

UN MODELLO GARCH MULTIVARIATO PER LA VOLATILITÀ DEI TASSI DI CAMBIO

Eduardo Rossi

Introduzione

La volatilità dei tassi di cambio è da alcuni anni al centro dell'analisi teorica ed empirica. I motivi di tale interesse sono essenzialmente due. Il primo è la grande rilevanza che il fenomeno ha assunto in questi ultimi tempi, grazie soprattutto alla crescente integrazione dei mercati finanziari internazionali. Il secondo è che la volatilità può essere interpretata come una misura del rischio di un'attività. Si deve aggiungere, inoltre, che i modelli strutturali costruiti negli anni scorsi per la spiegazione dell'andamento dei tassi di cambio nel breve e medio periodo, sono stati incapaci di cogliere l'accentuata variabilità (Baillie, MacMahon, 1989. Macdonald, Taylor, 1992). Una caratteristica, quest'ultima, non sorprendente se si tiene conto del fatto che questi modelli sono stati costruiti tenendo presente variazioni su orizzonti temporali più lunghi di quelli rilevanti per la volatilità.

L'interesse per la misurazione empirica della volatilità è stato accompagnato dalla valutazione delle misure teoriche della volatilità. In molti casi le misure di volatilità adottate nella letteratura empirica sono misure della variabilità osservata, di solito la varianza. Quest'ultima però non è altro che una semplice descrizione di una caratteristica osservata. Nei processi decisionali degli agenti, tuttavia, non entra la varianza non condizionata, bensì quella condizionata all'insieme disponibile, che misura il grado di incertezza temporale della variabile in considerazione (Lucas, 1982. Hodrick, 1989). La volatilità non è, inoltre, costante ma tende a variare nel tempo (Mandelbrot, 1963). Le analisi condotte sulle serie dei tassi di cambio hanno mostrato come queste siano caratterizzate da una sostanziale eteroschedasticità condizionale, nonlinearietà e distribuzioni che hanno code più spesse della distribuzione normale (leptocurtosi) (Boothe, Glassman, 1987. Hsieh, 1988). Questo significa che ci sono *cluster* di osservazioni con varianza più grande. Il processo *ARCH* (Autoregressive Conditional Heteroskedasticity) (Engle, 1982) e la sua generalizzazione, il processo *GARCH* (Generalized *ARCH*), costituiscono una parametrizzazione adatta a cogliere questo fenomeno. Le applicazioni del modello *ARCH* nelle sue diverse formulazioni, alle serie dei tassi di cambio sono ormai molto numerose.

Qui proponiamo un'applicazione del modello *GARCH* multivariato ad alcune serie di tassi di cambio settimanali della lira (lira-marco, lira dollaro, lira sterlina inglese) per il periodo 1973-1992. L'obiettivo è quello di stimare la volatilità, misurata con la varianza condizionata, dei tassi di cambio analizzati. In questo lavoro ripercorriamo, nel paragrafo 2, i risultati raggiunti nella letteratura sui modelli *ARCH* multivariati, presentando, nel paragrafo 3, un'applicazione alla stima della volatilità dei tassi di cambio sopra citati. Concludiamo, nel paragrafo 4, mostrando, con un esercizio, le capacità previsive di questi modelli.

Il modello GARCH multivariato

Il processo *ARCH* introdotto da Engle (1982) è uno dei possibili modi per modellare il comportamento delle serie finanziarie. Si tratta, infatti, di un processo che riproduce l'eteroschedasticità e la leptocurtosi delle serie finanziarie.

Il modello *ARCH* è un processo stocastico $\{\varepsilon_t\}$ con le seguenti caratteristiche:

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

$$(2) \quad z_t \text{ i.i.d.}, E(z_t) = 0 \quad Var(z_t) = 1$$

Per definizione ε_t è serialmente non correlato con media zero, ma la varianza condizionata di ε_t , uguale a h_t può variare nel tempo,

$$(3) \quad h_t = Var(\varepsilon_t | \Psi_{t-1})$$

dove Ψ_{t-1} rappresenta l'insieme informativo al tempo $t-1$. La varianza condizionata di ε_t , è una funzione variabile nel tempo, positiva e funzione dell'insieme informativo al tempo $t-1$.

Nella generalità delle applicazioni ε_t corrisponde all'innovazione nella media di un altro processo stocastico $\{y_t\}$:

$$(4) \quad y_t = g(x_{t-1}; b) + \varepsilon_t$$

dove $g(x_{t-1}; b)$ denota una funzione di x_{t-1} e del vettore dei parametri b , x_{t-1} appartiene all'insieme informativo al tempo $t-1$.

Una possibile specificazione di h_t è

$$(5) \quad h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 = \omega + \alpha(L) \varepsilon_t^2$$

cioè come funzione lineare dei valori passati del processo al quadrato dove $\omega > 0$ e $\alpha_i \geq 0$, dove L indica l'operatore ritardo. Questo modello è noto come modello *ARCH(q)* lineare. Il modello cattura la tendenza ai *cluster* di volatilità presente nelle serie finanziarie.

La prima estensione del modello *ARCH(q)* è stata il modello *GARCH* (Generalised *ARCH*) introdotto da Bollerslev (1986). Il processo *GARCH* è una struttura più parsimoniosa e flessibile del modello *ARCH*, infatti permette di esprimere un *ARCH* di ordine molto elevato usando un numero limitato di parametri e, di conseguenza, un numero ridotto di vincoli di non-negatività:

$$(6) \quad h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} = \omega + \alpha(L) \varepsilon_t^2 + \beta(L) h_t$$

dove $\alpha(0) = \beta(0) = 0$.

Al fine di assicurare l'esistenza di un processo ben definito tutti i parametri nella rappresentazione *AR* di ordine infinito

$$h_t = (1 - \beta(L))^{-1} \omega + (1 - \beta(L))^{-1} \alpha(L) \varepsilon_t^2.$$

devono essere non-negativi (Nelson, Cao, 1992, Drost, Nijman, 1993), dove si assume che le radici del polinomio $\beta(\lambda) = 1$ si trovino al di fuori del cerchio di raggio unitario.

Dato il modello :

$$(7) \quad y_t = g(x_{t-1}; b) + \varepsilon_t$$

$$(8) \quad \varepsilon_t = h_t^{1/2} z_t \quad z_t \sim i.i.d. D(0,1)$$

$$(9) \quad h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i}$$

D indica la distribuzione dei disturbi z_t . La distribuzione non condizionata di ε_t è caratterizzata da leptocurtosi, cioè da un eccesso di curtosì rispetto alla distribuzione normale. Nel caso di un *GARCH(1,1)*, ad esempio, la condizione:

$$3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2 < 1$$

implica l'esistenza del momento di ordine quarto di ε_t , che è pari a

$$E(\varepsilon_t^4) = 3\omega^2(1 + \alpha_1 + \beta_1)[(1 - \alpha_1 - \beta_1)(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2)]$$

il momento secondo è

$$E(\varepsilon_t^2) = \omega(1 - \alpha_1 - \beta_1)$$

Il coefficiente di eccesso di curtosi ($K-3$) è uguale a

$$K-3 = (E(\varepsilon_t^4) - 3E(\varepsilon_t^2)^2)E(\varepsilon_t^2)^{-2} = 6\alpha_1^2(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2)^{-1}$$

dove il denominatore è maggiore di zero per assunzione.

Il processo $GARCH(p,q)$ è stazionario in senso debole se e solo se

$$(10) \quad \sum_{i=1}^q \alpha_i + \sum_{i=1}^p \beta_i < 1 .$$

Quando

$$\sum_{i=1}^q \alpha_i + \sum_{i=1}^p \beta_i = 1$$

si ha il processo $IGARCH(p,q)$ (Integrated $GARCH$) ovvero esiste integrazione nella varianza condizionata. In tale modello la previsione, condizionata a Ψ_t , della varianza dipende da h_{t+1} . Il modello $IGARCH$, con o senza trend, fa parte di una più ampia classe di modelli, caratterizzati da persistenza nella varianza, nei quali l'informazione corrente rimane importante per le previsioni delle varianze condizionate future, per tutti gli orizzonti di previsione (Engle, Bollerslev, 1986; Lamoureux, Lastrapes, 1990; Nelson 1990).

Quando il processo $\{\varepsilon_t\}_{t=-\infty}^{\infty}$ è un vettore $(N \times 1)$, cioè $\varepsilon_t \in R^N$, si ha la versione multivariata del modello $GARCH$. Questo significa che si ha una formulazione $GARCH$ per ogni elemento della matrice varianza-covarianza condizionata. Si definisce processo multivariato $GARCH(p,q)$ ogni processo che permette la rappresentazione:

$$(11) \quad \varepsilon_t = H_t^{1/2} z_t$$

dove z_t è un vettore $(N \times 1)$ i.i.d., con $E(z_t) = 0$ $\text{Var}(z_t) = I$.

La matrice di varianze e covarianze H_t ($N \times N$) varia nel tempo, è definita positiva e misurabile con riferimento all'insieme informativo al tempo $t-1$. Questa matrice viene modellata secondo un processo $GARCH(p,q)$:

$$(12) \quad H_t = W + \sum_{i=1}^q A_i \varepsilon_{t-i} \varepsilon_{t-i}^\top A_i^\top + \sum_{i=1}^p B_i H_{t-i} B_{t-i}^\top.$$

Con questa scrittura appare chiaro che per assicurare la definitezza positiva di H_t è sufficiente che la matrice W sia definita positiva e che la matrice H_0 sia semidefinita positiva. Se supponiamo che W sia definita positiva allora applicando la scomposizione di Choleski il modello si può esprimere come:

$$(13) \quad H_t = CC^\top + \sum_{i=1}^q A_i \varepsilon_{t-i} \varepsilon_{t-i}^\top A_i^\top + \sum_{i=1}^p B_i H_{t-i} B_{t-i}^\top$$

Il numero di parametri, in questo caso, è uguale a $N(N+1)/2 + N^2(p+q)$. Per ottenere una scomposizione unica dobbiamo imporre che gli elementi della diagonale principale di C siano positivi. Il modello $GARCH(p,q)$ multivariato può essere anche espresso come:

$$(14) \quad vech(H_t) = W^* + \sum_{i=1}^q A_i^* vech(\varepsilon_{t-i} \varepsilon_{t-i}^\top) + \sum_{i=1}^p B_i^* vech(H_{t-i})$$

W^* è un vettore di dimensione $[(N(N+1)/2) \times 1]$, A_i^* e B_i^* sono matrici $[(N(N+1)/2) \times (N(N+1)/2)]$ e $vech$ indica l'operatore di vettorizzazione della parte triangolare inferiore di una matrice simmetrica ($N \times N$), tale da definire il vettore colonna $[(N(N+1)/2) \times 1]$. Il numero dei parametri da stimare in questo caso è uguale a $(1/2)N(N+1)[1+N(N+1)(p+q)/2]$.

Questa scrittura del modello $GARCH$ utilizzando l'operatore $vech$ non permette di analizzare quali siano le condizioni che assicurano la definitezza positiva della matrice H_t (Engle, Kroner, 1994).

E' possibile, però, risalire dalla rappresentazione (12) a quella vettoriale. Infatti fra le matrici dei coefficienti esistono le seguenti relazioni:

$$A_i^* = D_n^* (A_i + A_i^\top) D_n$$

$$B_i^* = D_n^* (B_i + B_i^\top) D_n$$

$$W^* = vech(CC^\top)$$

D_n è la matrice di duplicazione cioè:

$$\text{vec}(M) = D_n \text{vech}(M)$$

dove M è una matrice simmetrica e D_n^+ è l'inversa generalizzata della matrice di duplicazione (Duca, 1992).

L'estensione del modello GARCH al caso multivariato porta ad una crescita nel numero dei parametri da stimare, rendendo il modello difficilmente utilizzabile. E' necessario, quindi, introdurre alcuni vincoli. Questi devono essere compatibili con le condizioni di positività e permettere procedure di stima più snelle.

Un possibile approccio consiste nel considerare ogni elemento della matrice di varianze e covarianze condizionate come dipendente solo dai termini teorici o empirici con lo stesso indice cioè $h_{kl,t}$ dipende solo da $h_{kk,t-1}$ e da $\varepsilon_{k,t-1} \varepsilon_{l,t-1}$ (Bollerslev, Engle, Wooldridge, 1992). Questo vincolo si traduce con matrici A_i e B_i diagonali. Questa formulazione presenta due limitazioni (Gourieroux, 1990):

a) Il vincolo non è stabile rispetto alla composizione dei portafogli. Supponiamo che la serie osservata corrisponda ai guadagni in ogni periodo delle diverse attività e consideriamo un modello diagonale bidimensionale del tipo:

$$h_{11,t} = c_{11} + b_{11}h_{11,t-1}$$

$$h_{12,t} = c_{12} + b_{12}h_{12,t-1}$$

$$h_{22,t} = c_{22} + b_{22}h_{22,t-1}$$

Un portafoglio p costituito da α_1 unità della prima attività e da α_2 unità della seconda attività ha varianza condizionata uguale a:

$$h_{p,t} = \alpha_1^2 h_{11,t} + \alpha_2^2 h_{22,t} + 2\alpha_1\alpha_2 h_{12,t}$$

$$h_{p,t} = \alpha_1^2 c_{11} + \alpha_2^2 c_{22} + 2\alpha_1\alpha_2 c_{12} + \alpha_1^2 b_{11}h_{11,t-1} + \alpha_2^2 b_{22}h_{22,t-1} + 2\alpha_1\alpha_2 b_{12}h_{12,t-1}$$

Quest'ultima non può essere espressa come:

$$h_{p,t} = c_p + b_p h_{p,t-1}$$

non ha cioè la stessa forma delle varianze condizionali di ciascuna attività che entra nel portafoglio.

b) Il modello proposto non permette di prendere in considerazione le sostituzioni fra attività a livello del rischio. Per cui, ad esempio, il premio al rischio della prima attività è commisurato alla varianza condizionata h_{11} , la quale è indipendente dalla volatilità delle altre attività. Ciò significa che il rendimento della prima attività non è influenzato dal fatto che la seconda attività sia più volatile.

Adottando la parametrizzazione dell'eq.(13) ma imponendo matrici A_i , $i=1,\dots,q$, e B_i , $i=1,\dots,p$, diagonali il numero dei parametri si riduce a $N(N+1)/2+N(p+q)$. Con matrici A_i e B_i diagonali anche A_i^* e B_i^* sono diagonali, in particolare:

$$a_{ii}^* = a_{ii} a_{jj} \quad b_{ii}^* = b_{ii} b_{jj} \quad (1 \leq j \leq i \leq N)$$

(Magnus, 1988, p. 63).

Le procedure di stima dei parametri della media e della varianza condizionale sono basate sulla funzione di verosimiglianza associata alla legge normale. Questa verosimiglianza può essere utilizzata sia che la vera legge di distribuzione delle osservazioni sia normale sia che non lo sia (Gourieroux, Monfort, Trognon, 1984). In quest'ultimo caso sono diverse le procedure di stima delle covarianze dei parametri.

Consideriamo due processi multivariati, uno endogeno Y_t , $t=1,2,\dots$, di dimensione N ed uno esogeno Z_t , $t=1,2,\dots$ di dimensione L .

Definiamo con $\Psi_{t-1} = \{Z_t, Y_{t-1}, Z_{t-1}, Y_{t-2}, \dots, Y_1, Z_1\}$ l'insieme informativo disponibile all'inizio del periodo t . L'informazione disponibile all'inizio del periodo t è costituita allo stesso tempo dai valori presenti e ritardati delle variabili endogene e delle variabili esogene.

Si suppone che i primi due momenti condizionati appartengano a famiglie indicizzate da un numero finito K di parametri:

$$(15) \quad \begin{aligned} E(Y_t | \Psi_{t-1}) &= m_t(\Psi_{t-1}; \vartheta) = m_t(\vartheta) \\ Var(Y_t | \Psi_{t-1}) &= H_t(\Psi_{t-1}; \vartheta) = H_t(\vartheta) \end{aligned}$$

Supponiamo che la legge di distribuzione di $Y_t | \Psi_{t-1}$ non sia necessariamente normale, possiamo considerare la log-verosimiglianza corrispondente:

$$(16) \quad LogL_t(\vartheta) = \sum_{i=1}^T Logl_i(\vartheta)$$

dove T indica il numero di osservazioni e $Logl_i(\vartheta)$:

$$(17) \quad \begin{aligned} Logl_t(\theta) &= -\frac{1}{2} \log \det H_t(\theta) - \frac{1}{2} [Y_t - m_t(\theta)]' H_t^{-1}(\theta) [Y_t - m_t(\theta)] \\ Logl_t(\theta) &= -\frac{1}{2} \log \det H_t(\theta) - \frac{1}{2} \epsilon_t(\theta)' H_t^{-1}(\theta) \epsilon_t(\theta) \end{aligned}$$

dove $\epsilon_t(\theta) = Y_t - m_t(\theta)$ è il vettore degli errori teorici.

Lo stimatore di quasi-massima verosimiglianza è ottenuto massimizzando $LogL_t$, cioè esaminando le condizioni del primo ordine per un massimo:

$$(18) \quad \frac{\partial LogL_t(\hat{\theta}_t)}{\partial \theta} = \sum_{i=1}^T \frac{\partial Logl_t(\hat{\theta}_t)}{\partial \theta}$$

Introduciamo una notazione adatta al caso multivariato. Se $A(\theta)$ è una matrice (n,m) , θ_i è la i -esima componente del vettore θ , $i=1,..,k$, $\frac{\partial A(\theta)}{\partial \theta_i}$ rappresenta la matrice (n,m) i cui elementi sono le derivate degli elementi di $A(\theta)$ in rapporto a θ_i . Quando, invece, deriviamo la matrice $A(\theta)$ rispetto a tutte le componenti del vettore θ allora abbiamo

$$\frac{\partial A(\theta)}{\partial \theta} = \left[\text{vec}\left(\frac{\partial A(\theta)}{\partial \theta_1}\right), \dots, \text{vec}\left(\frac{\partial A(\theta)}{\partial \theta_k}\right) \right]$$

la quale è una matrice (nm,k) . Se m_t e H_t sono differenziabili rispetto a θ su tutto Ψ_{t-1} e se H_t è non singolare con probabilità uno per tutti i $\theta \in \Theta$, allora, utilizzando la nozione precedente, il gradiente corrispondente alla t -esima osservazione è:

$$(19) \quad \frac{\partial Logl_t(\theta)}{\partial \theta} = \frac{\partial m_t(\theta)'}{\partial \theta} H_t^{-1}(\theta) \epsilon_t(\theta) + \frac{1}{2} \frac{\partial H_t(\theta)'}{\partial \theta} [H_t^{-1}(\theta) \otimes H_t^{-1}(\theta)] \text{vec}[\epsilon_t(\theta) \epsilon_t(\theta)' - H_t(\theta)]$$

Gli stimatori ottenuti con il metodo della massima verosimiglianza hanno asintoticamente le proprietà standard di questa classe di stimatori. Per quanto riguarda gli stimatori di quasi massima verosimiglianza questi continuano ad essere consistenti ed asintoticamente distribuiti normalmente.

Se interpretiamo la serie dei gradienti valutati in corrispondenza di θ_0 (il valore vero del vettore dei parametri) come una differenza di martingala vettoriale allora avremo:

$$\begin{aligned}
 (20) \quad & E(\varepsilon_t(\vartheta_0) | \Psi_{t-1}) = 0 \\
 & E\{[\varepsilon_t(\vartheta_0)\varepsilon_t(\vartheta_0)' - H_t(\vartheta_0)] | \Psi_{t-1}\} = 0
 \end{aligned}$$

da cui si deduce:

$$(21) \quad E\left[\frac{\partial \text{Log}l_t(\vartheta_0)}{\partial \vartheta} | \Psi_{t-1}\right] = 0, \quad \forall t$$

Sotto ulteriori condizioni di regolarità sulle varianze condizionate 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, Wooldridge, 1992):

$$(22) \quad \sqrt{T}(\hat{\vartheta}_T - \vartheta_0) \xrightarrow{d} N(0, \Omega_0)$$

dove:

$$\begin{aligned}
 \Omega_0 &= A_0^{-1} B_0 A_0^{-1} \\
 A_0 &= E_{\vartheta_0} \left[-\frac{\partial^2 \text{Log}l_t(\vartheta_0)}{\partial \vartheta \partial \vartheta'} \right] \\
 B_0 &= E_{\vartheta_0} \left[\left(\frac{\partial \text{Log}l_t(\vartheta_0)}{\partial \vartheta} \right) \left(\frac{\partial \text{Log}l_t(\vartheta_0)}{\partial \vartheta} \right)' \right]
 \end{aligned}$$

In pratica le due matrici possono essere stimate in modo consistente rimpiazzando i valori attesi con le medie empiriche ed il valore vero del parametro ϑ_0 con lo stimatore $\hat{\vartheta}_T$. Queste approssimazioni sono:

$$\begin{aligned}
 \hat{A}_T &= -\frac{1}{T} \sum_{t=1}^T \frac{\partial^2 \text{Log}l_t(\hat{\vartheta}_T)}{\partial \vartheta \partial \vartheta'} \\
 \hat{B}_T &= \frac{1}{T} \sum_{t=1}^T \frac{\partial \text{Log}l_t(\hat{\vartheta}_T)}{\partial \vartheta} \frac{\partial \text{Log}l_t(\hat{\vartheta}_T)}{\partial \vartheta'}
 \end{aligned}$$

Ricordando che:

$$A_0 = E_{\theta_0} \left[-\frac{\partial^2 \text{Log} l_t(\theta_0)}{\partial \theta \partial \theta'} \right]$$

$$A_0 = E_{\theta_0} E_{\theta_0} \left[-\frac{\partial^2 \text{Log} l_t(\theta_0)}{\partial \theta \partial \theta'} \middle| \Psi_{t-1} \right]$$

Come stimatore consistente della matrice A_0 è stata proposta la quantità:

$$\tilde{A}_T = \frac{1}{T} \sum_{t=1}^T S_t(\hat{\theta}_T)$$

con

$$S_t(\theta_0) = E_{\theta_0} \left[-\frac{\partial^2 \text{Log} l_t(\theta_0)}{\partial \theta \partial \theta'} \middle| \Psi_{t-1} \right]$$

Nei nostri casi l'espressione esplicita di $a_t(\theta_0)$ è uguale a:

$$S_t(\theta_0) = \frac{\partial m_t(\theta_0)'}{\partial \theta} H_t^{-1}(\theta_0) \frac{\partial m_t(\theta_0)}{\partial \theta} + \frac{1}{2} \frac{\partial H_t(\theta_0)'}{\partial \theta} [H_t^{-1}(\theta_0) \otimes H_t^{-1}(\theta_0)] \frac{\partial H_t(\theta_0)}{\partial \theta}$$

Questa seconda approssimazione della matrice della matrice A_0 ha il vantaggio di essere calcolabile a partire dalle sole derivate prime dei momenti $\frac{\partial m_t(\theta_0)'}{\partial \theta}$ e $\frac{\partial H_t(\theta_0)'}{\partial \theta}$, non necessita, quindi, come \hat{A}_T del calcolo delle derivate seconde.

Un *GARCH* multivariato può essere persistente in varianza (Bollerslev, Engle, 1993). Indichiamo con Y_t il processo vettoriale stocastico ($N \times 1$) con i primi due momenti condizionali uguali a:

$$(23) \quad E_{t-1}(y_t) = M_t$$

$$(24) \quad \text{Var}_{t-1}(y_t) = H_t$$

dove M_t è $(N \times 1)$ e la matrice di varianze e covarianze H_t è definita positiva con probabilità uno per tutti i t .

Inoltre:

$$(25) \quad \varepsilon_t = y_t - M_t$$

denota il vettore delle innovazioni $(N \times 1)$ alla media condizionata.

La persistenza nella matrice di varianze e covarianze condizionate è naturalmente caratterizzata dal vettore stocastico $[(N(N+1)/2) \times 1]$

$$(26) \quad H_t^*(s) = E_s(\text{vech}(H_t)) - E_0(\text{vech}(H_t))$$

Il processo stocastico multivariato y_t è definito persistente in varianza se $\text{plim}(\{H_t^*(s)\}_i) \neq 0$ per $s > 0$ e $i = 1, \dots, N(N+1)/2$.

Per il modello $GARCH(p,q)$ scritto in forma vettoriale:

$$\text{vech}(H_t) = W + \sum_{i=1}^q A_i^* \text{vech}(\varepsilon_{t-i} \varepsilon_{t-i}^\top) + \sum_{i=1}^p B_i^* \text{vech}(H_{t-i})$$

la previsione ottima s periodi in avanti:

$$(27) \quad E_s(\text{vech}(H_t)) = W + \sum_{i=1}^q A_i E_s(\text{vech}(H_{t-i})) + \sum_{i=1}^p B_i E_s(\text{vech}(H_{t-i}))$$

dove $t-s \geq m = \max\{p, q\}$.

Riorganizzando i $\text{vech}(H_{t+s}), \dots, \text{vech}(H_{t+s-m+1})$ in un vettore $\text{vech}(H_t)$, $(n \times 1)$, $n = N(N+1)/2$, otteniamo un'equazione alle differenze stocastica che può essere riscritta in modo più conveniente come:

$$(28) \quad E_s(\text{vech}_n(H_t)) = F_n^* W$$

$$E_s(\text{vech}_n(H_t)) = F_n^* W + C_n E_s(\text{vech}(H_{t-1}))$$

La matrice F_n di dimensione $(N(N+1)/2 \times n)$ ha zero in tutti le posizioni eccetto lungo la diagonale principale dove gli elementi sono uguali a 1, la matrice C_n $(n \times n)$ è definita da:

$$C_n = \begin{vmatrix} A_1 + B_1 & A_2 + B_2 & \dots & A_{m-1} + B_{m-1} & A_m + B_m \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & \dots & I & 0 \end{vmatrix}$$

dove $A_i \equiv 0$ per $i > q$ e $B_i \equiv 0$ per $i > p$.

Il modello $GARCH(p,q)$ multivariato è stazionario in senso debole se e solo se tutte le radici del polinomio caratteristico,

$$(23) \quad |I - A(\lambda^{-1}) - B(\lambda^{-1})| = 0$$

stanno all'interno del cerchio unitario, in questo caso:

$$\text{plim}(\{H_i^*(s)\}_i) = 0 \quad \forall s > 0 \quad i = 1, 2, \dots, N(N+1)/2.$$

Se il processo è stazionario in senso debole allora la dipendenza dal valore corrente del processo svanisce con l'aumentare dell'orizzonte di previsione e la previsione ottimale si avvicinerà alla matrice di covarianza non condizionata.

La volatilità dei tassi di cambio: un'applicazione del modello GARCH multivariato

L'analisi empirica degli andamenti dei tassi di cambio

Vediamo adesso un'applicazione del processo $GARCH$ multivariato ai tassi di cambio. L'analisi multivariata permette di evidenziare ed analizzare le relazioni che esistono tra i corsi di diverse valute. In particolare, possiamo mettere in evidenza quali sono i canali attraverso i quali la volatilità caratteristica di ciascuna serie influenza le altre.

I dati utilizzati sono i tassi di cambio settimanali lira/marco, lira/dollaro e lira/sterlina dal 13 marzo 1973 al 31 maggio 1992 (1002 osservazioni). Il loro valore, riportato da *Il Sole 24ore*, è

la media fra la chiusura a pronti di Roma e quella di Milano. Per ovviare all'effetto week-end il giorno di rilevazione è il mercoledì¹.

I tassi di cambio sono misurati in lire per unità di valuta estera. Tutta l'analisi viene condotta utilizzando il logaritmo del tasso di cambio².

Dal grafico (vedi appendice) delle tre serie appaiono evidenti due elementi. Primo, il forte apprezzamento del dollaro alla metà anni '80 seguito poi da un altrettanto pronunciato deprezzamento. Secondo, il continuo apprezzamento del marco rispetto alla lira. L'andamento della sterlina appare caratterizzato da un forte apprezzamento tra la fine degli anni settanta e l'inizio degli anni ottanta.

Il campione è stato suddiviso in tre sottocampioni in corrispondenza di due importanti eventi: l'introduzione del Sistema Monetario Europeo il 13 marzo 1979 e l'accordo del Plaza del 22 settembre 1985 (il vertice dei G-7) che ha segnato l'inizio della svalutazione del dollaro, dopo la forte rivalutazione degli anni precedenti. Verranno, quindi, analizzati tre sottoperiodi: 1973-79, 1979-85, 1985-92.

Analisi della stazionarietà delle serie

Uno sguardo ai grafici (fig.1) dell'andamento dei tre tassi di cambio ci dice che ognuna delle tre serie non è stazionaria. Le serie appaiono processi non stazionari, cioè sono stazionari ed invertibili dopo un'appropriata differenziazione (Fuller, 1976. Wei, 1990).

Dobbiamo, quindi, accertare la presenza di radici unitarie e, visto nel caso del marco un andamento di crescita costante, se è presente una componente di trend.

Le funzioni di autocorrelazione campionarie, calcolate fino al ventesimo ritardo, sembrano indicare l'esistenza di non stazionarietà, come del resto è evidenziato dal fatto che sono tutte positive e non decrescono se non lievemente (fig.2). Le funzioni di autocorrelazione parziale hanno un valore molto elevato in corrispondenza del primo ritardo, mentre gli altri valori non sono significativamente diversi da zero (fig.3).

Questi comportamenti delle funzioni di autocorrelazione e di autocorrelazione parziale suggeriscono, in generale, una non stazionarietà omogenea del primo ordine, e in particolare un *random walk*. E' di importanza fondamentale selezionare correttamente il modello, poichè assumere che il modello contenga un trend può portare a modesti risulati in fase previsiva (Dickey, Bell, Miller, 1986). Se, infatti, da un lato l'eccessiva differenziazione porta alla non invertibilità, ma i parametri della serie potrebbero ancora essere stimati in modo consistente e non distorto; dall'altro, un'inappropriata eliminazione del trend porta a previsioni e intervalli di confidenza scorretti per ogni orizzonte previsionale.

L'evidenza empirica ha mostrato che molte serie finanziarie e di tassi di cambio mostrano un'eteroschedasticità che varia nel tempo (Baillie, Bollerslev, 1989a, 1989b, 1990a, 1990b.

Bollerslev, 1990, 1990b. Daddi, Tivegna, 1990. Diebold, 1988. Diebold, Nerlove 1989. Hsieh, 1989a, 1989b, Milhoj, 1987).

Phillips e Perron (Phillips, Perron, 1988; Campbell, Perron, 1991) hanno fornito una serie di test per le radici unitarie che sono robusti ad un'ampia varietà di situazioni caratterizzate da correlazione seriale ed eteroschedasticità (Kim, Schimdt, 1993).

I test di Phillips-Perron considerano due possibili meccanismi generatori del processo stocastico $\{y_t\}_0^\infty$:

$$(a) \quad y_t = y_{t-1} + u_t$$

$$(b) \quad y_t = \mu + y_{t-1} + u_t \quad t = 1, 2, \dots$$

Il primo processo specifica y_t come un processo stocastico senza *drift* con una radice unitaria, mentre il secondo considera la presenza di un *drift* fissato. Entrambi i processi sono non stazionari a causa della presenza di una radice unitaria.

Per sottoporre a test le due ipotesi (a) e (b), definiamo tre equazioni, tutte stimate con i minimi quadrati ordinari:

$$(24) \quad y_t = \hat{\alpha} y_{t-1} + \hat{u}_t$$

$$(25) \quad y_t = \mu^* + \alpha^* y_{t-1} + u_t^*$$

$$(26) \quad y_t = \tilde{\mu} + \tilde{\beta} (t - T/2) + \tilde{\alpha} y_{t-1} + \tilde{u}_t$$

Le ipotesi sottoposte a test sono le seguenti:

Modello (1)			
$H_0: \hat{\alpha} = 1$	DM	USD	ST
statistica $Z(t_{\alpha})$	3,94	1,51	1,15
Valori critici: 1%=-2.58, 5%=-1.95			

Modello (2)			
$H_0: \alpha^* = 1$	DM	USD	ST
statistica $Z(t_{\alpha^*})$	-3,42	-2,01	-1,66
Valori critici: 1%=-3,43, 5%=-2,86			
$H_0: \alpha^* = 1 \ \mu^* = 0$	DM	USD	ST
statistica $Z(\phi_1)$	15,5	3,44	2,09
Valori critici: 1%=-6,43, 5%=-4,59			

Modello (3)			
$H_0: \tilde{\alpha} = 1$	DM	USD	ST
statistica $Z(t_{\tilde{\alpha}})$	-1,85	-0,74	-1,48
Valori critici: 1%=-3,96, 5%=-3,41			
$H_0: \tilde{\alpha} = 1 \tilde{\beta} = 0 \ \tilde{\mu} = 0$	DM	USD	ST
statistica $Z(\phi_2)$	10,55	2,47	1,51
Valori critici: 1%=-6,09, 5%=-4,68			
$H_0: \tilde{\alpha} = 1 \ \tilde{\beta} = 0$	DM	USD	ST
statistica $Z(\phi_3)$	6,22	2,28	1,56
Valori critici: 1%=-8,27 5%=-6,25			

Il numero dei ritardi dei residui utilizzato nel test è pari a 4. Dall'esame dei test Phillips-Perron, i quali non rifiutano mai l'ipotesi di presenza di una radice unitaria, appare necessario differenziare le tre serie al fine di ottenere la stazionarietà. Questo risultato è ricorrente in molti studi empirici sui tassi di cambio (Corbae, Ouliaris, 1986).

Specificazione del modello

Per specificare correttamente il modello è stato calcolato lo *Schwarz Information Criterion* il quale, insieme con le usuali diagnostiche come la funzione di autocorrelazione, non rivela la presenza di una componente a media mobile in ognuna delle tre serie. Il numero dei ritardi considerato ottimale per le tre serie è pari a 3.

Il test di Ljung-Box, calcolato per i ritardi 6,12,18 sottopone a verifica l'ipotesi che i ritardi non siano autocorrelati. I residui sono così ottenuti:

$$(27) \quad \varepsilon_t = s_t^i - (\gamma_0 + \gamma_1 s_{t-1}^i + \gamma_2 s_{t-2}^i + \gamma_3 s_{t-3}^i) \quad i = \text{DM, USD, ST}$$

dove

$$(28) \quad s_t^i = \Delta \ln S_t^i \quad i = \text{DM, USD, ST}$$

Va però ricordato che nel caso di presenza di effetti *ARCH*, per i quali un apposito test è stato calcolato, il test Ljung-Box tende ad essere «conservativo» (Diebold, 1988).

Test Ljung-Box			
Ritardi	6	12	18
DM	6.9328	26.9630	37.0413
USD	1.2345	12.4616	15.8849
ST	6.0444	20.9569	25.3583

Il test di Ljung-Box è distribuito come una χ^2 con $(m-p-q)$ gradi di libertà dove m indica il numero di ritardi della correlazione ai quali si applica il test e $(p+q)$ è il numero dei parametri stimati. Rifiutiamo l'ipotesi nulla nel caso del marco per tutti i ritardi considerati, accettiamo l'ipotesi nulla per il dollaro per tutti i ritardi; nel caso della sterlina accettiamo l'ipotesi nulla per i 6 ritardi, mentre la rifiutiamo negli altri.

Infine per valutare le proprietà delle distribuzioni delle serie $s_t^i = \Delta \ln S_t^i$ è stato calcolato il test di Kiefer-Salmon per testare la normalità dei residui. Il test Kiefer-Salmon si basa sulla costruzione della statistica (Kiefer, Salmon, 1983):

$$(29) \quad KS = \frac{T}{6}(\hat{\mu}_3 - 3\hat{\mu}_1)^2 + \frac{T}{24}(\hat{\mu}_4 - 6\hat{\mu}_2 + 3)^2$$

$$KS = KS_1 + KS_2$$

La statistica è costruita con i residui standardizzati con la stima di massima verosimiglianza di σ^2 . KS è distribuita come una chi-quadro con due gradi di libertà sotto l'ipotesi nulla di normalità; può essere scomposta in due variabili indipendenti asintoticamente e distribuite come una χ^2 con un grado di libertà. La prima statistica, KS_1 , è un test LM per la *skewness* sotto l'ipotesi di normalità e la seconda, un test LM per la *curtosis*.

Test Kiefer-Salmon			
	DM	ST	USD
KS	66931.947	2049.648	4242.4641
KS_1	362.79152	0.32975520	12.126817
KS_2	66569.156	2049.318	4230.3373

Appare evidente dai risultati dei test che si deve rifiutare l'ipotesi di normalità. In particolare, i residui sono caratterizzati da un forte curtosi, cioè hanno una distribuzione caratterizzata da code spesse.

Infine, sottponiamo a verifica l'ipotesi di assenza di effetti *ARCH*. A questo scopo è stato calcolato un test LM sull'autocorrelazione dei residui al quadrato che è distribuito asintoticamente come una χ^2 con p gradi di libertà, dove p indica l'ordine dell'autocorrelazione.

I risultati sono i seguenti:

Test <i>LM</i> per effetti <i>ARCH</i>			
	DM	USD	ST
ARCH(1)	58.478475	122.33470	24.991538
ARCH(2)	60.567412	135.57151	25.206590
ARCH(3)	64.015392	140.20836	25.186069
ARCH(4)	69.507763	140.80197	25.256345
ARCH(5)	69.749255	140.92204	25.249596
ARCH(6)	75.409859	141.40879	25.436784
ARCH(7)	121.43973	145.01782	44.179661
ARCH(8)	218.57219	150.85121	136.08244
ARCH(9)	232.55179	152.63584	141.36413
ARCH(10)	233.03210	153.52069	141.38358
ARCH(11)	233.83793	154.40141	141.29206
ARCH(12)	238.78575	154.24769	141.27167

Come si può osservare il test rifiuta ampiamente l'ipotesi di assenza di effetti *ARCH* per tutti gli ordini per il quale è stato calcolato.

Gli effetti *ARCH* osservati sono consistenti con la curtosi osservata nelle tre serie dei tassi di cambio. I processi *ARCH*, infatti, possiedono densità non condizionali con code spesse, anche se le densità condizionali sono normali.

Dovremmo, quindi, modellare i disturbi con un processo *ARCH*(12). E' stato mostrato (Lee, 1991), però, che il test *LM* per l'ipotesi nulla di disturbi *white noise* contro l'ipotesi alternativa di disturbi *ARCH*(q) è equivalente ad un test che confronta l'ipotesi di disturbi *white noise* contro un'alternativa di disturbi *GARCH*(p,q). Questo implica che sotto l'ipotesi nulla di disturbi *white noise* gli effetti *ARCH*(q) e *GARCH*(p,q) sono alternative localmente equivalenti. La determinazione dell'ordine di grandezza di p e q nel processo *GARCH* non è un problema di immediata soluzione. Un modo per risolvere questo problema può essere quello di determinare prima il valore ottimale di q^* nel modello *ARCH* attraverso test *LR*. Un volta determinato il modello *ARCH*(q^*) lo si può confrontare, ad esempio, con il modello *GARCH*(1,1) adottando criteri quali l'*Akaike Information Criterion* o lo *Schwarz Information Criterion* (Hsieh, 1989a). Nel nostro caso l'applicazione di questa procedura ci ha portati alla scelta di un modello *GARCH*(1,1). Riassumendo, la formulazione univariata per le tre serie:

$$(30) \quad s_t^i = (\gamma_0 + \gamma_1 s_{t-1}^i + \gamma_2 s_{t-2}^i + \gamma_3 s_{t-3}^i) + \varepsilon_t \quad i = \text{DM, USD, ST}$$

dove

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

con

$$(31) \quad h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

$$(32) \quad s_t^i | \Psi_{t-1} \sim N(\gamma_0 + \sum_{j=1}^3 \gamma_j s_{t-j}^i, h_t) \quad i = \text{DM, USD, ST}$$

Per la positività di h_t devono valere i vincoli (Nelson, Cao, 1992):

$$\omega \geq 0, \alpha_1 \geq 0, \beta_1 \geq 0,$$

Il modello multivariato è costituito da una parte strutturale, quella relativa alla media condizionale, e da un processo *GARCH(1,1)* multivariato per la varianza condizionale degli errori.

La parte strutturale è la seguente:

$$(33) \quad Y = Xb + e$$

$$(34) \quad Y^i = (s_t^{DM}, s_t^{USD}, s_t^{ST})'$$

dove $s_t^i, i = \text{DM, USD, ST}$, è un vettore $(T \times 1)$, dove T è il numero delle osservazioni.

$$(35) \quad X = \text{diag}(X^{DM}, X^{USD}, X^{ST})$$

X è una matrice che ha dimensione $(NT \times (N(N^2+1)))$. In questo caso $N=3$.

$$(36) \quad X^i = (d, s_t^{DM}, s_t^{USD}, s_t^{ST}) \quad i = \text{DM, USD, ST}$$

d è un vettore di uno di dimensione $(T \times 1)$.

$$(37) \quad s^i = (s_{t-1}^i, s_{t-2}^i, s_{t-3}^i) \quad i = \text{DM, USD, ST}$$

s_{t-j}^j , $j=1,2,3$, è un vettore $(T \times 1)$, b è il vettore dei parametri di dimensione $N(N^2+1)$. La matrice X^i contiene le endogene ritardate fino a tre periodi, in accordo a quanto stabilito con l'analisi di specificazione del modello. Si tratta, in sostanza, di un modello *SURE* (Seemingly Unrelated Regression Equations) nel quale le equazioni solo apparentemente sono non collegate.

La matrice delle varianze e covarianze condizionate viene modellata secondo un *GARCH(1,1)* multivariato espresso nella forma che garantisce la definitezza positiva :

$$(38) \quad H_t = CC' + A_t \varepsilon_{t-1} \varepsilon_{t-1}' A_t' + B_t H_{t-1} B_t'$$

Per semplificare il modello e ridurre così il numero dei parametri assumiamo che le matrici A_t e B_t sono diagonali:

$$A = \text{diag}(\alpha_{11}, \alpha_{22}, \alpha_{33}), \quad B = \text{diag}(\beta_{11}, \beta_{22}, \beta_{33})$$

Questo modo di procedere permette di ridurre il numero dei parametri da stimare a 42 rispetto ai 54 del modello senza vincoli, semplificando così le già complesse procedure di calcolo.

Il modello scritto per esteso:

$$\begin{aligned} s_t^i &= \gamma_{i,c} + \gamma_{i,dm,1} s_{t-1}^{DM} + \gamma_{i,dm,2} s_{t-2}^{DM} + \gamma_{i,dm,3} s_{t-3}^{DM} \\ &+ \gamma_{i,usd,1} s_{t-1}^{USD} + \gamma_{i,usd,2} s_{t-2}^{USD} + \gamma_{i,usd,3} s_{t-3}^{USD} \\ &+ \gamma_{i,st,1} s_{t-1}^{ST} + \gamma_{i,st,2} s_{t-2}^{ST} + \gamma_{i,st,3} s_{t-3}^{ST} \quad i = \text{DM, USD, ST} \end{aligned}$$

le varianze e covarianze condizionate:

$$h_{dm,t} = c_{11}^2 + \alpha_{11}^2 \varepsilon_{dm,t-1}^2 + \beta_{11}^2 h_{dm,t-1}$$

$$h_{dm,us,t} = c_{21}c_{11} + \alpha_{11}\alpha_{22} \varepsilon_{dm,t-1} \varepsilon_{usd,t-1} + \beta_{11}\beta_{22} h_{dm,usd,t-1}$$

$$h_{usd,t} = c_{22}^2 + c_{21}^2 + \alpha_{22}^2 \varepsilon_{usd,t-1}^2 + \beta_{22}^2 h_{usd,t-1}$$

$$h_{dm,st,t} = c_{31}c_{11} + \alpha_{11}\alpha_{33} \varepsilon_{dm,t-1} \varepsilon_{st,t-1} + \beta_{11}\beta_{33} h_{dm,st,t-1}$$

$$h_{usd,st,t} = c_{22}c_{32} + c_{21}c_{31} + \alpha_{22}\alpha_{33} \varepsilon_{usd,t-1} \varepsilon_{st,t-1} + \beta_{22}\beta_{33} h_{usd,st,t-1}$$

$$h_{st,t} = c_{33}^2 + c_{32}^2 + c_{31}^2 + \alpha_{33}^2 \varepsilon_{st,t-1}^2 + \beta_{33}^2 h_{st,t-1}$$

Il modello completo, per ogni sottoperiodo, viene stimato con il metodo della massima verosimiglianza. Il modello di partenza, per ogni sottoperiodo, è stato ristretto sulla base del test *LR*.

La presenza di *outlier* dovuti a shock o a riallineamenti di una certa entità all'interno dello SME ha suggerito l'introduzione di *dummy*. Per il periodo 1973-79 è stata introdotta una *dummy* per cogliere un'ampia variazione dei tre tassi nel 1976. Per il periodo 1979-85 è stata inserita una *dummy* nell'equazione del tasso di cambio lira-marco tedesco per tenere conto del riallineamento del 5/10/1981, che è quello che provoca la variazione maggiore del logaritmo del tasso di cambio lira-marco. L'introduzione delle *dummy* permette di evitare il problema dell'integrazione nella varianza condizionale, ovvero che $\alpha_{11}^2 + \beta_{11}^2 \geq 1$, cioè di dare conto dell'improvvisa variazione del tasso altrimenti non catturabile.

Riportiamo qui le stime per l'ultimo sottoperiodo (1985-92). Le stime degli altri due sottoperiodi sono disponibili presso l'autore.

Modello completo, 1985-92

	Stima	Standard error	test t
c_{11}	0.0015774	0.00020939	7.5333
c_{21}	-0.0028872	0.000735	-3.9230
c_{22}	-0.0003992	0.000618	-0.64502
c_{31}	-2.287E-06	0.000270	-0.00847
c_{32}	-0.0012710	0.000426	-2.9778
c_{33}	-0.0001367	0.002740	-0.04989
α_{11}	0.54026	0.10157	5.3193
α_{22}	0.11335	0.041665	2.7206
α_{33}	0.21719	0.032970	6.5876
β_{11}	0.58403	0.12246	4.7692
β_{22}	0.97174	0.012356	78.643
β_{33}	0.96283	0.010544	91.316
$\gamma_{dm,c}$	7.5799E-05	0.000123	0.61210
$\gamma_{dm,dm,1}$	-0.078954	0.058059	-1.3599
$\gamma_{dm,dm,2}$	0.070405	0.047373	1.4862
$\gamma_{dm,dm,3}$	0.026516	0.044546	0.59524
$\gamma_{dm,usd,1}$	-0.0039295	0.009297	-0.42267
$\gamma_{dm,usd,2}$	0.013444	0.008840	1.5207
$\gamma_{dm,usd,3}$	0.0074437	0.008422	0.88378
$\gamma_{dm,st,1}$	0.0053684	0.013297	0.40372
$\gamma_{dm,st,2}$	-0.0040079	0.013229	-0.30296
$\gamma_{dm,st,3}$	0.0056510	0.013230	0.42714
dummy1	0.019597	0.002271	8.6288
dummy2	0.017004	0.002756	6.1687
dummy3	0.019550	0.001906	10.256

(segue)

$\gamma_{usd,c}$	-0.0005799	0.000767	-0.75590
$\gamma_{usd,dm,1}$	-0.055579	0.26531	-0.20949
$\gamma_{usd,dm,2}$	-0.52351	0.25744	-2.0335
$\gamma_{usd,dm,3}$	0.029930	0.26162	0.11440
$\gamma_{usd,usd,1}$	-0.0003831	0.057335	-0.00668
$\gamma_{usd,usd,2}$	0.061709	0.056424	1.0937
$\gamma_{usd,usd,3}$	-0.055850	0.055148	-1.0127
$\gamma_{usd,st,1}$	-0.056119	0.080938	-0.69337
$\gamma_{usd,st,2}$	0.096861	0.082068	1.1802
$\gamma_{usd,st,3}$	-0.028020	0.081701	-0.34296
$\gamma_{st,c}$	-7.878E-05	0.000435	-0.18083
$\gamma_{st,dm,1}$	0.30953	0.15312	-2.0214
$\gamma_{st,dm,2}$	0.028005	0.15063	-0.18592
$\gamma_{st,dm,3}$	0.21461	0.15217	1.4103
$\gamma_{st,usd,1}$	-0.012002	0.029707	-0.40401
$\gamma_{st,usd,2}$	0.015861	0.030630	0.51782
$\gamma_{st,usd,3}$	0.015087	0.030621	0.49272
$\gamma_{st,st,1}$	0.030717	0.054937	0.55913
$\gamma_{st,st,2}$	-0.0077983	0.056901	-0.13705
$\gamma_{st,st,3}$	0.026551	0.056433	0.47050

log-verosimiglianza=4427.6

Modello ristretto, 1985-92

	Stima	Standard error	test t
c_{11}	0.0015495	0.00021348	7.258
c_{21}	-0.0027516	0.000687	-4.0025
c_{22}	-0.0005671	0.000897	-0.63222
c_{31}	2.3734E-05	0.000257	0.092154
c_{32}	-0.0009702	0.000835	-1.1615
c_{33}	-0.000849	0.000944	-0.89853
α_{11}	0.51582	0.091998	5.6069
α_{22}	0.11134	0.039512	2.8178
α_{33}	0.21733	0.033089	6.5681
β_{11}	0.61196	0.11427	5.3555
β_{22}	0.97361	0.011381	85.545
β_{33}	0.96267	0.010770	89.386
$\gamma_{dm,c}$	8.4516E-05	0.000124	0.68017
$\gamma_{dm,dm,1}$	-0.081295	0.050233	-1.6184
$\gamma_{dm,dm,2}$	0.071910	0.046717	1.5393
$\gamma_{dm,usd,2}$	0.013457	0.008578	1.5686
dummy1	0.019604	0.002234	8.7741
dummy2	0.016991	0.002623	6.4756
dummy3	0.019323	0.001813	10.653
$\gamma_{usd,c}$	-0.0005504	0.000759	-0.72447
$\gamma_{usd,dm,2}$	-0.49972	0.25352	-1.9711
$\gamma_{usd,usd,2}$	0.057024	0.055416	1.0290
$\gamma_{usd,usd,3}$	-0.052851	0.046929	-1.1262
$\gamma_{usd,st,2}$	0.084587	0.074145	1.1408
$\gamma_{st,c}$	-8.901E-05	0.000437	-0.20363
$\gamma_{st,dm,1}$	-0.28479	0.14060	-2.0256
$\gamma_{st,dm,3}$	0.19749	0.14035	1.4071

Log-verosimiglianza=4425.7. Test LR=3,8

Previsioni.

Il modello stimato per il sottoperiodo 1985-92 è stato utilizzato in fase previsiva. Sono state calcolate le previsioni per gli ultimi 19 periodi, 1,2 e 3 passi in avanti, utilizzando il modello ristretto per il periodo 1985-92.

Il modello è stimato con un campione ridotto ($T=19$), cioè ogni previsione è ottenuta stimando il modello con le osservazioni disponibili fino al periodo di formazione della previsione.

Previsioni un periodo in avanti

	DM	USD	ST
ME	0.071192	-1.1005	-4.4704
MAE	0.9828	12.716	8.3809
RMSE	1.2777	15.137	13.289
IN1	0.072436	0.08654	0.53341
IN2	0.089318	0.05666	0.17439

Previsioni due periodi in avanti

	DM	USD	ST
ME	0.28525	4.3084	-8.1206
MAE	1.6858	18.562	11.428
RMSE	2.0517	23.543	17.593
IN1	0.1692	0.23211	0.71056
IN2	0.06461	0.07989	0.16059

Previsioni tre periodi in avanti

	DM	USD	ST
ME	0.4035	-8.1206	-11.964
MAE	2.1346	23.829	15.033
RMSE	2.4696	30.355	22.525
IN1	0.18903	0.34079	0.79585
IN2	10.0467	0.08154	0.14838

Dove le statistiche calcolate per valutare la performance previsiva sono le seguenti:

$$ME = \frac{1}{T} \sum_{i=1}^T (\hat{Y}_i - Y_i), \quad MAE = \frac{1}{T} \sum_{i=1}^T |\hat{Y}_i - Y_i|, \quad RMSE = \left[\frac{1}{T} \sum_{i=1}^T (\hat{Y}_i - Y_i)^2 \right]^{1/2},$$

$$IN1 = \frac{|ME|}{MAE}, \quad IN2 = \frac{(RMSE - MAE)}{MAE\sqrt{T} - MAE}$$

Le statistiche calcolate per i tre orizzonti di previsione indicano che sono assenti sia fenomeni di sottostima (quando $ME=MAE$) sia di sovrastima (quando $-ME=MAE$). Esistono invece degli errori di previsione di grandi dimensioni in alcuni periodi, come del resto ci si accorge confrontando i valori ottenuti per il MAE ed il RMSE. Questi errori di previsione sono spiegabili sulla base di eventi improvvisi che hanno turbato i mercati. I 19 periodi di previsione vanno dal 22 gennaio 1992 al 27 maggio 1992. Durante questo arco temporale vi è stata una continua oscillazione nelle aspettive sull'andamento dell'economia americana. A questo si devono aggiungere due scadenze elettorali, rispettivamente in Italia e nel Regno Unito.

Per quanto riguarda la previsione *out-of-sample*, ovvero le previsioni per i periodi non compresi nel campione inizialmente utilizzato, a tal fine è interessante costruire degli intervalli di previsione. Nel caso di processi autoregressivi stazionari con disturbi *GARCH* la varianza condizionale degli errori di previsione dipende dalla previsione della varianza condizionale dei disturbi del modello (Baillie, Bollersev, 1992). Se non si ipotizza la normalità condizionale dei disturbi, il quantile che deve essere usato nella costruzione degli intervalli deve essere stimato. In questo lavoro abbiamo ipotizzato che i disturbi siano condizionalmente normali, è quindi lecito adottare i quantili della normale standard. Per la stima della varianza degli errori di previsione si è adottato il metodo euristico proposto da Doan, Litterman e Sims (1984) basato sulla stima campionaria della varianza degli errori ottenuti con la falsa previsione³. Nel nostro

caso questo ha significato utilizzare gli errori di previsione ottenuti precedentemente con le previsioni 1,2 e 3 periodi in avanti per le tre valute.

Gli intervalli di confidenza con un livello di confidenza del 95% sono:

	DM	US	ST
1 Periodo	± 2.5	± 29.58	± 24.52
2 Periodi	± 3.98	± 45.36	± 30.58
3 Periodi	± 4.73	± 54.16	± 37.15

Nelle tabelle che seguono sono riportati: i valori previsti, effettivi, gli errori di previsione, il limite superiore ed inferiore dell'intervallo di previsione.

Previsioni 1 periodo in avanti

	Lim.sup.	Previsto	Effettivo	Lim.inf.	Errore
DM	755.87	753.37	755	750.87	-1.629
USD	1257.98	1228.4	1214.055	1198.82	14.32
ST	2241.42	2216.9	2209.35	2192.372	7.59

Previsioni 2 periodi in avanti

	Lim.sup.	Previsto	Effettivo	Lim.inf.	Errore
DM	757.7	753.72	755.945	749.74	-2.225
USD	1275.56	1230.2	1206.25	1184.83	23.956
ST	2247.48	2216.9	2209.55	2186.32	7.358

Previsioni 3 periodi in avanti

	Lim.sup.	Previsto	Effettivo	Lim.inf.	Errore
DM	758.47	753.74	756.255	749.01	-2.264
USD	1282.16	1228	1185.51	1173.83	42.513
ST	2253.95	2216.8	2206.52	2179.64	10.257

I risultati ottenuti sono abbastanza buoni tenuto conto che i periodi oggetto di previsione sono stati caratterizzati da una forte volatilità.

Si tratta, infatti, delle prime tre settimane del mese di giugno 1992. Con il referendum del 2 giugno 1992 i danesi rifiutarono il trattato di Maastricht e questo ebbe un effetto destabilizzante sui mercati valutari. Il 3 giugno ci fu un forte calo del dollaro provocato da aspettative negative sull'andamento dell'economia americana. Il 10 giugno aumentarono i tassi d'interesse italiani per fronteggiare movimenti speculativi; nella settimana dal 3 al 10 giugno ci furono forti oscillazioni dei cambi. Il 17 giugno aumentarono i tassi a breve in Italia ed il dollaro continuò la propria discesa mentre in Europa aumentò l'incertezza sul referendum indetto in Irlanda sul trattato di Maastricht.

La previsione per un arco di tempo caratterizzato da una componente di disturbo così forte appare difficile. Questo è evidente nel caso dei risultati ottenuti per il dollaro che, in particolare per la previsione a tre periodi, non sono molto significativi.

Conclusioni

L'analisi che è stata svolta è partita dalla considerazione che i modelli strutturali dei tassi di cambio, così come sono stati proposti e sottoposti a verifica empirica nella letteratura, paiono inadeguati a catturare il fenomeno della volatilità. I modelli ARCH, qui utilizzati nella versione multivariata, sono una generalizzazione della classe dei modelli ARMA. I processi ARCH sono pensati per modellare l'eteroschedasticità condizionale, questo costituisce una generalizzazione dell'intera classe dei processi stocastici lineari e stazionari in senso debole. Con il processo ARCH lineare la varianza condizionata dei disturbi varia nel tempo come funzione lineare dei disturbi passati al quadrato. In termini di valore atteso, la variabilità di oggi dipende linearmente dalla variabilità di ieri, in questo modo ampie variazioni tendono ad essere seguite da ampie variazioni, piccole variazioni vengono seguite da piccole variazioni di entrambi i segni.

Il modello utilizzato per l'analisi dei tre sottoperiodi in cui è stato diviso il campione (1973-79, 1979-85, 1985-1992) è costituito da un modello *VAR*(3) per le variabili in differenze con una specificazione multivariata *GARCH*(1,1) per la varianza condizionata dei disturbi. A tale formulazione si è giunti attraverso la consueta analisi di specificazione (funzioni di autocorrelazione ed autocorrelazione parziale) e l'analisi della stazionarietà delle serie. Per quanto riguarda la varianza condizionata dei disturbi l'effetto ARCH è stato individuato con l'applicazione di specifici test e lo stesso è stato fatto per la specificazione funzionale.

Il modello *GARCH* utilizzato è di tipo «diagonale» e per la stima è stata adottata una parametrizzazione che garantisce la definitezza positiva della matrice di varianze covarianze condizionate.

I risultati delle stime confermano che la modellistica *ARCH-GARCH* riesce a cogliere la volatilità dei tassi di cambio. La struttura multivariata adottata permette di analizzare le interdipendenze sempre maggiori, grazie alla crescente integrazione tra i mercati valutari. Questo allargamento della visuale rispetto alle analisi univariate ha permesso di evidenziare: in primo luogo, i canali attraverso cui i disturbi generati dall'arrivo di *news* o da errori nelle aspettative, relativi ad una specifica valuta, si diffondono sui mercati delle altre valute, in secondo luogo, come la volatilità tende a variare nel corso del tempo.

Con il modello stimato per l'ultimo periodo sono state condotte alcune previsioni *in-sample*, 1,2 e 3 passi in avanti per le ultime 19 osservazioni del campione. Queste hanno mostrato che il modello è ben specificato e che ha una buona affidabilità in fase previsiva. In particolare, gli errori di previsione sono imputabili a shock imprevedibili sui mercati valutari.

L'affidabilità previsiva del modello è stata confermata anche dalle previsioni per 1,2 e 3 passi in avanti non appartenenti al campione, ovvero per le prime tre settimane del mese di giugno 1992, un periodo caratterizzato da forti tensioni sui mercati valutari.

Il modello fornisce delle «buone» previsioni ad un passo mentre appare meno preciso per quelle a 2 e 3 periodi in avanti. Questo dipende dal meccanismo di formazione delle previsioni, che nel caso di modelli ARMA è di tipo iterativo.

* Desidero ringraziare Carlo Giannini e Alessandro Duca per gli aiuti, i suggerimenti e le correzioni. Gli eventuali rimanenti errori sono di mia esclusiva responsabilità.

Figure

Fig.1

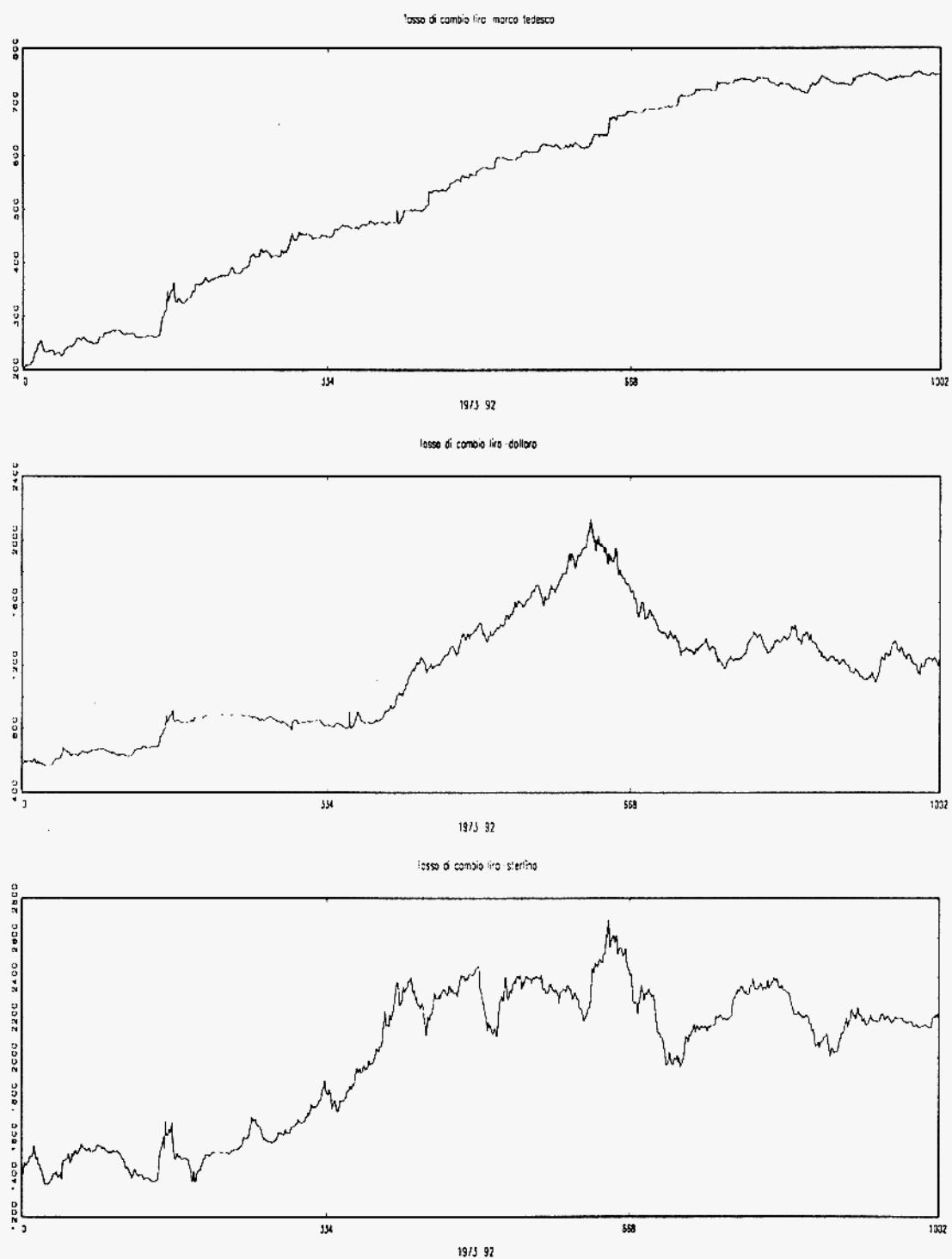
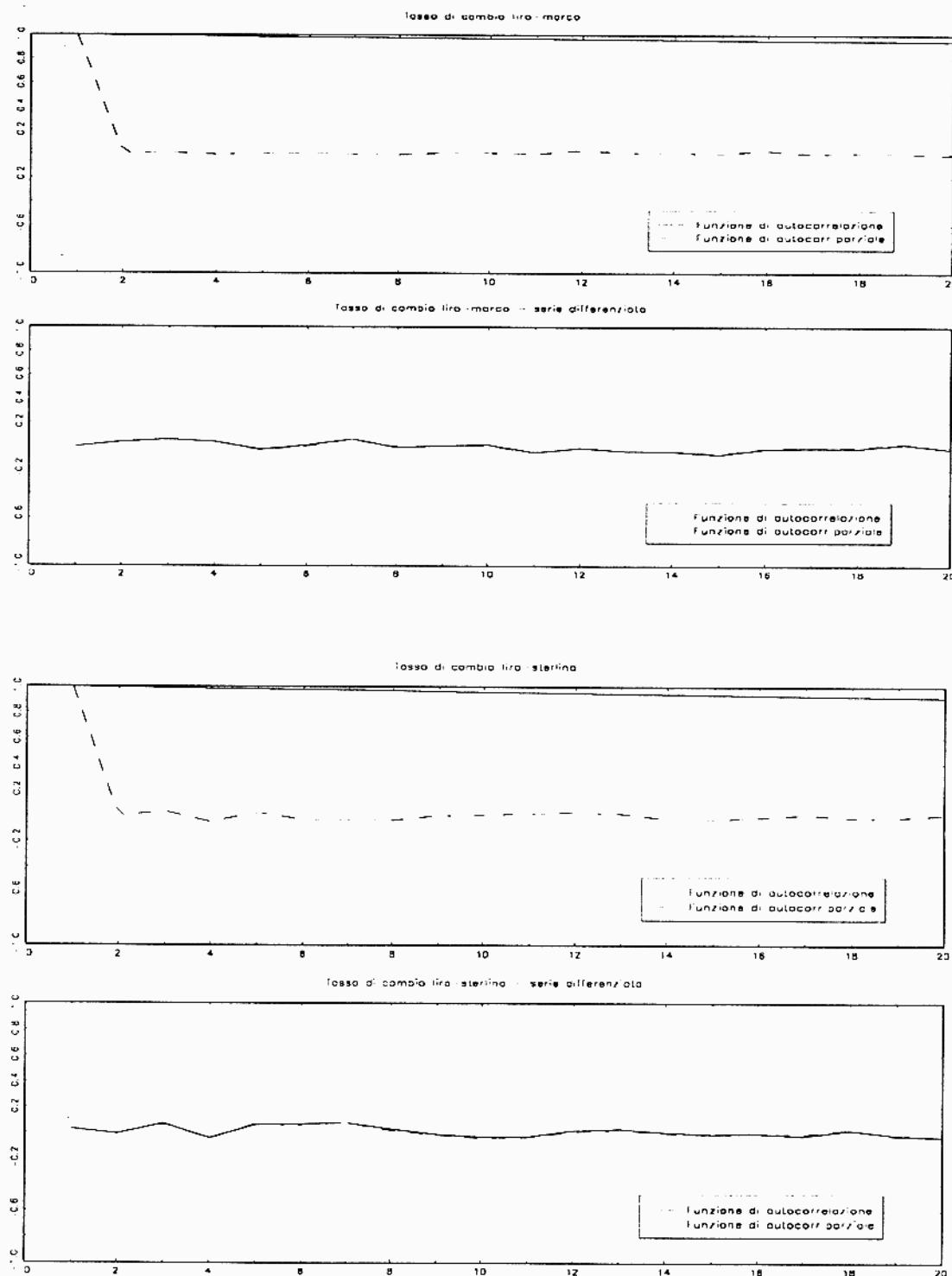


Fig.2



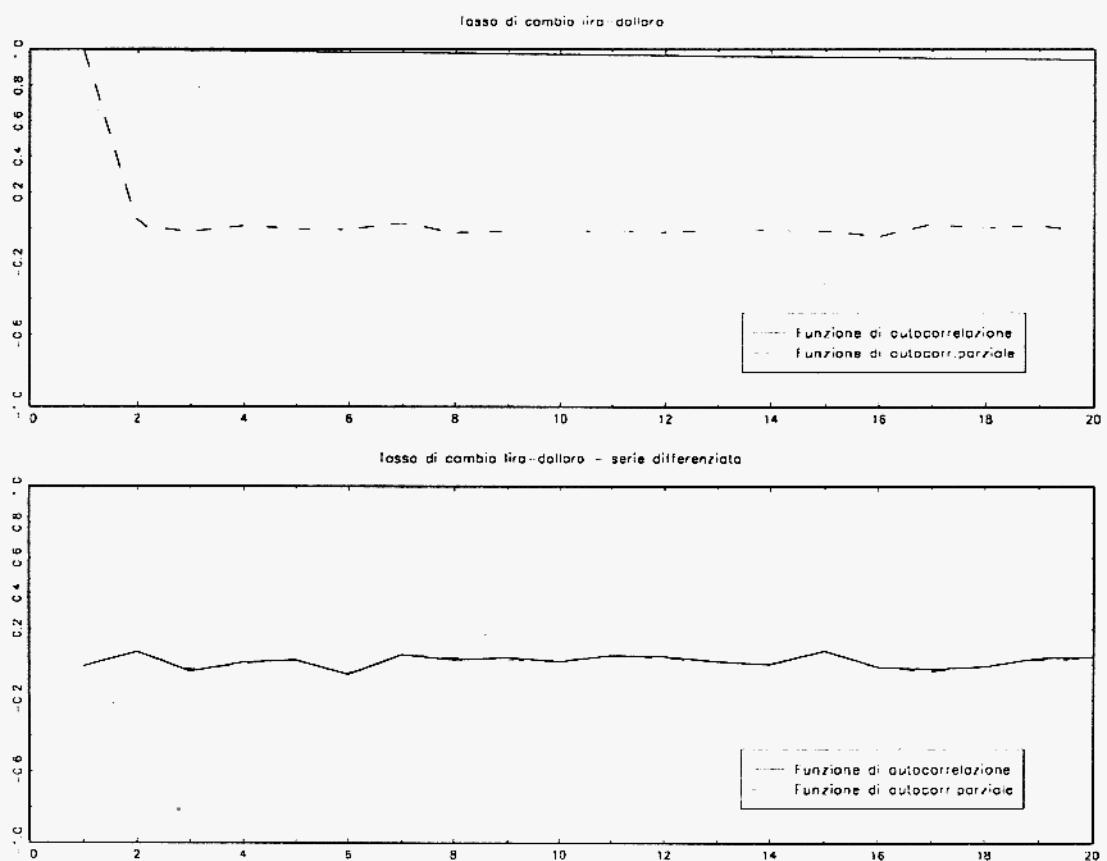


Fig.3

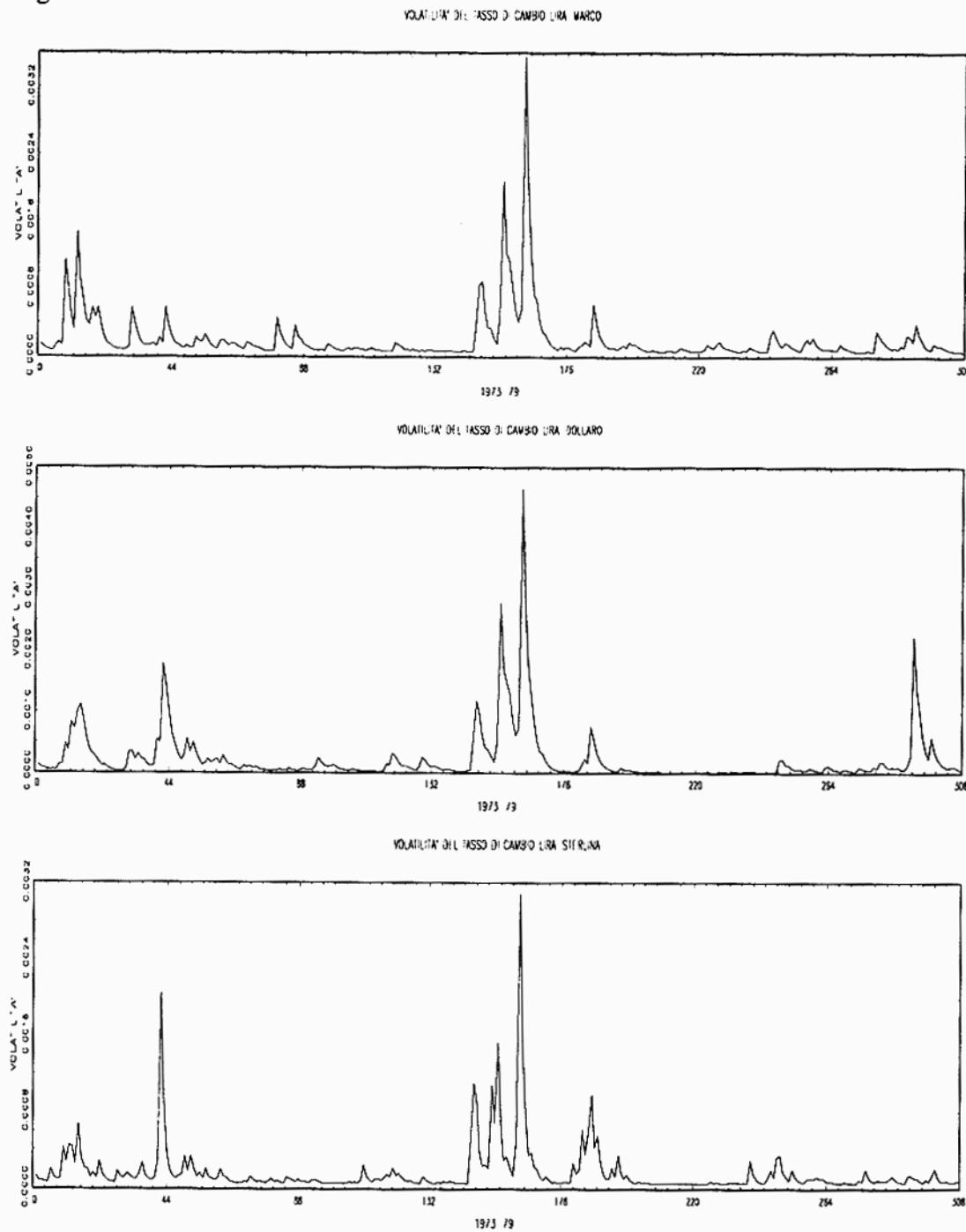


Fig.4

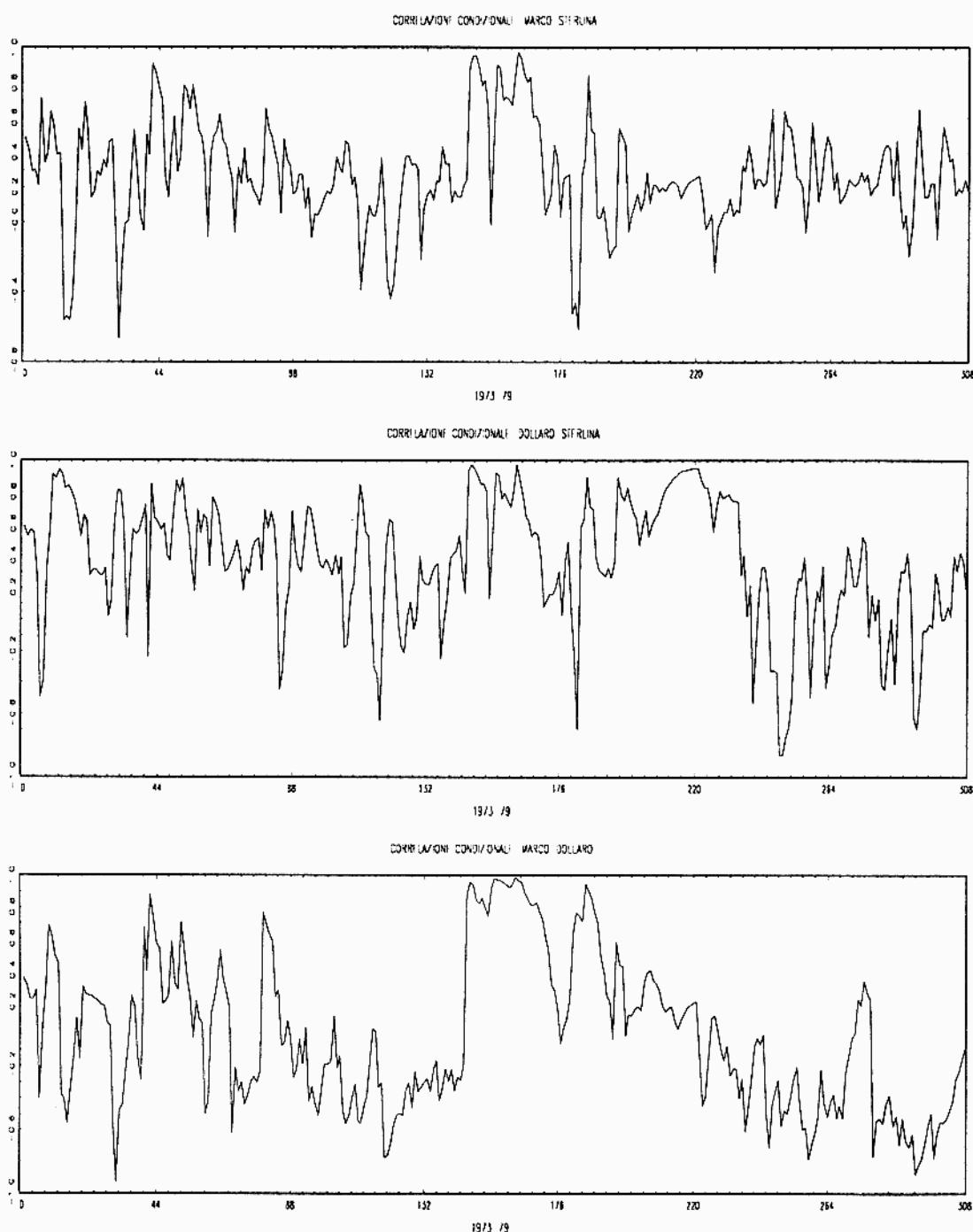


Fig.5

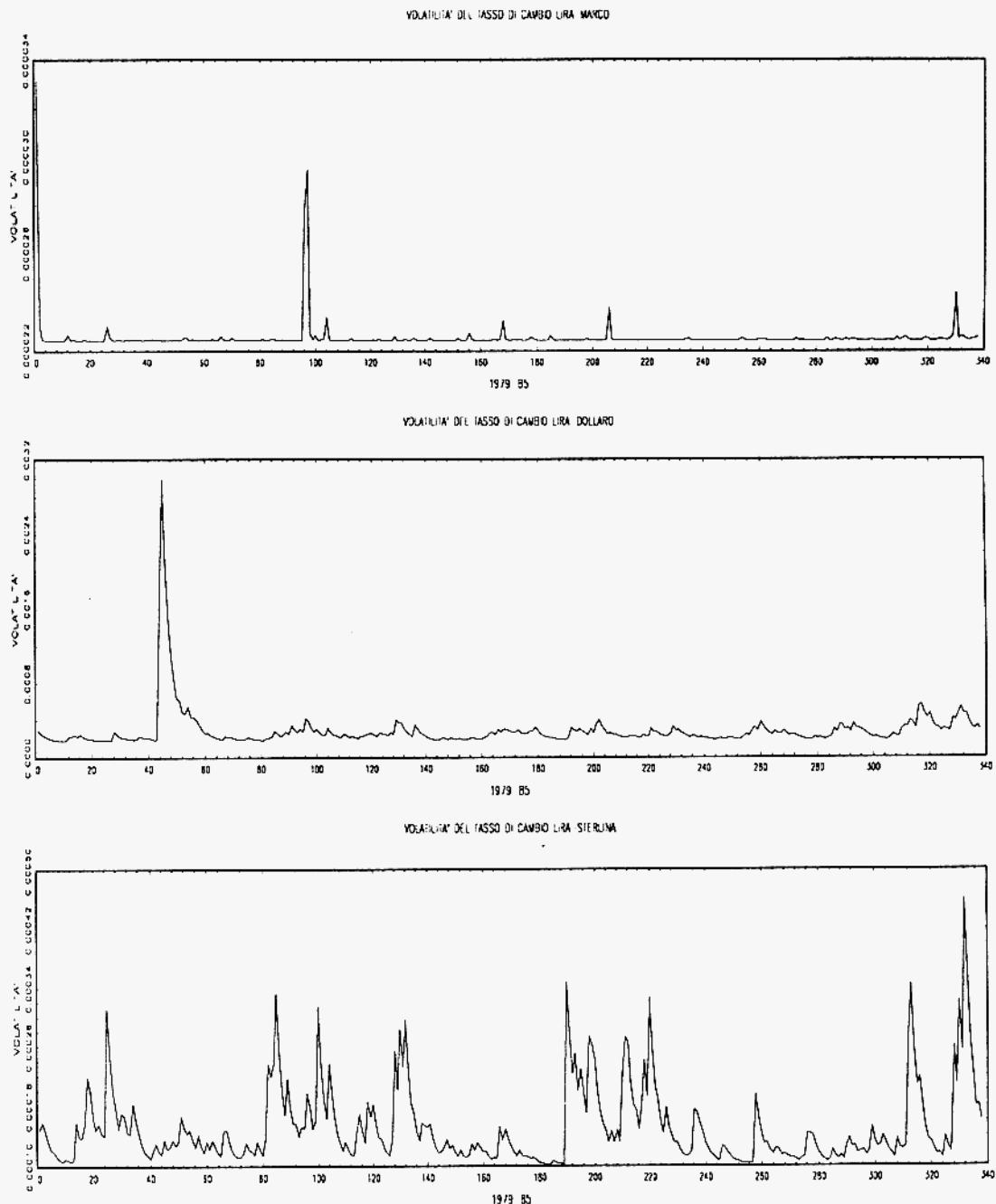


Fig.6

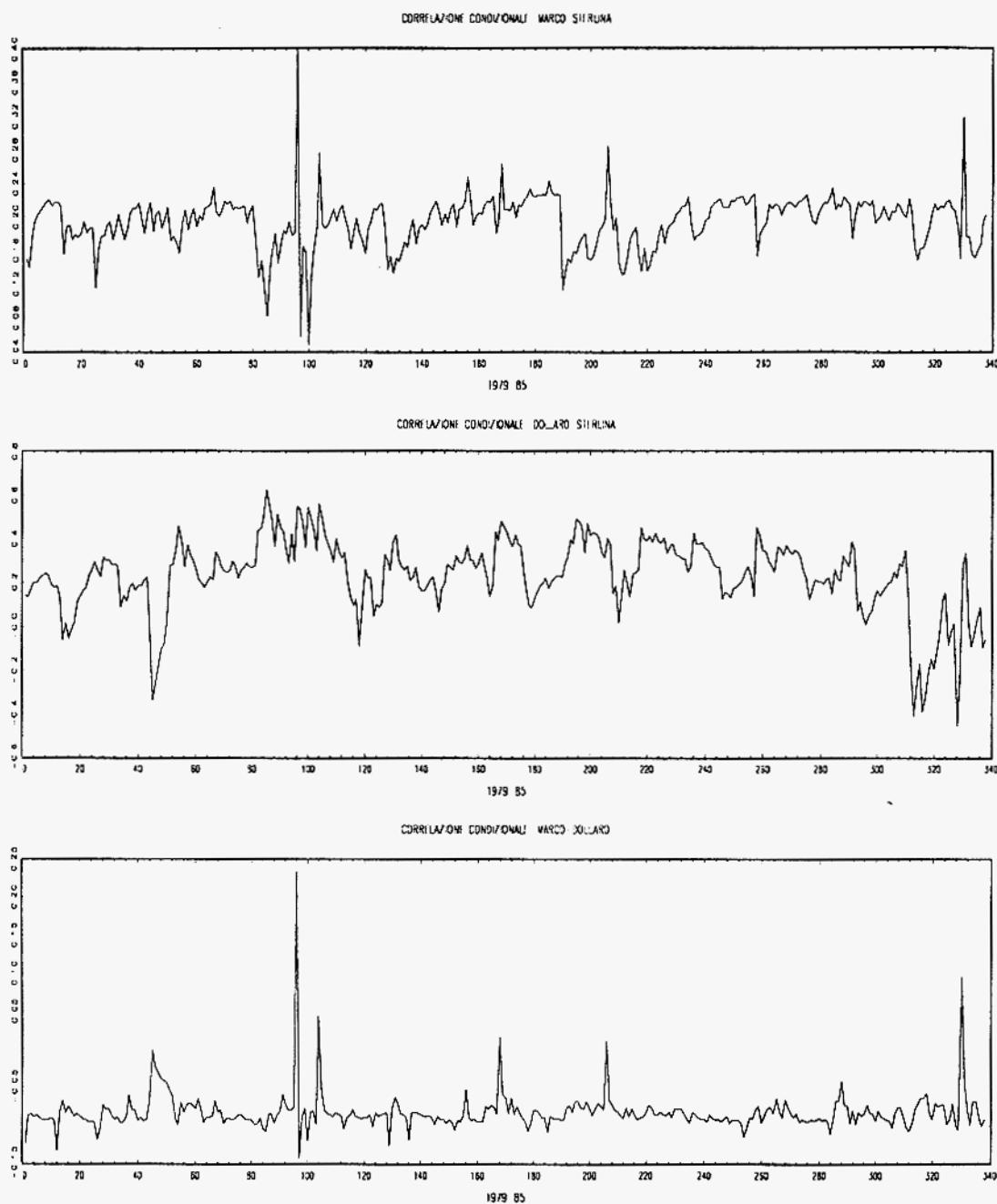


Fig.7

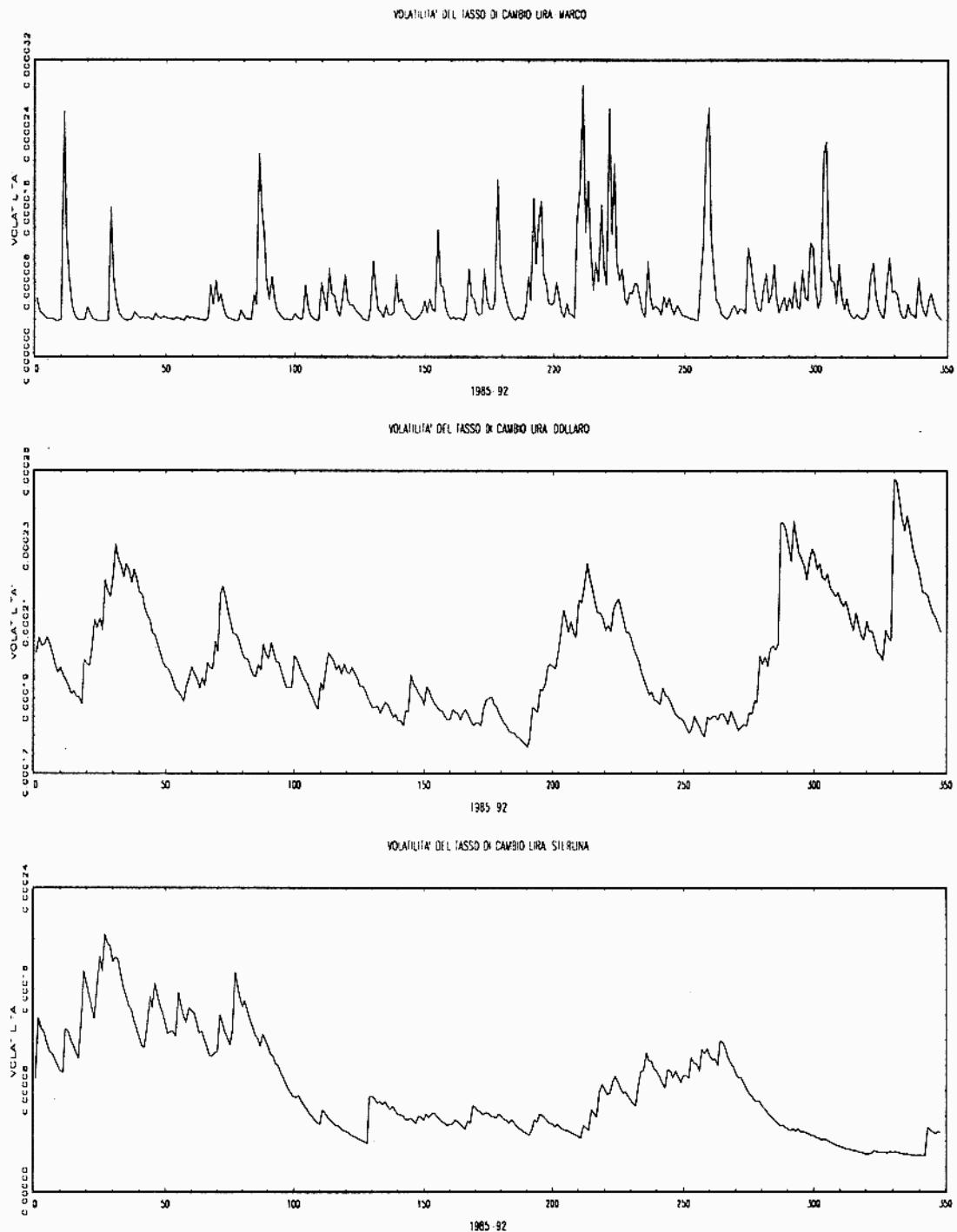
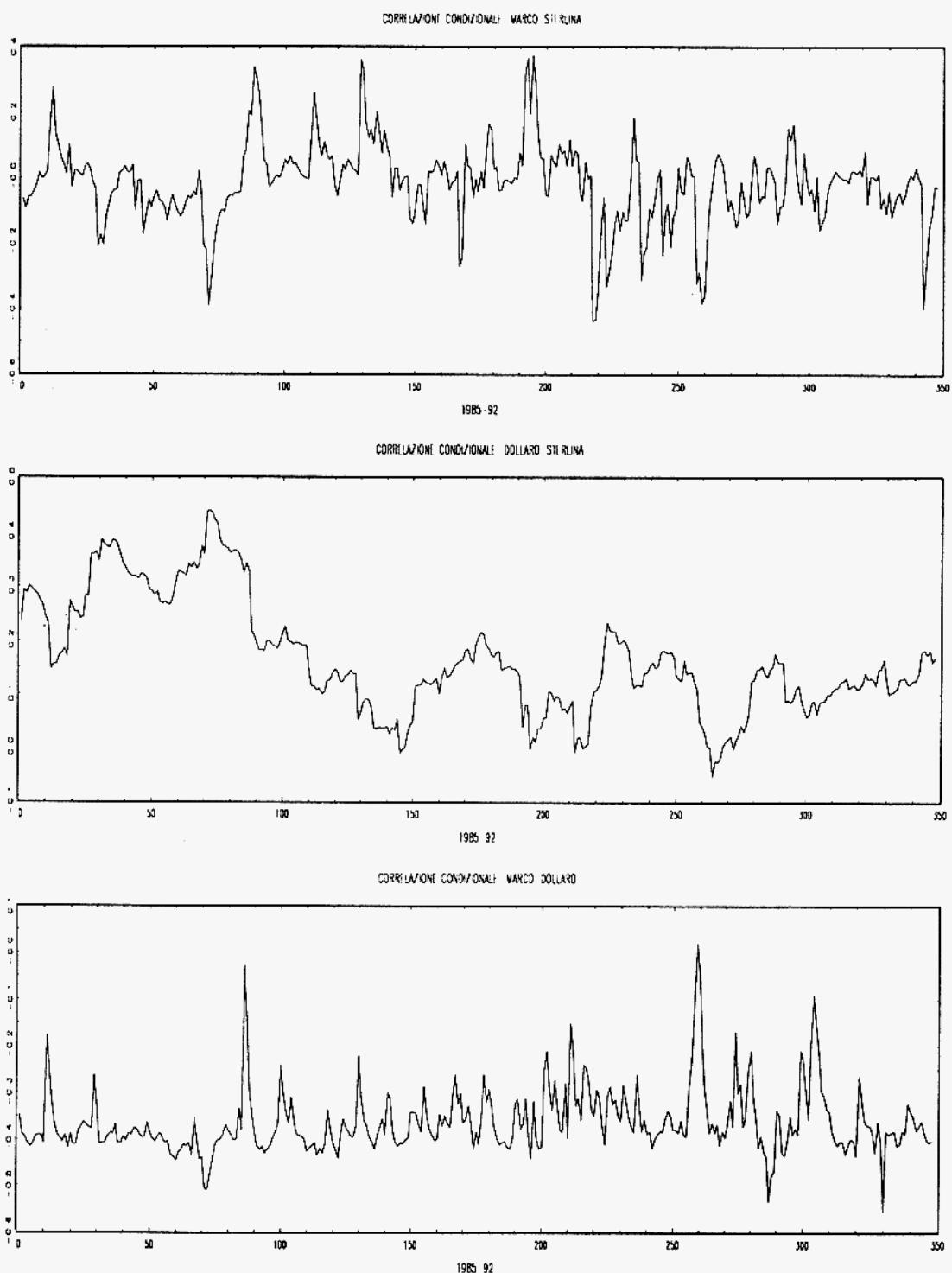


Fig.8



Bibliografia

- BAILLIE R.T., BOLLERSLEV T. (1989a) «The Message in daily exchange rates: a conditional variance tale». *Journal of Business and Economic Statistics*, 7, 297-305.
- BAILLIE R.T., BOLLERSLEV T. (1989b) «Common stochastic trends in a system of exchange rates». *The Journal of Finance*, 44, 167-181.
- BAILLIE R.T., BOLLERSLEV T. (1990a) «A multivariate generalized ARCH approach to modeling risk premia in forward foreign exchange rate markets». *Journal of International Money and Finance*, 16, 109-124.
- BAILLIE R.T., BOLLERSLEV T. (1990b) «Intra day and inter day volatility in foreign exchange rates». *Review of Economic Studies*, 58, 565-585.
- BAILLIE R.T., BOLLERSLEV T. (1992) «Prediction in dynamic models with time-dependent conditional variances». *Journal of Econometrics*, 52, 91-113.
- BAILLIE R.T., MCMAHON P.C. (1989) *The foreign exchange market. Theory and econometric evidence*. Cambridge, Cambridge University Press.
- BOLLERSLEV T. (1986) «Generalized autoregressive conditional heteroskedasticity». *Journal of Econometrics*, 31, 307-327.
- BOLLERSLEV T. (1990) «Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH approach». *Review of Economic Studies*, 498-505.
- BOLLERSLEV T., ENGLE R.F. (1993) «Common persistence in conditional variances». *Econometrica*, 61, 1, 167-186.
- BOLLERSLEV T., R.F. ENGLE, J.M. WOOLDRIDGE (1988) «A Capital asset pricing model with time-varying covariances». *Journal of Political Economy*, 96, 1, 116-131.
- BOLLERSLEV T., WOOLDRIDGE J. (1992) «Quasi maximum likelihood estimation and inference in dynamic models with time varying covariances». *Econometric Review*, 11, 143-172.
- BOOTHE P., GLASSMAN D. (1987) «The statistical distribution of exchange rates». *Journal of International Economics*, 22, 297-318.
- CAMPBELL J.Y., PERRON P. (1991) «Pitfalls and opportunities: what macroeconomists should know about unit roots», *NBER Macroeconomic Annual* (ed. by O.J. Blanchard-S. Fischer), 141-219.
- CORBAE D., OULIARIS S. (1986) «Robust tests for unit roots in the foreign exchange market». *Economics Letters*, 22, 375-380.
- DADDI P., TIVEGNA M. (1990) «Stima di equazioni di tasso di cambio: specificazioni GARCH mensili del cambio Dollaro-Marco». Centro di Ricerche Finanziarie, Gruppo IMI, Luglio, n.3.
- DICKEY A.D., BELL W.R., MILLER R.B. (1986) «Unit roots in time series models: tests and implications». *The American Statistician*, 55, 12-26.
- DIEBOLD F.X. (1988) *Empirical modeling of exchange rate dynamics*, Berlin, Springer Verlag.
- DIEBOLD F.X., NERLOVE M. (1989) «The dynamics of exchange rate volatility: a multivariate latent factor ARCH model». *Journal of Applied Econometrics*, 4, 1-21.

- DOAN T. LITTERMAN R.B., SIMS C.A., (1984) «Forecasting and conditional projection using realistic prior distributions». *Econometric Reviews*, 3, 1-100.
- DROST C.D., NIJMAN T.E. (1993), «Temporal aggregation of GARCH processes». *Econometrica*, 61, 909-27.
- DUCA A. (1992) *Modelli per l'analisi delle serie finanziarie un'applicazione ai tassi di cambio italiani della tecnica GARCH multivariati*. Tesi di laurea, Università di Ancona.
- ENGLE R.F. (1982) «Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation». *Econometrica*, 50, 987-1008.
- ENGLE R.F., BOLLERSLEV, T. (1986) «Modelling the persistence of conditional variances». *Econometric Reviews*, 5, 1-50, 81-87.
- ENGLE R.F., KRONER K.F. (1993) «Multivariate simultaneous generalized ARCH». Discussion Paper. Department of Economics. University of California, San Diego
- FULLER W.A. *Introduction to statistical time series*, New York: John Wiley.
- GOURIEROUX C. (1990) «Modèles ARCH». Mimeo.
- GOURIEROUX C., MONFORT A., TROGNON A. (1984) «Pseudo maximum likelihood theory». *Econometrica* 52, 681-700.
- HSIEH D.A. (1988) «The Statistical properties of daily foreign exchange rates: 1974-1983». *Journal of International Economics*, 24, 129-145.
- HSIEH D.A. (1989a) «Modeling heteroskedasticity in daily foreign exchange rates Changes». *Journal of Business & Economic Statistics*, 7, 3, 307-317.
- HSIEH D.A. (1989b) «Testing for nonlinear dependence in daily foreign exchange rate changes». *Journal of Business*, 62, 339-368.
- HODRICK R.J. (1989) «Risk, uncertainty, and exchange rates». *Journal of Monetary Economics*, 433-459.
- KIEFER K.G., SALMON M. (1983) «Testing normality in econometric models». *Economic Letters*, 11, 123-127.
- KIM K., SCHMIDT P. (1993) «Unit root tests with conditional heteroskedasticity». *Journal of Econometrics*, 59, 287-300.
- LAMOUREUX C.G., LASTRAPES W.D. (1990) «Persistence in variance, structural change and the GARCH model». *Journal of Business and Economic Statistics*, 8, 225-234.
- LASTRAPES W.D. (1989) «Exchange rate volatility and US monetary policy: an ARCH application». *Journal of Money, Credit, and Banking*, 21, 66-77.
- LEE J.H.H. «A Lagrange multiplier test for GARCH models». *Economics Letters*, 37, 265-271.
- LUCAS R.J. (1982) «Interest rates and currency prices in a two country world». *Journal of Monetary Economics*, 10, 335-59.
- MACDONALD R., TAYLOR M.P. (1992) «Exchange rate economics. A survey». *IMF Staff Papers*, 39, 1, 1-57.
- McFARLAND J.W., PETTIT R.R., SUNG S.K. (1982) «The distribution of foreign exchange price changes: trading day effects and risk measurement». *Journal of Finance*, 37, 693-715.
- MAGNUS J. (1988) *Linear structures*, London, Charles Griffin & Co.
- MANDELBROT B. (1963) «The variation of certain speculative prices». *Journal of Business*, 36, 394-419.

- MILHOJ A. (1987) «A conditional variance model for daily deviations of an exchange rate». *Journal of Business and Economic Statistics*, 5, 99-103.
- NELSON D.B. (1990) «Stationarity and persistence in the GARCH(1,1) model». *Econometric Theory*, 6, 318-334.
- NELSON D.B., CAO C.Q. (1992) «Inequality constraints in the univariate GARCH model». *Journal of Business & Economic Statistics*, 10, 229-235.
- PHILLIPS P.C.B., PERRON P. (1988) «Testing for a unit root in time series regression». *Biometrika*, 75, 335-46.
- WEI W.S. (1990) *Time series analysis*, Reading, Mass., Addison-Wesley.

Note:

¹ Per effetto week-end si intende il fatto che la relazione che esiste tra il venerdì ed il lunedì successivo differisce da quella tra due giorni lavorativi contigui. Questo è dovuto all'ammontare di informazioni che si accumula durante il giorno o i giorni non lavorativi. Per ovviare a questo problema si prende come dato settimanale quello del mercoledì, o, nel caso in cui il mercoledì non sia lavorativo si ricorre al dato del venerdì. I dati puntuali sono preferibili a quelli ottenuti costruendo una media dei dati giornalieri, in quest'ultimo caso infatti se il modello vero del tasso di cambio fosse un random walk, una media su base settimanale delle osservazioni presenterebbe correlazione seriale.

² Questo permette di evitare i problemi connessi con la diseguaglianza di Jensen (cioè $E[1/S_t] \neq 1/E[S_t]$ dove S_t è il tasso di cambio misurato in unità di moneta nazionale per moneta estera. Questo significa che, per esempio, il reciproco del tasso di cambio lira/marco è uguale al tasso di cambio marco/lira, ma lo stesso non è vero per i valori attesi); inoltre l'espressione $(1-L)\ln S_t$ (L è l'operatore ritardo, $x_{t-1} = Lx_t$) può essere interpretata come il cambiamento percentuale approssimato.

³ Quando si applicano metodi di quasi-massima verosimiglianza, ovvero quando si abbandona l'ipotesi di normalità condizionale dei disturbi, la distribuzione degli errori di previsione non è più deducibile dalla distribuzione dei disturbi. Questo significa che si devono stimare in modo non parametrico i quantili della distribuzione degli errori di previsione. Su questo punto l'autore collabora a una ricerca presso l'Università di Pavia.