WORKING PAPER n.03.01

January 2003

Second order causality: a new model and an application to the FIB30 returns and volume

M. Caporin[a]

[a] GRETA, Venice

Second order causality: a new model and an application to the FIB30 returns and volume

Massimiliano Caporin
Unviersità di Venezia and GRETA
mcaporin@unive.it

February 8, 2003

Abstract

In this paper we suggest a new bivariate GARCH model to detect second order causality. The main feature of this model is that it detect not only the causality existence but also its direction, a characteristic not yet taken into consideration. We present at first a survey on the definitions of causality in mean and in variance including a first result: we derived the parameter restrictions needed to verify mean and variance causality within a VAR-GARCH-M model. We consider then the current multivariate models used to detect second order causality, showing their drawback and then suggesting a new model that could be used to detect both causality existence and direction. Finally we applyied our model on a real case, using the FIB30 return and volume series. In this last part of our paper we present at first the preliminar analysis on the FIB30 data, in particular with respect to the study of the cyclical components of the market and on the memory structure of both returns and volumes. We consider ther the estimation of univariate models for both series. We shift then our attention to a multivariate setting comparing our GARCH model with causality to a simple CCC-GARCH. We show that information criteria clearly prefer our new specification.

In the last years there has been a growing interest in the study of the relation among prices and volumes, both from a theoretical point of view (as an example Blume, Easley and O'Hara, 1994) and from the empirical one, see among others Karpoff (1987). Most of the current empirical analysis considers different linear and non-linear specifications to verify and test the causal relations between prices (or returns) and volumes. However most of them consider only the mean, restricting their attention on Granger's causality definition or to the study of a asimultaneous relation. In the last decades with the emerging ARCH literature different specifications of conditional heteroskedastic have been taken into considerations and allow for a deeper analysis in applied studies on the causality topic, efforts that allow to adequately model the relation between returns, volumes and their volatility. These extensions can be thought both of one signle asset case that in a much more general framework. This interest on multivariate heteroskedastic models maybe coupled with the necessity of an extension of the causality concept, which must consider the spillover effect among variances, and the in-mean reaction of GARCH model components. This will be the object of our work.

After a brief review of the definitions of causality in mean and in variance (Engle and Granger (1986), as reviewed by Comte and Lieberman (2000)), in this paper we will analyze in detail the different approaches that have been used up to this moment to identify the presence of causality among variances. We will present two different approaches, the one of Cheung and Ng (1996) that is based on cross correlation analysis of the residuals of univariate models and the approaches based on multivariate GARCH formulations. All the up-to-date works in this field share a common problem: they can infer about the presence of causality but not on its direction, that is, given two assets A and B, assuming that there exist a causal relation among their variances, current model detect this relations but cannot tell us if an increase in the variance of A will imply an increase or a decrease in the variance of B. An interesting approach in this area is given by Hafner (2001), who provide a measure for causality in a multivariate GARCH framework, however its study do not directly include the causal relation into a model. In this paper we will try to solve this problem in a multivariate framework, considering an extension of multivariate GARCH models that could be used to test both the existence of causality among variances and its direction. The suggested formulation will be tested on an empirical basis, studying the relation between the returns and the volume of the FIB30 market, a future on the Italian stock exchange index. In this paper we will also analyze and identify the cyclical patterns of the FIB30 contract.

The plan of the paper is as follow: in section 1 we review the current theoretical framework on causality both for the mean and the variance while in section 2 we focus on different alternative models to verify second order causality. The remaining of the paper is devoted to the case study of the FIB30 contract, the future for the Italian stock market index. After a brief description of the market and of the contract in section 3 we analyze the dataset used and the procedure of extraction and filtration of the series of interest. We switch then to the univariate analysis of the returns and volume series of the FIB30, followed by the multivariate and causality analysis. Section 4 will conclude.

1 Causality in mean and in variance

This section is devoted to a brief review of the concept of causality in mean, causality in variance and to a first analysis that involve the techniques developed by Cheung and Ng (1996). We refer to the definition of causality given by Granger in its well know 1980 seminal paper. First of all a minimal notation is given.

Define X_t as the n-dimensional set of variables of interest at time t, this set can be partitioned into $X_t = \{X_{1,t}\ ,\ X_{2,t}\}$, that have, respectively, dimension n_1 and n_2, denote by $I(X_t) = I_t(X) = I(X)$ the information set (a sigma algebra generated by the variable of interest or in general an Hilbert space) for the whole variables and with $I(X_{1,t}) = I_t(X_1) = I(X_1)$ the information set given by the partition 1 (similarly for X_2).

Then we recall the following statement:

Proposition 1.1 *Granger (1980): X_2 does not cause X_1 in Granger sense, if and only if $E_t[X_{1,t}|I_{t-1}(X)] = E_t[X_{1,t}|I_{t-1}(X_1)]$.*

This will be denoted by $X_2 \overset{G}{\nrightarrow} X_1$

The violation of the previous condition will be referred as causality in the mean. However a contemporaneous bidirectional relation is not included in the Granger definition, in this case we can refer to Sims (1972) definition of bidirectional non-causality. This is stated as:

Proposition 1.2 *Sims (1972): there is no bidirectional causality between X_1 and X_2 if and only if $Cov[X_{1,t} - E_t[X_{1,t}|I_{t-1}(X)]\ , X_{2,t} - E_t[X_{2,t}|I_{t-1}(X)]] = 0$.*

This will be denoted by $X_2 \nleftrightarrow X_1$

However, in dealing with time varying conditional variances and causal relation among these quantities, we need an extension to this concept to consider causality among variances, the source of the spillover effects studied in financial markets. The extension is provided by the following proposition due to Engle, Granger and Robins (1986):

Proposition 1.3 *Engle, Granger and Robins (1986): X_2 does not second order cause X_1 in Granger sense, if and only if $E_t\left[(X_{1,t} - E_t[X_{1,t}|I_{t-1}(X)])^2 |I_{t-1}(X)\right] = E_t\left[(X_{1,t} - E_t[X_{1,t}|I_{t-1}(X)])^2 |I_{t-1}(X_1)\right]$.*

This will be denoted by $X_2 \overset{G^2}{\nrightarrow} X_1$

As we can see, the definition of second order non-causality does not presume any causal relation in the mean, however, this is not precisely a non-causality relation among variances. Comte and Lieberman in their 2000 paper on second order non-causality gave a different definition:

Proposition 1.4 *: X_2 does not cause X_1 in variance, if and only if $V_t[X_{1,t}|I_{t-1}(X)] = V_t[X_{1,t}|I_{t-1}(X_1)]$.*

This will be denoted by $X_2 \overset{G_V}{\nrightarrow} X_1$

The two authors gave also the following relation:

Remark 1.1 *: $X_2 \overset{G}{\nrightarrow} X_1 + X_2 \overset{G^2}{\nrightarrow} X_1 \Longleftrightarrow X_2 \overset{G_V}{\nrightarrow} X_1$*

Proof. By substitution and with a direct application of the law of iterated expectation. ■

This last remark allow us to note that non-causality in the variance exist if and only if there exist non-causality in the mean, moreover first and second order non-causality may combine in all possible pairs. This allow a sequential testing scheme, at first causality in the mean, if there is no relation, we can test for second order non causality. Moreover only if both tests lead to a no-relation result we can conclude that there is non-causality among variances.

From an empirical point of view we are also interested in verifying the previous relations, we will terefore now show what are the restrictions implied by first and second order non-causality in a very general framework, using as a reference model a VARMA-GARCH.

The benchmark model can be represented as follow: given the variables of interest X_t we consider the following VARMA(p,q)-GARCH($\overline{p}$,$\overline{q}$) model

$$\begin{aligned} X_t &= A(L) X_t + B(L) E_t \qquad E_t \tilde{} \, iid(\mathbf{0}, H_t) \\ H_t &= \boldsymbol{\omega} + C(L) H_t + D(L) [E_t E_t'] \end{aligned} \tag{1}$$

where $A(L) = \sum_{i=1}^{p} A_i L^i$, $B(L) = \sum_{i=1}^{q} B_i L^i$, $C(L) = \sum_{i=1}^{\overline{p}} C_i L^i$, $D(L) = \sum_{i=1}^{\overline{q}} D_i L^i$, and $\boldsymbol{\omega}$, A_i, B_i, C_i, D_i, are all square matrices of dimension n, while p, q, $\overline{p}$ and $\overline{q}$ are intereger numbers. Assume also that the model is stationary and invertible. A similar approach was also used by Comte and Lieberman (2000) and Boudjellaba, Dufour and Roy (1992 and 1994) in giving a set of parametric restrictions and tests for causality, we recall in the following their results. In testing first order noncausality it is convenient to transform the VARMA(p,q) into its AR(∞) representation (given invertibility assumption)

$$\begin{aligned} [B(L)]^{-1} [1 - A(L)] X_t &= W(L) X_t = E_t \\ where\ W(L) &= \sum_{i=0}^{\infty} W_i L^i \end{aligned} \tag{2}$$

then $X_2 \overset{G}{\nrightarrow} X_1$ if and only if $[W_i]_{12} = 0$ for all lags i, that is the coefficients that link the variables included in the 2 partitions on X, are identically equals to zero (Boudjellaba et al. 1992). As noted by Comte and Lieberman (2000) if we drop the GARCH part of the model and consider a simple VARMA process with constant variance-covariance matrix, we will always have second order noncausality. For the GARCH part, similarly with the VARMA, we convert the model into its ARCH(∞) representation:

$$H_t = [1 - C(L)]^{-1} \boldsymbol{\omega} - [1 - C(L)]^{-1} D(L) [E_t E_t'] = \overline{\boldsymbol{\omega}} + U(L) [E_t E_t']$$

$$where\ U(L) = \sum_{i=0}^{\infty} U_i L^i$$

For the moment we do not assume any of the traditional forms used in testing causality among variances, in particular the Vech or BEKK representations of Engle and Kroner (1995) since they can be obtained reparameterizing the previous relation, which is a much more general one. A causality restriction similar to the one of VARMA holds here: $X_2 \overset{G^2}{\nrightarrow} X_1$if and only if $[U_i]_{12} = 0$ for all lags i. In this framework noncausality of the first and of the second order can independently exist. If we have both then we have also variance noncausality. For an example of a VARMA(1,1)-GARCH(1,1) refer to the paper of Comte and Lieberman (2000). However none of the previous cited papers deal with the case of a VARMA-GARCH-M model, in which we add an additional source of causality, the one of variances on the mean of the process. In this case how are modified the conditions for first and second order noncausality? or more precisely, the implication of remark 3.1.1 are still valid or need an update? Let us consider the following extension of equation (1):

$$\begin{aligned} X_t &= A(L) X_t + B(L) E_t + GVech(H_t) \qquad E_t \tilde{\ } iid(\mathbf{0}, H_t) \\ H_t &= \boldsymbol{\omega} + C(L) H_t + D(L) [E_t E_t'] \end{aligned} \tag{3}$$

where the operator $Vech$ stacks the lower triangular element of H_t, therefore $Vech(H_t)$ is of dimension $r = n(n+1)/2$ and G of dimension $n \times r$. We start considering second order noncausality: given the previous definition, that do not presume any causal or noncausality relation on the mean, the restrictions are the same as in the previous case, that is rewriting the model in its ARCH(∞) representation, again $X_2 \overset{G^2}{\nrightarrow} X_1$ if and only if $[U_i]_{12} = 0$. The difference is in the first order noncausality: now we have dependence of returns from variance-covariance matrix, we can in principle distinguish two cases depending on the existence of second order noncausality. We can state the following:

Remark 1.2 *In a stationary and invertible VARMA-GARCH-M model, with* $\left|[G]_{i,j}\right| \geq 0$ $i = 1, ...n_1$ $j = 1, ...r$*, and at least one coefficient for which strict inequality hold, if there is second order causality there is also first order causality.*

Proof. Consider the condition for noncausality in the mean $E_t[X_{1,t}|I_{t-1}(X)] = E_t[X_{1,t}|I_{t-1}(X_1)]$, substituting $X_{1,t}$ with its expression from (3), we have

$$E_t\left[[A(L) X_t + B(L) E_t + GVech(H_t)]_{1,1} | I_{t-1}(X)\right]$$

the first two components maybe measurable with respect to the information set restricted to the past of X_1 but this is not true for H_1 given the presence of second order noncausality, H_1 is measurable only on the whole information set. This is true if at least one of the coefficients linking the variables in X_1 with the variance-covariance matrix is different from zero, in the opposite case we could write a restricted VARMA model for X_1 without the in-mean component, returning to a situation similar to the VARMA-GARCH approach. ■

Assume now that there is second order noncausality, in this case we can state the following:

Remark 1.3 *Consider a stationary and invertible VARMA-GARCH-M model, where, for simplicity of notation, we reorder the in-mean component as follows*

$$Vech\left(H_t\right) \to Vech\left(\left[H_t\right]^{T2}\right)$$

where with $[\cdot]^{T2}$ *we mean a transpose with respect to the secondary diagonal of a matrix. In this model,* $X_2 \overset{G}{\nrightarrow} X_1$ *if the following conditions are satisfied: i)* $[W_i]_{12} = 0$ *for all lags* i*; ii)* $[U_l]_{i,j} = 0$, $i = 1, ...n_1$, $j = 1, ...(r-n_1(n_1+1)/2)$, *for all lags* l*. Where* $W(L)$ *is defined as in (2) and* $U(L)$ *is defined as* $[B(L)]^{-1} G = \sum_{i=0}^{\infty} U_i L^i = U(L)$, *is a sequence of matrices of dimension* $n \times r$

Proof. Again referring to the measurability with respect to the information sets, violating one of the previous condition will imply non-measurability with respect to the restricted information set of X_1, i) concern with dependance from the variables included in X_2 while ii) is devoted to the dependance of X_1 only from its own variance covariance matrix. The reordering allow us to concentrate the element of the variance covariance matrix of X_1 at the end of the vector of in-mean effects. ■

Summarizing our results we have the following implications:

Remark 1.4 *In a stationary and invertible VARMA-GARCH-M the following relations hold*

i) $X_2 \overset{G}{\nrightarrow} X_1 + X_2 \overset{G^2}{\nrightarrow} X_1 \Longleftrightarrow X_2 \overset{G_V}{\nrightarrow} X_1$

ii) $X_2 \overset{G^2}{\to} X_1 \Longrightarrow X_2 \overset{G}{\to} X_1$ *if* $\left|[G]_{i,j}\right| \geq 0$ $i = 1, ...n_1$ $j = 1, ...r$ *with* $\left|[G]_{i,j}\right| > 0$ *for at least one* (i, j)

iii) $X_2 \overset{G}{\to} X_1 \nRightarrow X_2 \overset{G^2}{\to} X_1$

iv) $X_2 \overset{G^2}{\nrightarrow} X_1 \nRightarrow X_2 \overset{G}{\nrightarrow} X_1$

Now we will describe two different ways of detecting second order causality without inglobating it in a model.

1.1 A measure for variance causality

In Hafner (2001) we can find a set of measures of second order causality between X_1 and X_2. These measures are derived under the hypothesis that the whole system is driven by a multivariate GARCH process of the form

$$H_t = \boldsymbol{\omega} + \sum_{j=1}^{p} C_j H_{t-j} + \sum_{j=1}^{q} D_j \left[E_{t-j} E'_{t-j}\right] \tag{4}$$

where E_t is the vector of mean residuals. The model can be reformulated in a companion VARMA representation: define $V_t = E_t E'_t - H_t$, then

$$\left[E_t E'_t\right] = \boldsymbol{\omega} + \sum_{j=1}^{\max(p,q)} \left(C_j + D_j\right) \left[E_{t-j} E'_{t-j}\right] - \sum_{j=1}^{p} C_j V_{t-j} + V_t \tag{5}$$

The variance covariance matrix of the error term is defined as $\sum(V)$. Hafner (2001) suggest to estimate a full GARCH model including all variables, that is using X, and two lower dimension GARCH for X_1 and X_2. The variance covariance matrices of error term for the two restricted GARCH are labelled as $\sum_V^{X_1}$ and $\sum_V^{X_2}$. Moreover we can partition $\sum(V)$ in the following way

$$\sum(V) = \begin{bmatrix} \sum_{V(X_1)} & \sum_{V(X_1 X_2)} \\ \sum_{V(X_1 X_2)} & \sum_{V(X_2)} \end{bmatrix}$$

The causality measures are then defined as

$$\begin{aligned} a)\ \mathcal{CV}\left(X_2 \overset{G^2}{\to} X_1\right) &= \ln \frac{\left|\sum_V^{X_1}\right|}{\left|\sum_{V(X_1)}\right|} \\ b)\ \mathcal{CV}\left(X_1 \overset{G^2}{\to} X_2\right) &= \ln \frac{\left|\sum_V^{X_2}\right|}{\left|\sum_{V(X_2)}\right|} \\ c)\ \mathcal{CV}\left(X_1 \overset{G^2}{\longleftrightarrow} X_2\right) &= \ln \frac{\left|\sum_{V(X_2)}\right|\left|\sum_{V(X_2)}\right|}{\left|\sum_{V(X_1 X_2)}\right|} \\ d)\ \mathcal{CV}\left(X_1 \overset{L}{\longleftrightarrow} X_2\right) &= \ln \frac{\left|\sum_V^{X_1}\right|\left|\sum_V^{X_2}\right|}{\left|\sum_{V(X_1 X_2)}\right|} \end{aligned}$$

where $a)$ and $b)$ are second order causality measures between X_1 and X_2, $c)$ is a measure of instantaneous causality in volatility while $d)$ is a measure of linear dependence. In all cases above listed we have non-causality when the measure is zero, moreover the following relation holds

$$\mathcal{CV}\left(X_1 \overset{L}{\longleftrightarrow} X_2\right) = \mathcal{CV}\left(X_1 \overset{L}{\longleftrightarrow} X_2\right) + \mathcal{CV}\left(X_1 \overset{G^2}{\to} X_2\right) + \mathcal{CV}\left(X_1 \overset{G^2}{\longleftrightarrow} X_2\right)$$

In empirical analysis the statistic has to be estimated and under the assumption of normality of the error terms in the companion VARMA representation a standard likelihood ratio statistic can be used to test the null hypothesis of $\mathcal{CV}(\cdot) = 0$, the statistic will be computed as $n\widehat{\mathcal{CV}}(\cdot)$ with a presumed χ^2 distribution (not yet been proven). These measure could be very attractive, however they are of difficult computation in large scale systems given the number of parameters that have to be estimated even with low order GARCH. In small systems they could be of interest and also extended modifying the assumption on the underlying GARCH structure to allow constant or dynamic conditional correlation representations. However, the previous techniques just detect causality without including it in the model.

1.2 The approach of Cheung and Ng

In a recent paper, Cheung and Ng (1996) considered the problem of detecting second order causality and including it in a multivariate setting. They considered a bivariate model and tested the hypothesis of causality between the variances (time dependent and following a GARCH model) of the to series considered. They focused their attention on the cross-correlations between the standardized and squared standardized residuals of the univariate models they fitted on each series. Recall that the sample cross-correlations between two generic series X_t and Y_t can be expressed as follows

$$\rho_{XY}(k) = \frac{1}{T\sigma_X^2\sigma_Y^2}\sum_{i=1}^{T}(X_i - \mu_X)(Y_{i-k} - \mu_Y) \qquad k = 0, \pm1, \pm2...$$

where μ_X, μ_Y, σ_X^2, σ_Y^2 are respectively sample means and sample variances of the two series, and T is sample size. Moreover a result of Hannan (1970) shows that

$$\begin{pmatrix} \sqrt{T}\rho_{XY}(k) \\ \sqrt{T}\rho_{XY}(k') \end{pmatrix} \to N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) \qquad k \neq k' \tag{6}$$

therefore Cheung and Ng (1996) suggested to test the hypothesis of non-causality by a simple signifiativity test, comparing $\sqrt{T}\rho_{XY}(k)$ with the standard normal distribution. They provided also a theorem showing that the asymptotic limit of Hannan (1970)) (6) is valid also when the variables are not directly observed but have to be estimated (this is the case of the standardized residuals of an ARMA-GARCH model) and suggested also the use of a chi-square test statistics to verify causality from lag i up to a certain order k: this is simply

$$\Psi = T\sum_{j=i}^{k}[\rho_{XY}(j)]^2 \sim \chi^2(k - i + 1)$$

Whenever a significant cross-correlation is found across standardized residual or squared standardized residuals the model is corrected adding the relevant lagged terms in the univariate ARMA or GARCH equation. The methodology can be clarified with an example: consider two generic series X_t and Y_t which have been modeled by the following models

$$\begin{aligned} X_t &= \phi_{1,1}X_{t-1} + \theta_{1,2}\varepsilon_{t-2} + \varepsilon_t \\ \sigma_{X,t}^2 &= \omega_X + \alpha_{1,1}\varepsilon_{t-1}^2 + \beta_{1,1}\sigma_{X,t}^2 \end{aligned} \tag{7}$$

$$\begin{aligned} Y_t &= \phi_{2,2}Y_{t-2} + \theta_{2,1}\eta_{t-1} + \eta_t \\ \sigma_{Y,t}^2 &= \omega_Y + \alpha_{2,1}\eta_{t-1}^2 + \beta_{2,1}\sigma_{Y,t}^2 \end{aligned} \tag{8}$$

moreover the cross-correlations computed between $\tilde{\varepsilon}_t = \varepsilon_t/\sigma_{X,t}$ and $\tilde{\eta}_t = \eta_t/\sigma_{Y,t}$ resulted significative at lead 1 and lag 2 ($\tilde{\varepsilon}_t$ lead/lag $\tilde{\eta}_t$), while the ones computed between $\tilde{\varepsilon}_t^2$ and $\tilde{\eta}_t^2$ turned out to be relevant at lead 3 and lag 4. Following the suggestions of the cross-correlations functions the models (7) and (8) are updated to

$$\begin{aligned} X_t &= \phi_{1,1}X_{t-1} + \theta_{1,2}\varepsilon_{t-2} + \varepsilon_t + \psi_{1,1}Y_{t-1} \\ \sigma^2_{X,t} &= \omega_X + \alpha_{1,1}\varepsilon^2_{t-1} + \beta_{1,1}\sigma^2_{X,t} + \lambda_{1,3}X^2_{t-3} \end{aligned} \tag{9}$$

$$\begin{aligned} Y_t &= \phi_{2,2}Y_{t-2} + \theta_{2,1}\eta_{t-1} + \eta_t + \psi_{2,1}X_{t-2} \\ \sigma^2_{Y,t} &= \omega_Y + \alpha_{2,1}\eta^2_{t-1} + \beta_{2,1}\sigma^2_{Y,t} + \lambda_{2,4}Y^2_{t-4} \end{aligned} \tag{10}$$

Cheung an Ng (1996) suggest to add components based on the observations, as exogenous variables, both in the mean and variance equations of each of the series modelled. This allow to examine relations among a large number of series without considering large models, that may be very difficult to implement if we require a time varying conditional variance. However even the authors pointed out that this technique may be influenced by the structure of the univariate models employed and also crucially by the sample size. We would stress that this approach may be interesting in large scale analysis while in dealing with a small number of variable it will be necessarily dominated by a multivariate GARCH approach that allow to dynamically handle the causality relations that could be evidenced by the cross-correlations analysis. Moreover the corrections suggested by the authors act as in transfer function models, with an exogenous variable influencing the dynamics of another one, they do not consider multivariate modelling that could exploit different behavior of the model and surely affect parameter estimates. Therefore we recommend a careful use of this methodology preferring, whenever possible, the multivariate GARCH approach, that will be the object of next section.

2 Multivariate analysis and causality

Various works dealing with second order causality applied to test the relation among variance different GARCH-type models, see among other the paper of Comte and Lieberman (2000). Here we deal only with the application of this kind of models in causality testing among variances, presenting the different current approaches and a personal viewpoint.

2.1 Traditional models: BEKK, Vech and their drawback

Most empirical works dealing with second order causality considered multivariate GARCH in the BEKK and Vech representations. Engle and Kroner (1995) showed that the two formulations can be derived one from the other and viceversa with an adequate reparameterization. For the moment we assume that the specification chosen is the BEKK, represented as:

$$H_t = \boldsymbol{\omega} + \sum_{i=1}^{p} C_i H_{t-i} C_i' + \sum_{j=1}^{q} D_j E_{t-j} E_{t-j}' D_j'$$

where C_i and D_j are $n \times n$ matrices and $\boldsymbol{\omega}$ is a symmetric positive definite $n \times x$ matrix. The existence of any causal relation among the variances and covariances included in H_t imply that (at least some of) the off-diagonal coefficients of C_i and D_j are different from zero. If all the parameter matrices are diagonal the model collapse to a particular case in which all conditional variances and covariances follow a GARCH(p,q) process. Consider as an example the following bivariate BEKK-GARCH(1,1), constants are dropped for simplicity:

$$\begin{aligned} \begin{bmatrix} \sigma_{1,t}^2 & \sigma_{12,t} \\ \sigma_{12,t} & \sigma_{2,t}^2 \end{bmatrix} &= \dots \begin{bmatrix} \beta_{1,1} & \beta_{1,2} \\ 0 & \beta_{2,2} \end{bmatrix} \begin{bmatrix} \sigma_{1,t-1}^2 & \sigma_{12,t-1} \\ \sigma_{12,t-1} & \sigma_{2,t-1}^2 \end{bmatrix} \begin{bmatrix} \beta_{1,1} & 0 \\ \beta_{1,2} & \beta_{2,2} \end{bmatrix} + \\ & \begin{bmatrix} \alpha_{1,1} & 0 \\ 0 & \alpha_{2,2} \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1}^2 & \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} & \varepsilon_{2,t-1}^2 \end{bmatrix} \begin{bmatrix} \alpha_{1,1} & 0 \\ 0 & \alpha_{2,2} \end{bmatrix} \end{aligned}$$

note that the symmetry in the parameter matrices is not required since in any quadratic form the symmetry is given by the inner matrix. This simple model imply the following relations

$$\begin{aligned} \sigma_{1,t}^2 &= \beta_{1,1}^2\sigma_{1,t-1}^2 + 2\beta_{1,1}\beta_{1,2}\sigma_{12,t-1} + \beta_{1,2}^2\sigma_{2,t-1}^2 + \alpha_{1,1}^2\varepsilon_{1,t-1}^2 \\ \sigma_{12,t} &= \beta_{1,1}\beta_{2,2}\sigma_{12,t-1} + \beta_{1,1}\beta_{1,2}\sigma_{2,t-1}^2 + \alpha_{1,1}\alpha_{2,2}\varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \sigma_{2,t}^2 &= \beta_{2,2}^2\sigma_{2,t-1}^2 + \alpha_{2,2}^2\varepsilon_{2,t-1}^2 \end{aligned}$$

introducing causality from the second variable both an the first and on covariances, moreover causality from covariances to the first variable, this at a cost of a single additional parameter. This is the most important feature of the BEKK model, various nonlinear relations can be imposed with a limited number of parameters, that are also free of any constraints since are implemented in a quadratic form (think of positivity constraint for parameters, or the constraints needed to impose positive definiteness of the variance-covariance matrix). However the number of parameters greatly increase with the number of variables creating a series of problems on convergence of the estimation algorithm, reliability of the estimates and last but not least CPU time. An additional remark is also needed if we test causality with zero restrictions on parameter matrices: here the model postulate that causality among variances (exclude covariances for the moment) act only in one direction, that is the positive one. Consider the previous example: a shock to this system (in $\varepsilon_{2,t}$), that will affect only the second variable, causing an increase in its variances will cause necessarily an increase in the variance of the first series, the possibility that the first variances decrease is in principle not contemplated. This is a particular situation, maybe very difficult to realize in financial markets, but that we cannot a priori exclude. Moreover the same shock that affect $\sigma_{2,t}^2$ will move directly (through $\varepsilon_{2,t}$) and indirectly (through $\sigma_{2,t}^2$) the covariance and we are not able to exclude this second effect in such a model. All these point are related to the quadratic parameter structure that imply also another problem, only combinations of parameters are responsible for the non-linear relations between variables, we cannot therefore directly interpret the estimates of a BEKK formulation. Signifiativity tests run on the BEKK parameter are no more valid in testing the significance in the single equation GARCH.

The Vech formulation allow, as the BEKK, for shock transmission among variances and covariances, but the parameters increase with respect to the BEKK formulation. Moreover we face an additional problem, we must bound parameters in such a way they are all positive and that they guarantee the positive definiteness of the variance covariance matrix.

2.2 The CCC-GARCH and the approach of Brunetti and Gilbert

Consider now the constant correlation GARCH, this model can be easily generalized to allow for causality among variances. This model postulate that covariances depend on variances through a scale parameter and that variances follow a GARCH. We can therefore model the pure variance process in the following way

$$\begin{bmatrix} \sigma^2_{1,t} \\ \vdots \\ \sigma^2_{n,t} \end{bmatrix} = \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix} + \begin{bmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & \ddots & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{bmatrix} \begin{bmatrix} \varepsilon^2_{1,t-1} \\ \vdots \\ \varepsilon^2_{n,t-1} \end{bmatrix} \tag{11}$$

$$+ \begin{bmatrix} \beta_{1,1} & \cdots & \beta_{1,n} \\ \vdots & \ddots & \vdots \\ \beta_{n,1} & \cdots & \beta_{n,n} \end{bmatrix} \begin{bmatrix} \sigma^2_{1,t-1} \\ \vdots \\ \sigma^2_{n,t-1} \end{bmatrix} \tag{12}$$

where parameters will have to be bounded above zero to ensure positivity of variances. Stationarity of the process is then affected by this structure, and we need to impose additional restrictions. The model con be reformulated in a companion VARMA representation: define $\nu_{i,t} = \varepsilon^2_{i,t-1} - \sigma^2_{i,t}$ for $i = 1, 2...n$ we can write

$$\begin{bmatrix} \varepsilon^2_{1,t} \\ \vdots \\ \varepsilon^2_{n,t} \end{bmatrix} = \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix} + \left(\begin{bmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & \ddots & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{bmatrix} \right. \tag{13}$$

$$\left. + \begin{bmatrix} \beta_{1,1} & \cdots & \beta_{1,n} \\ \vdots & \ddots & \vdots \\ \beta_{n,1} & \cdots & \beta_{n,n} \end{bmatrix} \right) \begin{bmatrix} \varepsilon^2_{1,t-1} \\ \vdots \\ \varepsilon^2_{n,t-1} \end{bmatrix} +$$

$$- \begin{bmatrix} \beta_{1,1} & \cdots & \beta_{1,n} \\ \vdots & \ddots & \vdots \\ \beta_{n,1} & \cdots & \beta_{n,n} \end{bmatrix} \begin{bmatrix} \nu_{1,t-1} \\ \vdots \\ \nu_{n,t-1} \end{bmatrix} + \begin{bmatrix} \nu_{1,t} \\ \vdots \\ \nu_{n,t} \end{bmatrix}$$

and recalling that $\nu_{i,t}$ is a martingale difference sequence we can state that the stationarity conditions for (11) are exactly the same of a VARMA(1,1) model, we need that all the eigenvalues of the matrices be outside the unit root circle.

This model, as the BEKK of the Vech, detect only positive causality and imply an elevate number of parameters.

The approach of Brunetti and Gilbert (2000) can be embodied in this setup and viewed as an extension allowing long memory. They considered a particular bivariate CCC-GARCH model, allowing each series to include long memory and using a structure similar to (11). However, given the particular long memory structure, the parameter matrix that links the variances must be diagonal, as we show in Caporin (2002), ruling out any possible causal relation.

2.3 A new approach

A common drawback of traditional GARCH models applied to second order causality testing is the elevate number of parameters together with the necessary constraints to ensure positivity of conditional variances. These points lead to complex numerical evaluations in the estimation of parameters that transfer in increasing and often unrealistic CPU time. This lead to the choice of simple models that do not consider causality among variances, think of the traditional CCC-GARCH models applied in finance, a area where time really matter and cannot be wasted waiting for some results and extrapolations on the effect of a shock. We suggest here an alternative methodology that can be used in testing for the presence of second order causality. We try to solve the problem imposed by the constraint on parameters via a multiplicative effect between variances. This extension was suggested by a group of papers dealing with switching GARCH and treshold models, among these we mention Hamilton (1994) and McAleer (2001), and it is mainly derived from the ideas of the first author, that proposed a switching structure for ARCH models in a simple way, premultiplying the ARCH equation by a state dependant factor. In its framework the state variable was unonberved and driven by a Markow chain. Now if we presume the existence of a causal relation among variances and we model covariances via a constant conditional correlation structure we propose the following bivariate testing procedure for causality. Assume that the model can be represented as

$$\begin{aligned} \begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} &= \begin{bmatrix} \mu_{1,t}\left(I^{t-1}\right) \\ \mu_{2,t}\left(I^{t-1}\right) \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \\ & \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \sim iid \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{1,t} & \rho\sigma_{1,t}\sigma_{2,t} \\ \rho\sigma_{1,t}\sigma_{2,t} & \sigma^2_{2,t} \end{bmatrix} \right) \end{aligned} \tag{14}$$

where the mean dynamic is not specified and we allow for time dependence based on the information set up to time $t-1$ $\left(I^{t-1}\right)$ including therefore also

GARCH-in-mean effects. The variances are represented as

$$\sigma_{1,t}^2 = \exp\left[f_1\left(I^{t-1}\right)\right]\left[\omega_1 + \sum_{j=1}^{p}\beta_{1,j}\sigma_{1,t-j}^2 + \sum_{j=1}^{q}\alpha_{1,j}\varepsilon_{1,t-j}^2\right] \quad (15)$$

$$\sigma_{2,t}^2 = \exp\left[f_2\left(I^{t-1}\right)\right]\left[\omega_2 + \sum_{j=1}^{p}\beta_{2,j}\sigma_{2,t-j}^2 + \sum_{j=1}^{q}\alpha_{2,j}\varepsilon_{2,t-j}^2\right] \quad (16)$$

where we included a standard GARCH model but there are no constraint to consider any possible GARCH formulation such as FIGARCH, the leverage GARCH of Glosten, Jagannathan and Runkle (1993) or the asymmetric power ARCH of Ding, Granger and Engle (1993) as well as we do not constrain the two model to have the same structure. The causal relation is modeled by the functions $f_1\left(I^{t-1}\right)$ and $f_2\left(I^{t-1}\right)$ that depends on the information sets up to time $t-1$. We could in principle suggest two possible specifications of these functions (we report only the first $f_1(\cdot)$, the second is similar):

$$\begin{array}{lll} a) & f_1\left(I^{t-1}\right) & = \gamma_1\sigma_{2,t}^2 \\ b) & f_1\left(I^{t-1}\right) & = \gamma_1\varepsilon_{2,t-1}^2 \end{array} \quad (17)$$

where we considered a direct effect between variances (case a), recalling the variances at time t are computed conditionally on the information set up to time t-1. In the second case (b) we exploited the relation via the squared of mean residuals, that can be viewed as a proxy for the conditional variance. Unfortunately the case (a) is of difficult solution, given the contemporaneous relation between the two conditional variances. The system cannot be transformed in a "nice" reduced form causing a problem in the implementation: to calculate each of the two conditional variances for any time value t we need a numerical evaluation algorithm. Therefore we prefer the direct introduction of lagged innovations, or the introduction of lagged conditional variances, however we stress that in this last case we condition on the information set available up to time $t-2$ (recall that $\sigma_{1,t}^2$ depends on the information set up to time $t-1$). With such a parameterization a problem of stationarity arise, while the conditional variances are limited the same in not true for unconditional variances who can easily diverge to infinity, therefore we modify the causality function in the following way

$$c) \quad f_1\left(I^{t-1}\right) = \gamma_1 z_{2,t-1}^2 \quad (18)$$

We insert the squared standardized residual, this can be interpreted as the "effective" variance shock or innovation. Stationarity will be proofed in a while.

In this setup the multiplicative effect, driven also by the parameter γ_1, allow for positive and negative causality in the sense that an increase in the variance of the second series imply an increase in the variance of the first series only if the function $f_i(\cdot)$ is greater than 1 (the parameter greater than zero), otherwise we will have a decrease. Non causality is associated with a zero parameter. Moreover parameters need not to be constrained given the exponential formulation. Therefore, a significativity test on the parameter γ_1 will indicate the existence or not of a causal relation between the variances of the two series, while its sign can be interpreted as the causality direction. The most interesting feature of this setup is its flexibility however a drawback must be noted, it is again a problem of parameters, too many in high dimension systems.

We verify now the stationarity of our model. We are interested in computing the unconditional variance

$$E\left(\sigma_{1,t}^2\right) = E\left[\exp\left(\gamma_1 z_{2,t-1}^2\right)\left(\omega_1 + \beta_1 \sigma_{1,t-1}^2 + \alpha_1 \varepsilon_{1,t-1}^2\right)\right] \tag{19a}$$

where we considered a simple GARCH(1,1) but in general any GARCH structure can be employed. A similar equation, not necessarily with the same GARCH structure, can be written for the second variable of interest. We verify then stationarity for one series only. First of all we point out that $z_{2,t-1}^2$ is independent both from $\sigma_{1,t-1}^2$ and $\varepsilon_{1,t-1}^2$ (and any other past values of these two quantities) since $\sigma_{1,t-1}^2$ depends on the information set up to time $t-2$ while $\varepsilon_{1,t-1}^2$ depend on $z_{1,t-1}^2$ and on the past values of z_2^2 up to time $t-2$. Given that we can write

$$E\left(\sigma_{1,t}^2\right) = E\left[\exp\left(\gamma_1 z_{2,t-1}^2\right)\right] E\left[\omega_1 + \beta_1 \sigma_{1,t-1}^2 + \alpha_1 \varepsilon_{1,t-1}^2\right] \tag{20}$$

and recognizing in the first expected value the moment generating function of the variable $z_{2,t-1}^2$ we are almost done. We need an assumption on the distribution of $z_{2,t-1}^2$, for the sake of exposition let us consider a normal standardized residual. Therefore, recalling that the squared standardized residuals are distributed as a χ^2 with one degree of freedom, and that the moment generating function of the $\chi^2(k)$ distribution is

$$mgf(t) = \left[\frac{1}{1-2t}\right]^{k/2} \qquad t < 1/2$$

we have

$$E\left(\sigma_{1,t}^2\right) = \left[\frac{1}{1-2\gamma_1}\right]^{1/2}\left(\omega_1 + \beta_1 E\left[\sigma_{1,t-1}^2\right] + \alpha_1 E\left[\sigma_{1,t-1}^2\right]\right) \qquad \gamma_1 < 1/2 \tag{21}$$

given the independence between $\sigma_{1,t-1}^2$ and $z_{1,t-1}^2$. Therefore

$$\begin{aligned} E\left(\sigma_{1,t}^2\right) &= \frac{\omega_1}{\left(1-2\gamma_1\right)^{1/2}} + \frac{\alpha_1+\beta_1}{\left(1-2\gamma_1\right)^{1/2}} E\left[\sigma_{1,t-1}^2\right] \\ E\left(\sigma_{1,t}^2\right) &= \frac{\omega_1}{\left(1-2\gamma_1\right)^{1/2} - \alpha_1 - \beta_1} \end{aligned} \tag{22}$$

and the stationarity conditions are

$$\gamma_1 < 1/2 \quad \text{and} \quad \left(1-2\gamma_1\right)^{1/2} - \alpha_1 - \beta_1 > 0 \tag{23}$$

These are similar to the one of the GARCH(1,1) however much more narrower, at least when γ_1 is positive and close to its limit. When $\gamma_1 = 0$ the model collapse on a GARCH(1,1) and we get its well know stationarity restriction. The stationarity can be easily verified also for the IGARCH(1,1) case, and result to be $\gamma_1 < 0$, therefore we loose any interpretation of the sign of the coefficient. We suggest then to avoid the IGARCH(1,1) parameterization. The stationarity restrictions for the long memory GARCH specifications cannot be so easily verified, in fact we need a revision of the proofs of the ARCH(∞) stationarity conditions, as reviewed in the first chapter, to check if they could be extended in our framework. This analysis will be the object of a future research, therefore we will apply, for the moment, only the GARCH(1,1) parameterization of our model.

The model can be simply extended to dimension higher than 2, allowing a CCC structure across all covariances and extending (17) to

$$c\_m) \qquad f_i\left(I^{t-1}\right) = \sum_{\substack{j=1 \\ j\neq i}}^{n} \gamma_{i,j} z_{j,t-1}^2$$

the causality test is now on an elevate number of parameters. Therefore we recommend the application this structure in small systems, such as the volume-volatility study of this work of the causal relation existent across financial markets such as Europe, New York and Tokyo, in the setup suggested by Pojarlev and Polasek (2000). Additional possible bivariate (and then multivariate) extensions can be obtaining modifying the causality function as follows

$$\begin{array}{lrcl} c1) & f_1\left(I^{t-1}\right) & = & \gamma_1 z_{2,t-1} \\ c2) & f_1\left(I^{t-1}\right) & = & \gamma_{1,1} z_{2,t-1} + \gamma_{1,2}\left(\left|z_{2,t-1}\right| - E\left|z_{2,t-1}\right|\right) \\ c3) & f_1\left(I^{t-1}\right) & = & \gamma_{1,1} z_{2,t-1} + \gamma_{1,2} I_{z_{2,t-1}<0} z_{2,t-1} \\ & & \vdots & \end{array}$$

where we report only some of the possible formulae, mirroring the EGARCH and the GJR-GARCH. However we stress that the stationarity conditions should be recalculated for any of these specifications. A deeper analysis of stationarity of our model under different specifications of the causality function will be the object of future researches.

Another extension regards possible alternative functions $f_i(\cdot)$, we restricted our attention to the exponential, however, we can easily introduce a logistic function. In that case, however, the stationarity conditions should be evaluated numerically.

3 An application to real data: the FIB30 market and its memory and causality structure

In 1994 the Italian Stock Exchange moved to an automatic transaction system, by the end of that year a new market segment was created, to allow trading on derivatives and among these on the future on the stock market indices. In the Italian Exchange there exist four indices: the MIBTEL built on all the traded stocks, the MIB30 that consider only the thirty firms with higher capitalization and trading, the MIBR that consider the twenty stocks with higher capitalization among the ones excluded from the MIB30 and finally the NUMTEL the Italian correspondent of the NASDAQ, which group information, technology and communications firm recently entered into the market. In 1994 a future contract on the MIB30 index was launched; the objective of this operation was the increase in transactions and in the efficiency of the stock market, both objectives were reached. Contract characteristics are the followings (valid up to the end of 2001): there are four contract maturities in March, June, September and December; the contract clear the third Friday of the maturity month at 9:30 AM or next day of open market if Friday is holiday; there are four traded maturities all over the year; last day of negotiation is the maturity day; the closing price is fixed by the Clearing House with respect to the MIB30 index at 9:30 AM (opening price); regulation of the contract by cash the first open market day after maturity; contract nominal value is determined with respect to the MIB30 index, each point of the index is worth 5 euros and the minimum price movement is of 5 index points; transactions last from 9:15 AM up to 17:30 PM (this up to the end of 2001, from January 2002 the market close at 17:40 PM). It is worth to note that the main market open at 9:30, this mean that the future prices between 9:15 and 9:30 anticipate market movements since it respond to the information released during closed market periods (weekend or just the night), therefore we have an increase in volumes and high variations in returns in this limited period. Alternatively the operators could wait the opening of the main market before trading in the future to observe how the stocks react to information shocks. In both cases volume and prices (returns) of the future may be biased and present abnormal movements, the inclusion and exclusion of these 15 minutes together with the selection of outliers and the filtration of the cyclical components will be analyzed in the following section.

3.1 Data description, cleaning and filtration

The database used in this study was provided by the Italian stock exchange (Borsa Italiana S.p.A.). It contains approximately one year of transactions concluded on the FIB30 contract, the future on the Italian stock exchange index (MIB30), see previous section for details about the market and the index. The supplied data range from 13th march 2000 to 20th march 2001 for a total of 260 open market days. As specified in the previous section, contracts are traded with four different fixed maturities (mid March, June, September and December), with the possibility to trade on the next four maturity dates. This mean that at any given time there are, possibly, four contracts traded. In this analysis we concentrate our attention on the contract with the closest maturity, that is also the most traded one. It concentrate about the 95\% of the market, apart the last week of its life, when the trading in the next to maturity contract increase. This represent a problem in dealing with such a database, we have to check if is necessary to exclude the last days of a contract life to avoid noisy prices, that may be biased by the roll-over process. We will return on this problem in a few steps. The data provided are "transaction" data, that is they record all contracts concluded, in details for each contract the database contains: the current date; an identifier of the contract and its ISIN code, that are different for each maturity and instrument traded; the trading and clearing time, that is when contract is entered in the system and when it is matched with a counterpart; the contract number, price and finally the volume, the number of futures exchanged (see figure 1 for a snapshot of the row data).

This boil down, even if the Italian stock exchange market is not one of the biggest (it represents only 2% of worldwide stock markets capitalization), in an enormous amount of data, more than 20 millions of informations, with peaks of more than 15000 contracts signed in a trading day, see table 2 for a summary of descriptive statistics. However this dataset cannot be used for all market microstructure analysis since bid/ask spread, the sign of the contract and the identifier of the subject that trades the contract are missing. Even with this limitation the amount of information contained is still considerable, and useful for our purposes. Given that we concentrate on the prices of the closest maturity contract, we have to solve another problem: we must specify how we deal with the days of the rollover, with the change of maturity and with the price change in the traded contracts across maturity dates. It is well known that a future price get closer to the price of the underlying as it get closer to the maturity, while the price of the next-to-maturity contract include also expectations about the underlying and the interest rates movements. Different solutions for obtaining return series from future prices deal with this problem and suggest alternative strategies, from shifting the prices series of the delta between the two contract at the maturity date, or just adjusting return with a factor. We will not follow any of these suggestions but we will statistically test the relevance of the return in the maturity date: we will add an impulse dummy variable for each maturity date (four in our sample) and test its significativity. This additional variable will, eventually, exclude the return across the maturity date without introducing a bias on the structure of the underlying process.

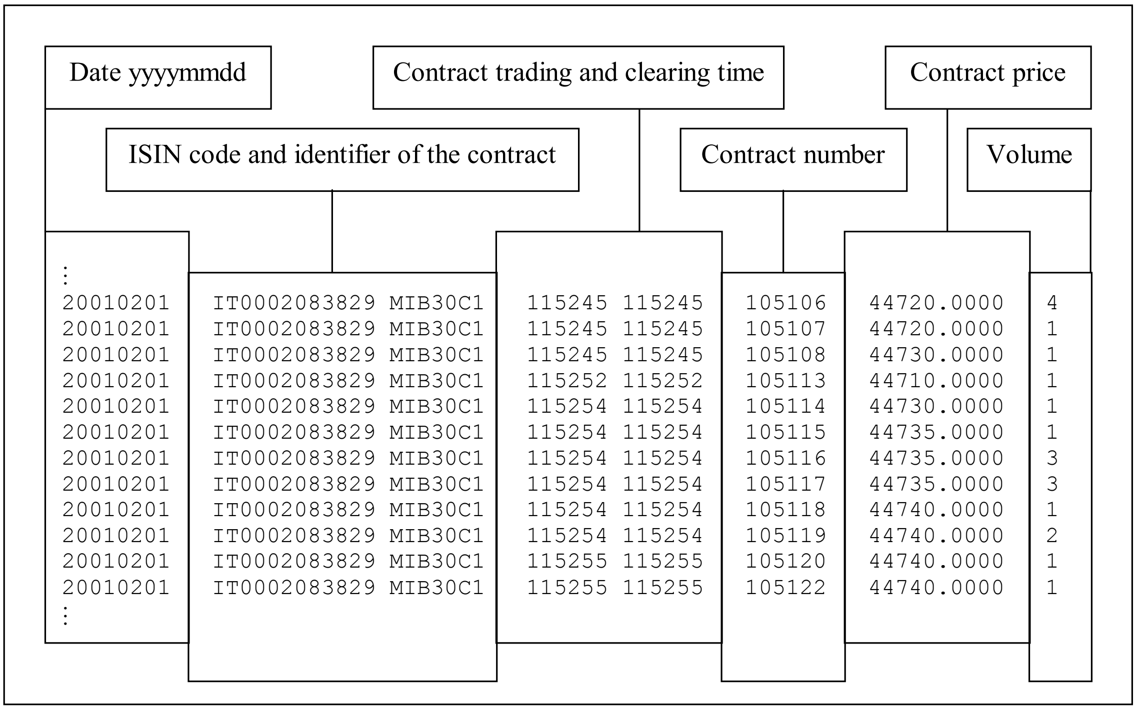

Figure 1: extract from the dataset provided by the Italian Exchange

In the database there has been recorded transaction series, for our analysis we need to transform them into equally spaced non-overlapping series. We decided to run our analysis on the 5 minute interval, a smaller tick frequency may not show the causality relation, while a longer one may destroy the microstructure and the correlation and causality between volume, returns and their volatility, resulting in a contemporaneous relation. Converting transaction data into five-minute observations is not a problem for the volume series; we have just to sum over the contract traded every 5 minutes. For the prices this is not so simple, we could in principle take the last price of all 5 minute intervals, take a weighted average over the 5 minutes, take the median, the simple mean and other combinations or means. However, our main problem is to avoid biases due to the averaging therefore, we chose to consider the last price of the interval, since it will include all the relevant information impacts, then we price all the volume traded in the interval at this last price.

Day	Motivation of exclusion
26 April 2000	Market was open up to 18 PM due to technical problems during the day
5 July 2000	Market was open up to 19 PM due to technical problems during the day
14 August 2000	Abnormal movements due to very low trading
28 August 2000	Market opened at 10 due to technical problems
19 February 2001	Market opened at 12:30 due to technical problems

Table 1: days removed from the sample and motivation

Another point concern the first 15 minutes of open market, from 9:15 to 9:30: we decided to compare the results obtained including and excluding the contract concluded within these 15 minutes to verify if they could bias both volume and price movements. In fact, these 15 minutes concentrate abnormal volume and price movements due to the impact of news diffused during closed market hours. In this view we introduce another dummy variable that exclude the first return of each day in order to test the importance of price movements from 9:15 to 9:30. A significative value in this dummy is expected since most of announcements of the central banks and of bigger firms are released when markets are closed. These announcements affect the underlying stock market index and the term structure of interest rates, resulting therefore in movements also in the future prices. A non-significative dummy will reveal that the news released do not heavily affect market prices probably because they are discounted well in advance by the whole market. Moreover we can verify the effect of transactions occurred between 9:15 and 9:30 observing how the coefficient of this dummy vary including and excluding these data. Some days in the sample are also completely deleted because the market showed anomalous trading or anomalous opening due to technical problems (for details see table 1).

Particular attention has been given to outlier detection and exclusion; we presume that they are mainly due to operational errors. Part of them are identified observing clearing time, abnormal high price sell contracts and low price buy contracts are generally not executed, this mean a clearing time of zero, but the problem is with abnormal low price sell contracts and high price buy contracts that may be automatically matched with counterparts once entered into the system. In this last case the outliers are detected by a procedure that test the presence of an operation of opposite sign in the next few seconds, and in that case delete both. However, the correction may not be the next operation, also there cannot be a correction at all. There remains therefore other outliers. Normally these are identified and deleted by a procedure that is run on a daily basis: at first the (daily) standard deviation of the log-returns is computed, then outliers are defined as the returns outside two bounds determined as three times the standard deviations (positive and negative), and are then removed; the procedure should also check for jumps in the series, that could be detected as outliers. We follow however this alternative procedure: assume that we have a presumed outlier p_t, then we take the 10 precedent contract prices and compute the mean $\mu_{t-1,t-10} = (10)^{-1} \sum_{i=1}^{i=10} p_{t-i}$; we remove this mean from the

10 contracts and the suspect price we are analyzing and consider the absolute value of these differences $d_{t-i} = \left| p_{t-i} - \mu_{t-1,t-10} \right| \; i = 0, 1...10$; we compare the presumed outlier difference d_t with the biggest difference among the 10 precedent prices $d_{10} = \max\left(d_{t-i} \quad i = 1...10\right)$, the contract is deleted if $d_t > 3d_{10}$ (this to allow shock effects during open market periods). The procedure control also for jumps in the price series that, following the previous methodology, may be detected as outliers in the jumping contract price. This procedure is based on the prices preceding any possible outlier, however to allow for an easier identification of jumps the same reasoning can be based on the 10 successive prices. We compared these two alternative methodologies in a limited sample and we get absolutely no difference on the identified outliers, therefore, given the gains that could be attained in the software implementation (these procedures are really time consuming in our database consisting of millions of points),we preferred basing our algorithm on the 10 successive prices.

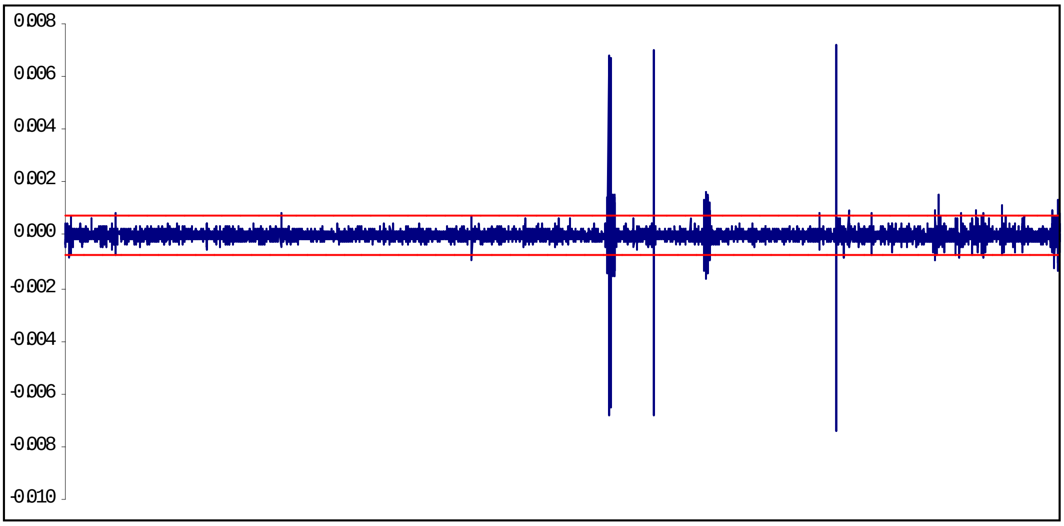

Figure 2: Log-returns and daily bounds for outlier detection

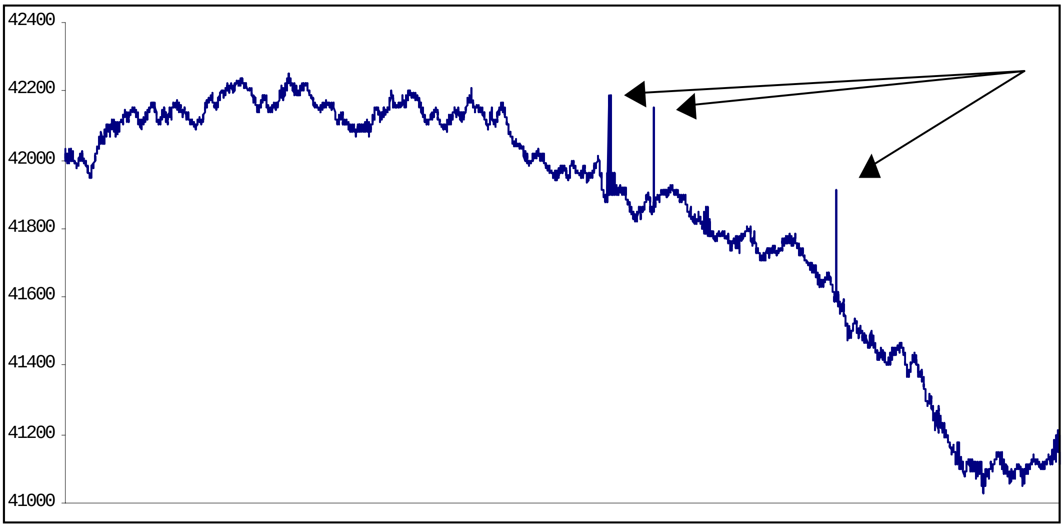

Figure 3: Prices with the outlier evidenced by our procedure

The choice of this approach was motivated by our interest in preserving extreme events, an example can be seen in Figure 2 where we report the observed log-returns for one day in the sample, the 20 February 2001. In the graph we report also the bounds for three times the standard deviations. In Figure 3 we report the prices with the outliers evidenced. Using our procedure only four outliers (the first on the left is a group of two outliers) are detected and deleted, while using the "standard" methodology much more point should have been deleted, incurring in the risk of eliminating extreme events. The difference among the two procedure is, however, really minimal, in the sense that we identified outliers in transaction data, that in the following step will be aggregated and averaged over a 5 minute interval, reducing therefore the possible effects of maintaining in the dataset an outlier or of excluding an extreme event.

Tables from 2 to 4 show a group of descriptive analysis that compares returns and volume across the quarters. For the returns we concentrate directly on the 5-minute series, derived from prices observed at the end of each of the 5-minute intervals. Table 3 compares a set of moments and extreme values between the return series that include or exclude the preopening period. As we can see the main changes appear on the sample skewness and kurtosis, this because we are introducing returns placed in the queues of the distribution: deleting the first 15 minutes of open market the first return of each day will be computed with respect to the last price of the previous day, resulting most of the times in a bigger value (in absolute terms) since the news released when markets were already closed will affect market specially between 9:15 and 9:30.

Quarters	17/03-15/06/2000			16/06-14/09/2000		
Contract	15/06	15/06 adj.	14/09	14/09	14/09 adj.	14/12
Number of days	62	60	62	64	61	64
Mean volume	17362	16171	858	11321	10554	799
Volume s.d.	3781	3594	2117	3161	2759	2548
Minimum	6586	5662	15	2624	6354	10
Maximum	27323	26016	12190	24801	21956	17631
Volume	1059058	970249	52335	724551	643765	51133
Quarters	15/09-14/12/2000			15/12/00-15/03/01		
Contract	14/12	14/12 adj.	15/03	15/03	15/03 adj.	14/06
Number of days	65	65	65	62	61	62
Mean volume	15728	14664	667	15902	14900	860
Volume s.d.	3540	3433	2070	5706	5310	2349
Minimum	6475	5777	5	4451	4088	8
Maximum	24514	23362	11196	37168	34698	13526
Volume	1022333	953163	43380	985907	908921	52455

Table 2: descriptive statistics of the daily volume across the quarters comparing the two most traded contracts and the effects of data cleaning (contracts recorder between 9:15 and 9:30 are excluded)

Quarters	17/03-15/06/2000		16/06-14/09/2000	
9:15-9:30	with	without	with	without
Mean	-4,62E-06	-4,45E-06	9,20E-07	9,49E-07
Standard dev.	0,00168	0,00173	0,00084	0,00086
Kurtosis	40,11631	50,92920	12,97777	15,05628
Skewness	-1,95357	-2,14094	0,58585	0,78776
Minimum	-0,03441	-0,03713	-0,00621	-0,00621
Maximum	0,01215	0,01828	0,01263	0,01263
Quarters	15/09-14/12/2000		15/12/00-15/03/01	
9:15-9:30	with	without	with	without
Mean	-1,13E-05	-1,16E-05	-2,61E-05	-2,69E-05
Standard dev.	0,00118	0,00119	0,00134	0,00135
Kurtosis	12,48523	14,27727	24,84416	24,67570
Skewness	0,01205	-0,23917	0,30956	-0,07730
Minimum	-0,01465	-0,01532	-0,01387	-0,01733
Maximum	0,01126	0,01256	0,02349	0,02095

Table 3: descriptive statistics of the 5 minute returns series across quarters comparing the effects of excluding or not the contracts concluded between 9:15 and 9:30

Quarters	17/03-15/06/2000		16/06-14/09/2000	
9:15-9:30	with	without	with	without
Mean	174,222	168,446	114,091	109,933
Standard dev.	132,624	125,727	104,451	99,130
Kurtosis	3,346	2,631	12,546	11,587
Skewness	1,528	1,411	2,531	2,427
Minimum	1	1	0	0
Maximum	1073	948	1324	1324
Sum	1034879	970249	688997	643765
Quarters	15/09-14/12/2000		15/12/00-15/03/01	
9:15-9:30	with	without	with	without
Mean	157,460	152,750	160,315	155,212
Standard dev.	130,888	126,293	141,693	135,341
Kurtosis	5,204	4,243	13,808	14,023
Skewness	1,776	1,702	2,386	2,306
Minimum	1	1	1	1
Maximum	1431	1147	2131	2131
Sum	1013256	953163	968141	908921

Table 4: descriptive statistics of the 5 minute volume series across the quarters comparing the effects of excluding or not the contracts concluded between 9:15 and 9:30

This is a typical behavior that creates abnormal volumes and trading during this limited period at the beginning of each day and may bias the underlying price and volume dynamic. In tables 2 and 4 we concentrated our attention on volume. Our first analysis is done in order to motivate the choice of the contract with closest maturity as the reference contract; table 2 explains this. We can see that the number of contracts traded in the future with next to closest maturity is around 5-8% of the most traded future. Moreover this behavior is stable during all the year considered. Table 2 allow to make a set of considerations on the volume traded in this derivative market, more that 3,8 millions contract were signed, with peaks of more than 37 thousand per day. The period with lower activity is the one that range from 16/06/00 to 14/09/00, not a surprising behavior. This pattern become much more clearly in figure 4, where the traded volume of the two contracts of table 2 is shown. In this same table we compare also the effect of the exclusion of the first 15 minutes and of the days listed in table 1. Table 4 is the volume descriptive analysis in the five minute intervals, showing the effects of removing preopening: as expected skewness and kurtosis goes in the direction of normality since we are deleting values that are in the queues of the distributions.

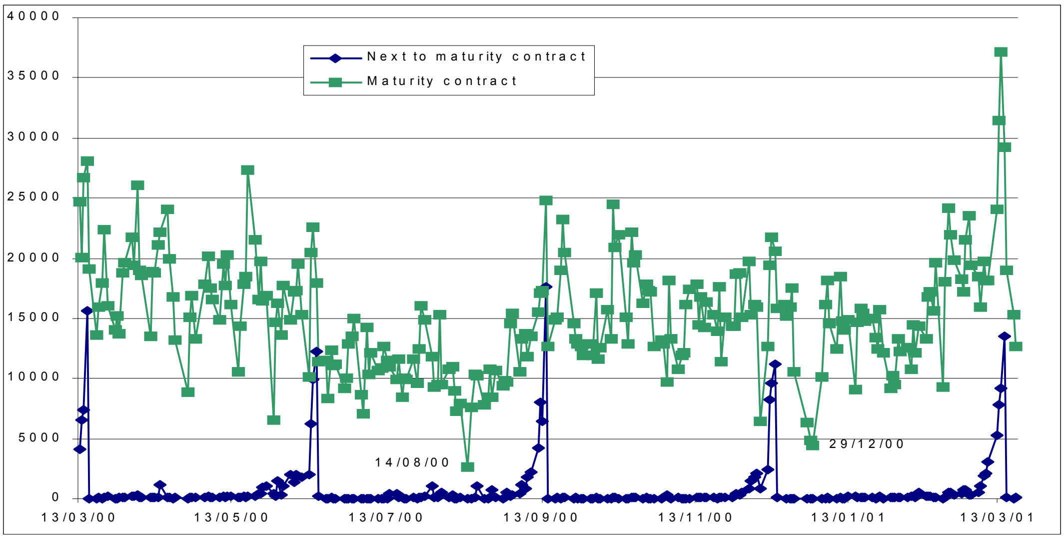

Figure 4: volume in the two most traded contracts

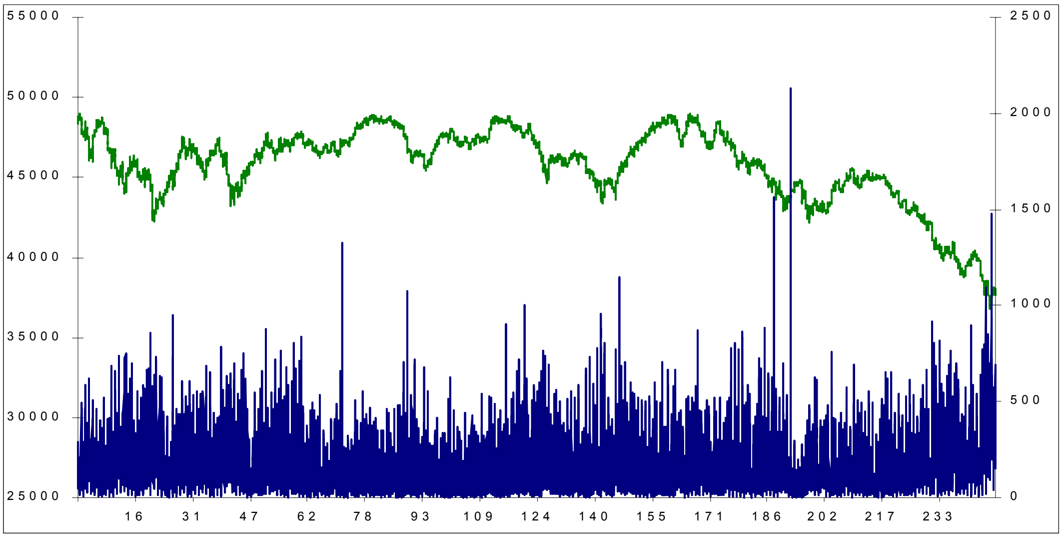

Figure 5: 5 minute prices and volumes

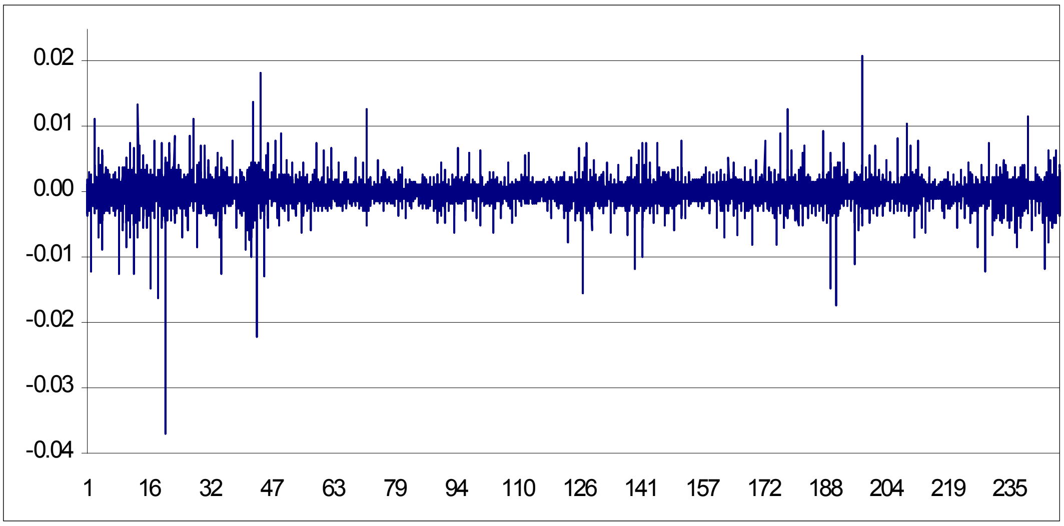

Figure 6: 5 minute returns

Given this introductory description a much more detailed analysis is required. We are interested in testing the presence of linear dynamics in the mean and in the volatility of both returns and volume, as well as to verify following the approach of Bollerslev and Andersen (1997) the presence of cyclical components. In order to perform these analysis we will make also use of autocorrelation functions computed on the series and on their absolute value, of test for correlation (Portmanteau) again computed on the series and its absolute value, and finally we will test for ARCH effects using the Engle lagrange multiplier test. We will study in detail at first the return series (contracts recorder from 9:30 up to 17:30) and in the following the volume.

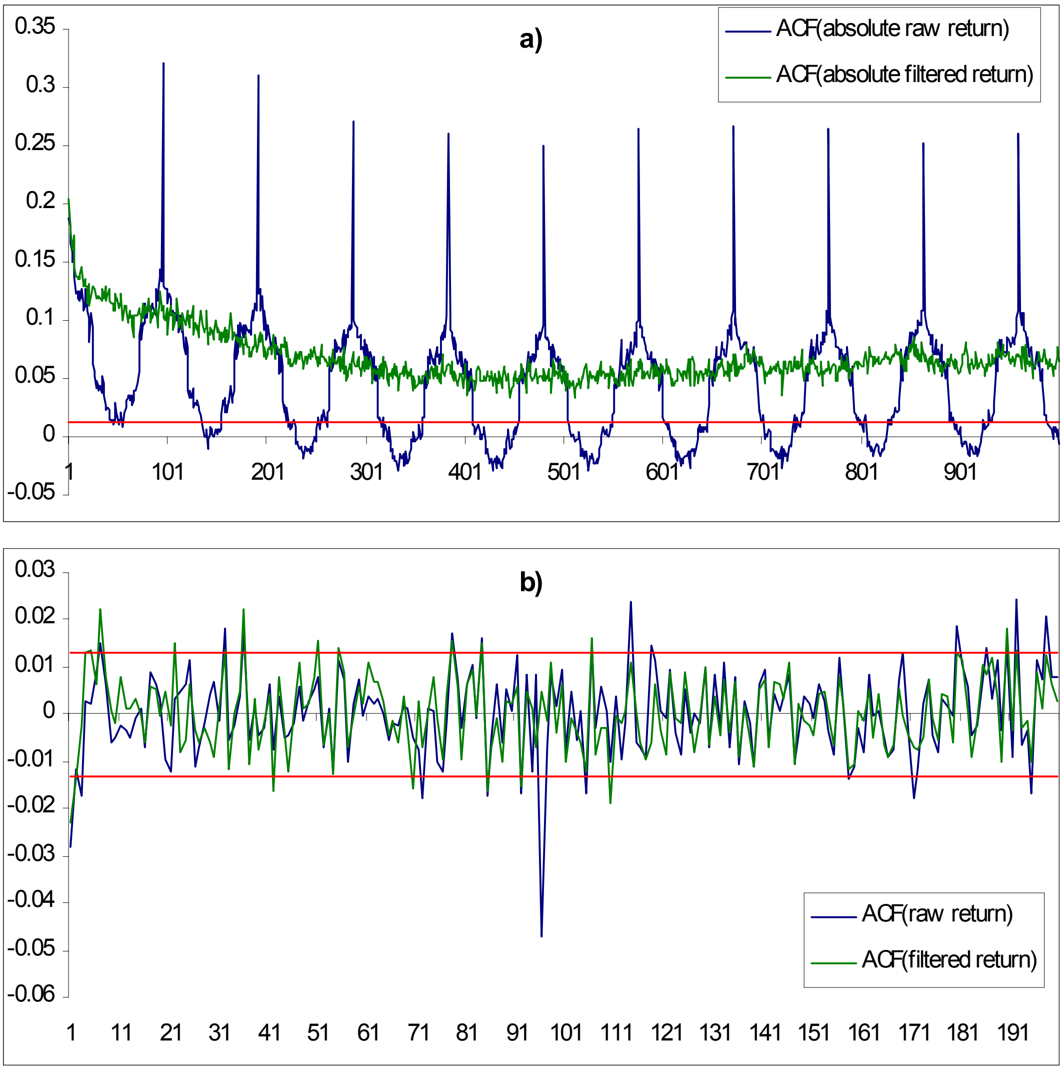

Figure 7: autocorrelation function of the 5 minute returns

Figure 7 shows the sample autocorrelations computed on 5-minute returns (200 lags considered) and on their absolute value (1000 lags considered). As expected the mean dynamic can be explained by a small order ARMA model, given the limited number of significative sample autocorrelations. The second panel clearly evidence a cyclical pattern, that is responsible also for the peak in panel a) at the 96th lag. The structure of the oscillation indicates a combination of different cyclical components acting at different frequencies (note the peaks at the top of the sinusoidal pattern). The very same observations can be derived for the series that include the contracts concluded between 9:15 and 9:30, the only difference is that the peaks are at the 99th lag and multiple. Following Andersen and Bollerslev (1997) in figure 8 and 9 we verify the presence of this periodical behavior. Note that the first 5-minute interval is not represented in both graphs given its very high value; plotting it the cyclical patter will not have been so evident.

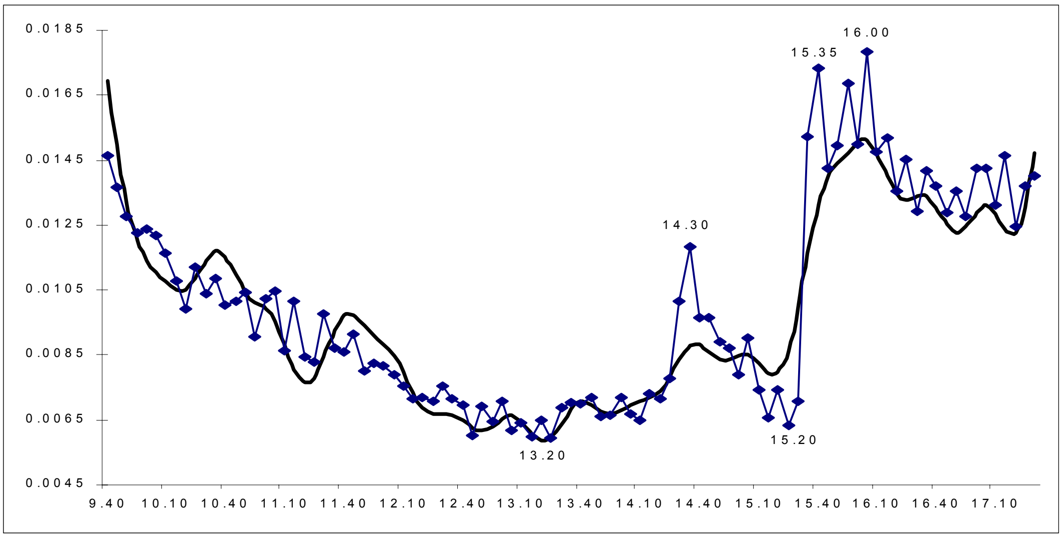

Figure 8: cyclical pattern in the absolute returns 9:35 - 17:30

The graphs represents the average absolute return across the five minute intervals in the sample: the open market day last for 8 hours (in figure 8, while in figure 9, 8 hours and 15 minutes), so 96 (99, this is the N index) ordered five minute intervals are considered, for each of these intervals we computed the mean across the days in the sample (247=T), and graphed them:

$$\hat{\mu}_n = \frac{1}{T}\sum_{t=1}^{T} x_{t,n} \qquad n = 1, ...N \tag{24}$$

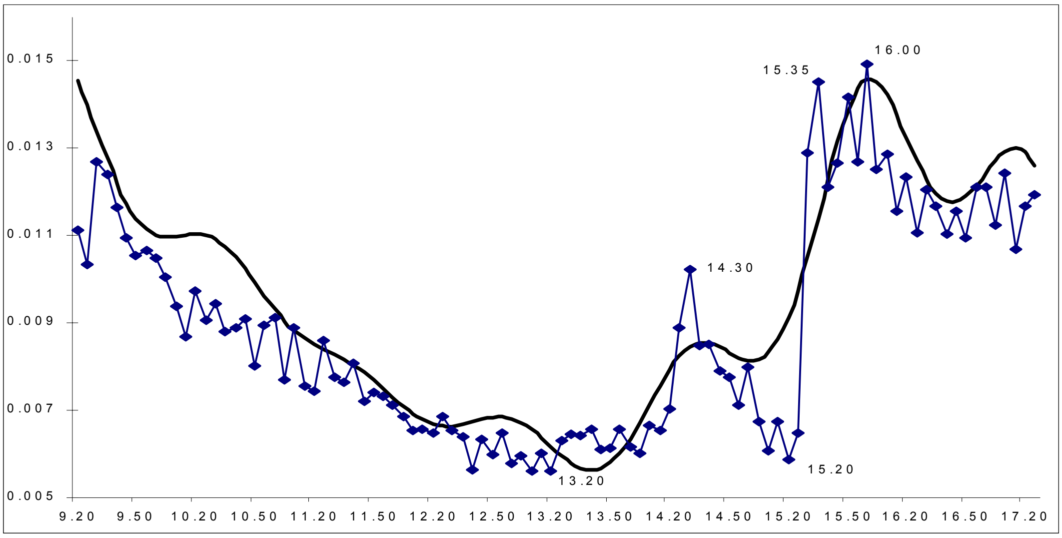

Figure 9: cyclical pattern in the absolute returns 9:20 - 17:30

The graphs clearly show a cyclical pattern, the well known U effect, returns are high when market opens and closes, they decrease during the day up to 2-3 PM and then increase. This effect is well documented for stock market returns and volatility. Two additional effects are noted: there is a peak after lunch, due to an increase in the trading after the break, that last for about an hour, than the trading decrease as quickly as it increased, until the first signals come from the American markets, in fact we can observe (note the timing in the graphs) that returns and volume react half an hour before the opening of the New York stock exchange. Both graphs 8 and 9 show the same pattern computed with (24), the difference is in the estimated pattern, the bold line. The seasonal component is filtered by the Flexible Fourier Form regression, see the appendix for a brief description of this technique. Flexible Fourier form is a deterministic filtration procedure, we do not consider any stochastic cyclical specification since the behavior of this periodic patter is very evident from data, moreover in a recent paper Beltratti and Morana (2000) compared the two methodologies and their choice was in the direction of the FFF, given the very close results. As in the cited papers the returns filtration is based on the following representation

$$r_{t,n} = E\left[r_{t,n}\right] + \frac{\sigma_{t(,n)} s_{t,n} Z_{t,n}}{\sqrt{N}} \tag{25}$$

where the return $r_{t,n}$ at day t and interval (5 minute) n is the sum of its expected value plus a term that depend on a daily (or intra-daily) volatility $\sigma_{t(,n)}$, on an error term $Z_{t,n}$ and on a function $s_{t,n}$ that explain the cyclical behavior. After a transformation on the data a Fourier regression is run, considering the harmonics reported in tables 5 and 6.

Parameter	Component correspondence	Time correspondence
α_1, α_2, α_3	Quadratic component	---
γ_1	Dummy for opening	---
γ_2, γ_3, γ_4	Dummies for maturity	---
δ_1, ϕ_1	Harmonic of period 1	8 hours (1 day - 96 5 minute intervals)
δ_2, ϕ_2	Harmonic of period 2	4 hours (48 intervals)
δ_3, ϕ_3	Harmonic of period 3	2 hours and 40 minutes (32 intervals)
δ_4, ϕ_4	Harmonic of period 4	2 hours (24 intervals)
δ_6, ϕ_6	Harmonic of period 6	1 hour and 20 minutes (16 intervals)
δ_8, ϕ_8	Harmonic of period 8	1 hour (12 intervals)
δ_{12}, ϕ_{12}	Harmonic of period 12	45 minutes (9 intervals)
δ_{16}, ϕ_{16}	Harmonic of period 16	30 minutes (6 intervals)

Table 5: harmonics and correspondence with time, 96 interval per day, 9:30-17:30

Parameter	Component correspondence	Time correspondence
α_1, α_2, α_3	Quadratic component	---
γ_1	Dummy for opening	---
γ_2, γ_3, γ_4	Dummies for maturity	---
δ_1, ϕ_1	Harmonic of period 1	1 day – 99 5 minute intervals
δ_2, ϕ_2	Harmonic of period 2	½ day – 49.5 5 minute intervals – 4h 7' 30''
δ_5, ϕ_5	Harmonic of period 5	19.2 5 minute intervals – 1h 36'
δ_6, ϕ_6	Harmonic of period 6	16 5 minute intervals – 1h 20'

Table 6: harmonics and correspondence with time, 99 interval per day, 9:15-17:30

The filtered returns are obtained by the following expression:

$$r_{t,n}^{f} = \frac{r_{t,n} - E\left[r_{t,n}\right]}{s_{t,n}} \tag{26}$$

Given the flexible representation of the FFF we added also a group of dummy variables to take into account and test the effect due to the first 15 minutes of open market and to the maturity dates: we added 3 dummies to consider the three maturity dates included in the sample (impulse dummy) and an additional one to consider the daily opening effect, in this last case the dummy get a value 1 for the first return of each day, a significant value of its associated coefficient imply that information flow during closed market period impact on prices on the first return of the day. A different value of the associated coefficient and significance may be used to infer on the relevance of the transactions realized between 9:15-9:30. The estimation results are summarized in table 7 and 8, were the parameters, standard deviation and significativity tests are considered. A graph of the estimated cyclical component is in figure 8 and 9, while a graph of the autocorrelation function of filtered data in figure 7 (again including or excluding the period 9:15-9:30 does not affect autocorrelation analysis). The harmonics listed in tables 5 and 6 have been chosen with a general to particular selection procedure, starting form a FFF regression with harmonics up to a period of 24 and iteratively deleting the less significant harmonic.

Parameter	Estimate	Std. Err.	t-value	Parameter	Estimate	Std. Err.	t-value
α_1	-7,501	0,938	-7,994	δ_3	-0,230	0,061	-3,774
α_2	-10,531	2,712	-3,884	ϕ_3	-0,086	0,034	-2,526
α_3	3,478	0,903	3,851	δ_4	-0,111	0,039	-2,871
γ_1	1,430	2,550	0,561	ϕ_4	0,101	0,030	3,417
γ_2	2,267	2,550	0,889	δ_6	0,069	0,027	2,575
γ_3	2,415	2,550	0,947	ϕ_6	-0,128	0,026	-4,911
γ_4	2,462	0,234	10,533	δ_8	-0,031	0,025	-1,282
δ_1	-1,309	0,530	-2,470	ϕ_8	0,110	0,025	4,455
ϕ_1	-0,405	0,079	-5,134	δ_{12}	-0,030	0,024	-1,241
δ_2	-0,569	0,132	-4,309	ϕ_{12}	0,046	0,024	1,922
ϕ_2	-0,171	0,044	-3,882	δ_{16}	0,014	0,024	0,576
				ϕ_{16}	0,067	0,024	2,826

Table 7: estimation results for the return cycle, period considered 9:30-17:30

Parameter	Estimate	Std. Err.	t-value	Parameter	Estimate	Std. Err.	t-value
α_1	-11,46041	0,33439	-34,27310	δ_1	1,03320	0,18654	5,53872
α_2	1,20278	0,95202	1,26339	ϕ_1	-0,25252	0,05067	-4,98332
α_3	-0,47152	0,31650	-1,48978	δ_2	0,00101	0,05080	0,01984
γ_1	0,37847	2,56010	0,14784	ϕ_2	-0,13173	0,03216	-4,09604
γ_2	2,81346	2,56010	1,09897	δ_5	0,09982	0,02421	4,12222
γ_3	2,13568	2,56010	0,83422	ϕ_5	0,09842	0,02462	3,99668
γ_4	2,80402	0,19666	14,25808	δ_6	0,12195	0,02376	5,13377
				ϕ_6	-0,08180	0,02413	-3,39027

Table 8: estimation results for the return cycle, period considered 9:15-17:30

Statistics	9:30-17:30		9:15-17:30	
	Raw	Filtered	Raw	Filtered
Mean	-1,056e-5	-9,661e-4	-1,032e-5	-0,00081
St. Dev.	0,00131	0,11353	0,00129	0,11756
Kurtosis	-1,223	-0,165	-0,938	-0,163
Skewness	50,772	6,831	41,303	6,981
Q(5)	29,342	25,178	10,372	24,325
Q(10)	38,923	39,770	18,523	39,721
Q(20)	46,335	44,511	25,837	45,740
Q(50)	84,085	98,332	63,290	103,646
Q(100)	219,632	176,500	152,049	178,770
$Q^2(5)$	57,562	2553,908	82,956	2624,742
Q^2 (10)	92,361	3848,511	154,235	3926,159
Q^2 (20)	152,718	5887,090	237,537	5872,836
Q^2 (50)	190,973	10751,499	302,345	10837,267
Q^2 (100)	1441,911	17024,438	1255,376	17018,321
JB	2266576,7	14713,698	1501913	16367
LM(2)	22,291	17,141	3,769	17,246
LM(5)	30,194	25,759	10,538	24,731
LM(10)	40,042	42,102	18,696	41,764
LM(20)	47,754	46,477	26,272	47,294

Table 9: moments and test of 5 minute returns

Finally in Table 7 we compare the moments and the tests for correlation and ARCH effects computed on the raw and on the filtered return series. After removing the periodic component what become evident is a long memory effect in the volatility of the 5 minute returns, see the smooth convergence toward zero of the autocorrelation of absolute returns, they are significant for all the lags graphed in Figure 7. The Portmanteau test compute on squared observations react to this emerging pattern while Engle LM test is not influenced. The filtration of periodic components shift also the distribution of the returns toward normality, see Figure 10, a kernel density estimate of 5 minute returns distribution before and after filtration, compared with a normal distribution. The graph plots standardized returns (filtered and not) to avoid scale problems. The results excluding the contracts signed between 9:15 and 9:30 are very similar.

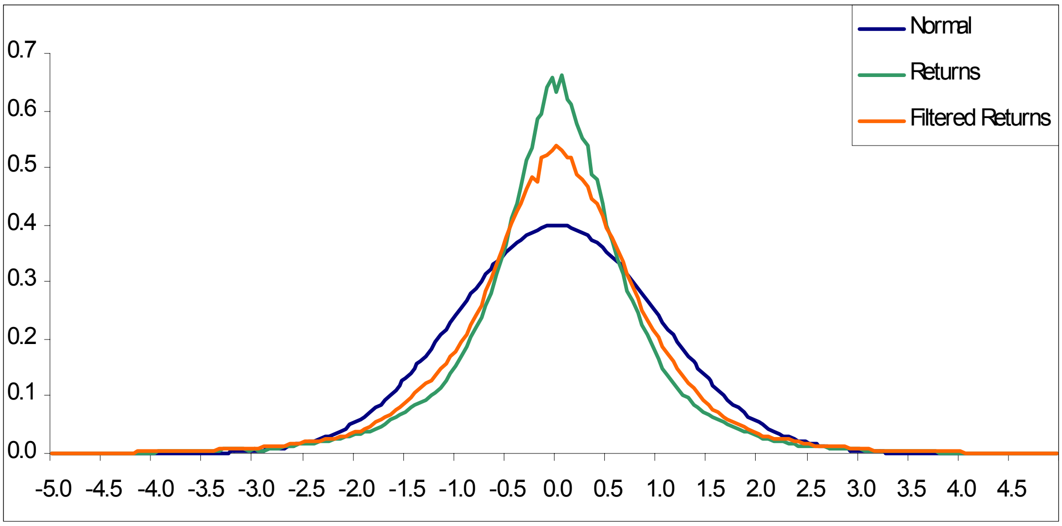

Figure 10: Kernel density estimate of 5 minute returns, period 9:15-17:30

Before turning to the analysis of the volume series a couple of comments on the estimated parameters of Table 6: the dummies reflecting the maturity date effect are not significant, probably they will have a greater impact in the mean than in the variance; as expected the dummy introduce to take into account the correction for contracts concluded between 9:15 and 9:30 is highly significative, therefore news impact affect market when it opens; all the periodic components included in the FFF regression resulted to be significant, and almost completely removing the periodic pattern from the data, this indicate that the choice of a deterministic procedure were correct.

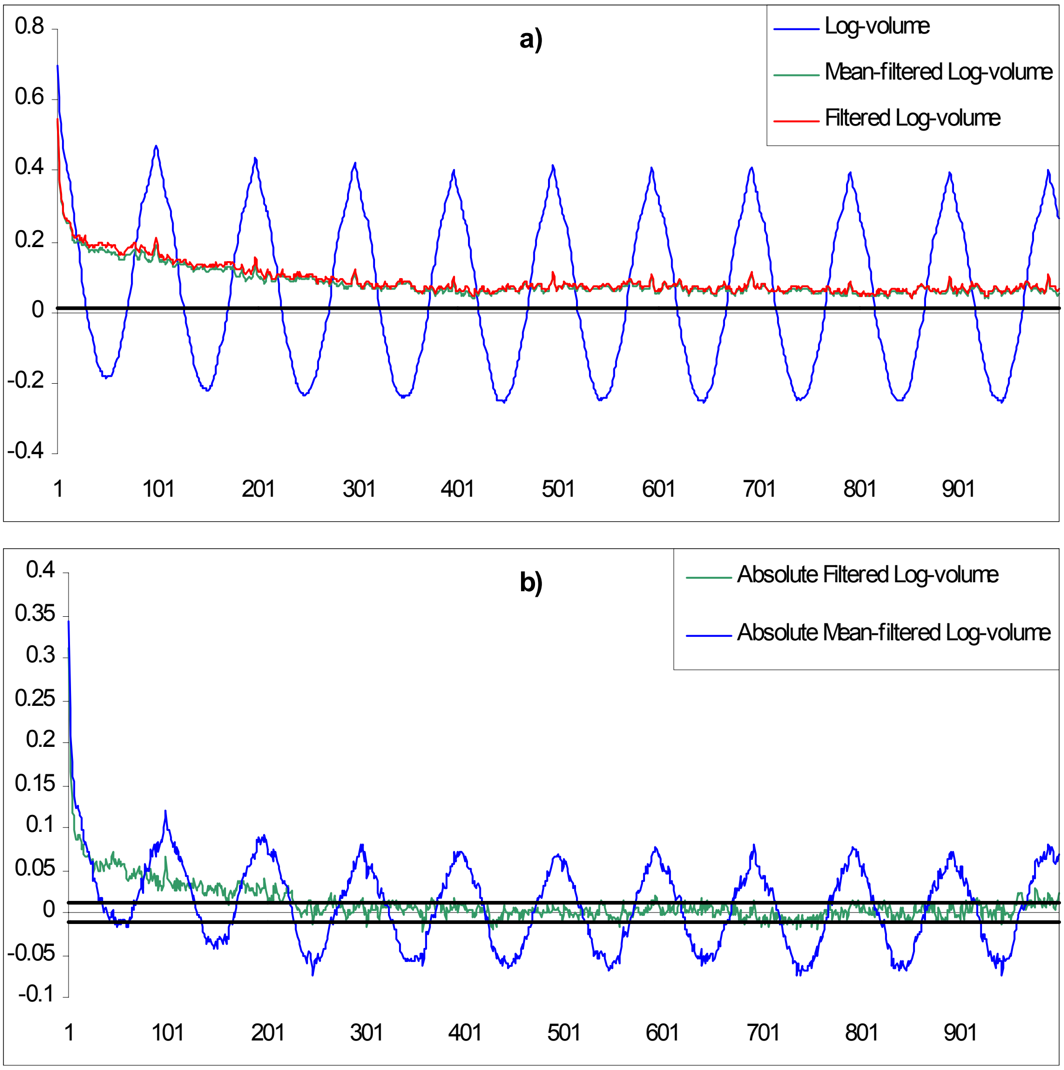

Figure 11: Autocorrelations of Volume

Let us now consider the volume series. Before all analysis the volume is rescaled with a log transformation, this operation will be necessary in a bivariate modelling view, to avoid problem in parameter estimation, moreover this will change the distribution of volumes from a lognormal (volume are positive by construction) to a normal setting. As for the returns we start considering autocorrelations depicted in Figure 11. Clearly the autocorrelations computed on volume and on its absolute value are identical.

The seasonal pattern is found here in the mean, as Figure 11 (panel a) shows, but also in the variance: this second pattern emerge once the cyclical component in the mean has been removed, as we can see from Figure 11 (panel b). Therefore, for the volume the flexible Fourier form is used both on the mean and on the variance, the representation used is the following

$$\begin{aligned} v_{t,n} &= E\left[v_{t,n}\right] + \frac{\sigma_{t(,n)} s_{t,n} Z_{t,n}}{\sqrt{N}} \\ E\left[v_{t,n}\right] &= \bar{s}_{t,n} + E\left[\tilde{v}_{t,n}\right] \end{aligned} \tag{27}$$

with two distinct seasonal components. The Harmonics used both in the mean and in the variance are the same used for the returns (see table 3) as well as the dummies. We tried to estimate a complete model in one step, but the elevate number of parameters gave problems in the convergence of the optimization algorithm, therefore we adopted a two step procedure: in the first stage we estimate and remove the cyclical component in the mean, while in a second step we deal with the seasonal pattern in the variance. The opposite filtration scheme, at first the seasonal pattern in the variance and then the one in the mean in not efficient. We considered this alternative approach and applied the FFF-technique on the variances, it turned out that the cyclical component were completely filtered and moreover the one on the mean resulted much more noisy. We presume that this is due both to the structure of the FFF technique and to the volume series values, all positive. Figures 12 and 13 show the seasonal patterns as they are found in the data and estimated by the FFF regression.

The periodic pattern in the mean is really similar to the one found in the variance of the returns, this clearly appear in Figure 14, where we superimpose the two cyclical components. We can find both the peak in the first afternoon and the abrupt increase in volumes in coincidence with New York opening. Interestingly the patter of volume volatility is of a very different shape, it peaks at noon and shows the widest variations in correspondence to New York opening and after 17 PM. This may not be so surprising: market activity is high early in the morning and near the end of the day, with an elevate number of contracts traded with high variation in prices, conversely at noon market activity is very low, prices do not move since trader wait for signals from the American market, therefore the number of contracts traded (the volume) per 5 minute intervals vary much more than in the morning or in the late afternoon. This particular behavior can explain the deterministic patterns of returns and volume.

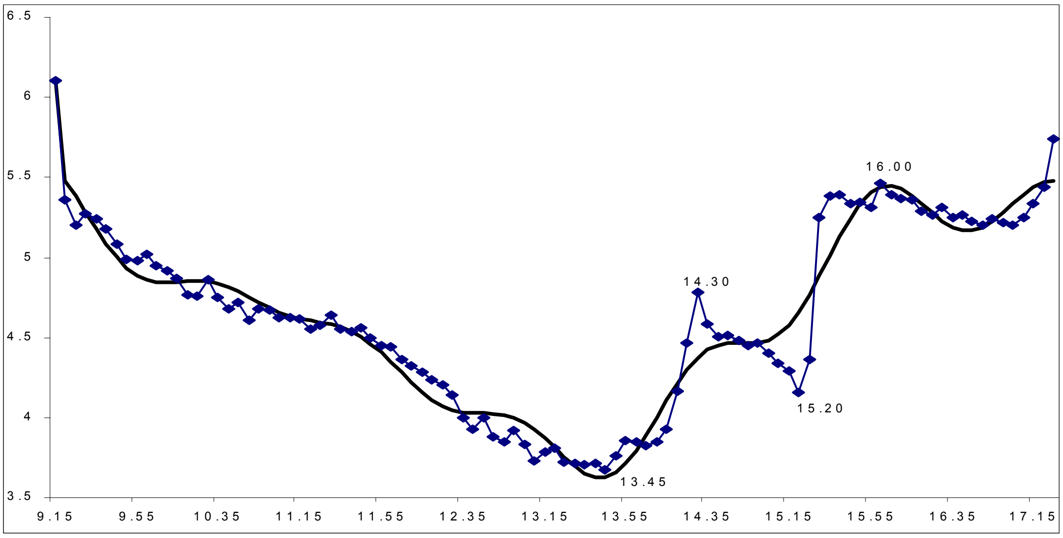

Figure 12: cyclical patter in the mean of the volume 9:20-17:30

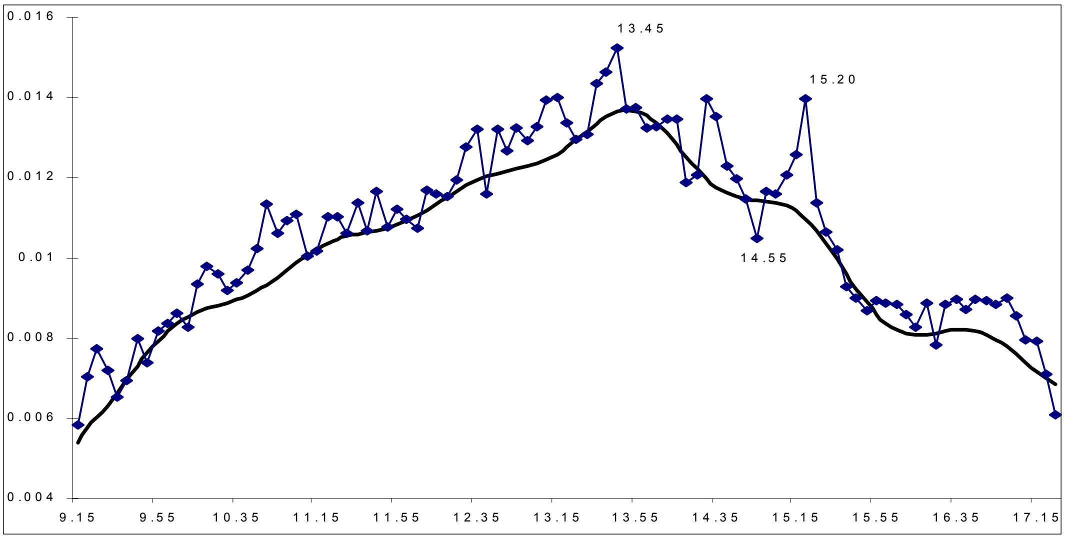

Figure 13: cyclical pattern in the variance of the volume 9:20-17:30

The estimated parameters of the two FFF regressions are reported in Table 8. As for the returns most of the harmonics used in the regression resulted to be significant, both in the mean and in the variance. The dummy for beginning of the day is significant in both regressions while the impulse dummies for the maturity dates are significant only for the variance, as if they affect only volatility and not the level of the traded contracts.

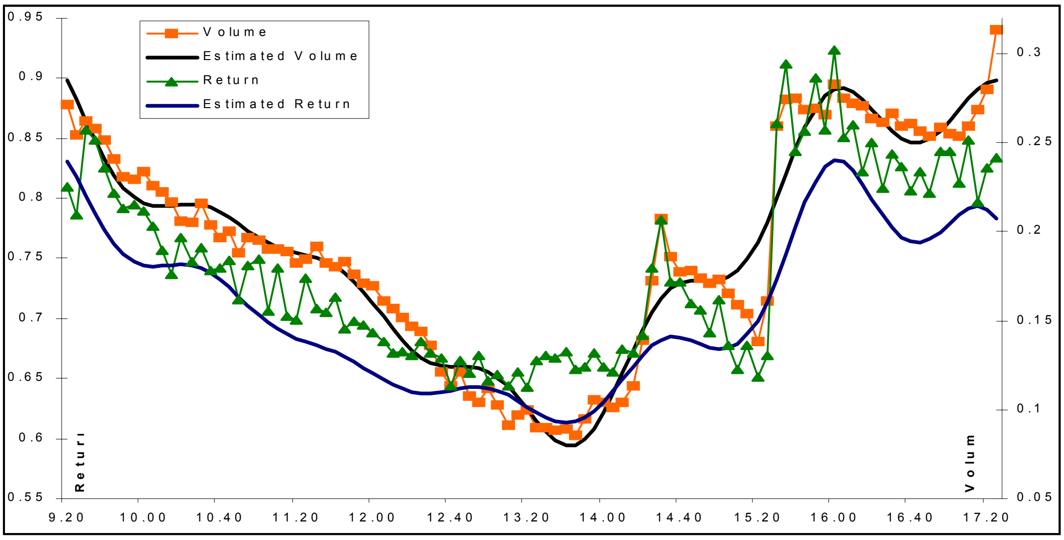

Figure 14: comparison of the cyclical patterns of returns volatility and volume mean

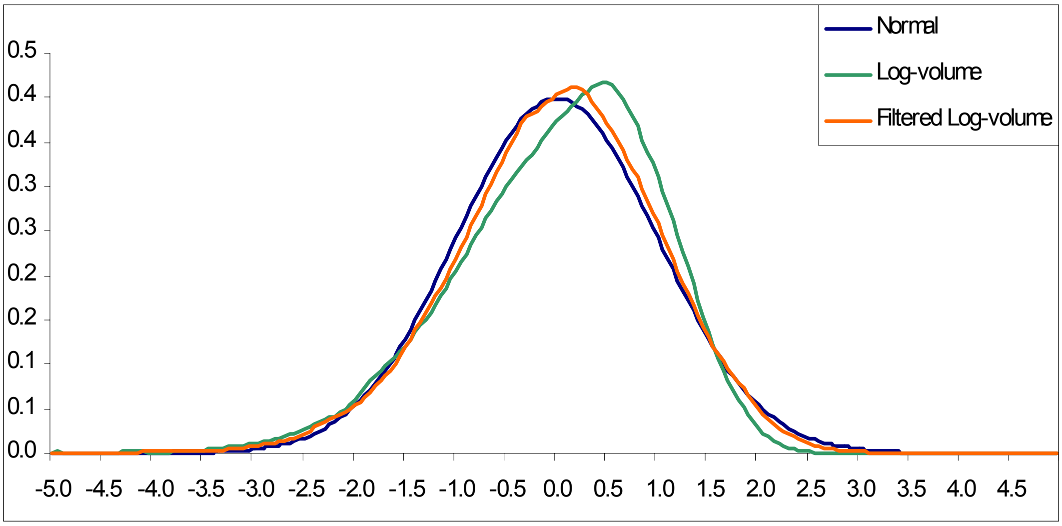

Figure 15: kernel density estimate of the log-volume

Similarly to the Returns, filtered volume is obtained by:

$$v_{t,n}^{f} = \frac{v_{t,n} - \bar{s}_{t,n} - E\left[\tilde{v}_{t,n}\right]}{s_{t,n}} \tag{28}$$

Consider now Table 12, where we report moments and tests for correlation and ARCH effects computed on the raw volume and after removing each of the two periodic component.

	Log-Volume - mean			Log-Volume - variance		
Paameter	Estimate	Std. Err.	t-value	Estimate	Std. Err.	t-value
α_1	7,807	0,281	27,738	-10,885	0,819	-13,290
α_2	-9,474	0,813	-11,647	14,721	2,367	6,219
α_3	3,193	0,271	11,786	-4,972	0,788	-6,306
γ_1	0,500	0,765	0,653	-53,645	2,226	-24,096
γ_2	-0,305	0,765	-0,399	-53,645	2,226	-24,096
γ_3	0,100	0,765	0,131	-53,645	2,226	-24,096
γ_4	-0,351	0,070	-5,007	0,561	0,204	2,751
δ_1	-1,172	0,159	-7,377	2,346	0,462	5,073
ϕ_1	-0,222	0,024	-9,404	0,071	0,069	1,032
δ_2	-0,585	0,040	-14,758	0,642	0,115	5,561
ϕ_2	-0,166	0,013	-12,541	0,116	0,038	3,024
δ_3	-0,148	0,018	-8,086	0,228	0,053	4,290
ϕ_3	-0,040	0,010	-3,926	-0,030	0,030	-1,007
δ_4	-0,136	0,012	-11,776	0,144	0,034	4,272
ϕ_4	0,049	0,009	5,557	-0,095	0,026	-3,705
δ_6	0,041	0,008	5,140	-0,032	0,023	-1,387
ϕ_6	-0,090	0,008	-11,522	0,007	0,023	0,292
δ_8	-0,004	0,007	-0,477	0,025	0,021	1,188
ϕ_8	0,055	0,007	7,405	-0,010	0,022	-0,462
δ_{12}	-0,016	0,007	-2,243	-0,004	0,021	-0,196
ϕ_{12}	0,012	0,007	1,711	-0,033	0,021	-1,582
δ_{16}	0,010	0,007	1,402	-0,029	0,021	-1,391
ϕ_{16}	0,045	0,007	6,374	0,003	0,021	0,160

Table 10: FFF regressions on log-volume, 9:30-17:30

Observing the table and the autocorrelations reported in Figure 8 we can note that a long memory effect is present both on the mean and on the variance of the volume. This effect is evident in the Portmanteau tests and in the Engle ARCH test. Interestingly Skewness and Kurtosis evidence a distribution very close to the normal a fact that is much more evident in Figure 15. A particular effect of filtration here becomes much more evident than in the return: the estimated seasonal component in the variance is removed by a division, since the values of the pattern are small compared to log-volume this affect the scale of the process (note the increase in the standard deviation). This can be explained considering that we removed two deterministic components and assuming that these components were compressing the stochastic dynamic generated by the underlying process. In the following analysis the resulting filtered volume in one more time rescaled, dividing it by 100.

	Log-Volume - mean			Log-Volume - variance		
Parameter	Estimate	Std. Err.	t-value	Estimate	Std. Err.	t-value
α_1	5,291	0,099	53,285	-7,032	0,291	-24,190
α_2	-1,804	0,283	-6,383	2,696	0,828	3,258
α_3	0,586	0,094	6,238	-0,835	0,275	-3,035
γ_1	0,049	0,760	0,064	-62,194	2,226	-27,945
γ_2	0,279	0,760	0,368	-62,194	2,226	-27,945
γ_3	-0,358	0,760	-0,471	-61,937	2,226	-27,829
γ_4	0,535	0,058	9,169	-0,069	0,171	-0,404
δ_1	0,381	0,055	6,886	-0,103	0,162	-0,635
ϕ_1	-0,152	0,015	-10,092	0,095	0,044	2,149
δ_2	-0,183	0,015	-12,137	0,030	0,044	0,674
ϕ_2	-0,176	0,010	-18,477	0,129	0,028	4,631
δ_5	0,050	0,007	7,004	-0,024	0,021	-1,160
ϕ_5	0,081	0,007	11,101	-0,045	0,021	-2,088
δ_6	0,102	0,007	14,433	-0,061	0,021	-2,955
ϕ_6	-0,055	0,007	-7,648	-0,009	0,021	-0,415

Table 11: FFF regressions on log-volume 9:15-17:30

	9:30-17:30			9:15-17:30		
Statistic	Raw	Mean filt.	Var. Filt.	Raw	Mean Filt.	Var. Filt.
Mean	4,618	-1,039e-13	0,018	4,647	-4,553e-14	-0,052
St. Dev.	0,936	0,763	71,988	0,942	0,758	73,603
Skewness	1,053	-0,285	-0,301	1,053	-0,289	-0,310
Kurtosis	1,137	3,853	3,511	1,137	3,866	3,439
Q(5)	39820,279	19185,399	19610,140	42157,828	19734,710	20377,917
Q(10)	62150,820	27802,570	28530,604	66140,347	28685,937	29768,113
Q(20)	83851,440	39924,860	41368,364	90418,897	40847,234	42945,089
Q(50)	93767,125	62361,783	66613,632	102471,400	65127,335	70888,597
Q(100)	174183,170	92785,812	103294,340	180922,470	97031,012	110257,680
$Q^2(5)$	40061,678	5418,088	5562,467	42274,666	5894,600	6000,630
Q^2 (10)	62074,589	7039,441	6849,840	65807,611	7706,708	7551,529
Q^2 (20)	83468,164	8718,860	8546,765	89803,130	9596,364	9627,444
Q^2 (50)	93201,321	9195,682	11105,105	101891,920	10188,584	12831,137
Q^2 (100)	174763,220	12307,430	14267,603	181726,110	13240,205	16805,901
JB	12188,761	1362,619	973,746	12569,571	1443,666	979,909
LM(2)	12020,309	7437,703	7582,311	12616,673	7694,255	7889,679
LM(5)	12492,956	7892,492	8042,759	13107,417	8157,775	8362,377
LM(10)	12572,968	8070,447	8227,787	13189,415	8345,803	8552,206
LM(20)	12622,467	8186,349	8351,431	13234,411	8461,731	8677,751

Table 12: Moments and tests of log-volume series

3.2 Univariate analysis

Consider now the two different series of filtered returns ($r^f_{t,n}$, in the following R_t) and volume ($v^f_{t,n}$, in the following V_t), as we can see from the ACF in figure 4 and 5, they show a different behavior. The returns clearly show a limited ARMA structure in the mean together with a long memory behavior in the variance, while volumes show long memory both in mean and in variance. In this section we try to fit a univariate model on these two series. The scope is to provide an initial base on which we can build up a multivariate model, as well as to have a set of starting value for its parameters. Tables 13 and 14 present a group of estimated models. For both series, in all tables, the first row indicates the order of the fitted model, respectively the orders p and q of the ARMA(p,q) or ARFIMA(p,d,q) filter and the orders m and l of the FIGARCH(m,d,l) filter. We considered model combinations up to the order 2 for p, q and m, while only up to 1 for l. An increase in the order resulted in non significative parameter estimates. In the first column of all tables we report the parameters used in the different specifications - inside the tables is reported the estimate and below it the correspondent standard error - and then in the order: the log-likelihood; the informations criteria of Akaike (AIC), Hannan-Quinn (HQ), Schwarz (BIC) and Shibata (SH); the Box-Pierce - Portmanteau - test for residual $Q(k)$ and squared residual correlation $Q^2(k)$, computed up to k lags; the sample Skewness and Kurtosis together with the Jarque-Bera normality test; the Engle Lagrange Multiplier test $LM(k)$ for residual ARCH effects computed up to lag k.

Model	(0,0)-(1,d,0)	(1,0)-(1,d,0)	(0,1)-(1,d,0)	(1,1)-(1,d,0)	(0,0)-(1,d,1)	(1,0)-(1,d,1)	(0,1)-(1,d,1)	(1,1)-(1,d,1)
μ	2.75E-05 0.00061	3.98E-05 0.00063	3.98E-05 0.00064	4.55E-05 0.00261	-4.51E-06 0.00353	8.06E-06 0.00355	8.24E-06 0.00367	1.32E-05 0.25241
φ		-0.02833 0.00898		0.36151 4.85485		-0.02991 0.02377		0.32795 0.04281
θ			0.02889 0.01853	0.39073 4.79923			0.03051 0.03646	0.35851 0.04354
ω	0.00153 0.00013	0.00152 0.00013	0.00152 0.00014	0.00153 0.00025	0.00069 0.00013	0.00069 0.00014	0.00069 0.00014	0.00069 2.22596
d	0.25586 0.01236	0.25603 0.01282	0.25597 0.01380	0.25579 0.03135	0.30957 0.01961	0.31017 0.01987	0.31016 0.02000	0.30976 0.01311
β_1	0.16933 0.01651	0.17120 0.01708	0.17113 0.02046	0.17039 0.03533	0.53173 0.05280	0.53356 0.05953	0.53365 0.05884	0.53312 0.01335
β_2								
ψ					0.31668 0.04808	0.31641 0.05471	0.31653 0.05365	0.31687 0.01109
LL	20268.976	20277.039	20277.191	20278.740	20303.557	20312.447	20312.616	20313.933
AIC	-1.70933	-1.70993	-1.70994	-1.70999	-1.71216	-1.71283	-1.71284	-1.71287
HQ	-1.70889	-1.70937	-1.70939	-1.70932	-1.71161	-1.71217	-1.71218	-1.71210
BIC	-1.70955	-1.71021	-1.71023	-1.71034	-1.71245	-1.71318	-1.71320	-1.71330
SH	-1.70933	-1.70993	-1.70994	-1.70999	-1.71216	-1.71283	-1.71284	-1.71287
Q(5)	18.85078	10.26376	10.36763	13.51954	18.95805	10.43115	10.60102	13.90824
Q(10)	36.29414	28.77150	28.91371	32.55014	36.21210	28.81657	29.03208	32.77107
Q(20)	41.57432	33.94126	34.08172	37.76260	41.11143	33.59037	33.80430	37.58477
Q(50)	71.29482	62.80330	62.92268	66.46945	70.66043	62.29166	62.48283	66.13738
Q(100)	134.17450	125.10899	125.20228	128.71044	133.50355	124.59651	124.75923	128.36427
$Q^2(5)$	23.98192	23.35388	23.31567	22.99517	7.68755	7.30511	7.28489	7.12123
$Q^2(10)$	32.90225	32.31422	32.27137	31.91545	23.39836	23.21221	23.19196	22.98272
$Q^2(20)$	41.90570	41.28031	41.23922	40.90941	30.61302	30.27016	30.24929	30.07850
$Q^2(50)$	72.61391	72.30590	72.30563	72.33763	59.10003	59.04754	59.06196	59.18873
$Q^2(100)$	122.99810	122.73769	122.75012	122.79710	103.47883	103.55817	103.57976	103.75905
Sk	-0.10058	-0.10346	-0.10358	-0.10519	-0.10986	-0.11273	-0.11291	-0.11481
K	5.42456	5.44365	5.44411	5.44764	5.45301	5.47354	5.47392	5.47518
JB	5887.646	5984.106	5986.565	6006.259	6040.208	6145.159	6147.340	6156.953
LM(2)	9.50025	0.66186	0.76004	2.70912	9.69072	0.90609	1.06615	3.34062
LM(5)	19.13298	10.18608	10.28797	13.33109	19.25091	10.34133	10.50924	13.68411
LM(10)	37.60637	28.52053	28.58866	31.26769	37.42796	28.41334	28.54356	31.33764
LM(20)	42.68345	33.72788	33.80572	36.48118	42.20497	33.31739	33.45759	36.26390

Table 13a - Returns univariate estimates, 9:30-17:30

Model	(0,0)-(2,d,0)	(1,0)-(2,d,0)	(0,1)-(2,d,0)	(1,1)-(2,d,0)	(0,0)-(2,d,1)	(1,0)-(2,d,1)	(0,1)-(2,d,1)	(1,1)-(2,d,1)
μ	7.64E-07	1.41E-05	1.40E-05	1.88E-05	-0.00001	0.00000	0.00000	0.00010
	0.00067	0.00063	0.00062	0.00032	0.00077	0.00059	0.00063	0.00002
ϕ		-0.03013		0.31752		-0.03014		0.31215
		0.00971		0.00974		0.01065		0.01387
θ			0.03072	0.34801			0.03088	0.34347
			0.00793	0.00990			0.00570	0.01369
ω	0.00116	0.00116	0.00116	0.00116	0.00069	0.00069	0.00073	0.00073
	0.00016	0.00014	0.00018	0.00003	0.00029	0.00007	0.00002	0.00001
d	0.29174	0.29220	0.29221	0.29202	0.30959	0.31044	0.30831	0.30717
	0.02743	0.02510	0.03090	0.01270	0.06685	0.04560	0.01589	0.00014
β_1	0.19927	0.20119	0.20120	0.20063	0.53171	0.53413	0.49996	0.50758
	0.03020	0.02830	0.03318	0.01164	0.13133	0.01487	0.01242	0.00447
β_2	0.06546	0.06634	0.06633	0.06591	0.00000	0.00000	0.00804	0.00534
	0.01475	0.01271	0.01723	0.00721	0.02262	0.00652	0.00220	0.00105
ψ					0.31666	0.31676	0.28468	0.29356
					0.11305	0.01459	0.01026	0.00465
LL	20299.419	20308.435	20308.591	20309.753	20303.248	20312.136	20312.153	20313.488
AIC	-1.71181	-1.71249	-1.71250	-1.71252	-1.71205	-1.71272	-1.71272	-1.71275
HQ	-1.71126	-1.71183	-1.71184	-1.71174	-1.71139	-1.71195	-1.71195	-1.71186
BIC	-1.71210	-1.71285	-1.71286	-1.71294	-1.71241	-1.71314	-1.71315	-1.71325
SH	-1.71181	-1.71249	-1.71250	-1.71252	-1.71205	-1.71272	-1.71272	-1.71275
Q(5)	18.55341	10.26593	10.44620	13.38770	18.96587	10.50184	10.67966	13.88587
Q(10)	35.88406	28.75246	28.97950	32.31248	36.22594	28.90185	29.12389	32.73771
Q(20)	40.76177	33.49980	33.72582	37.10219	41.12579	33.67477	33.88808	37.54254
Q(50)	70.35312	62.21079	62.41345	65.67977	70.67235	62.36522	62.53234	66.06296
Q(100)	133.37197	124.69673	124.87044	128.07073	133.51098	124.66132	124.84094	128.29017
Q(5)	9.22503	8.47474	8.46621	8.47691	7.68992	7.29941	7.20010	7.15259
Q(10)	22.55191	21.99316	21.98480	21.95299	23.41457	23.25141	22.98136	22.79164
Q(20)	30.09492	29.44087	29.43267	29.43317	30.63087	30.30976	30.07125	29.91734
Q(50)	58.84559	58.47477	58.50419	58.80339	59.10963	59.08427	58.88587	59.00288
Q(100)	105.09824	104.86301	104.89775	105.21418	103.49997	103.59309	103.59047	103.78650
Sk	-0.10663	-0.10963	-0.10979	-0.11155	-0.10972	-0.11259	-0.11255	-0.11702
K	5.39953	5.41609	5.41638	5.41761	5.45294	5.47388	5.46603	5.46926
JB	5778.282	5862.204	5863.831	5872.785	6039.634	6146.579	6108.179	6132.049
LM(2)	9.55314	1.02104	1.19174	3.26063	9.69294	0.96643	1.18284	3.47485
LM(5)	18.83094	10.16989	10.35067	13.18122	19.25825	10.40921	10.58024	13.65395
LM(10)	37.09723	28.32384	28.46372	30.91518	37.42925	28.47512	28.60107	31.28161
LM(20)	42.33707	33.20859	33.35947	35.83037	42.20431	33.37648	33.51036	36.21207

Table 13a - Returns univariate estimates, 9:30-17:30 (continued)

Model	(0,0)-(1,d,0)	(1,0)-(1,d,0)	(0,1)-(1,d,0)	(1,1)-(1,d,0)	(0,0)-(1,d,1)	(1,0)-(1,d,1)	(0,1)-(1,d,1)	(1,1)-(1,d,1)
μ	0.00013 0.00063	0.00014 0.00062	0.00014 0.00063	0.00016 0.19066	0.00013 0.00376	0.00014 0.00385	0.00014 0.00381	0.00015 0.21818
ϕ		-0.02565 0.00922		0.42192 0.01045		-0.02726 0.03642		0.40289 0.04587
θ			0.02627 0.00598	0.44907 0.00983			0.02795 0.04183	0.43134 0.04682
ω	0.00164 0.00014	0.00163 0.00014	0.00163 0.00014	0.00164 0.03274	0.00074 0.00016	0.00074 0.00017	0.00074 0.00014	0.00074 2.18177
d	0.25533 0.01327	0.25544 0.01262	0.25550 0.01336	0.25533 0.01811	0.30970 0.02161	0.31023 0.02024	0.31026 0.02156	0.30999 0.01327
β_1	0.17147 0.01743	0.17321 0.01659	0.17331 0.01723	0.17259 0.03018	0.53256 0.06133	0.53396 0.06735	0.53405 0.05588	0.53403 0.01368
β_2								
ψ					0.31468 0.05399	0.31406 0.06399	0.31412 0.05092	0.31495 0.01122
LL	20053.248	20060.122	20060.270	20062.447	20087.986	20095.655	20095.830	20097.919
AIC	-1.63989	-1.64037	-1.64038	-1.64048	-1.64265	-1.64319	-1.64321	-1.64329
HQ	-1.63946	-1.63983	-1.63984	-1.63983	-1.64211	-1.64255	-1.64256	-1.64254
BIC	-1.64010	-1.64064	-1.64066	-1.64082	-1.64292	-1.64354	-1.64355	-1.64371
SH	-1.63989	-1.64037	-1.64038	-1.64048	-1.64265	-1.64319	-1.64321	-1.64329
Q(5)	16.74298	9.01562	9.06510	11.83263	16.76302	9.15173	9.26263	12.62629
Q(10)	35.49032	28.58075	28.65834	31.92806	34.97897	28.24900	28.39478	32.23680
Q(20)	40.77848	33.67806	33.74751	36.98210	40.19574	33.26443	33.40103	37.20633
Q(50)	63.22973	55.81228	55.85915	58.91491	62.43762	55.21905	55.33140	58.93983
Q(100)	123.62569	116.22926	116.26453	119.34606	122.50444	115.42083	115.52293	119.14041
$Q^2(5)$	23.60086	22.98101	22.98773	22.68264	9.55847	9.27703	9.26226	9.03280
$Q^2(10)$	31.47417	31.00895	31.01970	30.65836	22.85976	22.88996	22.88194	22.57878
$Q^2(20)$	40.11776	39.66329	39.67018	39.25098	30.25413	30.14511	30.13167	29.76496
$Q^2(50)$	76.55835	76.08842	76.12834	76.29214	62.79184	62.62409	62.65019	62.87172
$Q^2(100)$	114.90379	114.13560	114.15018	114.11417	94.72366	94.25616	94.26204	94.36597
Sk	-0.11521	-0.11823	-0.11844	-0.12031	-0.12425	-0.12726	-0.12743	-0.12965
K	5.46592	5.48012	5.48055	5.48435	5.49406	5.50903	5.50938	5.51105
JB	6303.48130	6380.78930	6383.34300	6406.19250	6463.34210	6545.77870	6547.91470	6561.12520
LM(2)	8.53592	0.64727	0.69642	2.29473	8.60328	0.83048	0.93801	3.11289
LM(5)	16.89643	9.00721	9.06130	11.74673	16.91020	9.14342	9.25886	12.50059
LM(10)	36.79471	28.68870	28.72063	31.20123	36.21447	28.24845	28.34218	31.35756
LM(20)	41.90279	34.13719	34.17963	36.62043	41.30628	33.69240	33.79740	36.77267

Table 13b - Returns univariate estimates, 9:15-17:30

Model	(0,0)-(2,d,0)	(1,0)-(2,d,0)	(0,1)-(2,d,0)	(1,1)-(2,d,0)	(0,0)-(2,d,1)	(1,0)-(2,d,1)	(0,1)-(2,d,1)	(1,1)-(2,d,1)
μ	0.00013	0.00013	0.00014	0.00014	0.00013	0.00014	0.00014	0.00014
	0.00065	0.00065	0.00065	0.00059	0.00062	0.00059	0.00061	0.00124
ϕ		-0.02746		-0.00337		-0.02730		0.00985
		0.00874		0.00303		0.00995		0.02900
θ			0.02798	0.02461			0.02796	0.03792
			0.01047	0.00812			0.01561	0.03492
ω	0.00126	0.00125	0.00125	0.00125	0.00072	0.00074	0.00074	0.00074
	0.00018	0.00016	0.00016	0.00004	0.00015	0.00015	0.00016	0.00036
d	0.29027	0.29077	0.29076	0.29075	0.31256	0.31021	0.31025	0.31003
	0.02843	0.02596	0.02575	0.00587	0.01914	0.02041	0.01841	0.02373
β_1	0.20126	0.20323	0.20324	0.20329	0.54215	0.53399	0.53416	0.53259
	0.03177	0.02892	0.02899	0.00734	0.06714	0.06541	0.07374	0.26890
β_2	0.06248	0.06332	0.06329	0.06337	0.00000	0.00000	0.00000	0.00000
	0.01491	0.01350	0.01396	0.01060	0.00611	0.01032	0.00593	0.06798
ψ					0.32217	0.31411	0.31424	0.31270
					0.06320	0.05891	0.07279	0.25295
LL	20081.679	20089.447	20089.609	20089.589	20087.928	20095.647	20095.822	20095.884
AIC	-1.64213	-1.64268	-1.64270	-1.64261	-1.64256	-1.64311	-1.64312	-1.64305
HQ	-1.64159	-1.64204	-1.64205	-1.64186	-1.64191	-1.64236	-1.64237	-1.64219
BIC	-1.64241	-1.64303	-1.64304	-1.64303	-1.64290	-1.64352	-1.64354	-1.64353
SH	-1.64213	-1.64268	-1.64270	-1.64261	-1.64256	-1.64311	-1.64312	-1.64305
Q(5)	16.49265	9.08273	9.15565	9.16294	16.74581	9.17028	9.27547	9.28441
Q(10)	34.77389	28.25921	28.36156	28.36662	34.95693	28.27137	28.41023	28.42873
Q(20)	40.00478	33.28615	33.38079	33.38642	40.17120	33.29098	33.42074	33.43934
Q(50)	62.45029	55.43151	55.50264	55.51028	62.39380	55.23857	55.34440	55.35245
Q(100)	122.49036	115.57459	115.63336	115.64130	122.45837	115.44466	115.53990	115.54534
Q(5)	10.99589	10.24561	10.24395	10.27135	9.46179	9.27439	9.25730	9.30575
Q(10)	21.93702	21.48441	21.48296	21.51832	23.07331	22.90603	22.89638	22.92871
Q(20)	29.46079	28.93830	28.93224	28.96640	30.50280	30.15916	30.14421	30.17762
Q(50)	62.83596	62.23889	62.27477	62.29733	62.96523	62.63841	62.66256	62.72350
Q(100)	96.42160	95.50229	95.52033	95.53855	94.68516	94.24934	94.25272	94.32795
Sk	-0.12032	-0.12345	-0.12366	-0.12365	-0.12457	-0.12721	-0.12744	-0.12736
K	5.44622	5.45857	5.45878	5.45877	5.49969	5.50934	5.50974	5.50889
JB	6214.67600	6282.60960	6284.12750	6284.04860	6492.59660	6547.25890	6549.77030	6545.26060
LM(2)	8.48882	0.93029	1.00079	1.00743	8.56452	0.84663	0.94873	0.96108
LM(5)	16.63279	9.07422	9.15290	9.15930	16.90830	9.15388	9.26343	9.27441
LM(10)	36.02006	28.24624	28.30412	28.31169	36.19508	28.25537	28.34301	28.34838
LM(20)	41.13188	33.71257	33.77926	33.78685	41.28336	33.70020	33.79870	33.80827

Table 13b - Returns univariate estimates, 9:15-17:30 (continued)

Model	(0,0)-(1,d,0)	(1,0)-(1,d,0)	(0,1)-(1,d,0)	(1,1)-(1,d,0)	(0,0)-(1,d,1)	(1,0)-(1,d,1)	(0,1)-(1,d,1)	(1,1)-(1,d,1)
μ	0.11700 0.01583	0.11033 0.04535	0.11039 0.00062	0.11201 0.00104	0.11907 0.04866	0.11194 0.11734	0.11237 0.19111	0.11370 0.00484
d_1	0.35946 0.00456	0.32099 0.02441	0.32285 0.00764	0.33071 0.00763	0.35950 0.00545	0.32006 0.11521	0.32209 0.01159	0.32996 0.04102
ϕ		0.06434 0.03393		-0.22382 0.00812		0.06614 0.15584		-0.21216 0.04500
θ			-0.06318 0.00809	-0.27838 0.00776			-0.06487 0.02149	-0.26841 0.04064
ω	0.14206 0.00959	0.14145 0.00988	0.14146 0.02001	0.14159 0.02002	0.06363 0.03811	0.06253 0.05268	0.06252 0.01431	0.06259 0.03892
d_2	0.11194 0.00912	0.11231 0.00951	0.11230 0.02079	0.11218 0.02080	0.14333 0.02565	0.14564 0.03389	0.14566 0.01706	0.14550 0.01838
β_1	0.04800 0.01232	0.04793 0.01553	0.04775 0.00719	0.04743 0.00718	0.50761 0.25035	0.50926 0.35674	0.50924 0.08454	0.50907 0.00845
β_2								
ψ					0.42834 0.23784	0.42718 0.34015	0.42717 0.07758	0.42725 0.00812
LL	-20418.357	-20400.795	-20400.342	-20399.698	-20411.515	-20392.828	-20392.326	-20391.695
AIC	1.72262	1.72122	1.72118	1.72121	1.72213	1.72063	1.72059	1.72062
HQ	1.72317	1.72188	1.72185	1.72199	1.72279	1.72141	1.72136	1.72151
BIC	1.72233	1.72087	1.72083	1.72079	1.72177	1.72021	1.72016	1.72012
SH	1.72262	1.72122	1.72118	1.72121	1.72212	1.72063	1.72059	1.72062
Q(5)	34.57332	6.63460	6.32774	6.34487	35.99324	6.73005	6.37103	6.30414
Q(10)	45.98499	13.09564	12.87413	13.45068	47.27694	13.16592	12.88380	13.31902
Q(20)	70.60921	36.76616	36.47006	36.91686	72.07488	36.89140	36.53061	36.84973
Q(50)	120.86112	82.15293	82.11402	83.63370	122.18212	82.03702	81.94358	83.31116
Q(100)	297.40443	263.36231	263.36732	264.70790	297.13064	261.64435	261.56123	262.67555
$Q^2(5)$	5.66583	6.62650	6.68663	6.68673	2.40090	2.25109	2.24446	2.23926
$Q^2(10)$	11.35207	12.27045	12.30835	12.25793	12.16650	12.07145	12.03243	11.97545
$Q^2(20)$	18.25817	19.62016	19.69607	19.69548	18.46441	18.56711	18.56235	18.55531
$Q^2(50)$	61.79201	63.77407	63.76480	63.46354	59.33532	59.99969	59.90245	59.57979
$Q^2(100)$	146.17301	147.95737	147.77947	146.97363	137.27616	137.01438	136.72881	135.93867
Sk	0.02925	0.01293	0.01302	0.01441	0.03018	0.01350	0.01349	0.01501
K	3.58739	3.58419	3.58492	3.58729	3.58339	3.57990	3.58064	3.58301
JB	347.65511	338.50114	339.36509	342.40670	343.45380	333.68631	334.53138	337.60306
LM(2)	20.26017	1.16222	0.74158	0.90105	21.65544	1.31722	0.85177	0.93863
LM(5)	34.72364	6.58547	6.28051	6.37008	36.18900	6.67688	6.32125	6.32925
LM(10)	47.06863	12.95379	12.76945	13.56170	48.41181	12.99463	12.75475	13.41215
LM(20)	69.72182	36.49821	36.25337	36.81145	71.24155	36.62466	36.32169	36.76192

Table 14a - Volume univariate estimates, 9:30-17:30

Model	(0,0)-(2,d,0)	(1,0)-(2,d,0)	(0,1)-(2,d,0)	(1,1)-(2,d,0)	(0,0)-(2,d,1)	(1,0)-(2,d,1)	(0,1)-(2,d,1)	(1,1)-(2,d,1)
μ	0.11889 0.03696	0.11124 0.02457	0.11177 0.02672	0.11331 0.14670	0.11481 0.09962	0.10899 0.03834	0.10031 0.17679	0.11710 0.04138
d_1	0.35923 0.00454	0.32004 0.02083	0.32200 0.00419	0.32927 0.00754	0.35948 0.00571	0.31982 0.00805	0.32181 0.00738	0.32181 0.00860
ϕ		0.06571 0.02830		-0.19621 0.25200		0.06651 0.01057		0.01103 0.00333
θ			-0.06453 0.00989	-0.25286 0.24199			-0.06522 0.01690	-0.05397 0.01181
ω	0.12639 0.02457	0.12455 0.01645	0.12441 0.02249	0.12488 0.01690	0.06365 0.02429	0.06245 0.00988	0.06262 0.40831	0.06251 0.01116
d_2	0.12702 0.02733	0.12872 0.01882	0.12886 0.02574	0.12835 0.01824	0.14329 0.02238	0.14584 0.01253	0.14541 0.28797	0.14568 0.01268
β_1	0.06304 0.02501	0.06428 0.01576	0.06432 0.02550	0.06360 0.02975	0.50759 0.14547	0.50939 0.06413	0.50915 2.66091	0.50925 0.07502
β_2	0.02546 0.01423	0.02785 0.00753	0.02807 0.00894	0.02786 0.00223	0.00000 0.00487	0.00000 0.00325	0.00000 0.03818	0.00000 0.00486
ψ					0.42836 0.13353	0.42708 0.06363	0.42729 2.45407	0.42716 0.07491
LL	-20413.950	-20395.515	-20395.022	-20394.399	-20411.432	-20392.739	-20392.258	-20392.320
AIC	1.72233	1.72086	1.72082	1.72085	1.72220	1.72071	1.72067	1.72076
HQ	1.72299	1.72163	1.72159	1.72173	1.72298	1.72159	1.72155	1.72175
BIC	1.72198	1.72043	1.72039	1.72035	1.72178	1.72021	1.72017	1.72019
SH	1.72233	1.72086	1.72082	1.72085	1.72220	1.72071	1.72067	1.72076
Q(5)	35.62369	6.61926	6.29034	6.24518	35.99264	6.72174	6.34943	6.41426
Q(10)	47.01017	13.13015	12.87896	13.29630	47.26739	13.14125	12.83443	12.91766
Q(20)	71.70665	36.82492	36.49228	36.78221	72.06362	36.87191	36.49218	36.57261
Q(50)	121.83886	82.04839	81.97637	83.22912	122.16248	81.99170	81.86858	81.95053
Q(100)	297.45372	262.38620	262.32460	263.36101	297.13092	261.66597	261.67461	261.50855
Q(5)	3.08684	3.17812	3.19090	3.12164	2.40149	2.25916	2.24802	2.24158
Q(10)	11.03782	11.20409	11.21039	11.04000	12.16643	12.11521	12.01872	12.02995
Q(20)	17.41970	17.92278	17.96316	17.85122	18.47048	18.61755	18.56800	18.54731
Q(50)	59.68856	60.69757	60.63808	60.27640	59.33010	60.03347	59.87676	59.91608
Q(100)	139.70698	140.01049	139.74011	139.06606	137.27239	136.99482	136.71703	136.78377
Sk	0.02807	0.01158	0.01149	0.01271	0.03153	0.01460	0.01827	0.01154
K	3.58205	3.57899	3.57973	3.58203	3.58366	3.58007	3.58158	3.58011
JB	340.94961	332.26747	333.09941	335.97310	344.42831	334.12792	336.80876	333.53858
LM(2)	21.37830	1.20354	0.76802	0.86042	21.66284	1.32130	0.85321	0.89469905
LM(5)	35.81040	6.56798	6.24156	6.26440	36.18824	6.66687	6.29847	6.3631962
LM(10)	48.11912	12.96680	12.75600	13.37640	48.39892	12.96589	12.70401	12.779645
LM(20)	70.84015	36.56088	36.28537	36.68297	71.22661	36.61436	36.29428	36.352038

Table 14a - Volume univariate estimates, 9:30-17:30 (continued)

Model	(0,0)-(1,d,0)	(1,0)-(1,d,0)	(0,1)-(1,d,0)	(1,1)-(1,d,0)	(0,0)-(1,d,1)	(1,0)-(1,d,1)	(0,1)-(1,d,1)	(1,1)-(1,d,1)
μ	0.11649 0.01757	0.10936 0.09783	0.10983 0.02813	0.10964 0.00092	0.12124 0.18762	0.11284 0.02776	0.11343 0.05424	0.11270 0.00312
d_1	0.36078 0.00449	0.31820 0.02438	0.32109 0.00236	0.32225 0.00710	0.36096 0.00577	0.31753 0.00751	0.32058 0.00723	0.31936 0.04122
φ		0.07122 0.03344		-0.02554 0.00804		0.07262 0.00916		0.02663 0.04424
θ			-0.06866 0.00950	-0.09292 0.00706			-0.06995 0.00940	-0.04450 0.04143
ω	0.15456 0.00884	0.15348 0.00731	0.15354 0.00726	0.15345 0.00493	0.06640 0.08373	0.06521 0.02159	0.06521 0.02537	0.06520 0.03545
d_2	0.10602 0.00764	0.10672 0.00635	0.10667 0.00603	0.10675 0.00448	0.14269 0.07039	0.14504 0.02061	0.14505 0.02151	0.14509 0.01768
β_1	0.04226 0.01135	0.04238 0.00748	0.04219 0.00767	0.04229 0.00700	0.51160 0.46782	0.51305 0.12175	0.51301 0.15082	0.51304 0.00835
β_2								
ψ					0.42865 0.41178	0.42748 0.10973	0.42748 0.13835	0.42746 0.00808
LL	-21556.523	-21534.316	-21534.045	-21534.034	-21546.087	-21522.740	-21522.460	-21522.515
AIC	1.76351	1.76177	1.76175	1.76183	1.76274	1.76091	1.76088	1.76097
HQ	1.76404	1.76242	1.76240	1.76258	1.76338	1.76166	1.76164	1.76183
BIC	1.76323	1.76143	1.76141	1.76142	1.76239	1.76049	1.76047	1.76049
SH	1.76351	1.76177	1.76175	1.76183	1.76274	1.76091	1.76088	1.76097
Q(5)	40.99009	3.70419	3.97014	4.18340	42.34555	3.70311	3.95188	3.77998
Q(10)	57.46493	14.47699	14.80879	15.06835	58.70136	14.42208	14.72679	14.52103
Q(20)	88.97710	41.45387	41.95578	42.29846	90.58832	41.53727	42.03025	41.74294
Q(50)	133.98541	85.10423	85.32032	85.57649	135.55818	85.08629	85.27074	85.09235
Q(100)	316.97855	269.36125	269.31628	269.48933	318.74407	269.45968	269.35450	269.30201
$Q^2(5)$	7.25432	8.37280	8.42153	8.47204	3.40042	3.52405	3.57777	3.56345
$Q^2(10)$	15.05660	15.98653	16.04004	16.09594	14.21295	14.01503	14.05195	14.04657
$Q^2(20)$	23.51555	24.68344	24.76409	24.81630	22.25540	22.00387	22.06650	22.05259
$Q^2(50)$	65.78750	66.70121	66.66913	66.65604	58.80484	58.17203	58.11797	58.14016
$Q^2(100)$	149.65374	147.51783	147.39177	147.31459	132.79191	128.52072	128.33691	128.38551
Sk	0.01916	0.00138	0.00167	0.00201	0.02016	0.00233	0.00261	0.00263
K	3.47367	3.47478	3.47534	3.47554	3.46736	3.46822	3.46881	3.46862
JB	231.58662	229.68947	230.23249	230.43598	225.85845	223.41259	223.98374	223.80403
LM(2)	22.80587	0.87550	0.67777	0.66447	24.08530	1.00992	0.78289	0.81931
LM(5)	40.42655	3.71827	3.98163	4.19645	41.80659	3.71959	3.96542	3.79274
LM(10)	57.31673	14.44862	14.90384	15.20452	58.59437	14.37716	14.81023	14.55834
LM(20)	87.74866	41.92060	42.42509	42.75392	89.47994	42.02937	42.52746	42.24483

Table 14b - Volume univariate estimates, 9:15-17:30

Model	(0,0)-(2,d,0)	(1,0)-(2,d,0)	(0,1)-(2,d,0)	(1,1)-(2,d,0)	(0,0)-(2,d,1)	(1,0)-(2,d,1)	(0,1)-(2,d,1)	(1,1)-(2,d,1)
μ	0.11967 0.10759	0.11146 0.03370	0.11180 0.03364	0.11325 0.00071	0.11818 0.17706	0.11526 0.08719	0.10753 0.00077	0.11382 0.00094
d_1	0.36055 0.00452	0.31735 0.01964	0.32038 0.00404	0.32163 0.00556	0.36096 0.00630	0.31728 0.00782	0.32040 0.00359	0.32006 0.00387
ϕ		0.07250 0.02610		-0.02864 0.00962		0.07304 0.01037		0.01097 0.00784
θ			-0.06984 0.01007	-0.09716 0.00559			-0.07018 0.00385	-0.05952 0.00657
ω	0.13467 0.06532	0.13244 0.02968	0.13244 0.03772	0.13240 0.00780	0.06627 0.03330	0.06490 0.03746	0.06526 0.00836	0.06518 0.01061
d_2	0.12381 0.06677	0.12575 0.03118	0.12574 0.03931	0.12578 0.00930	0.14299 0.02879	0.14577 0.03104	0.14492 0.02075	0.14515 0.02525
β_1	0.06037 0.07387	0.06177 0.03323	0.06166 0.04582	0.06171 0.00691	0.51182 0.18770	0.51355 0.22066	0.51302 0.00709	0.51310 0.00784
β_2	0.02930 0.05790	0.03163 0.00942	0.03168 0.01771	0.03175 0.00550	0.00000 0.00276	0.00000 0.00441	0.00000 0.00881	0.00000 0.00694
ψ					0.42851 0.16734	0.42712 0.20145	0.42754 0.00489	0.42743 0.00717
LL	-21550.356	-21527.164	-21526.871	-21526.864	-21545.817	-21522.464	-21522.178	-21522.195
AIC	1.76308	1.76127	1.76125	1.76133	1.76280	1.76097	1.76094	1.76103
HQ	1.76373	1.76202	1.76200	1.76219	1.76355	1.76183	1.76180	1.76199
BIC	1.76274	1.76086	1.76083	1.76084	1.76238	1.76048	1.76046	1.76047
SH	1.76308	1.76127	1.76125	1.76133	1.76280	1.76097	1.76094	1.76103
Q(5)	42.01699	3.61687	3.88595	4.13091	42.35132	3.68617	3.92663	3.87653
Q(10)	58.45745	14.43315	14.75573	15.04284	58.70561	14.39118	14.68465	14.63437
Q(20)	90.04541	41.35573	41.85736	42.23590	90.58738	41.49075	41.96798	41.89947
Q(50)	135.03156	85.01889	85.20592	85.47801	135.55133	85.03492	85.21273	85.17439
Q(100)	318.58852	269.89655	269.79561	269.97987	318.71627	269.38949	269.37750	269.26931
Q(5)	5.25020	5.45214	5.47405	5.49123	3.43383	3.58433	3.58782	3.57761
Q(10)	14.73978	14.71480	14.73110	14.75338	14.29067	14.15679	14.05531	14.06048
Q(20)	22.69314	22.74768	22.78840	22.81732	22.34065	22.14920	22.07463	22.06878
Q(50)	61.24555	60.88454	60.80713	60.78960	58.85142	58.25160	58.11999	58.12739
Q(100)	138.86758	135.01865	134.81629	134.75630	132.78165	128.41941	128.38556	128.35084
Sk	0.01814	0.00037	0.00074	0.00041	0.02112	0.00122	0.00489	0.00234
K	3.46856	3.46992	3.47054	3.47067	3.46746	3.46797	3.46913	3.46864
JB	226.37264	224.99756	225.58710	225.70958	226.28338	223.14055	224.43286	223.81191
LM(2)	24.01884	0.91607	0.71869	0.72949	24.09038	1.03176	0.79469	0.80082
LM(5)	41.47744	3.63021	3.89679	4.14357	41.80698	3.70310	3.94023	3.89068
LM(10)	58.31201	14.39070	14.84258	15.17836	58.60167	14.35575	14.77810	14.70765
LM(20)	88.85737	41.83133	42.33892	42.70813	89.47743	42.00175	42.47681	42.40620

Table 14b - Volume univariate estimates, 9:15-17:30 (continued)

Analyzing the result of the univariate estimates and the corresponding information criteria and tests we choose the ARFIMA(1,d,1)-FIGARCH(1,d,1) for the returns and the ARFIMA(0,d,1)-FIGARCH(1,d,1) for the volume. These specifications were chosen both including and excluding the contracts concluded between 9:15 and 9:30. Our choice, given the results of the first chapter, was done only on the basis of the information criteria.

3.3 Multivariate analysis and causality

In this section we will apply the multivariate techniques to infer the causal relations between returns, volume and their variances. At first, using the univariate results, we will make use of the Cheung and Ng (1996) approach. Table 15 and graph 16 report the cross correlations on the two chosen parameterization, the ARMA(1,1)-FIGARCH(1,d,1) for the returns and the ARFIMA(0,d,1)-FIGARCH(1,d,1) for the volume. This analysis is performed using both the data from 9:15 to 17:30 and from 9:30 to 17:30.

	9:15-17:30				9:30-17:30			
Lead/Lag	V/R	V^2/R^2	V^2/R	V/R^2	V/R	V^2/R^2	V^2/R	V/R^2
-10	-0.01826	0.50718	-0.31762	-0.38667	0.90687	-2.21124	0.32729	1.67308
-9	0.78680	0.91194	-1.07272	0.27104	1.26709	-2.23262	-0.24770	2.53857
-8	-2.02765	-0.17769	-1.04216	-1.03738	-1.83346	-2.48006	-0.66144	3.04565
-7	-0.81417	1.74617	-0.59147	-0.30585	-1.11127	-1.94084	-0.92216	2.98660
-6	-1.20432	-0.63198	0.27625	1.90145	-0.72912	-2.85446	0.58373	5.70640
-5	0.12282	-0.46899	1.10970	-1.20836	0.57061	-3.61027	0.64820	2.95306
-4	0.85339	-0.04509	0.51313	-0.16476	0.39591	-2.93904	-0.45726	4.81035
-3	-2.30754	1.09615	-1.44858	1.56548	-2.69019	-1.30133	-1.81808	6.80716
-2	0.27484	0.27312	1.62994	0.80313	-0.41538	-2.15356	1.59608	6.42653
-1	-0.47272	-1.23728	-0.25092	4.66293	-0.81232	-3.22617	-0.53047	8.45160
0	-5.73078	36.08034	-5.37173	55.63877	-5.67801	21.05093	-5.09788	45.02831
1	-3.79714	-4.21649	-1.66454	14.81189	-4.20519	-5.66306	-2.10117	13.89140
2	-2.68483	-6.07146	0.25189	2.69746	-2.86569	-8.91897	0.04828	3.85557
3	-2.32943	-4.34449	1.60574	1.91942	-2.42342	-7.74508	1.14448	2.51904
4	-1.55644	-4.04129	-0.21084	-0.43554	-2.58295	-6.73280	-0.25177	0.36674
5	-2.92346	-2.78649	0.63012	2.15698	-2.17875	-5.98941	0.36503	1.32318
6	-0.38253	-3.32015	0.58343	0.54528	-0.33088	-6.55886	0.32054	0.69176
7	-1.42050	-2.18284	0.06042	0.19849	-1.40548	-5.80164	0.30047	0.70447
8	-0.84052	-1.84662	0.99365	-1.05743	-1.38084	-4.77306	1.09016	0.15303
9	-1.20721	-2.51346	-0.02200	0.23210	-0.10730	-6.27712	0.38175	1.21193
10	-1.10823	-1.58791	1.16384	-0.03475	-1.78093	-4.73428	1.26003	0.33786

Table 15: cross correlations

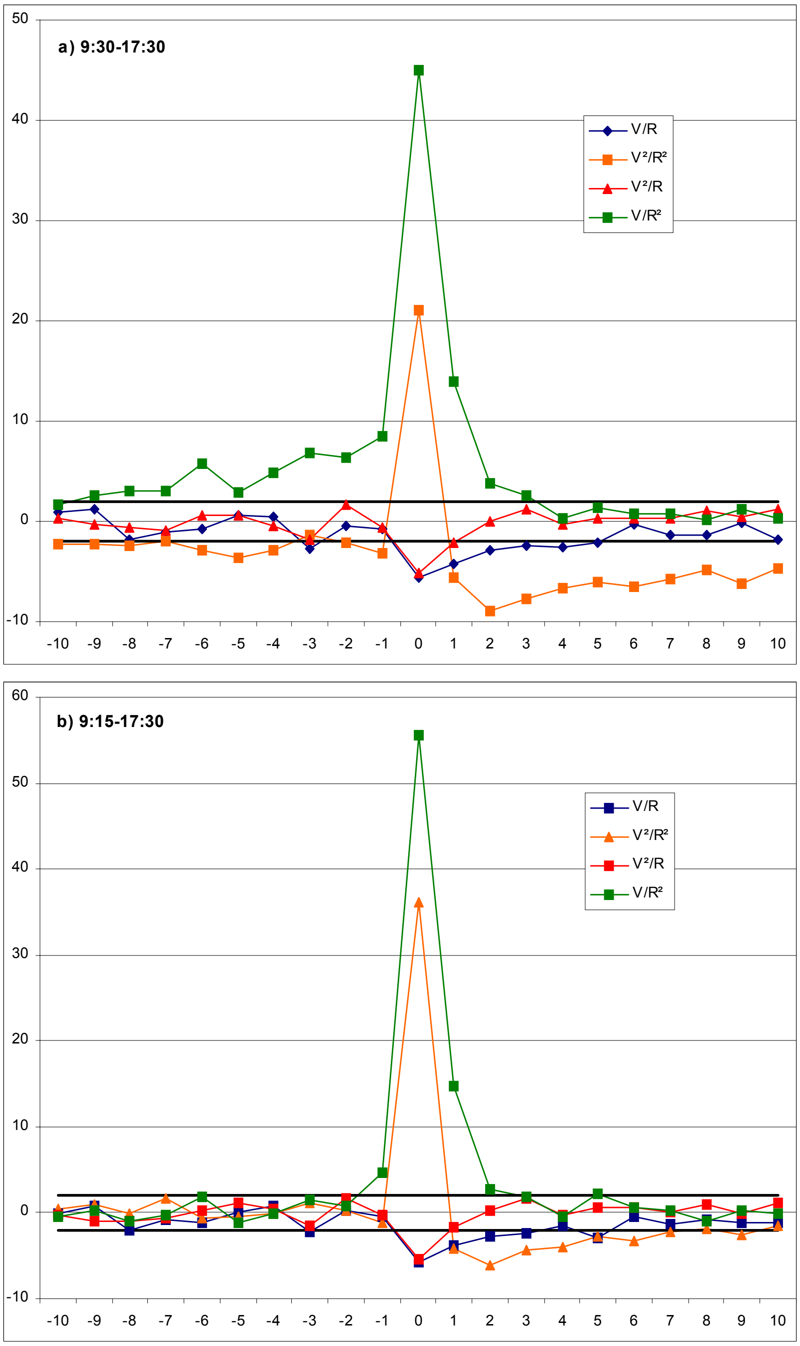

Figure 16: cross correlations

We considered not only the cross-correlations among the mean components and the variance components of returns and volume, but we extended the approach of Cheung and Ng (1996) also to the analysis of the mixed cross-correlations, that is among one mean and one variance component. This extension was motivated by the possibility of a GARCH-in-mean structure, whose presence can be suggested by significant mixed cross-correlations. We reported the cross-correlations in table 15, where R and R^2 (V and V^2) stand respectively for the mean component of the return (volume) and for its variance component. Recall that these cross correlations are distributed as a standardized normal variables. All the cross-correlations are calculated as if the second variable lead/lag the first. In the graphs we reported the level of the cross-correlations together with a 5% confidence bands. Observing both the table and the graphs we can note the strength of the contemporaneous effect among the variables, a fact that may suggest that the 5 minute interval is too long to correctly detect the causality relation. Apart this we note that the structure change including or excluding the 9:15-9:30 transactions, as if these strongly affect the structure of the causal relation, therefore we focus on the whole sample, the case of graph panel b) to include all possible source of information. Analyzing the cross-correlation a causal relation is evident, past returns values are relevant for the volume both in mean and in the variance. There is also evidence of a limited effect in the mixed cross-correlations. Given these results we tried to include all relevant lagged residuals and squared residuals and the to estimate the model. However the parameters correspondent to the cross-terms resulted to be not significant therefore we considered only a limited correction with the following representation (only for the full sample):

$$\begin{aligned}
R_t &= \mu_R + \varepsilon_{R,t} + \theta_R \varepsilon_{R,t-1} + \gamma_{R,1} V_t + \gamma_{R,2} V_{t-1} \\
\sigma^2_{R,t} &= \omega_R + \beta_R \sigma^2_{R,t-1} + \left[1 - \beta_R L - (1-L)^{d_R}(1-\psi_R L)\right] \varepsilon^2_{R,t} \\
&\quad + \gamma_{R,3} \varepsilon^2_{V,t} + \gamma_{R,4} \varepsilon^2_{V,t-1} \\
(1-L)^{d_1}(V_t - \mu_V) &= +\varepsilon_{V,t} + \theta_V \varepsilon_{V,t-1} + \gamma_{R,1} R_t + \gamma_{R,2} R_{t-1} \\
\sigma^2_{V,t} &= \omega_V + \beta_V \sigma^2_{V,t-1} + \left[1 - \beta_V L - (1-L)^{d_V}(1-\psi_V L)\right] \varepsilon^2_{V,t} \\
&\quad + \gamma_{R,3} \varepsilon^2_{R,t} + \gamma_{R,4} \varepsilon^2_{R,t-1}
\end{aligned} \tag{29}$$

The inclusion of the terms in the return mean is justified by the significance of the third lead, the inclusion of additional terms caused only nonsignificant parameters. We estimated again the model, the results are in table 16.

Parameters	Returns	Volume
μ	0.00026 0.00062	-0.02856 0.00024*
d_1		0.31929 0.00008*
θ	0.02800 0.00621*	-0.06946 0.00011*
γ_1	0.00059 0.00099	-0.01007 0.00004*
γ_2	0.00001 0.00099	-0.19015 0.00097*
ω	0.00072 0.00009*	0.13931 0.00025*
d_2	0.31104 0.01448*	0.10262 0.00014*
β	0.53426 0.03058*	0.03327 0.00017*
ψ	0.31358 0.02678*	1e-38 0.00008
γ_3	8e-41 0.00010	0 0.00009
γ_4	0.00003 0.00010	0.98449 0.01053*
LL	20096.289	-21366.165
AIC	-1.64292	1.74850
HQ	-1.64184	1.75968
BIC	-1.64354	1.74779
SH	-1.64292	1.74850
Sk	-0.12906	-0.05921
Kt	5.50942	3.47353
JB	6551.544	257.029
Q(5)	9.424	6.662
Q(10)	28.667	17.716
Q(20)	33.675	45.149
Q(50)	55.604	89.785
Q(100)	115.95	276.234
$Q^2(5)$	9.224	3.106
$Q^2(10)$	22.886	10.165
$Q^2(20)$	30.080	26.940
$Q^2(50)$	62.639	81.771
$Q^2(100)$	94.206	192.151

Table 16: estimated model with correction for causality

Significant parameters are marked with a star. We can note that only in the volume equation we obtain significant estimates, as if the evidence of the cross-correlation analysis is detected by the model only in limited cases, where the relation has its maximum strength.

The previous estimations were a first attempt to causality testing, now we will turn to a pure multivariate setting. At this point we estimate a VARFIMA(1,d,0) with residuals following a multivariate constant conditional correlation FIGARCH, that will be used as a benchmark model. The representation we use is the following, where r_t is the log-return series and v_t is the volume series, both filtered from deterministic components:

$$\begin{aligned}
\begin{bmatrix} R_t \\ \tilde{V}_t \end{bmatrix} &= \begin{bmatrix} \mu_R \\ \mu_V \end{bmatrix} + \begin{bmatrix} \phi_{1,1} & \phi_{1,2} \\ \phi_{2,1} & \phi_{2,2} \end{bmatrix} \begin{bmatrix} R_{t-1} \\ \tilde{V}_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{R,t} \\ \varepsilon_{V,t} \end{bmatrix} \\
\tilde{V}_t &= (1-L)^{d_{1,2}} V_t \\
\begin{pmatrix} \varepsilon_{R,t} \\ \varepsilon_{V,t} \end{pmatrix} &\sim F\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{R,t} & \sigma_{RV,t} \\ \sigma_{RV,t} & \sigma^2_{V,t} \end{bmatrix} \right) \\
\sigma^2_{R,t} &= \omega_1 + \beta_1 \sigma^2_{R,t-1} + \left[1 - \beta_1 L - (1-\psi_1 L)(1-L)^{d_{2,1}}\right] \varepsilon^2_{R,t} \\
\sigma^2_{V,t} &= \omega_2 + \beta_2 \sigma^2_{V,t-1} + \left[1 - \beta_2 L - (1-\psi_2 L)(1-L)^{d_{2,2}}\right] \varepsilon^2_{V,t} \\
\sigma_{RV,t} &= \rho \sigma_{R,t} \sigma_{V,t}
\end{aligned} \tag{30}$$

The model is estimated including data recorded from 9:15 and 9:30. Results are in Table 17. As we can observe our result contradict the theory: there is evidence of the return causing the volume in the mean but not the opposite. However we can explain this behavior reasoning on market structure. In the FIB30 market, the market makers quote prices for different maturities and we assume that they are as well as dealers, informed traders. Therefore as soon as a news impact on the market they react immediately changing prices. Differently brokers and other traders, less informed that market makers, react observing price changes quoted by the firsts and adjust then their positions. Using this sequence we can expect that changes in prices, responsible of changes in returns, will have a significant impact on the market volume.

Parameters i=1,2	Whole sample		02/01/02-15/03/02	
	Returns	Volume	Returns	Volume
μ_i	-0.00006 0.00060	0.10390 0.04851**	-0.00098 0.00317	0.25761 0.08650*
$d_{1,i}$		0.31751 0.00808*		0.31944 0.01752*
$\phi_{i,1}$	-0.02714 0.00699*	-0.11804 0.03113*	-0.01533 0.02742	-0.12402 0.05953*
$\phi_{i,2}$	0.00073 0.00101	0.07261 0.01079*	-0.00110 0.00409	0.10186 0.02426*
ω_i	0.00075 0.00008*	0.03912 0.00646*	0.00012 0.00148	0.14921 0.02351*
$d_{2,i}$	0.31319 0.01433*	0.16887 0.01549*	0.89112 0.24031*	0.08892 0.02126*
β_i	0.53139 0.02948*	0.68221 0.04202*	0.89312 0.09565*	0.01883 0.02839
ψ_i	0.30858 0.02589	0.58549 0.04396*	0.09571 0.15126	
ρ	-0.03625 0.00635*		-0.08009 0.03043	
AIC	-0.11433		-0.02897	
HQ	-0.11261		-0.02211	
BIC	-0.11541		-0.03388	
SH	-0.11433		-0.02898	
Q(5)	8.932	4.213	8.524	1.629
Q(10)	28.303	15.330	22.868	8.251
Q(20)	33.365	42.602	29.862	23.235
Q(50)	55.497	86.060	51.983	48.294
Q(100)	115.731	270.169	100.756	130.971
$Q^2(5)$	9.411	6.543	4.698	1.484
$Q^2(10)$	21.794	15.599	11.810	5.625
$Q^2(20)$	28.645	23.804	20.890	24.365
$Q^2(50)$	61.259	58.715	46.827	61.304
$Q^2(100)$	93.059	128.313	105.199	116.434

Table 17: bivariate model for returns and volume

All the previous observations take into consideration only causality in mean, we have not yet analyzed second order causality within a multivariate framework. Given the complexity of the involved models and the CPU time needed to perform the estimation, from now on our results are based on a restricted sample. We consider the whole records of data (from 9:15 to 17:30) for a limited period of time, ranging from 2 January 2001 up to 15 March 2001, a total of 53 days, 5247 observations. To get a benchmark model for test on residual we re-estimated the baseline model, VARFIMA(1,d,0) with a CCC-FIGARCH, this is again reported in table 17. In this case the Volume GARCH structure has been modified into a FIGARCH(1,d,0), since the FIGARCH(1,d,1) performed very poorly. Although the β in the estimated FIGARCH(1,d,0) for the volume resulted non significant we decided to maintain it in the model, this to preserve a limited short memory pattern within the volume series.

In the estimate of the model we applied numerical optimizations algorithm that allow for nonlinear constraints on parameters. We used the BFGS optimization algorithm for the first up to the third iteration, then the procedure switch to the Newton method. The estimates obtained by this methodology for the benchmark model are reported in table 17. In Annex 4 we report all the different specifications estimated for the variance components.

Given the benchmark estimates the following step is the analysis of multivariate GARCH models that consider also the causal relation among variances. To be specific we estimated the extensions previously introduces, that is adding in the return (or volume) GARCH equation the lagged variance or squared residuals of the volume (return). This was done both with a simple linear term, as if it was an exogenous term or via an exponential term that multiply the equation (see formula **??**). Given that the model structure change we tried a wide range of possible specifications for the GARCH terms, starting form the GARCH(1,1) up to the FIGARCH(1,d,1). All the results are reported in the appendix, while here we focus on the most promising ones.

Specification	No causality		Exogenous Residuals		Exogenous Variance	
	Return	Volume	Return	Volume	Return	Volume
	Figarch(1,d,1)	Figarch(1,d,0)	Figarch(1,d,1)	Figarch(0,d,0)	Figarch(1,d,1)	Figarch(0,d,0)
Parameters	Estimate	Sd. Err.	Estimate	Sd. Err.	Estimate	Sd. Err.
VARFI(1,1)	-0.00098	0.00317	-0.00108	0.00138	-0.00105	0.00137
	0.25761	0.08650	0.25052	0.08058	0.25244	0.08024
	0.31944	0.01752	0.31993	0.01804	0.31974	0.01768
	-0.01533	0.02742	-0.01538	0.01869	-0.01524	0.02087
	-0.00110	0.00409	-0.00128	0.00281	-0.00124	0.00253
	-0.12402	0.05953	-0.12225	0.06054	-0.12202	0.05830
	0.10186	0.02426	0.10055	0.02383	0.10171	0.02364
Return	0.00012	0.00148	0.00002	0.00007	0.00000	0.00017
	0.89112	2.71137	0.96960	0.17348	0.95616	0.28148
	0.89312	1.31620	0.92879	0.06539	0.92281	0.11376
	0.09571	1.46300	0.04987	0.11706	0.05800	0.17707
Volume	0.14921	0.02351	0.16124	0.01650	0.16045	0.01643
	0.08892	0.02126	0.07877	0.01401	0.07946	0.01407
	0.01883	0.02839				
Causality			0.00016	0.00016	0.00027	0.00050
			0.00000	0.18917	0.00000	0.01894
Correlation	-0.08008	0.03044	-0.08077	0.02003	-0.08073	0.02031
Log-likelihood	87.414		88.618		87.940	
AIC	-0.02897		-0.02905		-0.02878	
HQ	-0.02211		-0.02174		-0.02147	
BIC	-0.03388		-0.03434		-0.03407	
SH	-0.02898		-0.02907		-0.02880	
	Return	Volume	Return	Volume	Return	Volume
Q(5)	8.524	1.629	8.678	1.647	8.729	1.676
Q(10)	22.868	8.251	23.356	8.254	23.349	8.286
Q(20)	29.862	23.235	30.680	23.184	30.552	23.184
Q(50)	51.983	48.294	52.214	48.128	52.391	48.148
Q(100)	100.756	130.971	100.378	131.217	100.748	131.117
$Q^2(5)$	4.698	1.484	4.640	1.724	4.557	1.728
$Q^2(10)$	11.810	5.625	11.079	5.860	11.239	5.869
$Q^2(20)$	20.890	24.365	19.142	24.705	19.562	24.724
$Q^2(50)$	46.827	61.304	44.689	62.826	45.284	62.835
$Q^2(100)$	105.199	116.434	102.225	118.437	102.944	118.432

Table 18.a: bivariate model estimates with causality components

Specification	No causality		Exponential Causality	
	Return	Volume	Return	Volume
	Figarch(1,d,1)	Figarch(1,d,0)	Figarch(1,d,1)	Garch(1,1)
Parameters	Estimate	Sd. Err.	Estimate	Sd. Err.
VARFI(1,1)	-0.00098	0.00317	-0.00122	0.00134
	0.25761	0.08650	0.33331	0.10461
	0.31944	0.01752	0.33207	0.01845
	-0.01533	0.02742	-0.02107	0.01599
	-0.00110	0.00409	-0.00156	0.00249
	-0.12402	0.05953	-0.13724	0.05985
	0.10186	0.02426	0.10747	0.02427
Return	0.00012	0.00148	0.00017	0.00006
	0.89112	2.71137	0.07249	0.01404
	0.89312	1.31620	0.92341	0.01481
	0.09571	1.46300		
Volume	0.14921	0.02351	0.01391	0.01044
	0.08892	0.02126	0.03002	0.01749
	0.01883	0.02839	0.95826	0.04313
Causality			-0.00611	0.00706
			-0.03655	0.01096
Correlation	-0.08008	0.03044	-0.08374	0.01944
Log-likelihood	87.414		105.821	
AIC	-0.02897		-0.03593	
HQ	-0.02211		-0.02862	
BIC	-0.03388		-0.04122	
SH	-0.02898		-0.03595	
	Return	Volume	Return	Volume
Q(5)	8.524	1.629	10.502	2.192
Q(10)	22.868	8.251	26.422	8.101
Q(20)	29.862	23.235	32.867	24.111
Q(50)	51.983	48.294	55.710	47.222
Q(100)	100.756	130.971	100.711	128.061
$Q^2(5)$	4.698	1.484	6.743	10.741
$Q^2(10)$	11.810	5.625	12.330	16.797
$Q^2(20)$	20.890	24.365	18.877	33.713
$Q^2(50)$	46.827	61.304	42.428	63.528
$Q^2(100)$	105.199	116.434	97.129	119.462

Table 18.b: bivariate model estimate with exponential causality

Table 18 is divided in two panels, in both we report in the first two columns the benchmark model, the remaining are devoted to the causality estimates, in panel a with the introduction of an exogenous variable in the variance equation, while in panel b with the introduction of a multiplicative exponential factor.

Recalling equation (30) the variance component has been modified into

$$\begin{aligned}
\sigma^2_{R,t} &= \omega_1 + \beta_1 \sigma^2_{R,t-1} + \gamma_1 f\left(V_{t-1}\right) + \\
& \left[1 - \beta_1 L - \left(1 - \phi_1 L\right)\left(1 - L\right)^{d_{2,1}}\right] \varepsilon^2_{R,t} \\
\sigma^2_{V,t} &= \omega_2 + \beta_2 \sigma^2_{V,t-1} + \gamma_2 f\left(R_{t-1}\right) + \\
& \left[1 - \beta_2 L - \left(1 - L\right)^{d_{2,2}}\right] \varepsilon^2_{V,t}
\end{aligned} \tag{31}$$

for panel a, where $f\left(V_{t-1}\right) = z^2_{V,t-1}$ and $f\left(R_{t-1}\right) = z^2_{R,t-1}$. For panel b the estimated variance equations are

$$\begin{aligned}
\sigma^2_{R,t} &= \exp\left(\gamma_1 f\left(V_{t-1}\right)\right)\left(\omega_1 + \alpha_1 \varepsilon^2_{R,t-1} + \beta_1 \sigma^2_{R,t-1}\right) \\
\sigma^2_{V,t} &= \exp\left(\gamma_2 f\left(R_{t-1}\right)\right)\left(\omega_2 + \alpha_2 \varepsilon^2_{V,t-1} + \beta_2 \sigma^2_{V,t-1}\right)
\end{aligned} \tag{32}$$

where $f\left(V_{t-1}\right)$ and $f\left(R_{t-1}\right)$ are substituted as before. In the table parameters are reported in the following order: for the VARFI(1,d) μ_R μ_V $d_{1,2}$ $\phi_{1,1}$ $\phi_{1,2}$ $\phi_{2,1}$ $\phi_{2,2}$; for the return ω_1 α_1 β_1; for the volume ω_2 α_2 β_2 both models have a GARCH(1,1) specification, as suggested in the theoretical presentation of the model. The remaining three parameters are, in the order, γ_1 γ_2 ρ. Finally note that the log-likelihoods and the information criteria (to be minimized) can be directly compared across the different specifications. First of all we can note that the causality parameters are non-significant in panel a, while in panel b the causality effect of return variance on volume variance became significant, as we were expecting. Moreover the likelihood clearly increase and all the information criteria prefer the second fitted model. Interestingly the tests on residuals are very close between the two models, this could be expected in the mean since both maintain the same structure, however some difference was expected in the variance because the first model include a long memory behavior while the second is a modified GARCH model, a short memory one. The two models return very close Box-Pierce statistics for elevate lags (50-100), while for the lags up to 20 it seems that the causality GARCH model cannot explain all the correlation. In principle the reverse should appear, the GARCH do not explain high lag correlation since it is a short memory one, however this result has been obtained. We conjecture that, at least in this case, the long memory behavior of the volume and return volatility is generated by the causality structure and an increase of the lags of the GARCH structure should be considered in order to explain the residual correlation for the lags 2-5. Moreover an interesting comparison could be the one of these results with a causality FIGARCH model. These estimations will be the object of future researches.

4 Conclusions

Within this work we focused on second order causality an on its application on a real case. We reviewed the current definitions concerning causality among variances providing a first result, the parameters restrictions that have to be satisfied in order to obtain non-causality among mean and variances in a VAR-GARCH-M model. We considered then the current techniques used to detect and estimate a causal relation within variances pointing out their main drawback, relative to the impossibility of preserving the sign of the relation. In fact, in the GARCH models, even if multivariate, we have to satisfy some parameter positivity constraints or to use quadratic formulations that give non interpretable results, this in order to preserve the positivity of variances and the semi-positive definiteness of the variance-covariance matrix. We suggest therefore an extension, actually only in the bivariate case, presenting a GARCH-type model that can be used for the detection of second order causality and that determine also its direction. We applied then our model on a real case, studying the relation among the returns and the volume of the FIB30, the future on the Italian stock market index. In this paper we present also all the preliminary analysis concerning the study of the cyclical component of the series, and of the estimation of univariate models for both return and volume. We then compare our model with a simple benchmark without causality. We show that the increase in the likelihood is really evident, indicating a clear preference for our formulation. For the suggested model we provided the stationarity restrictions, but at the moment only for the GARCH(1,1) and IGARCH(1,1) case. For the second these restrictions imply that a negative causality relation (if it exist), therefore we suggest a careful application and interpretation of the integrated model. We mentioned in the chapter some possible extensions of our model, both modifying the GARCH part (for example to a FIGARCH structure) or the causality relation including terms that considers not only the modulus of the shocks but also their sign, or that allow for different distributions, far from the normality case, and finally changing the exponential representation to a logistic case. All these possibilities and their extension to multivariate models of order greater that two will be the object of future researches, that will focus both on the required stationarity conditions and to the application of such models in real cases for example in the volatility spillover study, in contagion models, or in portfolio pricing framework.

References

[1] Blume L., D. Easley and M. O'Hara, 1994, Market statistics and technical analysis: the role of the volume, *Journal of Finance*

[2] Bollerslev T., 1990, Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH approach, *Review of Economic Studies*, 72, 498-505

[3] Boudjellaba H, J. Dufour and R. Roy, 1992, Testing causality between two vectors in multivariate autoregressive moving average models, *Journal of the American Statistical Association*, 87 (420), 1082-1090

[4] Boudjellaba H, J. Dufour and R. Roy, 1994, Simplified conditions for noncausality between vectors in multivariate ARMA models, *Journal of Econometrics*, 63, 271-287

[5] Brunetti C. and C.L. Gilbert, 2000, Bivariate FIGARCH and fractional cointegration, *Journal of Empirical Finance*, 7, 509-530

[6] Caporin M., 2002, Long memory conditional heteroskedasticity and second order causality, PhD dissertation, University of Venice

[7] Cheung Y. and L.K. Ng, 1996, A causality-in-variance test and its application to financial market prices, *Journal of Econometrics*, 72, 33-48

[8] Comte F. and O. Lieberman, 2000, Second-order noncausality in multivariate GARCH processes, *Journal of Time Series Analysis*, 21 (5), 535-557

[9] Engle R.F. and K.F. Kroner, 1995, Multivariate simultaneous generalized ARCH, *Econometric Theory*, 11, 122-150

[10] Engle R.F. and K. Sheppard, 2001, Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH, University of California, San Diego, Discussion pape 2001-15

[11] Granger C.W.J., 1980, Testing for causality: a personal viewpoint, *Journal of Economics Dynamic and Control*, 2 (4), 329-352

[12] Granger C.W.J., 1988, Some recent developments in a concept of causality, *Journal of Econometrics*, 39, 199-211

[13] Granger C.W.J., R.P. Robins and R.F. Engle, 1986, Wholesale retail prices: bivariate time series modeling with forecastable error variances. In *Model reliability*, Boston MA, MIT press

[14] Hamilton J.D. and R. Susmel, 1994, Autoregressive conditional heteroskedasticity and changes in regime, *Journal of Econometrics*, 64, 307-333

[15] Hafner, C.M., 2001, Fourth moments of multivariate GARCH processes, working paper

[16] Hannan E. J., 1970, Multiple time series, Wiley, New York

[17] Karpoff J, 1987, The relation between price changes and trading volume: a survey, Journal of Financial and Quantitative Analysis, 22, 109-126

[18] LeBaron B., 1992, Persistence of the Dow Jones index on rising volumes, Working paper, Wisconsin Madison - Social System

[19] Li W.K., S. Ling and M. McAleer, 2001, A survey of recent theoretical results for time series models with GARCH errors, working paper

[20] Pojarlev M. and W. Polasek, 2000, Portfolio construction using multivariate time series forecasts, Working paper, Institue for Statistics and Econometrics, University of Basel

[21] Sims C.A., 1972, Money, income and causality, American Economic Review, 62, 540-552

[22] Teyssière G., 1998, Multivariate long-memory ARCH modelling for high frequency exchange rates, University of London, working paper

[23] Teyssière G., Modelling exchange rates volatility with multivariate long-memory ARCH processes, GREQAM working paper

5 Appendix

5.1 Filtering the cyclical component

We apply the Fast Fourier Form modeling, as in Andersen and Bollerslev (1997). This techniques is due to Gallant (1981, 1982) and is just a kind of regression on Fourier frequencies, trying to filter the periodic component via a mixture of harmonics. Assuming the following representation

$$x_{t,n} = E\left[x_{t,n}\right] + \frac{\sigma_{t(,n)} s_{t,n} Z_{t,n}}{\sqrt{N}} \tag{33}$$

that is the return at time t (day), intraday period n is equal to its expected value plus an heteroskedastic (daily or intradaily in that case denominator will be 1) error. Acting as in Bollerslev and Mikkelsen we estimate on unfiltered data the conditional variances with an MA(1)-GARCH(1,1), then we apply the following log-transformation

$$y_{t,n} = 2\ln\left[\left|x_{t,n} - E\left[x_{t,n}\right]\right|\right] - \ln\sigma^2_{t(,n)} + \ln N = \ln s^2_{t,n} + \ln Z^2_{t,n} \tag{34}$$

The resulting relation can be also viewed as the sum of a cyclical function plus a noise

$$y_{t,n} = f\left(\sigma_{t(,n)}, n, \theta\right) + \eta_{t,n} \tag{35}$$

Following the cited papers we specified the function as

$$f\left(\sigma_{t(,n)}, n, \theta\right) = \sum_{j=0}^{J} \sigma^j_{t(,n)} \left[\begin{array}{c} \alpha_{1,j} + 2\alpha_{2,j} n/\left(N+1\right) + 6\alpha_{3,j} n^2/\left(\left(N+1\right)\left(N+2\right)\right) \\ + \sum_{i=i}^{I} \gamma_{i,j} I_{n=d_i} + \sum_{l=1}^{L} \left(\delta_{l,j}\cos\frac{nl2\pi}{N} + \phi_{l,j}\sin\frac{nl2\pi}{N}\right) \end{array} \right] \tag{36}$$

It is the sum of three different components: a quadratic term, a set of dummies and a group of harmonics. All are multiplied by a scaling factor that involve volatility. In this paper, given the elevate number of harmonic considered J is set in all cases identically equal to 0, this imply no scaling factor. The seasonal function is evaluated and parameters are estimated with OLS, given a preliminar estimate of daily conditional volatility. Given the estimated seasonal function it is necessary to reconstruct the cyclical component of equation (33), a point solved using

$$\hat{s}_{t,n} = \frac{T \exp\left(\hat{f}_{t,n}/2\right)}{\sum_{t=1}^{T/N} \sum_{n=1}^{N} \exp\left(\hat{f}_{t,n}/2\right)} \tag{37}$$

Given this component the variance filtering is performed with

$$\tilde{x}_{t,n} = x_{t,n}/\hat{s}_{t,n} \tag{38}$$

For the purposes of our analysis we are also interested in a filtration of a cyclical component in the mean and in the variance, the whole model is then

$$y_{t,n} = x_{t,n} - \bar{s}_{t,n} = E\left[x_{t,n} - \bar{s}_{t,n}\right] + \frac{\sigma_{t(,n)} s_{t,n} Z_{t,n}}{\sqrt{N}} \tag{39}$$

where $\bar{s}_{t,n}$ follow directly (36) with $J = 0$. Then the log-transformation correspondent to (34) become

$$y_{t,n} = 2\ln\left[\left|y_{t,n} - E\left[y_{t,n}\right]\right|\right] - \ln\sigma^2_{t(,n)} + \ln N = \ln s^2_{t,n} + \ln Z^2_{t,n} \tag{40}$$

therefore, given the estimated parameters filtration is performed using

$$\tilde{x}_{t,n} = \left(x_{t,n} - \bar{s}_{t,n}\right)/\hat{s}_{t,n} \tag{41}$$

chosen frequencies 1-2-3-4-6-8-12-16, equivalent respectively to 1 day (8 hrs.), 4 hrs, 2 hrs and 40', 2 hrs, 1 hr, 40' and 30'.

5.2 Multivariate GARCH representations

In this appendix we summarize the mainly used formulation of multivariate GARCH and FIGARCH processes, showing, whenever possible, statistical properties, features and drawbacks.

Let we start with a bit of notation: we focus on an n-dimensional process X_t, defining by $I_t(X)$ the information set of X at time t, we assume the following hold

$$X_t | I_{t-1}(X) \, \tilde{} \, iid \; D\left(\boldsymbol{\mu}_t, H_t\right) \tag{42}$$

where $D(\mu_t, H_t)$ is a non-specified multivariate distribution with time dependent mean H_t and time dependent variance covariance matrix H_t. This formulation nest all multivariate GARCH representations that will be introduced in a few steps, and allow also ti specify a multivariate ARMA process for the mean, as well as the in-mean effects of the variances. In this appendix vector variables and matrices are denoted by bold or uppercase letters, while scalar are denoted by lowercase letters. We will also make use of the $Vech(\cdot)$ matrix operator that stacks the lower triangular portion of a matrix, and we will define the residuals of the mean as $E_t = \boldsymbol{\mu}_t - X_t$. In a general framework a multivariate GARCH process can be represented as

$$H_t = \boldsymbol{\omega} + C(L) H_t + D(L) [E_t E_t'] \tag{43}$$

$$C(L) = \sum_{i=1}^{p} C_i L^i \qquad D(L) = \sum_{i=1}^{q} D_i L^i \tag{44}$$

a relation that, even if useful for theoretical purposes, turn out to contain redundant relations, given the symmetry of the variance covariance matrix and of the cross product of the mean residuals. This observation, together with the need of formulations that couple flexibility, limited number of parameters and are also easy to handle with software packages (think of the restrictions needed to impose: positivity of conditional variances; invertibility of the variance-covariance matrix; stationarity of the process) leaded to a number of different specifications of the possible relation among variances in a multivariate framework. In the following we will present the Vech and BEKK representation of Engle and Kroner (1995), the CCC and DCC representations of Bollerslev (1992) and Engle and Sheppard (2001), their extension to allow for in-mean and long memory effects.

5.2.1 The Vech representations

These formulations derive from the work of Engle and Kroner (1995), the multivariate GARCH is represented as

$$Vech(H_t) = Vech(\boldsymbol{\omega}) + C(L) Vech(H_{t-1}) + D(L) Vech(E_t E_t') \tag{45}$$

where $\boldsymbol{\omega}$ is a positive definite matrix of dimension $n \times n$, $C(L) = \sum_{i=1}^{p} C_i L^i$, $D(L) = \sum_{i=1}^{q} D_i L^i$, C_i and D_i, are square matrices of dimension $r \times r$ and $r = n(n+1)/2$. The number of parameters in this formulation are $r \times [1 + r(p+q)]$ this is the main constraint on the estimation and application of this model, as an example consider the bivariate model, the parameters are here 21. However in this case Engle and Kroner (1995) show that the Vech-GARCH is stationary if and only if all the eigenvalues of $C(1) + D(1)$ are less than one in modulus. A restricted case of the Vech-GARCH is its diagonal representation, the D-Vech-GARCH: it is define diagonal as it assume that all parameter matrices are diagonal. This boil down to a model that parameterize all conditional variances and covariances as univariate GARCH(p,q) processes. The total number of parameters reduce to $3r$, for the bivariate case 9. A much greater problem that is not solved restricting to the diagonal version is the positive definiteness of the H_t matrix, very difficult to check, impose and control in the optimization routines.

5.2.2 The BEKK representation

This formulation was suggested by Baba, Engle, Kraft and Kroner in a preliminary version of Engle and Kroner (1995). The main feature is that we do not need to impose any restriction on parameters to get positive definiteness of the H_t matrix, given its quadratic structure. The fundamental BEKK-GARCH model is

$$H_t = \boldsymbol{\omega} + \sum_{i=1}^{p} C_i H_{t-i} C_i' + \sum_{j=1}^{q} D_j E_{t-j} E_{t-j}' D_j' \tag{46}$$

where C_i and D_j are $n \times n$ matrices and $\boldsymbol{\omega}$ is a symmetric positive definite $n \times x$ matrix. In a general formulation the number of parameters are here $r + n \times n \times (p+q)$, for the bivariate case the number drop to only 11. Positive definiteness of the variance covariance matrix is controlled by the constant matrix $\boldsymbol{\omega}$, whose positive definiteness is often obtained assuming this factorization $\boldsymbol{\omega} = \Gamma\Gamma'$ where Γ is a lower triangular matrix. The BEKK formulation and the Vech are strictly related as shown in Engle and Kroner (1995), in particular the stationarity condition of the BEKK model is very similar to the one of the Vech representation, we will not present it here to avoid the introduction of additional notation, if interested please refer to the cited paper.

5.2.3 The Constant Conditional Correlation GARCH

This form was introduced by Bollerslev (1990) trying to reduce the number of parameters of the Vech representations. He suggested the following:

$$\begin{aligned} H_t &= \begin{bmatrix} \sigma_{1,t}^2 & \sigma_{12,t} & \cdots & \sigma_{1n,t} \\ \sigma_{12,t} & \sigma_{2,t}^2 & & \sigma_{2n,t} \\ \vdots & & \ddots & \vdots \\ \sigma_{1n,t} & \sigma_{2n,t} & \cdots & \sigma_{n,t}^2 \end{bmatrix} \qquad (47) \\ \sigma_{i,t}^2 &= \omega_i + \sum_{j=1}^{p} \beta_{i,j} \sigma_{i,t-j}^2 + \sum_{j=1}^{q} \alpha_{i,j} \varepsilon_{i,t-j}^2 \qquad i = 1...n \\ \sigma_{ij,t} &= \rho_{ij} \sigma_{i,t} \sigma_{j,t} \qquad i,j = 1...n \ , \ i \neq j \end{aligned}$$

The main difference between this formulation and the previous is in the assumption of constant correlation among variables. The total number of parameters is not $(p+q+1)n + n(n+1)/2$, exactly 7 in the bivariate case. Positive definiteness of the variance covariance matrix is now controlled by the correlation matrix (for the conditional variances the usual GARCH constraints for positivity of conditional variances are required), since we can rewrite

$$H_t = diag\left(\sigma_{1,t}\sigma_{2,t}...\sigma_{n,t}\right)\begin{bmatrix} 1 & \rho_{1,2} & \cdots & \rho_{1,n} \\ \rho_{1,2} & 1 & & \vdots \\ \vdots & & \ddots & \rho_{1,n-1} \\ \rho_{1,n} & \cdots & \rho_{1,n-1} & 1 \end{bmatrix} diag\left(\sigma_{1,t}\sigma_{2,t}...\sigma_{n,t}\right) \tag{48}$$

where with $diag\left(\sigma_{1,t}\sigma_{2,t}...\sigma_{n,t}\right)$ we mean a diagonal matrix with the given elements, and the correlation matrix can be factorized similarly to the constant of the BEKK-GARCH to impose its positive definiteness.

5.2.4 The Dynamic Conditional Correlation formulations

This alternative representation try to add some limited dynamics to the CCC-GARCH and come from a recent paper of Engle and Sheppard (2001). The idea behind this GARCH model derive from the CCC representation, in that it assume the following factorization of the H_t matrix

$$H_t = diag\left(\sigma_{1,t}\sigma_{2,t}...\sigma_{n,t}\right) R_t diag\left(\sigma_{1,t}\sigma_{2,t}...\sigma_{n,t}\right) \tag{49}$$

where the conditional variances are parameterized as in the CCC-GARCH and R_t is a dynamic correlation matrix satisfying

$$\begin{aligned} R_t &= \left(Q_t^*\right)^{-1} Q_t \left(Q_t^*\right)^{-1} \\ Q_t &= \left[1-\alpha(1)-\beta(1)\right]+\alpha(L)\,\varepsilon_t\varepsilon_t'+\beta(L)\,Q_t \\ \varepsilon_t &= E_t' diag\left(\sigma_{1,t}\sigma_{2,t}...\sigma_{n,t}\right) \\ \alpha(L) &= \sum_{i=1}^{\overline{q}}\alpha_i L^i \qquad \beta(L)=\sum_{i=1}^{\overline{p}}\beta_i L^i \\ Q_t^* &= diag(\sqrt{q_{11,t}},\sqrt{q_{22,t}},...\sqrt{q_{nn,t}}) \end{aligned} \tag{50}$$

that is just a particular kind of multivariate GARCH on the correlations. Positive definiteness of the DCC-GARCH is controlled by the correlation function and depend on a set of restrictions, namely the same positivity restriction of the univariate GARCH(p,q), $\alpha(1)+\beta(1)<1$. Engle and Sheppard (2001) provide also proofs for consistency and asymptotic normality of a two step estimator based in a first stage on the estimation of the n univariate GARCH for the conditional variances, and in a second step on the estimation of the Dynamic correlation structure. In this model the number of parameters is $(p+q+1)\times n+(\overline{p}+\overline{q})$, and in the bivariate case are 8 if $\overline{p}=\overline{q}=1$ (the simplification introduced by the model are useful with larger models).

A natural extension to this model is to allow for a block structure on the dynamic correlation GARCH, in fact the structure proposed by Engle and Sheppard (2001) presume that all the correlation among the considered variables follow the same dynamics, this may not be the case, consider as an example a stock market, with the assets grouped in homogeneous categories (energy, food, chemistry...) we may assume different patterns of correlation inside the groups and between the groups. This consideration is on the basis of our extension of the DCC-GARCH to include this possibility. We therefore introduce the Block-DCC-GARCH reformulating the dynamic correlation equation in the following way:

$$\begin{aligned} Q_t &= \boldsymbol{\gamma} + \alpha(L)\,\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' + \beta(L)\,Q_t \qquad (51) \\ \boldsymbol{\gamma} &= \Gamma\Gamma' \\ \alpha(L) &= \sum_{i=1}^{\overline{q}} \boldsymbol{\alpha}_i L^i \qquad \beta(L) = \sum_{i=1}^{\overline{p}} \boldsymbol{\beta}_i L^i \end{aligned}$$

where $\boldsymbol{\gamma}$, $\boldsymbol{\alpha}_i$, $\boldsymbol{\beta}_i$ (full) and Γ (lower triangular) are square matrices of dimension $n \times n$ with the following structure: assume we group the n variables in w sets of dimension $m_1, m_2 ... m_w$, with $i(y)$ we indicate a column vector of dimension y filled with ones, then

$$\boldsymbol{\alpha}_i = \begin{bmatrix} \alpha_{i,11} i(m_1) i(m_1)' & \alpha_{i,12} i(m_1) i(m_2)' & \cdots & \alpha_{i,w1} i(m_1) i(m_w)' \\ \alpha_{i,12} i(m_2) i(m_1)' & \alpha_{i,22} i(m_2) i(m_2)' & & \alpha_{i,w2} i(m_2) i(m_w)' \\ \vdots & & \ddots & \vdots \\ \alpha_{i,w1} i(m_w) i(m_1)' & \alpha_{i,w2} i(m_w) i(m_2)' & \cdots & \alpha_{i,ww} i(m_w) i(m_w)' \end{bmatrix} \qquad (52)$$

and in similar way for $\boldsymbol{\beta}_i$ and $\boldsymbol{\gamma}$. Clearly this representation to be competitive require a small number of groups. In this new model the number of parameters are $(p+q+1) \times n + (\overline{p} + \overline{q} + 1) \times w$, this formulation is not applyable to the bivariate case.

5.2.5 Extension to In-Mean components

Any of the previous multivariate GARCH can be used to include in-mean effects, we specify this relation using (42) as

$$\mathbf{E}[X_t | I_{t-1}(X)] = \boldsymbol{\mu}_t = \Phi(L) X_t + \Theta(L) E_t + G Vech(H_t) \qquad (53)$$

where $\Phi(L) = \sum_{i=1}^{p} \Phi_i L^i$, $\Theta(L) = \sum_{i=0}^{q} \Theta_i L^i$, allowing for a general VARMA formulation, and G is a matrix of dimension $n \times n(n+1)/2$.

5.2.6 Extension to multivariate GARCH with long memory

In this case a group of models has been extended to consider FIGARCH effects. The main works are the one of Teyssière (1998 and 2000) and Brunetti and Gilbert (2000). They just admit that a FIGARCH structure can be used instead of the traditional GARCH in the CCC-GARCH (Brunetti and Gilbert) or in the Diagonal Vech-GARCH (Teyssière). As an example consider the CCC-GARCH, equation (47) is modified into

$$\begin{aligned}\sigma^2_{i,t} &= \omega_i + \sum_{j=1}^{p}\beta_{i,j}\sigma^2_{i,t-j} + \left[1-\sum_{j=1}^{p}\beta_{i,j}L^j - \left(1-\sum_{j=1}^{q}\alpha_{i,j}L^j\right)(1-L)^{d_i}\right]\varepsilon^2_{i,t} \qquad i=1...n \\ \sigma_{ij,t} &= \rho_{ij}\sigma_{i,t}\sigma_{j,t} \qquad i,j=1...n\ ,\ i\neq j\end{aligned}$$

where the constraints for stationarity are modified given the presence of the long memory coefficient, while positive definiteness is controlled as in the previous cases. The Vech-FIGARCH of Teyssière is represented as

$$Vech\,(H_t) = Vech\,(\boldsymbol{\omega}) + Vech\,(A\,(L)) \circledast Vech\,(E_t)$$

where $\circledast$ is the element by element matrix multiplication operator and the matrices $A\,(L)$ and $\boldsymbol{\omega}$ have generic element in the form

$$A_{i,j}\,(L) = \frac{\left[1-\left(1-\sum_{l=1}^{q}\alpha_{ij,l}L^l\right)(1-L)^{d_i}\right]}{1-\sum_{m=1}^{p}\beta_{ij,m}L^m}$$

and

$$\omega_{i,j} = \frac{\tilde{\omega}_{ij}}{1-\sum_{m=1}^{p}\beta_{ij,m}}$$

therefore a FIGARCH structure can be specified for all conditional variances and covariances.

5.2.7 Brunetti-Gilbert bivariate FIGARCH(1,d,1)

We consider this model as a special case, it is a kind of CCC-FIGARCH with a slight modification, to introduce causality among variances. This has been done assuming that a bivariate formulation hold for the two conditional variances, represented as:

$$(\mathbf{I}-\Phi L)\begin{bmatrix}(1-L)^{d_1} & 0 \\ 0 & (1-L)^{d_2}\end{bmatrix}\begin{bmatrix}\varepsilon^2_{1,t} \\ \varepsilon^2_{2,t}\end{bmatrix} = \boldsymbol{\omega} + (\mathbf{I}-\mathbf{B}L)\begin{bmatrix}\varepsilon^2_{1,t}-\sigma^2_{1,t} \\ \varepsilon^2_{2,t}-\sigma^2_{2,t}\end{bmatrix} \tag{54}$$

where the matrices Φ and $\mathbf{B}$ are full and not necessarily symmetric. In this case Brunetti and Gilbert (??) gave a set of constraints that have to be satisfied for the positivity condition. We will now identify this constraints. Expand formula (54) as follows

$$\begin{aligned}\begin{bmatrix}\sigma^2_{1,t} \\ \sigma^2_{2,t}\end{bmatrix} &= \begin{bmatrix}\omega_1 \\ \omega_2\end{bmatrix} + \begin{bmatrix}b_{11} & b_{12} \\ b_{21} & b_{22}\end{bmatrix}\begin{bmatrix}\sigma^2_{1,t-1} \\ \sigma^2_{2,t-1}\end{bmatrix} + \qquad (55) \\ &\left(\begin{matrix}\begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix} - \begin{bmatrix}b_{11} & b_{12} \\ b_{21} & b_{22}\end{bmatrix}L - \\ \left(\begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix} - \begin{bmatrix}\phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22}\end{bmatrix}L\right)\begin{bmatrix}(1-L)^{d_1} & 0 \\ 0 & (1-L)^{d_2}\end{bmatrix}\end{matrix}\right)\begin{bmatrix}\varepsilon^2_{1,t} \\ \varepsilon^2_{2,t}\end{bmatrix}\end{aligned}$$

then we rearrange to derive the multivariate ARCH(∞) representation.

$$\begin{aligned}
\left[\begin{array}{c}\sigma_{1,t}^2\\ \sigma_{2,t}^2\end{array}\right] &= \left[\begin{array}{c}\omega_1\\ \omega_2\end{array}\right]\left(\left[\begin{array}{cc}1&0\\0&1\end{array}\right]-\left[\begin{array}{cc}b_{11}&b_{12}\\b_{21}&b_{22}\end{array}\right]L\right)^{-1}+\\
&\left(\begin{array}{c}\left[\begin{array}{cc}1&0\\0&1\end{array}\right]-\left(\left[\begin{array}{cc}1&0\\0&1\end{array}\right]-\left[\begin{array}{cc}b_{11}&b_{12}\\b_{21}&b_{22}\end{array}\right]L\right)^{-1}\\
\left(\left[\begin{array}{cc}1&0\\0&1\end{array}\right]-\left[\begin{array}{cc}\phi_{11}&\phi_{12}\\\phi_{21}&\phi_{22}\end{array}\right]L\right)\left[\begin{array}{ll}(1-L)^{d_1}&0\\0&(1-L)^{d_2}\end{array}\right]\end{array}\right)\left[\begin{array}{c}\varepsilon_{1,t}^2\\ \varepsilon_{2,t}^2\end{array}\right]
\end{aligned}$$

and more compactly

$$\left[\begin{array}{c}\sigma_{1,t}^2\\ \sigma_{2,t}^2\end{array}\right]=\boldsymbol{\omega}\left(\mathbf{I}-\mathbf{B}L\right)^{-1}+\left\{\mathbf{I}-\left(\mathbf{I}-\mathbf{B}L\right)^{-1}\left(\mathbf{I}-\Phi L\right)\left[\begin{array}{ll}(1-L)^{d_1}&0\\0&(1-L)^{d_2}\end{array}\right]\right\}\left[\begin{array}{c}\varepsilon_{1,t}^2\\ \varepsilon_{2,t}^2\end{array}\right] \tag{56}$$

A sufficient condition for positivity of conditional variances is the positivity of all coefficients in this ARCH(∞) formulation. The coefficient can be recursively computed as follows: consider at a first stage the following polynomial

$$\begin{aligned}
\mathbf{H}(L) &= (\mathbf{I}-\Phi L)\left[\begin{array}{ll}(1-L)^{d_1}&0\\0&(1-L)^{d_2}\end{array}\right]=\sum_{j=0}^{\infty}\mathbf{h}_jL^j \qquad (57)\\
\mathbf{h}_0 &= \mathbf{I}\\
\mathbf{h}_j &= \left[\begin{array}{cc}\pi_{1,j}-\phi_{11}\pi_{1,j-1} & -\phi_{12}\pi_{2,j-1}\\ -\phi_{21}\pi_{1,j-1} & \pi_{2,j}-\phi_{22}\pi_{2,j-1}\end{array}\right]\\
(1-L)^{d_i} &= \sum_{k=0}^{\infty}\pi_{k,i}L^k
\end{aligned}$$

adding then the remainder

$$\begin{aligned}
\left(\mathbf{I}-\mathbf{B}L\right)^{-1}\mathbf{H}(L) &= \mathbf{W}(L)=\sum_{j=0}^{\infty}\mathbf{w}_jL^j \qquad (58)\\
\mathbf{H}(L) &= \mathbf{W}(L)\left(\mathbf{I}-\mathbf{B}L\right)
\end{aligned}$$

and expanding and solving in the matrices $\mathbf{w}_j$ we get a recursive formula

$$\begin{aligned}
\mathbf{w}_0 &= \mathbf{I} \qquad (59)\\
\mathbf{w}_j &= \mathbf{h}_j+\mathbf{B}\mathbf{w}_{j-1}
\end{aligned}$$

Given the structure of the model to ensure positivity in both conditional variances I have to impose positivity restrictions on the matrices of coefficients of

$\mathbf{I} - \mathbf{W}(L)$ this is just for the first and second lag:

$$-\mathbf{w}_1 = -\begin{bmatrix} \pi_{1,1} - \phi_{11} & -\phi_{12} \\ -\phi_{21} & \pi_{2,1} - \phi_{22} \end{bmatrix} - \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \tag{60}$$

$$\begin{aligned} -\mathbf{w}_2 &= -\begin{bmatrix} \pi_{1,2} - \phi_{11}\pi_{1,1} & -\phi_{12}\pi_{2,1} \\ -\phi_{21}\pi_{1,1} & \pi_{2,2} - \phi_{22}\pi_{2,1} \end{bmatrix} - \\ &\quad + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \pi_{1,1} - \phi_{11} + b_{11} & -\phi_{12} + b_{12} \\ -\phi_{21} + b_{21} & \pi_{2,1} - \phi_{22} + b_{22} \end{bmatrix} \end{aligned} \tag{61}$$

From the first lag coefficients we can derive a first set of constraints, namely

$$\begin{aligned} -\pi_{1,1} + \phi_{11} - b_{11} &\geq 0 \\ -\pi_{2,1} + \phi_{22} - b_{22} &\geq 0 \\ \phi_{21} - b_{21} &\geq 0 \\ \phi_{12} - b_{12} &\geq 0 \end{aligned} \tag{62}$$

exactly the same presented by Brunetti and Gilbert. However from the second lag we can derive the following additional restrictions

$$\begin{aligned} -\pi_{1,2} + \phi_{11}\pi_{1,1} + b_{11}(-\pi_{1,1} + \phi_{11} - b_{11}) + b_{12}(\phi_{21} - b_{21}) &\geq 0 \\ -\pi_{2,2} + \phi_{22}\pi_{2,1} + b_{22}(-\pi_{2,1} + \phi_{22} - b_{22}) + b_{21}(\phi_{12} - b_{12}) &\geq 0 \\ \phi_{12}\pi_{2,1} + b_{11}(\phi_{12} - b_{12}) + b_{12}(-\pi_{2,1} + \phi_{22} - b_{22}) &\geq 0 \\ \phi_{21}\pi_{1,1} + b_{22}(\phi_{21} - b_{21}) + b_{21}(-\pi_{1,1} + \phi_{11} - b_{11}) &\geq 0 \end{aligned} \tag{63}$$

that together considered with (62) imply $0 \geq \phi_{12} \geq b_{12} \geq 0$ and $0 \geq \phi_{21} \geq b_{21} \geq 0$. Additional lags are not investigated since, given the restrictions implied by first and second lag coefficients we just obtain the restrictions given in the simple univariate FIGARCH case. Therefore the model collapse on the CCC-FIGARCH if we impose all positivity constraints and is useless in testing for causality.

5.3 Multivariate GARCH estimates

The following tables report the estimates of a group of alternative models with extensions for causality. In all cases we report in the first row the GARCH specification adopted for returns and volumes, while the in-mean specification is constant and include a VAR(1) and a long memory component for the volumes mean. The parameters for the mean are reported in the following order (referring to the bivariate benchmark model): μ_R return mean; μ_V volume mean; $d_{1,2}$ volume long memory coefficient; VAR parameters, in the order $\phi_{1,1}$, $\phi_{1,2}$, $\phi_{2,1}$, $\phi_{2,2}$. For the variance component we distinguish the specifications for Returns and Volumes, however in all cases, independently from the series, the parameters follow this order: ω_i, $d_{2,i}$, α_i, β_i, ψ_i. Please note that the order is maintained even if some parameter does not belong to the estimated specification. After the GARCH part we report the causality parameters, first for the return equation, and then for the volume equation. Finally the correlation is reported. For all the parameters the first column report the estimated coefficient, the second the standard error. Recall that these are heteroskedastic consistent quasi maximum likelihood standard errors. After the parameters section, all the tables report the log-likelihood value, four different information criteria and the Box-Pierce residual autocorrelation tests values for a group of lags. This last test is computed both on standardized residuals than on squared standardized residuals, and both for the returns and the volume.