C O N T R I B U T I   D I   R I C E R C A   C R E N O S

CLUSTERING HETEROSKEDASTIC TIME SERIES BY MODEL-BASED PROCEDURES

Edoardo Otranto

WORKING PAPERS

2008/01

CENTRO RICERCHE ECONOMICHE NORD SUD
(CRENoS)
UNIVERSITÀ DI CAGLIARI
UNIVERSITÀ DI SASSARI

Il CRENoS è un centro di ricerca istituito nel 1993 che fa capo alle Università di Cagliari e Sassari ed è attualmente diretto da Raffaele Paci. Il CRENoS si propone di contribuire a migliorare le conoscenze sul divario economico tra aree integrate e di fornire utili indicazioni di intervento. Particolare attenzione è dedicata al ruolo svolto dalle istituzioni, dal progresso tecnologico e dalla diffusione dell'innovazione nel processo di convergenza o divergenza tra aree economiche. Il CRENoS si propone inoltre di studiare la compatibilità fra tali processi e la salvaguardia delle risorse ambientali, sia globali sia locali.
Per svolgere la sua attività di ricerca, il CRENoS collabora con centri di ricerca e università nazionali ed internazionali; è attivo nell'organizzare conferenze ad alto contenuto scientifico, seminari e altre attività di natura formativa; tiene aggiornate una serie di banche dati e ha una sua collana di pubblicazioni.

www.crenos.it
info@crenos.it

CRENoS – CAGLIARI
VIA SAN GIORGIO 12, I-09100 CAGLIARI, ITALIA
TEL. +39-070-6756406; FAX +39-070- 6756402

CRENoS - SASSARI
VIA TORRE TONDA 34, I-07100 SASSARI, ITALIA
TEL. +39-079-2017301; FAX +39-079-2017312

Titolo: CLUSTERING HETEROSKEDASTIC TIME SERIES BY MODEL-BASED PROCEDURES

Clustering Heteroskedastic Time Series by Model-Based Procedures

Edoardo Otranto
Dipartimento di Economia, Impresa e Regolamentazione,
University of Sassari and CRENoS, Via Torre Tonda 34, 07100 Sassari, Italy,
E-mail: eotranto@uniss.it; Tel. +39 079 2017317; Fax +39 079 2017312

Abstract

Financial time series are often characterized by similar volatility structures, often represented by GARCH processes. The detection of clusters of series displaying similar behavior could be important to understand the differences in the estimated processes, without having to study and compare the estimated parameters across all the series. This is particularly relevant dealing with many series, as in financial applications. The volatility of a time series can be characterized in terms of the underlying GARCH process. Using Wald tests and the AR metrics to measure the distance between GARCH processes, it is possible to develop a clustering algorithm, which can provide three classifications (with increasing degree of deepness) based on the heteroskedastic patterns of the time series. The number of clusters is detected automatically and it is not fixed a priori or a posteriori. The procedure is evaluated by simulations and applied to the sector indexes of the Italian market

Key Words: Agglomerative algorithm, AR metrics, Cluster analysis, GARCH models, Wald test
JEL: C02, C19, C22

1 Introduction

The topic of classification of time series has recently received a lot of contributions, in particular in time series data mining (see, for example, Agrawal et al., 1994), computer science (Gray and Markel, 1976), economic time series (Caiado et al., 2006). Liao (2005) provides an extensive review of studies in clustering and discrimination of time series. In particular, he distinguishes three major categories of approaches to time series clustering: 1) *raw-data-based approaches*, in which the series compared are considered as normally sampled at the same interval; 2) *features-based approaches*, in which the series are compared using some selected features; 3) *model-based methods*, where the time series are considered similar when the models characterizing them are similar.

The approach proposed in this work belongs to the third category; in particular it follows the tradition of AR processes to capture the similarity among time series, as in Piccolo (1990), Maharaj (1996, 1999, 2000), Xiong and Yeung (2002) (see Piccolo, 2007, and Corduas and Piccolo, 2008, for a review). Most of these studies are devoted to capturing the structure of the mean of the process hypothesized as generator of the data, whereas little attention was put on the variance. This is a correct approach when dealing with classifications based on ARMA models and in presence of homoskedastic variance (for example, the clustering methods based on the AR metrics proposed by Piccolo, 1990); in fact, in this case, the variance is a function of the process parameters, so that it is implicitly considered in the classification. Dealing with heteroskedastic time series, in which the (conditional) variance follows a stochastic process (typically a GARCH process; Engle, 1982, Bollerslev, 1986), the comparison of the dynamics of the variances is fundamental. This is particularly important if dealing with financial time series, when the investor has a very large investment universe (hundreds of stocks) and s/he would like to have groups of series with similar characteristics (similar unconditional variance, similar dynamics, etc.). Moreover, the volatility of a return is generally considered as a proxy of the risk of the same return; in other words the classification of returns of several assets is equivalent to classifying the assets in clusters with similar risk. Furthermore movements in a time series could help to forecast the movements of a similar time series. In fact, the financial time series are generally subject to co-movements and similar volatility structures, due to the strong reciprocal influence among financial markets (see, for example, Bollerslev et al., 1994) and the increasing integration among markets (Gallo and Otranto, 2007b). Generally, turmoil periods are transmitted from a market to another. The classification of financial time series in homogeneous clusters for similar volatility structures could be an important objective for the financial analysts. Such a result could be useful, for example, in the analysis of the spillover effects among markets, where the shocks affecting a market can influence the behavior of another market (see, for example, Gallo and Otranto, 2007a and 2007b).

In this paper we propose a clustering procedure based on simple statistical tools. In particular, we consider the squared disturbances of the returns of a financial time series as the volatility of the series. Then, we use the GARCH representation of the conditional variance to derive the model underlying the squared disturbances. From this model we can separate the volatility in a constant part and a time-varying part; this subdivision can have an appealing interpretation, in particular when we use the volatility to represent the risk of the asset. The constant part of the volatility is measured in a natural way, whereas we measure the time-varying part extending the idea of distance between AR models (Piccolo, 1990) to the GARCH family. The constant part could be interpreted as the lower bound of volatility (minimum expected risk) of a certain series, whereas the unconditional volatility represents the expected risk and seems more interesting for the investor. We classify the series with similar unconditional volatility and similar time-varying volatility using classical Wald statistics. The use of the results of a statistical test (the p-value) to classify time series was successfully performed by Maharaj (1996, 1999, 2000), dealing with AR models. Her approach is different from the one proposed here because she calculates the p-value for every pair of series and uses these results in an algorithm which follows the principles of hierarchical clustering. In our approach we consider a starting benchmark series and then apply an agglomerative algorithm. In particular we can obtain three different levels of clustering, depending on the degree of deepness chosen. An interesting characteristic of this procedure is that, unlike the main agglomerative algorithms, the number of clusters is detected automatically and it is not determined by the user.

In the next section we deal with the statistical tools we need to develop the clustering algorithm, which will be described in section 3; this algorithm is based on a sequence of statistical tests, described in section 4. In section 5 we show some simulations studies, using various GARCH data generating processes, to evaluate the performance of the clustering algorithm, whereas in section 6 we apply the procedure to classify the sectors of the Italian financial market. Final remarks will conclude the paper.

2 Some Statistical Tools

Let us consider a time series y_t; we suppose that it is the sum of a constant term and a heteroskedastic disturbance:

$$\begin{array}{l} y_t = \mu + \varepsilon_t, \qquad t = 1, ..., T \\ \text{and} \\ \varepsilon_t = h_t^{1/2} u_t \end{array} \tag{1}$$

where u_t are i.i.d. Normal disturbances with mean zero and variance one. The conditional variance h_t follows a GARCH(p,q) process (Bollerslev, 1986), as:

$$h_t = \gamma + \alpha_1 \varepsilon_{t-1}^2 ... + \alpha_p \varepsilon_{t-p}^2 + \beta_1 h_{t-1} + ... + \beta_q h_{t-q}, \tag{2}$$

with $\gamma > 0, 0 \leq \alpha_i < 1, 0 \leq \beta_j < 1$ $(i = 1, ...p, j = 1, ..., q)$, $\left(\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j\right) < 1$.

In this section we will characterize the volatility of a time series as (1) in terms of squared disturbances, describing some statistical tools to develop a clustering procedure.

2.1 Unconditional, Minimum and Time-varying Volatilities

The volatility of a financial time series is not observable and its study is generally made using a proxy, such as the squared returns, the intra-daily range, the realised volatility. In the framework described by equation (1) the disturbances ε_t (and the observed time series y_t) have the heteroskedastic behavior represented by the GARCH process (2). Then the squared disturbances ε_t^2 can represent the unobserved volatility of y_t. In our approach we do not need to create a proxy of the volatility, but only to derive a statistical model for it.

From equation (2), after simple algebra, we can represent the time series ε_t^2 by an ARMA(p^*, q) model, with $p^* = max(p, q)$:

$$\varepsilon_t^2 = \gamma + \sum_{i=1}^{p^*}(\alpha_i + \beta_i)\varepsilon_{t-i}^2 - \sum_{j=1}^{q}\beta_j(\varepsilon_{t-j}^2 - h_{t-j}) + (\varepsilon_t^2 - h_t), \tag{3}$$

where $(\varepsilon_t^2 - h_t)$ are mean zero errors, uncorrelated with past information, $\alpha_i = 0$ for $i > p$ if $p^* = q$ and $\beta_i = 0$ for $i > q$ if $p^* = p$. Substituting recursively in (3) the errors with their ARMA(p^*, q) expression, we obtain the AR(∞) representation:

$$\varepsilon_t^2 = \frac{\gamma}{1 - \sum_{j=1}^{q}\beta_j} + \sum_{k=1}^{\infty}\pi_k\varepsilon_{t-k}^2 + \left(\varepsilon_t^2 - h_t\right). \tag{4}$$

From the ARMA expression (3) it is easy to obtain the AR coefficients π_k. As it is well known, indicating with ϕ_i the generic AR coefficient and θ_j the generic MA coefficient of an ARMA(p,q) model, the recursive formula (see, for example, Brockwell and Davis, 1996):

$$\pi_k - \sum_{j=1}^{q}\theta_j\pi_{k-j} = \phi_k, \quad k = 0, 1, ..., \tag{5}$$

provides the sequence of the coefficients π_k. In (5) $\phi_0 = 1$, $\phi_i = 0$ for $i > p$ and $\pi_k = 0$ for $k < 0$. From (3), the previous relationship is equivalent to:

$$\pi_k = (\alpha_k + \beta_k) - \sum_{j=1}^{q}\beta_j\pi_{k-j}. \tag{6}$$

From (4) the expected volatility at time $t+1$, given the information available at time t, is given by:

$$E_t(\varepsilon_{t+1}^2) = \frac{\gamma}{1 - \sum_{j=1}^{q}\beta_j} + \sum_{k=1}^{\infty}\pi_k\varepsilon_{t-k}^2. \tag{7}$$

Let us note that the expected volatility $E_t(\varepsilon_{t+1}^2)$ can be split in two positive parts: a constant part $\gamma/(1 - \beta_1 - ... - \beta_q)$, which can represent the *minimum expected volatility* of the return; a *time-varying* part ($\sum_{k=1}^{\infty}\pi_k\varepsilon_{t-k}^2$), which depends on the past history of the volatility and can vary along the time. Dealing with financial time series, it is likely that returns referred to shares with high risk or funds with a large risky component will show a high constant part. Anyway, the investor is interested in a general measure of the risk and not in its lower bound; this measure is given by the unconditional expected value of ε_{t+1}^2 (*unconditional volatility*), given by:

$$E(\varepsilon_{t+1}^2) = \frac{\gamma}{(1 - \sum_{j=1}^{q} \beta_j)(1 - \sum_{k=1}^{\infty} \pi_k)}. \tag{8}$$

Series with similar unconditional volatility can have different dynamics, characterized by different time-varying volatilities.

2.2 Distance between GARCH Processes

Piccolo (1990) proposed a metrics measuring the distance between two ARMA models, based on the comparison of the coefficients of their AR(∞) representation. This tool has had a large success in several fields; a review of its properties and applications with several references can be found in Piccolo (2007) and Corduas and Piccolo (2008).

Indicating with π_{1j} and π_{2j} the coefficients at lag j of the two models, the general form of this metrics is given by:

$$d = \left[\sum_{k=1}^{\infty} (\pi_{1k} - \pi_{2k})^2 \right]^{1/2}. \tag{9}$$

This metrics requires only that the two ARMA processes are invertible. Following Otranto (2004), this metrics can be applied to the AR(∞) structures of the squared disturbances, expressed in (4); the invertibility is assured by the constraints on the coefficients of the GARCH model. A higher distance implies a more different dynamics between the two series; a distance equal to zero implies that the two series follow the same dynamics. It is important to note that a distance equal to zero does not mean that the series have the same volatility, but only that the time-varying part of the volatility is the same.

2.3 Equal Volatility Structures

If two series have the same unconditional (and/or minimum) volatility and the same time-varying volatility, it is not possible to conclude that they are generated by the same GARCH process. In fact the equality of the two unconditional volatilities and the two time-varying volatilities is relative to two nonlinear combinations of the coefficients. The case of equality of the two data generating processes (which we define *equal volatility structure*) is obtained if and only if the constant γ and the coefficients α_i and β_j ($i = 1, ...p$; $j = 1, ..., q$) are the same for the two GARCH models. This is a stronger relationship with respect to the equality of the unconditional volatilities and the time-varying volatilities. It is obvious that equal volatility structure of s time series implies equal unconditional, minimum and time-varying volatilities (but not vice versa).

3 The Clustering Algorithm

The statistical tools illustrated in the previous section can help in the development of an algorithm which creates clusters of series with homogeneous volatility. In particular, the definition of unconditional and time-varying volatility could suggest two levels of clustering: one based only on the unconditional volatility and the other based on both kinds of volatilities. Finally, a more accurate classification can be obtained distinguishing, within the groups with equal unconditional and time-varying volatilities, the series with equal volatility structure.

In sum, we can obtain three levels of clustering. Deeper clustering could imply small groups and a large number of clusters. The steps of the clustering algorithm can be synthesized in the following way:

1. First Level (clusters with equal unconditional volatility):
   (a) order the series in terms of increasing unconditional volatility and choose an initial benchmark;
   (b) insert in the same cluster the series with unconditional volatility not significantly different from the benchmark volatility; the series with minimum unconditional volatility which does not enter in this cluster is considered the benchmark for the successive cluster;
   (c) go on until no series remain;
2. Second Level (clusters with equal unconditional volatility and equal time-varying volatility):
   (a) in each first level cluster, order the series by increasing time-varying volatility; the series with minimum time-varying volatility of each cluster is the benchmark of the cluster;
   (b) form sub-clusters with series having equal time-varying volatility with respect to the benchmark; the series with minimum time-varying volatility which does not enter in a sub-cluster is considered the benchmark for the successive sub-cluster.
   (c) go on until no series remain;
3. Third Level (clusters with equal volatility structure)
   (a) in each second level sub-cluster verify the equal volatility structure (by some test) for each pair of series;
   (b) if all the p-values are less than the nominal size of the test, the series considered have different volatility structures; otherwise, select the pair of series with maximum p-value and repeat the test of equal volatility structure adding a series to the pair selected;

(c) select the group of series with maximum p-value and go on until the hypothesis of equal volatility structure is rejected;

(d) repeat steps 3.(b) and 3.(c) for the remaining series;

(e) go on until no series remain.

Note that, in the first two levels, the clusters have a natural order, with the series with smallest unconditional volatility (first level clustering) or the series with smallest unconditional and time-varying volatilities (second level clustering) belonging to the first group and the series with highest unconditional (and time-varying) volatility to the latest group.

To provide this classification we need to define an initial benchmark series. In step 1.(a) the choice is rather natural if we define an hypothetical series with null unconditional volatility. This case is verified when the series y_t is constant (in (1) the stochastic part is absent).

In step 2.(a) we have to establish what is the minimum time-varying volatility. This is a less obvious idea because the time-varying volatility is an infinite weighted sum of unobserved random variables (see equation (7)). In this case we can use the relationship between the distance (9) and the time-varying volatility, described in section 2.2. In particular, a null time-varying volatility is obtained when each π coefficient is equal to zero. From (6) it is easy to deduce that this is obtained when $\alpha_i = 0$ for each $i = 1, ..., p$. If we do not impose $\beta_j = 0$ for $j = 1, ..., q$, the conditional variance h_t of the process (2) is constant from a certain t; so, having to choose a benchmark, we impose also $\beta_j = 0$ for every j, obtaining a constant conditional variance for each t. In other words, we characterize the case of null time-varying volatility by a series with constant conditional variance. In this way, in step 4 of the clustering algorithm, the arrangement of the series in terms of increasing time-varying volatility is obtained calculating the distance of each series r from the case of constant volatility:

$$\left[\sum_{k=1}^{\infty} \pi_{rk}^2\right]^{1/2}. \tag{10}$$

In the third level clustering we verify if the series belonging to the same cluster can be considered as generated by the same data generating process. For this case the order is based on a p-value approach; its use in a clustering algorithm is justified because the p-value is a measure of similarity and satisfies properties of a semi-metric (see Maharaj, 1999).

In the next section we describe the tests to apply the clustering procedure.

4 Wald Statistics to Perform the Clustering Procedure

The algorithm proposed to cluster the series in groups with similar volatility requires verifying what series have similar unconditional volatility, what series have similar time-varying volatility and what series have the same volatility structure. We can obtain this result verifying a set of constraints on the GARCH parameters. This section will illustrate the test to be used for each clustering level.

Let us consider n time series, each one following the model (1)-(2). Three points have to be emphasized. First, the series can follow different dynamics in (1), for example they can have an ARMA expression or can depend on some observed variables; what we need is only that the conditional variance of ε_t follows a GARCH(p,q) process. Second, the orders p and q of the GARCH processes are not necessarily the same for each time series. Third, the results are valid also in multivariate modeling, assuming a Dynamic Conditional Correlation structure as in Engle (2002); in fact, in this case the estimation of the volatility part of this model can be performed estimating n univariate GARCH(p,q) models (in the final section we will comment more in detail this point).

Let us call $\widehat{\gamma}_r$, $\widehat{\alpha}_{r,i}$, $\widehat{\beta}_{r,j}$ the maximum likelihood estimates of the parameters of the n GARCH models; the index r is referred to the time series ($r = 1, ..., n$), whereas $i = 1, ..., p_r$, $j = 1, ..., q_r$, where p_r and q_r are the orders of the GARCH process underlying the $r-th$ series.

4.1 Testing the Hypothesis of Equal Unconditional Volatility

From (8) the estimation of the unconditional volatility can be expressed as:

$$\frac{\widehat{\gamma}_r}{(1-\sum_{j=1}^{q_r}\widehat{\beta}_{r,j})(1-\sum_{k=1}^{\infty}\widehat{\pi}_k)}; \tag{11}$$

s series (let us indicate them with $1, 2, ..., s$) have the same unconditional volatility if the following hypothesis is verified:

$$\begin{aligned} &H_0 : \frac{\gamma_1}{(1-\sum_{j=1}^{q_1}\beta_{1,j})(1-\sum_{k=1}^{\infty}\pi_{1,k})} = \frac{\gamma_2}{(1-\sum_{j=1}^{q_2}\beta_{2,j})(1-\sum_{k=1}^{\infty}\pi_{2,k})} = ... \\ &... = \frac{\gamma_s}{(1-\sum_{j=1}^{q_s}\beta_{s,j})(1-\sum_{k=1}^{\infty}\pi_{s,k})}. \end{aligned} \tag{12}$$

A simple test to verify this hypothesis is the classical Wald test. Let

$$\widehat{\boldsymbol{\delta}} = \left[\frac{\widehat{\gamma}_1}{(1-\sum_{j=1}^{q_1}\widehat{\beta}_{1,j})(1-\sum_{k=1}^{\infty}\widehat{\pi}_{1,k})}, ..., \frac{\widehat{\gamma}_s}{(1-\sum_{j=1}^{q_s}\widehat{\beta}_{s,j})(1-\sum_{k=1}^{\infty}\widehat{\pi}_{s,k})}\right]'$$

the vector of maximum likelihood estimators of the s unconditional volatilities; for the invariance property it is asymptotically Normally distributed with mean:

$$\boldsymbol{\delta} = \left[\frac{\gamma_1}{(1-\sum_{j=1}^{q_1}\beta_{1,j})(1-\sum_{k=1}^{\infty}\pi_{1,k})}, ..., \frac{\gamma_s}{(1-\sum_{j=1}^{q_s}\beta_{s,j})(1-\sum_{k=1}^{\infty}\pi_{s,k})}\right]'$$

and covariance matrix given by $\boldsymbol{G}\widehat{\boldsymbol{\Lambda}}\boldsymbol{G}'$, where $\widehat{\boldsymbol{\Lambda}}$ is a block diagonal matrix with blocks constituted by the covariance matrices $\widehat{\boldsymbol{\Lambda}}_r$ ($r = 1, ...s$), and $\boldsymbol{G}$ is the matrix of derivatives of $\widehat{\gamma}_r/(1-\sum_{j=1}^{q_r}\widehat{\beta}_{r,j})(1-\sum_{k=1}^{\infty}\widehat{\pi}_{r,k})$, for each r, with respect to

$$\boldsymbol{\theta} = (\gamma_1, \alpha_{1,1}, ..., \alpha_{1,p_1}, \beta_{1,1}, ..., \beta_{1,q_1}, ..., \gamma_s, \alpha_{s,1}, ..., \alpha_{s,p_s}, \beta_{s,1}, ..., \beta_{s,q_s})'.$$

The null hypothesis can be rewritten as:

$$\boldsymbol{A}\boldsymbol{\delta} = \boldsymbol{0}, \tag{13}$$

where $\boldsymbol{A}$ is the $(s-1)\times s$ matrix:

$$\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix}.$$

The Wald test statistic for the null hypothesis (13) is given by:

$$W_A = (\boldsymbol{A}\widehat{\boldsymbol{\delta}})'(\boldsymbol{A}\boldsymbol{G}\widehat{\boldsymbol{\Lambda}}\boldsymbol{G}'\boldsymbol{A}')^{-1}(\boldsymbol{A}\widehat{\boldsymbol{\delta}}). \tag{14}$$

W_A is asymptotically distributed as a central chi-square random variable with $(s-1)$ degrees of freedom.

Recalling the clustering algorithm described in section 3, the first benchmark in step 1.(a) is given by an hypothetical series with null unconditional volatility. Labeling with 1 the series with smallest unconditional volatility, the first hypothesis to be verified is:

$$H_0 : \frac{\gamma_1}{(1-\sum_{j=1}^{q_1}\beta_{1,j})(1-\sum_{k=1}^{\infty}\pi_{1,k})} = 0. \tag{15}$$

In this case, in (14) we put $\widehat{\boldsymbol{\delta}} = \widehat{\gamma}_1/(1-\sum_{j=1}^{q_1}\widehat{\beta}_{1,j})(1-\sum_{k=1}^{\infty}\widehat{\pi}_{1,k})$, $\widehat{\boldsymbol{\Lambda}} = \widehat{\boldsymbol{\Lambda}}_1$ and $\boldsymbol{A} = 1$. If the null hypothesis is not rejected, the series belonging to the first cluster have null unconditional volatility.

4.2 Testing the Hypothesis of Equal Time-varying Volatility

Concerning the hypothesis of equal time-varying volatility, we can use the interpretation of the distance given in section 2.2 and, in particular, expression (10) as a measure of the time-varying volatility. Similarly to the case of equal unconditional volatility, s series have the same time-varying volatility if the following null hypothesis is verified:

$$\left[\sum_{k=1}^{\infty}\pi_{1,k}^2\right]^{1/2} = \left[\sum_{k=1}^{\infty}\pi_{2,k}^2\right]^{1/2} = \ldots = \left[\sum_{k=1}^{\infty}\pi_{s,k}^2\right]^{1/2}. \tag{16}$$

The hypothesis (16) can be verified using the Wald test statistic:

$$W_B = (\boldsymbol{A}\widehat{\boldsymbol{\eta}})'(\boldsymbol{A}\boldsymbol{V}\widehat{\boldsymbol{\Lambda}}\boldsymbol{V}'\boldsymbol{A}')^{-1}(\boldsymbol{A}\widehat{\boldsymbol{\eta}}), \tag{17}$$

where $\widehat{\boldsymbol{\eta}}$ is the vector containing the s maximum likelihood estimates of the time varying volatilities, $\boldsymbol{V}$ is the matrix of the derivatives of $\left[\sum_{k=1}^{\infty}\widehat{\pi}_{r,k}^2\right]^{1/2}$ with respect to $\boldsymbol{\theta}$. W_B is asymptotically distributed as a central chi-square random variable with $(s-1)$ degrees of freedom. All the considerations made in the previous sub-section about the invariance of the estimators and the first hypothesis of null volatility can be extended to the case of time-varying volatility.

4.3 Testing the Hypothesis of Equal Volatility Structure

As said in section 2.3, if two series have the same unconditional volatility and the same time-varying volatility, it is not possible to conclude that they have the same volatility structure, because we have tested separately two hypotheses concerning the equality of nonlinear combinations of the parameters. The test of equality of the volatility structure was developed by Otranto and Triacca (2007), who extend the idea of equivalence among ARMA processes, developed by Steece and Wood (1985), to the GARCH case. In this framework we want to test the hypotheses:

$$H_0: \begin{array}{l} \gamma_1 = \gamma_2 = ... = \gamma_s, \\ \alpha_{1,i} = \alpha_{2,i} = ... = \alpha_{s,i}, \quad i = 1, ...p \\ \text{and} \\ \beta_{1,j} = \beta_{2,j} = ... = \beta_{s,j}, \quad j = 1, ..., q. \end{array} \tag{18}$$

by the Wald statistic

$$W_C = (\boldsymbol{B}\widehat{\boldsymbol{\theta}})'(\boldsymbol{B}\widehat{\boldsymbol{\Lambda}}\boldsymbol{B}')^{-1}(\boldsymbol{B}\widehat{\boldsymbol{\theta}}), \tag{19}$$

where $q = max(q_1, q_2, ..., q_s)$, $p = max(p_1, p_2, ...p_s)$, $\beta_{r,j} = 0$ for $j > q_r$ and $\alpha_{r,i} = 0$ for $i > p_r$ $(r = 1, ..., s)$. The matrix $\boldsymbol{B}$ is built to represent the linear constraints; its form for each single case is intuitive, whereas its general expression is formally heavy. We show it in the final Appendix for people interested to computational and programming aspects. W_C follows the central chi-squared distribution with degrees of freedom equal to the number of constraints (rows of B).

5 Simulation Studies

In this section we verify the goodness of the clustering algorithm with some simulation studies. We have performed the simulation experiments with many GARCH data generating processes (DGPs). To save space we show only the results relative to six DGPs; the considerations that will be made are valid in general also for the cases not illustrated here (anyway the results are available on request). The six DGPs are given by (uv indicates the value of the unconditional volatility, whereas tvv the value of the time-varying volatility, using equations (8) and (10) respectively):

- M_1: GARCH(1,1) with coefficients $\gamma = 0.1$, $\alpha_1 = 0.5$, $\beta_1 = 0.2$, $uv = 0.214$, $tvv = 0.510$;
- M_2: GARCH(1,1) with coefficients $\gamma = 0.1$, $\alpha_1 = 0.1$, $\beta_1 = 0.5$, $uv = 0.214$, $tvv = 0.115$;
- M_3: ARCH(1) with coefficients $\gamma = 0.5$, $\alpha_1 = 0.6$, $uv = 1.25$, $tvv = 0.6$;
- M_4: GARCH(1,1) with coefficients $\gamma = 0.482$, $\alpha_1 = 0.5$, $\beta_1 = 0.4$, $uv = 1.25$, $tvv = 0.546$;

Table 1: Simulation results: reject percentage (corresponding to sizes 0.01 and 0.05) applying the test for equal unconditional volatility to clusters composed by k series generated by the model indicated in the first column and 1 series generated by the model indicated in the first row.

		The series to be added to the cluster is generated by:							
		the same model		M_2		M_4		M_6	
Model	k	1%	5%	1%	5%	1%	5%	1%	5%
M_1	1	0.4	2.3	2.2	5.3	100	100	99.8	99.9
	2	0.7	2.1	2.9	6.2	100	100	99.4	99.7
	3	0.6	1.6	1.9	4.1	99.8	100	99.5	99.8
	4	0.4	2.0	2.0	4.0	99.9	100	99.7	99.8
M_3	1	0.0	0.5	98.7	99.8	90.0	96.5	0.3	3.4
	2	0.1	0.7	99.9	100	93.7	98.0	0.7	4.0
	3	0.0	0.1	99.9	100	91.2	97.3	0.8	2.8
	4	0.0	0.7	100	100	88.5	96.7	0.8	2.6
M_5	1	8.0	13.5	98.0	99.4	76.7	93.0	4.4	8.8
	2	5.8	10.1	100	100	92.9	97.8	2.7	6.5
	3	4.6	7.1	100	100	96.1	98.6	2.8	5.1
	4	4.0	9.8	99.9	99.9	97.0	98.6	1.7	4.1

- M_5: GARCH(2,1) with coefficients $\gamma = 2$, $\alpha_1 = 0.2$, $\alpha_2 = 0.2$, $\beta_1 = 0.1$, $uv = 3.492$, $tvv = 0.270$;
- M_6: GARCH(1,2) with coefficients $\gamma = 1.692$, $\alpha_1 = 0.4$, $\beta_1 = 0.1$, $\beta_2 = 0.2$, $uv = 3.492$, $tvv = 0.410$.

The pairs (M_1, M_2), (M_3, M_4), (M_5, M_6) have the same unconditional volatility. M_1 and M_2 have different time-varying volatility, whereas M_5 and M_6 have a time varying volatility slightly different; finally tvv is practically the same for M_3 and M_4 (the difference is equal to 0.054).

The length of each simulated series is $T = 750$, which is the time span used in the application on real data shown in the next section. In the estimation of the GARCH models we have noted that the models with larger number of parameters (in particular M_5) show larger variability in the estimates with respect the low order GARCH. This result is obvious, but affects the final simulation results.

We have performed several experiments using the simulated data to verify the performance of the Wald test in this framework and to evaluate the quality of the clusters obtained with the proposed procedure.

5.1 Evaluation of the Size and Power of the Wald Tests

The aim of the first group of experiments is to evaluate the performance of the test procedures illustrated in section 4. In the first one we hypothesize a first level clustering framework. In practice we suppose to have a cluster composed by n ($n = 1, ...4$) series generated by the same model (M_1, M_3 or M_5), for which the testing procedure has verified that they have the same unconditional volatility; we verify if adding a new series (generated by the same model or by M_2, M_4, M_6) the test procedure will assign correctly the series to this cluster. The hypotheses are tested at the nominal sizes of 1% and 5%; for each case we have replicated 1000 simulations. The results are shown in Table 1; it seems that the test procedure has a good performance, assigning correctly the series to the clusters, also when they are generated by different GARCH processes, but with equal unconditional volatility. In the case of Model M_5 we note that the size of the test is not correctly respected (there is a certain propensity to reject the null hypothesis), probably due to the wider variability of the estimates; anyway the results seem acceptable. In the other two cases, the size seems underestimated, with the extreme situations of model M_3 where the null hypothesis is always correctly not rejected in several scenarios at nominal size of 0.01. Anyway, it is a valid test because the power is very high in all the cases (also with peaks of 100%).

The second experiment is similar to the first one, but we verify the hypothesis of equal time-varying volatility by the test procedure shown in section 4.2, hypothesizing a second level clustering. In this case we hypothesize clusters with equal unconditional volatility and equal time-varying volatility, so that, when the cluster is formed by series generated by model M_i ($i = 1, 3, 5$), the series to be added can be generated by the same model M_i or by the model with equal unconditional volatility (M_2 for $i = 1$, M_4 for $i = 3$, M_6 for $i = 5$). The results of this experiment are shown in Table 2 and they seem to be consistent with the values of tvv of the DGPs. In general, when the series added to the cluster is generated by the same model, the reject percentage seems sufficiently similar to the nominal size. When the series added is generated by the other model with equal uv, the test procedure rejects a very high percentage of cases when the two DGPs have very different tvv (M_1 and M_2); it rejects a percentage similar to the size of the test in the case of DGPs having equal tvv (M_3 and M_4); it rejects a small percentage of cases when the two DGPs have similar (but not equal) tvv (M_5 and M_6). Obviously, the percentage of the last case reduces increasing the number of the series generated by the same model M_5.

The third experiment is aimed at evaluating the performance of the test of equal volatility structure described in section 4.3. In this case the clustering framework is similar to that one of the second experiment, but, dealing with clusters with the same uv and tvv, only the case with models M_3 and M_4 was considered. It is important to note that in this case the number of constraints to be tested is larger than the other two tests, evaluating the equality of each coefficients of the DGPs. The test procedure has a practically perfect performance (Table 3); in this case the presence of a large coefficient β_1 in M_4, which is not present in M_3, favors this result. Anyway, also in other experiments with other DGPs, not illustrated here, the power of the test is very good.

In conclusion, it seems that the use of the Wald test in the clustering procedure is supported by the simulations. The tendency to reject the null hypothesis arisen in some particular case seems due to the variability in the estimates of models with a larger

Table 2: Simulation results: reject percentage (corresponding to sizes 0.01 and 0.05) applying the test for equal time-varying volatility to clusters composed by k series generated by the model indicated in the first column and 1 series generated by the model indicated in the corresponding row.

The series to be added to the cluster is generated by:					
		the same model		other model	
Model	k	1%	5%	1%	5%
M_1	1	1.5	6.0	97.9	99.6
(other	2	0.6	2.9	98.3	99.5
model	3	0.6	2.0	98.1	99.2
M_2)	4	0.5	1.1	98.7	99.7
M_3	1	1.5	5.9	1.8	8.2
(other	2	0.5	2.8	1.3	5.2
model	3	0.5	2.3	2.1	4.5
M_4)	4	0.4	0.8	0.9	3.5
M_5	1	2.7	6.7	13.8	30.6
(other	2	1.8	3.9	9.1	25.9
model	3	1.8	2.9	6.6	18.5
M_6)	4	1.4	2.5	7.6	16.8

Table 3: Simulation results: reject percentage (corresponding to sizes 0.01 and 0.05) applying the test for equal volatility structure to clusters composed by k series generated by the model indicated in the first column and 1 series generated by the model indicated in the corresponding row.

The series to be added to the cluster is generated by:					
		the same model		M_4	
Model	k	1%	5%	1%	5%
	1	1.3	5.4	100	100
M_3	2	0.9	3.3	100	100
	3	0.4	2.2	100	100
	4	0.7	1.6	100	100

number of coefficients. Note also that the iterative test procedure implies an increasing number of equalities under the null hypotheses (12), (16), (18); theoretically this could cause a decrease in the power of the test. This fact is not verified in our experiments and the results are consistent with those obtained by Chenowet et al. (2004), who show that, in presence of large data set, the Wald test verifying the equality of ARMA models (or the equality of coefficients of different ARMA models) does not suffer from this kind of problem.

5.2 Evaluation of the Quality of the Clustering Procedure

An important task is to verify if the algorithm proposed is able to create an adequate subdivision of the universe in homogeneous clusters. For this purpose a number of cluster validation indexes has been proposed; these indexes express the quality of a given clustering (see for example Theodoridis and Koutroumbas, 1998, or Jain et al., 1999, for a review). We consider three indexes which are intuitively plausible and largely diffused in the clustering literature:

- *C index* (Hubert and Schultz, 1976); it is calculated as:

$$C = \frac{S - S_{min}}{S_{max} - S_{min}}. \tag{20}$$

  Let us indicate with a the number of all pairs of series where both series are included in the same cluster; S is the sum of distances between series in the a pairs; S_{min} and S_{max} are the sum of the a smallest distances and the sum of the a largest distances respectively, considering all possible pairs in the universe of series. This index falls in the interval $[0, 1]$, assuming small values when the quality of the clustering is good.

- *Davies-Bouldin index* (Davies and Bouldin, 1979); it is defined as:

$$DB = \frac{1}{m} \sum_{i=1}^{m} \max_{j(j \neq i)} \frac{\delta_i + \delta_j}{d(c_i, c_j)},$$

  where m is the number of clusters, δ_r is the average distance of all series in cluster r to their cluster center c_r, $d(c_i, c_j)$ is the distance between the cluster centers c_i and c_j. The DB index falls in the interval $[0, \infty]$ and assumes small values when the quality of the cluster is good.

- *Dunn index* (Dunn, 1974): it is given by:

$$D = \frac{d_{min}}{d_{max}},$$

  where d_{min} is the smallest distance between two series from different clusters and d_{max} is the largest distance of two series from the same cluster. The D index falls in the interval $[0, \infty]$ and assumes large values when the quality of the cluster is good.

Table 4: Simulation results: comparison of C_r and C_e and synthesis of the distribution of C_e calculated on clusters of equal unconditional volatility relative to n series generated by models M_1-M_6 (1000 simulations).

	percentages of		distribution of C_e				
	cases with $C_r = C_e$	cases with $C_r > C_e$	Min	Q_1	Me	Q_3	Max
		$n = 10$					
1%	61.9	7.9	0.00	0.00	0.01	0.05	0.32
5%	64.9	12.3	0.00	0.00	0.01	0.03	0.32
		$n = 20$					
1%	35.8	11.9	0.00	0.01	0.03	0.05	0.17
5%	46.4	16.2	0.00	0.01	0.02	0.04	0.17
		$n = 30$					
1%	23.1	15.2	0.00	0.02	0.03	0.04	0.11
5%	33.8	20.3	0.00	0.01	0.02	0.03	0.11

In the following experiments we will use the Euclidean distance in the computation of the three indexes; moreover, showing results with similar interpretation, we will illustrate only those relative to the C index which, being limited in the interval $[0, 1]$, can be easily interpreted.

The aim of the first experiment is to evaluate the quality of the first-level clustering, based on the Wald test procedure. For this purpose we have generated, for each DGP, 5 series, obtaining a universe of 30 series. Then we have chosen randomly a number n of series (n =10, 20, 30) and applied the clustering procedure. Finally we have calculated the three indexes for the estimated clustering and the *real* one; the last one is obtained inserting in the same cluster the series generated by DGPs with the same unconditional variance. In this way the *real* clustering is a valid benchmark, but it is possible that the estimated clustering could have a better quality because the uv considered is the estimated one and not the theoretical uv of the DGP. In the left part of Table 4 we show the percentage of cases (on 1000 replications) in which we have obtained the same C index for the estimated and the real clustering (call them C_e and C_r respectively), and the percentage of cases in which the estimated clustering shows a better quality with respect to the real one ($C_r > C_e$). We notice that there is a good percentage of cases with equal quality for $n = 10$ (more than 60%) and the share decreases when the size of the universe increases; on the other side the percentage of cases in which the estimated clustering is better than the true one increases when the universe increases. Furthermore, the cases in which the estimated clustering has a smaller C index with respect to the true one, show in general a high quality. This aspect can be seen in the right part of Table 4, which shows the minimum, the first quartile, the median, the second quartile and the maximum of the empirical distribution of C_e are shown; we can notice that more than 75% of cases has a C_e index smaller than 0.1 and the maximum is very small. Furthermore, the maximum value decreases when the size of the universe of the series

Table 5: Simulation results: comparison of C_r and C_e and synthesis of the distribution of C_e calculated on clusters of equal time-varying volatility relative to n series generated by pairs of models (1000 simulations).

	percentages of		distribution of C_e				
	cases with $C_r = C_e$	cases with $C_r > C_e$	Min	Q_1	Me	Q_3	Max
		Series generated by M_1 and M_2; $n = 6$					
1%	93.8	1.9	0.00	0.00	0.00	0.00	0.32
5%	91.5	3.5	0.00	0.00	0.00	0.00	0.16
		Series generated by M_1 and M_2; $n = 10$					
1%	86.9	2.2	0.00	0.00	0.00	0.00	0.17
5%	84.1	3.8	0.00	0.00	0.00	0.00	0.17
		Series generated by M_3 and M_4; $n = 6$					
1%	2.1	97.2	0.00	0.00	0.00	0.00	0.53
5%	2.0	94.9	0.00	0.00	0.00	0.00	0.61
		Series generated by M_3 and M_4; $n = 10$					
1%	0.00	99.5	0.00	0.00	0.00	0.00	0.47
5%	0.00	98.6	0.00	0.00	0.00	0.00	0.51
		Series generated by M_5 and M_6; $n = 6$					
1%	11.7	73.1	0.00	0.00	0.00	0.00	0.55
5%	12.6	66.5	0.00	0.00	0.00	0.15	0.57
		Series generated by M_5 and M_6; $n = 10$					
1%	1.3	73.9	0.00	0.00	0.00	0.20	0.58
5%	1.8	65.9	0.00	0.00	0.10	0.22	0.54

to be clustered increases.

A similar experiment was conducted to evaluate the quality of the second level clustering. In this instance we have distinguished three cases, each one composed by two models with the same unconditional variance. As in the previous experiment, 5 series from each DGP were generated and then n of them extracted to compose the universe; in this case we have chosen $n = 6$ to guarantee the existence of two DGPs in the universe, and $n = 10$, using all the series generated. The results of this experiment are synthesized in Table 5. The first case considers series generated by M_1 and M_2, which possess two different tvv, so that the benchmark constituted by the *real* clustering groups the series correctly. The results for this case are quite positive. In fact, in more than 90% of cases, for $n = 6$, the estimated clustering has the same quality of the benchmark and its quality decreases slightly for $n = 10$. The pair (M_3,M_4) has an opposite behavior, being constituted by series with the same uv and tvv; in more than 95% of cases the empirical clustering is better than the benchmark. Finally, the case with M_5 and M_6 shows that more than 73% of cases have a better quality with respect to the benchmark, considering a nominal size equal to 1%, and 66% of cases in correspondence of a nominal size of 1%. The distribution of C_e demonstrates the good quality of the estimated clustering,

Table 6: Simulation results: comparison of C_r and C_e and synthesis of the distribution of C_e calculated on clusters of equal volatility structure relative to n series generated by models M_3 and M_4 (1000 simulations).

	percentages of		distribution of C_e				
	cases of $C_r = C_e$	cases of $C_r > C_e$	Min	Q_1	Me	Q_3	Max
		$n = 6$					
1%	97.7	1.2	0.00	0.00	0.00	0.00	0.43
5%	89.9	3.6	0.00	0.00	0.00	0.00	0.48
		$n = 10$					
1%	96.6	1.3	0.00	0.00	0.00	0.01	0.16
5%	88.8	4.1	0.00	0.00	0.00	0.01	0.30

with few cases in which the index is more than 0.2.

The final simulation experiment was conducted to evaluate the quality of the third level clustering. Also in this case, having to consider only clusters with equal uv and equal tvv, we limit the experiment to the M_3 and M_4 DGPs. The two DGPs are clearly different and the real clustering is a useful benchmark for the evaluation; in this case a large percentage of estimated clustering (more tan 96% in correspondence of a nominal size of 1% and around 90% in correspondence of a nominal size of 5%) shows a quality equal to the benchmark.

6 An Example:Classifying the Italian Sectorial Indexes

A typical problem in financial investments is the evaluation of the stocks in terms of their degree of risk. This is particularly important in asset allocation problems or in the management of financial funds. The application we illustrate is the classification of the twenty sectorial indexes of the Italian Mibtel general index. They can be divided in three main sectors:

FINANCE: Banks (Fba), Finance Holdings (Ffh), Finance Misc. (Ffm), Finance Services (Ffs), Insurance (Fin), Real Estate (Fre);

INDUSTRIAL: Cars (Ica), Chemicals (Ich), Construction (Ico), Electronics (Iel), Food (Ifo), Industrial Misc. (Iim), Minerals Metals (Imm), Paper (Ipa), Plant Machine (Ipm), Textile Clothing (Itc);

SERVICE: Distribution (Sdi), Media (Sme), Public Utility (Spu), Transport Tourism (Stt).

These series were studied by Billio et al. (2006), who deduce similar correlation dynamics within the main sectors. Anyway similar correlation dynamics do not imply similar degrees of risk. For this purpose we apply our procedure, adopting the widely accepted idea that the risk can be represented by the volatility of the series (see, for example, Arnott and Fabozzi, 1988), modeled by GARCH processes. We consider the returns of the series from 23 March 2004 to 27 February 2007 (daily data, 750 observations; source: Yahoo finance) and estimate a different (1)-(2) model for each one, selecting the order of each GARCH by the AIC criterion. We have also used the BIC criterion and, of course, the models are more parsimonious; anyway, the final results in terms of clustering are similar.

Tables 7-9 show the estimates of the twenty GARCH models, sorted by the three main sectors. Note that for nine of the twenty series the AIC criterion selects the GARCH(1,3) model. Many parameters are not significantly different from zero; we have erased, in the estimation procedure, those identically equal to zero, which cause problems in the computation of standard errors. In the same tables we show the estimates of the unconditional volatility and the time-varying volatility, obtained by (8) and (10) respectively. The index with largest uv is Ffm, which is composed by a mixture of financial products. The other indexes with unconditional volatility greater than 1 are the financial services (in the finance sector) and paper and cars (in the industrial sector). In the service sector, Sdi has an unconditional volatility which is around twice the unconditional volatility of the other indexes belonging to the same sector, but it is less than one. The time-varying volatility shows a limited range between 0.109 (papers) and 0.341 (public utility services). It is interesting to notice that there is a low correlation between the constant and the time-varying volatilities (the correlation coefficient is equal to 0.21) because they express two different characteristics of the volatility. The unconditional volatility can be viewed as the expected global risk of the series and depends on the nature of the financial series studied. On the other side the time-varying volatility captures only the movements of the volatility along the time and depends on the turmoil and quiet periods which alternate; the similarity of this kind of volatility among the twenty series is a consequence of the similar movements of the financial time series, which react in a similar way to shocks and other transmission mechanisms, such as spillover and contagion effects (see, for example, Forbes and Rigobon, 2002, Gallo and Otranto, 2007a and 2007b).

To apply the clustering procedure proposed we order the indexes in terms of increasing unconditional volatility. In Table 10 we show the null hypotheses verified and the corresponding p-values to obtain the clusters. Fin is the index with the smallest unconditional volatility and we verify if it is not significantly different from a constant volatility equal to zero, which represents our initial benchmark (hypothesis (15)). Considering a nominal size equal to 0.01, we reject the null, so that Fin is the new benchmark. The new null hypothesis consists of the equality of unconditional volatility between Fin and Ico (the successive series with minimum unconditional volatility). We do not reject this hypothesis and continue adding a new series (Fre). We go on until the null (12) is rejected; this happens when the series Spu is added to the null hypothesis. In this case we close the first cluster and Spu is the new benchmark for the second cluster. Then we restart the procedure with the new benchmark and the remaining series. Finally, we obtain four clusters: cluster 1 contains the series with the smallest unconditional volatility (but different from zero), whereas cluster 4 is that one with maximum unconditional volatility. The classification is illustrated in Table 11; we will comment the results at the end of this section.

Table 7: Estimates of GARCH models for the returns of the Italian financial indexes, unconditional volatility (uv) and time-varying volatility (tvv); standard errors in parentheses.

	μ	γ	α_1	α_2	β_1	β_2	β_3	uv	tvv
Fba	0.101	0.082	0.148		0.526		0.150	0.277	0.185
	(0.024)	(0.032)	(0.050)		(0.374)		(0.043)	(0.002)	(0.002)
Ffh	0.147	0.099	0.293		0.265		0.294	0.276	0.325
	(0.025)	(0.026)	(0.057)		(0.092)		(0.092)	(0.002)	(0.003)
Ffm	0.081	1.726	0.250	0.154		0.236	0.046	3.509	0.304
	(0.073)	(0.331)	(0.059)	(0.042)		(0.155)	(0.108)	(0.264)	(0.003)
Ffs	0.137	0.415	0.258		0.316	0.089	0.177	1.185	0.275
	(0.051)	(0.107)	(0.055)		(0.160)	(0.200)	(0.137)	(0.029)	(0.003)
Fin	0.075	0.056	0.154		0.438		0.302	0.236	0.197
	(0.024)	(0.021)	(0.049)		(0.148)		(0.129)	(0.002)	(0.002)
Fre	0.155	0.081	0.210		0.181	0.490		0.272	0.243
	(0.026)	(0.023)	(0.043)		(0.153)	(0.159)		(0.002)	(0.002)

Table 8: Estimates of GARCH models for the returns of the Italian industrial indexes, unconditional volatility (uv) and time-varying volatility (tvv); standard errors in parentheses.

	μ	γ	α_1	α_3	β_1	β_2	β_3	uv	tvv
Ica	0.141	0.522	0.285		0.188		0.242	1.142	0.300
	(0.044)	(0.146)	(0.059)		(0.088)		(0.109)	(0.028)	(0.003)
Ich	0.049	0.158	0.139		0.603			0.435	0.175
	(0.027)	(0.069)	(0.038)		(0.129)			(0.004)	(0.002)
Ico	0.137	0.055	0.129	0.040	0.768			0.260	0.243
	(0.027)	(0.027)	(0.035)	(0.066)	(0.095)			(0.003)	(0.003)
Iel	0.099	0.174	0.138		0.282		0.288	0.443	0.154
	(0.028)	(0.059)	(0.045)		(0.183)		(0.171)	(0.003)	(0.002)
Ifo	0.130	0.158	0.143		0.255		0.438	0.562	0.173
	(0.033)	(0.049)	(0.045)		(0.127)		(0.113)	(0.005)	(0.002)
Iim	0.044	0.116	0.101		0.793			0.596	0.167
	(0.036)	(0.038)	(0.028)		(0.051)			(0.007)	(0.001)
Imm	0.058	0.102	0.076		0.794			0.519	0.125
	(0.032)	(0.046)	(0.030)		(0.074)			(0.008)	(0.002)
Ipa	-0.003	0.417	0.097		0.507	0.182		1.422	0.109
	(0.221)	(0.295)	(0.049)		(0.367)	(0.342)		(0.029)	(0.002)
Ipm	0.150	0.187	0.169		0.724			0.753	0.246
	(0.042)	(0.053)	(0.042)		(0.059)			(0.011)	(0.002)
Itc	0.109	0.130	0.201		0.244	0.123	0.242	0.380	0.213
	(0.028)	(0.041)	(0.047)		(0.227)	(0.186)	(0.116)	(0.003)	(0.002)

Table 9: Estimates of GARCH models for the returns of the Italian service indexes, unconditional volatility (uv) and time-varying volatility (tvv); standard errors in parentheses.

	μ	γ	α_1	α_2	α_3	β_1	β_2	β_3	u.v	tvv
Sdi	0.055	0.250	0.181			0.543		0.071	0.731	0.221
	(0.038)	(0.107)	(0.056)			(0.169)		(0.147)	(0.027)	(0.003)
Sme	0.048	0.150	0.081	0.161		0.237		0.269	0.362	0.173
	(0.026)	(0.063)	(0.042)	(0.054)		(0.143)		(0.140)	(0.004)	(0.003)
Spu	0.072	0.231	0.314	0.117	0.062				0.456	0.341
	(0.022)	(0.027)	(0.067)	(0.052)	(0.050)				(0.004)	(0.004)
Stt	0.085	0.103	0.228				0.620		0.316	0.290
	(0.026)	(0.035)	(0.053)				(0.087)		(0.003)	(0.003)

Table 10: Test procedure to verify the equality of unconditional volatilities: null hypothesis (H_0) and corresponding p-value .

H_0	$p-value$
Fin=0	0.000
Fin=Ico	0.724
Fin=Ico=Fre	0.830
Fin=Ico=Fre=Ffh	0.912
Fin=Ico=Fre=Ffh=Fba	0.957
Fin=Ico=Fre=Ffh=Fba=Stt	0.914
Fin=Ico=Fre=Ffh=Fba=Stt=Sme	0.784
Fin=Ico=Fre=Ffh=Fba=Stt=Sme=Itc	0.505
Fin=Ico=Fre=Ffh=Fba=Stt=Sme=Itc=Ich	0.186
Fin=Ico=Fre=Ffh=Fba=Stt=Sme=Itc=Ich=Iel	0.028
Fin=Ico=Fre=Ffh=Fba=Stt=Sme=Itc=Ich=Iel=Spu	0.008
Spu=Imm	0.556
Spu=Imm=Ifo	0.540
Spu=Imm=Ifo=Iim	0.544
Spu=Imm=Ifo=Iim=Sdi	0.451
Spu=Imm=Ifo=Iim=Sdi=Ipm	0.181
Spu=Imm=Ifo=Iim=Sdi=Ipm=Ica	0.004
Ica=Ffs	0.857
Ica=Ffs=Ipa	0.459
Ica=Ffs=Ipa=Ffm	0.000

Table 11: Classification in clusters of equal unconditional volatility and equal time-varying volatility. The series between bars form sub-clusters with equal volatility structure.

				CLUSTER 1					
|Fba	Fin	Itc	Iel	Ffh	Sme|	|Fre	Stt|	|Ich	Ico|
				CLUSTER 2					
|Iim	Sdi	Ipm	Imm|	Ifo	Spu				
				CLUSTER 3					
|Ffs	Ica|	Ipa							
				CLUSTER 4					
Ffm									

The second level classification can be made considering the time-varying volatility. For each cluster we consider as initial benchmark that one with minimum time-varying volatility (Iel for cluster 1, Imm for cluster 2, Ipa for cluster 3); then we verify the null hypothesis of equal time-varying volatility (16), adding series with increasing tvv (similarly to the previous analysis relative to the unconditional volatility). In this case the second level clustering is equal to the first level, in the sense that all the series having equal uv have also equal tvv.

Finally, we verify if, within the four clusters, the series have equal volatility structure, to provide the third level classification. Applying the procedure based on the p-value and described in section 3, we obtain the final classification showed in Table 11, where the series with equal volatility structure are included between bars, providing a deeper subdivision in sub-groups.

In conclusion, Table 11 shows that the twenty series of the Italian sector indexes can be grouped in four homogeneous clusters, with similar unconditional and time-varying volatilities. The first kind of volatility is that one which discriminates the series, whereas the time-varying volatility seems similar in the full data set. It is likely that, being series relative to the same country, the degree of risk is different, but the dynamics is very similar. The first group contains ten of the twenty series, characterized by an unconditional volatility included between 0.2 and 0.45. Three GARCH DGPs seem to have generated these ten series: six series (Fba, Fin, Itc, Iel, Ffh, Sme) are generated by a GARCH(1,3) process, with constant γ around 0.1, the coefficient α_1 between 0.1 and 0.2, the coefficient β_2 and β_3 not significantly different from zero. In practice these series could be considered as generated by a GARCH(1,1) process. In the same first (and second) level cluster two series (Fre and Stt) are generated by a GARCH(1,2) process, with $\gamma \simeq 0.1$, $\alpha_1 \simeq 0.2$, β_1 equal to zero and β_2 large (around 0.5), and two series (Ich and Ico) by a GARCH(1,1) model with coefficient $\alpha_1 \simeq 0.1$ and a large β_1 (the coefficients α_2 and α_3 relative to the series Ico are not significantly different from zero).

The second cluster includes the series of industrial and service sectors with unconditional volatility between 0.45 and 0.75. Two series (Ifo and Spu) are generated respectively by a GARCH(1,3) process, with large β_1 and β_3 coefficients, and an ARCH(3) process, and four series (Iim, Sdi, Ipm and Imm) generated by the same model, a GARCH(1,1) with large coefficient β_1 (around 0.7).

The third cluster is formed by the series with unconditional volatility around 1.2; two series (Ffs and Ica) are generated by a GARCH(1,3) model with γ around 0.5, α_1 coefficient around 0.27 and small β coefficients; the remaining series (Ipa) is generated by a GARCH(1,2) model, different from that one of cluster 2 for the presence of larger γ and β_1, and smaller α_1 and β_2.

Finally, cluster 4 contains only the series Ffm, which has an unconditional volatility more than twice that one of Ipa, which is the series with highest constant volatility in cluster 3. The series Ffm is generated by a GARCH(2,3) process with large γ and α coefficients.

If we use the clustering generated by the test procedure considering a nominal size for the tests equal to 0.05, the results are very similar: the only difference is that Iel belongs to the second cluster and not to the first. In this case, the choice could be based on the quality index illustrated in section 5. The clustering at 1% nominal size provides the following indexes:

$$C = 0.016, \quad DB = 0.324, \quad D = 0.043,$$

whereas the clustering at 5% nominal size will provide:

$$C = 0.020, \quad DB = 0.352, \quad D = 0.027.$$

All the three indexes indicate the (slightly) better performance of the first clustering; moreover, the very low value of index C confirms the high quality of the subdivision obtained.

7 Final Remarks

In this paper we have proposed a clustering algorithm to group heteroskedastic time series with similar volatility patterns, based on a particular decomposition of the squared disturbances of the series. A natural field of application of this procedure is the clustering of financial time series based on their volatility structure.

A first classification level is based only on the unconditional volatility, which can be interpreted as the expected constant risk of the returns; the second classification level is based on the unconditional and time-varying volatility, which depends also on the history of the time series; finally a third level is based on the full volatility structure. The three levels have an increasing degree of deepness, implying a more accuracy in the detection of the similar characteristics. On the other side, a deeper classification could provide a large number of clusters, often composed by only one series, so the choice of the level depends on the type of classification we need. Anyway, in our application, the first and second level clustering provide the same clustering, and the third level provides useful information about the DGPs without an excessive splitting.

The first level could be useful in fund management because the series will be grouped taking into account their expected risk, excluding from the analysis all the historical factors that have caused some shocks. In this case another useful information could derive by the minimum expected risk, given by the first term in the right part of equation (7). The second level classifications could be more useful in the historical analysis, for example to evaluate if the reactions to turmoil periods are similar or not. The third level classification seems important when the user is interested in the differences in the GARCH estimated models, without having to study and compare the estimated parameters across all the series.

We have used the GARCH(p,q) models to represent the volatility of the series and we have developed the approach deriving the models underlying the squared disturbances; in particular, in the application proposed, the squared disturbances are equivalent to the squared returns and have the advantage to not calculate a proxy to represent the volatility. Anyway, we could consider, instead of squared returns, the absolute returns, the range or other proxies of the volatility, maintaining the flavor of the procedure. For example, we can model the range in terms of VAR (as in Gallo and Otranto, 2007a and 2007b) and obtain the corresponding $AR(\infty)$ representation; the steps of the procedure are the same.

As said in section 4, our approach works in a univariate framework; this is made because the extension to the multivariate case will be unfeasible when we deal with a large number of series. Anyway, if we suppose a multivariate framework à la Engle (2002) the results will be the same; in fact, if we assume a Dynamic Conditional Correlation structure with n GARCH models for the volatility part, the likelihood can be split in two parts: one relative to the volatility part and one relative to the correlation part. In our approach we are interested only on the first part, which can be estimated maximizing n separate likelihoods relative to the n univariate GARCH models (Engle, 2002). Furthermore, the covariance matrix of the ML estimators of the GARCH parameters is a block diagonal matrix with elements of each block obtained by the Bollerslev and Wooldridge (1992) robust covariance matrix of each GARCH model (Engle and Sheppard, 2001), which corresponds to the matrix $\widehat{\mathbf{\Lambda}}$ used in section 4.

Acknowledgments

I would like to thank two anonymous referees and an associate editor for their detailed reading of the paper and many suggestions which lead to the clarification of the focus of the paper. Financial support from Italian MIUR under Grant 2006137221_001 is gratefully acknowledged.

Appendix A: General Formulation of C in (19)

Let $p_r^* = max(p_{r-1}, p_r)$ and $q_r^* = max(q_{r-1}, q_r)$ $(r = 2, ...k)$. Furthermore let $\boldsymbol{I}_h$ the $h \times h$ identity matrix and $\mathbf{0}_{v,z}$ the $v \times z$ matrix with all the elements equal to zero. Let us define:

$$\boldsymbol{A}_{p_r^*} = \begin{cases} \begin{bmatrix} \boldsymbol{I}_{p_{r-1}} \\ \mathbf{0}_{p_r^*-p_{r-1},p_{r-1}} \end{bmatrix} & if \quad p_r^* > p_{r-1} \\ \boldsymbol{I}_{p_{r-1}} & if \quad p_r^* = p_{r-1} \end{cases},$$

$$\boldsymbol{A}^-_{p_r^*} = \begin{cases} \begin{bmatrix} -\boldsymbol{I}_{p_r} \\ \mathbf{0}_{p_r^*-p_r,p_r} \end{bmatrix} & if \quad p_r^* > p_r \\ \boldsymbol{I}_{p_r} & if \quad p_r^* = p_r \end{cases},$$

$$\boldsymbol{B}_{q_r^*} = \begin{cases} \begin{bmatrix} \boldsymbol{I}_{q_{r-1}} \\ \mathbf{0}_{q_r^*-q_{r-1},q_{r-1}} \end{bmatrix} & if \quad q_r^* > q_{r-1} \\ \boldsymbol{I}_{q_{r-1}} & if \quad q_r^* = q_{r-1} \end{cases},$$

and

$$\boldsymbol{B}^-_{q_r^*} = \begin{cases} \begin{bmatrix} -\boldsymbol{I}_{q_r} \\ \mathbf{0}_{q_r^*-q_r,q_r} \end{bmatrix} & if \quad q_r^* > q_r \\ \boldsymbol{I}_{q_r} & if \quad q_r^* = q_r \end{cases}$$

then the matrix $\boldsymbol{C}$ is given by:

$$\begin{bmatrix} 1 & \mathbf{0}_{1,p_1} & \mathbf{0}_{1,q_1} & -1 & \mathbf{0}_{1,p_2} & \mathbf{0}_{1,q_2} & \cdots & 0 & \mathbf{0}_{1,p_{k-1}} & \mathbf{0}_{1,q_{k-1}} & 0 & \mathbf{0}_{1,p_k} & \mathbf{0}_{1,q_k} \\ \mathbf{0}_{p_2^*,1} & \boldsymbol{A}_{p_2^*} & \mathbf{0}_{p_2^*,q_1} & \mathbf{0}_{p_2^*,1} & \boldsymbol{A}^-_{p_2^*} & \mathbf{0}_{p_2^*,q_2} & \cdots & \mathbf{0}_{p_2^*,1} & \mathbf{0}_{p_2^*,p_{k-1}} & \mathbf{0}_{p_2^*,q_{k-1}} & \mathbf{0}_{p_2^*,1} & \mathbf{0}_{p_2^*,p_k} & \mathbf{0}_{p_2^*,q_k} \\ \mathbf{0}_{q_2^*,1} & \mathbf{0}_{q_2^*,p_1} & \boldsymbol{B}_{q_2^*} & \mathbf{0}_{q_2^*,1} & \mathbf{0}_{q_2^*,p_2} & \boldsymbol{B}^-_{q_2^*} & \cdots & \mathbf{0}_{q_2^*,1} & \mathbf{0}_{q_2^*,p_{k-1}} & \mathbf{0}_{q_2^*,q_{k-1}} & \mathbf{0}_{q_2^*,1} & \mathbf{0}_{q_2^*,p_k} & \mathbf{0}_{q_2^*,q_k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \mathbf{0}_{1,p_1} & \mathbf{0}_{1,q_1} & 0 & \mathbf{0}_{1,p_2} & \mathbf{0}_{1,q_2} & \cdots & 1 & \mathbf{0}_{1,p_{k-1}} & \mathbf{0}_{1,q_{k-1}} & -1 & \mathbf{0}_{1,p_k} & \mathbf{0}_{1,q_k} \\ \mathbf{0}_{p_k^*,1} & \mathbf{0}_{p_k^*,p_1} & \mathbf{0}_{p_k^*,q_1} & \mathbf{0}_{p_k^*,1} & \mathbf{0}_{p_k^*,p_2} & \mathbf{0}_{p_k^*,q_2} & \cdots & \mathbf{0}_{p_k^*,1} & \boldsymbol{A}_{p_k^*} & \mathbf{0}_{p_k^*,q_{k-1}} & \mathbf{0}_{p_k^*,1} & \boldsymbol{A}^-_{p_k^*} & \mathbf{0}_{p_k^*,q_k} \\ \mathbf{0}_{q_k^*,1} & \mathbf{0}_{q_k^*,p_1} & \mathbf{0}_{q_k^*,q_1} & \mathbf{0}_{q_k^*,1} & \mathbf{0}_{q_k^*,p_2} & \mathbf{0}_{q_k^*,q_2} & \cdots & \mathbf{0}_{q_k^*,1} & \mathbf{0}_{q_k^*,p_{k-1}} & \boldsymbol{B}_{q_k^*} & \mathbf{0}_{q_k^*,1} & \mathbf{0}_{q_k^*,p_k} & \boldsymbol{B}^-_{q_k^*} \end{bmatrix}$$

The statistic (19) follows the chi-square distribution with $[s-1+\sum_{r=2}^{s}(p_r^*+q_r^*)]$ degrees of freedom.

References

[1] Agrawal, R., Faloutsos, C. and Swami, A., 1994. Efficient similarity search in sequence databases. Lecture Notes in Computer Science, 69–84.

[2] Arnott, R. and Fabozzi, F.J., 1988. Asset Allocation: a Handbook of Portfolio Policies, Strategies and Tactics. Probus Publishing Company: Chicago, Illinois.

[3] Billio, M., Caporin, M. and Gobbo, M., 2006. Flexible Dynamic Conditional Correlation Multivariate GARCH models for asset allocation. Applied Financial Economics Letters, 2, 123-130.

[4] Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31, 307–327.

[5] Bollerslev, T., Chou, R.Y. and Kroner, K.F., 1992. ARCH modeling in finance: A review of the theory and empirical evidence. Journal of Econometrics, 52, 5–59.

[6] Bollerslev T., Engle R.F. and Nelson D., 1994. ARCH models, in Handbook of Econometrics, Vol. IV (Engle, R.F. and McFadden, D.L. eds.), 2959–3038.

[7] Bollerslev, T. and Wooldridge, J.M., 1992. Quasi-Maximum Likelihood estimation and inference in dynamic models with time varying covariances. Econometric Reviews, 11, 143-172.

[8] Brockwell, P.J. and Davis, R.A., 1996. Introduction to Time Series and Forecasting, Springer-Verlag, New York.

[9] Caiado, J., Crato, N. and Peña, D., 2006. A Periodogram-based metric for time series classification. Computational Statistics and Data Analysis, 50, 2668–2684.

[10] Chenoweth, T., Hubata, R. and St. Louis, R., 2004. The Power of tests for equivalent ARMA models: the implications for practitioners. Empirical Economics, 29, 281–292.

[11] Corduas, M. and Piccolo, D., 2008. Time series clustering and classification by the Autoregressive metric. Computational Statistics and Data Analysis, 52, 1860-1872.

[12] Davies, D. and Bouldin, D., 1979. A cluster separation measure. IEEE Trans. Pattern Recognit. Machine Intell. 1 (2), 224 227.

[13] Dunn, J., 1974. Well separated clusters and optimal fuzzy partitions. J. Cybernet. 4, 95104.

[14] Engle, R., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. Econometrica, 50, 987–1008.

[15] Engle, R., 2002. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. Journal of Business and Economic Statistics, 20, 339–350.

[16] Engle, R. and Sheppard, K., 2001. Theoretical and empirical properties of dynamic conditional correlation Multivariate GARCH. NBER Working Paper, 8554, National Bureau of Economic Research.

[17] Forbes, K. J. and Rigobon, R., 2002. No contagion, only interdependence: measuring stock market comovements. The Journal of Finance, 67, 2223–2261.

[18] Gallo, G. M. and Otranto, E., 2007a. Volatility transmission across markets: a multi-chain Markov switching model, Applied Financial Economics, 17, 659–670.

[19] Gallo, G. M. and Otranto, E., 2007b. Volatility spillovers, interdependence and comovements: a Markov Switching approach, Computational Statistics and Data Analysis, doi: 10.1016/j.csda.2007.09.016.

[20] Gray, A. and Markel, J., 1976. Distance measures for speech processing. IEEE Transactions on Acoustic, Speech and Signal Processing, ASSP-24, 380–391.

[21] Hubert, L. and Schultz, J., 1976. Quadratic assignment as a general data-analysis strategy. Br. J. Math. Stat. Psychol. 29, 190 241.

[22] Jain, A., Murty, M. and Flynn, P., 1999. Data clustering: a review. ACM Computational Surveys 31, 264323.

[23] Liao, T., 2005. Clustering time series data: a survey. Pattern Recognition, 38, 1857–1874.

[24] Maharaj, E. A., 1996. A significance test for class ifying ARMA models. Journal of Statistical Computation and Simulation, 54, 305–331.

[25] Maharaj, E. A., 1999. Comparison and classification of stationary multivariate time series. Pattern Recognition, 32, 1129–1138.

[26] Maharaj, E. A., 2000. Clusters of time series. Journal of Classification, 17, 297–314.

[27] Otranto, E., 2004. Classifying the markets volatility with ARMA distance measures. Quaderni di Statistica, 6, 1–19.

[28] Otranto, E. and Triacca, U., 2007. Testing for equal predictability of stationary ARMA processes, Journal of Applied Statistics, forthcoming.

[29] Piccolo, D., 1990. A distance measure for classifying ARIMA models. Journal of Time Series Analysis, 11, 153–164.

[30] Piccolo, D., 2007. Statistical Issues on the AR Metric in Time Series Analysis. Proceedings of the SIS 2007 intermediate conference ”Risk and Prediction”, 221–232.

[31] Steece, B. and Wood, S., 1985. A test for the equivalence of k ARMA models. Empirical Economics, 10, 1-11.

[32] Theodoridis, S. and Koutroumbas, K.K., 1998. Pattern Recognition. Academic Press.

[33] Xiong, Y. and Yeung, D.-Y., 2002. Mixtures of ARMA models for model-based time series clustering. Proceedings of the IEEE International Conference on Data Mining.

Ultimi Contributi di Ricerca CRENoS

I Paper sono disponibili in: http://www.crenos.it

07/16 *Sergio Lodde,* "Specialization and Concentration of the Manufacturing Industry in the Italian Local Labor Systems"
07/15 *Giovanni Sulis,* "Gender Wage Differentials in Italy: a Structural Estimation Approach"
07/14 *Fabrizio Adriani, Luca G. Deidda, Silvia Sonderegger,* "Over-Signaling Vs Underpricing: the Role of Financial Intermediaries In Initial Public Offerings"
07/13 *Giovanni Sulis,* "What Can Monopsony Explain of the Gender Wage Differential In Italy?"
07/12 *Gerardo Marletto,* "Crossing the Alps: Three Transport Policy Options"
07/11 *Sergio Lodde* "Human Capital And Productivity Growth in the Italian Regional Economies: a Sectoral Analysis"
07/10 *Axel Gautier, Dimitri Paolini,* "Delegation, Externalities and Organizational Design"
07/09 *Rinaldo Brau, Antonello E. Scorcu, Laura Vici,* "Assessing visitor satisfaction with tourism rejuvenation policies: the case of Rimini, Italy"
07/08 *Dimitri Paolini,* "Search and the firm's choice of the optimal labor contract"
07/07 *Giacomo Carboni,* "Shape of U.S. business cycle and long-run effects of recessions"
07/06 *Gregory Colcos, Massimo Del Gatto, Giordano Mion and Gianmarco I.P. Ottaviano,* "Productivity and firm selection: intra-vs international trade"
07/05 *Silvia Balia,* "Reporting expected longevity and smoking: evidence from the share"
07/04 *Raffaele Paci, Stefano Usai,* "Knowledge flows across European regions"
07/03 *Massimo Del Gatto, Giordano Mion and Gianmarco I.P. Ottaviano,* "Trade Integration, firm selection and the costs of non-europe"
07/02 *Vittorio Pelligra,* "Intentions, trust and frames: a note on sociality and the theory of games"
07/01 *Oliviero A. Carboni and Giuseppe Medda,* "Government size and the composition of public spending in a neoclassical growth model"
06/17 *Susana Iranzo, Fabiano Schivardi and Elisa Tosetti,* "Skill dispersion and firm productivity: an analysis with employer-employee matched data"
06/16 *Luigi Guiso and Fabiano Schivardi,* "What determines entrepreneurial clusters?"
06/15 *Vittorio Pelligra,* "Trust Responsiveness: on the dynamics of fiduciary interactions"
06/14 *Rinaldo Brau, Alessandro Lanza and Francesco Pigliaru,* "How fast are small tourist countries growing? The 1980-2003 evidence"
06/13 *Emanuela Marrocu and Raffaele Paci,* "The effects of public capital on the productivity of the italian regions"
06/12 *Raffaele Paci and Stefano Usai,* "Agglomeration economies and growth-the case of Italian local labour systems, 1991-2001"
06/11 *Fabrizio Ariani and Luca G. Deidda,* "The Monopolist's blues"

www.crenos.it