

UNIVERSITA' DEGLI STUDI DI PADOVA
FACOLTA' DI SCIENZE STATISTICHE



TESI DI LAUREA IN STATISTICA ECONOMIA E FINANZA

**STIMA DELLA VOLATILITA' NEI MERCATI
FINANZIARI CON DATI INFRA-GIORNALIERI:
ALCUNI CONFRONTI**

RELATORE: CH.MO PROF. SILVANO BORDIGNON

LAUREANDA: ALESSIA BORGATO

ANNO ACCADEMICO 2005 – 2006

Ai miei genitori

Indice

Premessa

Capitolo 1: I dati ad alta frequenza nei mercati finanziari

1.1 Introduzione

1.2 Variabili d'interesse per lo studio di serie storiche finanziarie

1.2.1 Tempo

1.2.2 Prezzo

1.2.3 Rendimenti

1.2.4 Volatilità

1.2.5 Spread

1.3 Le caratteristiche statistiche dei dati finanziari ad alta frequenza

1.3.1 La distribuzione del processo dei rendimenti

1.3.2 La memoria del processo

Capitolo 2: Gli stimatori della volatilità

2.1 Introduzione

2.2 Stimatori giornalieri

2.2.1 Lo stimatore Simple

2.2.2 Gli stimatori Extreme Value

2.3 Modello di studio proposto

2.3.1 Lo stimatore VARHAC

2.3.1.1 Costruzione dello stimatore

Capitolo 3: La serie di dati

3.1 Introduzione

3.2 I mercati finanziari

3.2.1 Il mercato azionario

3.3 L'indice S&P 500

3.3.1 Analisi preliminare dei dati

Capitolo 4: Applicazione degli stimatori all'indice S&P 500

4.1 Analisi dei dati

4.2 Bontà degli stimatori

4.2.1 Criterio 1: Confronto con le stime della volatilità semplice

4.2.2 Criterio 2: Normalità dei rendimenti standardizzati

4.2.3 Criterio 3: Varianza dei rendimenti standardizzati

Capitolo 5: Conclusioni

5.1 Confronto tra gli stimatori considerati

Appendice

Bibliografia

Premessa

Il lavoro presentato in questa tesi si propone di analizzare un concetto fondamentale per lo studio delle serie storiche finanziarie: la volatilità giornaliera realizzata che contraddistingue i dati ad alta frequenza.

La disponibilità di serie temporali ad elevata numerosità campionaria ha indotto i ricercatori, di vari settori, a sviluppare nuove tecniche di analisi cercando di sfruttare la massima quantità di informazione possibile nel modo più efficiente. La possibilità di sfruttare questa enorme mole di informazioni costituisce un indubbio vantaggio, tuttavia, l'impiego di tali dati pone dei problemi. Oltre a quello di gestire e organizzare numerose osservazioni (si può arrivare ad alcune migliaia di registrazioni nell'arco di una sola giornata borsistica) si pone quello del trattamento di osservazioni a frequenza irregolare, che si riflette sulle variabili utilizzate per le analisi delle serie: rendimenti, volatilità, ecc..

Concentriamo l'attenzione sulla stima della volatilità e sulle assunzioni del processo generatore dei dati, confrontando alcuni stimatori della volatilità giornaliera frequentemente utilizzati con uno stimatore, denominato VARHAC (Bollen and Inder, 2002), che usa un insieme meno restrittivo di assunzioni, tenendo conto di alcune caratteristiche tipiche dei rendimenti infra-giornalieri, quali eteroschedasticità e autocorrelazione.

La tesi è organizzata in cinque capitoli. Il primo descrive le caratteristiche principali dei dati ad alta frequenza, considerando sia aspetti economici che statistici. Il secondo si sofferma sul concetto di volatilità riassumendo diversi stimatori utilizzati in letteratura e presentando un nuovo approccio per la stima della volatilità: lo stimatore VARHAC. Il terzo presenta la serie dei dati utilizzati per l'applicazione delle tecniche di stima, l'indice S&P 500. Nel quarto capitolo è comparato lo stimatore VARHAC con gli stimatori citati usando l'indice S&P 500 e dopo un'analisi preliminare se ne

verifica la bontà attraverso alcuni criteri. Infine, l'ultimo capitolo riassume i confronti tra gli stimatori e conclude.

Capitolo 1: I dati ad alta frequenza nei mercati finanziari

1.1 Introduzione

La risorsa dei dati ad alta frequenza sono i mercati finanziari. La forma originale dei prezzi di mercato sono i dati *tick by tick*. Ogni *tick* è un'unità logica di informazione, come una quota o un prezzo di una transazione, e ogni giorno d'apertura del mercato genera centinaia e migliaia di *ticks*. Per natura questi dati sono spazati irregolarmente nel tempo.

I dati ad alta frequenza sono uno degli "oggetti" principali di ricerca per coloro che sono interessati a capire gli andamenti di un indice o di un'azione nel mercato finanziario, in quanto osservando i dati e avendo molte informazioni a disposizione permette ai professionisti di determinare le loro decisioni commerciali.

La maggior parte degli studi condotti nel passato citati nella letteratura trattano dati a bassa frequenza spazati regolarmente nel tempo. Le ragioni principali sono due. La prima, era abbastanza costoso in termini di denaro e di tempo, raccogliere, unire e immagazzinare dati ad alta frequenza perché i dati finanziari disponibili erano giornalieri o settimanali. La seconda, la maggior parte degli strumenti statistici era sviluppata per serie temporali omogenee.

Con lo sviluppo delle tecnologie informatiche, la disponibilità dei dati è diventato un problema minore perché si sono ridotti i costi di rilevazione e memorizzazione dei dati e gli scambi elettronici ci permettono di negoziare in tutto il mondo in tempo reale permettendo di controllare l'evoluzione dei prezzi e l'andamento di un mercato. Grazie lo sviluppo di tecniche informatiche sempre più sofisticate, e ad elevata diffusione, si è permesso ad un gran numero di utenti di avere a disposizione dati finanziari ad alta frequenza, consentendo ai ricercatori di sviluppare nuovi metodi di analisi e modificare quelli esistenti.

Oggi esistono alcuni circuiti telematici (per esempio *Tenfore* e *Reuters*) che sono pagati dagli utenti per rilevare le quotazioni in tempo reale. Questi circuiti non raccolgono i dati come serie storiche, quindi alcune società di ricerca acquistano i dati dai circuiti telematici e si incaricano di ordinarli e filtrarli a seconda delle esigenze di analisi. Una di queste società di ricerca internazionale è la *Olsen Ltd* fondata nel 1985 dall'economista Richard Olsen. La *Olsen Ltd* è una dei principali fornitori di servizi e tecnologie *e-finance*, inclusi dati ad alta frequenza, modelli avanzati di previsioni, di amministrazione dei rischi e di supporto alle decisioni di investimento. Inizialmente questa società aveva lo scopo di raccogliere e riordinare gli insiemi di quotazioni *tick by tick* relative ai mercati finanziari mondiali. Questo lavoro ha permesso di fornire nuove intuizioni sul meccanismo dei mercati finanziari e sviluppare una nuova teoria finanziaria sviluppando nuovi metodi di analisi e di previsione dei dati ad alta frequenza. I suoi prodotti e i suoi servizi sono utilizzati da più di 200 istituti finanziari in tutto il mondo. Le ricerche della *Olsen Ltd* sono basate su insiemi di dati ad alta frequenza raccolti, a partire dal 1986, prima dai circuiti telematici *Reuters*, *Knight-Ridder* e *Telerate*, poi da GTIS e *Tenfore*. Le più importanti innovazioni introdotte dalla società sono le procedure di filtraggio dei dati, la legge di scala per il mercato degli scambi, gli studi sulle proprietà di lunga memoria della volatilità, lo sviluppo di nuovi modelli della volatilità e di software per l'analisi di previsione e lo sviluppo di servizi *on-line* dei *tick* nei vari mercati finanziari.

L'analisi di serie finanziarie ad alta frequenza ha fatto emergere la nozione di mercato eterogeneo: l'asimmetria di informazione e il differente scopo di investimento dei partecipanti, porta quest'ultimi a comportarsi in maniera diversa secondo le varie notizie che giungono sul mercato causando complesse dinamiche.

L'uso di dati ad alta frequenza permette di studiare l'influenza delle notizie sui prezzi: ogni operatore interpreta a suo modo le notizie, portando variazioni di rilievo nell'attività del mercato, provocando diverse variazioni

dei prezzi, questo porta il mercato ad essere caratterizzato da periodi di alta volatilità (in cui gli scambi sono frequenti) e periodi di bassa volatilità (dove si registrano pause fra gli scambi). La conoscenza di queste variazioni è molto importante per gli investitori che richiedono rendimenti più elevati come compensazione per detenere attività più rischiose.

1.2 Variabili d'interesse per lo studio di serie storiche finanziarie

Per un'adeguata analisi dei dati ad alta frequenza è opportuno definire alcune variabili di interesse.

1.2.1 Tempo

Una delle tendenze più moderne della finanza consiste nello studio di osservazioni registrate in tempo reale sui mercati. A differenza degli studi tradizionali, nei quali si tende a considerare dati misurati ad intervalli equispaziati e a bassa frequenza, dette serie omogenee, nel caso dell'alta frequenza si registra ogni singola transazione che avviene sul mercato.

La modernizzazione degli strumenti informatici, che hanno permesso raccolta, gestione e archiviazione di un'enorme quantità di dati, e lo sviluppo di moderne teorie sulla microstruttura del mercato sono coincise con innovazioni nel campo della finanza. L'impiego di tale quantità di dati e la conseguente quantità di informazione a disposizione ha dimostrato che tralasciando i movimenti infra-giornalieri del prezzo si trascurano informazioni rilevanti ai fini della stima e della previsione. Si evidenzia, difatti, che il prezzo derivante da meccanismi di transazioni è molto sensibile allo scorrere del tempo facendo cadere l'assunzione che non ritiene significativa la variazione di prezzo in "piccoli" intervalli di tempo.

Una delle proprietà fondamentali dei dati ad alta frequenza è la loro spaziatura temporale: le osservazioni rilevate non sono necessariamente equispaziate nel tempo, le contrattazioni avvengono ad intervalli casuali tra loro perché sono collegate alla volatilità del prezzo.

Il tempo è uno strumento per capire l'attività finanziaria: lunghe durate di transazione sono indicative che il mercato sta aspettando segnali, informazioni, più rilevanti; mentre scambi continui indicano l'arrivo di notizie importanti.

1.2.2 Prezzo

Il prezzo al tempo t_j , $x(t_j)$, è definito come:

$$x(t_j) = [\log p_{ask}(t_j) + \log p_{bid}(t_j)] / 2 \quad (1.1)$$

dove t_j è la sequenza di passi temporali non ugualmente spaziat.

Il prezzo è la più importante variabile rilevata in finanza e dipende sia dalla struttura del mercato che dai fornitori dei dati.

La determinazione del prezzo avviene tramite gli agenti di cambio che concentrano, in un determinato momento, tutti gli ordini di compravendita affluiti dalla clientela: il prezzo *bid* (p_{bid}) è il prezzo al quale si è disposti a comprare il titolo, mentre il prezzo *ask* (p_{ask}) è il prezzo al quale si è disposti a venderlo. Il prezzo $x(t_j)$ è basato sulla media geometrica tra il prezzo *bid* e il prezzo *ask* piuttosto che sui singoli prezzi *ask* e *bid*: questo è il metodo migliore per approssimare il vero valore del prezzo.

Un risultato rilevante negli studi dei dati infra-giornalieri è l'uso di una scala temporale corretta. Contrariamente ai dati giornalieri e a bassa frequenza, i dati *tick by tick* sono spaziat irregolarmente nel tempo t_j , istante di tempo in cui avviene la transazione, momento nel quale si acquisisce l'informazione. La maggior parte delle analisi statistiche utilizza serie equispaziate nel tempo t_i , quindi per ottenere il valore del prezzo al

tempo t_i si utilizza il metodo dell'interpolazione lineare tra il prezzo all'istante precedente t_{j-1} e quello immediatamente successivo t_j , dove $t_{j-1} < t_i < t_j$. Il metodo dell'interpolazione lineare permette di costruire una serie omogenea regolarmente spaziata e, come citato in Muller et al. (1990), è il metodo più appropriato per un processo casuale che presenta incrementi identicamente e indipendentemente distribuiti.

1.2.3 Rendimenti

Il rendimento al tempo t_i , $r(t_i)$, è definito come:

$$r(t_i) = r(\Delta t; ; t_i) = [x(t_i) - x(t_i - \Delta t)] \quad (1.2)$$

dove $x(t_i)$ è la sequenza di logaritmi dei prezzi definiti nell'equazione 1.1 e Δt è un intervallo di tempo prefissato (5 minuti, 1 ora, 1 giorno, ...).

Il rendimento è solitamente considerato una variabile più adatta per le analisi del prezzo in quanto è una misura di successo di un investimento e gli investitori cercano di massimizzarlo nel periodo più breve e perché statisticamente la sua distribuzione presenta un andamento stazionario e più simmetrico di quello del prezzo.

1.2.4 Volatilità

Nella determinazione del prezzo di un'azione, il grado di incertezza dell'attività finanziaria è un elemento determinante, tanto che posso ricavare dalla velocità delle quotazioni le volatilità implicite, corrispondenti ad una valutazione della volatilità da parte dei mercati. La volatilità è considerata la principale misura del rischio e statisticamente è associata ad una misura di variabilità: la varianza (o scarto quadratico medio) dei rendimenti su un periodo storico.

La volatilità realizzata $v(t_i)$ al tempo t_i , è definita come:

$$v(t_i) = v(\Delta t, n, p; t_i) = \left[\frac{1}{n} \sum_{j=1}^n |r(\Delta t; t_{i-n+j})|^p \right]^{1/p} \quad (1.3)$$

dove i rendimenti equispaziati r sono definiti dall'equazione 1.2, e n è il numero delle osservazioni dei rendimenti. Ci sono due intervalli di tempo: l'intervallo di tempo prefissato tra i rendimenti, Δt , e il periodo campionario nel quale la volatilità è misurata, $n\Delta t$. L'esponente p è spesso posto pari a 2, tale che v^2 è la varianza dei rendimenti sotto lo zero, o a 1.

Un esempio, la misura della volatilità giornaliera come media della volatilità dei rendimenti giornalieri su un campione annuale: usando $p=2$, $\Delta t = 1$ giorno lavorativo, $n=250$ giorni lavorativi in un anno, $n\Delta t = 250$ periodo campionario.

Nei prossimi capitoli, saranno approfonditi questa variabile e alcuni stimatori della volatilità.

1.2.5 Spread

Lo spread relativo al tempo t_i , $s(t_i)$, è definito come la differenza tra le quote *bid* e *ask*:

$$s(t_j) = \log p_{ask}(t_j) - \log p_{bid}(t_j) \quad (1.4)$$

dove j è l'indice della serie temporale originale non omogenea.

Lo spread relativo è preferito all'equazione dello spread nominale,

$$s^*(t_j) = p_{ask}(t_j) - p_{bid}(t_j), \quad (1.5)$$

in quanto è un'equazione adimensionale. Gli spread relativi dai diversi mercati possono essere direttamente confrontati con i diversi cambi.

Questo indice relativo è una misura dell'efficienza del mercato: in borsa gli individui hanno preferenze dipendenti dal loro livello di informazione, lo spread serve all'operatore di borsa per compensare le perdite che subisce nel negoziare con gli individui informati e i guadagni che ottiene nello scambiare con i non informati.

1.3 Le caratteristiche statistiche dei dati finanziari ad alta frequenza

1.3.1 La distribuzione del processo dei rendimenti

Le serie temporali con dati ad alta frequenza nei mercati finanziari presentano distribuzioni leptocurtiche, implicando code più pesanti nei rendimenti di quelle di una distribuzione normale. Un campione di rendimenti con alta curtosi tende a contenere un numero più alto di valori estremi.

Per descrivere queste distribuzioni sono stati impiegati diversi modelli. Quello più adatto che è riuscito a cogliere le leptocurtosi dei rendimenti è stato il modello GARCH introdotto da Bollerslev (1986).

1.3.2 La memoria del processo

Una delle sfide poste dallo studio di dati ad alta frequenza nella finanza è di costruire modelli che possano spiegare empiricamente il comportamento dei dati con ogni frequenza, sia che questa sia misurata in minuti o in mesi.

I rendimenti di un'attività finanziaria sono generalmente incorrelati e non indipendenti. Questa dipendenza si manifesta con l'autocorrelazione nei momenti secondi dei rendimenti che, condizionatamente all'informazione disponibile, variano continuamente nel tempo. I ricercatori hanno introdotto nuove classi di modelli eteroschedastici ognuno dei quali ha un'opportuna equazione per la varianza condizionata σ_t^2 .

Per studiare la funzione di autocorrelazione della serie della volatilità e il suo raggrupparsi in periodi di alta volatilità e di bassa volatilità, sono stati sviluppati i modelli *Autoregressive Conditional Heteroskedastic Model* (ARCH) nel 1982 da Engle e la sua generalizzazione *Generalized*

Autoregressive Conditional Heteroskedastic Model (GARCH) nel 1986 da Bollerslev.

Il modello ARCH modella la serie temporale in questo modo:

$$y_t = x_t \beta + \varepsilon_t \quad (1.6)$$

dove x_t è l'insieme di variabili che influenzano la media condizionata di y_t e ε_t è il termine di errore con media zero e varianza condizionata σ_t^2 .

Questa varianza condizionata al tempo t è espressa come funzione lineare dei termini di errore ritardati al quadrato:

$$\sigma_t^2 = \gamma + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \quad (1.7)$$

con $\alpha_i \geq 0$ per assicurare la positività di σ_t^2 .

In questo modo la variabilità al tempo t dipende linearmente da quella rilevata i periodi precedenti, così larghe variazioni tendono essere seguite da larghe variazioni, e piccole da piccole variazioni. L'aspetto negativo dell'utilizzo dei modelli ARCH è che sono necessari un numero notevole di parametri α_i per poter catturare la lunga memoria che i dati manifestano. Il modello GARCH, generalizzazione del modello ARCH, supera questo problema riproducendo la parsimonia del modello ARMA rispetto alle rappresentazioni AR e MA in termini di numero di parametri utilizzati. Questo modello introduce i valori ritardati della varianza condizionata, è un modello in cui la varianza al tempo t è combinazione lineare di q ritardi dei residui al quadrato e di p ritardi della varianza condizionata, cioè:

$$\sigma_t^2 = \gamma + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (1.8)$$

Il modello è più parsimonioso e presenta una struttura dei ritardi più flessibile.

Gli studi su serie di dati ad alta frequenza evidenziano la presenza di memoria lunga per la volatilità.

Capitolo 2: Gli stimatori della volatilità

2.1 Introduzione

Lo studio della stima del rischio giornaliero, o volatilità, nei mercati finanziari è rilevante poiché il concetto di rischio è trattato come una variabile latente inosservabile e come un utile costrutto teorico per spiegare il comportamento dei rendimenti attivi.

Lo scopo di questo lavoro è di sviluppare un efficiente approccio per la stima della volatilità giornaliera utilizzando tutte le transazioni infra-giornaliere di cui si dispone. Trattando la volatilità giornaliera come un'osservazione, si ha un potente strumento per analizzare numerosi problemi nell'economia finanziaria.

Molti autori hanno utilizzato la nozione di rischio per stimare la volatilità giornaliera utilizzando dati ad alta frequenza. Beckers (1983), Anderson (1995), Parkinson (1980) e Rogers & Satchell (1991) hanno proposto stimatori della volatilità giornaliera utilizzando i prezzi di apertura, di chiusura, il prezzo più alto e quello più basso giornalieri. Schwert (1990), Hsieh (1991), Andersen e Bollerslev (1998) e Andersen et al. (2001) hanno proposto, invece, stimatori efficienti della volatilità giornaliera incondizionata basati su serie di rendimenti infra-giornalieri.

L'analisi dei rendimenti di serie temporali ad alta frequenza hanno mostrato come questi contengono preziose informazioni per la misurazione della volatilità a livello giornaliero e come, i modelli standard ARCH stimati, forniscono accurate previsioni della volatilità.

Inoltre, Andersen et al. (2001) utilizzano il concetto di volatilità realizzata per esaminare la *cross-correlation* nei vari rendimenti. Moosa e Bollen (2001) utilizzano la volatilità realizzata per verificare la relazione tra la volatilità e il tempo di scadenza nei mercati dei *futures* e per verificare l'errore sistematico nelle stime del *Value at Risk*.

2.2 Stimatori giornalieri

Sono riportati di seguito alcuni estimatori della volatilità giornaliera citati in letteratura.

Il rendimento giornaliero è definito come:

$$r_t = \ln(p_t) - \ln(p_{t-1}) \quad (2.1)$$

dove p_t è il prezzo di chiusura del giorno t e r_t è ottenuto da una distribuzione normale con volatilità variabile nel tempo tale che

$$r_t \sim N(0, \sigma_t^2). \quad (2.2)$$

Il nostro intento è di stimare il parametro della volatilità giornaliera σ_t .

2.2.1 Lo stimatore Simple

Lo stimatore Simple si basa sull'osservazione che il valore atteso della distribuzione normale è dato da $E[|r_t|] = \sigma_t \sqrt{2/\pi}$. Questo risultato permette di calcolare lo stimatore della volatilità semplice come:

$$\hat{\sigma}_{(Simple),t} = \frac{|r_t|}{\sqrt{2/\pi}} \quad (2.3)$$

Questo stimatore utilizza una sola osservazione per stimare ogni σ_t giornaliera e ciò suggerisce la sua inefficienza. Solo se l'equazione 2.2 è verificata, allora lo stimatore semplice della volatilità sarà non distorto.

2.2.2 Gli estimatori Extreme Value

Gli estimatori *Extreme Value* preparano alcuni usi delle informazioni nei dati infra-giornalieri, utilizzando i prezzi giornalieri di apertura e di chiusura e i valori estremi delle transazioni giornaliera cioè il prezzo più alto e quello più basso. Si definisce H_t il prezzo più alto e L_t il prezzo più basso del giorno t .

Parkinson, nel 1980, sviluppa lo stimatore della volatilità giornaliera PARK, basato sull'assunzione che i prezzi infra-giornalieri delle azioni seguono un processo generatore di tipo Browniano geometrico. Il corrispondente stimatore della volatilità giornaliera PARK, sotto questi assunti, è definito come:

$$\hat{\sigma}_{(PARK),t}^2 = \frac{(\ln(H_t) - \ln(L_t))^2}{4 \ln(2)} \quad (2.4)$$

Garman e Klass, nel 1980, presentano uno stimatore che essi affermano sia più efficiente di quello PARK. Anch'esso assume un processo generatore dei prezzi di tipo Browniano geometrico, e oltre al logaritmo del rapporto valori estremi considera anche i logaritmi del rapporto tra i prezzi di chiusura. Lo stimatore della volatilità GK è definito come:

$$\hat{\sigma}_{(GK),t}^2 = 0.5 \left(\ln \left(\frac{H_t}{L_t} \right) \right)^2 - 0.39 \left(\ln \left(\frac{P_t}{P_{t-1}} \right) \right)^2 \quad (2.5)$$

Andersen et al.(2001) propongono uno stimatore della volatilità che somma i quadrati dei rendimenti infra-giornalieri, lo stimatore della volatilità *Sum of Square Returns* (SSR) è definito come:

$$\hat{\sigma}_{(SSR),t}^2 = \sum_{i=1}^{n_t} r_{i,t}^2 \quad (2.6)$$

dove n_t è il numero dei rendimenti infra-giornalieri ed è l' i -esimo rendimento del giorno t di contrattazione. Lo stimatore SSR non richiede che la serie dei rendimenti infra-giornalieri sia omoschedastica, ma si basa sull'assunto che i rendimenti siano incorrelati. Per affrontare il problema delle autocorrelazioni nelle serie dei rendimenti infra-giornalieri, Andersen et al. (2001b) stimano un modello MA(1) sulla serie dei rendimenti infra-giornalieri e usano i residui di tale modello stimato nell'applicazione dell'equazione 2.6.

2.3 Modello di studio proposto

Gli stimatori descritti precedentemente si basano su assunzioni restrittive del processo generatore dei dati per poter essere utilizzati sui dati ad alta frequenza. Lo stimatore VARHAC (Bollen and Inder, 2002) proposto di seguito, invece, utilizza un set più povero di assunzioni per la stima della volatilità giornaliera

Precedenti studi hanno evidenziato certe caratteristiche dei rendimenti infra-giornalieri. Andersen e Bollerslev (1997), studiando i rendimenti infra-giornalieri ogni 5 minuti dello S&P 500 Index Futures, propongono una forma funzionale flessibile di Fourier come metodo appropriato per modellare la stagionalità infra-giornaliera della volatilità. Roll (1984) mostra come i prezzi fluttuano tra il bid e l'ask dello spread inducendo autocorrelazione di primo ordine nella serie dei rendimenti infra-giornalieri. Hasbrouck e Ho (1987) propongono un modello ARMA per il processo dei rendimenti infra-giornalieri, constatando "un'autocorrelazione negativa al primo ritardo, seguita da autocorrelazioni positive di grandezza decrescente statisticamente significative dal quinto ritardo". Admati e Pfleiderer (1988) sviluppano un modello teorico dove transazioni concentrate in un certo istante erano il risultato di un comportamento commerciale strategico da parte dei negozianti informati. Conseguenza di questo modello, applicato ai fini previsivi, è che a periodi prevedibili di volatilità infra-giornaliera, si presentano periodi di volatilità non prevedibile come risposta all'arrivo di nuove informazioni sul mercato.

Si conclude da questi studi che ogni giorno di transazioni è caratterizzato sia da periodi di volatilità prevedibile che non prevedibile. Conseguentemente, ogni giorno può essere caratterizzato da una differente struttura di autocorrelazione, a periodi di concentrata attività finanziaria seguono periodi di alta volatilità. La presenza di autocorrelazioni a ritardi maggiori a uno, rende inappropriati gli stimatori della volatilità giornaliera, visti nella sezione 2.2.

2.3.1 Lo stimatore VARHAC

Lo stimatore VARHAC considera queste proprietà dei rendimenti infra-giornalieri e usa tutti i dati infra-giornalieri disponibili.

Si assume l'osservazione di n_t rendimenti infra-giornalieri nel giorno t , si definisce $r_{i,t}^* = n_t^{1/2} r_{i,t}$ cosicché il rendimento giornaliero è calcolato come:

$$r_t = \sum_{i=1}^{n_t} r_{i,t} = n_t^{-1/2} \sum_{i=1}^{n_t} r_{i,t}^* \quad (2.7)$$

Il parametro da stimare è la varianza dei rendimenti giornalieri $\sigma_t^2 = E[r_t^2]$. Questo parametro può essere espresso come funzione dei dati infra-giornalieri:

$$\sigma_t^2 = E[r_t^2] = E\left[\left(\sum_{i=1}^{n_t} r_{i,t}\right)^2\right] \quad (2.8)$$

e può essere riscritta come:

$$\sigma_t^2 = \frac{1}{n_t} \sum_{i=1}^{n_t} \sum_{j=1}^{n_t} E[r_{i,t}^* r_{j,t}^*] \quad (2.9)$$

per calcolare questo parametro utilizziamo lo stimatore VARHAC proposto da den Haan e Levin nel 1996.

Prima di adottare questo stimatore è necessario verificare che soddisfi gli assunti richiesti per un valido utilizzo. Le condizioni, riferite da den Haan e Levin, sotto le quali lo stimatore VARHAC può essere utilizzato sono:

$$\sup_{i \geq 0} n_t E[r_{i,t}^2] < +\infty \quad (2.10)$$

$$\sigma^2 = \lim_{n_t \rightarrow \infty} \sum_{i=1}^{n_t} E[r_{i,t}^2] > 0 \quad (2.11)$$

$$\sum_{j=1}^{\infty} \sup_{i \geq 1} |E[r_{i,t} r_{i+j,t}]| < +\infty. \quad (2.12)$$

con $r_{i,t}$ sequenza di scalari a media zero.

La condizione 2.10 richiede che i rendimenti infra-giornalieri abbiano una varianza finita e la condizione 2.11 esclude la possibilità che troppi valori

abbiano varianza uguale a zero. La condizione 2.12 controlla i gradi di dipendenza dalle osservazioni: questo implica che i rendimenti $r_{i,t}$ sono asintoticamente indipendenti, e che la somma delle autocovarianze non sia costante nel tempo. Queste condizioni permettono di avere un grado di eterogeneità e autocorrelazione, e forniscono condizioni sufficienti per una valida stima della serie temporale utilizzando lo stimatore VARHAC.

Seguendo den Haan e Levin (1998), si applica lo stimatore VARHAC per ottenere una stima consistente ed efficiente di σ_t^2 usando tutte le informazioni infra-giornaliere di cui si dispone.

2.3.1.1 Costruzione dello stimatore

I passi per costruire questo stimatore sono i seguenti:

Step 1 Selezione degli ordini di ritardo per ogni giorno. Considerata una serie di osservazioni sulle transazioni dei rendimenti in uno specifico giorno lavorativo t , e avendo dati per un campione S di giorni di lavorativi, tale che $1 < t < S$. Per ognuno dei S giorni, si stima il seguente modello con il metodo dei minimi quadrati:

$$r_{i,t} = \sum_{k=1}^K \hat{\alpha}_{k,t}(K) r_{i-k,t} + \hat{e}_{i,t}(K) \quad \text{per } i = K + 1, \dots, n_t, \quad (2.13)$$

per ogni ordine di ritardo possibile $K=1, \dots, \bar{K}$. Per $K=0$ fissiamo $\hat{e}_{i,t}(0) = r_{i,t}$.

Poi si calcola la quantità:

$$BIC(K, t) = \ln \left(\frac{\sum_{i=k+1}^{n_t} \hat{e}_{i,t}^2(K)}{n_t} \right) + \frac{\ln(n_t)K}{n_t}. \quad (2.14)$$

Per ogni giorno, l'ordine di ritardo ottimale K_t è scelto come il valore di K che minimizza il BIC (K, t) .

Step 2 Calcolo della stima di σ_t^2 . Per ogni giorno, l'ordine di ritardo selezionato K_t e le stime di $\hat{\alpha}_{k,t}(K_t)$ e $\hat{e}_{i,t}^2(K_t)$, ricavate dall'equazione 2.13, possono essere usate per stimare:

$$\hat{\sigma}_t^2 = \frac{n_t \hat{\sigma}^2}{\left[1 - \sum_{k=1}^K \hat{\alpha}_{k,t}(K_t)\right]^2} \quad \text{con} \quad \hat{\sigma}^2 = \sum_{i=K_t+1}^{n_t} \hat{e}_{i,t}(K_t)^2 / n_t \quad (2.15)$$

Lo stimatore definito nell'equazione 2.15 è detto stimatore della volatilità VARHAC, $\hat{\sigma}_{(VARHAC),t}^2$.

Capitolo 3: La serie di dati

3.1 Introduzione

Lo scopo di questo studio è quello di analizzare la stima della volatilità giornaliera realizzata nei mercati finanziari con dati infra-giornalieri confrontando i diversi stimatori presentati nel capitolo 2. Questa applicazione avviene sulla serie temporale dell'indice *Standard and Poor's 500* (S&P 500), un indice che si riferisce ad un mercato nazionale, quello statunitense. La serie di dati è stata fornita dalla *Olsen Ltd* che le ha raccolte attraverso i circuiti telematici. Il periodo di riferimento per l'analisi va dal 01 Marzo 2000 al 28 Febbraio 2005 e i dati sono stati rilevati ogni 5 minuti secondo l'EST (*Eastern Standard Time*). Il tempo di scambio considerato è dalle 13:30 alle 20:30 per il periodo in cui vige l'ora solare e dalle 14:30 alle 21:30 per l'ora legale. Il dataset originale conteneva 525888 osservazioni ogni 5 minuti per l'intero arco giornaliero divise in 365 giorni per 5 anni. Per l'analisi compiuta abbiamo costruito un nuovo dataset di 100081 osservazioni ogni 5 minuti divisa in 1255 giorni lavorativi, si sono esclusi dall'originale i fine settimana e 50 giorni di festa infrasettimanale dove si verifica assenza di contrattazione, e i record delle transazioni al di fuori del tempo di scambio.

3.2 I mercati finanziari

I dati ad alta frequenza sono l'informazione diretta dai mercati. Le quantità quotate sono i prezzi, ma i mercati forniscono altre informazioni come il volume delle transazioni, non sono invece disponibili informazioni dettagliate sui partecipanti che spesso preferiscono mantenere l'anonimato.

I prodotti finanziari scambiati nei mercati finanziari si distinguono secondo l'attività finanziaria negoziata. Una distinzione principale è data dai mercati principali e i mercati dei prodotti derivati.

I mercati principali comprendono il mercato azionario e il mercato obbligazionario, entrambi sono mercati ad orizzonte nazionale ma differiscono in quanto il primo fornisce informazioni e serie storiche relative ai corsi azionari, ai volumi di scambio e agli indici sintetici di borsa; il secondo, invece fornisce informazioni sui prezzi delle obbligazioni e serie storiche per tassi d'interesse con varie scadenze. I mercati principali comprendono anche il mercato dei cambi su valuta, mercato ad orizzonte internazionale, che fornisce informazioni e serie storiche sui tassi di cambio.

Il mercato dei prodotti derivati comprende il mercato dei *futures* e il mercato delle opzioni. È il mercato dove sono negoziati prodotti finanziari il cui rendimento è determinato dall'andamento di un prodotto finanziario principale.

Un'altra classificazione per i mercati finanziari è tra mercati centralizzati nella forma di scambi o borse, e mercati decentralizzati interbancari (*over-the-counter*, OTC), dove i partecipanti individuali effettuano transazioni direttamente senza intermediari. I dati dai mercati OTC sono raccolti e forniti in tempo reale dai circuiti telematici, come *Reuters*, *Bloomberg*, o *Bridge*. I dati dai mercati centralizzati sono disponibili dalle fonti stesse e qualche volta direttamente dagli scambi. Gli archivi di dati storici sono raccolti in *database* da società specializzate nei dati ad alta frequenza, come *Olsen Ltd*, che inizialmente li riuniva in tempo reale da differenti fonti.

3.2.1 Il mercato azionario

I mercati azionari sono la maggior risorsa dei dati ad alta frequenza, poiché gli azionari scambiano titoli di diverso tipo. I dati ad alta frequenza sono principalmente prodotti durante le ore di apertura e di chiusura degli

scambi. In alcuni dei mercati principali, ci sono anche transazioni fuori del normale orario di apertura, che producono alcuni dati aggiuntivi.

I dati ad alta frequenza sono disponibili dai seguenti mercati azionari.

- Azioni di società individuali. Questi dati sono fortemente determinati dallo specifico comportamento di una singola azienda e da qualche tendenza generale del mercato e dell'economia. I dividendi disponibili e quelli pagati influiscono sui prezzi delle azioni, e la serie dei prezzi può essere capita considerando tutti questi eventi.
- Indici azionari. Sono calcolati come la somma dei prezzi di singole azioni concordate secondo una determinata formula. Il paniere delle azioni comprende importanti azioni di specifici Paesi o determinati settori industriali. Il paniere e i pesi sono adattati secondo i periodi, seguendo i cambiamenti dimensionali delle società. Gli indici azionari rappresentano un largo segmento di un'economia piuttosto delle azioni di società individuali e il loro comportamento è meno erratico dello studio di singole azioni. I dati ad alta frequenza per i principali indici azionari sono disponibili e interessanti oggetti di ricerca. Dovuti alle loro definizioni matematiche, essi mostrano spesso autocorrelazioni positive dei rendimenti a ritardi superiori ai 15 minuti. Questo potrebbe essere una conseguenza della struttura dei ritardi tra le principali azioni in testa e le azioni meno "liquide" del paniere.

3.3 L'indice S&P 500

La società di analisi finanziaria *Standard and Poor's* fornisce alcuni indici dell'andamento del mercato azionario statunitense, come lo *Standard and Poor's 500*. Questi indici attribuiscono ai titoli facente parti del paniere un peso pari al rapporto tra la capitalizzazione del titolo e la capitalizzazione complessiva del paniere dei titoli che entrano nell'indice.

L'indice S&P 500 è calcolato sulla base delle azioni, quotate presso il New York Stock Exchange (NYSE) e l'American Stock Exchange (ASE), di 500 titoli delle più importanti società americane ad elevata capitalizzazione. Le

500 azioni che compongono l'indice sono emesse da società operanti in diversi settori: industriale, energetico, finanziario, informatico e telecomunicazioni tra i principali. Questo indice è il più noto tra gli indici azionari di Wall Street, la Borsa di New York: le sue variazioni sono un parametro principale per giudicare l'andamento di una giornata borsistica. Inizialmente l'indice S&P 500, inaugurato il 4 Marzo del 1957, era composto di 425 imprese industriali, 25 del settore dei trasporti e 50 del settore pubblico e rappresentavano circa il 90% delle industrie quotate nella Borsa di New York. Dalla fine degli anni Ottanta, il paniere di riferimento dell'indice è cambiato, le 500 imprese sono periodicamente aggiornate secondo la loro importanza e del loro peso sul mercato. Difatti, questo indice è il principale indicatore dell'economia statunitense e molte delle compagnie, oggi, possono non essere più presenti a causa di fusioni o fallimenti, mentre altre sono entrate grazie allo sviluppo di nuovi settori che inizialmente erano inesistenti, si pensi al settore dell'informatica. L'indice S&P 500 è calcolato come media degli andamenti borsistici ponderata con il peso di ogni azienda, che è calcolato in funzione sia del numero di azioni in circolazione che delle contrattazioni giornaliere.

3.3.1 Analisi preliminare

Lo studio prende come riferimento lo S&P 500 per gli anni tra il 2000 e il 2005. Di seguito riportiamo i grafici della serie infra-giornaliera, dei rendimenti e dei rendimenti assoluti, e la tabella che riassume alcune statistiche preliminari, condotte sulla serie dei rendimenti.

Il grafico 1 mostra un andamento irregolare della serie e una volatilità erratica nel tempo, il grafico 2 evidenzia ampie fluttuazioni nella prima metà della serie mentre la parte terminale della serie ha una variabilità più stabile. Il grafico 3 mostra la volatilità.

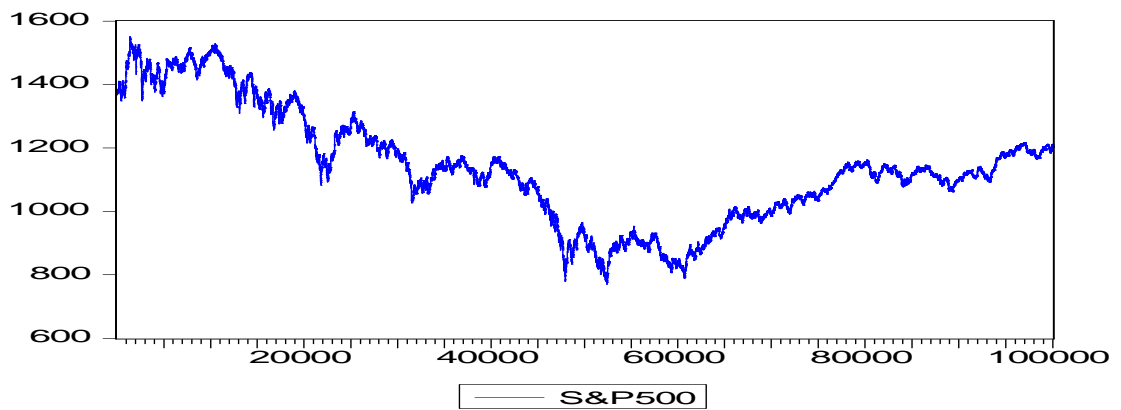


Grafico 3.1. Serie dell'indice S&P 500 (01/03/2000 – 28/02/2005)

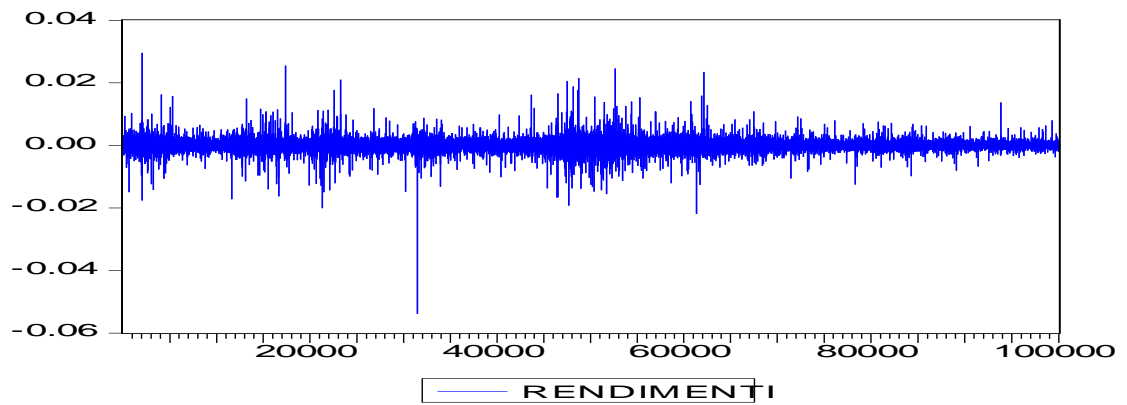


Grafico 3.2. Rendimenti della serie dell'indice S&P 500 (01/03/2000 – 28/02/2005)

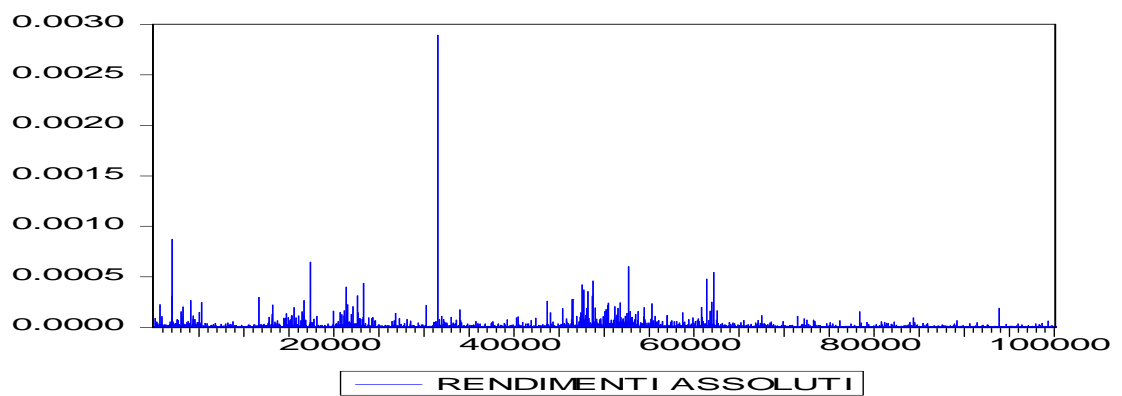


Grafico 3.3. Rendimenti assoluti della serie dell'indice S&P 500 (01/03/2000 – 28/02/2005)

La tabella 3.1, riportata di seguito, mostra le principali proprietà statistiche dei dati finanziari ad alta frequenza, quali la non normalità dei dati.

La non normalità dei dati è dimostrata dall'indice di simmetria, di curtosi e dal test di Jarque-Bera. L'indice di simmetria è calcolato come

$$SK = \frac{1}{T} \sum_{t=1}^T (y_t - \bar{y})^3 \quad \text{e l'indice di curtosi come: } k = \frac{\frac{1}{T} \sum_{t=1}^T (y_t - \bar{y})^4}{(\hat{\sigma}^2)^2}.$$

La maggior parte dei rendimenti infra-giornalieri mostrano leptocurtosi ($k > 3$) e spesso distribuzione asimmetrica ($SK \neq 0$). Nel nostro studio, la serie dei rendimenti presenta una leggera asimmetria, $SK = -0.427958$, e leptocurtosi, $k = 66.11163$. Il test di normalità Jarque-Bera è costruito basandosi sugli indici di simmetria e di curtosi campionari:

$$JB = \frac{T-1}{6} \left[SK^2 + \frac{1}{4} (\hat{k} - 3)^2 \right] \sim \chi^2_2.$$

Questo test verifica simultaneamente se la simmetria e la curtosi sono coerenti con i valori che dovrebbero assumere sotto l'ipotesi nulla di normalità. Dalla tabella, $JB = 1661.123$ e il test rifiuta chiaramente l'ipotesi di normalità per la serie in esame.

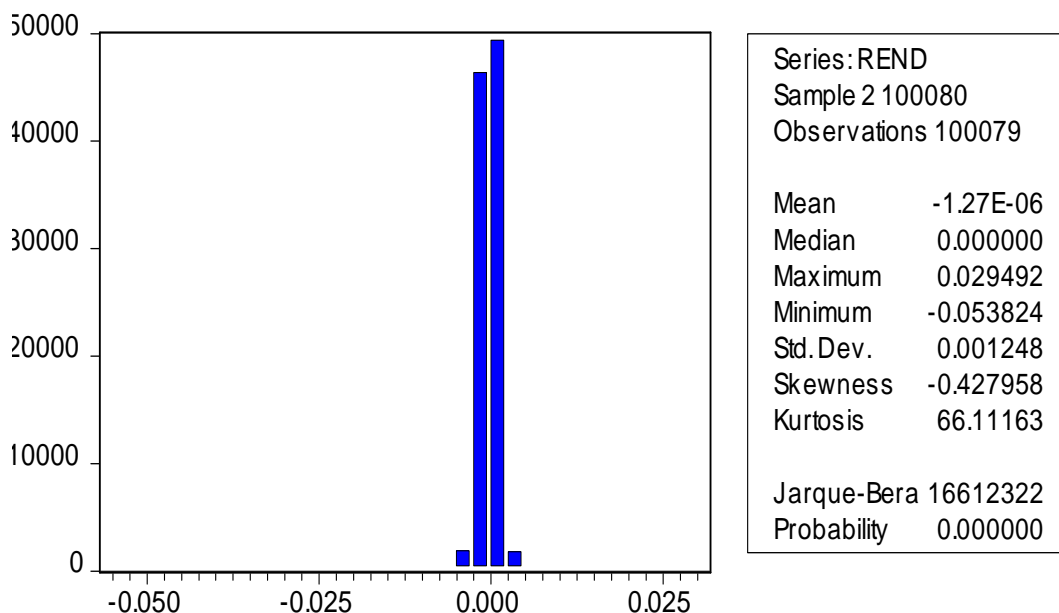


Tabella 3.1. Statistiche descrittive riassuntive della serie dei rendimenti

Un'altra proprietà dei dati finanziari ad alta frequenza è la correlazione seriale dei livelli nulla o prossima a zero; calcolando il correlogramma della

serie dei rendimenti: $\hat{\rho}_j = \left[\frac{1}{T-j} \sum_{t=j+1}^T (y_t - \bar{y})(y_{t-j} - \bar{y}) \right] / \hat{\sigma}^2$ con $j=0,1,\dots$ si

ottiene $\hat{\rho}_j$ che misura la “memoria” dei rendimenti, in altre parole la correlazione tra y_t e y_{t-1} . Si verifica la significatività di tale correlazione utilizzando il test di Ljung-Box:

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_r = 0$$

$$H_1 : \rho_1 \neq 0 \text{ and / or } \rho_2 \neq 0 \dots \rho_r \neq 0$$

$$Q = T(T+2) \sum_{j=1}^T \frac{\hat{\rho}_j^2}{(T-j)} \sim \chi^2_{(r)}.$$

Nel nostro caso, il grafico 3.4 (in appendice) mostra la funzione di autocorrelazione dei rendimenti e il test Q che conduce a rifiutare H_1 .

Inoltre, nelle serie con dati ad alta frequenza, spesso vi è la presenza di correlazione seriale nella serie dei rendimenti al quadrato, vedi grafico 3.5 (in appendice). Questo può essere verificato dal test di Ljung-Box che ci conduce a rifiutare H_0 . La correlazione tra i quadrati dei rendimenti è un sintomo della proprietà che la volatilità non è costante nel tempo.

Capitolo 4: Applicazione dello stimatore VARHAC all'indice S&P 500

in questo capitolo, saranno applicati gli stimatori della volatilità, presentati nel capitolo 2, ai dati delle transazioni dello S&P 500 per il periodo da marzo 2000 a febbraio 2005.

4.1 Analisi dei dati

I grafici delle volatilità giornaliere dei 1255 giorni di scambio sono stimate dai dati infra-giornalieri usando lo stimatore della volatilità giornaliera Simple, PARK, e GK. Lo stimatore SSR è stato calcolato in ogni giorno di scambio usando i dati delle transazioni ogni 5 minuti, i rendimenti infra-giornalieri sono stati filtrati stimando un processo MA(1).

La figura 4.1, riportata di seguito, mostra i grafici delle volatilità stimate compresa la stima della volatilità VARHAC. Un' ispezione visiva della figura 4.1 rivela che le stime della volatilità Simple sono molto più variabili degli altri stimatori, questo probabilmente è dovuto all'inefficienza di questo stimatore. Tutte le stime sembrano muoversi secondo una strada comune nel tempo, e sembrano distinguere distintamente i periodi di alta e bassa volatilità negli stessi periodi. Gli stimatori PARK e GK, di andamento simile, sembrano mostrare gradi di rumore più numerosi che gli stimatori della volatilità SSR e VARHAC. Questo è dovuto probabilmente perché gli stimatori della volatilità PARK e GK utilizza informazioni infra-giornaliere molto piccole, e conseguentemente sono anche inefficienti.

Le stime della volatilità giornaliera sembrano mostrare *outlier* il 16/01/2002, giorno di alta e inusuale volatilità, anche se il rendimento giornaliero di questi giorno non è alto.

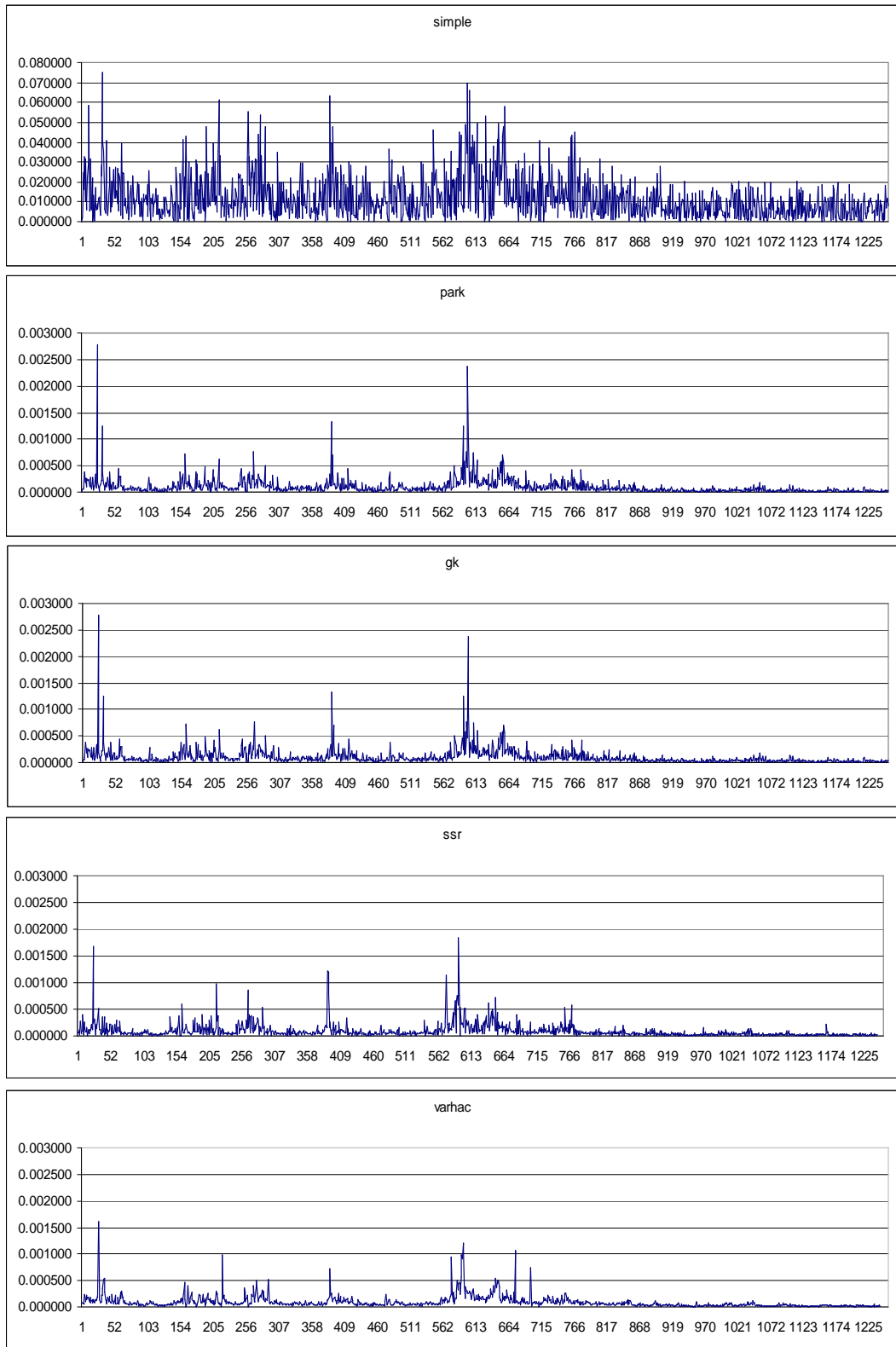


Figura 4.1. Stime della volatilità giornaliera dell'indice S&P 500 (01/03/2000 – 28/02/2005)

4.2 Bontà degli stimatori

Per valutare quali di queste stime sia la migliore, dato che la vera volatilità è inosservabile, si costruiscono i seguenti criteri di riferimento per verificare la bontà dei differenti stimatori.

4.2.1 Criterio 1: Confronto con le stime della volatilità semplice

Ricordando che lo stimatore della volatilità Simple è uno stimatore non distorto della volatilità giornaliera se i rendimenti giornalieri sono incorrelati, segue che la media delle stime della volatilità utilizzando lo stimatore Simple sui 1255 giorni di scambi può essere uno stimatore non distorto delle medie delle vere volatilità giornaliere σ_t . Conseguentemente, se un altro stimatore produce una media delle stime significativamente differente dalla media ottenuta usando lo stimatore della volatilità semplice, allora chiaramente questi altri stimatori mostrano un errore sistematico. Questa ipotesi è verificata dal confronto tra $\bar{X}_{(Simple)} = \frac{1}{S} \sum \hat{\sigma}_{(Simple),t}$ e $\bar{X}_{(.)} = \frac{1}{S} \sum \hat{\sigma}_{(.),t}$ usando un test semplice per la differenza di due medie, dove $\hat{\sigma}_{(.),t}$ è una delle stime della volatilità realizzata giornaliera proposti nel capitolo 2.

4.2.2 Criterio 2: Normalità dei rendimenti standardizzati

Il modello per i rendimenti giornalieri può essere scritto come $r_t = \sigma_t z_t$ dove $z_t \sim iid \mathcal{N}(0,1)$.

I rendimenti standardizzati per un dato stimatore della volatilità sono definiti come:

$$z_{(\cdot),t} = \frac{r_t}{\hat{\sigma}_{(\cdot),t}} \quad (4.1)$$

Se σ_t é stimata adeguatamente, $z_{(\cdot),t}$ potrebbe essere distribuita normalmente con varianza unitaria. Possiamo verificare la normalità utilizzando il test di Jarque-Bera sulle S osservazioni per $z_{(\cdot),t}$ per ogni stimatore scelto.

4.2.3 Criterio 3: Varianza dei rendimenti standardizzati

Dall'equazione 4.1, è chiaro che ogni errore sistematico inerente a uno stimatore della volatilità giornaliera dovrebbe produrre rendimenti standardizzati con varianza differente per ognuno. Un test è compiuto sulla varianza dei rendimenti standardizzati per ognuno degli stimatori della volatilità. La statistica test è data da:

$$\frac{\bar{Y}_{(\cdot)} - 1}{\hat{\sigma}_{Y_{(\cdot)}} / \sqrt{S}}, \quad (4.2)$$

dove $Y_{(\cdot)} = (z_{(\cdot),t} - \bar{z}_{(\cdot)})^2$, $\bar{Y}_{(\cdot)}$ e $\hat{\sigma}_{Y_{(\cdot)}}$ sono la media e la deviazione standard del campione di $Y_{(\cdot)}$. Questa statistica segue una distribuzione normale standard se la vera varianza è pari a 1.

La tabella 4.1, riportata di seguito, contiene i risultati per i dati dello S&P 500 rilevati per ognuno dei 3 criteri utilizzando tutti gli stimatori proposti. La statistica t nella terza riga dei risultati mostra che H_0 è rigettata per tutti gli stimatori della volatilità.

Il test di normalità di Jarque-Bera, condotto sui rendimenti standardizzati, suggerisce che gli stimatori PARK e SSR non producono rendimenti normali standardizzati. La condizione di normalità non può essere rifiutata per lo stimatore GK e per lo stimatore VARHAC.

L'ultima riga della tabella 4.1, indica che non si può rifiutare, ai limiti del livello di significatività, l'ipotesi che i rendimenti standardizzati per gli stimatori della volatilità PARK, GK hanno varianza pari a 1. Lo stimatore SSR ha una varianza significativamente diversa da 1, questo risultato indica che lo stimatore SSR è uno stimatore sistematicamente influenzato dalla volatilità giornaliera. Una conclusione di questo genere può essere spiegata notando che il processo filtrato MA(1) non si aggiusta adeguatamente per alti ordini di autocorrelazioni nelle serie di rendimenti delle transazioni. Anche lo stimatore VARHAC rifiuta l'ipotesi nulla.

Riassumendo, nel dataset in studio, nessun stimatore della volatilità passa sia il bias test (criterio 1 e 3) e il test di normalità condizionata (criterio 2). Lo stimatore della volatilità VARHAC, che passa solo il criterio 2, può comunque essere ritenuto uno stimatore robusto e efficiente, in quanto graficamente modella bene la serie facendo pieno uso dei dati dei rendimenti infra-giornalieri a disposizione.

	Stimatori della volatilità				
	Simple VARHAC	PARK	GK	SSR	
<i>Stime della volatilità</i>					
Mean	0.011635	0.009439	0.008817	0.008474	0.008689
Standard Deviation	0.010554	0.005384	0.004966	0.004858	0.004239
t-Test for difference between means		6.46	8.44	11.40	9.32
<i>Rendimenti Standardizzati</i>					
Mean		0.016820	0.036378	0.046727	0.023406
Standard Deviation		1.042204	1.390820	1.805644	1.436882
Skewness		-0.001901	0.083279	0.331387	0.155681
Kurtosis		1.678914	2.556632	8.637988	3.606033
Test Jarque-Bera		91.19107	11.72056	1561.627	24.19758
Z _(t) Test for unit variance		3.378692	3.389966	8.8795625	7.896805

Tabella 4.1. Applicazione dei tre criteri ai differenti stimatori della volatilità
I dati in grassetto indicano valori significativi al livello 1%

Capitolo 5: Conclusioni

5.1 Confronto tra gli stimatori considerati

Gli stimatori della volatilità presentati, nascono da uno studio di fatti stilizzati che riguardano la microstruttura del mercato, in particolare l'autocorrelazione e l'eteroschedasticità nei rendimenti infra-giornalieri. Da quest'analisi, è stato proposto un nuovo stimatore della volatilità giornaliera, lo stimatore VARHAC. Tutti gli stimatori della volatilità, proposti nei capitoli, sono stati applicati sullo S&P 500. Per valutare lo stimatore più adeguato sono stati proposti tre criteri. L'uso dello stimatore VARHAC è ragionevole perché impone pochi assunti sulla natura del processo generatore dei dati per i rendimenti infra-giornalieri, e fa uso di tutte le informazioni e di tutti i dati che si hanno a disposizione: utilizza tutta la popolazione e non un campione dei rendimenti infra-giornalieri. Le stime VARHAC basate sui rendimenti delle transazioni possono essere viste come la "vera" osservazione della volatilità.

Appendice

Di seguito sono riportati alcuni grafici discussi nei capitoli precedenti.

Capitolo 3. La serie dei dati

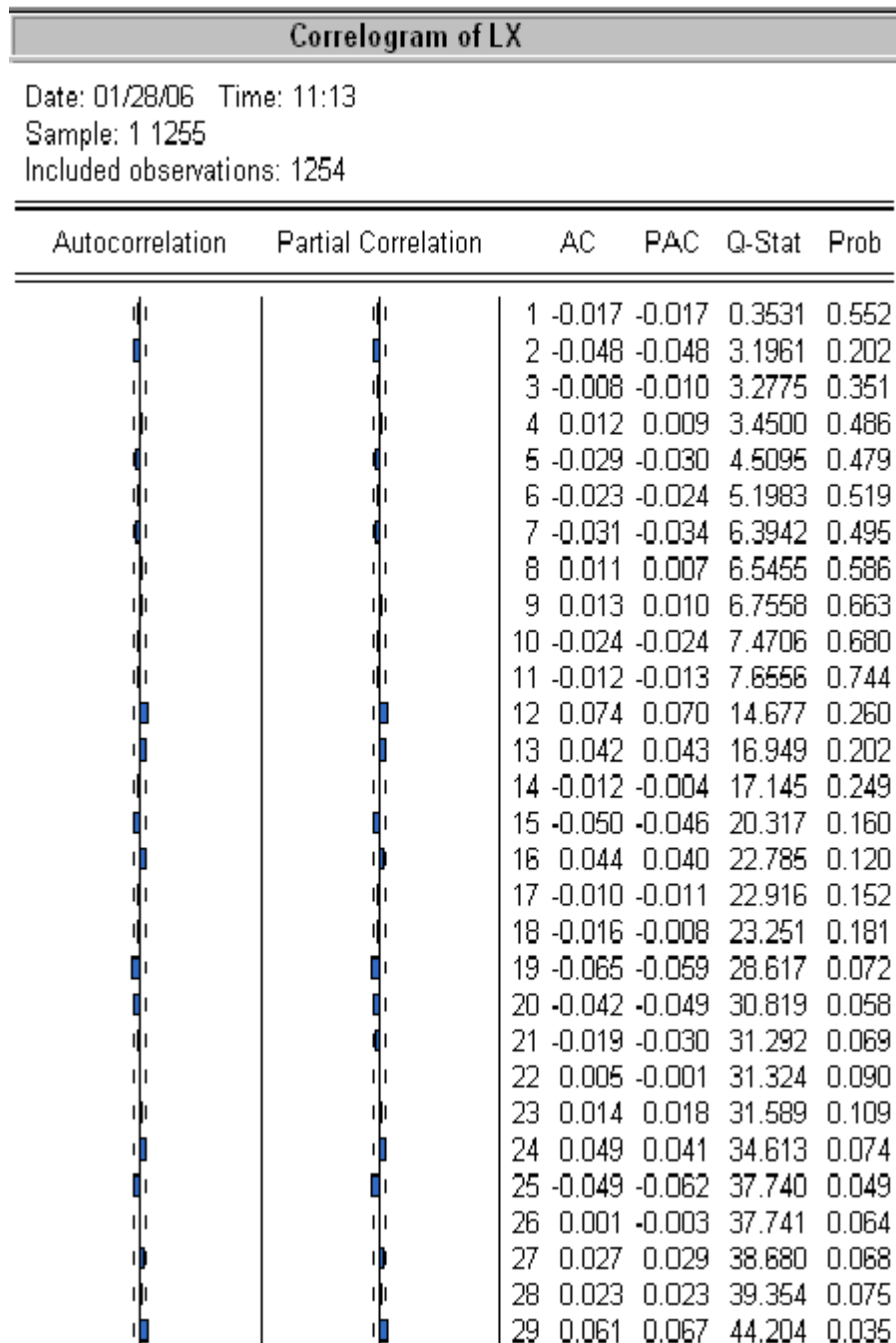


Grafico 3.4. Correlogramma della serie dei rendimenti dell'indice S&P 500 (01/03/2000 – 28/02/2005)

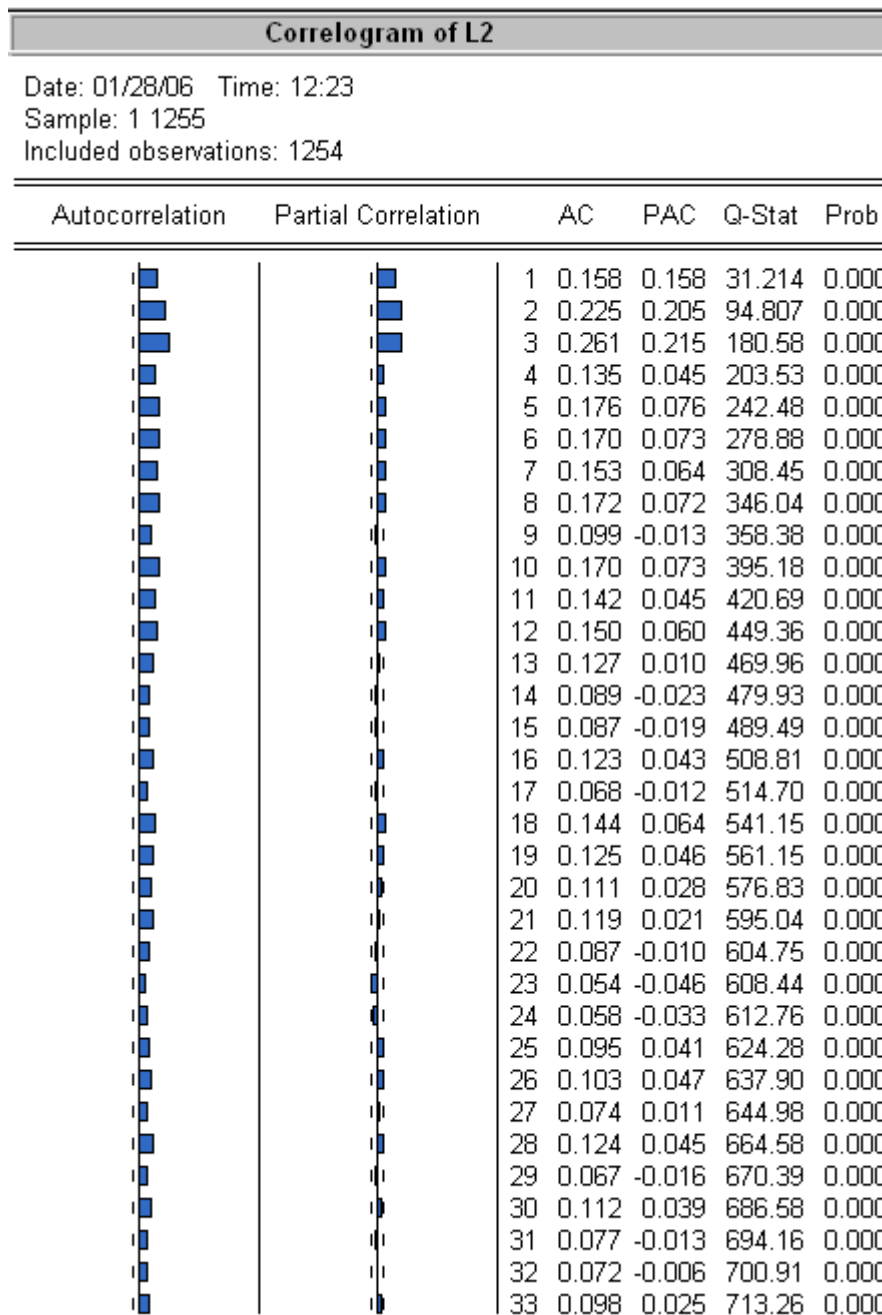


Grafico 3.5. Correlogramma della serie dei rendimenti al quadrato dell'indice S&P 500 (01/03/2000 – 28/02/2005)

Bibliografia

Bollen, B., Inder, B., 2002. Estimating daily volatility in financial markets utilizing intraday data. *Journal of Empirical Finance* 9, 551-562.

Daracogna, M., Gençay, R., Muller, U., Olsen, R. B., Pictet, O., 2001. *An Introduction to High-Frequency Finance*. Academy Press.

Gallo, G.M., Pacini, B., 2002. *Metodi quantitative per i mercati finanziari*. Carocci Editore.

Cecconi, M., Lombardi, M.J., 2001. Dati finanziari ad alta frequenza. *Scienza & Business*, Anno III, n. 9-10.

www.olsen.ch