

Dipartimento di Economia e Management

Cattedra di Statistica

Cosa è successo alla volatilità durante la crisi
finanziaria?
Analisi statistica attraverso il modello HAR-RV

Prof. Gianluca Cubadda

RELATORE

Antonio Domenichelli
Matr. 212051

CANDIDATO

Anno Accademico 2018/2019

Sommario

Introduzione	3
Capitolo 1: Caratteristiche delle serie finanziarie.....	5
1.1 Fatti stilizzati dei rendimenti	5
1.1.1 Assenza di autocorrelazione lineare	5
1.1.2 Distribuzione	6
1.2 Fatti stilizzati della volatilità	7
1.2.1 <i>Volatility Clustering</i> e non-stazionarietà	7
1.3 <i>Long Memory</i>	8
Capitolo 2: I modelli per la previsioni delle serie finanziarie.....	11
2.1 Il modello ARCH.....	11
2.2 La classe di modelli ARCH	13
Capitolo 3: Il modello	14
3.1 <i>Notation</i>	15
3.2 <i>Motivations</i>	17
3.3 Il modello HAR-RV	19
Capitolo 4: Analisi empirica.....	20
4.1 Database.....	20
4.2 Software e metodologia	21

4.3 Analisi.....	22
Conclusioni	27
Bibliografia	29
Indice delle figure	32

Introduzione

L'applicazione più diffusa delle serie storiche nel contesto finanziario è lo studio dei prezzi di determinati *assets*. Data l'incertezza dei prezzi futuri, occorre descriverli con distribuzioni probabilistiche. L'analisi diretta dei prezzi risulta molto difficile, poiché sono altamente correlati e sono soggetti a fluttuazioni rilevanti nel corso del tempo, quindi è più conveniente lavorare con i cambiamenti dei prezzi, ossia la volatilità dei prezzi diventa importante per capire come questi si comportano.

Nella letteratura recente è stata dimostrata l'importanza di adoperare con dati ad alta frequenza, e stimare da questi la volatilità infra-giornaliera come somma dei rendimenti al quadrato (quelli adoperati nella nostra analisi hanno una frequenza di dieci minuti). Questa misura prende il nome di *realized volatility*. Il modello che viene esaminato, elaborato da Fulvio Corsi, usa la *realized volatility* sotto forma di modello a cascata, prendendo spunto dalla *Heterogeneous Market Hypothesis*. La forza del modello in questione risiede nella sua capacità di comportarsi come un modello di memoria lunga, senza propriamente appartenere a quella famiglia.

Nel primo capitolo dell'elaborato, vengono esaminate le caratteristiche, chiamate *fatti stilizzati*, delle serie storiche finanziarie di rendimenti e della volatilità, nello specifico la loro distribuzione, l'assenza di auto-correlazione lineare, il *volatility clustering* e la *long memory*. Nel secondo capitolo, si fa

riferimento ai principali modelli di previsione per le serie storiche finanziarie. Nel terzo capitolo viene spiegato il modello in questione e le motivazioni da cui si è partiti per la sua elaborazione. Nell'ultima parte viene proposta un'analisi empirica per individuare i cambiamenti nella struttura di *long memory* e analizzare, attraverso il modello, i comportamenti dei diversi agenti del mercato.

Capitolo 1: Caratteristiche delle serie finanziarie

Per comprendere il significato del modello che andremo ad analizzare, è importante esaminare quali sono i principali *fatti stilizzati* delle serie storiche finanziarie, ossia le evidenze empiriche che si sono dimostrate così importanti da venire accettate trasversalmente nella letteratura.

Da un punto di vista strettamente statistico, si nota come quelle che possono sembrare variazioni casuali in realtà hanno in comune diverse proprietà statistiche, anche confrontando rendimenti di *assets* molto diversi tra loro.

Queste proprietà comuni sono chiamate *fatti stilizzati* e sono analizzate in questo capitolo.

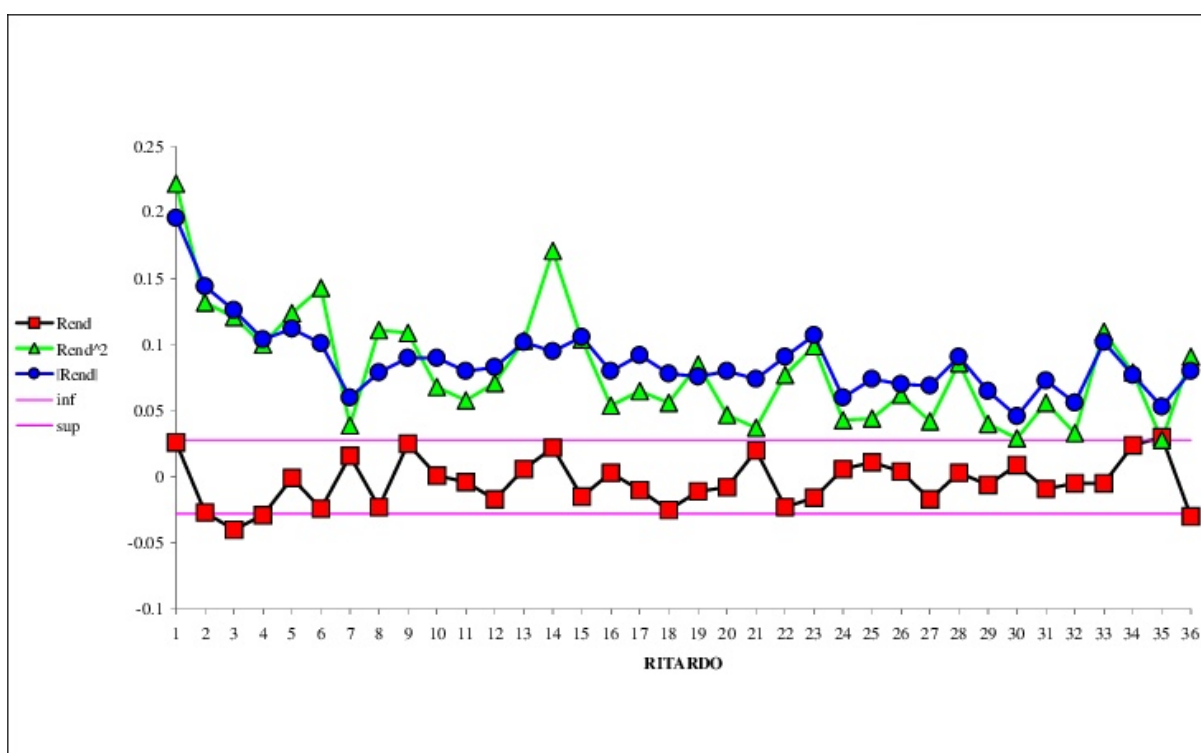
1.1 Fatti stilizzati dei rendimenti

1.1.1 Assenza di autocorrelazione lineare

Nell'analisi statistica delle serie finanziarie, l'attenzione è quasi sempre rivolta alle variazioni dei prezzi e quindi ai rendimenti e non ai prezzi in sé, dal momento che questi ultimi non sono stazionari. Ciò nonostante, numerosi studi hanno provato a verificare la presenza di autocorrelazione nei rendimenti finanziari (SEWELL, 2011), arrivando a definire l'autocorrelazione dei rendimenti non troppo significativa. Tuttavia, considerando intervalli temporali inferiori, i

rendimenti in serie tendono a mostrare un'autocorrelazione negativa (in tal caso, andrebbero analizzati fattori altri, e.g. comportamenti degli investitori *et alia*).

Al contrario, invece, l'autocorrelazione dei valori assoluti e dei quadrati dei rendimenti risulta essere più significativa e mostra una forte persistenza che perdura nel tempo (Corsi, 2009).



1 Autocorrelazione dei rendimenti, dei valori assoluti e dei quadrati dei rendimenti. (Technical indicators and volatility in financial time series - Scientific Figure on ResearchGate., s.d.)

1.1.2 Distribuzione

La distribuzione dei rendimenti, considerando differenti orizzonti temporali, mostra *fat tails* e *tail crossover*; ciò vale a dire che la funzione di densità di probabilità dei rendimenti di cui sopra è leptocurtica (CORSI, 2009). Nei primi studi effettuati sulle distribuzioni dei rendimenti, queste sono considerate

approssimativamente normali. Con il tempo, il numero degli studi effettuati nei confronti di tale tematica è aumentato, ed hanno confermato la presenza di un elevato indice di curtosi che cresceva proporzionalmente con l'aumento della frequenza di dati.

Considerata l'elevata curtosi, possiamo concludere che i modelli di distribuzione normale non sono più adatti a rendere ragione circa la distribuzione dei ricavi finanziari, né tantomeno a giustificare situazioni ove la probabilità di ottenere rendimenti più distanti rispetto alla media sia maggiore che alla distribuzione normale. Infatti, l'ipotesi di rendimenti distribuiti secondo una Gaussiana standard è stata rigettata e sostituita con l'ipotesi di distribuzioni *heavy-tailed* (MCALEER, 2008).

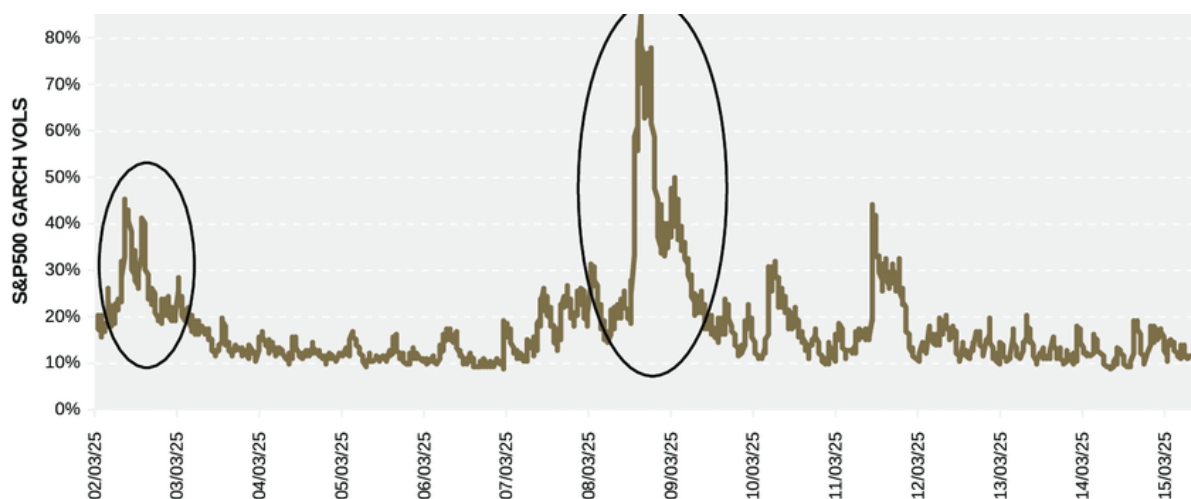
1.2 Fatti stilizzati della volatilità

1.2.1 *Volatility Clustering* e non-stazionarietà

Un ulteriore *fatto stilizzato* dei rendimenti finanziari è la loro non-stazionarietà. La stazionarietà è una proprietà delle serie storiche, tale per cui vengono soddisfatte determinate condizioni di invarianza (WALTER, 2015), ovvero se le sue proprietà statistica non variano nel tempo.

Nelle serie storiche finanziarie, sebbene la media dei rendimenti sia vicina allo zero, lo scarto quadratico subisce un fenomeno chiamato *volatility clustering*, ossia che ampie variazioni tendono ad essere seguite da altrettanto ampie

variazioni, e che, analogamente, variazioni meno ampie tendono ad essere seguite da variazioni meno ampie (MANDELBROT, 1967). Questo fenomeno è collegato all'autocorrelazione dei quadrati del paragrafo 1.1.1; variazioni di prezzo in termini assoluti correlate tra loro implicano che ampie variazioni sono susseguite da ampie variazioni, portando al fenomeno del *volatility clustering* appena descritto. Questa proprietà empirica è stata dimostrata da (GRANGER, 1996).



2 S&P 500 Volatility Clustering (Currency Hedge Design: Accounting for Uncertain Correlation and Volatility, s.d.).

1.3 Long Memory

Formalmente, un processo di memoria lunga si definisce tale se:

Definizione: se $\rho(k)$ è il coefficiente di correlazione al lag k , sia X_t un processo stazionario per cui è valido ciò che segue. Esiste un numero reale $\alpha \in (0, 1)$ e una costante $c_p > 0$ tale che:

$$\lim_{k \rightarrow \infty} \frac{p(k)}{c_p k^{-\alpha}} = 1$$

Allora X_t prende il nome di processo stazionario a memoria lunga, a dipendenza a lungo raggio o a forte dipendenza (SEWELL M. , 2011).

Sostanzialmente, la principale differenza tra processi a memoria corta e processi a memoria lunga sta nella loro funzione di autocorrelazione: processi a memoria corta sono rappresentati da funzioni di autocorrelazione con un decadimento esponenziale; mentre, al contrario, processi a memoria lunga tendono ad avere un decadimento iperbolico.

L'analisi della presenza di memoria lunga nelle serie finanziarie rappresenta uno studio chiave nell'ambito dei modelli per la loro previsione e, col tempo, ha portato allo sviluppo di diversi processi alternativi in grado di considerarla. GRANGER e DING hanno mostrato come non solo i processi *fractionally integrated* possono godere della proprietà di lunga memoria, bensì questa è suscettibile di essere rintracciata anche in altri processi (GRANGER, *Varieties of long memory models*, 1996). *Fractional integration* è un espediente matematico conveniente, ma manca di una chiara interpretazione economica. Difatti, i modelli frazionatamente integrati sono non banali da stimare e non facilmente estendibili a processi multivariati; ed infine, questi tipi di modelli non sono in grado di

riprodurre il comportamento *multiscaling* rinvenuto in alcuni dati empirici (CORSI, 2009)¹.

Esiste, inoltre, un approccio alternativo che considera le caratteristiche del comportamento di *long memory* (*multiscaling*) osservate nei dati come solo un comportamento apparente, generato da un processo che non gode della proprietà di lunga memoria, difatti un semplice modello additivo, e.g. la somma di tre modelli AR(1), mostra un *decaying memory pattern* tale da suscitare ambiguità se paragonato ad un andamento iperbolico (LEBARON, 2001).

Dal momento che questi ultimi (*simple component models*) sono più facili da stimare e da interpretare, il modello che andremo ad analizzare ha seguito questo punto di vista alternativo.

¹ Prove del comportamento *multiscaling* (*multifractal*) nei dati finanziari sono state trovate (anche se mancanti di un esplicito riferimento alla terminologia *multiscaling* o *multifractal*) da DING; GRANGER; ENGLE (1993), ANDERSEN; BOLLERSLEV (1997), LOBAT; SAVIN (1998), e DACOROGNA; GENÇAY; MULLER; OLSEN; PICTET (2001).

Capitolo 2: I modelli per la previsioni delle serie finanziarie

Nel capitolo precedente sono stati analizzati i principali *fatti stilizzati* delle serie finanziarie conosciuti fino al giorno d'oggi; in questo capitolo, verranno invece affrontati i modelli utilizzati per prevedere le serie storiche finanziarie, in particolare la loro volatilità, per ciò che nella fattispecie concerne i modelli ARCH.

2.1 Il modello ARCH

Il primo modello ARCH, ovvero *AutoRegressive Conditional Heteroskedasticity*, è stato elaborato da Engle. Il presupposto fondamentale è stato quello di analizzare il fenomeno della *volatility clustering* come una forma di eteroschedasticità nei rendimenti finanziari. L'elaborazione dei modelli ARCH ha aperto ad una nuova classe di processi eventualmente non lineari né in media né in varianza in grado di spiegare il fenomeno della volatilità variabile nelle serie finanziarie e le diverse peculiarità connesse. In questo modello, la varianza assume questo tipo di andamento, con ε_t termine di disturbo:

$$\sigma_t^2 = \alpha + \sum_{i=1}^m \beta_i \varepsilon_{t-1}^2$$

, con $\alpha > 0$ e $\beta_i \geq 0$.

Da questo modello si può subito notare che, se un certo shock ε_{t-1}^2 è elevato, questo implica che la varianza condizionata ad un istante successivo a (t-i) sarà elevata a sua volta, conseguendone che shock elevati tendono ad essere seguiti da altrettanto elevate variazioni di prezzo (in valore assoluto), descrivendo quel fenomeno già sottolineato dal nome di *volatility clustering*.

I vantaggi che derivano da tale modello rappresentano in primo luogo la possibilità di riprodurre il fenomeno della *volatility clustering*, oltre che la presenza di code grosse nella distribuzione.

È però necessario fare una considerazione: il modello assume che shock positivi o negativi abbiano lo stesso effetto sulla volatilità in quanto essa dipende dal quadrato degli shock precedenti; in pratica, invece, è noto nella letteratura finanziaria che questo non sempre corrisponde al vero, e molto spesso è presente una risposta diversa rispetto a shock positivi o negativi, soprattutto nei mercati azionari (PALOMBA, 2004).

2.2 La classe di modelli ARCH

Da questo modello iniziale (R.ENGLE, 1982) si sono sviluppati poi una serie di modelli che, prendendo spunto dal primo, si focalizzano su diversi aspetti delle serie finanziarie; per citarne alcuni: GARCH, IGARCH, EGARCH e FIGARCH.

Capitolo 3: Il modello

Viene proposto un *additive cascade model* con componenti di volatilità definite in diversi periodi di tempo. Questa *volatility cascade* porta ad un semplice modello di tipo AR², con la caratteristica di considerare le diverse componenti di volatilità realizzate su diversi orizzonti temporali, generate dalle azioni dei diversi partecipanti al mercato, e quindi chiamato modello *Heterogeneous Autoregressive model of Realized Volatility* (HAR-RV). Nonostante la semplicità della sua struttura e l'assenza di veri e propri legami a lunga memoria, i risultati della simulazione mostrano che il modello HAR-RV raggiunge con successo lo scopo di riprodurre le principali caratteristiche empiriche dei rendimenti finanziari (*long memory, fat tails*) e, malgrado non appartenga alla famiglia dei modelli di *long memory*, riesce a riprodurre la stessa persistenza della volatilità osservata nei dati (CORSI, 2009).

Inoltre, questo modello è caratterizzato da una struttura a più componenti, che rispecchiano la natura eterogenea dei comportamenti all'interno del mercato finanziario. Questo tema verrà ripreso nei paragrafi a seguire.

² Modello statistico per le serie storiche, che mette in relazione il valore presente di una variabile con i suoi valori ritardati, per tenere conto della possibile dipendenza statistica tra osservazioni corrispondenti a istanti diversi (LEORATO, s.d.).

3.1 Notation

In questa sezione viene analizzata la volatilità realizzata e i suoi aggregati nei diversi orizzonti temporali. La volatilità realizzata è un nuovo concetto crescente nella letteratura finanziaria. Esso deriva dalla varianza realizzata, ed è stata introdotta da BANDORFF-NIELSEN e SHEPPARD. Viene spesso utilizzato per misurare la variabilità dei prezzi dei rendimenti infra-giornalieri (O. E. BARNDORFF-NIELSEN, 2001).

La volatilità realizzata è una stima *ex-post* della variazione di rendimento. La misura della volatilità realizzata è la somma delle realizzazioni di rendimento al quadrato finemente campionate su un intervallo di tempo fisso (ANDERSEN B. , 2008; ANDERSEN B. D., 2003).

Per calcolare la volatilità realizzata, è necessario calcolare prima di tutto il logaritmo dei rendimenti dello stesso, ossia $r_t = \log(P_t) - \log(P_{t-1})$; si usa il logaritmo dei rendimenti per alcuni benefici sia teorici che in termini di algoritmi (R. HUDSON, 2010), dopodiché, la volatilità realizzata di N rendimenti è:

$$RV_t = \sqrt{\sum_{i=1}^N r_t^2}$$

Molti studi recenti evidenziano i vantaggi derivanti dall'uso di dati ad alta frequenza per calcolare la volatilità dei rendimenti finanziari; tra questi vanno citati, ANDERSEN; BOLLERSLEV; DIEBOLD (2001), BARNDORFF-NIELSEN;

NICOLATO; SHEPARD (2002). Ipotizziamo che volessimo calcolare la volatilità nell'arco di una giornata: in questi articoli si consiglia di studiarla attraverso la stima dei rendimenti infra-giornalieri al quadrato, prendendo come riferimento i rendimenti a cinque, dieci o trenta minuti (MEDDAHI, 2002). Diventa importante quindi la volatilità realizzata, poiché possiamo raccogliere i rendimenti infra-giornalieri all'interno della volatilità realizzata giornaliera.

Il nostro modello necessita di comparare diverse misure di volatilità realizzata definite su diversi orizzonti temporali più lunghi di quello giornaliero; queste ultime vengono “normalizzate”, ossia, semplicemente, si fa una media delle quantità giornaliere³.

Nella fattispecie, per ciò che riguarda tale modello, si rendono necessari i periodi di riferimento mensile e settimanale, calcolando di questi solo i giorni lavorativi; i.e. una settimana lavorativa corrisponde a cinque giorni; invece, un mese lavorativo corrisponde a ventidue giorni. Per indicare la volatilità realizzata settimanale, la simbologia è: RV_t^w ; mentre, invece, per quella mensile è: RV_t^m .

$$RV_t^{(w)} = \frac{1}{5} \sum_{i=1}^4 RV_i^{(d)}$$

³ Si noti che a causa della disuguaglianza di Jensen, la volatilità aggregata, come qui definita, non può essere interpretata esattamente come la volatilità realizzata nello specifico intervallo di tempo. Tuttavia, la differenza è irrilevante nelle applicazioni empiriche (CORSI, 2009).

$$RV_t^{(m)} = \frac{1}{22} \sum_{i=1}^{21} RV_i^{(d)}$$

3.2 Motivations

L'idea che traina il modello è la cosiddetta *Heterogeneous Market Hypothesis*, introdotta da MULLER nel 1993. I principi soggiacenti a tale ipotesi sono le proprietà della volatilità, ossia la sua correlazione positiva con la struttura del mercato (*presence, activity* e *volume*), e inoltre, la volatilità come indicatore della persistenza delle tendenze. Queste proprietà della volatilità portarono all'ipotesi di eterogeneità dei mercati, che contrastava l'assunto che tutti gli agenti del mercato reagissero alle notizie in maniera omogenea.

La *Heterogeneous Market Hypothesis* è caratterizzata dalle seguenti interpretazioni dei risultati empirici:

1. I diversi attori del mercato eterogeneo hanno diversi orizzonti temporali e frequenze di negoziazione. Dal lato delle alte frequenze di negoziazione vi sono gli operatori che di solito chiudono tutte le loro posizioni aperte prima della sera; dal lato delle basse frequenze di negoziazione sono le banche centrali, le organizzazioni commerciali e, ad esempio, gli investitori dei fondi pensione con le loro coperture valutarie. Le diverse frequenze di negoziazione indicano chiaramente reazioni diverse alle stesse notizie, nello stesso mercato. Il mercato è eterogeneo, con una struttura *fractal* degli orizzonti temporali dei

partecipanti, in quanto costituito da componenti a breve, medio e lungo termine. Ognuno di questi componenti si configura per un diverso tempo di reazione alle notizie, in relazione al proprio orizzonte temporale e alla caratteristica frequenza di negoziazione.

2. In un mercato omogeneo, più agenti sono presenti, più velocemente il prezzo dovrebbe avvicinarsi al "valore di mercato reale", sul quale tutti gli agenti con una "aspettativa razionale" sono d'accordo. Pertanto, la volatilità dovrebbe essere correlata negativamente con la *presence* e l'*activity* del mercato. In un mercato eterogeneo, è probabile che attori diversi si accontentino di prezzi diversi e decidano di eseguire le loro operazioni in situazioni di mercato diverse: i.e. creano volatilità. Ciò si riflette nella correlazione positiva empiricamente riscontrata della volatilità e della presenza sul mercato (U. MULLER, 1993).

L'eterogeneità nei mercati finanziari può insorgere per diversi motivi. Quello che però riguarda più da vicino il nostro modello è la differenza negli orizzonti temporali di riferimento, ossia operatori ad alte frequenze ed operatori a basse frequenze, a causa di ciò gli operatori creano diverse volatilità. Nello specifico possiamo individuarne tre tipi: 1) gli investitori a breve termine che lavorano con frequenze di *trading* giornaliera o ancora più alte, 2) gli investitori a medio termine che operano con frequenze settimanali e 3) gli operatori a lungo termine che hanno frequenze operative mensili o superiori.

Possiamo dedurre che gli investitori a breve termine saranno influenzati dalle decisioni e dalle posizioni degli operatori a lungo termine, perché essi determinano le tendenze e i rischi futuri che creeranno nuove volatilità a breve termine; d'altra parte, gli operatori a lungo termine non saranno influenzati da quelli a breve termine. Ciò che emerge è un modello di volatilità a cascata dalle basse frequenze alle alte frequenze (CORSI, 2009)⁴.

3.3 Il modello HAR-RV

Quindi, il modello HAR-RV è un *additive cascade* di volatilità parziali su orizzonti temporali di un giorno, una settimana e un mese che seguono un processo auto-regressivo. Il modello si presenta così:

$$RV_{t+1d}^{(d)} = c + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + \omega_{t+1d}$$

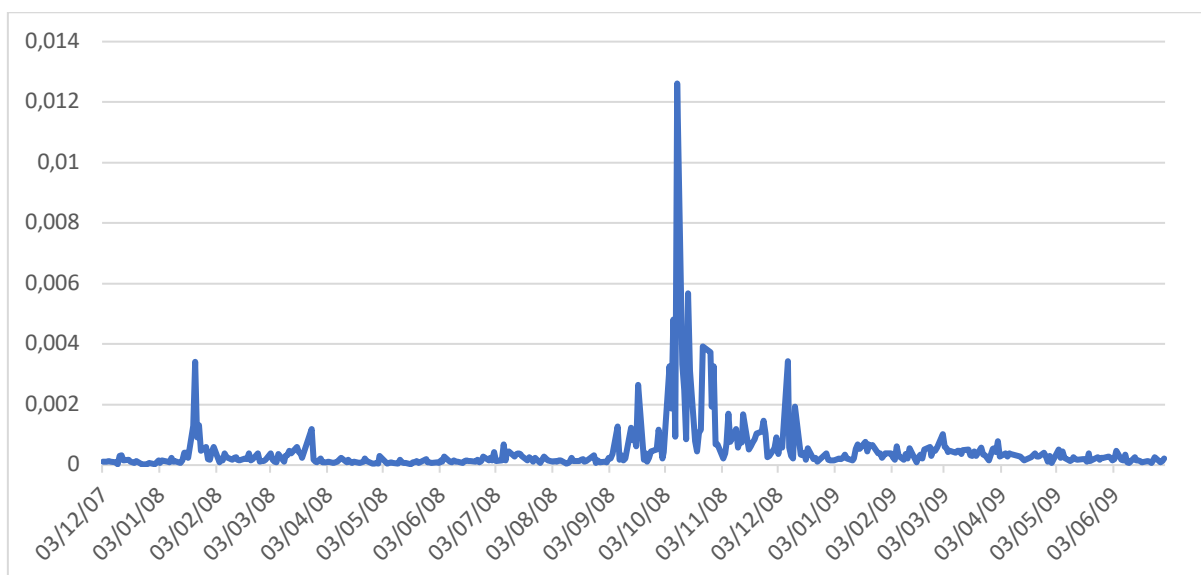
L'interpretazione economica è che ogni componente della *volatility cascade* corrisponde ad un agente del mercato che crea delle aspettative per la volatilità futura, basandosi sulle osservazioni della volatilità realizzata attuale. Rispetto ad altri modelli econometrici utilizzati per prevedere la volatilità futura, l'HAR-RV è relativamente semplice e parsimonioso in termini di potenza di modellazione della memoria di volatilità a lungo termine.

⁴ Prendendo a prestito dal modello KOLMOGOROV di turbolenza idrodinamica, processi a cascata moltiplicativi a cascata per la volatilità sono stati proposti da GHASHAGHAIE *et al.* (1996).

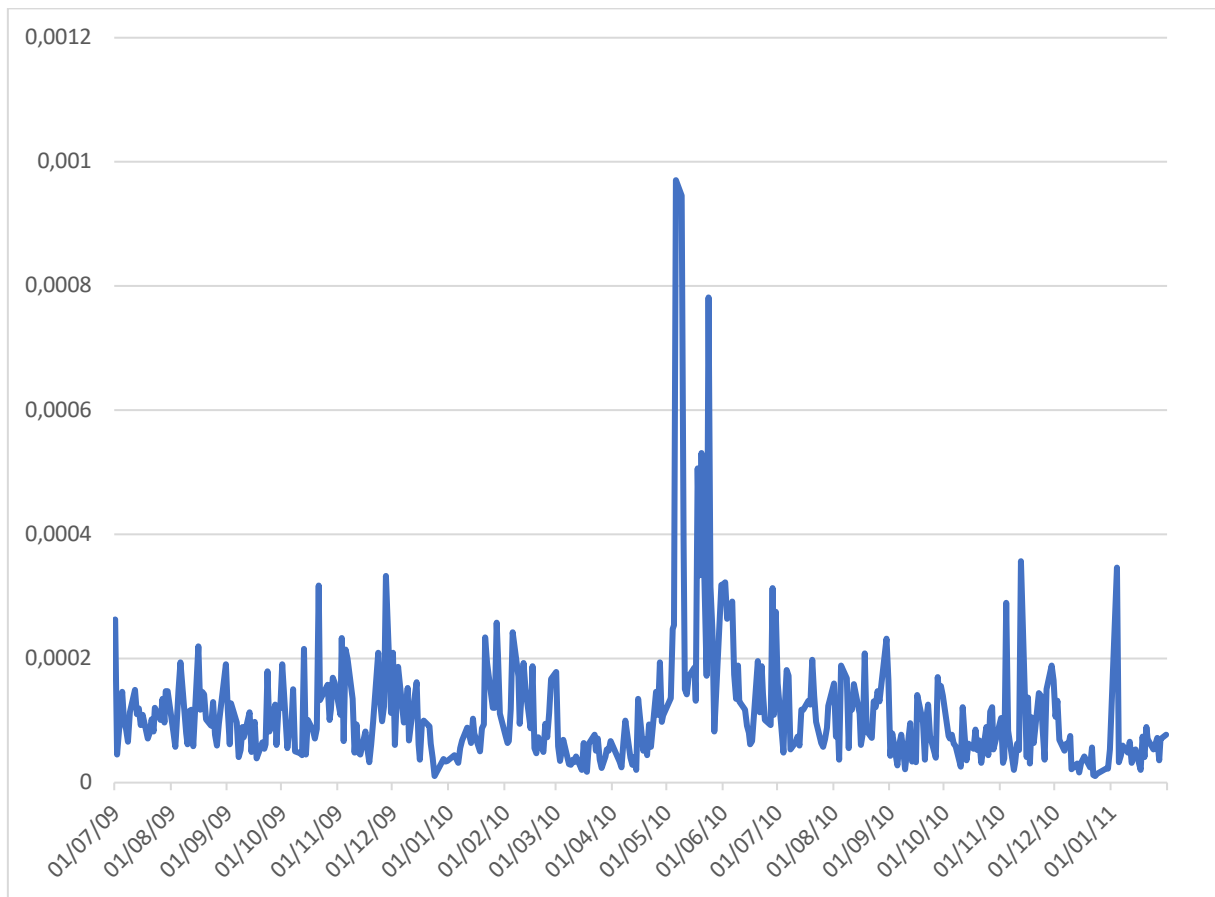
Capitolo 4: Analisi empirica

4.1 Dataset

I dati per l'analisi empirica sono stati presi dalla *Realized Library* dell'*Oxford-Man Institute of Quantitative Finance*. I dati che interessavano l'analisi riguardavano il periodo della crisi finanziaria e successivo; quindi, per quanto riguarda la crisi finanziaria, sono stati presi i dati dal dicembre del 2007 (03/12/2007) fino a giugno del 2009 (30/06/2009); come periodo successivo sono stati scelti i 18 mesi successivi (per avere due periodi di uguale misura), ovvero dal luglio del 2009 (01/07/2009) al gennaio 2011 (31/01/2011). Di questi periodi ci interessa la *Realized Volatility* a 10 minuti dei titoli FTSE.



3 RV a 10 minuti del FTSE100 durante la crisi finanziaria.



4a RV a 10 minuti del FTSE100 nei mesi successivi alla crisi finanziaria.

4.2 Software e metodologia

I software adoperati nell'analisi sono stati: *Gretl*, e il software *RStudio* in minor parte. Per verificare la presenza di cambiamenti nella struttura di *long memory* all'interno del modello HAR-RV, abbiamo effettuato un *test di Chow*. Questo test, che prende il nome dallo statistico GREGORY CHOW, misura la stabilità strutturale dei parametri di un modello stimato attraverso regressione lineare. Esso serve per verificare la presenza di un *break strutturale*, dal quale possiamo evidenziare un cambiamento netto dei parametri stessi della regressione. Materialmente si concretizza così:

poniamo t come ipotetica data di rottura e creiamo una *variabile dummy* "D(0)", "D(1)" e "D(2)" all'interno del periodo.

Allora, partendo dalla regressione:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon$$

la rimodifichiamo ottenendo:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \lambda_1 * D(0) + \lambda_2 * D(1)x_{1t} + \lambda_3 * D(2)x_{2t} + \varepsilon$$

verificando l'ipotesi nulla che i tre lambda siano uguali a 0. In caso negativo, saremo in presenza di una rottura strutturale nel panel. Nel nostro caso, alla *variabile dummy* creata è stato attribuito il valore 1 per il periodo della crisi finanziaria, ed il valore 0 per il periodo successivo. .

4.3 Analisi

Per effettuare il test di Chow ai nostri dati abbiamo riprodotto con il software *Gretl* il modello HAR-RV:

Modello 1: OLS, usando le osservazioni 2008-01-07:2011-01-31 (T = 777)
 Variabile dipendente: rv10

	coefficiente	errore std.	rapporto t	p-value	
const	0,135379	0,0588586	2,300	0,0217	**
LRV10_w	0,607797	0,0801985	7,579	1,00e-13	***
LRV10_m	0,186078	0,0809419	2,299	0,0218	**
rv10_1	0,0423876	0,0449599	0,9428	0,3461	
Media var. dipendente	0,000287	SQM var. dipendente	0,000667		
Somma quadr. residui	0,000231	E.S. della regressione	0,000546		
R-quadro	0,330964	R-quadro corretto	0,328368		
F(3, 773)	127,4648	P-value(F)	4,40e-67		
Log-verosimiglianza	4736,466	Criterio di Akaike	-9464,932		
Criterio di Schwarz	-9446,310	Hannan-Quinn	-9457,768		
rho	-0,000686	Durbin-Watson	2,001366		

Note: SQM = scarto quadratico medio; E.S. = errore standard

Escludendo la costante, il p-value è massimo per la variabile 11 (rv10_1)

5Modello HAR-RV con RVa 10 min

I dati raccolti sono stati immessi in una regressione lineare come quello descritta nel paragrafo 3.3:

$$RV_{t+1d}^{(d)} = c + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + \omega_{t+1d}$$

La volatilità realizzata sia settimanale che mensile è stata ricavata da *Gretl*, sono indicate nell'immagine come: LRV10_d, LRV10_w, e LRV10_m; il modello OLS è stato ricavato dall'apposita funzione di *Gretl*.

Dai dati nell'immagine 5, si evince che i tre aggregati di volatilità realizzata dei tre differenti orizzonti temporali sono tutti altamente significativi, fatta eccezione per il coefficiente del regressore giornaliero. Difatti, gli altri sono tutti significativi in un intervallo di confidenza del 5%.

Tramite il software *RStudio*, abbiamo poi creato una variabile *Dummy* che attribuisse il valore 1 al periodo della crisi e valore 0 al periodo successivo, e abbiamo moltiplicato la variabile *Dummy* per i valori di volatilità realizzata raccolti in precedenza; abbiamo quindi creato una regressione con questi nuovi valori. Questa ulteriore regressione viene costruita per effettuare il test di Chow e per fare un'analisi dei coefficienti.

Modello 2: OLS, usando le osservazioni 2008-01-07:2011-01-31 (T = 777)
 Variabile dipendente: rv10

	coefficiente	errore std.	rapporto t	p-value
const	0,148036	0,530919	0,2788	0,7805
dummy	-0,0450873	0,534488	-0,08436	0,9328
LRV10_w	0,166738	0,640845	0,2602	0,7948
LRV10_m	0,203512	0,730065	0,2788	0,7805
LRV10dummy_d	-0,355519	0,373863	-0,9509	0,3419
LRV10dummy_w	0,447285	0,645921	0,6925	0,4888
LRV10dummy_m	-0,0621005	0,734975	-0,08449	0,9327
rv10_1	0,390187	0,371107	1,051	0,2934
Media var. dipendente	0,000287	SQM var. dipendente		0,000667
Somma quadr. residui	0,000229	E.S. della regressione		0,000546
R-quadro	0,334702	R-quadro corretto		0,328646
F(7, 769)	55,26764	P-value(F)		5,16e-64
Log-verosimiglianza	4738,643	Criterio di Akaike		-9461,285
Criterio di Schwarz	-9424,042	Hannan-Quinn		-9446,958
rho	0,000068	Durbin-Watson		1,999841

Note: SQM = scarto quadratico medio; E.S. = errore standard

Escludendo la costante, il p-value è massimo per la variabile 2 (dummy)

6. Modello HAR-RV con RV a 10mi e la variabile Dummy

Abbiamo adesso la regressione per svolgere un'analisi sulla significatività dei coefficienti. Per effettuarla è stata adoperata la funzione di *Gretl* "TEST

OMIT”, nello specifico l’opzione “ELIMINAZIONE SEQUENZIALE DELLE VARIABILI USANDO *p-value* A DUE CODE”. Questa funzione del software permette di omettere le variabili dalla meno significativa fino a isolare le variabili significative al 10%, attuando sulla regressione un c.d. processo di riduzione. Si parte quindi isolando la variabile *dummy* con un *p-value* di 0,9328 (come si evince dalla figura 6) e così via. I risultati sono riportati di seguito:

Eliminazione sequenziale usando alfa (a due code) = 0,10

```
Scartata dummy          (p-value 0,933)
Scartata LRV10_w        (p-value 0,697)
Scartata LRV10dummy_m   (p-value 0,253)
Scartata LRV10dummy_d   (p-value 0,225)
Scartata rv10_1         (p-value 0,294)
```

Test sul modello 1

Ipotesi nulla: i parametri della regressione valgono zero per le variabili dummy, LRV10_w, LRV10dummy_m, LRV10dummy_d, rv10_1
 Statistica test: $F(5, 769) = 0,806997$, *p-value* 0,544778
 L'omissione delle variabili ha migliorato 3 dei 3 criteri di informazione.

Modello 2: OLS, usando le osservazioni 2008-01-07:2011-01-31 (T = 777)
 Variabile dipendente: rv10

	coefficiente	errore std.	rapporto t	p-value	
const	0,110226	0,0597164	1,846	0,0653	*
LRV10_m	0,151428	0,0821259	1,844	0,0656	*
LRV10dummy_w	0,646536	0,0611893	10,57	1,81e-24	***
Media var. dipendente	0,000287	SQM var. dipendente		0,000667	
Somma quadr. residui	0,000231	E.S. della regressione		0,000546	
R-quadro	0,331211	R-quadro corretto		0,329483	
F(2, 774)	191,6582	P-value(F)		2,44e-68	
Log-verosimiglianza	4736,609	Criterio di Akaike		-9467,219	
Criterio di Schwarz	-9453,252	Hannan-Quinn		-9461,846	
rho	0,040298	Durbin-Watson		1,919382	

Note: SQM = scarto quadratico medio; E.S. = errore standard

Test per l'aggiunta di variabili -

Ipotesi nulla: i parametri valgono zero per le variabili LRV10_w
 Statistica test: $F(1, 773) = 0,98158$
 con *p-value* = $P(F(1, 773) > 0,98158) = 0,322119$

Dai risultati che abbiamo ottenuto è possibile trarre le seguenti conclusioni: notiamo come la variabile significativa durante la crisi, nella figura indicata come LRV10_w, è la media settimanale della *Realized Volatility*. Ciò indica che in quel periodo era forte la presenza di investitori con orizzonti temporali di riferimento medio-brevi, i quali lavoravano con frequenze di negoziazione settimanali. Invece, al di là della crisi, ciò che risulta più significativo è il coefficiente della volatilità mensile: è forte la presenza di investitori istituzionali che hanno orizzonti temporali di riferimento più lunghi e quindi frequenze di negoziazione pressoché mensili.

Di conseguenza, possiamo dedurre che la crisi finanziaria ha portato dei cambiamenti nella struttura di *long memory* e, soprattutto, tramite questo modello, abbiamo notato come la crisi abbia modificato gli approcci al mercato da parte dei diversi attori. L'eterogeneità degli attori dei mercati finanziari è la base di partenza teorica di questo modello, difatti da quest'ipotesi deriva la sua struttura, come spiegato nei paragrafi precedenti.

Conclusioni

Il modello HAR-RV è ampiamente diffuso ed usato, poiché, come spiegato dallo stesso autore, in maniera parsimoniosa e non complicata riproduce il comportamento della volatilità a lunga memoria. Si ispira infatti all'ipotesi eterogenea del mercato, e porta ad un semplice modello di tipo AR di *realized volatility*, costruito come *additive cascade model*. Ciò non rientra formalmente nella categoria dei modelli a memoria lunga, però riproduce con successo la modellizzazione di quel tipo di volatilità. Così come scrive l'autore di questo modello, esso si presta benissimo alla propria estensione in varie dimensioni, rendendolo utile per ulteriori analisi e ulteriori regressori.

L'efficacia del modello è stata, naturalmente, già dimostrata. L'obiettivo di questo elaborato, invece, è quello di vedere cosa è accaduto nei mercati finanziari durante la crisi finanziaria. Il mercato di riferimento nel nostro caso è stato il FTSE; in particolare, per ciò che strettamente concerneva la *realized volatility* a 10 minuti di quest'ultimo. Il risultato interessante è stato il cambiamento di orizzonte temporale di riferimento; gli investitori istituzionali, durante la crisi finanziaria, hanno agito di meno. Il coefficiente di *realized volatility* più significativo è infatti risultato quello settimanale.

Per il conseguimento di questo risultato è stata essenziale la *Oxford-Manchester Realized Library*, dalla quale ho avuto accesso ai dati del mercato che

è stato analizzato, ed il software open-source *Gretl* che è stato adoperato per ottenere le regressioni e svolgere il test dei coefficienti significativi.

Bibliografia

Andersen, B. (2008). Realized Volatility. *Federal Reserve Bank of Chicago*.

Andersen, B. D. (2001). The Distribution of Stock Returns Volatilities. *Journal of Financial Econometrics*, 43-76.

Andersen, B. D. (2001a). The Distribution of Exchange Rate Volatility. *Journal of the American Statistical Association*, 42-55.

Andersen, B. D. (2003). Modelling and Forecasting Realized Volatility. *Econometrica*, 579-625.

Barndorff-Nielsen, N. S. (2002). Some Recent Developments in Stochastic Volatility Modelling. *Quantitative Finance*, 11-23.

Barndorff-Nielsen, S. (2001). Econometric analysis of realised volatility and its use in estimating stochastic volatility models.

Corsi. (2004). Simple Long Memory Models of Realized Volatility. *Working Papers*.

Corsi. (2009). A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 174-196.

Currency Hedge Design: Accounting for Uncertain Correlation and Volatility.

(s.d.). Tratto da ResearchGate.net: <https://www.researchgate.net/figure/S->

P-500-VOLATILITY-CLUSTERING-

90_fig4_314822126/actions#caption

Engle. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 987-1007.

Engle. (1993). Statistical Models for Financial Volatility. *Financial Analyst Journal*, 72-78.

Engle, G. (2007). A Multiple Indicators Model for Volatility Using Intra-day Data. *Journal of Econometrics*, 3-27.

Fama. (1965). The Behavior of Stock Market Prices. *Journal of Business*, 34-105.

Ghashghaie, B. P. (1996). Turbulent Cascades in Foreign Exchange Markets. *Nature* 381, 761-770.

Granger, D. (1996). Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics*.

Granger, D. (1996). Varieties of long memory models. *Journal of Econometrics*, 61-77.

Hudson, G. (2010). Calculating and Comparing Security Returns is Harder than you Think: A Comparison between Logarithmic and Simple Returns.

LeBaron. (2001). Stochastic volatility as a simple generator of apparent financial power laws and long memory. *Quantitative Finance*, 621-631.

Leorato. (s.d.). *Dizionario di Economia e Finanza*. Tratto da http://www.treccani.it/enciclopedia/modello-autoregressivo_%28Dizionario-di-Economia-e-Finanza%29/

Mandelbrot. (1967). The Variation of Some Other Speculative Prices. *The Journal of Business*, 393-413.

Mastro. (2014). *Forecasting Realized Volatility: ARCH-Type Models VS The HAR-RV Models*. Tratto da https://papers.ssrn.com/https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2519107

McAleer, M. (2008). Realized Volatility: A Review. *Econometric Review*, 10-45.

Meddahi. (2002). A Theoretical Comparison between Integrated and Realized Volatility. *CIRANO Working Papers*.

Muller, D. D. (1993). Facts and intrinsic time - A challenge to econometricians. *39th International AEA Conference on Real Time Econometrics*.

Oxford-Man Institute of Quantitative Finance Realized Library. (s.d.). Tratto da Oxford-Man Institute of Quantitative Finance: <http://www.oxford-man.ox.ac.uk>

Palomba. (2004). Modelli ARCH.

Poon, G. (2003). Forecasting Volatility in Financial Markets: a Review. *Journal of Economic Literature*, 478-539.

Sewell. (2011). Characterization of financial time series. *RN*.

Sewell, M. (2011). Characterization of financial time series. *RN*.

Stock, W. (2011). *Introduction to Econometrics*. Addison-Wesley.

Technical indicators and volatility in financial time series - Scientific Figure on ResearchGate. (s.d.). Tratto da ResearchGate:

https://www.researchgate.net/figure/Figura-4-Autocorrelazione-dei-rendimenti-dei-valori-assoluti-e-dei-quadrati-dei_fig2_266311905

Walter. (2015). Stationarity. In W. Enders, *Applied Econometric Time Series* (p. 51-55). Wiley.

1 Autocorrelazione dei rendimenti, dei valori assoluti e dei quadrati dei rendimenti. (Technical indicators and volatility in financial time series - Scientific Figure on ResearchGate., s.d.)	6
2 S&P 500 Volatility Clustering (Currency Hedge Design: Accounting for Uncertain Correlation and Volatility, s.d.).	8
3 RV a 10 minuti del FTSE100 durante la crisi finanziaria.....	20
4a RV a 10 minuti del FTSE100 nei mesi successivi alla crisi finanziaria.	21
5Modello HAR-RV con RVa 10 min	23
6. Modello HAR-RV con RV a 10mi e la variabile Dummy	24
7. OMIT TEST.....	25