore di Cholesky presenta un p-value maggiore di 0.10 quando si considera un modello con costante e trend. Emergono risultati contrastanti dal test KPSS che conferma la necessità di una specificazione più complessa

¹⁵L'ordine dei titoli nel vettore dei rendimenti è SP, 10Y, 30d. Come nel caso cross-country, invertendo l'ordine dei titoli le caratteristiche delle serie dei fattori di Cholesky, derivanti dalla procedura descritta in precedenza, non cambiano.

¹⁶Si utilizzano i prezzi di chiusura dei future sulle obbligazioni disponibili su Datastream. Mediante l'utilizzo dei contratti future, i rendimenti obbligazionari possono essere calcolati direttamente.

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.5: Statistiche descrittive Varianze e Covarianze mercato americano

Varianze realizzate dataset mercato americano									
Simbolo	Media	Dev.St	Asimm.	Curt.	ρ_1	ρ_2	ADF		KPSS
							c	c-t	
SP	0.0026	0.0047	6.99	1.77	0.71	0.46	0.0000	0.0000	0.0817
10Y	0.0004	0.0007	14.47	236.56	0.07	0.03	0.0000	0.0000	0.0825
30d	0.0000	9.3e-6	3.79	20.00	0.52	0.45	0.0000	0.0000	0.0904

Covarianze realizzate dataset mercato americano									
Simbolo	Media	Dev.St	Asimm.	Curt.	ρ_1	ρ_2	ADF		KPSS
							c	c-t	
SP-10Y	-0.0001	0.0005	-2.64	17.04	0.56	0.59	0.0005	0.0001	0.1246 ^c
SP-30d	-9.7e-6	6.6e-5	-4.70	39.13	0.43	0.37	0.0277	0.0335	0.1572 ^b
10Y-30d	2.4e-5	2.7e-5	2.45	9.53	0.42	0.39	0.0024	0.0048	0.1845 ^b

a, b e c indicano un KPSS significativo rispettivamente all'1%, 5% e 10%

Tabella 2.6: Statistiche descrittive fattori di Cholesky mercato americano

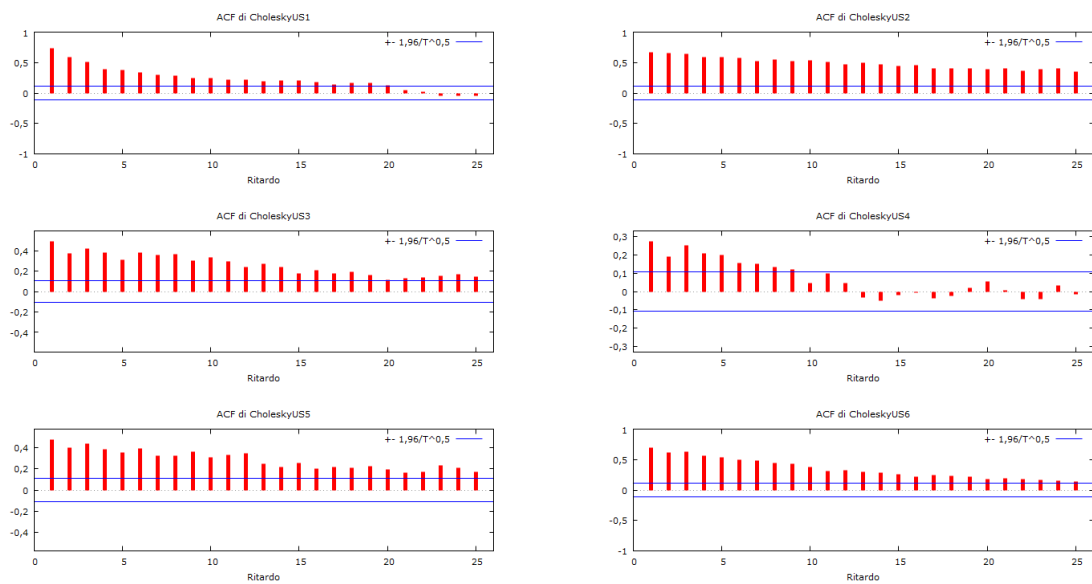
Fattori di Cholesky dataset mercato americano									
Simbolo	Media	Dev.St	Asimm.	Curt.	ρ_1	ρ_2	ADF		KPSS
							c	c-t	
Cholesky 1b	0.0443	0.0262	2.96	13.81	0.74	0.60	0.0005	0.0029	0.1208 ^c
Cholesky 2b	-0.0007	0.0086	-0.05	0.01	0.67	0.66	0.0815	0.0001	0.1518 ^b
Cholesky 3b	6.2e-6	0.0010	0.01	5.89	0.49	0.38	0.0261	0.1120	0.1872 ^b
Cholesky 4b	0.0158	0.0076	6.31	70.87	0.28	0.19	0.0000	0.0001	0.0864
Cholesky 5b	0.0011	0.0009	0.76	1.12	0.48	0.39	0.0001	0.0000	0.2975 ^a
Cholesky 6b	0.0013	0.0011	2.18	7.88	0.70	0.61	0.0108	0.0034	0.1339 ^c

a, b e c indicano un KPSS significativo rispettivamente all'1%, 5% e 10%

per le serie dei fattori di Cholesky.

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Figura 2.2: Autocorrelogramma delle serie dei fattori di Cholesky della matrice varianze e covarianze realizzate del dataset sul mercato americano, RCOVUS.



2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Per l'analisi del mercato statunitense è stato utilizzato un insieme di 11 variabili esplicative. Alle variabili in uso anche nel dataset cross-country, come l'inflazione, la crescita della produzione industriale e le due variabili finanziarie (crescita *dividend yield* e *earning price*), si aggiungono ulteriori variabili finanziarie, principalmente legate alla prevedibilità dei rendimenti extra azionari.

In particolare, nel dataset sul mercato americano vengono inseriti i fattori di rischio di Fama e French (1993), l'High Book Minus Low (HML) e lo Small Minus Big (SMB), utilizzati per misurare gli extra-rendimenti delle imprese a bassa capitalizzazione rispetto a quelle ad alta capitalizzazione e delle azioni value rispetto alle azioni growth. Viene incluso anche il rendimento extra del mercato americano (MKT), per tenere conto dell'effetto leva, per il quale i rendimenti negativi tendono a prevedere una volatilità maggiore. La lista delle variabili esplicative comprende anche il fattore Short Term Reversal (STR), il quale tiene conto della *mean reversion* dei rendimenti azionari. In particolare, questo fattore coglie la componente dei rendimenti azionari non spiegata dai fondamentali delle azioni, si veda Nagel (2012).

Tra le variabili esplicative sono presenti anche il term spread (T-S), il quale risulta essere un buon predittore del rendimento obbligazionario, e il tasso T-Bill (T-B), predittore del rendimento extra azionario. Infine, si tiene conto anche di una misura del rischio di credito obbligazionario americano mediante il default spread (DEF), pari alla differenza tra il rendimento di un'obbligazione con rating BAA e una con rating AAA. Secondo modelli come quello di Merton (1974), il rischio di credito tende a salire quando il grado di leva finanziaria aumenta, pertanto la volatilità è correlata ad un maggior grado di leva finanziaria e di conseguenza ad un maggiore rischio di credito.

Una descrizione delle variabili esogene utilizzate viene presentata nella tabella 2.7, la tabella 2.8 presenta alcune statistiche descrittive delle serie delle determinanti.

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.7: Descrizione Variabili Esogene mercato americano

Simbolo	Variabile	Origine Dati	
		Descrizione	Fonte
D-P	Crescita Dividend Yield Ratio S&P 500	Crescita dividendi degli ultimi 12 mesi relativi ai prezzi correnti di mercato; indice S&P500	Sito di Robert Shiller
E-P	Crescita Earning Price Ratio S&P 500	Crescita guadagni degli ultimi 12 mesi relativi ai prezzi correnti di mercato; indice S&P500	Sito di Robert Shiller
MKT	Rendimento extra mercato US	Fattore di mercato di Fama-French: rendimento mercato azionario U.S. meno il tasso T-Bill a un mese	Sito di Kenneth French
SMB	Fattore di dimensione	Fattore SMB di Fama-French: Rendimento azionario delle piccole imprese meno rendimento azionario grandi imprese	Sito di Kenneth French
HML	Fattore di valore	Fattore HML di Fama-French: Rendimento azioni value meno rendimento azioni growth	Sito di Kenneth French
STR	Fattore Short Term Reversal	Fattore Short Term Reversal di Fama-French: Rendimento azioni con rendimento mese precedente basso meno rendimento azioni con rendimento mese precedente alto	Sito di Kenneth French
$\Delta T-B$	Tasso T-Bill	Differenza prima tasso T-Bill a 3 mesi	Datastream
$\Delta T-S$	Term Spread	Differenza prima della differenza tra tassi a lungo termine e a breve termine	Datastream
ΔDEF	Default Spread	Differenza prima della misura di rischio di default data dalla differenza tra rendimenti di obbligazioni BAA e AAA	Datastream
ΔIP	Crescita Produzione Industriale US	Differenza prima della produzione industriale	Federal Reserve Bank of St. Louis
ΔINF	Crescita Inflazione US	Differenza prima del tasso di inflazione	Datastream

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.8: Statistiche descrittive determinanti macroeconomiche e finanziarie mercato americano

La tabella riporta media, deviazione standard, asimmetria e curtosi per ogni variabile, in aggiunta vengono riportate le autocorrelazioni campionarie fino al secondo ordine (ρ_1 e ρ_2). Le ultime due colonne riportano il p -value del test di radice unitaria Augmented Dickey-Fuller (ADF) con costante (c) e con costante e trend (c-t).

Simbolo	Media	Dev.St	Asimm.	Curt.	ρ_1	ρ_2	ADF		KPSS
							c	c-t	
DP	-0.0016	0.0361	0.91	5.41	0.25	0.03	0.0000	0.0000	0.0648
EP	-0.0019	0.0716	1.89	36.47	0.59	0.39	0.0000	0.0001	0.0272
MKT	0.6317	4.2832	-0.64	1.18	0.07	-0.02	0.0000	0.0000	0.0818
SMB	0.1094	3.2256	0.84	8.26	-0.06	0.04	0.0000	0.0000	0.0749
HML	0.1638	3.0489	0.14	3.09	0.14	0.04	0.0000	0.0000	0.0691
STR	0.1797	3.5367	0.26	5.21	-0.05	-0.10	0.0000	0.0000	0.1214 ^c
Δ TB	-0.0002	0.0019	-1.12	3.08	0.47	0.27	0.0388	0.0205	0.0427
Δ TS	8.3e-6	0.0023	0.28	0.99	0.29	-0.01	0.0000	0.0000	0.0335
Δ DEF	1.3e-5	0.0011	1.50	22.94	0.45	0.10	0.0000	0.0000	0.0286
Δ IP	0.0016	0.0063	-1.67	8.94	0.23	0.28	0.0000	0.0001	0.0817
Δ INF	-1.6e-5	0.0028	-0.33	2.86	-0.17	-0.29	0.0000	0.0000	0.0174

a, b e c indicano un KPSS significativo rispettivamente all'1%, 5% e 10%

2.3.4 Risultati della stima del modello lineare cross-country

Prima di passare alla valutazione delle previsioni, vengono discussi brevemente i risultati della stima dei due modelli presentati sui due diversi dataset.

Per quanto riguarda il modello vettoriale sul dataset cross-country, al modello (2.1.6) viene aggiunta una dummy riferita all'Ottobre 1987. La dummy è volta a cogliere l'effetto del "lunedì nero" del 19 Ottobre, nel quale i mercati mondiali hanno fatto registrare ingenti perdite in poche ore.

I criteri di informazione per la scelta del numero dei ritardi indicano un VAR(1), mentre viene considerato unicamente il primo ritardo delle esogene. Il modello da stimare diventa dunque:

$$P_t = c + A_1 P_{t-1} + D X_{t-1} + \gamma Black_t + \varepsilon_t \quad (2.3.1)$$

dove $Black_t$ è la dummy riferita al lunedì nero. La tabella 2.9 presenta i risultati della stima Cholesky-VAR(1) sul campione cross-country, insieme ai test diagnostici. Dalle stime emerge che gli effetti spillover tra i fattori di Cholesky sono positivi e significativi. Il test Ljung-Box (L-B) evidenzia la presenza di autocorrelazione dei residui in diverse equazioni, il risultato è confermato dal test L-B nel multivariato (si veda Hosking (1981) per ulteriori dettagli). In aggiunta, il test ARCH sulle singole equazioni rifiuta l'ipotesi nulla di assenza di eteroschedasticità condizionale in diverse equazioni. I test diagnostici suggeriscono, dunque, una specificazione diversa da quella lineare, la stima di un non modello lineare punta a superare questo tipo di problemi.

In generale, il modello Cholesky-VAR sull'intero campione cross-country mostra che le determinanti macroeconomiche hanno un impatto statisticamente significativo sulle serie dei fattori di Cholesky della matrice delle covarianze realizzate, il risultato è confermato dal test di Wald sull'azzeramento dei coefficienti delle esogene che rifiuta in maniera decisa l'ipotesi nulla di non significatività degli stessi. Le determinanti di origine finanziaria, come il tasso di crescita del dividend yield e del price earning ratio, non sembrano avere un impatto significativo sulle stesse, ma si ricorda che lo scopo di questo lavoro è quello di verificare la capacità previsiva delle variabili esogene e non la Granger-casualità delle stesse in senso stretto. Il segno e la dimensione dei singoli coefficienti sono di difficile interpretazione, vista la natura delle serie dei fattori derivanti da una parametrizzazione della matrice originale¹⁷. Dal punto di vista della stima, i risultati migliorano se le variabili

¹⁷Per le problematiche legate all'uso della scomposizione di Cholesky si veda l'appendice A.

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

esogene sono considerate senza il ritardo, tuttavia la mancanza di disponibilità di dati per alcune variabili esplicative rende questo modello non applicabile ai fini previsionali.

Tabella 2.9: Stima VAR dataset cross-country

Stime OLS usando le osservazioni 02:1969-10:2015 (T = 562)										
	Cholesky1a	Cholesky2a	Cholesky3a	Cholesky4a	Cholesky5a	Cholesky6a	Cholesky7a	Cholesky8a	Cholesky9a	Cholesky10a
const	0.0223 ^a		0.0057 ^a	0.0048 ^c	0.0288 ^a	0.0046 ^a	0.0038 ^b	0.0186 ^a	0.0079 ^a	0.0149 ^a
Cholesky1a_1	0.4726 ^a		0.1248 ^a	0.1319 ^b			0.1127 ^b	0.0869 ^c		0.0726 ^c
Cholesky2a_1			-0.1478 ^b	0.1475 ^b						
Cholesky3a_1	0.1751 ^b	0.1536 ^b	0.2326 ^a		0.1550 ^c					
Cholesky4a_1			0.0878 ^b	0.4735 ^a					0.2704 ^a	
Cholesky5a_1	-0.0796 ^b	0.0827 ^a			0.4120 ^a			-0.0647 ^b		
Cholesky6a_1		0.1319 ^b	0.1869 ^a	0.1816 ^b					0.2245 ^a	
Cholesky7a_1							0.1765 ^a			
Cholesky8a_1					-0.1218 ^b		-0.0727 ^c	0.5174 ^a	-0.1403 ^a	
Cholesky9a_1			0.1572 ^a	0.2523 ^a	0.2235 ^a	0.1161 ^a	0.0827 ^b	-0.1333 ^a	0.3193 ^a	
Cholesky10a_1			-0.1142 ^b	-0.1199 ^c						0.5243 ^a
ΔInflUS_1	-0.0595	0.1128	-0.1269	0.2326	-0.0869	-0.2540	-0.0889	-0.0535	-0.1480	0.1407
ΔInflUK_1	0.0004	0.0004	0.0014	0.0009	0.0005	0.0003	-0.0004	0.0021 ^c	0.0009	0.0010
ΔInflGER_1	-0.0004	-0.0009	0.0011	-0.0006	0.0028	0.0002	0.0012	-0.0011	-0.0002	-0.0030 ^c
ΔInflJPN_1	0.0003	0.0009	-7.8e-5	-0.0006	-0.0003	-0.0022 ^b	-0.0001	-0.0018 ^c	0.0005	0.0001
IPUS_1	-0.3374 ^a	0.0059	-0.1298	-0.1789	-0.4196 ^a	-0.2704 ^a	-0.0715	-0.2497 ^a	-0.0226	-0.0376
IPJPN_1	-0.0006	0.0004	-3.2e-5	0.0004	0.0006	0.0002	0.0003	-0.0006	-0.0002	0.0001
IPEU_1	-0.0001	-0.0002	0.0009	0.0008	-0.0006	-0.0002	0.0008	-0.0005	0.0006	-0.0008
d_DY_1	0.0239	0.0332 ^c	-0.0224	-0.0164	0.0365	0.0160	0.0265	0.0550 ^a	-0.0139	-0.0098
d_EP_1	-0.0216 ^c	-0.0141	-0.0076	-0.0059	-6.0e-5	-0.0029	-0.0205 ^c	0.0002	-0.0176 ^b	-0.0032
Black	0.2301 ^a	0.0173 ^a	0.0942 ^a	0.1018 ^a	0.1736 ^a	0.1173 ^a	0.0329 ^b	0.0644 ^a	0.0567 ^a	0.0702 ^a
SSR	0.1373	0.1054	0.0930	0.1818	0.2444	0.0842	0.1064	0.1099	0.1151	0.0825
E.S.	0.0159	0.0139	0.0131	0.0183	0.0213	0.0125	0.0140	0.0143	0.0146	0.0124
R ²	0.5840	0.1370	0.4466	0.5192	0.4457	0.2807	0.1487	0.4576	0.4657	0.3550
L-B(5)	0.0546	0.9020	0.1480	0.0003	0.0278	0.0664	0.9860	0.0000	0.0445	0.0000
L-B(10)	0.0036	0.6560	0.0175	0.0000	0.0003	0.2410	0.8760	0.0000	0.0023	0.0000
ARCH(5)	0.0000	0.7130	0.0076	0.0494	0.7543	0.0000	0.6830	0.4823	0.0000	0.6880
ARCH(10)	0.0003	0.9544	0.0084	0.0000	0.5197	0.0001	0.9018	0.5918	0.0000	0.2890
LL	1536.35	1610.64	1645.63	1457.73	1374.66	1673.61	1607.81	1598.85	1586.00	1679.33
AIC	-3030.70	-3179.28	-3249.26	-2873.45	-2707.32	-3305.21	-3173.61	-3155.69	-3130.00	-3316.66
BIC	-2939.78	-3088.35	-3158.34	-2782.53	-2616.40	-3214.29	-3082.69	-3064.77	-3039.08	-3225.74
LL	16565.464									
LB	0.0000									
Wald	0.0003									

a, b e c indicano i livelli di significatività rispettivamente al 1%, 5% e 10%

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

2.3.5 Risultati della stima del modello lineare americano

Il modello sul dataset americano segue l'equazione (2.3.1), fatta eccezione per l'uso della dummy intervallare *Black*, poiché l'evento non è compreso nel periodo considerato. Anche in questo campione, i criteri informativi per la scelta dei ritardi di un VAR indicano un VAR(1). La tabella 2.10 riporta i risultati della stima Cholesky-VAR(1). I problemi inerenti l'interpretazione dei coefficienti delle determinanti dei fattori di Cholesky, evidenziati per il dataset cross-country, rimangono in essere anche per il campione sul mercato americano. Tuttavia, si può notare come sia presente e significativo un effetto spillover tra i diversi fattori di Cholesky, infatti l'aumento di un fattore comporta, nella maggior parte dei casi in cui il relativo coefficiente è significativo, un aumento del fattore dipendente. Il test di Wald conferma la significatività congiunta dei coefficienti delle determinanti macroeconomiche e finanziarie. Le statistiche test Ljung-Box sulle singole equazioni e quella multivariata confermano quanto emerso dal dataset cross-country e suggeriscono una specificazione alternativa a quella presentata; il modello non lineare, presentato nel corso di questo capitolo, ha anche l'obiettivo di ovviare a questo problema. Poiché questo lavoro ha l'obiettivo di valutare il modello in ambito previsionale, non viene approfondito il significato economico dei coefficienti derivanti dalla stima.

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.10: Stima VAR dataset mercato americano

Stime OLS usando le osservazioni 1988:11-2015:06 (T = 320)						
	Cholesky1b	Cholesky2b	Cholesky3b	Cholesky4b	Cholesky5b	Cholesky6b
const	0.0184 ^a			0.0095 ^a	0.0006 ^a	0.0003 ^b
Cholesky1b_1	0.6181 ^a	-0.0427 ^b		0.0752 ^a		
Cholesky2b_1		0.4749 ^a	0.0395 ^a		0.0145 ^c	
Cholesky3b_1		1.1570 ^b	0.2361 ^a			
Cholesky4b_1				0.1629 ^a		
Cholesky5b_1		1.1892 ^a	0.1252 ^b		0.4314 ^a	0.1422 ^a
Cholesky6b_1		0.7699 ^b	-0.1350 ^a		0.0816 ^c	0.6404 ^a
D-P_1	-0.1058 ^a	-0.0281 ^c	-0.0002	0.0070	0.0002	-0.0006
E-P_1	-0.0127	0.0017	0.0005	-5.6e-6	-0.0004	0.0003
MKT_1	-0.0014 ^a	-0.0002	7.1e-6	-4.9e-6	-1.7e-5	-1.6e-7
SMB_1	0.0005 ^c	-0.0001	-2.2e-5	4.8e-5	2.1e-5	-1.4e-5
HML_1	0.0001	-0.0002	-3.5e-5 ^b	-9.9e-5	3.2e-5 ^c	-3.1e-6 ^b
STR_1	-0.0006 ^b	-4.1e-5	1.9e-6	-0.0001	1.1e-5	-3.5e-5
ΔT-B_1	0.5547	-0.0870	0.0069	0.2285	0.0257	-0.0082
ΔT-S_1	0.3925	0.0268	0.0124	0.3502 ^c	-0.0037	0.0066
ΔDEF_1	0.0549	0.3636	-0.0425	-1.0918 ^b	0.1078	-0.0163
ΔIP_1	-0.5612 ^a	0.0553	0.0104	-0.1491 ^b	-0.0004	-0.0159 ^c
ΔInf_1	-0.4206	0.1373	0.0241	-0.1677	-0.0359	-0.0326 ^b
SSR	0.0814	0.0116	0.0002	0.0153	0.0002	0.0002
E.S. regressione	0.0164	0.0062	0.0009	0.0071	0.0008	0.0008
R ²	0.6290	0.5125	0.4466	0.1784	0.2902	0.5285
L-B(5)	0.0183	0.0000	0.0259	0.2550	0.0045	0.0002
L-B(10)	0.0335	0.0000	0.0002	0.4190	0.0000	0.0000
ARCH(5)	0.0106	0.2103	0.0000	0.9970	0.0004	0.0000
ARCH(10)	0.1205	0.2456	0.0005	0.9990	0.0333	0.0000
LL	870.19	1182.27	1819.23	1137.82	1838.63	1836.85
AIC	-1704.39	-2358.53	-3602.46	-2239.64	-3641.25	-3637.69
BIC	-1636.56	-2260.70	-3534.63	-2171.80	-3573.42	-3564.89
LL	8784.9962					
LB	0.0000					
Wald	126.078					
	(0.0000)					

a, b e c indicano i livelli di significatività rispettivamente al 1%, 5% e 10%

2.3.6 Risultati della stima del modello non lineare cross-country

La stima del modello non lineare viene preceduta dai test sulla presenza di break strutturali presentati nella sezione 2.2.1, ai fini della scelta del numero di regimi nel modello non lineare. La tabella 2.11 riporta i test $\sup F_T(l)$, UD_{max} , WD_{max} e $\sup F_T(l+1 | l)$. Il test $\sup F_T(l)$ rifiuta la nulla di assenza di break per ogni ritardo compreso tra 1 e 8, per ogni serie esclusa quella del primo fattore di Cholesky (quando l'ipotesi alternativa implica la presenza di 1 o 2 break). Questo suggerisce la presenza di almeno un break strutturale nelle serie dei fattori.

Il test $\sup F_T(1 | 0)$ è, per la maggior parte delle serie, altamente significativo e la statistica test associata è superiore a quella degli altri test sequenziali. In aggiunta, le statistiche test sequenziali confermano un numero di break non superiore a 3. Poiché la maggior parte delle serie presenta un numero di break pari a uno e per il principio della parsimonia, in questo lavoro si utilizza un modello non lineare con un unico break e due regimi.

Il modello non lineare utilizzato per descrivere la relazione tra le variabili esogene e i fattori di Cholesky è quello descritto nella (2.2.8). Si è scelto un numero di ritardi, p , pari a quello suggerito dai criteri di informazione per il VAR. Il modello, per un numero di regimi pari a 2, diventa quello descritto nella (2.2.9).

Seguendo la procedura descritta nel paragrafo 2.2, una volta scelto il numero di regimi presenti nel modello, è necessario testare la linearità dello stesso prima di procedere alla stima. Il test di linearità si accompagna alla scelta della variabile di transizione che guida le dinamiche della funzione di transizione logistica. Poiché la teoria economica non fornisce alcuna indicazione nella scelta di una variabile di transizione unica per il modello, si procede ad un test di linearità equazione per equazione, considerando come potenziali variabili di transizione tutte le variabili dipendenti e le esogene. Viene scelta come variabile di transizione dell'equazione quella che produce il p -value minore nel test di linearità in ambito univariato. Se due o più equazioni condividono la medesima variabile di transizione, viene utilizzato un test congiunto di non linearità del modello, assumendo che le dinamiche non lineari siano guidate da un'unica variabile di transizione (si veda nel dettaglio l'appendice B).

I p -value della statistica test per ogni singola equazione e per ogni potenziale variabile di transizione, s_t , sono riportati nella tabella 2.12. In diverse equazioni, il p -value minore è ottenuto in corrispondenza del primo ritardo del fattore di Cholesky, Cholesky1a_1. In

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

alcune equazioni, il test rifiuta la linearità in maniera forte anche per il primo ritardo del terzo fattore di Cholesky, Cholesky3a_1. Dalla tabella si può notare, in maniera non sorprendente, che per diverse serie il p -value minore è associato all'uso del ritardo della variabile dipendente come variabile di transizione.

Si conduce, infine, un test congiunto sulla linearità dell'intero modello, presupponendo che ci sia un'unica variabile di transizione, rispettivamente Cholesky1a_1 e Cholesky3a_1, come individuato dai test in ambito univariato. La tabella 2.13 riporta i risultati dei test congiunti. Come in Yang e Teräsvirta (2013), sono riportati anche il Λ di Wilks e il test F di Rao, insieme ai test LM e F. Dai risultati dei test non emerge in maniera univoca una variabile di transizione preferita. Data la tipologia di parametrizzazione della matrice delle covarianze che ha portato alle serie dei fattori di Cholesky (i.e. la prima serie dell'ordinamento assume importanza maggiore rispetto alle altre) e vista l'importanza del mercato azionario americano nel mondo della finanza globale, si è deciso di stimare un modello non lineare che abbia come unica variabile di transizione il primo ritardo del primo fattore di Cholesky, i.e. Cholesky1a_1.

Tabella 2.11: Test di break strutturali multipli nelle serie dei fattori di Cholesky con trimming pari a 0.1 e $M = 8$.

	Cholesky1a	Cholesky2a	Cholesky3a	Cholesky4a	Cholesky5a	Cholesky6a	Cholesky7a	Cholesky8a	Cholesky9a	Cholesky10a
$\sup F_T(1)$	6.7970	40.7047*	49.7174*	66.3543*	62.5332*	14.3831*	46.3050*	11.4870*	100.9959*	56.5339*
$\sup F_T(2)$	3.2822	20.8722*	28.6484*	33.6436*	40.1811*	10.2714*	24.7518*	9.6653*	80.2022*	33.8727*
$\sup F_T(3)$	12.5422*	15.1497*	23.2421*	31.9377*	48.6514*	6.0876	16.4801*	12.0343*	55.2779*	36.6723*
$\sup F_T(4)$	22.6816*	11.4142*	20.4423*	34.1888*	22.3317*	6.4435*	9.8088*	7.2069*	44.1243*	31.2609*
$\sup F_T(5)$	15.9205*	9.5319*	18.8746*	27.2869*	32.7154*	7.6890*	10.6969*	12.6069*	35.3998*	29.4016*
$\sup F_T(6)$	13.6838*	7.5919*	16.4963*	25.9239*	32.7597*	6.6404*	9.0466*	11.7153*	34.3533*	37.4604*
$\sup F_T(7)$	11.9656*	8.5264*	14.3747*	23.8696*	31.0934*	6.4064*	11.8825*	10.8161*	31.3035*	32.4488*
$\sup F_T(8)$	12.0627*	6.1649*	12.4883*	18.4508*	24.6087*	5.3594*	10.1086*	11.0987*	24.6059*	24.1035*
UD_{max}	22.6816	40.7047	49.7174	66.3543	62.5332	14.3830	46.3050	12.6069	100.9959	56.5339
WD_{max}	34.2293	40.7047	49.7147	66.3543	66.8911	14.3831	46.3050	28.2118	100.9959	71.0187
UD_{max} Determined breaks	4	1	1	1	1	1	1	5	1	1
WD_{max} Determined breaks	4	1	1	1	7	1	1	8	1	6
$\sup F_T(1 0)$	6.7970	40.7047*	49.7147*	66.3543*	62.5332*	14.3831*	46.3050*	11.4870*	100.9959*	56.5339*
$\sup F_T(2 1)$	31.7192*	2.8990*	7.7482	6.7158	15.8806*	8.1329	5.5429	12.1289*	42.1075*	18.4710*
$\sup F_T(3 2)$	29.2231*	2.0233	6.6074	10.8004	4.5939	8.0583	2.1734	11.8920*	10.6149	43.4135*
$\sup F_T(4 3)$	6.4395	0.9795	4.6995	9.3167	2.3839	7.2543	4.1242	2.4317	10.6296	9.9936
$\sup F_T(5 4)$	3.6125	6.0493	3.2513	6.1974	4.6041	7.9336	11.9426	28.6944*	1.5193	20.9190*
$\sup F_T(6 5)$	4.6189	0.8836	2.9402	6.2473	2.1360	3.4719	1.2284	5.2786	0.8242	9.1168
$\sup F_T(7 6)$	0.1207	3.1109	1.2736	3.3479	2.1204	1.1275	1.1656	5.2608	0.0833	2.4940
$\sup F_T(8 7)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.5384	0.0000	0.0000
Sequential F-statistics determined break	0	1	1	1	2	1	1	3	2	3
Sequential F-statistics largest breaks	3	1	1	1	2	1	1	5	2	3

* Significativo al 5%

Tabella 2.12: Test non linearità cross-country

s_t	Test LM									
	Cholesky1a	Cholesky2a	Cholesky3a	Cholesky4a	Cholesky5a	Cholesky6a	Cholesky7a	Cholesky8a	Cholesky9a	Cholesky10a
Cholesky1a_1	3.8836**	2.9884*	1.7470	2.5015	5.0742*	3.2797	2.5329*	3.9119	1.5821	4.4206*
	0.0000	0.0000	0.0020	0.0000	0.0000	0.0000	0.0000	0.0000	0.0096	0.0000
Cholesky2a_1	2.2960	1.5685	1.1940	1.6002	2.9941	1.0726	1.3272	2.5705	1.7174	2.9171
	0.0000	0.0109	0.1822	0.0082	0.0000	0.3486	0.0759	0.0000	0.0027	0.0000
Cholesky3a_1	4.5004*	2.7694	2.0992**	2.3485	5.0238**	4.6652*	2.0483	4.1400**	1.5984	3.7294
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0083	0.0000
Cholesky4a_1	3.6855	2.5749	1.9764	3.0204**	4.4647	3.2664	2.2963	3.4470	2.2225**	3.6594
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Cholesky5a_1	3.3765	2.9553**	1.3838	1.8708	4.8094	2.6353	2.1506	3.2187	1.5668	4.1129**
	0.0000	0.0000	0.0500	0.0000	0.0000	0.0000	0.0000	0.0000	0.0110	0.0000
Cholesky6a_1	3.4588	2.8907	2.0079	1.9305	4.4363	2.9919	2.3508	2.8489	2.1922	3.6610
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Cholesky7a_1	3.2009	2.5740	1.9491	2.4062	4.6264	3.0081	2.4483**	3.1844	1.9610	4.0290
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Cholesky8a_1	3.6465	1.7222	1.8108	1.6494	4.1175	2.7836	1.8085	4.1712*	1.6601	3.9890
	0.0000	0.0026	0.0011	0.0052	0.0000	0.0000	0.0011	0.0000	0.0047	0.0000
Cholesky9a_1	3.6798	1.9408	2.0171	2.1648	3.4054	2.8839	1.5594	3.0220	2.4994*	2.9502
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0118	0.0000	0.0000	0.0000
Cholesky10a_1	3.5517	1.0610	1.4906	1.1115	3.0038	1.7923	1.9781	3.9638	1.6595	3.3272
	0.0000	0.3678	0.0213	0.2879	0.0000	0.0013	0.0000	0.0000	0.0047	0.0000
Δ InflUS_1	1.4056	1.2962	1.1641	0.7912	1.6053	2.0558	1.0562	1.5384	0.8957	1.7081
	0.0423	0.0944	0.2167	0.8415	0.0078	0.0000	0.3760	0.0141	0.6728	0.0030
Δ InflUK_1	1.2327	1.2772	1.3992	1.2793	1.4871	1.7355	1.4236	2.6711	0.6233	1.9289
	0.1436	0.1074	0.0444	0.1059	0.0219	0.0023	0.0367	0.0000	0.9779	0.0003
Δ InflGER_1	1.8792	1.8959	1.3487	1.3397	2.1164	2.2210	1.4764	3.8488	1.0916	1.3676
	0.0005	0.0004	0.0650	0.0694	0.0000	0.0000	0.0239	0.0000	0.3182	0.0565
Δ InflJPN_1	0.8158	1.0440	1.8655	1.6075	1.3112	0.9838	0.8500	3.6907	0.9239	1.3187
	0.8067	0.3970	0.0000	0.0077	0.0850	0.5069	0.7529	0.0000	0.6205	0.0806
IPUS_1	2.7045	2.3258	2.0867	1.5793	3.2899	4.0113**	2.2154	4.0051	0.9752	2.2861
	0.0000	0.0000	0.0000	0.0084	0.0000	0.0000	0.0000	0.0000	0.5254	0.0000
d_DY_1	2.6257	1.0986	1.7333	1.6596	2.3913	1.7768	1.7878	2.8108	1.8611	2.3686
	0.0000	0.3039	0.0019	0.0039	0.0000	0.0012	0.0011	0.0000	0.0005	0.0000
d_EP_1	3.2303	1.0717	2.2071*	2.5103**	2.3000	1.7317	2.0086	2.5068	1.5989	2.3580
	0.0000	0.3476	0.0000	0.0000	0.0000	0.0019	0.0000	0.0000	0.0070	0.0000

* e ** indicano i due p-value minori

Tabella 2.13: Test non linearità cross-country congiunto

Test	Cholesky1a_1	Cholesky3a_1
LM	539.6178 ^a	585.4010 ^a
F	2.9818 ^a	3.2348 ^a
Wilks	175.9930 ^a	156.3957 ^a
Rao	12.0884 ^a	10.5372 ^a

La tabella 2.14 riporta i risultati della stima di un modello con due regimi la cui variabile di transizione, s_t , è unica e pari a *Cholesky1a_1*. Le esogene sono le stesse utilizzate per il modello (2.3.1), fatta eccezione della dummy riferita all'Ottobre 1987 e alle due variabili non significative nella stima lineare, *IPJPN_1* e *IPEU_1*. La dimensione e il segno dei coefficienti, così come il livello di significatività, variano sensibilmente da equazione ad equazione e tra i vari regressori. Confrontando le singole equazioni del modello non lineare con quelle del modello lineare, in termini di criteri informativi e log-verosimiglianza, non emergono differenze sostanziali tra i due modelli. Analizzando le stime in tabella 2.14, emerge che solo alcune variabili esogene sembrano impattare sulle serie dei fattori di Cholesky quando $G(s_t; \gamma, c) = 1$. In particolare, non risultano mai significativi i coefficienti associati ai tassi di inflazione. In aggiunta, i coefficienti associati ai ritardi delle variabili dipendenti, in $G(s_t; \gamma, c) = 1$, sono spesso significativi e di segno positivo.

Dalle stime emergono dinamiche non lineari, ma risulta difficile separarle da una specificazione lineare ad un singolo regime. In aggiunta, nel multivariato il modello presenta alcune criticità legate al numero di regimi da utilizzare per l'intero modello e alla scelta delle variabili di transizione che guidano le dinamiche della funzione di transizione. Infine, la tipologia del modello non lineare, come suggerito in ambito previsionale da Nason e Diebold (1990), potrebbe essere di natura diversa (e.g. diversa funzione di transizione, diversa specificazione della non linearità). Per questi motivi, si rimanda all'analisi delle previsioni per verificare praticamente l'utilità del modello non lineare sui fattori di Cholesky della matrice delle covarianze realizzate.

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.14: Stima QML modello non lineare cross-country

ϕ	VLSTAR									
	Chol1a	Chol2a	Chol3a	Chol4a	Chol5a	Chol6a	Chol7a	Chol8a	Chol9a	Chol10a
constant	0.0123	0.0622 ^a	0.1246 ^a	0.1745	0.0102	-0.0462	-0.0288	0.0093 ^c	0.1032	0.0093
Chol1a_1	-11.2416	-3.5097 ^a	-36.6538 ^a	5.0666	55.5962 ^b	74.4554 ^a	-10.8239 ^a	11.3703	3.1297	33.5109
Chol2a_1	6.7558	30.5306 ^a	9.4181 ^a	28.0490 ^c	64.7448 ^b	114.8945 ^a	19.5622 ^a	12.1793	9.8239	20.4444
Chol3a_1	2.8946	-19.6988 ^a	97.1312 ^a	-10.8721	-50.9699 ^c	2.6975	42.4437 ^a	-28.2247	-7.1255	-46.9898
Chol4a_1	-7.7056	-90.8633 ^a	-36.1191 ^a	6.7688	0.5195	-14.8578	-41.5563 ^a	8.9040	4.0460	28.4650
Chol5a_1	-19.5763	-10.1647 ^a	-19.3472 ^a	-35.6322 ^a	-77.8218 ^b	-86.0867 ^a	-10.0242 ^a	-5.6953	-10.0089	-25.1445
Chol6a_1	65.1792	-24.6669 ^a	17.3610 ^a	54.3903 ^a	63.4385 ^c	17.3850	19.0859 ^a	1.9119	16.0133	25.5936
Chol7a_1	-15.3072	137.3015 ^a	21.9867 ^a	-22.0551	0.3923	-39.9635	-24.3255 ^a	6.8898	-5.8667	-9.2872
Chol8a_1	4.4783	-174.7444 ^a	65.2408 ^a	-37.6736 ^a	-38.2856	42.5595 ^a	56.5854 ^a	-14.0131	4.5298	-22.2674
Chol9a_1	12.8647	109.8826 ^a	7.2756 ^a	33.9457	0.4155	-26.9283	-0.2820	-5.5418	-4.9149	2.9948
Chol10a_1	-26.7552	194.5704 ^a	-22.1939 ^a	36.7866 ^a	3.7109	-39.4128	-18.0236 ^a	-7.1201	0.7790	-26.1414
$\Delta \ln \text{US}_1$	-0.7535	78.4895 ^a	37.2327 ^a	128.8588	-32.7054	50.5249	243.0778 ^a	-50.8923	90.1022	-88.9067
$\Delta \ln \text{UK}_1$	-0.5221	8.4176 ^a	1.8254 ^a	-0.9172	0.0053	-1.6631	-0.1203	0.6033	0.6239	-1.8122
$\Delta \ln \text{GER}_1$	0.4636	-2.6771 ^a	2.8422 ^a	-2.8682	0.6248	0.6638	-2.2449 ^b	0.4392	-1.3383	0.6617
$\Delta \ln \text{JPN}_1$	-0.1723	4.1311 ^a	0.3443 ^a	0.9552	-0.4280	-2.3877	-0.0723	-0.1828	0.1691	0.8098
IPUS_1	-67.5849	-1.4608	129.3954 ^a	-20.8608	-153.5030 ^b	-110.0541	82.9551 ^a	-71.5798	21.9978	-106.9432
d_DY_1	14.6386	39.0227 ^a	-12.7707 ^a	-3.0647	14.9995	-17.9205	0.3139	4.6489	-1.3174	-2.1915
d_EP_1	2.8807	-23.3162 ^a	-10.2307 ^a	3.0780	0.1449	11.8225	0.6221	2.0501	-1.3291	8.5139
ψ										
constant	0.0123 ^a	0.0009	0.0082 ^a	0.0037	0.0102	0.0015	0.0052 ^b	0.0093 ^a	0.0079 ^a	0.0093
Chol1a_1	0.8300 ^a	-0.0329	0.0708	0.1636	0.0679	0.1075	0.0487	-7.8982	0.0264	0.0745
Chol2a_1	-0.0759	0.0174	-0.0453	-0.1766 ^c	-0.2120	-0.0491	0.0028	-9.1601	-0.0387	-0.1097
Chol3a_1	0.0265	0.0855 ^c	0.1442 ^b	-0.0845	0.2905	-0.1520 ^c	-0.0630	20.8928	0.0376	0.0044
Chol4a_1	0.1220	0.0319	0.1262 ^a	0.4974 ^a	0.0000	0.0811 ^b	0.0652	-6.5048	0.2390 ^a	-0.0284
Chol5a_1	0.0086	0.1027 ^a	0.0058	0.1196 ^c	0.9030 ^a	0.0580	-0.0148	4.1355	0.0429	0.0973
Chol6a_1	-0.2015	0.1056 ^b	0.1154 ^c	0.0819	-0.2434	0.0565	-0.0917	-1.3609	0.1174	-0.1130
Chol7a_1	0.0902	-0.0441	-0.0386	-0.0503	-0.0571	0.0703	0.1891 ^a	-5.1607	0.0132	0.0851
Chol8a_1	-0.0575	-0.0212	-0.0034	-0.0157	0.0963	-0.0614	-0.0707 ^c	11.2906	-0.1825 ^a	0.0817
Chol9a_1	-0.0990	0.0027	0.1658 ^a	0.1715 ^b	0.1969	0.1291 ^b	0.1110	3.9227	0.3445 ^a	-0.0753
Chol10a_1	0.0768	0.0087	-0.1127 ^b	-0.1914 ^b	-0.0316	-0.0068	0.0679	5.2788	0.0295	0.6203 ^a
$\Delta \ln \text{US}_1$	-0.0153	-0.0032	-0.2074	-0.0279	-0.1479	-0.3375	-0.1867	37.6126	-0.4440	0.3187
$\Delta \ln \text{UK}_1$	0.0016	-0.0001	0.0013	0.0013	-0.0005	0.0007	-0.0004	-0.4441	-0.0011	0.0064
$\Delta \ln \text{GER}_1$	-0.0022	-0.0002	0.0002	0.0017	0.0010	0.0018	0.0019	-0.3268	0.0026	-0.0051
$\Delta \ln \text{JPN}_1$	0.0010	0.0004	0.0001	-0.0016	0.0028	-0.0004	0.0001	0.1329	0.0003	-0.0020
IPUS_1	0.0497	0.1623 ^b	-0.0864	-0.0373	0.7512	-0.0306	0.0280	52.8532	-0.0526	0.3969
d_DY_1	-0.0208	0.0410 ^b	-0.0270	-0.0104	-0.0543	0.0378	0.0268	-3.3592	-0.0044	0.0274
d_EP_1	-0.0349	-0.0231	0.0004	-0.0184	0.0033	-0.0177	-0.0200 ^c	-1.5428	-0.0169	-0.0559
γ	21.0278 ^c	109.5624 ^a	234.9665 ^a	43.6172 ^c	16.8712 ^b	34.0930 ^b	99.6815 ^a	0.4857 ^c	41.9345	19.8353 ^c
c	0.3187 ^a	0.1578 ^a	0.1180 ^a	0.2062 ^a	0.3503 ^a	0.2660 ^a	0.1447 ^a	0.6540 ^c	0.1803 ^a	0.3295 ^c
σ^2	0.0172	0.0125	0.0130	0.0177	0.0202	0.0120	0.0128	0.0138	0.0142	0.0117
LL	1484.74	1662.04	1641.69	1468.66	1394.26	1686.48	1647.72	1606.33	1592.23	1700.72
AIC	-2889.48	-3244.08	-3203.38	-2857.32	-2708.52	-3292.95	-3215.44	-3132.65	-3104.46	-3321.45
BIC	-2716.29	-3070.89	-3030.19	-2684.13	-2535.33	-3119.76	-3042.25	-2959.46	-2931.27	-3148.26
R ²	0.4998	0.2817	0.4388	0.5376	0.4831	0.3129	0.2616	0.4719	0.4774	0.4024
SSR	0.1651	0.0877	0.0943	0.1748	0.2279	0.0804	0.0923	0.1070	0.1125	0.0764

a, b e c indicano i livelli di significatività rispettivamente al 1%, 5% e 10%

2.3.7 Risultati della stima del modello non lineare americano

Anche nel dataset americano la stima del modello non lineare viene preceduta dal test sulla presenza di break strutturali e da quello di linearità delle serie. I risultati dei test sui break strutturali sono riportati nella tabella 2.15. I test $\sup F_T$ sono significativi per ogni l compresa tra 1 e 8 per la maggior parte delle serie, si presume che sia presente almeno un break. Viene utilizzato il test $\sup F_T(l+1 | l)$ per la scelta del numero di break del modello. I test $\sup F_T(l+1 | l)$ suggeriscono un diverso numero di break tra le varie serie, per questo si è scelto il numero di break minimo che emerge dal test sulla serie di *Cholesky5b*, per cui si considerano un solo break e due regimi.

Seguendo la procedura illustrata nella sezione 2.2, si conduce un test di linearità sulle singole equazioni per determinare la variabile di transizione che guida ogni singola funzione di transizione. Come si può notare dalla tabella 2.16, a differenza del dataset cross-country, le equazioni non sembrano condividere la stessa variabile di transizione. Risultano anche frequenti i casi in cui l'ipotesi nulla di linearità del modello non può essere rifiutata. Sulla base dei risultati dei test di linearità, vengono scelte variabili di transizioni diverse per ogni singola equazione, considerando quelle variabili che permettono di ottenere il p -value minore nel test di linearità. Per la serie *Cholesky1b*, il p -value minore è ottenuto utilizzando come variabile di transizione *DP_1*, mentre per il secondo fattore di Cholesky il rifiuto dell'ipotesi nulla è massimo se la variabile di transizione è *STR_1*. Per i fattori di *Cholesky3b, 4b, 5b* e *6b* il p -value è minimo quando la variabile di transizione è rispettivamente *Cholesky5b_1*, *Cholesky2b_1*, *ΔTB_1* e *Cholesky3b_1*.

La tabella 2.17 riporta i risultati della stima non lineare con due regimi, dove le variabili di transizione, $s_{it}, \forall i = 1, \dots, m$, variano da equazione ad equazione e sono rispettivamente *DP_1*, *STR_1*, *Cholesky5b_1*, *Cholesky2b_1*, *TB_1* e *Cholesky3b_1*. I regressori sono gli stessi utilizzati per la stima del modello lineare nella (2.3.1) fatta eccezione per la variabile dummy *Black* e le determinanti non significative nel modello lineare, come *EP_1*, *MKT_1*, *SMB_1*, *STR_1*, *TB_1* e *TS_1*. I coefficienti risultano spesso significativi all'1%, ma segno e dimensione dei coefficienti sono di difficile interpretazione data l'alta diversità tra le varie equazioni e le varie determinanti. In particolare, il modo in cui le singole variabili impattano sulle diverse serie dei fattori, cambia da serie a serie per segno e dimensione. La variabile *DP_1* sembra essere quella che maggiormente aiuta a prevedere le serie dei fattori di Cholesky, mentre il ritardo del primo fattore di Cholesky, *Cholesky1b_1*, risulta il fattore autoregressivo maggiormente significativo. Si rimanda alla sezione sulla

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

valutazione delle previsioni per una profonda analisi dell'utilità della stima non lineare a fini previsionali.

Tabella 2.15: Test di break strutturali multipli nelle serie dei fattori di Cholesky con trimming pari a 0.1 e $M = 8$.

	Cholesky1b	Cholesky2b	Cholesky3b	Cholesky4b	Cholesky5b	Cholesky6b
$\sup F_T(1)$	12.3099*	165.8932*	41.4791*	11.8554*	230.4767*	25.8489*
$\sup F_T(2)$	1.3148	115.7399*	39.7831*	7.5482	166.7899*	69.7394*
$\sup F_T(3)$	5.9676	112.1361*	27.6863*	14.3018*	119.5718*	42.8663*
$\sup F_T(4)$	12.2636*	63.3609*	22.0339*	10.3794*	98.1970*	37.2904*
$\sup F_T(5)$	16.0126*	71.2631*	17.8066*	13.4534*	83.5568*	40.0983*
$\sup F_T(6)$	14.2618*	68.0604*	19.5641*	12.0599*	71.1777*	34.5118*
$\sup F_T(7)$	12.9614*	67.6917*	17.5385*	10.5579*	61.0617*	21.5994*
$\sup F_T(8)$	10.6106*	57.4277*	17.5602*	8.0746*	53.9123*	21.1503*
UD_{max}	16.0126	165.8932	41.4791	14.3018	230.4767	69.7394
WD_{max}	27.8838	165.8932	45.7103	22.8637	230.4767	80.1299
UD_{max} Determined breaks	5	1	1	3	1	2
WD_{max} Determined breaks	7	1	2	6	7	2
$\sup F_T(1 0)$	12.3099*	165.8932*	41.4791*	11.8554*	230.4767*	25.8489*
$\sup F_T(2 1)$	28.8479*	18.6003*	14.4284*	6.0335	8.6147	103.0461*
$\sup F_T(3 2)$	13.8178*	36.1066*	13.3281*	27.2606*	8.9599	8.5844
$\sup F_T(4 3)$	37.9037*	9.1114	8.9078	8.5855	9.0190	5.8679
$\sup F_T(5 4)$	21.1499*	35.1576*	7.2981	6.7242	8.9187	7.9815
$\sup F_T(6 5)$	2.5347	8.3607	14.4743*	2.0785	0.9885	3.2214
$\sup F_T(7 6)$	2.0579	12.0887	3.6535	2.0514	0.4887	0.0000
$\sup F_T(8 7)$	0.0000	0.0000	0.0049	0.0000	0.0000	0.0000
Sequential F-statistics determined break	5	3	3	1	1	2
Sequential F-statistics largest breaks	5	5	1	3	1	2

* Significativo al 5%

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.16: Test non linearità mercato americano

s_t	Test LM					
	Cholesky1b	Cholesky2b	Cholesky3b	Cholesky4b	Cholesky5b	Cholesky6b
Cholesky1b_1	3.9371	1.3764	1.5850	2.4144	1.4400	1.4817
	0.0000	0.0623	0.0131	0.0000	0.0397	0.0292
Cholesky2b_1	3.0030	1.5259	2.8876	4.3515*	1.7228	1.4706
	0.0000	0.0208	0.0000	0.0000	0.0041	0.0317
Cholesky3b_1	1.8367	1.2401	3.0109	2.4438	2.2603**	2.3000*
	0.0015	0.1493	0.0000	0.0000	0.0000	0.0000
Cholesky4b_1	3.0137	1.6271	1.1297	2.5178**	1.4498	1.0757
	0.0000	0.0093	0.2726	0.0000	0.0371	0.3518
Cholesky5b_1	3.4040	1.5799	3.4894*	1.8628	1.7990	2.0485**
	0.0000	0.0136	0.0000	0.0012	0.0021	0.0002
Cholesky6b_1	3.1874	1.4945	2.1982	2.1789	1.6935	1.6574
	0.0000	0.0265	0.0000	0.0000	0.0053	0.0072
D-P_1	4.5492*	1.7759	1.5759	1.6934	1.7237	1.2947
	0.0000	0.0026	0.0141	0.0053	0.0041	0.1069
E-P_1	3.3005	1.4585	1.1338	2.2430	0.9016	1.1510
	0.0000	0.0347	0.2669	0.0000	0.6585	0.2446
MKT_1	4.5214**	1.9970**	1.8505	2.3333	1.2860	0.8864
	0.0000	0.0003	0.0013	0.0000	0.1129	0.6854
SMB_1	2.1585	1.2556	1.4317	2.1450	0.8462	1.1994
	0.0000	0.1362	0.0422	0.0000	0.7530	0.1887
HML_1	2.6552	1.8492	1.3735	2.3380	0.9612	1.1728
	0.0000	0.0013	0.0636	0.0000	0.5501	0.2181
STR_1	2.6357	1.9999*	1.2674	1.4173	1.9371	1.6144
	0.0000	0.0003	0.1267	0.0468	0.0006	0.0103
$\Delta T-B_1$	2.0649	1.9296	3.3759**	1.2391	2.2793*	1.9937
	0.0002	0.0006	0.0000	0.1502	0.0000	0.0004
$\Delta T-S_1$	3.1478	1.9903	1.8576	1.0985	1.3948	1.6096
	0.0000	0.0004	0.0013	0.3168	0.0549	0.0107
ΔDEF_1	3.6465	1.3536	1.5729	1.0583	1.0730	1.0082
	0.0000	0.0728	0.0143	0.3797	0.3559	0.4652
ΔIP_1	3.2844	1.5717	1.8272	2.2252	1.3051	1.4663
	0.0000	0.0145	0.0016	0.0000	0.1000	0.0328
ΔINF_1	2.0533	1.0240	1.3444	1.4132	1.0866	1.5353
	0.0002	0.4376	0.0774	0.0482	0.3348	0.0194

* e ** indicano i due p -value minori

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.17: Stima QML modello non lineare mercato americano

	VLSTAR					
ϕ	Chol1b	Chol2b	Chol3b	Chol4b	Chol5b	Chol6b
const	0.0212 ^a	0.0267	-0.0042	0.0058	-0.0017	0.0004
Chol1b_1	0.2498 ^b	-0.5432	23.8993 ^a	-4.5718	-12.3584 ^a	2.6684 ^a
Chol2b_1	-0.1316	-4.5001	148.6505 ^a	11.1597	12.4103	-2.2815
Chol3b_1	-4.2170	-2.1707	-89.6949 ^a	0.1920	-64.9950	52.5956
Chol4b_1	0.4461	-5.9751	-49.1120 ^a	2.9087	24.8712 ^a	-8.9418 ^a
Chol5b_1	-5.1440	13.3846	44.4366 ^b	-0.2916	81.2314	42.7141
Chol6b_1	-1.3189	10.0700	384.2828 ^a	-0.9440	8.9257 ^a	247.3420 ^a
D-P_1	0.5899 ^a	-2.7639	8.9104 ^a	-12.5243 ^a	-4.2909 ^a	-4.1393 ^a
HML_1	-0.0005	0.0031 ^c	0.0785 ^a	-0.2180 ^a	-0.0026	0.0411 ^a
Δ DEF_1	5.1310 ^a	-11.6683	-322.1030 ^a	-0.8479	98.4343 ^a	30.4675
Δ IP_1	-1.3794 ^a	2.2291	105.2431 ^a	-2.1425	5.6609	3.3597
Δ INF_1	0.1552	3.6958	-44.9149 ^a	2.2377	-59.7560 ^a	-49.1858 ^a
ψ						
const	0.0122 ^a	-0.0014	0.0023	0.0161	0.0014	0.0001
Chol1b_1	0.7102 ^a	-0.0383 ^b	-11.9315 ^a	6.9707	4.3624 ^a	-2.3482 ^a
Chol2b_1	0.1667	0.4704 ^a	-74.1557 ^a	-16.3963	-4.3612	2.0732
Chol3b_1	-0.8832	1.1472 ^b	44.9020 ^a	-0.9914	22.8405	-47.5255 ^c
Chol4b_1	0.0697	0.0093	24.5195 ^a	-3.8162	-8.7882 ^a	7.8985 ^a
Chol5b_1	-0.4630	1.1796 ^a	-22.4371 ^b	0.9220	-28.0930	-37.3746
Chol6b_1	-0.3344	0.8043 ^a	-191.8912 ^a	1.4925	-3.0473 ^a	-216.8520 ^a
D-P_1	0.1690 ^a	-0.0084	-4.4540 ^a	18.7628 ^a	1.5180 ^a	3.6508 ^a
HML_1	0.0005	-0.0002	-0.0392 ^a	0.3255 ^a	0.0010	-0.0362 ^a
Δ DEF_1	-1.6867 ^c	0.3046	160.7535 ^a	1.1141	-34.9134 ^a	-26.8953
Δ IP_1	0.1804	0.1152	-52.5223 ^a	2.8630	-1.9983	-2.9641
Δ INF_1	-0.0715	-0.0509	22.3844 ^a	-2.8298	21.1573 ^a	43.3414 ^a
γ	122.9803 ^a	1.1339 ^a	0.7600 ^a	0.1608	1.0351 ^a	1.6270 ^a
c	0.0130 ^b	10.9669 ^a	0.9150 ^a	-2.4996	1.0051 ^a	0.0780 ^a
σ^2	0.0113	0.0059	0.0008	0.0070	0.0007	0.0007
LL	979.12	1188.41	1841.98	1135.91	1852.05	1854.22
AIC	-1900.51	-1874.81	-1812.80	-1830.50	-1812.80	-1806.97
BIC	-1749.78	-1724.08	-1662.07	-1679.77	-1662.07	-1656.23
R ²	0.8122	0.5309	0.4212	0.1685	0.3473	0.5379
SSR	0.0412	0.0114	0.0002	0.0155	0.0002	0.0002

a, b e c indicano i livelli di significatività rispettivamente al 1%, 5% e 10%

2.4 Previsioni

Per una valutazione delle previsioni out-of-sample, si ricorre al metodo delle previsioni *rolling window*. Il metodo rolling window permette il confronto tra modelli *nested* e risulta essere un buon metodo per cogliere l'eterogeneità dei dati e la variabilità dei parametri. Nel metodo rolling window la lunghezza del campione da stimare, d , è fissa, la data di inizio e fine del sottocampione successivo al primo considerato sono aumentate di un'unità per un totale di $\tilde{N} = T - d + 1$ sottocampioni. Ad esempio, se si considera una finestra di tempo che va dal 01.1969 al 10.2006, si usano le osservazioni contenute in questa finestra per ottenere la previsione ad un passo, ovvero 11.2006. In maniera analoga per prevedere la matrice varianze e covarianze del 12.2006 si utilizza la finestra di tempo che va dal 02.1969 all'11.2006 e così via. I parametri dei modelli vengono stimati ad ogni passo.

Sulla base di altri lavori simili presenti in letteratura (si veda ad esempio Halbleib-Chiriac e Voev (2011)), in questo lavoro si utilizza una finestra lunga circa due terzi del campione, la lunghezza del vettore delle previsioni prodotte è di circa un terzo del campione totale. Di conseguenza il periodo out-of-sample per il dataset cross-country va dal Dicembre 1999 all'Ottobre 2015, per un totale di 191 osservazioni, mentre quello del dataset americano va dal Settembre 2006 al Giugno 2015, per un totale di 106 osservazioni.

Per ogni sottocampione di ogni dataset vengono stimati sia il modello lineare che quello non lineare e si producono le previsioni ad un passo in maniera iterativa. Una volta prodotta la stima ad un passo, viene ricomposta la previsione della varianza e della covarianze e si calcola l'errore di previsione della stima, $e_{\tilde{n}}$, come

$$e_{\tilde{n}} = y_{d+\tilde{n}} - \hat{y}_{\tilde{n}}, \quad \tilde{n} = 1, \dots, N \quad (2.4.1)$$

dove $y_{d+\tilde{n}}$ è il valore effettivo della variabile oggetto della previsione e $\hat{y}_{\tilde{n}}$ è la previsione a un passo del sottocampione \tilde{n} -esimo. Si ha pertanto una matrice $\tilde{n} \times \tilde{N}$ di errori di previsione ad un passo, e_t .

Come evidenziato da Bauer e Vorkink (2011) e Halbleib-Chiriac e Voev (2011), fare previsioni su una trasformazione non-lineare di $RCOV_t$, come i fattori di Cholesky, e invertire la trasformazione per ottenere una previsione di $RCOV_t$, può portare ad errori. Entrambi i lavori presentano un approccio teorico per la correzione della distorsione soggetto ad assunzioni stringenti. Infatti, la correzione presentata da Bauer e Vorkink (2011) rimane valida in maniera accurata esclusivamente sotto normalità e in un contesto di regressione lineare, mentre quella presentata da Halbleib-Chiriac e Voev (2011) dipende dal momento secondo dell'errore di previsione di P_t e di conseguenza dal modello generatore dei dati.

Queste restrizioni, unite ad un'analisi accurata dei dati, hanno portato Halbleib-Chiriac e Voev (2011) alla conclusione che una correzione delle previsioni non sia necessaria per questa tipologia di dati. Sulla base di quanto presentato dai precedenti lavori, in questa tesi non viene presentata una correzione delle previsioni, problematica che viene lasciata per ulteriori ricerche future.

2.4.1 Valutazione delle previsioni

In questa sezione vengono analizzati diversi metodi per la valutazione delle previsioni della volatilità. La valutazione della capacità previsiva fuori dal campione di riferimento rappresenta un aspetto cruciale per la selezione di un modello e si basa su metodi diretti o indiretti. I metodi diretti prevedono una valutazione statistica delle previsioni per decidere un ordinamento tra modelli concorrenti, mentre i metodi indiretti prevedono la valutazione delle previsioni in un'ottica di ottimizzazione di portafoglio. Questa sezione si occupa principalmente dei metodi diretti, nel terzo capitolo si analizzano i metodi indiretti.

Un problema comune che caratterizza il confronto tra le previsioni di volatilità è la natura latente della variabile di interesse, questo rende la valutazione e il confronto delle previsioni piuttosto complicati (si veda anche Patton (2011)) e risulta necessario l'uso di una proxy. Tipicamente, questo problema viene risolto mediante l'utilizzo di uno stimatore corretto di volatilità, come il quadrato dei rendimenti. È noto, però, che il quadrato dei rendimenti sia una proxy *noisy* della volatilità latente. Per questo, l'attenzione si è spostata sull'uso di uno stimatore corretto della volatilità come la varianza realizzata. La varianza realizzata è uno stimatore della varianza condizionale più efficiente rispetto al quadrato dei rendimenti, infatti si ha che $E[(r_t^2 - \sigma_t^2)^2 | I_{t-1}] = 2\sigma_t^4$, mentre per la varianza realizzata $E[(RV_t^m - \sigma_t^2)^2 | I_{t-1}] = 2\sigma_t^4/m$, dove σ_t^2 è la vera varianza non osservabile e m il numero di periodi intragiornalieri. L'uso della misura realizzata come proxy della varianza condizionale non osservata permette di utilizzare i metodi di valutazione e di confronto delle previsioni presenti in letteratura, come la regressione di Mincer e Zarnowitz (1969) e il test di Diebold e Mariano (1995). Tuttavia, l'uso di una proxy, seppure corretta, non sempre conduce allo stesso risultato ottenuto in presenza di una variabile osservata. Infatti, Andersen e Bollerslev (1998) e Andersen *et al.* (2005) dimostrano che i test perdono di potenza in presenza di una proxy. In aggiunta, Hansen e Lunde (2006) hanno evidenziato la presenza di una distorsione nell'ordinamento di due o più modelli, in termini previsionali, quando si utilizza una proxy della volatilità. In seguito, Patton (2011) ha contribuito a questa letteratura dimostrando che l'uso della varianza realizzata come proxy di volatilità

produce test e ordinamenti meno distorti rispetto ad altre proxy, come ad esempio il quadrato dei rendimenti.

Quando si dispone di previsioni derivanti da due o più modelli concorrenti, la valutazione delle previsioni della volatilità si basa sull'ordinamento delle previsioni determinato dalle funzioni di perdita. Diversi lavori si sono occupati di definire le condizioni necessarie per ottenere un ordinamento consistente. Patton (2011) ha determinato le condizioni necessarie e sufficienti sulla forma funzionale della funzione di perdita in ambito univariato. Laurent *et al.* (2013) estendono nel multivariato l'analisi delle condizioni che una funzione di perdita deve soddisfare per garantire il medesimo ordinamento ottenuto in presenza della vera matrice delle covarianze. La scelta della funzione di perdita influenza i risultati dei test di superiorità delle previsioni, come il test Diebold e Mariano (1995) e il Giacomini e White (2006), per questo si analizzano nel dettaglio le proprietà delle principali funzioni di perdita.

Sia $E[r_t | I_{t-1}] = 0$ e sia $E[r_t^2 | I_{t-1}] = \sigma_t^2$ la vera varianza condizionale non osservabile dei rendimenti r_t e sia h_t una previsione derivante da un modello previsionale, la funzione di perdita è data da $L(\sigma_t^2, h_t)$. In ambito multivariato la variabile di interesse diventa la matrice delle covarianze condizionali, $\Sigma_t = E[r_t r_t' | I_{t-1}]$, e la funzione di perdita è data da $L(\Sigma_t, H_t)$.

L'approccio comune prevede il confronto delle volatilità sulla base della perdita attesa utilizzando la vera varianza latente. Poiché la volatilità non è osservabile, si usa una proxy della stessa, rispettivamente $\hat{\sigma}_t^2$ in ambito univariato e $\hat{\Sigma}_t$ in ambito multivariato, di conseguenza le funzioni di perdita diventano $L(\hat{\sigma}_t^2, h_t)$ e $L(\hat{\Sigma}_t, H_t)$. Patton (2011) si occupa di definire le proprietà che una funzione di perdita deve avere in presenza di una proxy di volatilità. In particolare, Patton (2011) definisce una funzione di perdita "robusta" se, utilizzando una proxy della volatilità, si ottiene lo stesso ordinamento tra due modelli concorrenti ottenuto in presenza della vera varianza condizionale. Per cui, in termini di funzione di perdita attesa, $E[L(\sigma_t^2, h_t)]$, dati due modelli concorrenti, k e j , si ha

$$E[L(\sigma_t^2, h_{k,t})] \leq E[L(\sigma_t^2, h_{j,t})] \Leftrightarrow E[L(\hat{\sigma}_t^2, h_{k,t})] \leq E[L(\hat{\sigma}_t^2, h_{j,t})]. \quad (2.4.2)$$

La condizione (2.4.2) è garantita se

$$\frac{\partial^2 L(\sigma_t^2, h_t)}{(\partial \sigma_t^2)^2} \quad (2.4.3)$$

esiste e non dipende da h_t (cfr. Laurent e Violante (2012)).

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Patton (2011) definisce, poi, una classe di funzioni di perdita, in ambito univariato, omogenee¹⁸ e robuste all'uso di una proxy. Se si considera un grado di omogeneità pari a $b + 2$, la classe delle funzioni di perdita è data da

$$L(\hat{\sigma}_t^2, h_t; b) = \begin{cases} \frac{1}{(b+1)(b+2)}(\hat{\sigma}_t^{2b+4} - h_t^{b+2}) - \frac{1}{b+1}h_t^{b+1}(\hat{\sigma}_t^2 - h_t), & b \neq -1, -2 \\ h_t - \hat{\sigma}_t^2 + \hat{\sigma}_t^2 \log \frac{\hat{\sigma}_t^2}{h_t}, & b = -1 \\ \frac{\hat{\sigma}_t^2}{h_t} - \log \frac{\hat{\sigma}_t^2}{h_t} - 1, & b = -2 \end{cases} \quad (2.4.4)$$

Per $b = 0$ si ottiene la funzione di perdita MSE (*mean squared error*), $L(\hat{\sigma}_t^2, h_t) = (\hat{\sigma}_t^2 - h_t)^2$, mentre per $b = 2$ si ottiene una funzione di perdita del tipo QLIKE (*quasi-likelihood*), $L(\hat{\sigma}_t^2, h_t) = \log h + \frac{\hat{\sigma}_t^2}{h_t}$. Patton e Sheppard (2009) hanno dimostrato che il test Diebold-Mariano e West (DMW) è più potente quando si utilizza una funzione di perdita del tipo QLIKE rispetto alla funzione MSE, in ambito univariato.

Nel caso multivariato, la condizione necessaria per un ordinamento robusto delle previsioni in presenza di una proxy è data da

$$E[L(\Sigma_t, H_{k,t})] \leq E[L(\Sigma_t, H_{j,t})] \Leftrightarrow E[L(\hat{\Sigma}_t, H_{k,t})] \leq E[L(\hat{\Sigma}_t, H_{j,t})], \quad (2.4.5)$$

e la condizione sufficiente per assicurare la (2.4.5) diventa

$$\frac{\partial^2 L(\Sigma_t, H_t)}{\partial \sigma_{k,t} \partial \sigma_{j,t}}, \quad \forall k, j = 1, \dots, N(N+1)/2 \quad (2.4.6)$$

dove $\sigma_{k,t}$ è il k -esimo elemento di $\sigma_t = \text{vech}(\Sigma_t)$. Una funzione di perdita L è robusta se e solo se assume la forma seguente:

$$L(\hat{\Sigma}_t, H_t) = \tilde{C}(H_t) + \tilde{C}(\Sigma_t) + C'(H_t) \text{vech}(\hat{\Sigma}_t - H_t), \quad (2.4.7)$$

dove

$$C(H_t) = \begin{bmatrix} \frac{\partial \tilde{C}(H_t)}{\partial h_{1,t}} \\ \vdots \\ \frac{\partial \tilde{C}(H_t)}{\partial h_{K,t}} \end{bmatrix}, \quad C'(H_t) = \begin{bmatrix} \frac{\partial \tilde{C}(H_t)}{\partial h_{1,t} \partial h_{1,t}} & \cdots & \frac{\partial \tilde{C}(H_t)}{\partial h_{1,t} \partial h_{K,t}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \tilde{C}(H_t)}{\partial h_{K,t} \partial h_{1,t}} & \cdots & \frac{\partial \tilde{C}(H_t)}{\partial h_{K,t} \partial h_{K,t}} \end{bmatrix}$$

dove $C(\cdot)$ e $C'(\cdot)$ sono rispettivamente il gradiente e l'hessiana di $\tilde{C}(\cdot)$ rispetto agli elementi di H_t .

La (2.4.7) viene ridefinita da Laurent *et al.* (2013) sulla base dell'errore di previsione, per cui diventa

$$L(\hat{\Sigma}_t, H_t) = L(\hat{\Sigma}_t - H_t) = \text{vech}(\hat{\Sigma}_t - H_t)' \hat{\Lambda} \text{vech}(\hat{\Sigma}_t - H_t) \quad (2.4.8)$$

¹⁸Una funzione di perdita, L , è omogenea di ordine k se $L(a\hat{\sigma}_t^2, ah_t) = a^k L(\hat{\sigma}_t^2, h_t)$, $\forall a > 0$.

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

dove $\hat{\Lambda}$ è una matrice definita positiva di pesi associati agli elementi della matrice degli errori di previsione. La funzione di perdita nella (2.4.8) comprende diverse funzioni di perdita sia nello spazio vettoriale, come la distanza Euclidea e la distanza Euclidea pesata, che nello spazio matriciale, come la distanza di Frobenius e la distanza di Stein. In questa sezione vengono analizzati diversi test per la valutazione delle previsioni che utilizzano le funzioni di perdita appena presentate.

Per valutare l'accuratezza delle previsioni di un modello vengono utilizzate molteplici misure e diversi test, sia nell'univariato che nel multivariato.

Un metodo diffuso per valutare la precisione delle previsioni della volatilità è la regressione Mincer-Zarnowitz (MZ), si veda Mincer e Zarnowitz (1969). Questo approccio prevede la stima dei coefficienti di una regressione lineare, del tipo

$$\sigma_t^2 = \alpha + \beta h_t + \varepsilon_t \quad (2.4.9)$$

dove σ_t^2 è la vera varianza condizionale, α è una costante e β il coefficiente associato alla previsione al tempo t . Una previsione è ottimale se $\sigma_t^2 = h_t$, di conseguenza l'ipotesi di previsione ottimale si traduce in $H_0 : \alpha = 0 \cup \beta = 1$. Poiché la volatilità è una variabile latente, si utilizza una proxy e la (2.4.9) diventa

$$\hat{\sigma}_t^2 = \alpha + \beta h_t + \varepsilon_t \quad (2.4.10)$$

La (2.4.10) fornisce indicazioni sulla precisione delle previsioni mediante il test congiunto sui coefficienti e attraverso l' R^2 che diventa un indicatore di quanto la previsione sia correlata con la proxy della volatilità.

L'uso della regressione Mincer-Zarnowitz può essere esteso anche in ambito multivariato. Un approccio semplicistico prevede la stima della regressione (2.4.10) per ogni singolo elemento della matrice delle covarianze, per cui

$$\hat{\sigma}_{ij,t} = \alpha_{ij} + \beta_{ij} h_{ij,t} + \varepsilon_{ij,t} \quad (2.4.11)$$

dove $\hat{\sigma}_{ij}$ è l'elemento ij della matrice delle covarianze realizzate $\hat{\Sigma}$, per $i = 1, \dots, N(N+1)/2$ e $j = 1, \dots, N(N+1)/2$. Questo approccio diventa di difficile trattazione quando il numero delle serie è relativamente alto. Un'alternativa è un test congiunto sul modello multivariato:

$$vech(\hat{\Sigma}_t) = \alpha + diag(\beta)vech(H_t) + \varepsilon_t \quad (2.4.12)$$

dove α e β sono vettori di parametri di dimensione $N(N+1)/2 \times 1$, $vech(\cdot)$ è l'operatore che riporta in un unico vettore gli elementi della matrice triangolare bassa di una matrice

quadrata e $diag(\cdot)$ è l'operatore che trasforma un vettore di dimensione $k \times 1$ in una matrice $k \times k$ con gli elementi del vettore lungo la diagonale. Il test congiunto di $\alpha = 0$ e $\beta = 1$ viene condotto sulla (2.4.12). Patton e Sheppard (2009) evidenziano la possibilità di risultati distorti del test in campioni finiti, per questo suggeriscono l'imposizione di un vincolo sui parametri, tale per cui $\alpha = \alpha_i$ e $\beta = \beta_i, \forall i = 1, \dots, N(N + 1)/2$.

Alcune delle misure più utilizzate sono applicabili nel caso in cui modelli differenti hanno come oggetto il medesimo dataset e necessitano del solo errore di previsione per essere calcolate. Si tratta, dunque, di misure che fanno uso dell'errore di previsione e delle sue trasformazioni, come l'errore quadratico medio (*Mean Squared Error*, MSE), la radice quadrata dell'errore quadratico medio (*Root Mean Squared Error*, RMSE) e l'errore medio assoluto (*Mean Absolute Error*, MAE). I primi due sono i più diffusi in letteratura per la loro importanza teorica nell'ambito statistico, nonostante questo alcuni autori suggeriscono che sia più adeguato utilizzare il MAE, il quale risulta meno sensibile a valori estremi dell'errore di previsione.

L'errore medio assoluto (MAE) misura l'accuratezza delle previsioni mediante la media della grandezza degli errori di previsione, data dal valore assoluto di ogni errore. Il MAE è calcolato come

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i|$$

dove e_i è l'errore di previsione, per $i \in \tilde{n}$. Le misure MSE e RMSE sono utili per previsioni di modelli sulla stessa serie. Poiché dipendono strettamente dall'errore di previsione, ne segue che più piccola è la misura e migliore risulta la previsione. Nell'univariato seguono le seguenti formule:

$$MSE = \frac{1}{n} \sum_{i=1}^n e_i^2$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2}$$

dove n è il numero di errori di previsione. Nel multivariato si utilizza il RMSE basato sulla norma di Frobenius¹⁹ della matrice $n \times \tilde{n}$ dell'errore di previsione calcolato come

$$RMSE = \frac{\|e\|_2}{\sqrt{(n \cdot \tilde{n}) - 1}}$$

¹⁹La norma di Frobenius di una matrice è una norma Euclidea, costruita sulla matrice A , di dimensione $m \times n$, pari alla radice della somma dei quadrati degli elementi di una matrice, definita come $\|A\|_2 = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2}$

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Anche in ambito multivariato è preferibile, nel confronto tra gli RMSE di due modelli, il modello con un RMSE minore.

Per il confronto tra due modelli concorrenti si considerano due test basati sul confronto tra le funzioni di perdita del modello di riferimento e quello concorrente, ovvero il test Diebold-Mariano (1995), DM, e il Giacomini-White (2006), GW. I due test condividono l'ipotesi nulla di assenza di differenza nell'accuratezza delle previsioni, mentre l'ipotesi alternativa prevede che uno dei due modelli produca previsioni migliori dell'altro. Il test di Diebold e Mariano (1995) è considerato il primo tentativo di formalizzare il confronto delle previsioni di due modelli alternativi, unitamente al test di West (1996). Il test DM si basa su assunzioni fatte direttamente sulla differenza delle funzioni di perdita dell'errore di previsione.

Si consideri una funzione di perdita robusta, come quelle definite in questa sezione, il test DM si basa sul differenziale delle funzioni di perdita dei modelli k e j , per cui

$$d_t = L(\hat{\sigma}_t^2, h_{k,t}) - L(\hat{\sigma}_t^2, h_{j,t}), \quad (2.4.13)$$

nel caso univariato e

$$d_t = L(\hat{\Sigma}_t, H_{k,t}) - L(\hat{\Sigma}_t, H_{j,t}), \quad (2.4.14)$$

nel caso multivariato, dove $L(\hat{\Sigma}_t, H_{k,t})$ è una funzione di perdita come la distanza Euclidea tra vettori e la distanza di Frobenius tra matrici.

L'ipotesi nulla del test è un'uguale capacità previsionale dei due modelli, per cui $H_0 : E[d_t] = 0$. Il test assume la seguente forma

$$DM = \sqrt{T} \frac{\bar{d}}{\sqrt{w}} \xrightarrow{d} N(0, 1) \quad (2.4.15)$$

dove

$$\bar{d} = \frac{1}{T} \sum_{t=1}^T d_t$$

e

$$w = \lim_{t \rightarrow \infty} VAR(\sqrt{T}\bar{d})$$

è la sua varianza asintotica, generalmente stimata mediante la varianza campionaria.

Nel presente lavoro, si utilizza, in ambito univariato, la funzione di perdita MSE, per cui $L(\hat{\sigma}_t^2, h_t) = (\hat{\sigma}_t^2 - h_t)^2$, mentre nel multivariato si utilizzano la distanza di Frobenius tra matrici, per cui $L_F(\hat{\Sigma}_t, H_t) = \sum_{ij} (\hat{\sigma}_{ij,t} - h_{ij,t})^2$, e la distanza Euclidea tra vettori, per la quale $L(\hat{\sigma}_t, h_t) = \sum_{k=1}^{N(N+1)/2} (\hat{\sigma}_{k,t} - h_{k,t})^2$.

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

In seguito al lavoro di Diebold e Mariano (1995) si sono sviluppati diversi test per la valutazione delle previsioni out-of-sample di un modello rispetto ad uno concorrente, in un lavoro recente Giacomini e White (2006) hanno proposto di misurare la superiore capacità previsiva di un modello rispetto ad un altro mediante la perdita attesa condizionale al set informativo al tempo t , I_t , piuttosto che mediante la perdita non condizionale (come in Diebold e Mariano (1995) e in West (1996)). Il test Giacomini-White è il candidato ideale per valutare previsioni derivanti da uno schema rolling window, mentre non consente l'uso di uno schema ricorsivo. Più specificatamente, essi mostrano che nel caso in cui le previsioni siano costruite mediante uno schema rolling window con una finestra fissa d , come nel presente lavoro, allora l'ipotesi nulla è data da

$$H_0 : E[(Y_{t+1} - f_t(\hat{\beta}_{1t}))^2 - (Y_{t+1} - g_t(\hat{\beta}_{2t}))^2 \mid I_t] = 0 \quad (2.4.16)$$

dove $f_t(\hat{\beta}_{1t})$ e $g_t(\hat{\beta}_{2t})$ sono due modelli di previsione per la media condizionale della variabile di interesse Y_{t+1} , data una funzione di perdita quadratica.

Il fatto che il test si basi principalmente sulle previsioni derivanti da uno schema rolling window, consente di eliminare alcune assunzioni necessarie in altri test come il Diebold e Mariano (1995) e West (1996). In particolare, l'ipotesi di stazionarietà delle osservazioni non risulta più necessaria e il test diventa applicabile ad una classe di modelli più ampia, inclusi tutti i modelli lineari e non lineari, semi-parametrici o non parametrici, *nested* e non *nested*.

Per i modelli di tipo *nested*, Clark e West (2007) suggeriscono un modello di uguale capacità previsiva tra i modelli. Il test si basa sull'uso del *mean squared prediction error* (MSPE), definito come $MSPE_i = \sum_t^T (\hat{\sigma}_t^2 - h_{i,t})^2$, dove $i = 1, 2$. Se si definisce $\hat{f}_t = MSPE_1 - MSPE_2 + T^{-1} \sum_t^T (h_{1,t} - h_{2,t})^2$ e la sua media $\bar{f} = T^{-1} \sum_t^T \hat{f}_t$, allora il test aggiustato per il quadrato della differenza tra le previsioni è calcolato come

$$CW = \sqrt{T} \bar{f} / \sqrt{\sigma_{\hat{f}_t}^2} \quad (2.4.17)$$

dove $\sigma_{\hat{f}_t}^2 = T^{-1} \sum_t^T (\hat{f}_t - \bar{f})^2$. L'ipotesi nulla del test è $MSPE_1 = MSPE_2$, mentre l'ipotesi alternativa è $MSPE_2 < MSPE_1$, dunque il CW si presenta come un test unilaterale²⁰.

Un'importante limitazione dei test à la Diebold-Mariano è che essi rendono possibile un confronto unicamente bivariato, non è dunque possibile confrontare più di due modelli contemporaneamente. La letteratura contiene diversi approcci che cercano di risolvere

²⁰In questo lavoro non vengono presentati i risultati del test CW, ma sono disponibili su richiesta a a.bucci@univpm.it

questo problema, come il test di abilità previsiva superiore di Hansen (2005) e il model confidence set (MCS) di Hansen *et al.* (2011b). La procedura MCS definisce, per un dato livello di confidenza, l'insieme di modelli che contiene le migliori previsioni out-of-sample. In questo modo vengono confrontati tutti i modelli contemporaneamente, eliminando le specificazioni con le migliori previsioni che fanno parte dell'insieme di modelli inizialmente considerati. La procedura di Hansen consiste in una sequenza di test che permette di costruire un insieme di modelli "superiori", SSM (*superior set of models*), dove l'ipotesi nulla di uguale capacità previsiva non è rifiutata ad un certo livello di confidenza. La procedura parte da un insieme iniziale di modelli \hat{M}^0 di dimensione m che comprende tutte le specificazioni di interesse e restituisce, per un dato livello di confidenza $1 - \alpha$, un insieme ristretto, appunto l'SSM, $\hat{M}_{1-\alpha}^*$ di dimensione $m^* \leq m$. Il MCS procede in maniera sequenziale, eliminando ad ogni passo il modello peggiore, fino a quando non viene accettata l'ipotesi nulla di uguale capacità previsiva per tutti i modelli appartenenti al SSM.

La procedura MCS dipende fortemente dall'ordinamento imposto dalla funzione di perdita. In questo lavoro si utilizzano la funzione di perdita MSE per le singole serie e la distanza di Frobenius unitamente alla distanza Euclidea per l'insieme delle serie delle covarianze.

2.4.2 Risultati previsioni

In questa sezione vengono analizzati i risultati dei metodi di valutazione delle previsioni introdotti nella sezione precedente.

In primo luogo, nella tabella 2.18, sono riportati i RMSE del modello cross-country per ogni singola equazione. I modelli con il minor RMSE sono, per la maggior parte delle serie, quelli non lineari, seguiti dal modello lineare che utilizza la scomposizione di Cholesky della matrice delle covarianze e le variabili esogene come regressori.

Dalla tabella 2.19 dei RMSE nel mercato americano, in ambito univariato, si nota che il minore RMSE è spesso associato al modello CholVARExo, mentre il modello HEAVY e il modello BEKK presentano RMSE per le singole serie decisamente superiori agli altri modelli.

Nella tabella 2.20 sono riportati i RMSE basati sulla distanza di Frobenius (definita nella sezione precedente) per i vari modelli dalla vera varianza realizzata, nei due distinti dataset. Come si può notare, per il dataset cross-country l'RMSE è minimo per i due modelli non lineari mentre la specificazione Cholesky-VAR con e senza esogene non si

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.18: RMSE cross-country

Il modello CholVARExo utilizza la procedura Cholesky-VAR con l'uso di variabili esogene, mentre il modello Chol-VAR non prevede l'uso di esogene. NLChExo e NLCh rappresentano le previsioni derivanti dai modelli non lineari, il modello VARExo non ricorre all'uso della scomposizione di Cholesky della matrice varianze e covarianze e utilizza le determinanti, mentre il VAR non comprende l'uso di esogene. Infine il logVARExo comprende le previsioni derivanti da un modello equivalente a quello di Bauer e Vorkink (2011) con esogene, mentre il logVAR è privo di determinati, il DCC include le previsioni rolling window derivanti da un modello DCC, il BEKK quelle derivanti dal modello omonimo e l'HEAVY rappresenta il modello introdotto da Noureldin *et al.* (2011) in ambito multivariato.

	SP	SP-JPN	SP-UK	SP-DAX	JPN	JPN-UK	JPN-DAX	UK	UK-DAX	DAX
CholVARExo	0.0038	0.0026	0.0023	0.0033	0.0084	0.0039	0.0041	0.0038	0.0035	0.0055**
NLChExo	0.0026**	0.0022**	0.0023	0.0037	0.0051**	0.0029**	0.0039	0.0028**	0.0033*	0.0059
NLCh	0.0025*	0.0019*	0.0021*	0.0033**	0.0043*	0.0026*	0.0039**	0.0026*	0.0036	0.0053*
CholVAR	0.0046	0.0024	0.0026	0.0036	0.0079	0.0039	0.0041	0.0041	0.0039	0.0058
VARExo	0.0047	0.0028	0.0026	0.0037	0.0089	0.0044	0.0045	0.0041	0.0039	0.0058
VAR	0.0052	0.0027	0.0028	0.0040	0.0089	0.0044	0.0045	0.0044	0.0042	0.0062
logVARExo	0.0033	0.0031	0.0023	0.0034	0.0082	0.0035	0.0038*	0.0033	0.0033**	0.0055
logVAR	0.0050	0.0024	0.0028	0.0039	0.0077	0.0040	0.0042	0.0043	0.0042	0.0059
DCC	0.0057	0.0025	0.0029	0.0042	0.0078	0.0041	0.0043	0.0043	0.0046	0.0062
BEKK	0.0716	0.0201	0.0938	0.0165	0.0539	0.0255	0.0216	0.1295	0.0171	0.0158
HEAVY	0.0036	0.0026	0.0022**	0.0025*	0.0094	0.0052	0.0046	0.0029	0.0060	0.0045

* e ** indicano i modelli con minor RMSE

Tabella 2.19: RMSE mercato americano

	SP	SP-10Y	SP-30d	10Y	10Y-30d	30d
CholVARExo	0.0047	0.0009	0.000113	0.000272**	3.000036	6.7e-6
NLChExo	0.0034*	0.0006**	0.000074**	0.000767	0.000034	5.2e-6**
NLCh	0.0042**	0.0003*	0.000067*	0.000226*	0.000028	4.0e-6*
CholVAR	0.0054	0.0006	0.000083	0.000279	0.000027	5.8e-6
VARExo	0.0055	0.0006	0.000082	0.000528	0.000026	5.9e-6
VAR	0.0054	0.0006	0.000084	0.000459	0.000027	6.1e-6
logVARExo	0.0077	0.0029	0.000673	0.000638	0.000179	4.5e-5
logVAR	0.0059	0.0006	0.000083	0.000277	0.000025**	5.9e-6
DCC	0.0071	0.0007	0.000082	0.000334	0.000023*	8.6e-6
BEKK	0.1001	0.0223	0.000316	0.023900	0.000217	0.00180
HEAVY	0.0069	0.0049	0.003425	0.013509	0.001181	0.000102

* e ** indicano modelli con minor RMSE

Tabella 2.20: RMSE basato sulla norma di Frobenius dell'errore di previsione

	Cross-Country	US
CholVARExo	0.0308	0.0073
NLChExo	0.0299	0.0068
NLCh	0.0274	0.0068
CholVAR	0.0304	0.0044
VARExo	0.0351	0.0090
VAR	0.0348	0.0049
logVARExo	0.0314	0.0109
logVAR	0.0310	0.0057
DCC	0.0347	0.0113
BEKK	0.3419	0.0440
HEAVY	0.4118	0.0703

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

discosta molto dal risultato dei precedenti modelli, performando meglio di tutte le altre specificazioni.

Per il dataset americano il minore RMSE è dato dalle previsioni del modello CholVAR.

I risultati della regressione Mincer-Zarnowitz, riportati nella tabella 2.21, non forniscono indicazioni rilevanti sulla bontà delle previsioni dei modelli considerati. Il test congiunto sui parametri delle singole equazioni accetta raramente l'ipotesi nulla ad un livello di confidenza superiore al 10%.

La tabella 2.22 riporta i risultati della Mincer-Zarnowitz nel dataset americano. Rispetto al dataset cross-country, il test congiunto accetta l'ipotesi nulla in un numero maggiore di casi. Tuttavia, per la maggior parte delle serie e dei modelli l'ipotesi nulla di capacità previsionale ottimale viene rifiutata ad ogni livello di significatività. In definitiva, la regressione MZ evidenzia l'incapacità dei metodi impiegati nel prevedere la matrice delle covarianze realizzate.

Tabella 2.21: Mincer-Zarnowitz cross-country

$\hat{\alpha}$ e $\hat{\beta}$ sono rispettivamente la costante e il coefficiente associato alla previsione stimati, mentre MZ è il risultato del test congiunto che ha per ipotesi nulla $H_0 : \hat{\alpha} = 0 \wedge \hat{\beta} = 1$.

		SP	SP-JPN	SP-UK	SP-DAX	JPN	JPN-UK	JPN-DAX	UK	UK-DAX	DAX
CholVARExo	$\hat{\alpha}$	0.0003	0.0004	0.0006	0.0009	0.0030	0.0007	0.0007	0.0007	0.0010	0.0018
	$\hat{\beta}$	1.0836	0.4795	0.9103	0.8900	0.4387	0.6534	0.6390	0.9961	0.9640	0.9002
	M-Z	0.0779	0.0000	0.0041	0.0027	0.0000	0.0031	0.0009	0.0543	0.0008	0.0003
CholVAR	$\hat{\alpha}$	-0.0003	0.0002	0.0003	0.0007	0.0026	0.0005	0.0005	0.0005	0.0008	0.0016
	$\hat{\beta}$	1.3991	0.9018	1.2796	1.0977	0.5676	0.9302	0.8524	1.1514	1.1063	1.0047
	M-Z	0.0005	0.5941	0.0002	0.0030	0.0008	0.3360	0.3332	0.0139	0.0004	0.0004
VARExo	$\hat{\alpha}$	0.0005	0.0005	0.0006	0.0010	0.0032	0.0008	0.0009	0.0008	0.0012	0.0016
	$\hat{\beta}$	0.9315	0.3084	0.8529	0.7751	0.3440	0.4242	0.4337	0.7688	0.7186	0.7069
	M-Z	0.5313	0.0000	0.0181	0.0036	0.0000	0.0000	0.0000	0.0264	0.0001	0.0004
VAR	$\hat{\alpha}$	0.0006	0.0005	0.0006	0.0012	0.0036	0.0010	0.0010	0.0010	0.0013	0.0019
	$\hat{\beta}$	0.8971	0.3107	0.8982	0.7161	0.2685	0.3313	0.3950	0.7100	0.6743	0.6527
	M-Z	0.5620	0.0000	0.0799	0.0050	0.0000	0.0000	0.0000	0.0213	0.0002	0.0003
logVARExo	$\hat{\alpha}$	0.0011	0.0004	0.0009	0.0013	0.0030	0.0005	0.0007	0.0008	0.0011	0.0018
	$\hat{\beta}$	0.9119	0.3647	0.7297	0.8080	0.4753	0.8482	0.7262	1.0550	1.0021	0.9837
	M-Z	0.0001	0.0000	0.0000	0.0000	0.0000	0.0867	0.0047	0.0002	0.0000	0.0000
logVAR	$\hat{\alpha}$	-0.0008	0.0002	0.0000	0.0004	0.0017	0.0001	0.0003	-0.0003	0.0001	0.0004
	$\hat{\beta}$	2.1915	1.0524	1.9430	1.7151	0.9221	1.5572	1.2585	1.9236	1.7645	1.5388
	M-Z	0.0000	0.3526	0.0000	0.0000	0.0309	0.0231	0.1187	0.0000	0.0000	0.0000
DCC	$\hat{\alpha}$	-0.0054	-0.0004	-0.0023	-0.0017	0.0005	-0.0010	-0.0001	-0.0022	-0.0008	0.0003
	$\hat{\beta}$	4.4123	0.9351	2.8519	2.7641	1.2355	1.8643	1.0965	2.1322	2.3037	1.2483
	M-Z	0.0000	0.0248	0.0000	0.0000	0.0419	0.1286	0.9368	0.0000	0.0000	0.0089
BEKK	$\hat{\alpha}$	0.0033	0.0007	0.0018	0.0025	0.0049	0.0014	0.0016	0.0029	0.0031	0.0050
	$\hat{\beta}$	0.0000	0.0000	-0.0000	0.0000	-0.0000	-0.0000	-0.0000	0.0000	-0.0000	0.0000
	M-Z	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
HEAVY	$\hat{\alpha}$	-0.0000	0.0005	-0.0002	-0.0003	0.0048	0.0013	0.0017	0.0004	0.0033	0.0002
	$\hat{\beta}$	1.3403	0.7018	1.9184	1.7113	0.1221	-3.9095	-0.6782	0.7077	-0.2762	0.9140
	M-Z	0.0000	0.0409	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.2117
NLExo	$\hat{\alpha}$	-0.0000	0.0005	0.0004	0.0013	-0.0017	-0.0002	0.0009	-0.0005	0.0008	0.0026
	$\hat{\beta}$	1.0729	0.7096	1.5728	1.1886	1.8215	2.0885	0.8144	1.6237	1.4946	0.7798
	M-Z	0.0969	0.0000	0.0000	0.0000	0.0000	0.0000	0.0101	0.0000	0.0000	0.0000
NL	$\hat{\alpha}$	-0.0010	-0.0005	-0.0003	0.0002	-0.0018	-0.0006	-0.0024	-0.0006	-0.0003	-0.0012
	$\hat{\beta}$	1.6477	2.7784	2.3618	2.4547	1.8330	1.9596	5.4189	1.6852	2.6232	2.4742
	M-Z	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.22: Mincer-Zarnowitz US

$\hat{\alpha}$ e $\hat{\beta}$ sono rispettivamente la costante e il coefficiente associato alla previsione stimati, mentre MZ è il risultato del test congiunto che ha per ipotesi nulla $H_0 : \hat{\alpha} = 0 \wedge \hat{\beta} = 1$.

		SP	SP-10Y	SP-30d	10Y	10Y-30d	30d
CholVARExo	$\hat{\alpha}$	0.0006	-0.0003	-0.0000	0.0001	0.0000	0.0000
	$\hat{\beta}$	0.8299	0.2925	0.3266	0.7772	0.3160	0.4898
	M-Z	0.0256	0.0000	0.0000	0.0096	0.0000	0.0000
CholVAR	$\hat{\alpha}$	0.0004	-0.0003	-0.0000	0.0001	-0.0000	0.0000
	$\hat{\beta}$	0.9689	0.5129	0.5849	0.9333	0.8758	1.0600
	M-Z	0.8111	0.0001	0.0571	0.0266	0.2909	0.0000
VARExo	$\hat{\alpha}$	0.0011	-0.0003	-0.0000	0.0003	-0.0000	-0.0000
	$\hat{\beta}$	0.7590	0.4728	0.5450	0.1675	0.8842	0.8182
	M-Z	0.0062	0.0000	0.0001	0.0000	0.0041	0.0000
VAR	$\hat{\alpha}$	0.0007	-0.0003	-0.0000	0.0003	-0.0000	-0.0000
	$\hat{\beta}$	0.8818	0.5169	0.5419	0.2387	1.2646	1.0870
	M-Z	0.3855	0.0006	0.0383	0.0000	0.0024	0.0000
logVARExo	$\hat{\alpha}$	0.0015	-0.0004	-0.0000	0.0003	0.0000	0.0000
	$\hat{\beta}$	0.4903	0.0645	0.0422	0.2277	0.0211	0.0256
	M-Z	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
logVAR	$\hat{\alpha}$	-0.0002	-0.0002	-0.0000	-0.0000	-0.0000	0.0000
	$\hat{\beta}$	1.3375	0.6197	0.8531	1.3265	1.2872	1.3742
	M-Z	0.0409	0.0072	0.4931	0.0002	0.3550	0.0000
DCC	$\hat{\alpha}$	0.0001	-0.0004	-0.0000	0.0156	-0.0000	0.0000
	$\hat{\beta}$	1.8395	0.9285	0.6546	-38.3476	1.2529	0.3724
	M-Z	0.0034	0.0000	0.2150	0.0043	0.0004	0.0000
BEKK	$\hat{\alpha}$	0.0040	-0.0005	-0.0000	0.0004	0.0000	0.0000
	$\hat{\beta}$	-0.0020	-0.0005	0.0003	-0.0002	0.0000	0.0000
	M-Z	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
HEAVY	$\hat{\alpha}$	-0.0078	-0.0031	-0.0014	-0.0003	0.0000	0.0000
	$\hat{\beta}$	5.9980	0.5913	0.4081	-0.0505	0.0128	0.0097
	M-Z	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
NLExo	$\hat{\alpha}$	-0.0004	-0.0002	-0.0000	0.0003	0.0000	0.0000
	$\hat{\beta}$	1.3106	0.5473	0.2625	0.2585	0.2919	0.2927
	M-Z	0.0049	0.0000	0.0000	0.0000	0.0000	0.0000
NL	$\hat{\alpha}$	-0.0008	-0.0003	-0.0002	0.0001	0.0000	0.0000
	$\hat{\beta}$	1.5737	0.5192	0.2484	0.5191	0.2844	0.4684
	M-Z	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000

L'accuratezza delle previsioni viene valutata anche mediante i test di confronto tra due modelli, dapprima vengono utilizzate le funzioni di perdita delle previsioni delle singole equazioni, poi si utilizzano quelle in ambito multivariato. Il pannello A della tabella 2.23 presenta i test DM per le previsioni di varianze e covarianze realizzate per il dataset cross-

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

country, nel quale la specificazione benchmark è quella Cholesky-VAR con variabili esogene. Il test rifiuta l'ipotesi nulla solo in alcuni casi e sempre a favore del modello benchmark. In un unico altro caso il test rifiuta l'ipotesi nulla a favore del modello di confronto, per la serie SP-DAX e a favore del modello HEAVY. Nonostante il test non indichi un chiaro rifiuto dell'ipotesi nulla per la maggior parte delle serie, si può trarre comunque informazione dal segno della statistica DM. Infatti, un segno negativo della statistica indica che le previsioni del modello benchmark producono, in media, una perdita minore delle previsioni del modello di confronto. Come si evince dalla tabella, la statistica presenta quasi esclusivamente il segno negativo, ad indicare una propensione per le previsioni derivanti dalla specificazione Cholesky-VAR seppure non statisticamente significativa.

Nel pannello B della tabella 2.23 sono riportate le statistiche GW sul dataset cross-country. Il test Giacomini-White rifiuta spesso l'ipotesi nulla di uguale capacità previsiva e quasi sempre a favore della specificazione Cholesky-VAR con esogene. Il segno delle statistiche conferma quanto emerge dall'RMSE multivariato e dal test Diebold-Mariano. Il test presenta, nella maggior parte dei casi, un segno negativo, ad indicare una migliore capacità previsiva da parte del modello benchmark. In particolare, è forte la superiorità delle previsioni del modello Cholesky-VAR con esogene rispetto ai due modelli che non considerano l'uso della scomposizione di Cholesky e alle previsioni dei modelli DCC e BEKK.

La tabella 2.24 riporta i risultati dei test DM e GW nel dataset cross-country, quando si utilizzano come funzioni di perdita la distanza di Frobenius e la distanza Euclidea. La tabella conferma le indicazioni derivanti dai risultati dei due test sulle singole serie. Le statistiche test risultano significative solo nei confronti del modello benchmark, CholVA-RExo, con i modelli VARExo, VAR, DCC, BEKK e HEAVY e sempre a favore del modello benchmark. Nel confronto con il modello non lineare senza esogene, le statistiche test sono significative a favore di questo modello. Negli altri casi i test non evidenziano una differenza significativa tra le previsioni ottenute dagli altri modelli, il segno negativo della statistica test induce a pensare ad una migliore, seppure non significativa, accuratezza delle previsioni derivanti dal modello di riferimento.

Tabella 2.23: Valutazione previsioni out-of-sample cross-country

Il pannello A riporta il Diebold e Mariano (1995) test tra le previsioni dei modelli presenti nella prima colonna e quelle derivanti da un modello Cholesky-VAR con esogene, mentre nel pannello B sono mostrati i risultati di un Giacomini e White (2006) test. Le performance delle previsioni sono valutate per tutte le serie delle varianze e covarianze realizzate.

Pannello A: Test Diebold-Mariano (Rolling window = 370 osservazioni)										
	SP	SP-JPN	SP-UK	SP-DAX	JPN	JPN-UK	JPN-DAX	UK	UK-DAX	DAX
CholVAR	-1.3605	0.4266	-0.7245	-0.6804	0.4359	-0.1482	-0.1355	-1.0252	-0.9441	-0.6042
	0.1753	0.6702	0.4696	0.4971	0.6634	0.8823	0.8924	0.3066	0.3463	0.5464
VARExo	-1.8387 ^c	-1.0024	-1.0307	-0.8326	-0.7333	-1.4626	-1.6444	-1.5280	-1.2128	-0.5948
	0.0675	0.3174	0.3040	0.4061	0.4643	0.1452	0.1017	0.1282	0.2267	0.5527
VAR	-1.5098	-0.2639	-0.9519	-0.8999	-0.4633	-0.9367	-1.0414	-1.4642	-1.1831	-0.8407
	0.1328	0.7921	0.3423	0.3693	0.6437	0.3501	0.2989	0.1448	0.2382	0.4016
logVARExo	0.6042	-0.5964	0.0541	-0.0456	0.2127	0.9592	0.8197	0.7715	0.4369	-0.3068
	0.5464	0.5516	0.9569	0.9637	0.8318	0.3387	0.4133	0.4414	0.6627	0.7594
logVAR	-1.6498	0.3391	-1.1166	-1.0963	0.5081	-0.1999	-0.2263	-1.3380	-1.3630	-0.8042
	0.1006	0.7349	0.2656	0.2743	0.6120	0.8417	0.8212	0.1825	0.1745	0.4223
DCC	-1.7980 ^c	0.0786	-1.2024	-1.4826	0.4231	-0.3830	-0.5064	-1.3773	-1.8157 ^c	-1.2516
	0.0738	0.9374	0.2307	0.1398	0.6720	0.7021	0.6132	0.1700	0.0709	0.2123
BEKK	-1.1591	-1.3746	-1.0448	-1.4778	-1.0248	-1.2045	-1.0486	-1.0324	-1.4071	-1.6459
	0.2478	0.1708	0.2975	0.1411	0.3068	0.2299	0.2957	0.3032	0.1610	0.1014
HEAVY	0.6336	0.0423	0.7562	2.5445 ^b	-0.9443	-1.2479	-1.0966	0.6640	-3.0678 ^a	1.6817 ^c
	0.5271	0.9663	0.4505	0.0117	0.3462	0.2136	0.2742	0.5075	0.0025	0.0943
NL	1.6076	1.6060	1.0448	0.3741	1.6674 ^c	1.3848	0.7405	1.7720 ^c	-0.8269	0.3486
	0.1096	0.1099	0.2974	0.7087	0.0971	0.16771	0.4599	0.0779	0.40931	0.7277
Pannello B: Test Giacomini-White (Rolling window = 370 osservazioni)										
	SP	SP-JPN	SP-UK	SP-DAX	JPN	JPN-UK	JPN-DAX	UK	UK-DAX	
CholVAR	-1.6754 ^c	0.6348	0.1987	0.7718	0.8076	0.4413	0.2875	-0.2654	-0.7160	-0.9438
	0.0939	0.5255	0.8425	0.4402	0.4193	0.6590	0.7737	0.7907	0.4740	0.3453
VARExo	-4.5022 ^a	-3.3026 ^a	-2.3974 ^b	-1.4472	-3.2922 ^a	-4.0512 ^a	-3.9089 ^a	-4.4680 ^a	-1.3218	-0.7549
	0.0000	0.0010	0.0165	0.1478	0.0010	0.0001	0.0001	0.0000	0.1862	0.4503
VAR	-3.7931 ^a	-1.5208	-0.9221	-0.3988	-2.6186 ^a	-2.7686 ^a	-3.6284 ^a	-4.8880 ^a	-2.0982 ^b	-1.7865 ^c
	0.0001	0.1283	0.3565	0.6900	0.0088	0.0056	0.0003	0.0000	0.0359	0.0740
logVARExo	0.2971	-0.7538	-0.5126	-1.4576	1.0877	1.3431	0.9315	0.0484	-1.5875	-2.4459 ^b
	0.7664	0.4509	0.6082	0.1450	0.2767	0.1793	0.3516	0.9614	0.1124	0.0145
logVAR	-1.6875 ^c	0.6700	-0.7538	-0.8010	1.0023	0.6560	0.4761	-1.3305	-2.4745 ^b	-1.0678
	0.0915	0.5029	0.4510	0.4232	0.3162	0.5119	0.6340	0.1833	0.0133	0.2856
DCC	-2.7023 ^a	-2.8896 ^a	-2.0754 ^b	-1.8463 ^c	0.3252	-1.3017	-1.2937	-3.7200 ^a	-3.2403 ^a	-2.3787 ^b
	0.0069	0.0039	0.0380	0.0648	0.7451	0.1930	0.1958	0.0002	0.0012	0.0174
BEKK	-2.0214 ^b	-2.5194 ^a	-2.0262 ^b	-2.2276 ^b	-2.0645 ^b	-2.7900 ^a	-2.2384 ^b	-2.1900 ^b	-2.5510 ^a	-3.7023 ^a
	0.0216	0.0059	0.0214	0.0129	0.0195	0.0026	0.0126	0.0143	0.0054	0.0001
HEAVY	-0.5266	1.7811 ^c	2.0926 ^b	3.7096 ^a	-5.8945 ^a	-4.2780 ^a	-2.8646 ^a	0.9465	-6.7238 ^a	1.4244
	0.5985	0.0749	0.0364	0.0002	0.0000	0.0000	0.0042	0.3439	0.0000	0.1543
NL	2.7906 ^a	-0.3147	-0.6109	-2.8914 ^a	2.7144 ^a	1.7170 ^c	-0.1984	2.4531 ^b	-1.6701 ^c	-1.4382
	0.0053	0.7529	0.5413	0.0038	0.0066	0.0859	0.8427	0.0142	0.0949	0.1504

a, b e c indicano i livelli di significatività rispettivamente al 1%, 5% e 10%

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.24: Test in ambito multivariato, dataset cross-country

	Diebold-Mariano		Giacomini-White	
	Frobenius	Euclidea	Frobenius	Euclidea
CholVAR	0.2689	0.2079	0.0829	0.0459
	0.7883	0.8355	0.9340	0.9635
VARExo	-3.4346 ^a	-3.4275 ^a	-11.3314 ^a	-11.2968 ^a
	0.0007	0.0008	0.0000	0.0000
VAR	-2.9057 ^a	-3.0548 ^a	-15.0955 ^a	-15.8906 ^a
	0.0041	0.0026	0.0000	0.0000
logVARExo	-0.5946	-0.4439	-1.3821	-1.4466
	0.5526	0.6576	0.1686	0.1497
logVAR	-0.1305	-0.0739	-0.0202	-0.0069
	0.8963	0.9411	0.9839	0.9945
DCC	-2.9802 ^a	-2.7328 ^a	-8.5253 ^a	-7.5502 ^a
	0.0033	0.0069	0.0000	0.0000
BEKK	-1.5858	-1.5957	-3.8078 ^a	-3.7854 ^a
	0.1144	0.1122	0.0000	0.0000
HEAVY	-7.9076 ^a	-7.0616 ^a	-21.7943 ^a	-19.8866 ^a
	0.0000	0.0000	0.0000	0.0000
NL	1.9979 ^b	2.4622 ^b	3.9060 ^a	5.9208 ^a
	0.0471	0.0147	0.0001	0.0000

a, b e c indicano i livelli di significatività rispettivamente al 1%, 5% e 10%

Nel pannello A della tabella 2.25 sono riportate le statistiche DM per il dataset americano. I risultati si presentano simili a quelli ottenuti per il dataset cross-country; sono infatti pochi i casi di differenze significative dall'ipotesi nulla. In aggiunta, il segno della statistica DM risulta negativo per la maggior parte dei confronti con altri modelli, ad indicare una migliore capacità previsiva della specificazione Cholesky-VAR con esogene.

La tabella 2.25, per il GW sul mercato americano, mostra risultati simili alla tabella del dataset cross-country. Allo stesso tempo, il test Giacomini-White rifiuta meno spesso l'ipotesi nulla e presenta meno raramente un segno positivo. Le maggiori differenze a favore del modello benchmark si presentano sempre nei confronti delle previsioni derivanti dalle specificazioni senza scomposizione di Cholesky, dal modello DCC, dal modello BEKK e dal modello HEAVY.

Anche per il dataset americano vengono riportati i risultati dei test in ambito multivariato, in tabella 2.26. I test confermano quanto emerge in ambito univariato, in particolare il test GW rifiuta l'ipotesi nulla di uguale capacità previsiva per tutti i confronti e sempre a favore del modello benchmark.

Le indicazioni derivanti dai test DM e GW, sia in ambito univariato che multivariato, suggeriscono che il modello introdotto in questo lavoro preveda meglio di quelli esistenti

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

in letteratura, in particolar modo il vantaggio è evidente nei confronti dei modelli classici come il DCC e il BEKK e di un modello recente come il modello HEAVY.

Tabella 2.25: Valutazione previsioni out-of-sample mercato americano

Pannello A: Test Diebold-Mariano (Rolling window = 214 osservazioni)						
	SP	SP-10Y	SP-30d	10Y	10Y-30d	30d
CholVAR	-0.5604	0.9387	0.8782	-0.1923	0.7797	0.6273
	0.5764	0.3500	0.3818	0.8479	0.4373	0.5318
VARExo	-1.2450	1.0246	1.0018	-1.7841 ^c	0.7889	0.6247
	0.2159	0.3079	0.3187	0.0773	0.4319	0.5335
VAR	-0.6792	0.9687	0.8544	-1.1612	0.7410	0.4242
	0.4985	0.3349	0.3948	0.2481	0.4603	0.6723
logVARExo	-1.4346	-1.1092	-1.0690	-0.9857	-1.0217	-1.0232
	0.1543	0.2699	0.2875	0.3265	0.3092	0.3085
logVAR	-0.8186	0.9423	0.8850	-0.1413	0.8212	0.5779
	0.4149	0.3482	0.3781	0.8879	0.4134	0.5646
DCC	-1.3161	0.4988	0.9109	-2.1129 ^b	0.9023	-0.9679
	0.1909	0.6189	0.3644	0.0369	0.3689	0.3353
BEKK	-1.9307 ^c	-1.5383	-1.8589 ^c	-1.5564	-1.8362 ^c	-1.4552
	0.0562	0.1269	0.0658	0.1226	0.0691	0.1486
HEAVY	-1.3255	-5.0151 ^a	-6.8899 ^a	-4.7752 ^a	-4.5222 ^a	-4.9556 ^a
	0.1879	0.0000	0.0000	0.0000	0.0000	0.0000
NL	0.5257	1.7019 ^c	1.3934	0.8475	0.6027	1.3043
	0.6002	0.0917	0.1664	0.3986	0.5480	0.1950
Pannello B: Test Giacomini-White (Rolling window = 214 osservazioni)						
	SP	SP-10Y	SP-30d	10Y	10Y-30d	30d
CholVAR	-0.7453	1.1759	0.9910	-0.4151	0.7286	0.8643
	0.4561	0.2397	0.3217	0.6781	0.4662	0.3874
VARExo	-2.6716 ^a	0.7809	0.4588	-2.9661 ^a	-0.5634	-3.2597 ^a
	0.0075	0.4349	0.6464	0.0030	0.5731	0.0011
VAR	-1.1537	1.2516	0.6927	-1.6480 ^c	-0.5156	-2.6631 ^a
	0.2486	0.2107	0.4885	0.0994	0.6062	0.0077
logVARExo	-1.5169	-1.2857	-1.1740	-0.9116	-0.9529	-1.0487
	0.1293	0.1985	0.2404	0.3620	0.3407	0.2943
logVAR	-1.0733	0.9664	1.3633	0.0771	1.8329 ^c	1.4357
	0.2831	0.3338	0.1728	0.9386	0.0668	0.1511
DCC	-1.9984 ^b	-0.5543	1.6844 ^c	-4.2552 ^a	0.3455	-3.3399 ^a
	0.0457	0.5794	0.0921	0.0000	0.7297	0.0008
BEKK	-2.2981 ^b	-2.1082 ^b	-2.4795 ^b	-2.0723 ^b	-2.5085 ^b	-2.0952 ^b
	0.0216	0.0350	0.0132	0.0382	0.0121	0.0362
HEAVY	-1.9741 ^b	-8.7716 ^a	-11.6748 ^a	-7.2030 ^a	-6.7741 ^a	-7.8961 ^a
	0.0484	0.0000	0.0000	0.0000	0.0000	0.0000
NL	-0.5341	0.9542	-0.4045	-3.7113 ^a	-4.3409 ^a	-3.9488 ^a
	0.5933	0.3400	0.6859	0.0002	0.0000	0.0000

α , b e c indicano i livelli di significatività rispettivamente al 1%, 5% e 10%

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.26: Test in ambito multivariato, dataset mercato americano

	Diebold-Mariano		Giacomini-White	
	Frobenius	Euclidea	Frobenius	Euclidea
CholVAR	-0.6032	-0.6562	-5.8681 ^a	-5.9224 ^a
	0.5477	0.5131	0.0000	0.0000
VARExo	-3.0064 ^a	-2.9208 ^a	-9.3262 ^a	-8.6052 ^a
	0.0033	0.0043	0.0000	0.0000
VAR	-1.0496	-1.1076	-5.8561 ^a	-5.8958 ^a
	0.2963	0.2705	0.0000	0.0000
logVARExo	-1.4490	-1.4822	-2.3122 ^b	-2.3724 ^b
	0.1503	0.1413	0.0227	0.0195
logVAR	-0.9135	-0.9650	-4.3888 ^a	-4.7030 ^a
	0.3631	0.3367	0.0000	0.0000
DCC	-2.1460 ^b	-2.1118 ^b	-4.858 ^a	-4.5950 ^a
	0.0342	0.0371	0.0000	0.0000
BEKK	-2.6229 ^b	-2.6298 ^a	-6.5417 ^a	-6.5584 ^a
	0.0100	0.0098	0.0000	0.0000
HEAVY	-13.2154 ^a	-12.8528 ^a	-14.9589 ^a	-14.5366 ^a
	0.0000	0.0000	0.0000	0.0000
NL	-0.5089	-0.5467	-4.3694 ^a	-4.4357 ^a
	0.6119	0.5857	0.0000	0.0002

a, *b* e *c* indicano i livelli di significatività rispettivamente al 1%, 5% e 10%

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Nel corso di questa sezione si analizzano anche le previsioni out-of-sample dei modelli non lineari. La tabella 2.27 riporta i risultati, nel dataset cross-country, dei test DM e GW, nei quali il modello previsionale di riferimento è il VLSTAR con esogene sui fattori di Cholesky della matrice delle covarianze realizzate, presentato nel paragrafo 2.2. Le statistiche DM, nel pannello A, evidenziano come non ci siano differenze significative tra le previsioni dei modelli considerati. Tuttavia si può notare che la statistica test presenta un segno negativo nel confronto con tutti i modelli, fatta eccezione per il modello non lineare senza esogene che sembra prevedere meglio rispetto al modello di riferimento e del modello HEAVY, per alcune serie.

Il pannello B della 2.27 riporta i risultati del test GW sulle singole serie del dataset cross-country. Il rifiuto dell'ipotesi nulla di uguale capacità previsionale è evidente per la maggior parte delle serie a favore del modello benchmark. Come nel caso del test DM, il test rifiuta l'ipotesi nulla quando il modello alternativo è il modello non lineare senza esogene e per la serie SP-DAX. Infine, sembra che per le serie DAX, SP-JPN e JPN-DAX non ci siano differenze, in termini previsionali, tra il modello di riferimento e gli altri modelli selezionati.

La tabella 2.28 riporta i risultati dei test DM e GW quando il modello di riferimento è il NLExo e quando la funzione di perdita è rispettivamente la distanza di Frobenius, tra la vera matrice delle covarianze realizzate e quella prevista dai modelli considerati in questo lavoro, e la distanza Euclidea, tra i vettori di dimensione $N(N + 1)/2 \times 1$ di covarianze. Il test DM rifiuta l'ipotesi nulla nel confronto con il VARExo, con il DCC e con il modello HEAVY, per entrambe le funzioni di perdita e sempre a favore del modello benchmark, negli altri casi la statistica test presenta un segno negativo, fatta eccezione per il confronto con il modello non lineare senza esogene, NL. Il test GW rifiuta l'ipotesi nulla nella quasi totalità dei casi, sempre a favore del modello benchmark. In particolare, il p -value è minore nel confronto con il modello HEAVY, con il modello CholVARExo, con il modello DCC e con quello BEKK. L'unico modello che sembra prevedere meglio del NLExo è il modello non lineare senza le esogene.

Tabella 2.27: Valutazione previsioni non lineari out-of-sample cross-country

Pannello A: Diebold-Mariano Test										
	SP	SP-JPN	SP-UK	SP-DAX	JPN	JPN-UK	JPN-DAX	UK	UK-DAX	DAX
NL	0.1508	0.3699	1.1505	1.5745	2.0131 ^b	1.7013 ^c	0.03402	1.6945 ^c	-1.0763	0.8551
	0.8803	0.7118	0.2514	0.1170	0.0455	0.0905	0.97289	0.0918	0.2832	0.3936
CholVARExo	-1.1166	-0.5664	-0.3534	1.2354	-1.5765	-1.2959	-0.1984	-1.5960	-0.4203	0.5302
	0.2656	0.5718	0.7242	0.2182	0.1166	0.1966	0.8429	0.1121	0.6747	0.5966
CholVAR	-1.2941	-0.2498	-0.7928	0.2053	-1.1977	-0.9898	-0.2111	-1.4191	-0.7416	0.1861
	0.1972	0.8029	0.4289	0.8376	0.2325	0.3235	0.8331	0.1575	0.4593	0.8526
VARExo	-1.5412	-0.7882	-0.9209	0.0128	-1.6641 ^c	-1.4086	-0.8273	-1.6690 ^c	-0.8594	0.1737
	0.1249	0.4316	0.3582	0.9898	0.0977	0.1606	0.4091	0.0968	0.3912	0.8623
VAR	-1.4169	-0.5456	-0.9602	-0.3409	-1.3962	-1.1738	-0.6775	-1.5659	-0.9524	-0.1295
	0.1582	0.5860	0.3382	0.7336	0.1643	0.2419	0.4989	0.1190	0.3421	0.8971
logVARExo	-0.9275	-0.7808	-0.1548	0.8995	-1.4469	-1.4664	0.2574	-1.4168	-0.0704	0.5248
	0.3548	0.4359	0.8772	0.36953	0.1496	0.1442	0.7972	0.1582	0.9439	0.6004
logVAR	-1.4794	-0.2692	-1.1653	-0.2653	-1.0817	-0.9977	-0.2568	-1.5281	-1.0443	-0.0013
	0.1407	0.7881	0.2454	0.7911	0.2807	0.3197	0.7976	0.1281	0.2977	0.9989
DCC	-1.6587 ^c	-0.4084	-1.2824	-0.7154	-1.1367	-1.0726	-0.3963	-1.5777	-1.4083	-0.1616
	0.0988	0.6834	0.2013	0.4752	0.2571	0.2848	0.6923	0.1163	0.1607	0.8718
BEKK	-1.1591	-1.3747	-1.0448	-1.4778	-1.0248	-1.2045	-1.0486	-1.0324	-1.4071	-1.6459
	0.2478	0.1708	0.2975	0.1411	0.3068	0.2299	0.2957	0.3032	0.1610	0.1014
HEAVY	-1.1037	-0.3782	0.4989	3.0822 ^a	-1.8800 ^c	-1.2721	-0.7484	-0.1541	-2.6074 ^a	1.3949
	0.2711	0.7057	0.6184	0.0024	0.0616	0.20490	0.4551	0.8777	0.0098	0.1647
Pannello B: Giacomini-White Test										
	SP	SP-JPN	SP-UK	SP-DAX	JPN	JPN-UK	JPN-DAX	UK	UK-DAX	DAX
NL	0.1889	1.3121	1.9379 ^c	1.9594 ^c	2.8273 ^a	2.4709 ^b	0.7750	1.4206	-1.2765	0.6813
	0.8502	0.1895	0.0526	0.0501	0.0047	0.0135	0.4383	0.1554	0.2018	0.4957
CholVARExo	-2.7906 ^a	0.3147	0.6109	2.8914 ^a	-2.7144 ^a	-1.7170 ^c	0.1984	-2.4531 ^b	1.6701 ^c	1.4382
	0.0053	0.7530	0.5413	0.0038	0.0066	0.0860	0.8427	0.0142	0.0949	0.1504
CholVAR	-3.1488 ^a	0.7758	0.7111	2.7639 ^a	-2.5706 ^b	-1.2648	0.3747	-2.1666 ^b	0.9422	1.0577
	0.0016	0.4379	0.4770	0.0057	0.0102	0.2060	0.7079	0.0303	0.3461	0.2902
VARExo	-5.1126 ^a	-0.8138	-0.7754	1.4523	-4.2893 ^a	-3.4366 ^a	-1.3479	-4.2672 ^a	0.5615	0.8928
	0.0000	0.4158	0.4381	0.1464	0.0000	0.0006	0.1777	0.0000	0.5744	0.3720
VAR	-4.2307 ^a	-0.4291	-0.1080	1.7424 ^c	-4.1577 ^a	-2.9317 ^a	-1.2443	-3.9537 ^a	0.0534	0.3966
	0.0000	0.6679	0.9140	0.0814	0.0000	0.0034	0.2134	0.0001	0.9574	0.6916
logVARExo	-2.7164 ^a	-0.1582	0.0823	2.0039 ^b	-2.1870 ^b	-1.2873	0.4587	-2.7023 ^a	1.1048	1.0545
	0.0066	0.8743	0.9344	0.0451	0.0287	0.1980	0.6465	0.0069	0.2693	0.2916
logVAR	-3.1001 ^a	0.8481	-0.1558	2.0446 ^b	-1.7935 ^c	-1.2134	0.4923	-2.8646 ^a	-0.0666	0.8813
	0.0019	0.3964	0.8762	0.0409	0.0729	0.2250	0.6225	0.0042	0.9469	0.3781
DCC	-3.6420 ^a	-1.8428 ^c	-1.5014	1.1016	-2.9825 ^a	-2.7319 ^a	-0.5623	-4.5934 ^a	-0.9664	0.2290
	0.0003	0.0654	0.1333	0.2706	0.0029	0.0063	0.5739	0.0000	0.3338	0.8189
BEKK	-1.3905	-1.6887 ^c	-1.2293	-1.7413 ^c	-1.3337	-1.6375	-1.4360	-1.2526	-1.7763 ^c	-2.1117 ^b
	0.1644	0.0913	0.2190	0.0816	0.1823	0.1015	0.1510	0.2103	0.0757	0.0347
HEAVY	-3.2171 ^a	1.4857	2.6683 ^a	5.1725 ^a	-7.7443 ^a	-3.5749 ^a	-1.3517	-0.2824	-4.7800 ^a	2.1762 ^b
	0.0013	0.1374	0.0076	0.0000	0.0000	0.0004	0.1765	0.7776	0.0000	0.0295

^a, ^b e ^c indicano i livelli di significatività rispettivamente al 1%, 5% e 10%

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.28: Test in ambito multivariato (modello non lineare), dataset cross-country

	Diebold-Mariano		Giacomini-White	
	Frobenius	Euclidea	Frobenius	Euclidea
CholVARExo	-0.3273	-0.7009	-3.3484 ^a	-2.7910 ^a
	0.7438	0.4842	0.0009	0.0058
CholVAR	-0.1619	-0.5685	-0.9512	-1.5837
	0.8715	0.5704	0.3427	0.1149
VARExo	-1.7425 ^c	-2.0735 ^b	-3.2832 ^a	-4.2970 ^a
	0.0830	0.0395	0.0012	0.0000
VAR	-1.5575	-1.9153 ^c	-2.8401 ^a	-3.8852 ^a
	0.1210	0.0569	0.0050	0.0001
logVARExo	-0.5332	-0.8382	-2.3101 ^b	-2.6756 ^a
	0.5945	0.4029	0.0219	0.0081
logVAR	-0.3905	-0.7188	-1.5421	-1.5659
	0.6966	0.4732	0.1247	0.1190
DCC	-1.6916 ^c	-1.8810 ^c	-3.2952 ^a	-3.8874 ^a
	0.0924	0.0615	0.0012	0.0001
BEKK	-1.5858	-1.5957	-3.8078 ^a	-3.7854 ^a
	0.1144	0.1122	0.0000	0.0000
HEAVY	-3.9461 ^a	-4.1532 ^a	-14.5851 ^a	-16.3327 ^a
	0.0001	0.0001	0.0000	0.0000
NL	1.4305	1.4793	3.3269 ^a	3.3657 ^a
	0.1542	0.1407	0.0011	0.0009

a, b e c indicano i livelli di significatività rispettivamente al 1%, 5% e 10%

Nella tabella 2.29 sono riportate le statistiche test DM e GW per le previsioni del modello non lineare con esogene nel confronto con quelle dei restanti modelli nel dataset americano. Le statistiche del test DM, riportate nel pannello A della tabella 2.29, non risultano mai significative, fatta eccezione per il confronto con le previsioni dei modelli DCC, BEKK e HEAVY, sempre a favore del modello non lineare. Il pannello B della stessa tabella riporta i risultati del test GW. Dalla tabella risulta che le previsioni del modello NLExo siano migliori degli altri modelli solo per alcune serie delle covarianze.

La tabella 2.30 riporta le statistiche test DM e GW in ambito multivariato. I risultati divergono sensibilmente da quanto emerge per le singole serie. Si rifiuta spesso l'ipotesi nulla di uguale capacità previsionale a favore dei modelli confrontati. In particolare, è forte il rifiuto della nulla a favore del modello CholVARExo e del modello NL. Il modello benchmark non sembra prevedere meglio, in ambito multivariato, degli altri modelli, fatta esclusione dei modelli HEAVY e BEKK.

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.29: Valutazione previsioni non lineari out-of-sample mercato americano

Pannello A: Diebold-Mariano Test						
	SP	SP-10Y	SP-30d	10Y	10Y-30d	30d
NL	-1.4046	1.2077	0.5033	1.1162	1.2364	1.4708
	0.16308	0.2299	0.6158	0.2669	0.2191	0.1443
CholVARExo	-1.3860	-0.9878	-1.0684	1.0693	-0.2208	-0.9608
	0.1686	0.3255	0.2878	0.2874	0.8256	0.3388
CholVAR	-0.9561	-0.6545	-0.3110	1.1002	1.0375	-0.3744
	0.34122	0.5142	0.7564	0.2737	0.3019	0.7089
VARExo	-1.4091	-0.5729	-0.3842	0.6193	1.1363	-0.5395
	0.1617	0.5679	0.7016	0.5371	0.2584	0.5907
VAR	-1.0546	-0.5650	-0.3600	0.7902	1.0080	-0.5659
	0.2940	0.5732	0.7196	0.4311	0.3157	0.5727
logVARExo	-1.4960	-1.1043	-1.0704	0.3097	-1.0042	-1.0268
	0.1376	0.2719	0.2869	0.7574	0.3176	0.3068
logVAR	-1.0626	-0.2212	-0.3076	1.1020	1.1781	-0.3935
	0.2904	0.8254	0.7590	0.2729	0.2414	0.6947
DCC	-1.3840	-3.2845 ^a	-0.3018	1.0216	1.3916	-1.9508 ^c
	0.1693	0.0014	0.7634	0.3093	0.1669	0.0537
BEKK	-2.3008 ^b	-2.1123 ^b	-2.4744 ^b	-2.0657 ^b	-2.5046 ^b	-2.0949 ^b
	0.0214	0.0347	0.0133	0.0389	0.0123	0.0362
HEAVY	-1.8281 ^b	-17.4642 ^a	-23.6377 ^a	-6.0739 ^a	-5.5515 ^a	-6.9568 ^a
	0.0703	0.0000	0.0000	0.0000	0.0000	0.0000
Pannello B: Giacomini-White Test						
	SP	SP-10Y	SP-30d	10Y	10Y-30d	30d
NL	-1.8060 ^c	0.1095	0.8014	0.4303	-1.8829 ^c	0.4835
	0.0709	0.9128	0.4229	0.6669	0.0597	0.6287
CholVARExo	-0.2449	-0.8780	1.6405	1.8521 ^c	2.6384 ^a	5.1439 ^a
	0.8066	0.3799	0.1009	0.0640	0.0083	0.0000
CholVAR	-2.4998 ^b	0.2881	2.2219 ^b	1.9307 ^c	2.3584 ^b	3.3006 ^a
	0.0124	0.7733	0.0263	0.0535	0.0184	0.0010
VARExo	-1.7369 ^c	-0.2815	1.6354	-0.0969	1.4508	0.3630
	0.0824	0.7783	0.1020	0.9228	0.1468	0.7166
VAR	-2.4304 ^b	0.3663	1.9610 ^b	0.7581	1.2735	0.4126
	0.0151	0.7141	0.0499	0.4484	0.2028	0.6799
logVARExo	-1.2437	-1.2882	-0.9818	0.7942	-0.5675	-0.7061
	0.2136	0.1977	0.3262	0.4271	0.5703	0.4801
logVAR	-2.0713 ^b	0.3147	2.5085 ^b	2.1212 ^b	3.0572 ^a	3.6078 ^a
	0.0383	0.7530	0.0121	0.0339	0.0022	0.0003
DCC	-2.2440 ^b	-1.4245	3.0130 ^a	0.6164	1.7849 ^c	-1.3592
	0.0248	0.1543	0.0026	0.5376	0.0743	0.1741
BEKK	-1.9307 ^c	-1.5386	-1.8589 ^c	-1.5564	-1.8361 ^c	-1.4552
	0.0562	0.1269	0.0658	0.1226	0.0691	0.1486
HEAVY	-2.2255 ^b	-8.6329 ^a	-12.9426 ^a	-7.8976 ^a	-6.7583 ^a	-5.7526 ^a
	0.0260	0.0000	0.0000	0.0000	0.0000	0.0000

^a, ^b e ^c indicano i livelli di significatività rispettivamente al 1%, 5% e 10%

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.30: Test in ambito multivariato (modello non lineare), dataset mercato americano

	Diebold-Mariano		Giacomini-White	
	Frobenius	Euclidea	Frobenius	Euclidea
CholVARExo	2.1460 ^b	2.1118 ^b	4.8583 ^a	4.5950 ^a
	0.0342	0.0371	0.0000	0.0000
CholVAR	1.7798 ^c	1.7298 ^c	3.1202 ^a	2.9579 ^a
	0.0779	0.0866	0.0023	0.0038
VARExo	0.9673	0.9579	2.2569 ^b	2.2278 ^b
	0.3356	0.3403	0.0261	0.0280
VAR	1.9863 ^b	1.9248 ^c	3.9329 ^a	3.7211 ^a
	0.0496	0.0569	0.0002	0.0003
logVARExo	0.1303	0.3485	0.0256	0.1710
	0.8966	0.7282	0.9796	0.8646
logVAR	2.0184 ^b	1.9494 ^c	4.0402 ^b	3.8119 ^a
	0.0461	0.0539	0.0001	0.0002
DCC	1.1460	1.1118	1.8583 ^c	0.5950
	0.2544	0.2687	0.0659	0.5531
BEKK	-2.6298 ^a	-2.6093 ^b	-6.4555 ^a	-6.4641 ^a
	0.0098	0.0104	0.0000	0.0000
HEAVY	-6.2154 ^a	-5.8528 ^a	-15.4892 ^a	-15.9770 ^a
	0.0000	0.0000	0.0000	0.0000
NL	2.3472 ^b	2.3079 ^b	5.5815 ^a	5.3701 ^a
	0.0208	0.0229	0.0000	0.0000

a, b e c indicano i livelli di significatività rispettivamente al 1%, 5% e 10%

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Infine, viene analizzato il confronto tra tutti i modelli introdotti in questo lavoro mediante il Model Confidence Set. La tabella 2.31 riporta il MCS per il dataset cross-country, con un livello di confidenza pari a $\alpha = 0,1$ per le singole serie, quando la funzione di perdita è il MSE. I modelli scartati dal SSM risultano spesso essere il BEKK, il DCC e il modello HEAVY, mentre nell'ordinamento dei modelli facenti parte del SSM, le prime posizioni sono occupate principalmente dal modello NL, da quello NLExo e dal modello CholVARExo.

La tabella 2.32 riporta il MCS in ambito multivariato, quando le funzioni di perdita sono la distanza di Frobenius e quella Euclidea. Gli unici modelli non compresi nel SSM sono il modello BEKK e il modello HEAVY, mentre le prime quattro posizioni del SSM sono occupate dai modelli che fanno uso della scomposizione di Cholesky e sono rispettivamente i modelli non lineari, NL e NLExo, e i modelli lineari, CholVAR e CholVARExo.

Tabella 2.31: MCS cross-country univariato

Ranking	SP	SP-JPN	SP-UK	SP-DAX	JPN	JPN-UK	JPN-DAX	UK	UK-DAX	DAX
1	NL	NL	NL	HEAVY	NL	NL	logVARExo	NL	NLExo	HEAVY
2	NLExo	NLExo	HEAVY	NL	NLExo	NLExo	NL	NLExo	logVARExo	NL
3		CholVAR	NLExo	CholVARExo	logVAR	logVARExo	NLExo	HEAVY		CholVARExo
4		logVAR	logVARExo	logVARExo	DCC	CholVARExo	CholVARExo	logVARExo		logVARExo
5		DCC	CholVARExo	CholVAR	CholVAR	CholVAR	CholVAR	CholVARExo		CholVAR
6		HEAVY		VARExo	logVARExo	logVAR	logVAR	VARExo		VARExo
7		CholVARExo		NLExo	CholVARExo	DCC		CholVAR		NLExo
8		VAR		logVAR	VARExo	VARExo		DCC		logVAR
9		VARExo		VAR	VAR	VAR		logVAR		VAR
10				DCC	HEAVY			VAR		DCC
Modelli Scartati										
1	BEKK	BEKK	BEKK	BEKK	BEKK	BEKK	BEKK	BEKK	BEKK	BEKK
2	DCC	logVARExo				HEAVY	HEAVY		HEAVY	
3	VAR		logVAR				VARExo		DCC	
4	logVAR		VAR				VAR		VAR	
5	VARExo		VARExo				DCC		logVAR	
6	CholVAR		CholVAR						VARExo	
7	CholVARExo								CholVAR	
8	HEAVY								NL	
9	logVARExo								CholVARExo	

Tabella 2.32: MCS cross-country multivariato

Ranking	Funzione di perdita	
	Frobenius	Euclidea
1	NL	NL
2	NLE _{exo}	NLE _{exo}
3	CholVAR	CholVAR
4	CholVARE _{exo}	CholVARE _{exo}
5	logVAR	logVAR
6	logVARE _{exo}	logVARE _{exo}
7	DCC	DCC
8	VAR	VAR
9	VARE _{exo}	VARE _{exo}
Modelli Scartati		
1	BEKK	BEKK
2	HEAVY	HEAVY

La procedura MCS viene applicata anche al dataset americano, i risultati sono riportati in tabella 2.33. In maniera non sorprendente, i modelli maggiormente scartati dal SSM sono il modello BEKK e il modello HEAVY, insieme a uno dei modelli presentati in questo lavoro, il CholVARE_{exo}, al logVARE_{exo} e al DCC. Le prime posizioni dell'ordinamento imposto dalla procedura MCS sono occupate principalmente dai modelli non lineari, NL e NLE_{exo}.

In ambito multivariato, la procedura sottolinea la scarsa capacità previsionale dei modelli BEKK e HEAVY in un confronto multiplo, mentre il modello migliore appartenente al SSM risulta essere il CholVARE_{exo}.

Nonostante in letteratura sia emersa la scarsa capacità previsiva dei modelli non lineari in un'analisi out-of-sample, sembra che in questo contesto il modello non lineare riesca a prevedere meglio di quelli lineari, cogliendo le dinamiche non lineari che si instaurano tra la volatilità e le determinanti macroeconomiche e finanziarie, almeno per quanto concerne il dataset cross-country.

In generale, i modelli introdotti in questo capitolo sembrano produrre previsioni più accurate di quelle dei modelli esistenti e dei modelli che non fanno uso di variabili determinanti esogene. Un risultato positivo è evidenziato dai risultati ottenuti dai modelli che comportano la scomposizione di Cholesky della matrice delle covarianze realizzate, dando indicazione di quanta importanza assuma la scelta di trasformare la matrice delle covarianze. Per un'ulteriore analisi sulla capacità previsionale dei modelli, nell'ultimo capitolo si utilizzano le matrici delle covarianze previste in un'ottimizzazione di portafoglio e in ambito di gestione del rischio.

2. MODELLO DI PREVISIONE DELLA VOLATILITÀ CON VARIABILI ESOGENE

Tabella 2.33: MCS mercato americano univariato

Ranking	SP	SP-10Y	SP-30d	10Y	10Y-30d	30d
1	NLExo	NL	NL	NL	DCC	NL
2	NL	NLExo	NLExo	CholVARExo	logVAR	NLExo
3	CholVARExo	logVAR	VARExo	logVAR	VARExo	CholVAR
4	CholVAR	VAR	CholVAR	CholVAR	CholVAR	logVAR
5	VAR	CholVAR	DCC		VAR	VARExo
6	VARExo	VARExo	logVAR		NL	VAR
7	logVAR	DCC	VAR			CholVARExo
Modelli Scartati						
1	BEKK	BEKK	BEKK	BEKK	BEKK	BEKK
2	logVARExo	HEAVY	HEAVY	HEAVY	HEAVY	HEAVY
3	DCC	logVARExo	logVARExo	NLExo	logVARExo	logVARExo
4	HEAVY	CholVARExo	CholVARExo	logVARExo	CholVARExo	DCC
5				VARExo	NLExo	
6				VAR		
7				DCC		

Tabella 2.34: MCS mercato americano multivariato

Ranking	Funzione di perdita	
	Frobenius	Euclidea
1	NLExo	NLExo
2	CholVARExo	CholVARExo
3	NL	NL
4	CholVAR	CholVAR
5	VAR	VAR
6	logVAR	logVAR
Modelli Scartati		
1	BEKK	BEKK
2	HEAVY	HEAVY
3	DCC	DCC
4	logVARExo	logVARExo
5	VARExo	VARExo

3 Analisi empirica dei modelli

Il capitolo precedente presentava una valutazione statistica della precisione delle previsioni. Al fine di sottolineare l'utilità della procedura descritta nel secondo capitolo, si valutano i modelli presentati nel secondo capitolo con riferimento all'ottimizzazione di portafoglio e alla misurazione del rischio.

3.1 Ottimizzazione di portafoglio

Diversi studi valutano l'utilizzo della volatilità realizzata per l'ottimizzazione di portafoglio. Fleming *et al.* (2003), ad esempio, hanno analizzato le performance dell'utilizzo della matrice delle covarianze realizzate nel problema di ottimo di portafoglio, evidenziandone i vantaggi. Kyj *et al.* (2009), Halbleib-Chiriac e Voev (2011) e Hautsch *et al.* (2015), confermano l'utilità della matrice delle covarianze realizzate per le decisioni di investimento, in particolare quando si utilizza un portafoglio *Global Minimum Variance* (GMV).

I pesi di portafoglio ottimali per un GMV si ottengono considerando unicamente la matrice delle covarianze, evitando di specificare un modello per il valore atteso dei rendimenti dei titoli.

L'obiettivo della *portfolio theory*, introdotta da Markowitz (1952) con il modello media-varianza, è quello di massimizzare il rendimento di un portafoglio, dato un livello di rischio, o di minimizzare il rischio, dato un livello di rendimento atteso. Poiché alla base della teoria c'è il trade-off tra rischio e rendimento, si rende necessaria un'adeguata scelta della misura di rischio. In questo contesto si inseriscono le previsioni derivanti dalla procedura presente nel secondo capitolo, in sostituzione delle misure di volatilità costanti o derivanti da un modello di eteroschedasticità condizionale.

In questo lavoro vengono confrontate le performance dei portafogli Global Minimum Variance (GMV) e Maximum Sharpe Ratio, utilizzando le previsioni di volatilità derivanti dai modelli presentati nel secondo capitolo.

3.1.1 Valutazione delle performance di portafoglio

I portafogli valutati in questa sezione derivano da un'ottimizzazione GMV, che consente di ovviare al problema della prevedibilità della media condizionale dei rendimenti, e da una massimizzazione dello Sharpe Ratio. I pesi dei portafogli sono ribilanciati²¹ mensilmente

²¹Spinu (2015) evidenzia i vantaggi della strategia di un portafoglio costantemente ribilanciato rispetto ad un approccio *buy and hold*.

sulla base delle previsioni delle covarianze ottenute secondo una rolling window, come nella sezione 2.4.

La valutazione delle performance di un portafoglio si basa sull'osservazione di due elementi, il rendimento conseguito dal portafoglio e il livello di rischio che ha caratterizzato il portafoglio gestito. Per questo vengono utilizzati gli indicatori di performance *risk-adjusted*.

La più comune misura di rendimento aggiustato per il rischio è lo *Sharpe Ratio* (SR), definito come il rapporto tra il rendimento in eccesso del portafoglio rispetto al tasso privo di rischio e la deviazione standard del portafoglio e dato da

$$SR = \frac{\mu_p - r_f}{\sigma_p} \quad (3.1.1)$$

dove r_f è il tasso di un'attività priva di rischio, $\mu_p - r_f$ rappresenta il tasso di rendimento extra del portafoglio rispetto al tasso privo di rischio e σ_p è la deviazione standard del rendimento di portafoglio. Uno Sharpe ratio maggiore implica una performance migliore di portafoglio.

In letteratura sono stati presentati diversi test per fare inferenza sugli Sharpe ratio. Tradizionalmente è stato utilizzato il test introdotto da J. D. Jobson (1981) e Memmel (2003), tuttavia il test è applicabile solo in caso di rendimento extra distribuito in maniera normale. Il complesso test utilizzato da Ledoit e Wolf (2008) (LW) consente di tenere conto della non-normalità dei rendimenti e della presenza di autocorrelazione e di eteroschedasticità. Il test parte dall'assunzione che esistano due vettori di rendimenti da confrontare e che la loro distribuzione sia una normale bivariata. Siano μ_1 e μ_2 le medie dei due vettori di rendimenti e sia Σ la matrice delle covarianze, il test LW si basa sull'ipotesi nulla che la differenza tra gli Sharpe ratio dei due modelli sia nulla, per cui

$$H_0 : \Delta = SH_1 - SH_2 = \frac{\mu_1}{\sigma_1} - \frac{\mu_2}{\sigma_2} = 0 \quad (3.1.2)$$

poiché i veri momenti dei rendimenti non sono osservabili vengono usati i momenti campionari, per cui la (3.1.2) diventa

$$H_0 : \hat{\Delta} = \hat{S}H_1 - \hat{S}H_2 = \frac{\hat{\mu}_1}{\hat{\sigma}_1} - \frac{\hat{\mu}_2}{\hat{\sigma}_2} = 0 \quad (3.1.3)$$

Per ottenere i valori di $\hat{\Delta}$, Ledoit e Wolf (2008) specificano un particolare metodo bootstrap, mentre la deviazione standard $s(\hat{\Delta})$ viene calcolata utilizzando una matrice di tipo HAC (si veda Andrews (1991)), robusta sia alla presenza di eteroschedasticità che di autocorrelazione. Nello specifico Ledoit e Wolf (2008) suggeriscono di utilizzare il kernel spettrale quadratico di Andrews e Monahan (1992).

L'indice di Treynor, poi, viene calcolato in maniera simile allo Sharpe ratio, ma prevede l'utilizzo dell'indice di rischio sistematico, β_p , al posto dell'indice di rischio generale, σ_p , per cui

$$T = \frac{\mu_p - r_f}{\beta_p} \quad (3.1.4)$$

dove β_p è il beta del portafoglio di investimento, equivalente al rapporto tra la covarianza tra il rendimento del portafoglio con quello del mercato e la varianza del rendimento di mercato (i.e. $\beta_p = \frac{\sigma_{pm}}{\sigma_m^2}$). L'indice di Treynor presuppone che il rischio specifico dei titoli sia stato eliminato mediante una buona diversificazione e deriva direttamente dal *Capital Asset Pricing Model* (CAPM), definito come

$$E(r_p) = r_f + \beta_p [E(r_m) - r_f]. \quad (3.1.5)$$

Arrangiando i termini si ha

$$E(r_m) - r_f = \frac{E(r_p) - r_f}{\beta_p}, \quad (3.1.6)$$

il termine a destra è l'indice di Treynor del portafoglio e quello a sinistra può essere inteso come l'indice di Treynor per il mercato.

Negli studi empirici le condizioni di equilibrio dettate dalla (3.1.5) sono rimpiazzate da un'equazione per i rendimenti ex post, data da

$$r_{pt} = \alpha + r_{ft} + \beta(r_{mt} - r_{ft}) + \varepsilon_t \quad (3.1.7)$$

dove ε è un termine di errore white noise e l'intercetta α è detta Alpha di Jensen (si veda Jensen (1968)). Dalla (3.1.7) l'Alpha di Jensen è stimato come

$$\alpha = E(r_p) - [r_f + \beta_p(r_m - r_f)]. \quad (3.1.8)$$

Se si considerano due portafogli con il medesimo benchmark, un Alpha di Jensen maggiore indica una migliore performance di portafoglio.

Nel tempo, nonostante alcuni svantaggi legati all'uso della sola misura di rischio sistematico, l'Alpha di Jensen è diventata la misura di performance maggiormente utilizzata negli studi accademici (cfr. Aragon e Ferson (2007)).

3.1.2 Risultati

In questa sezione vengono presentati i risultati dei modelli utilizzati nel capitolo precedente. Il portafoglio *Equal* è ottenuto mediante una strategia *naïve* con pesi pari a $1/m$, dove m

è il numero di titoli di portafoglio. Per la costruzione dello Sharpe Ratio, dell'indice di Treynor e dell'alpha di Jensen si considera un tasso r_f pari a zero²². Mentre il benchmark di mercato, r_m , è l'indice MSCI World.

Cross-country La tabella 3.1 riporta alcune statistiche descrittive dei portafogli considerati insieme agli indicatori di performance descritti nella sezione precedente e al Value-at-Risk (VaR), condizionale e non, al 95% e al 99%.

I pesi dei portafogli nel pannello A sono ottenuti mediante un'ottimizzazione GMV non vincolata, mentre nel pannello B i pesi sono ottenuti mediante una massimizzazione dello Sharpe Ratio. Il portafoglio BH-Chol segue una strategia buy-and-hold (i pesi sono stimati solo una volta e vengono tenuti per l'intero campione), gli altri portafogli assumono un ribilanciamento mensile. Il campione di riferimento combacia con la porzione di campione out-of-sample del capitolo 2 e va da dicembre 1999 a ottobre 2015. Dalla tabella emergono diversi risultati. In primo luogo, è evidente che i portafogli ottenuti mediante una strategia buy-and-hold, l'Equal e il BH-Chol, abbiano rendimenti medi superiori ai portafogli ottenuti da una ottimizzazione GMV. Questo risultato è in linea con quanto emerge dal lavoro di DeMiguel *et al.* (2009), che evidenzia come le strategie di ottimizzazione di portafoglio basate su previsioni della matrice delle covarianze non possono avere performance consistentemente superiori a quelle di una strategia naïve. L'unico portafoglio che registra rendimenti medi superiori ai portafogli con pesi non ribilanciati è il portafoglio benchmark, NLExo. Il portafoglio CholVARExo, ottenuto da un'ottimizzazione GMV con l'uso della matrice delle covarianze derivante dalla procedura introdotto in questo lavoro (denominato da qui in avanti come "portafoglio benchmark" e contrassegnato da un asterisco), presenta rendimenti medi di poco inferiori ai due con pesi costanti. I portafogli ottenuti mediante la massimizzazione dello SR presentano rendimenti medi sensibilmente superiori a quelli ottenuti mediante GMV. La tecnica Cholesky-VAR con esogene fornisce le migliori performance assolute in termini di deviazione standard di portafoglio e di VaR, condizionale e non. In particolare, è evidente il vantaggio di una stima della volatilità del tipo Cholesky-VAR nei confronti di un modello più classico come il DCC. Questo conferma la bontà, anche in un'ottica economica, della procedura introdotta nel secondo capitolo. La deviazione standard del portafoglio NLExo, invece, è in linea con i risultati degli altri portafogli. Le misure di rischio dei portafogli del pannello B, in maniera non sorpren-

²²È noto il problema dello *zero lower bound* per i tassi di interesse nominali dei titoli governativi (si veda Swanson e Williams (2013)).

dente, sono sensibilmente più elevate di quelle ottenute mediante la GMV. In particolare, si può notare che i portafogli derivanti dai modelli introdotti in questo lavoro, registrano performance in linea con gli altri modelli quando l'obiettivo dell'ottimizzazione non è la minimizzazione della varianza di portafoglio.

Gli indicatori di performance evidenziano le migliori performance dei portafogli benchmark. In particolare, il portafoglio NLExo presenta il maggiore Sharpe Ratio in assoluto e il maggiore indice di Treynor. Gli indicatori di performance dei portafogli nel pannello B sono mediamente superiori a quelli dei portafogli nel pannello A, visto il differente problema di ottimo. Infine, il modello HEAVY genera portafogli con deviazione standard e VaR maggiori in termini assoluti, mentre gli altri portafogli presentano performance differenti a seconda dell'indice considerato.

Nell'ultima colonna della tabella 3.1 vengono presentati i p -value del test di ipotesi sull'uguaglianza degli Sharpe Ratio, introdotto nel paragrafo 3.1.1, tra la strategia $1/m$ e i portafogli ottimizzati. Un p -value inferiore al 10% evidenzia una differenza tra due Sharpe Ratio, viceversa per un p -value elevato. Gli unici portafogli che mostrano una differenza statistica, in termini di Sharpe Ratio, rispetto alla strategia naïve sono i portafogli con media dei rendimenti negativa.

Tabella 3.1: Performance Portafogli cross-country

	$\mu_p(\%)$	$\sigma_p(\%)$	SR	Treynor	alpha	VaR(%)		CVaR(%)		Ledoit-Wolf
						95%	99%	95%	99%	
Equal	0.1522	4.6233	0.032912	0.001629	0.000252	7.756	10.908	9.689	12.474	-
Pannello A: GMV										
BH-Chol	0.1536	4.5240	0.033947	0.001659	0.000277	7.595	10.678	9.485	12.211	0.8726
CholVARExo*	0.1128	3.8694	0.029148	0.001489	0.000098	6.477	9.114	8.094	10.426	0.8835
CholVAR	-0.0492	4.0224	-0.012239	-0.000625	-0.001562	6.567	9.308	8.248	10.671	0.0445
VARExo	0.1005	5.3289	0.018872	0.001049	-0.000297	8.866	12.498	11.093	14.303	0.7358
VAR	0.0198	4.1119	0.004821	0.000245	-0.000899	6.783	9.586	8.502	10.979	0.4954
logVARExo	0.0769	3.9811	0.019317	0.000980	-0.000297	6.625	9.338	8.289	10.687	0.6554
logVAR	-0.0399	3.9843	-0.010026	-0.000506	-0.001472	6.513	9.229	8.179	10.579	0.0637
DCC	0.0749	4.4329	0.016905	0.000826	-0.000484	7.366	10.388	9.218	11.889	0.3043
BEKK	0.0339	4.8279	0.007017	0.000380	-0.000873	7.974	11.264	9.991	12.899	0.8603
HEAVY	-2.8064	38.7921	-0.072359	0.165477	-0.027834	45.008	78.437	57.210	82.583	0.1871
NLExo*	0.1953	4.1011	0.047614	0.002406	0.000849	6.941	9.736	8.655	11.126	0.4220
NL	0.0898	4.1664	0.046867	0.001069	-0.000243	6.943	9.782	8.684	11.194	0.4516
Pannello B: Maximum Sharpe Ratio										
CholVARExo*	0.0819	5.7097	0.014335	0.000797	-0.000578	9.474	13.365	11.859	15.299	0.4782
CholVAR	0.1178	5.6738	0.020754	0.001158	-0.000205	9.450	13.317	11.821	15.239	0.6380
VARExo	0.1656	5.6603	0.029253	0.001603	0.000251	9.476	13.333	11.841	15.252	0.8815
VAR	0.2587	5.5235	0.046840	0.002655	0.001262	9.344	13.108	11.652	14.980	0.6069
logVARExo	0.1496	5.6058	0.026690	0.001486	0.000127	9.370	13.190	11.713	15.090	0.7996
logVAR	0.1422	5.6278	0.025270	0.001399	0.000041	9.399	13.234	11.751	15.142	0.7581
DCC	0.2140	5.3795	0.039787	0.002232	0.000836	9.062	12.729	11.310	14.552	0.7953
BEKK	0.1863	5.8201	0.035691	0.001954	0.000567	9.073	13.030	11.254	14.799	0.8537
HEAVY	0.0733	5.6133	0.013052	0.000717	-0.000656	9.306	13.132	11.652	15.034	0.4083
NLExo*	0.0826	5.5956	0.014765	0.000822	-0.000539	9.287	13.099	11.625	14.996	0.5192
NL	0.0735	5.7148	0.014456	0.000719	-0.000654	9.474	13.368	11.862	15.305	0.4616

* indica i portafogli benchmark

Mercato americano La tabella 3.2 riporta le statistiche dei portafogli nel dataset americano. I portafogli nel pannello A sono ottenuti mediante un GMV vincolato, escludendo la possibilità di vendite allo scoperto e limitando i singoli pesi a 0.5, mentre nel pannello B sono riportate le statistiche descrittive dei portafogli ottenuti con una massimizzazione dello Sharpe Ratio, limitando i singoli pesi a 0.7. La scelta di vincolare i pesi dei portafogli è dovuta alla natura dei dati, in particolare alla differenza, in termini di rischiosità, tra i 3 titoli considerati (e.g. si vedano le medie delle varianze realizzate mensili nella tabella 2.5).

Il portafoglio con diversificazione *naïve* presenta performance in termini di media dei rendimenti, indice di Treynor e alpha di Jensen superiori a tutti i portafogli del pannello A, a conferma di quanto sottolineato da DeMiguel *et al.* (2009). In termini assoluti, la media dei rendimenti più alta si registra col portafoglio DCC derivante da una massimizzazione dello Sharpe Ratio, mentre dal punto di vista del rischio, la migliore performance è registrata dal portafoglio che usa le previsioni del modello VAR sulle varianze e covarianze realizzate dal pannello A. Lo Sharpe Ratio maggiore si ha, invece, nel portafoglio dello stesso modello ma con l'uso di esogene, VARExo. I portafogli benchmark presentano performance in linea con gli altri portafogli, ma non i risultati non riescono a confermare l'utilità dell'approccio Cholesky-VAR nelle decisioni di investimento nella scelta tra azioni e obbligazioni americane. Tuttavia, gli indicatori di performance variano leggermente tra i vari portafogli e non risultano decisivi nella scelta del modello appropriato per la volatilità nel multivariato. I portafogli del pannello B presentano indici di rischio decisamente superiori a quelli del pannello A, tale differenza è dovuta alla diversa funzione obiettivo del problema di ottimo di portafoglio.

L'ultima colonna della tabella 3.2 riporta i p -value del test di Ledoit e Wolf (2008). Il test non evidenzia differenze sostanziali, nel pannello A, tra il portafoglio naïve e quelli ottimizzati, dal punto di vista dello Sharpe Ratio. Nel pannello B la differenza con lo Sharpe Ratio del portafoglio naïve, misurate dal p -value del test Ledoit-Wolf, sono significative per diversi modelli.

In generale, la valutazione indiretta delle previsioni evidenzia l'abilità dei modelli introdotti in questo lavoro di generare portafogli migliori di quelli esistenti in letteratura, in termini di rischio-rendimento. Quando le attività che compongono il portafoglio presentano differenze sostanziali nei momenti delle distribuzioni dei rendimenti, i modelli introdotti nel corso di questo lavoro performano in maniera simile ai modelli esistenti in letteratura.

Tabella 3.2: Performance Portafogli mercato americano

	$\mu_p(\%)$	$\sigma_p(\%)$	SR	Treynor	alpha	VaR(%)		CVaR(%)		Ledoit-Wolf
						95%	99%	95%	99%	
Equal	0.2087	1.4942	0.139688	0.095422	0.002069	2.666	3.685	3.291	4.191	-
Pannello A: GMV										
BH-Chol	0.0813	0.5810	0.139934	-0.057223	0.000824	1.037	1.433	1.279	1.629	0.9986
CholVARExo*	0.1015	0.5240	0.193737	-0.156995	0.001020	0.963	1.321	1.182	1.499	0.5753
CholVAR	0.0922	0.5232	0.176309	-0.115191	0.000929	0.953	1.309	1.171	1.487	0.7214
VARExo	0.1930	0.7135	0.270489	-1.061091	0.001931	1.367	1.853	1.665	2.095	0.1367
VAR	0.1060	0.5041	0.210316	-0.161773	0.001065	0.935	1.279	1.146	1.449	0.4722
logVARExo	0.0885	0.5213	0.169831	-0.230028	0.000888	0.946	1.303	1.164	1.478	0.7467
logVAR	0.0888	0.5201	0.170897	-0.113521	0.000895	0.944	1.299	1.162	1.475	0.7628
DCC	0.0941	0.5108	0.184277	-0.113721	0.000948	0.934	1.283	1.148	1.456	0.6404
BEKK	0.0545	0.7009	0.077787	-0.031733	0.000559	1.208	1.685	1.500	1.923	0.5815
HEAVY	0.1861	1.4447	0.128763	0.093960	0.001845	2.562	3.547	3.166	4.036	0.1266
NLExo*	0.0813	0.5388	0.150896	-0.106991	0.000819	0.968	1.335	1.193	1.517	0.9024
NL	0.0964	0.5224	0.184510	-0.140668	0.000969	0.956	1.312	1.174	1.489	0.5928
Pannello B: Maximum Sharpe Ratio										
CholVARExo*	-0.0110	3.1215	-0.00353	-0.001487	-0.000169	5.123	7.251	6.428	8.308	0.0013
CholVAR	0.2266	3.1217	0.07259	0.025726	0.002197	5.361	7.489	6.666	8.547	0.1177
VARExo	-0.1728	2.9417	-0.05875	-0.017039	-0.001808	4.666	6.671	5.895	7.668	0.0014
VAR	0.0616	3.1029	0.01985	0.008994	0.000562	5.165	7.279	6.462	8.331	0.0055
logVARExo	-0.0859	3.0519	-0.02816	-0.011136	-0.000920	4.934	7.014	6.209	8.048	0.0007
logVAR	0.1682	3.0880	0.05445	0.018049	0.001608	5.247	7.352	6.538	8.398	0.0593
DCC	0.2515	2.4151	0.10415	0.108240	0.002497	4.224	5.870	5.233	6.688	0.6938
BEKK	0.2581	2.4211	0.10659	0.322751	0.002575	4.241	5.891	5.252	6.711	0.6874
HEAVY	0.1568	1.2190	0.12860	0.049772	0.001543	2.162	2.993	2.671	3.406	0.7465
NLExo*	0.0254	2.9876	0.00849	0.003763	0.000201	4.939	6.976	6.188	7.988	0.0216
NL	0.2283	2.8796	0.00882	0.023709	0.002207	4.965	6.927	6.168	7.903	0.3511

* indica il portafoglio benchmark

3.2 Gestione del rischio

Un possibile criterio per valutare le previsioni di volatilità consiste nel confronto delle previsioni del Value-at-Risk (VaR).

Il VaR e le misure a questo collegate, come il Conditional-Value-at-Risk (CVaR), sono uno standard operativo nell'ambito delle istituzioni finanziarie (banche, fondi di investimento, etc.).

Queste misure sono basate sui quantili della distribuzione, in particolare sulle perdite estreme. Il vantaggio di queste misure è che esse richiedono poche informazioni per essere determinate, inoltre fanno uso di metodi statistici con assunzioni minime per ottenere una stima del contributo del singolo titolo al sistema. La loro stima è possibile mediante modelli econometrici come quelli presentati nel corso del secondo capitolo.

Il VaR è definito come la massima perdita di un portafoglio ad un dato livello di probabilità, α , in un intervallo di tempo definito, ovvero

$$P(r_t \leq -VaR_\alpha) = \alpha \quad (3.2.1)$$

dove r_t è il rendimento del titolo al tempo t . Il VaR equivale, dunque, alla massima perdita del titolo nell'intervallo di confidenza α .

Esistono diversi metodi per prevedere il VaR in letteratura, come il metodo media-varianza, il metodo delle simulazioni storiche, quello delle simulazioni di Monte Carlo e l'Extreme Value Theory. Kuester *et al.* (2006) contiene una raccolta e un confronto dei diversi metodi.

Il metodo utilizzato in questo lavoro si basa sul contributo di Giot e Laurent (2004) nella previsione del VaR mediante la volatilità realizzata. Siano r_t i rendimenti mensili al tempo t del singolo titolo, si assume che

$$r_t = \sqrt{h_t}z_t, \quad z_t \sim F \quad (3.2.2)$$

dove h_t è la varianza condizionale mensile, z_t è una variabile i.i.d a varianza unitaria e F è la funzione di ripartizione dei rendimenti. La previsione ad un passo del VaR è data da

$$\hat{VaR}_{t+1|t}^\alpha = -\hat{F}^{-1}\sqrt{h_t}$$

assumendo che h_t sia nota, condizionatamente all'informazione al tempo t . Ad esempio, utilizzando un approccio GARCH, si può prevedere la varianza condizionale ad un passo e, assumendo una certa distribuzione per F , si può definire un quantile della distribuzione e, quindi, il VaR.

Quando si utilizza una proxy della volatilità, si può prevedere la varianza condizionale al tempo t usando una appropriata specificazione per le dinamiche della volatilità, $rv_{t+1|t}$. Si assume allora che la varianza condizionale del rendimento sia funzione di $rv_{t+1|t}$ e di un vettore di parametri non noti ϕ , per cui

$$h_t = f\left(rv_{t+1|t} \mid \phi\right) \quad (3.2.3)$$

dove f è una qualche funzione di $rv_{t+1|t}$.

Il VaR presenta, tuttavia, alcuni problemi. Ai fini dell'ottimizzazione di portafoglio, la più rilevante è quella della non-additività. Tra le proprietà che una misura di rischio deve avere, presentate da Artzner *et al.* (1999), quella di sub-additività implica che la diversificazione riduca il rischio²³. Per due portafogli, X e Y , invece il VaR ottenuto dalla combinazione dei due può essere maggiore della somma dei VaR dei singoli portafogli, $VaR(X+Y) > VaR(X) + VaR(Y)$. In aggiunta, il VaR non dà alcuna informazione sulle perdite che eccedono la perdita massima ad un dato livello di confidenza, per questo si è diffusa una nuova misura di rischio denominata Conditional VaR o Expected Shortfall.

Adrian e Brunnermeir (2011) hanno proposto il CVaR, che misura la perdita attesa di un titolo condizionata al fatto di trovarsi sulla coda della distribuzione a sinistra del VaR. Sia L una variabile continua rappresentante le perdite di un titolo, il CVaR è definito come

$$CVaR_\alpha = E[L \mid L \geq VaR_\alpha]. \quad (3.2.4)$$

In altri termini, il CVaR equivale alla media delle perdite che si trovano nella coda a destra del VaR (si veda la figura 3.1).

Se i rendimenti sono distribuiti in maniera normale, si ha

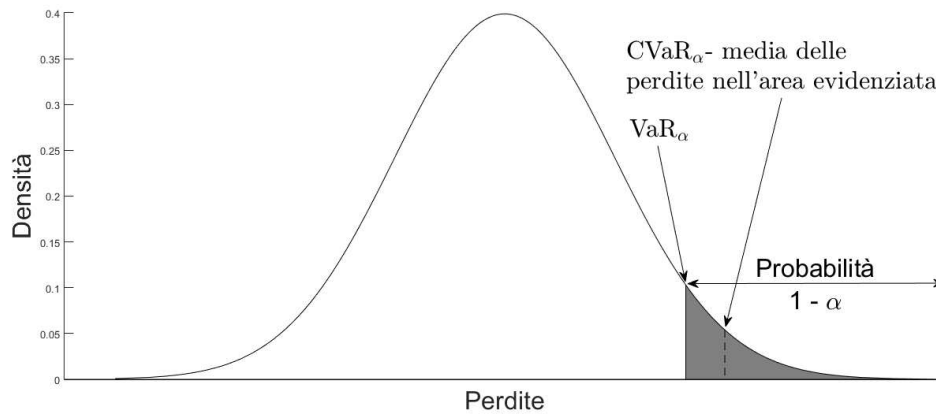
$$CVaR_\alpha = -\left(\mu - \sigma \frac{\Phi(N^{-1}(\alpha))}{\alpha}\right), \quad (3.2.5)$$

dove $\Phi(N^{-1}(\alpha))$ è la funzione di ripartizione del quantile α . Se si ipotizza che la media dei rendimenti sia nulla, la formula si riduce a

$$CVaR_\alpha = \sigma \frac{\Phi(N^{-1}(\alpha))}{\alpha}. \quad (3.2.6)$$

Se la distribuzione dei rendimenti è normale, per $\alpha \rightarrow 0$, allora $VaR_\alpha \rightarrow CVaR_\alpha$ (cfr. Barr (2013)).

²³Quando vengono aggregati due portafogli, la misura di rischio del portafoglio aggregato deve essere minore o uguale della somma delle singole misure di rischio per i singoli portafogli.

Figura 3.1: Conditional Value at Risk

Al contrario del VaR, il CVaR è una misura di rischio coerente secondo le definizioni di Artzner *et al.* (1999). La proprietà di sub-additività permette alla misura di rischio di cogliere al meglio gli effetti della diversificazione di portafoglio.

In questo lavoro si confrontano sia i VaR che i CVaR previsti mediante i diversi approcci.

3.2.1 Valutazione delle previsioni di VaR e CVaR

Per valutare le previsioni di VaR e CVaR derivanti da un modello, sono necessari dei metodi di valutazione ex-post delle previsioni. Le stime di VaR e CVaR presentate in questa tesi sono valutate usando due test: il test non condizionale di Kupiec (1995) e quello condizionale di Christoffersen (1998). Al fine di determinare i relativi test, si considerano i rendimenti mensili delle attività finanziarie e si considerano i casi in cui i rendimenti mensili abbiano ecceduto il VaR stimato o meno. La variabile indice è calcolata come segue

$$\eta_t = \begin{cases} 1 & \text{se } r_t < -VaR \\ 0 & \text{se } r_t \geq -VaR, \end{cases} \quad (3.2.7)$$

dove 1 indica un'eccezione e 0 indica un rendimento minore del VaR (o del CVaR per quanto riguarda la misura condizionale). Le eccezioni sono poi sommate e divise per il numero totale di stime out-of-sample del VaR per ottenere una misura empirica.

Il test dell'unconditional coverage è stato introdotto da Kupiec (1995) e ha come ipotesi nulla che la frequenza delle eccezioni empiricamente rilevate, ϕ , sia coerente con quella

teorica desiderata, α . La statistica test è la seguente

$$LR_{uc} = -2 \left[\ln(\alpha^x (1 - \alpha^{N-x})) - \ln(\phi^x (1 - \phi)^{N-x}) \right] \sim \chi^2(1), \quad (3.2.8)$$

dove x è il numero delle eccezioni, N è la dimensione del campione. Dunque, un rifiuto dell'ipotesi nulla implica che la dimensione del VaR empirico è significativamente differente dal VaR nominale.

Il test Kupiec (1995) è applicabile solo quando le eccezioni sono indipendenti tra loro, quando le eccezioni non sono indipendenti e si presentano a grappoli, si utilizza il test di Christoffersen (1998). Il test introdotto da Christoffersen (1998) è il metodo più utilizzato per valutare le performance dei modelli di VaR. Sia la probabilità di due eccezioni consecutive pari a

$$p_{ij} = P(\eta_t = 1 \mid \eta_{t-1} = j),$$

dove η è una funzione indice come nella (3.2.7). Due eccezioni sono definite indipendenti se non avvengono in due giorni consecutivi. Christoffersen (1998) suggerisce un test di verosimiglianza con ipotesi nulla che le osservazioni siano indipendenti nel tempo e come ipotesi alternativa che le osservazioni seguano una catena di Markov a due stadi. Si assuma che le violazioni e le non-violazioni seguano una catena di Markov con la seguente funzione di transizione

$$\Pi = \begin{pmatrix} \pi_{0,0} & \pi_{1,0} \\ \pi_{0,1} & \pi_{1,1} \end{pmatrix} \quad (3.2.9)$$

dove 1 rappresenta una violazione e 0 rappresenta una non-violazione. Siano $n_0, n_1, n_{00}, n_{01}, n_{10}, n_{11}$ gli stadi o le transizioni del processo stocastico di Markov. Allora,

$$\pi_{00} = \frac{n_{00}}{n_{00} + n_{01}}, \quad \pi_{01} = \frac{n_{01}}{n_{00} + n_{01}} \quad (3.2.10)$$

$$\pi_{10} = \frac{n_{10}}{n_{10} + n_{11}}, \quad \pi_{11} = \frac{n_{11}}{n_{10} + n_{11}}. \quad (3.2.11)$$

Sia inoltre $\pi_0 = n_0/N, \pi_1 = n_1/N$, allora il test LR può essere definito come segue

$$LR_{ind} = -2 \ln \left[(\pi_0^{n_0} \pi_1^{n_1}) - \ln(\pi_{00}^{n_{00}} \pi_{01}^{n_{01}} \pi_{10}^{n_{10}} \pi_{11}^{n_{11}}) \right] \sim \chi^2(1). \quad (3.2.12)$$

Per un livello di confidenza del 95%, se $LR_{ind} > LR_{critical} = 3.841$, l'ipotesi nulla viene rifiutata e le violazioni risultano non indipendenti. Dunque, il modello non passa il test di indipendenza.

Spesso, gli investitori non sono interessati nelle performance di un singolo modello per il VaR, ma anche in un confronto con altri modelli VaR. La letteratura ha prodotto diversi

modelli per classificare due o più modelli, come la *quadratic probability score function* (Lopez (1998)) e la *quadratic score function* (Blanco e Ihle (1999)). Entrambe le funzioni permettono di misurare le performance di un modello relativamente ad un altro.

Lopez (1998) ha introdotto la funzione *quadratic probability score* come misura della performance dei modelli per il VaR (e il CVaR) ed è definita come

$$QPS = \frac{2}{n} \sum_{t=1}^n (C_t - p)^2, \quad (3.2.13)$$

dove n è il numero di osservazioni, p è la probabilità attesa di una violazione (i.e. la perdita attuale è più grande del VaR/CVaR stimato). C_t è una funzione di perdita. Lopez (1998) utilizza una funzione di perdita definita come segue

$$C_t = \begin{cases} 1 & \text{se } L_t > VaR_t \\ 0 & \text{se } L_t \leq VaR_t. \end{cases} \quad (3.2.14)$$

La funzione QPS assume valori compresi tra 0 e 2. Sotto condizioni generali, stime accurate del VaR (CVaR) generano il minor score possibile, dunque un minore QPS indice una migliore performance in presenza di violazioni.

Accanto alla funzione QPS, un metodo diffuso per la valutazione le previsioni del VaR è il Root mean square error. Sia VaR_t , il VaR (CVaR) stimato e L_t la perdita reale, l'indicatore RMSE è definito come

$$RMSE = \sqrt{E[(VaR_t - L_t)]} = \sqrt{\frac{1}{n} \sum_{t=1}^n (VaR_t - L_t)^2}. \quad (3.2.15)$$

Un RMSE minore implica una maggiore capacità previsionale del modello.

3.2.2 Risultati empirici

Questa sezione presenta i risultati empirici dei modelli per il VaR e il CVaR. I modelli sono calcolati per il VaR e per il CVaR al 99% ($\alpha = 0.99$) e il periodo di mantenimento è un mese. Sono state utilizzate le previsioni rolling window di volatilità prodotte nel corso del secondo capitolo, al fine di tenere conto della flessibilità dei parametri della stima.

I risultati per ogni modello di VaR, nel dataset cross-country, sono presentati nella tabella 3.3 dove sono confrontati il test di Christoffersen, il *quadratic probability score* (QPS) e il RMSE. Quando la statistica test è significativa, il modello di VaR non ha superato il test di Christoffersen.

Il test di Christoffersen, LR_{uc} , evidenzia come tutti i modelli falliscano nello stimare in maniera adeguata le perdite future. In particolare, è forte il rifiuto dell'ipotesi nulla per i modelli che non ricorrono alla scomposizione di Cholesky (i.e. VAR, VARExo) e per il modello BEKK. Anche il QPS è maggiore per questi modelli. In generale, i modelli che esibiscono il minore QPS e il minor numero di violazioni effettive sono il modello logVAR, il modello logVARExo, i modelli non lineari e il modello HEAVY. Mentre i modelli lineari introdotti in questo lavoro sembrano avere performance peggiori di quelle degli altri modelli, in termini di previsione del VaR. L'indicatore RMSE, infine, evidenzia ancora una volta l'utilità della procedura Cholesky-VLSTAR, del modello logVAR e del modello HEAVY in termini previsionali, mentre i modelli DCC e BEKK mostrano le performance peggiori in termini di RMSE.

La tabella 3.4 riporta i risultati per il CVaR. I modelli che presentano il numero minore di eccezioni sono il logVARExo e il logVAR, il modello HEAVY e il modello non lineare senza esogene presentano un numero di eccezioni non troppo differente da questi modelli. Come nel caso del VaR, i modelli che presentano la peggiore performance, in termini di numero di eccezioni, sono il modello BEKK e il modello VARExo. Questi risultati sono confermati anche dal test di Christoffersen, che rifiuta l'ipotesi nulla per ogni modello, dal QPS e dal RMSE.

Tabella 3.3: Valutazione modelli VaR, cross-country

Modello VaR	Violazioni previste	Violazioni effettive	Christoffersen Test			Performance relativa	
			LR_{uc}	LR_{ind}	LR_{cc}	QPS	RMSE
CholVARExo	1	21	64.5072 ^a	1.9948	66.5019 ^a	0.2157	0.0905
CholVAR	1	20	59.5538 ^a	1.9948	61.5486 ^a	0.2054	0.0641
VARExo	1	30	113.4620 ^a	1.9948	115.4568 ^a	0.3081	0.0815
VAR	1	24	79.9930 ^a	1.9948	81.9878 ^a	0.2465	0.0804
logVARExo	1	14	32.3847 ^a	1.9948	34.3795 ^a	0.1439	0.1236
logVAR	1	11	20.7816^a	1.9948	22.7763^a	0.1131	0.0521
DCC	1	16	40.9126 ^a	1.9948	42.9073 ^a	0.1644	0.1215
BEKK	1	38	162.4704 ^a	1.9948	164.4652 ^a	0.3901	9.6113
HEAVY	1	13	28.3475 ^a	1.9948	30.3423 ^a	0.1336	0.0362
NLExo	1	16	40.9125 ^a	1.9948	42.9073 ^a	0.1644	0.0682
NL	1	14	32.3847 ^a	1.9948	34.3795 ^a	0.1439	0.0621

^a indica un livello di significatività rispettivamente all'1%

Tabella 3.4: Valutazione modelli CVaR, cross-country

Modello VaR	Violazioni previste	Violazioni effettive	Christoffersen Test			Performance relativa	
			LR_{uc}	LR_{ind}	LR_{cc}	QPS	RMSE
CholVARExo	1	17	74.7309 ^a	1.9948	76.7257 ^a	0.2362	0.0328
CholVAR	1	19	69.5677 ^a	1.9948	71.5624 ^a	0.2259	0.0367
VARExo	1	27	137.3926 ^a	1.9948	139.3874 ^a	0.3491	0.0374
VAR	1	21	96.3386 ^a	1.9948	98.3334 ^a	0.2773	0.0405
logVARExo	1	9	28.3475^a	1.9948	30.3423	0.1336	0.0308
logVAR	1	11	29.4123 ^a	1.9948	31.7542 ^a	0.1385	0.0260
DCC	1	13	69.5677 ^a	1.9948	71.5624 ^a	0.2259	0.0186
BEKK	1	49	236.6458 ^a	1.9948	238.6406 ^a	0.5030	16.4132
HEAVY	1	12	40.9125 ^a	1.9948	42.9073 ^a	0.1644	0.0365
NLExo	1	15	49.9874 ^a	1.9948	51.9821 ^a	0.1849	0.0419
NL	1	12	36.576 ^a	1.9948	38.5710 ^a	0.1541	0.0382

^a indica un livello di significatività rispettivamente all'1%

Nella tabella 3.5 sono riportati i risultati dei test sulle previsioni del VaR per il mercato americano. Come nel caso del dataset cross-country, il test di Christoffersen rifiuta per ogni modello l'ipotesi nulla e il numero di violazioni effettive si discosta di molto da quelle originariamente supposte dal modello VaR. Le peggiori performance, in termini di test di Christoffersen e numero di violazioni effettive, sono mostrate dal modello BEKK e dal modello VARExo. Il QPS conferma quanto emerge dal test di Christoffersen e evidenzia come i modelli che fanno uso della trasformazione logaritmica della matrice delle covarianze presentino performance migliori rispetto agli altri modelli. Infine, l'RMSE minore è associato al modello HEAVY. I modelli non lineari e i modelli logVARExo e logVAR mostrano un RMSE simile e migliore degli altri modelli.

I risultati nella tabella 3.6 sul CVaR confermano quanto emerge dalla tabella 3.5. I risultati confermano la bontà dei modelli sul logaritmo della matrice delle covarianze per quanto riguarda il test di Christoffersen e il numero di eccezioni. I modelli con le statistiche test più grandi sono il modello HEAVY e il modello CholVAR. Un indicatore RMSE minore è, invece, associato ai modelli non lineari e al modello CholVARExo. Il DCC e il modello HEAVY presentano un Root mean square error maggiore.

L'analisi sulle previsioni del VaR e del CVaR conferma l'utilità, in termini previsionali, dei modelli non lineari introdotti in questo lavoro, mentre risultano modeste le performance dei modelli CholVARExo e CholVAR. Le performance dei modelli classici della letteratura sulla volatilità nel multivariato si rivelano di nuovo le peggiori nell'insieme dei modelli considerati.

Tabella 3.5: Valutazione modelli VaR, US

Modello VaR	Violazioni previste	Violazioni effettive	Christoffersen Test			Performance relativa	
			LR_{uc}	LR_{ind}	LR_{cc}	QPS	RMSE
CholVARExo	1	32	166.3779 ^a	1.9905	168.3685 ^a	0.5919	0.0341
CholVAR	1	27	129.9660 ^a	1.9905	131.9565 ^a	0.4994	0.0347
VARExo	1	40	229.2359 ^a	1.9905	231.2265 ^a	0.7398	0.0322
VAR	1	28	137.0583 ^a	1.9905	139.0489 ^a	0.5179	0.0361
logVARExo	1	32	166.3779 ^a	1.9905	168.3685 ^a	0.5919	0.0411
logVAR	1	24	109.2959^a	1.9905	111.2864^a	0.4439	0.02744
DCC	1	31	158.9094 ^a	1.9905	160.8999 ^a	0.5734	0.01440
BEKK	1	34	181.5822 ^a	1.9905	183.5727 ^a	0.6289	0.3383
HEAVY	1	52	333.1137 ^a	1.9905	335.1042 ^a	0.9618	0.0118
NLExo	1	30	151.5321 ^a	1.9905	153.5226 ^a	0.5549	0.0277
NL	1	28	137.0583 ^a	1.9905	139.0489 ^a	0.5179	0.0239

^a indica un livello di significatività rispettivamente all'1%

Tabella 3.6: Valutazione modelli CVaR, US

Modello VaR	Violazioni previste	Violazioni effettive	Christoffersen Test			Performance relativa	
			LR_{uc}	LR_{ind}	LR_{cc}	QPS	RMSE
CholVARExo	1	16	158.9094 ^a	1.9905	160.8999 ^a	0.5734	0.0074
CholVAR	1	19	109.2959 ^a	1.9905	111.2864 ^a	0.4439	0.0089
VARExo	1	20	205.0351 ^a	1.9905	207.0256 ^a	0.6843	0.0078
VAR	1	14	89.5973 ^a	1.9905	91.5878 ^a	0.3885	0.0087
logVARExo	1	13	89.5973 ^a	1.9905	91.5878 ^a	0.3885	0.0083
logVAR	1	11	83.2634^a	1.9905	85.2539^a	0.3700	0.0078
DCC	1	24	173.9361 ^a	1.9905	175.9266 ^a	0.6104	0.0093
BEKK	1	15	92.1307 ^a	1.9905	94.1212 ^a	0.4216	0.0088
HEAVY	1	41	262.6255 ^a	1.9905	264.616 ^a	0.8138	0.0105
NLExo	1	18	137.0583 ^a	1.9905	139.0489 ^a	0.5179	0.0078
NL	1	14	116.0821 ^a	1.9905	118.0727 ^a	0.4625	0.0073

^a indica un livello di significatività rispettivamente all'1%

4 Conclusioni

In questo lavoro, sono stati analizzati e sviluppati modelli econometrici per le dinamiche delle covarianze realizzate.

In analogia a quanto fatto con la volatilità realizzata di un singolo mercato, si è cercato di capire il ruolo predittivo svolto dalle grandezze macroeconomiche.

A differenza di quanto viene fatto nella letteratura sui modelli per le covarianze realizzate, le applicazioni presentate in questa tesi sono basate su dati mensili.

La procedura introdotta in questa tesi prevede la scomposizione di Cholesky della matrice delle covarianze, per garantire che le previsioni siano definite positive. I fattori di Cholesky vengono poi modellati in maniera lineare, con un VAR, e in maniera non lineare, con un modello VLSTAR. Nei modelli vengono incluse delle variabili macroeconomiche e finanziarie scelte tra un insieme di possibili previsori individuati in letteratura e ne viene valutato il ruolo in termini previsionali. È utile sottolineare che tale approccio non consente di definire un modello strutturale per la volatilità ma permette di valutarne le performance previsive. Per questa ragione, il lavoro si concentra sull'analisi delle previsioni out-of-sample, derivanti da uno schema rolling window, dei modelli introdotti e utilizza sia metodi diretti che indiretti di valutazione delle previsioni.

Il confronto diretto delle previsioni con quelle dei modelli esistenti in letteratura evidenzia il vantaggio della procedura Cholesky-VAR e Cholesky-VLSTAR. In particolare, nel dataset cross-country è evidente la superiorità dei modelli Cholesky-VAR e Cholesky-VLSTAR con e senza esogene nei confronti di modelli esistenti in letteratura come DCC, BEKK e HEAVY. Il confronto con i modelli derivanti da una trasformazione logaritmica della matrice delle covarianze realizzate, invece, evidenzia la superiorità della scomposizione di Cholesky in termini previsionali tra le due tecniche di parametrizzazione della matrice delle covarianze. Anche nel dataset sul mercato americano i metodi diretti di valutazione mostrano la migliore accuratezza delle previsioni dell'approccio Cholesky-VAR e Cholesky-VLSTAR.

L'uso di variabili macroeconomiche e finanziarie sembra migliorare le previsioni quando si utilizza un modello lineare, mentre in ambito non lineare le migliori previsioni sembrano essere prodotte da un modello Cholesky-VLSTAR senza esogene, per la maggior parte delle serie. Infine, il Model Confidence Set sottolinea l'importanza di una parametrizzazione della matrice delle covarianze realizzate in una previsione out-of-sample.

Per verificare il valore economico e statistico delle strategie proposte in questa tesi,

si analizzano le performance dei portafogli ottenuti dalle previsioni generate dai modelli considerati e si analizzano le performance dei modelli nella previsione del Value-at-Risk (condizionale e non). Nello specifico, si confrontano le performance dei portafogli ottenuti mediante due tecniche di ottimizzazione con la GMV e la massimizzazione dello Sharpe Ratio. Sulla base delle statistiche descrittive delle serie dei rendimenti di tali portafogli, è evidente che, nel dataset cross-country, le migliori performance in termini di rischio siano quelle ottenute da un modello Cholesky-VAR con esogene ottenuto da una ottimizzazione GMV. Il modello Cholesky-VLSTAR con esogene, ottenuto da una ottimizzazione GMV, presenta uno Sharpe Ratio superiore a tutti i modelli confrontati, mentre in termini di rischio mostra performance in linea con quelle degli altri modelli. Quando il problema di ottimo diventa la massimizzazione dello Sharpe Ratio, il modello Cholesky-VLSTAR con esogene riesce a performare meglio degli altri modelli in termini di rischio, fatta eccezione per il modello DCC.

Nel dataset sul mercato americano i modelli introdotti in questo lavoro presentano performance in linea con gli altri modelli, sia in un'ottimizzazione GMV che in un'ottimizzazione Maximum Sharpe Ratio. La mancanza di utilità dell'approccio introdotto in questa tesi, nel dataset sul mercato americano, può essere ricondotta alla diversità dei rendimenti considerati in questo dataset.

Per quanto concerne l'utilizzo dei diversi approcci analizzati in questo lavoro nella gestione del rischio, i modelli che sembrano presentare le performance migliori sono i modelli non lineari e quelli derivanti da una parametrizzazione logaritmica della matrice delle covarianze, mentre risultano in linea con gli altri modelli i risultati riguardanti i modelli lineari.

Questo lavoro risulta complementare a quello di Halbleib-Chiriac e Voev (2011), che viene esteso mediante l'uso di determinanti macroeconomiche e finanziarie e di un modello non lineare. In aggiunta, questa tesi può essere considerata un'estensione in ambito multivariato dei lavori di Paye (2012), Christiansen *et al.* (2012) e Ghalanos e Rossi (2016). L'uso di un modello non lineare per la volatilità, in ambito multivariato, rappresenta un'innovazione assoluta nella letteratura di riferimento.

Esistono approcci alternativi a quello presentato in questo lavoro per modellare la matrice delle covarianze, oltre ai classici modelli per l'eteroschedasticità condizionale e i modelli di volatilità stocastica; di particolare interesse risultano il modello MIDAS e il Realized Garch. Recentemente, si è assistito all'utilizzo di reti neurali per fare previsioni in ambito finanziario. In particolare, le Support Vector Machine (SVM), introdotte da

Vapnik *et al.* (1997) rappresentano una tecnica promettente per prevedere la volatilità (si veda Chen *et al.* (2010), Ou e Wang (2010) e Li (2014)).

Sono possibili diverse estensioni del presente lavoro per lavori di ricerca futuri. In primo luogo i due nuovi approcci sono stati applicati ad un numero di attività finanziarie contenuto, una possibile estensione futura è analizzare tali modelli quando il numero di titoli è elevato. In secondo luogo, un lavoro di ricerca futuro potrebbe tenere conto di determinanti macroeconomiche e finanziarie differenti, come ad esempio di alcuni strumenti di politica economica. Infine, questo lavoro riguarda solo le previsioni un passo in avanti, un'analisi più approfondita potrebbe analizzare i modelli introdotti per più passi in avanti.

Riferimenti bibliografici

- ADRIAN, T. E BRUNNERMEIR, M. K. (2011). *CoVaR*. Working paper.
- AÏT-SAHALIA, Y., MYKLAND, P. A. E ZHANG, L. (2006). *Comment*. Journal of Business and Economic Statistics, 24: 162–167.
- ALEXANDER, C. E CHIBUMBA, A. M. (1996). *Multivariate orthogonal factor GARCH*. Discussion Paper in mathematics, University of Sussex.
- AMIHUD, Y. E MENDELSON, H. (1987). *Trading mechanisms and stock returns: an empirical investigation*. Journal of Finance, 42: 533–553.
- ANDERSEN, T. G. (2009). *Stochastic volatility*. In *Encyclopedia of Complexity and Systems Science*. Springer Verlag.
- ANDERSEN, T. G. E BOLLERSLEV, T. (1997). *Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns*. Journal of Finance, 52: 975–1005.
- (1998). *Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts*. International Economic Review, 39: 885–905.
- ANDERSEN, T. G., BOLLERSLEV, T., DIEBOLD, F. X. E EBENS, H. (2001a). *The distribution of realized stock return volatility*. Journal of Financial Economics, 61: 43–76.
- ANDERSEN, T. G., BOLLERSLEV, T., DIEBOLD, F. X. E LABYS, P. (2000). *Exchange rate returns standardized by realized volatility are (nearly) Gaussian*. Multinational Finance Journal, 4: 159–179.
- (2001b). *The distribution of exchange rate volatility*. Journal of American Statistical Association, 96: 42–55.
- (2003). *Modeling and Forecasting Realized Volatility*. Econometrica, 71: 579–625.
- ANDERSEN, T. G., BOLLERSLEV, T. E MEDDAHI, N. (2005). *Correcting the Errors: Volatility Forecast Evaluation Using High-Frequency Data and Realized Volatilities*. Econometrica, 73(1): 279–296.
- ANDERSON, H. M. E VAHID, F. (1998). *Testing multiple equation systems for common nonlinear components*. Journal of Econometrics, 84: 1–36.

- ANDREWS, D. W. K. (1991). *Heteroskedasticity and autocorrelation consistent covariance matrix estimation*. *Econometrica*, 59: 817–858.
- ANDREWS, D. W. K. E MONAHAN, J. C. (1992). *An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator*. *Econometrica*, 60(4): 953–966.
- ARAGON, G. O. E FERSON, W. E. (2007). *Portfolio performance evaluation*. *Foundations and Trends in Finance*, 2(2): 83–190.
- ARTZNER, P., DELBAEN, F., EBER, J.-M. E HEATH, D. (1999). *Coherent measures of risk*. *Mathematical Finance*, 9: 203–228.
- ASAI, M., MCALEER, M. E YU, J. (2006). *Multivariate stochastic volatility: A review*. *Econometric Reviews*, 25: 145–175.
- ASGHARIAN, H., CHRISTIANSEN, C. E HOU, A. J. (2014). *Macro-Finance Determinants of the Long-Run Stock-Bond Correlation: The DCC-MIDAS Specification*. Working paper.
- ASGHARIAN, H., HOU, A. J. E JAVED, F. (2013). *The Importance of the Macroeconomic Variables in Forecasting Stock Return Variance: A GARCH-MIDAS Approach*. *Journal of Forecasting*, 32: 600–612.
- BACK, K. (1991). *Asset pricing for general processes*. *Journal of Mathematical Economics*, 20: 371–395.
- BACON, D. W. E WATTS, D. G. (1971). *Estimating the transition between two intersecting straight lines*. *Biometrika*, 58: 525–534.
- BAI, J. E PERRON, P. (1998). *Estimating and Testing Linear Models With Multiple Structural Changes*. *Econometrica*, 66: 47–78.
- (2003a). *Computation and Analysis of Multiple Structural Change Models*. *Journal of Applied Econometrics*, 18: 1–22.
- (2003b). *Critical values for multiple structural change tests*. *Econometrics Journal*, 6: 72–78.
- BANDI, F. M. E RUSSELL, J. R. (2005). *Realized covariation, realized beta and microstructure noise*. Unpublished paper, Graduate School of Business, University of Chicago.

- (2007). *Microstructure noise, realized volatility, and optimal sampling*. Unpublished paper, Graduate School of Business, University of Chicago.
- BARNDORFF-NIELSEN, O. E., HANSEN, P. R., LUNDE, A. E SHEPHARD, N. (2008). *Designing realized kernels to measure the ex post variation of equity prices in the presence of noise*. *Econometrica*, 76: 1481–1536.
- BARNDORFF-NIELSEN, O. E. E SHEPHARD, N. (2002). *Econometric analysis of realised volatility and its use in estimating stochastic volatility models*. *Journal of the Royal Statistical Society*, 64: 253–280.
- (2004). *Power and Bipower Variation with Stochastic Volatility and Jumps*. *Journal of Financial Econometrics*, 2: 1–37.
- BARR, D. G. (2013). *Value at Risk*. Bank of England, Centre for Central Banking Studies.
- BARUNÍK, J. E ČECH, F. (2016). *On the modelling and forecasting multivariate realized volatility: Generalized Heterogeneous Autoregressive (GHAR) model*. *Journal of Forecasting*.
- BAUER, G. H. E VORKINK, K. (2011). *Forecasting multivariate realized stock market volatility*. *Journal of Econometrics*, 160: 93–101.
- BAUWENS, L., HAFNER, C. E LAURENT, S. (2012). *Handbook of Volatility Models and Their Applications*. Wiley.
- BECKER, R., CLEMENTS, A. E O’NEILL, R. (2010). *A Cholesky-MIDAS model for predicting stock portfolio volatility*. Working paper, Centre for Growth and Business Cycle Research Discussion Paper Series.
- BLACK, F. (1976). *Noise*. *Journal of Finance*, 41: 529–543.
- BLANCO, C. E IHLE, G. (1999). *How Good is Your VaR? Using Backtesting to Assess System Performance*. *Financial Engineering News*, 11(8): 1–2.
- BOLLEN, B. E. E INDER, B. (2002). *Estimating daily volatility in financial markets utilizing intraday data*. *Journal of Empirical Finance*, 9: 551–562.
- BOLLERSLEV, T. (1986). *Generalized autoregressive conditional heteroskedasticity*. *Journal of Econometrics*, 31: 307–327.

- (1990). *Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model*. *The Review of Economics and Statistics*, pp. 498–505.
- (2009). *Glossary to ARCH (GARCH)*. Working paper, Duke University.
- BOLLERSLEV, T., ENGLE, R. F. E NELSON, D. B. (1994). *ARCH models*. In *Handbook of Econometrics*. Elsevier Science, Amsterdam.
- BOLLERSLEV, T., ENGLE, R. F. E WOOLDRIDGE, J. M. (1988). *A capital asset pricing model with time-varying covariances*. *The Journal of Political Economy*, pp. 116–131.
- BOLLERSLEV, T. E WOOLDRIDGE, J. M. (1992). *Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances*. *Econometric Reviews*, 11: 143–172.
- BONATO, M. (2009). *Estimating the degrees of freedom of the Realized Volatility Wishart Autoregressive model*. Working paper.
- BONATO, M., CAPORIN, M. E RANALDO, A. (2009). *Forecasting realized (co)variances with a block structure Wishart autoregressive model*. Working papers, Swiss National Bank.
- CAMACHO, M. (2004). *Vector smooth transition regression models for US GDP and the composite index of leading indicators*. *Journal of Forecasting*, 23: 173–196.
- CHEN, L. (1996). *Stochastic Mean and Stochastic Volatility: A Four-Dimensional Term Structure of Interest Rates and Its Application to the Pricing of Derivative Securities*. *Financial Markets, Institutions, and Instruments*, 5: 1–88.
- CHEN, S., HÄRDLE, W. K. E JEONG, K. (2010). *Forecasting Volatility with Support Vector Machine-Based GARCH Model*. *Journal of Forecasting*, 29: 406–433.
- CHRISTENSEN, B. E PRABHALA, N. (1998). *The relation between implied and realized volatility*. *Journal of Financial Economics*, 37: 125–150.
- CHRISTIANSEN, C., SCHMELING, M. E SCHRIMPF, A. (2012). *A comprehensive look at financial volatility prediction by economic variables*. *Journal of Applied Econometrics*, 27: 956–977.
- CHRISTOFFERSEN, P. F. (1998). *Evaluating interval forecasts*. *International Economic Review*, 39: 841–862.

- CLARK, T. E. E WEST, K. D. (2007). *Approximately normal tests for equal predictive accuracy in nested models*. *Journal of Econometrics*, 138(1): 291–311.
- COLACITO, R., ENGLE, R. F. E GHYSELS, E. (2011). *A Component Model for Dynamic Correlations*. *Journal of Econometrics*, 164: 45–59.
- CONRAD, C. E LOCH, K. (2014). *Anticipating Long-Term Stock Market Volatility*. *Journal of Applied Econometrics*, Forthcoming.
- CORSI, F. (2009). *A simple approximate long-memory model of realized volatility*. *Journal of Financial Econometrics*, 7: 174–196.
- DE POOTER, M., MARTENS, M. E VAN DIJK, D. (2008). *Predicting the Daily Covariance Matrix for S&P 100 Stocks Using Intraday Data - But Which Frequency to Use?* *Econometric Reviews*, 27: 199–229.
- DEMIGUEL, V., GARLAPPI, L. E UPPAL, R. (2009). *Optimal Versus Naive Diversification: How Inefficient is the 1-N Portfolio Strategy?* *Review of Financial Studies*, 22(5): 1915–1953.
- DIEBOLD, F. X. E MARIANO, R. S. (1995). *Comparing predictive accuracy*. *Journal of Business and Economic Statistics*, 13: 253–263.
- DIEBOLD, F. X. E YILMAZ, K. (2009). *Measuring Financial Asset Return and Volatility Spillovers, with Application to Global Equity Markets*. *Economic Journal*, 119: 158–171.
- EBENS, H. (1999). *Realized Stock Index Volatility*. Working Paper No. 420, Department of Economics, Johns Hopkins University, Baltimore.
- ENGLE, R. F. (1982). *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation*. *Econometrica*, 50: 987–1007.
- (2002). *Dynamic conditional correlation: a simple class of multivariate GARCH models*. *Journal of Business and Economic Statistics*, 20: 339–350.
- ENGLE, R. F. E GALLO, G. M. (2006). *A multiple indicators model for volatility using intra-daily data*. *Journal of Econometrics*, 131: 3–27.
- ENGLE, R. F., GHYSELS, E. E SOHN, B. (2009). *On the Economic Sources of Stock Market Volatility*. Working paper.

- (2013). *Stock Market Volatility and Macroeconomic Fundamentals*. *Review of Economics and Statistics*, 95: 776–797.
- ENGLE, R. F. E KRONER, K. F. (1995). *Multivariate simultaneous generalized ARCH*. *Econometric Theory*, 11: 122–150.
- ENGLE, R. F., NG, V. K. E ROTHSCCHILD, M. (1990). *Asset pricing with a factor ARCH covariance structure: empirical estimates for treasury bills*. *Journal of Econometrics*, 45: 213–238.
- ENGLE, R. F. E RANGEL, J. G. (2008). *The Spline-GARCH Model for Low-Frequency Volatility and Its Global Macroeconomic Causes*. *Review of Financial Studies*, 21: 1187–1222.
- ENGLE, R. F. E SHEPPARD, K. (2001). *Theoretical and empirical properties of dynamical conditional correlation model multivariate GARCH*. UCSD Discussion.
- EPPS, T. W. (1979). *Comovements in Stock Prices in the Very Short Run*. *Journal of the American Statistical Association*, 74: 291–298.
- FAMA, E. E FRENCH, K. (1989). *Business Conditions and Expected Returns on Stock and Bonds*. *Journal of Financial Economics*, 25: 23–49.
- (1993). *Common Risk Factors in the Returns on Stocks and Bonds*. *Journal of Financial Economics*, 33: 3–56.
- FLEMING, J., KIRBY, C. E OSTDIEK, B. (2003). *The economic value of volatility timing using realized volatility*. *Journal of Financial Economics*, 67: 473–509.
- FORNARI, F. E MELE, A. (1997). *Sign- and volatility-switching ARCH models: Theory and applications to international stock markets*. *Journal of Applied Econometrics*, 12: 49–65.
- GHALANOS, A. E ROSSI, E. (2016). *Stock market volatility and economic variables: A nonlinear approach*. Preliminary version.
- GHYSELS, E., HARVEY, A. E RENAULT, E. (1996). *Stochastic volatility*. In *Statistical models in finance*, pp. 119–191. Amsterdam: North-Holland.
- GHYSELS, E., RUBIA, A. E VALKANOV, R. (2009). *Multi-Period Forecasts of Volatility: Direct, Iterated, and Mixed-Data Approaches*. Working paper.

- GHYSELS, E., SANTA-CLARA, P. E VALKANOV, R. (2004). *The MIDAS touch: Mixed data sampling regression models*. Working paper, UNC and UCLA.
- (2006). *Predicting volatility: getting the most out of return data sampled at different frequencies*. *Journal of Econometrics*, 131: 59–95.
- GIACOMINI, R. E WHITE, H. (2006). *Tests of conditional predictive ability*. *Econometrica*, 74: 1545–1578.
- GIOT, P. E LAURENT, S. (2004). *Modelling daily Value-at-Risk using realized volatility and ARCH type models*. *Journal of Empirical Finance*, 11: 379–398.
- GOLUB, G. H. E VAN LOAN, C. F. (1996). *Matrix Computations*. Johns Hopkins University Press.
- GONZALO-RIVERA, G. (1998). *Smooth transition GARCH models*. *Studies in Nonlinear Dynamics and Econometrics*, 3: 61–78.
- GOURIEROUX, C. S., JASIAK, J. E SUFANA, R. (2009). *The Wishart Autoregressive process of multivariate stochastic volatility*. *Journal of Econometrics*, 150: 167–181.
- HAGERUD, G. (1997). *A New Non-Linear GARCH Model*. Ph.D. Thesis, Stockholm School of Economics.
- HALBLEIB-CHIRIAC, R. (2007). *Nonstationary Wishart Autoregressive Model*. Working Paper, CoFE, University of Konstanz.
- HALBLEIB-CHIRIAC, R. E VOEV, V. (2011). *Modelling and Forecasting Multivariate Realized Volatility*. *Journal of Applied Econometrics*, 26: 922–947.
- HANSEN, P. R. (2005). *A test for superior predictive ability*. *Journal of Business and Economics Statistics*, 23: 365–380.
- HANSEN, P. R., HUANG, Z. E SHEK, H. H. (2011a). *Realized GARCH: A Joint Model for Returns and Realized Measures of volatility*. *Journal of Applied Econometrics*, 27: 877–906.
- HANSEN, P. R. E LUNDE, A. (2004). *An unbiased measure of realized variance*. Unpublished manuscript, Stanford Univeristy.

- (2005). *A realized variance for the whole day based on intermittent data*. *Journal of Financial Econometrics*, 3: 525–554.
- (2006). *Consistent ranking of volatility models*. *Journal of Econometrics*, 131(1-2): 97–121.
- HANSEN, P. R., LUNDE, A. E. NASON, J. M. (2011b). *The Model Confidence Set*. *Econometrica*, 79: 435–497.
- HANSEN, P. R., LUNDE, A. E. VOEV, V. (2014). *Realized Beta GARCH: A Multivariate GARCH Model With Realized Measures of Volatility*. *Journal of Applied Econometrics*, 29: 774–799.
- HARRIS, L. (1990). *Estimation of stock variance and serial covariance from discrete observations*. *Journal of Financial and Quantitative Analysis*, 25: 291–306.
- (1991). *Stock price clustering and discreteness*. *Review of Financial Studies*, 4: 389–415.
- HARVEY, A., RUIZ, E. E SHEPHARD, N. (1994). *Multivariate stochastic variance models*. *Review of Economic Studies*, 61: 247–264.
- (1996). *Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns*. *Journal of Business and Economic Statistics*, 14: 429–434.
- HAUTSCH, N., KYJ, L. M. E MALEC, P. (2015). *Do High-Frequency Data Improve High Dimensional Portfolio Allocations?* *Journal of Applied Econometrics*, 30(2): 263–290.
- HEIDEN, M. D. (2015). *Pitfalls of the Cholesky decomposition for forecasting multivariate volatility*. Working paper.
- HESTON, S. L. (1993). *A closed form solution for options with stochastic volatility with applications to bond and currency options*. *The Review of Financial Studies*, 6: 327–343.
- HOSKING, J. (1981). *Equivalent Forms of the Multivariate Portmanteau Statistic*. *Journal of the Royal Statistical Society*, 43.
- HUBRICH, K. E TERÄSVIRTA, T. (2013). *Thresholds and Smooth Transitions in Vector Autoregressive Models*. CREATES Research Paper, Aarhus University.
- J. D. JOBSON, B. M. K. (1981). *Performance Hypothesis Testing with the Sharpe and Treynor Measures*. *The Journal of Finance*, 36(4): 889–908.

- JACOD, J. E PROTTER, P. (1998). *Asymptotic error distributions for the Euler method for stochastic differential equations*. *Annals of Probability*, 26: 267–307.
- JENSEN, M. C. (1968). *The Performance of Mutual Funds in the Period 1945-1964*. *Journal of Finance*, 23: 389–416.
- JIN, X. E MAHEU, J. M. (2012). *Modelling Realized Covariances and Returns*. *Journal of Financial Econometrics*, 11: 335–369.
- KOCK, A. B. E TERÄSVIRTA, T. (2011). *Forecasting with nonlinear time series models*. CREATES Research Paper, Aarhus University.
- KUESTER, K., MITTNIK, S. E PAOLELLA, M. S. (2006). *Value-at-risk prediction: a comparison of alternative strategies*. *Journal of Financial Econometrics*, 4: 53–89.
- KUPIEC, P. (1995). *Techniques for verifying the accuracy of risk measurement models*. *Journal of Derivatives*, 2: 73–84.
- KYJ, L., OSTDIEK, B. E ENSOR, K. (2009). *Realized Covariance Estimation in Dynamic Portfolio Optimization*. Working Paper.
- LAURENT, A. C., KOCK, A. B. E MEDEIROS, M. C. (2014). *Estimation and Forecasting of Large Realized Covariance Matrices and Portfolio Choice*. CREATES Research Paper, Aarhus University.
- LAURENT, S., ROMBOUITS, J. V. E VIOLANTE, F. (2013). *On Loss Functions and Ranking Forecasting Performances of Multivariate Volatility Models*. *Journal of Econometrics*, 173(1): 1–10.
- LAURENT, S. E VIOLANTE, F. (2012). *Volatility forecasts evaluation and comparison*. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(1): 1–12.
- LEDOIT, O. E WOLF, M. (2008). *Robust Performance Hypothesis Testing with the Sharpe Ratio*. IEW - Working Papers 320, Institute for Empirical Research in Economics - University of Zurich.
- LI, Y. (2014). *Estimating and Forecasting APARCH-Skew-t Models by Wavelet Support Vector Machines*. *Journal of Forecasting*, 33(4): 259–269.
- LOPEZ, J. A. (1998). *Regulatory evaluation of Value-at-Risk models*. *Federal Reserve Bank of New York Economic Policy Review*, 4(3): 119–124.

- LUUKKONEN, R., SAIKKONEN, P. E TERÄSVIRTA, T. (1988). *Testing Linearity Against Smooth Transition Autoregressive Models*. *Biometrika*, 75: 491–499.
- MAHEU, J. M. E MCCURDY, T. H. (2002). *Nonlinear features of realized volatility*. *Review of Economics and Statistics*, 84: 668–681.
- MARKOWITZ, H. (1952). *Portfolio selection*. *Journal of Finance*, 7: 77–91.
- MARTENS, M., DE POOTER, M. E VAN DIJK, D. J. (2004). *Modeling and Forecasting S&P 500 Volatility: Long Memory, Structural Breaks and Nonlinearity*. Tinbergen Institute Discussion Paper No. 04-067/4.
- MCÁLEER, M. E MEDEIROS, M. (2008). *A multiple regime smooth transition Heterogeneous Autoregressive model for long memory and asymmetries*. *Journal of Econometrics*, 147: 104–119.
- MELE, A. (2007). *Asymmetric stock market volatility and the cyclical behavior of expected returns*. *Journal of Financial Economics*, 86: 446–478.
- (2008). *Understanding Stock Market Volatility - A Business Cycle Perspective*. Working Paper.
- MEMMEL, C. (2003). *Performance Hypothesis Testing with the Sharpe Ratio*. *Finance Letters*, 1(1): 21–23.
- MERTON, R. C. (1974). *On the Pricing of Corporate Debt: The Risk Structure of Interest Rates*. *Journal of Finance*, 29: 449–470.
- (1980). *On estimating the expected return on the market: An explanatory investigation*. *Journal of Financial Economics*, 8: 323–361.
- MINCER, J. A. E ZARNOWITZ, V. (1969). *The Evaluation of Economic Forecasts*. In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, NBER Chapters, pp. 3–46. National Bureau of Economic Research, Inc.
- NAGEL, S. (2012). *Evaporating Liquidity*. *The Review of Financial Study*, 25: 2005–2039.
- NASON, J. A. E DIEBOLD, F. X. (1990). *Nonparametric Exchange Rate Prediction*. *Journal of International Economics*, 28(3-4): 315–332.

- NELSON, D. B. (1990). *Stationarity and persistence in the GARCH(1,1) model*. *Econometric Theory*, 6: 318–334.
- NOURELDIN, D., SHEPHARD, N. E SHEPPARD, K. (2011). *Multivariate high-frequency-based volatility (HEAVY) models*. *Journal of Applied Econometrics*, 27: 907–933.
- OOMEN, R. C. (2001). *Using High Frequency Stock Market Index Data to Calculate, Model & Forecast Realized Return Variance*. Working paper.
- (2005a). *Properties of bias-corrected realized variance under alternative sampling schemes*. *Journal of Financial Econometrics*, 3: 555–577.
- (2005b). *Properties of realized variance under alternative sampling schemes*. *Journal of Business and Economic Statistics*, 24: 219–237.
- OU, P. E WANG, H. (2010). *Financial Volatility Forecasting by Least Square Support Vector Machine Based on GARCH, EGARCH and GJR Models: Evidence from ASEAN Stock Markets*. *International Journal of Economics and Finance*, 2: 51–64.
- PATTON, A. J. (2011). *Volatility forecast comparison using imperfect volatility proxies*. *Journal of Econometrics*, 160(1): 246 – 256.
- PATTON, A. J. E SHEPPARD, K. (2009). *Evaluating Volatility and Correlation Forecasts*. Springer Berlin Heidelberg.
- PAYE, B. S. (2012). *Dèjà vol: Predictive regressions for aggregate stock market volatility using macroeconomic variables*. *Journal of Financial Economics*, 106: 527–546.
- PINHEIRO, J. C. E BATELS, D. M. (1996). *Unconstrained parametrizations for variance-covariance matrices*. *Statistics and Computing*, 6: 289–296.
- POTERBA, J. M. E SUMMERS, L. H. (1986). *The persistence of volatility and stock market fluctuations*. *The American Economic Review*, 76: 1142–1151.
- POURAHMADI, M. (1999). *Joint mean-covariance models with applications to longitudinal data: unconstrained parameterisation*. *Biometrika*, 86: 677–690.
- PROTTER, P. E. (1992). *Stochastic Integration and Differential Equations: A New Approach*. New York: Springer-Verlag.

- RABEMANANJARA, R. E ZAKOIAN, J. M. (1993). *Threshold ARCH Models and Asymmetries in Volatility*. Journal of Applied Econometrics, 8: 31–49.
- RICHARDSON, M. E STOCK, J. H. (1989). *Drawing inference from statistics based on multiyear asset returns*. Journal of Financial Economics, 25: 323–348.
- ROSS, S. A. (1976). *The arbitrage theory of capital asset pricing*. Journal of Economic Theory, 13: 341–360.
- SCHWERT, W. G. (1989). *Why does stock market volatility change over time*. Journal of Finance, 44: 1115–1153.
- (1990). *Indexes of U.S. stock prices from 1802 to 1987*. Journal of Business, 63: 399–426.
- SHARPE, W. F. (1964). *Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk*. Journal of Finance, 19: 425–442.
- SHEPHARD, N. (1996). *Statistical aspects of ARCH and stochastic volatility*. In *Time Series Models in Econometrics, Finance and Other Fields*.
- SHEPHARD, N. E SHEPPARD, K. (2010). *Realising the future: forecasting with high frequency based volatility (HEAVY) models*. Journal of Applied Econometrics, 25: 197–231.
- SHIRYAEV, A. N. (1999). *Essentials of Stochastic Finance: Facts, Models, Theory*. World Scientific.
- SILVENNOINEN, A. E TERÄSVIRTA, T. (2008). *Multivariate GARCH Models*. Working paper.
- SOWELL, F. (1992). *Maximum likelihood estimation of fractionally integrated time series models*. Journal of Econometrics, 53: 165–188.
- SPINU, F. (2015). *Buy-and-hold versus constantly rebalanced portfolios: A theoretical comparison*. Journal of Asset Management, 16: 79–84.
- SWANSON, E. T. E WILLIAMS, J. C. (2013). *Measuring the Effect of the Zero Lower Bound on Yields and Exchange Rates in the U.K. and Germany*. Federal Reserve Bank of Francisco, Working Paper Series.

- TAYLOR, S. J. (1994). *Modeling Stochastic Volatility: A Review and Comparative Study*. *Mathematical Finance*, 4: 183–204.
- TAYLOR, S. J. E XU, X. (1997). *The incremental volatility information in one million foreign exchange quotations*. *Journal of Empirical Finance*, 4: 317–340.
- TERÄSVIRTA, T., TJØ STHEIM, D. E GRANGER, W. (2010). *Modelling Nonlinear Economic Time Series*. Oxford University Press.
- TERÄSVIRTA, T. E YANG, Y. (2014). *Specification, estimation and evaluation of vector smooth transition autoregressive models with applications*. CREATES Research Paper, Aarhus University.
- THEIL, H. E BOOT, J. C. G. (1962). *The final form of econometric equation systems*. *Review of International Statistical Institute*, 30: 136–152.
- TSAY, R. S. (2005a). *Analysis of Financial Time Series*. Wiley.
- (2005b). *Testing and modeling multivariate threshold models*. *Journal of the American Statistical Association*, 93: 1188–1202.
- VAN DER WEIDE, R. (2002). *GO-GARCH: a multivariate generalized orthogonal GARCH model*. *Journal of Applied Econometrics*, 17: 549–564.
- VAPNIK, V., GOLOWICH, S. E. E SMOLA, A. (1997). *Support vector method for function approximation, regression estimation, and signal processing*. *Advances in Neural Information Processing Systems*, 9: 281–287.
- WELCH, I. E GOYAL, A. (2008). *A comprehensive look at the empirical performance of equity premium prediction*. *Review of Financial Studies*, 21(4): 1455–1508.
- WEST, K. (1996). *Asymptotic inference about predictive ability*. *Econometrica*, 64: 1067–1084.
- WOLD, S. (1976). *Spline Functions in Data Analysis*. *Technometrics*, 16(1): 1–11.
- YANG, Y. E TERÄSVIRTA, T. (2013). *Linearity and misspecification tests for vector smooth transition regression models*. CREATES Research Paper, Aarhus University.

- ZELLNER, A. (1962). *An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests of Aggregation Bias*. Journal of the American Statistical Association, 57(298): 348–368.
- ZHANG, L. (2006). *Estimating Covariation: Epps Effect, Microstructure Noise*. Working paper.
- ZHANG, L., MYKLAND, P. A. E AÏT-SAHALIA, Y. (2005). *A tale of two time scales: determining integrated volatility with noisy high frequency data*. Journal of the American Statistical Association, 100: 1394–1411.
- ZHOU, B. (1996). *High frequency data and volatility in foreign-exchange rates*. Journal of Business and Economic Statistics, 14: 45–52.

A

Parametrizzazioni della matrice delle covarianze

In generale, una matrice varianze e covarianze Σ , di dimensione $n \times n$, data da

$$\Sigma = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

è simmetrica e semi-definita positiva per ogni x tale per cui

$$x^T \Sigma x \geq 0.$$

In questa Appendice verrà considerato solo il caso di una matrice definita positiva, ignorando tutti gli altri casi, dovuti a combinazioni lineari delle variabili casuali sottostanti. Gli elementi non ripetuti della matrice possono essere racchiusi in un vettore θ di dimensione $n(n+1)/2$. La matrice può essere scritta come

$$\Sigma = L'L \tag{A.0.1}$$

dove $L = L(\theta)$ è una matrice $n \times n$ a rango pieno ottenuta da un vettore di dimensione $n(n+1)/2$ di parametri non ristretti. Qualunque Σ definita da una L a rango pieno come in (A.0.1) è definita positiva.

Scelte differenti di L portano a differenti parametrizzazioni di Σ . In questa appendice sono considerate due classi di L : una basata sulla scomposizione di Cholesky di Σ e un'altra basata sulla scomposizione spettrale di Σ . La seguente matrice varianze e covarianze viene usata a scopo dimostrativo

$$\Sigma = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix} \tag{A.0.2}$$

A.1 Parametrizzazione di Cholesky

Poiché Σ è definita positiva, può essere fattorizzata come $\Sigma = L'L$, dove L è una triangolare alta.

Un problema con la parametrizzazione di Cholesky è che il fattore di Cholesky non risulta unico, infatti se L è un fattore di Cholesky di Σ , allora lo è ogni matrice ottenuta

moltiplicando un sottoinsieme di righe di L per -1 . Questo ha implicazioni sull'identificazione dei parametri, in quanto 2^n differenti θ possono rappresentare la stessa matrice Σ . Potrebbero sorgere problemi nell'ottimizzazione di una funzione obiettivo quando differenti soluzioni di ottimo sono vicine nello spazio dei parametri.

Un altro problema con la parametrizzazione di Cholesky è la mancanza di una relazione diretta tra θ e gli elementi di Σ . Questo rende difficile da interpretare le stime di θ e ottenere gli intervalli di confidenza per le varianze e le covarianze in Σ basate sugli intervalli di confidenza per gli elementi di θ . L'unica eccezione è data da $|L_{11}| = \sqrt{|\Sigma_{11}|}$, così gli intervalli di confidenza su $|\Sigma_{11}|$ possono essere ottenuti dagli intervalli di confidenza su $[L_{11}]$, dove $[A]_{ij}$ denota il ij -esimo elemento di Σ . Modificando in maniera appropriata le colonne e le righe di Σ si possono derivare gli intervalli di confidenza per tutte le varianze basati sugli intervalli di confidenza per gli elementi di L . Il più grande vantaggio di questa parametrizzazione, oltre alla garanzia di previsioni di Σ definite positive, è che è computazionalmente stabile e semplice.

A.2 Parametrizzazione logaritmica di una matrice

Questa parametrizzazione è basata sulla scomposizione spettrale della matrice varianze e covarianze Σ . Poiché Σ è definita positiva, essa ha n autovalori positivi λ . Sia U la matrice ortogonale di autovettori ortonormali di Σ e $\Lambda = \text{diag}(\lambda)$, si può scrivere

$$\Sigma = U\Lambda U'.$$

Ponendo

$$L = \Lambda^{1/2}U'$$

nell'equazione $\Sigma = L'L$, dove $\Lambda^{1/2}$ rappresenta la matrice diagonale con $[\Lambda^{1/2}]_{ii} = \sqrt{[\Lambda]_{ii}}$, si ha una fattorizzazione di Σ basata sulla scomposizione spettrale. La trasformazione logaritmica di Σ è definita come $\log(\Sigma) = U \log(\Lambda) U'$, dove $\log(\Lambda) = \text{diag}[\log(\Lambda)]$. Σ e $\log(\Sigma)$ condividono gli stessi autovettori. La matrice $\log(\Sigma)$ può assumere qualsiasi valore nello spazio di $n \times n$ matrici simmetriche.

La parametrizzazione logaritmica della matrice restituisce una mappatura uno a uno tra gli elementi θ di $\log(\Sigma)$ e Σ , non sorgono pertanto i problemi di identificazione riscontrati nella scomposizione di Cholesky. Il vettore di parametri θ , in maniera simile alla parametrizzazione di Cholesky, non ha una diretta interpretazione rispetto agli elementi originali della matrice varianze e covarianze Σ . Per definire $\log(\Sigma)$ si ricorre all'uso dell'algoritmo Schur-Fréchet.

B

Test linearità congiunto

Quando il modello VLSTAR presenta un'unica variabile di transizione, $s_{1t} = s_{2t} = \dots = s_{\tilde{n}t} = s_t$, può essere applicato un test congiunto di linearità. Il test è una generalizzazione nel multivariato del test di Luukkonen *et al.* (1988).

Si consideri un modello VLSTAR a due regimi

$$y_t = B_1 z_t + G_t B_2 z_t + \varepsilon_t. \quad (\text{B.0.1})$$

L'ipotesi nulla del test di linearità è data da $H_0 : \gamma_j = 0, j = 1, \dots, \tilde{n}$, per cui $G_t \equiv (1/2)/\tilde{n}$ e la (B.0.1) sono lineari. Quando è vera l'ipotesi nulla, emerge un problema di identificazione del parametro c_j nella funzione di transizione e della combinazione lineare $B_1 + (1/2)B_2$, che può essere risolto mediante un'espansione di Taylor intorno a $\gamma_j = 0$.

L'approssimazione della funzione logistica con un'espansione di Taylor di primo ordine è data da

$$\begin{aligned} G(s_t; \gamma_j, c_j) &= (1/2) + (1/4)\gamma_j(s_t - c_j) + r_{jt} \\ &= a_j s_t + b_j + r_{jt} \end{aligned} \quad (\text{B.0.2})$$

dove $a_j = \gamma_j/4$, $b_j = 1/2 - a_j c_j$ e r_j è il resto dell'espansione. Se la matrice G_t viene specificata come

$$\begin{aligned} G_t &= \text{diag}\{a_1 s_t + b_1 + r_{1t}, \dots, a_{\tilde{n}} s_t + b_{\tilde{n}} + r_{\tilde{n}t}\} \\ &= A s_t + B + R_t \end{aligned} \quad (\text{B.0.3})$$

dove $A = \text{diag}(a_1, \dots, a_{\tilde{n}})$, $B = \text{diag}(b_1, \dots, b_{\tilde{n}})$ e $R_t = \text{diag}(r_{1t}, \dots, r_{\tilde{n}t})$. y_t può essere riscritta come

$$\begin{aligned} y_t &= B_1 z_t + (A s_t + B + R_t) B_2 z_t + \varepsilon_t \\ &= (B_1 + B B_2) z_t + A B_2 z_t s_t + R_t B_2 z_t + \varepsilon_t \\ &= \Theta_0 z_t + \Theta_1 z_t s_t + \varepsilon_t^* \end{aligned} \quad (\text{B.0.4})$$

dove $\Theta_0 = B_1 + B_2' B$, $\Theta_1 = B_2' A$ e $\varepsilon_t^* = R_t B_2 + \varepsilon_t$. Sotto l'ipotesi nulla, $\Theta_0 = B_1$ e $\Theta_1 = 0$ e il modello (B.0.4) è lineare, con $\varepsilon_t^* = \varepsilon_t$. Ne segue che il test Lagrange multiplier (LM), sotto l'ipotesi nulla, è ricavato dallo score

$$\frac{\partial \log L(\tilde{\theta})}{\partial \Theta_1} = \sum_{t=1}^T z_t s_t (y_t - \tilde{B}_1 z_t)' \tilde{\Omega}^{-1} = S(Y - Z \tilde{B}_1) \tilde{\Omega}^{-1}, \quad (\text{B.0.5})$$

dove

$$Y = \begin{bmatrix} y'_1 \\ \vdots \\ y'_T \end{bmatrix}, \quad Z = \begin{bmatrix} z'_1 \\ \vdots \\ z'_T \end{bmatrix}, \quad S = \begin{bmatrix} z'_1 s_1 \\ \vdots \\ z'_T s_T \end{bmatrix}$$

e dove \tilde{B}_1 e $\tilde{\Omega}$ vengono stimati dal modello in H_0 . Se $P_Z = Z(Z'Z)^{-1}Z'$ è la matrice di proiezione di Z , il test LM è specificato come segue

$$LM = tr\{\tilde{\Omega}^{-1}(Y - Z\tilde{B}_1)'S[S'(I_T - P_Z)S]^{-1}S'(Y - Z\tilde{B}_1)\}. \quad (\text{B.0.6})$$

Sotto l'ipotesi nulla, la statistica test si distribuisce come una χ^2 con $\tilde{n}(p \cdot \tilde{n} + k)$ gradi di libertà.