CONTRIBUTI DI RICERCA CRENOS

CLASSIFICATION OF VOLATILITY IN PRESENCE OF CHANGES IN MODEL PARAMETERS

Edoardo Otranto

WORKING PAPERS

2011/13

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

CRENOS was set up in 1993 with the purpose of organising the joint research effort of economists from the two Sardinian universities (Cagliari and Sassari) investigating dualism at the international and regional level. CRENoS' primary aim is to improve knowledge on the economic gap between areas and to provide useful information for policy intervention. Particular attention is paid to the role of institutions, technological progress and diffusion of innovation in the process of convergence or divergence between economic areas. To carry out its research, CRENoS collaborates with research centres and universities at both national and international level. The centre is also active in the field of scientific dissemination, organizing conferences and workshops along with other activities such as seminars and summer schools.
CRENoS creates and manages several databases of various socio-economic variables on Italy and Sardinia. At the local level, CRENoS promotes and participates to projects impacting on the most relevant issues in the Sardinian economy, such as tourism, environment, transports and macroeconomic forecasts.

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

Title: CLASSIFICATION OF VOLATILITY IN PRESENCE OF CHANGES IN MODEL PARAMETERS

First Edition: July 2011

Classification of Volatility in Presence of Changes in Model Parameters*

Edoardo Otranto
Dipartimento di Scienze Cognitive e della Formazione and CRENoS,
University of Messina,
Via Concezione 6, 98121 Messina, Italy;
E-mail: eotranto@unime.it

Abstract

The classification of volatility of financial time series has recently received a lot of contributions: in particular using model based clustering algorithms. Recent works have evidenced how volatility structure can vary along time, with gradual or abrupt changes in the coefficients of the model. We wonder if these changes can affect the classification of series in terms of similar volatility structure. We propose to classify the level of the unconditional volatility obtained from Multiplicative Error Models with the possibility of changes in the parameters of the model in terms of regime switching or time varying smoothed coefficients. They provide different unconditional volatility structures with a proper interpretation, useful to represent different situations of interest. The different methodologies are coherent with each other and provide a common synthetic pattern. The procedure is experimented on fifteen stock indices volatilities.

Keywords: clustering, AMEM, Markov switching, smooth transition, unconditional volatility
JEL Classification: C22, C38, C58

1 Introduction

The classification of time series assumes a particular importance in the financial analysis because it is frequently linked to the volatility of financial indices or assets; the relevance of this point is clear if we underline that volatility is a proxy of risk, so the detection of assets with similar risk is an important tool for traders and for the creation of balanced portfolios. Moreover, from a statistical point of view, the detection of similar volatility patterns is fundamental, in a multivariate framework, to construct models with the same

*Financial support from Italian MIUR under Grant 20087Z4BMK_002 is gratefully acknowledged.

set of parameters driving many time series which possess a similar dynamics (see, for example, Otranto, 2010).

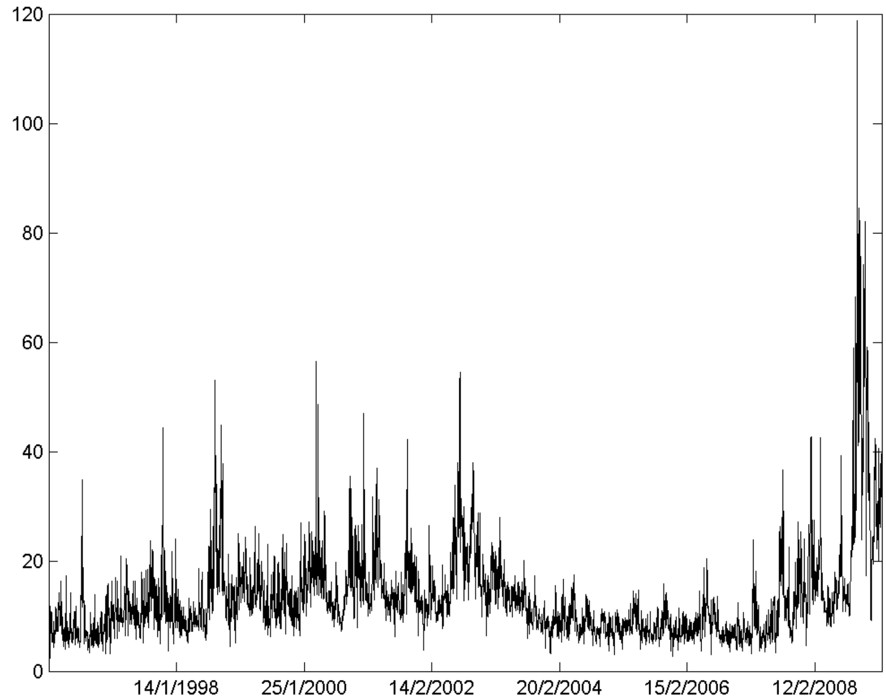
Volatility is a time varying concept. The problem of its measurement seems to have reached a mature stage: it is a diffused opinion that realized volatility is the best measure, being less subject to noise with respect to other measures, such as, for example, the squared returns (Andersen et al., 2000, 2003). The real problem, in a clustering framework, is to identify an indicator (or more) which characterizes the time series of volatility (see, for example, Wang et al., 2006). In our case, a simple way to represent the volatility level of a certain financial time series is the unconditional volatility, which is a measure derived from the model fitted to represent the dynamics of the conditional volatility (such as the family of GARCH model of Engle, 1982, and Bollerslev, 1986, or the most recent Multiplicative Error Model of Engle, 2002b). For this reason, in spite of the existence of several approaches in time series clustering (for a review see Liao, 2005), this kind of problem has a natural solution in a model based framework.

The typical model based approach to clustering time series hypothesizes that the series are generated by ARMA models and then a clustering algorithm, based on some distance measure specific to this framework, is applied to group models with similar characteristics. In time series analysis, distance measures were developed by Piccolo (1990), who proposed a distance between invertible ARMA models (AR metrics), whereas Corduas and Piccolo (2008) have studied the statistical properties of this distance, implementing a clustering algorithm for time series; Maharaj (1999) and (2000) extends the AR metrics to the case of dependent time series; Planas and Depoutot (2002) propose a distance between the filters used to extract the unobserved components of a time series; Caiado et al. (2006) suggest a periodogram-based metrics; D'Urso and Maharaj (2009) propose an autocorrelation-based fuzzy approach. In this framework, procedures based on tests of equality of parameters characterizing the ARMA models are frequently used; for example, Maharaj (1996, 1999, 2000) calculates the p-value of the statistic measuring the equality of the coefficients of every pair of the series analyzed and uses these results in an algorithm which follows the principles of hierarchical clustering. Otranto (2008) considers a starting benchmark series and then applies an agglomerative algorithm based on a test of the equality of some characteristics of the volatility with respect to a benchmark; an interesting property of this procedure is that the number of clusters is detected automatically and it is not determined by the user.

The classification of financial time series is a topic developed only recently. For example, Pattarin et al. (2004) combine different statistical techniques to develop an algorithm to classify mutual funds; Otranto (2008) proposes an extension of the AR metrics of Piccolo (1990) to the GARCH case, and develops a clustering algorithm to classify financial assets; Otranto (2010) defines the Dynamic Conditional Correlation (DCC) distance between the conditional correlation of a pair of series, to develop flexible and parsimonious DCC models (Engle, 2002a).

The model-based approach has been developed under the assumption that the parameters of the models are constant along the time. Such hypothesis is very strong, in particular in the analysis of volatility (see, for example, Mikosch and Stărică, 2004). The volatility of time series is frequently subject to changes in regime; for example in Figure 1 we show the dynamics of the realized kernel volatility of the S&P500 index from January 3,

Figure 1: Realized kernel volatility of S&P500 index.



1996 to February 27, 2009.[1] It is clear how the level of volatility changes along time with particular events as the Russian crisis of August 1998, the dot-com bubble (which has its peak on March 2000), the 2001 recession and, in particular, the long latest crisis which began in July 2007, starting with the collapse of the subprime mortgage industry; these events cause jumps in the dynamics of the series, implying data generating processes with changes in the coefficients. Gallo and Otranto (2011) have analyzed the series illustrated in Figure 1, proposing the Asymmetric Multiplicative Error Model with Markov Switching (AMEM-MS), extending the Asymmetric MEM (AMEM) of Engle and Gallo (2006) to the Markov Switching case. The AMEM-MS with three regimes fits the data better than the classical AMEM and captures the autocorrelation present in the residuals of the AMEM. The practical implications of capturing the changes in volatility are relevant: it is an important task for investors and traders, because it will imply changes in the portfolio composition along time, depending on the risk aversion of the financial operator (for a statistical analysis of this case see Coretto et al., 2011).

Given this empirical evidence, does the volatility clustering change when the regime is changed? More generally, if the model provides the possibility of changing parameters, how can we classify the series and will the results be different with respect to a model with fixed coefficients? To analyze these two perspectives, we propose the use of two separate models, the AMEM-MS of Gallo and Otranto (2011) and a new Smooth Transition AMEM (ST-AMEM), which introduces the change in the constant of the AMEM

[1]Data are taken from the *Oxford-Man Institute's realised library* version 0.1 (Heber et al., 2009).

according to a smooth transition function. It is worth noting that the two models respond to a different way to compare the volatilities of time series. The AMEM-MS provides an inference on the regime which could differ when time series change; for example, the period belonging to the regime interpreted as the one of high volatility for S&P500, could not be identical to the high volatility periods of DAX. In this case we would be interested in classifying the series in the presence of a certain regime (for example, following Gallo and Otranto, 2011, during periods of low, high or very high volatility). It is important to interpret the regime in a very precise way, so that the regime can be compared across the series.

Using the ST-AMEM, we could compare the model for each time unit; for example, it could be interesting to compare the models in correspondence of the day after a financial shock or an important world event, such as the terroristic attack of September 11. Moreover, the use of parameters that change along time explicitly considers the dependence on the particular time span, changing the estimation of the coefficients in correspondence with particular periods or events, whereas a constant coefficient model would make a sort of average of the coefficients, so that it is not robust to the time span considered for the estimation, depending on the level and the frequency of quiet and turmoil periods.

We will apply a clustering procedure based on statistical tests, using the unconditional volatilities derived by the AMEM, the AMEM-MS and the ST-AMEM. This approach has some similarities with the first part of the clustering procedure proposed by Otranto (2008), based on GARCH models. In our view the approach proposed here is simpler and more suitable in terms of volatility classification. It is simpler because we use the definition of unconditional volatility of an AMEM directly, whereas Otranto (2008) bases his algorithm on the AR(∞) representation of the squared returns, derived by the GARCH representation, implying a lot of coefficients. It is more suitable because the GARCH approach of Otranto (2008) is based on squared returns, which are a noisy estimator of the squared volatility, whereas here we deal with some proxy of the volatility which is robust to noise, as the realized kernel volatility (Barndorff-Nielsen et al., 2008).

The paper is organized as follows. In the next section we will describe briefly the models that will be used, whereas in Section 3 we will develop the clustering algorithm to obtain groups with similar volatility. The method will be applied in Section 4, using 15 series of realized kernel volatility relative to the main financial indices; the main purpose of this section is to illustrate how to perform the clustering procedure and the different information provided using different modelling approaches. Some final remarks will conclude the paper.

2 AMEM and Unconditional Volatilities

The MEM with asymmetric effects, called AMEM, is discussed in Engle and Gallo (2006), who develop the basic MEM idea of Engle (2002b). The volatility x_t of a certain financial time series is hypothesized to be the product of a time varying scale factor μ_t, representing the conditional mean of x_t, which follows a GARCH-type dynamics with threshold (Zakoïan, 1994), and a positive valued error ε_t:

$$
\begin{aligned}
&x_t = \mu_t \varepsilon_t, \qquad \varepsilon_t | \mathbf{\Psi}_{t-1} \sim Gamma(a, 1/a) \text{ for each } t = 1, \ldots, T \\
&\mu_t = \omega + \alpha x_{t-1} + \beta \mu_{t-1} + \gamma D_{t-1} x_{t-1} \\
&D_t = \begin{cases} 1 & \text{if} \quad r_t < 0 \\ 0 & \text{if} \quad r_t \geq 0 \end{cases}
\end{aligned}
\tag{2.1}
$$

where $\mathbf{\Psi}_t$ represents the information available at time t and r_t is the corresponding return observed at time t. The coefficient γ is a parameter introduced to take into account the different reactions of markets in the presence of negative returns, which, in general, imply higher level of the expected volatility at the next time period. The Gamma distribution is used because volatility is positive for each time; we hypothesize that this distribution depends only on a single parameter a, providing a mean and a variance of the conditional error equal to 1 and $1/a$ respectively. Correspondingly, the conditional mean and variance of x_t are μ_t and μ_t^2/a respectively. For the sake of simplicity, in the second equation of (2.1) we have used an AMEM(1,1), which is the most used specification of the AMEM, but it is possible to extend the results to a generical AMEM(p,q). In (2.1) $\omega > 0$, $(\alpha, \beta, \gamma) \geq 0$ to ensure a sufficient condition for the positiveness of μ_t, and the constraint $(\alpha + \beta + \gamma/2) < 1$ is imposed for stationarity. The unconditional volatility, analogously to the GARCH case (Bollerslev, 1986) is given by:

$$
u = \frac{\omega}{1 - \alpha - \beta - \gamma/2} \tag{2.2}
$$

We call u the *long time unconditional volatility*, in the sense that it is referred to the full time span analyzed with the AMEM.

The recent success of this approach, an alternative to the GARCH methodology, is due to the possibility to directly model non negative processes, common in finance, as the processes representing the volatility. This fact allows to model the volatility (which is not observable) using the proxies not affected by noise, as the realized volatility (Andersen et al., 2000, 2003) or, following the most recent studies, proxies which consider the presence of microstructure, as the realized kernel volatility (Barndorff-Nielsen et al., 2008); a review of parametric and nonparametric measures of volatility is given in Andersen et al. (2010). This is an enormous advantage with respect to the GARCH methodology, which, in general, deals with squared returns, not possessing the properties described above; at the same time it is possible to use softwares for GARCH estimation to estimate an AMEM (see Engle, 2002b).

The presence of a parameter of asymmetry in (2.1) is not sufficient to capture the different phases of volatility, which characterize a series such as the one shown in Figure 1. Recently, Gallo and Otranto (2011), analyzing the same S&P500 series, have disputed the idea that these phases, which, from a visual inspection, can be classified as quiet periods, turmoil phases and brief abnormal peaks, correspond to changes in regimes in the AMEM process generating the data. They introduce parameters that can change, according to a discrete Markov chain, along the lines of a Markov Switching (MS) model (Hamilton, 1990). The AMEM-MS we will use in this framework is defined as:

$$\begin{aligned} x_t &= \mu_{t,s_t}\varepsilon_t, \qquad \varepsilon_t|\mathbf{\Psi}_{t-1} \sim Gamma(a_{s_t}, 1/a_{s_t}) \quad \text{for each } t \\ \mu_{t,s_t} &= \omega + \textstyle\sum_{i=1}^{n} k_i I_{s_t} + \alpha x_{t-1} + \beta\mu_{t-1,s_{t-1}} + \gamma D_{t-1}x_{t-1} \end{aligned} \tag{2.3}$$

where s_t is a discrete latent variable which ranges in $[1,\ldots,n]$, representing the regime at time t. It is hypothesized that s_t follows a Markovian dynamics, represented by a transition probability matrix with elements $p_{\tau\iota} = Pr[s_t = \iota|s_{t-1} = \tau]$ and $p_{\tau n} = 1 - \sum_{j=1}^{n-1} p_{\tau j}$ $(\tau, \iota = 1,\ldots,n)$.

Some considerations about the other coefficients could help to interpret the regimes. Firstly, I_{s_t} is an indicator equal to 1 when $s_t \leq i$ and 0 otherwise; $k_i \geq 0$ and $k_1 = 0$. In this way the constant in regime 1 is given by ω, in regime 2 by $(\omega + k_2)$, in regime 3 by $(\omega + k_2 + k_3)$ and so on. A more general model would consider, in the second equation of (2.3), also the coefficients α, β and γ as regime dependent; anyway, if the only switching coefficient is the constant, the reparameterization used ensures that the level of the volatility (from low to high) increases with the label of the regime, and, as a consequence, the same happens for the *unconditional volatility within the regime* ι, defined as:

$$u_{MS}^{\iota} = \frac{\omega + \sum_{i=1}^{\iota} k_i}{1 - \alpha - \beta - \gamma/2} \tag{2.4}$$

Secondly, also the a coefficient is considered as switching because it includes a larger range of values in the periods of high volatility with respect to the periods of lower volatility. For example, if we think that the several peaks present in the series can be considered prevalently as periods of high volatility, we note that values between 30 and 120 can belong to this regime, whereas the quiet periods show values contained in the range 0-20. This consideration implies different variances within each regime, characterized by smaller and higher values of a in high and low volatility periods respectively.

Thirdly, as observed in Gallo and Otranto (2011), it is useful to impose a particular reparameterization for β in (2.3) to guarantee a degree of coherence between the regime and the level of volatility. In practice, to avoid frequent and not realistic changes in regime, they reparameterize $\beta = \beta^* - \alpha - \gamma/2$, with $(\alpha + \gamma/2) \leq \beta^* < 1$; so they suggest to estimate β^* and then obtain β from it. We refer to Gallo and Otranto (2011) for details about the derivation of this particular reparameterization.

The use of the AMEM-MS with the previous parameters specification is fundamental to classify all the series analyzed in n regimes which can be interpreted in the same way (from low to high volatility), even if different series are classified in different regimes at the same time t. For example, a spillover effect which starts from a *dominant* market would influence the other markets with some lags, so the change in regime will happen at different dates. For this reason and for the purposes illustrated in the first section, we need the regimes to be interpreted in a clear manner.

Another perspective is to consider that the unconditional volatility could change at each instant in time. In practice, an alternative (more flexible) approach could be obtained considering an AMEM in which the constant coefficient is time varying in each period and not only within a certain regime. Of course, such hypothesis requires some reparameterization of ω. A parsimonious and diffused idea considers smooth transitions between regimes, introducing the smooth transition functions, along the lines of Chan and Tong (1986) for the simpler autoregressive case. This kind of models has been adopted in the financial framework extending it to the GARCH case (Gonzales-Rivera, 1998, Anderson et al., 1999, Lanne and Saikkonen, 2005). We propose a similar extension to the AMEM case, providing a smooth transition function depending on the lagged value of the volatility. A simple form is given by:

$$\begin{aligned} & x_t = \mu_t \varepsilon_t, \qquad \varepsilon_t | \Psi_{t-1} \sim Gamma(a, 1/a) \text{ for each } t \\ & \mu_t = \omega_1 + \omega_2 g_t + \alpha x_{t-1} + \beta \mu_{t-1} + \gamma D_{t-1} x_{t-1} \\ & g_t = (1 + exp(-\delta(x_{t-1} - c)))^{-1} \end{aligned} \tag{2.5}$$

where g_t is the smoothed function and ω_1, ω_2, δ, c are new unknown coefficients. We call model (2.5) the Smooth Transition AMEM (ST-AMEM). Unlike the AMEM-MS case, in this model we do not need to consider the a coefficient also as time-varying, because the constant changes in each point in time, adapting to the current level of the volatility. The unconditional volatility will change at each point in time and will be equal to:

$$u_{ST}(t) = \frac{\omega_1 + \omega_2 g_t}{1 - \alpha - \beta - \gamma/2} \tag{2.6}$$

We call the series $u_{ST}(t)$ the *time-varying unconditional volatility*.

The three models considered, the AMEM, the AMEM-MS and the ST-AMEM, can be estimated by maximum likelihood, obtaining also the asymptotic covariance matrix of the estimators. For the invariance property, the maximum likelihood estimators of the unconditional volatility will be obtained substituting in (2.2), (2.4), (2.6), the maximum likelihood estimators of the parameters on which the unconditional volatility depends. Also, the standard errors of the unconditional volatilities can be obtained by:

$$\boldsymbol{d}' \boldsymbol{\Sigma} \boldsymbol{d} \tag{2.7}$$

where $\boldsymbol{\Sigma}$ is the asymptotic *sandwich* covariance matrix of the maximum likelihood estimator of the vector of parameters $\boldsymbol{\theta}$ and $\boldsymbol{d}$ is the vector of the partial derivatives of the unconditional volatility with respect to the elements of $\boldsymbol{\theta}$.

3 The Clustering Procedure

Let us suppose to have s financial time series; for each one we estimate an AMEM model as (2.1), so that we obtain s estimated unconditional volatilities (2.2). Our purpose is to obtain clusters of financial time series with similar unconditional volatility levels. The procedure proposed is a test-based agglomerative algorithm, similar to that one proposed in Otranto (2008), which has the advantage of automatically providing the number of groups forming the clustering.

The procedure is based on a simple Wald test to verify the equality of r unconditional volatilities; the null hypothesis is:

$$H_0 : u^{(1)} = u^{(2)} = \cdots = u^{(r)} \tag{3.1}$$

where the superscript is referred to the time series. The null hypothesis (3.1) can be rewritten as:

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{0}, \tag{3.2}$$

where $\boldsymbol{u}$ is the vector containing the r unconditional volatilities and $\boldsymbol{A}$ is the $(r-1)\times r$ 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 statistic to test the null hypothesis (3.2) is given by:

$$W = (\boldsymbol{A}\widehat{\boldsymbol{u}})'(\boldsymbol{A}\boldsymbol{D}\widehat{\boldsymbol{H}}\boldsymbol{D}'\boldsymbol{A}')^{-1}(\boldsymbol{A}\widehat{\boldsymbol{u}}). \tag{3.3}$$

where $\widehat{\boldsymbol{H}}$ is a block diagonal matrix with blocks constituted by the estimated covariance matrices of the vector of parameters and $\boldsymbol{D}$ is the matrix composed by the row vectors $\boldsymbol{d}^{(i)\prime}$ of the partial derivatives of each function $u^{(i)}$ $(i = 1, \ldots, r)$ with respect to the full set of parameters $\boldsymbol{\theta} = (\omega^{(1)}, \alpha^{(1)}, \beta^{(1)}, \gamma^{(1)}, \ldots, \omega^{(r)}, \alpha^{(r)}, \beta^{(r)}, \gamma^{(r)})'$. W is asymptotically distributed as a central chi-square random variable with $(r-1)$ degrees of freedom.

The clustering algorithm is composed by the following steps:

1. order the series by increasing unconditional volatility and choose the series with lowest u as benchmark to form the first cluster;

2. verify if the successive series has an unconditional volatility not significantly different from the one of the benchmark, using the Wald test at a fixed size, based on the statistic (3.3) with $r = 2$;

3. if the null hypothesis is accepted, add the successive series, put $r = r + 1$ and verify the null hypothesis (3.2) using the Wald test; repeat the step 3 until the null hypothesis is rejected, put the last series as the benchmark for the successive cluster and go to step 2.;

4. continue until no series remain.

Notice that the number of clusters is selected automatically from the procedure, depending on the number of times in which the null hypothesis (3.2) is rejected. Moreover the interpretation of the clusters is quite simple: the first cluster obtained represents the group of series with lowest (and similar) unconditional volatility and so on until the last group, which is the one with highest unconditional volatility.

The choice of the test size is subjective; Otranto (2008), in a similar procedure, suggests a size equal to 0.01, based on a set of simulation experiments. This choice is crucial

Table 1: Realized kernel volatility indices object of the analysis

Index	Symbol	Starting date	# observations
Dow Jones Industrials	DJ	3 Jan 1996	3261
CAC 40	CA	3 Jan 1996	3301
FTSE 100	FT	21 Oct 1997	2844
Spanish IBEX	IB	3 Jan 1996	3270
NASDAQ 100	NA	3 Jan 1996	3262
Italian MIBTEL	IM	4 Jul 2000	2176
S&P 400 Midcap	SP4	3 Jan 1996	3258
Nikkei 250	NI	8 Jan 1996	3160
Russell 3000	R3	3 Jan 1996	3262
Russell 1000	R1	3 Jan 1996	3262
Russell 2000	R2	3 Jan 1996	3264
Milan MIB30	MM	3 Jan 1996	3289
German DAX	DA	3 Jan 1996	3296
S&P TSE	SPT	4 Jan 1999	2529
S&P 500	SP5	3 Jan 1996	3263

in the final clustering performance and could produce some puzzling cases; for this reason we will base the choice of the test size on a quality clustering index, that will be described in the next section, using an empirical example.

The previous procedure is described for the AMEM case, in which a single unconditional volatility is estimated for each time series. In the AMEM-MS case we can calculate n different unconditional volatilities, one for each state, and apply n distinct clustering procedures as the previous one.[2] For example, having $n = 3$, as in the example of the next section, the states can be interpreted as low, high and very high volatility, and we will obtain three different classifications of the s time series, conditional on the regime.

Finally, it is possible to obtain a classification of the volatility for each point in time, applying the clustering algorithm at a given time t to the unconditional volatilities (2.6); this information could be particularly useful to evaluate how the behavior of a single index or asset changes in terms of volatility, in correspondence with particular events.

4 An Application to Realized Kernel Volatilities

In order to show how the classification of the volatility level changes considering the different perspectives illustrated in section 2, we apply the clustering algorithm, with the alternative three models, to a set of realized kernel volatilities relative to the 15 main stock indices in the world, available at the Oxford-Man Institutes realised library. The original data are relative to the daily realized kernel variance; we consider, as usual, the annualized volatility, which is the squared root of the variance multiplied by 252 (and multiplying the result by 100). The indices present similar patterns, showing an enormous jump in the

[2]Obviously, in the AMEM-MS case, the vector of coefficients $\boldsymbol{\theta}$ contains also the constants $k_j^{(i)}$ ($j = 2, \ldots, n$; $i = 1, \ldots, r$) (see equation (2.4)).

Table 2: Descriptive Statistics for 15 realized kernel volatility indices

Index	Mean	Median	Min	Max	st.dev.	skewness	kurtosis
DJ	12.90	11.15	2.92	123.14	7.80	3.82	30.20
CA	15.81	13.62	2.91	95.81	9.33	2.47	13.58
FT	13.37	11.75	3.03	91.95	8.06	2.77	17.31
IB	14.31	12.73	2.91	82.61	8.35	2.26	12.65
NA	15.71	12.86	2.37	124.71	10.27	2.35	13.36
IM	11.37	9.27	2.20	86.35	7.82	2.71	15.83
SP4	11.02	9.13	2.06	100.96	8.25	3.92	25.60
NI	15.01	13.85	3.70	85.69	6.98	2.65	18.30
R3	12.18	10.26	2.33	114.73	8.02	3.57	25.22
R1	12.52	10.61	2.18	115.86	8.17	3.47	24.17
R2	10.51	8.85	1.59	113.06	8.43	3.62	24.91
MM	14.41	12.17	2.58	96.66	9.02	2.44	13.31
DA	18.17	15.39	2.76	133.58	11.28	2.54	15.00
SPT	11.72	9.27	2.40	98.46	8.29	3.44	21.40
SP5	13.36	11.39	2.39	118.75	8.61	3.41	23.12

last part of the series, corresponding to the last world economic crises, as the $S\&P500$ index depicted in Figure 1. The series considered are illustrated in Table 1; we show the symbols that we will use hereafter, the period covered by the data set, the number of observations considered for each one. All the series end on 27 February 2009, but the starting date is different, so that also the number of observations is different. This fact does not imply technical problems in our clustering algorithm because the classification is based on the unconditional volatility of each time series, or, in other words, on the parameters of the models used, which are estimated separately for each time series in a univariate framework. The length of the series affects the degree of uncertainty in the estimation of the variance of the unconditional volatility; anyway, the minimum length available is 2176 (IM index), which is sufficiently large to guarantee the consistency of the estimators. From another point of view, it is clear that the consideration of different spans could imply a different parameter estimation in the AMEM case; for example, if a series starts at the beginning of 2007, the period considered is one of turmoil and the coefficients would highlight high unconditional volatility, whereas if the time span includes the data from 1996, quiet and turmoil periods alternate and the unconditional volatility would represent a sort of average of the different levels of volatility. This is not the case of our data set, in which most of the series start at the beginning of 1996 and only SPT and IM start three and three years and half after, respectively, so that they include different levels of volatility.

The data set considered have similar characteristics, as shown in Table 2, where the main descriptive statistics are illustrated. All the series show a clear positive asymmetry with very high peaks, in particular the American indices (DJ, NA, SP4, R3, R1, R2 and SP5) and the German DAX, with maximum values more than 100. The presence of a positive skewness is evidence in favour of the consideration of asymmetric models, whereas the large differences between the minimum and the maximum, jointly with the very large

Table 3: Unconditional volatilities estimated by the AMEM and corresponding standard errors (in parentheses)

	DJ	CA	FT	IB	NA	IM	SP4	NI
unc. vol.	12.24	14.53	12.37	13.44	14.60	9.99	10.30	14.50
	(0.03)	(0.06)	(0.35)	(0.07)	(0.10)	(0.05)	(0.04)	(0.04)
	R3	R1	R2	MM	DA	SPT	SP5	
unc. vol.	11.31	11.66	9.73	13.39	16.64	10.62	12.56	
	(0.03)	(0.03)	(0.08)	(0.10)	(0.09)	(0.03)	(0.05)	

kurtosis, justify the consideration of distinct regimes or time-varying coefficients.

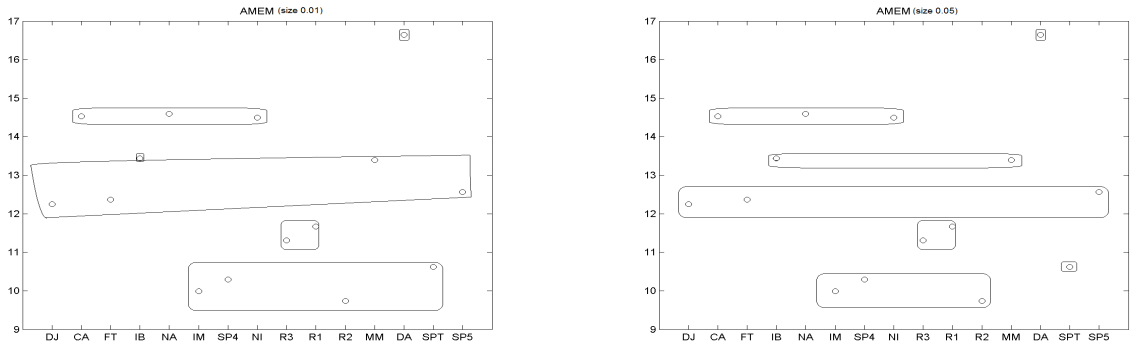
4.1 Clustering the long time unconditional volatilities

The first experiment that we propose is relative to the clustering of the unconditional volatilities of the 15 financial indices obtained with the estimation of 15 univariate AMEM's. To save space we do not show the estimation of the 5 parameters of model (2.1) for each time series, but directly the estimation of the unconditional volatilities (Table 3) with the corresponding standard errors.

The clustering procedure is based on a statistical test applied iteratively, so the choice of the size of the test is crucial for the final clustering. In fact, in spite of the good performance of the proposed Wald test in terms of power and respect of the nominal size (as shown in the simulation experiments of Otranto, 2008), also when we verify the equality of several unconditional volatilities simultaneously, the change of the nominal size could change the assignment of a series to a group rather than another. For example, considering the data in Table 3, let us perform the clustering algorithm with two different size levels: 0.01 and 0.05. In Figure 2 we show the two different classifications (the y axis represents the level of the unconditional volatility). The first classification (the one with size equal to 0.01) provides six clusters, whereas the second one (with size equal to 0.05) implies seven groups. The differences consist only of the classification of SPT and MM. SPT is included in the group of lowest volatility in the first classification, whereas it constitutes the only element of the second group in the second clustering. MM belongs to the group formed by DJ, FT and SP5 in the first clustering, whereas it constitutes a separate group with IB in the second clustering. At a simple visual inspection, the second classification seems more plausible then the first one, in particular for the location of MM; in fact its level of unconditional volatility is in line with the level of IB, but, in the first clustering they belong to distinct groups, providing a *leaning form* of the cluster, whereas in the second clustering the series belonging to the same group are sufficiently aligned.

From this empirical evidence we can deduce that a criterion for the choice of the best size could be the quality of the clustering, so that we can choose the size which provides the highest quality, based on some measure. For this purpose several quality indices have been proposed in literature; see Theodoridis and Koutroumbas, 1998, or Jain et al., 1999, for a review. We choose the C-index (Hubert and Schultz, 1976), which is largely diffused in the clustering literature and possesses the nice characteristic to be limited to the interval $[0, 1]$ (the lower the index the higher the quality), so it can be easily interpreted. It is

Figure 2: Clustering of the long time unconditional volatilities, based on the AMEM, with two different levels of size.



calculated as:

$$C = \frac{S - S_{min}}{S_{max} - S_{min}}. \tag{4.1}$$

Let us indicate with ν the number of all pairs belonging to series which are both included in the same cluster; S is the sum of distances between series in the ν pairs (we will use the Euclidean distance); S_{min} and S_{max} are the sum of the ν smallest distances and the sum of the ν largest distances respectively, considering all possible pairs in the universe of series.

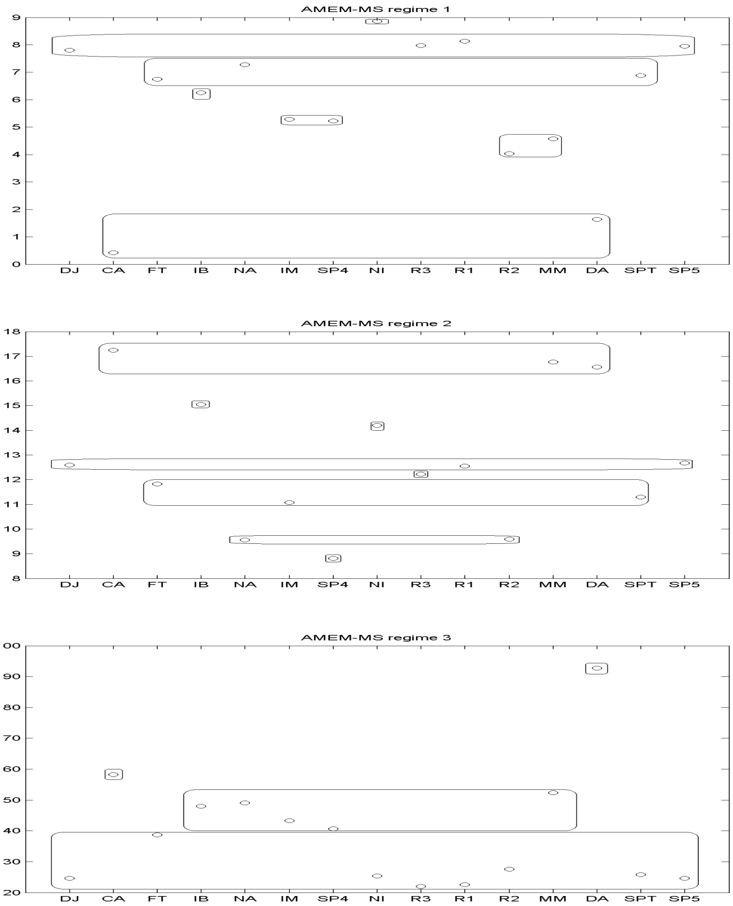
Using the C-index (4.1) in the previous example, it is equal to 0.025 in the clustering with size 0.01 and 0.004 in the clustering with size 0.05 (reaching also a very high quality). Moreover, performing the clustering with all the levels of size between 0.01 and 0.10 (with steps of 0.01), the lowest C-index is equal to 0.004, obtained for the size levels included in the interval [0.02,0.10]. In practice, the criterion adopted selects the clustering shown in the right part of Figure 2; we can notice as the series with lowest unconditional volatility levels are IM, SP4 and R2; it is interesting to note that SP4 is calculated on 400 companies of the mid-cap equities sector, whereas R2 are the last 2000 companies among the first 3000 companies with larger capitalization (considered in R3). In practice, the indices with no blue chips companies are present in the group with lowest volatility, whereas the most important European indices (DAX and CAC) show the highest volatility.

These results are relative to the full spans illustrated in Table 1, so they include periods characterized by regimes with different degree of conditional volatility (from low to high); but what happens within each regime? Is there some peculiar behaviour of some index that is hidden in the long time classification? We can answer this question by repeating the clustering procedure within each regime.

4.2 Clustering the unconditional volatilities within each regime

In this second stage we estimate an AMEM-MS, as (2.3), for each time series, obtaining different unconditional volatilities, one for each regime, as indicated in (2.4). The number n of regimes is generally chosen a priori; it can be identified preliminarily (for example,

Figure 3: Clustering of unconditional volatilities within each regime, based on the AMEM-MS.



using the nonparametric Bayesian procedure of Otranto and Gallo, 2002) or ex post (for example, using loss functions, as in Psaradakis and Spagnolo, 2003). In our case, in order to compare the unconditional volatilities within each regime, we need to select the same number of regimes for each series; at a visual inspection (we recall that all the series have a similar behaviour to the one illustrated in Figure 1), it is easy to notice the presence of alternating periods of low and then higher volatility with fewer periods with very high volatility. For this reason we put $n = 3$; this choice is also consistent with the results of Gallo and Otranto (2011), who select the number of regimes for the series SP5 of Figure 1, using AIC and BIC.

After this, we apply the clustering algorithm for u^1_{MS}, u^2_{MS} and u^3_{MS} respectively, using, for each case, the criterion based on the C-index to select the best size of the test; in particular the highest quality performance for regimes 1 and 2 is obtained in the size interval [0.08-0.1] (C-index equal to 0.009 and 0.012 respectively) and for regime 3 in the size interval [0.01-0.06] (C-index equal to 0.036). The results are illustrated in Figure 3.

The details of the behaviour of the series in each regime provides new information, which is able to distinguish the characteristics of time series classified in the same group in terms of long time unconditional volatility. Let us fix the attention on the series previously considered, the least and the most volatile respectively. We have noted that IM, SP4 and R2 belong to the group with the lowest long time unconditional volatility; in terms of unconditional volatility within the regimes we notice that they are classified in groups with medium unconditional volatility in regime 1, whereas, increasing the regime, IM seems to stay in groups of medium unconditional volatility, whereas SP4 and R2 are among the less volatile series. In practice, recalling the composition of SP4 and R2, the role of mid capitalization indices is evident in the regimes of high and very high conditional volatilities, when the turmoil spreads across the markets, differently from IM, which shows a more volatile pattern, especially in regime 3, where it belongs to the same group as NA.

Considering the four series with the highest long time unconditional volatility (CA, NA, NI, DA), it is possible to notice that the two European indices behave the opposite way to the other two. CA and DA belong to the group with the lowest unconditional volatility in regime 1, but their volatility increases dramatically in the other two regimes (in particular DA in regime 3). NA and NI are more volatile in the regime of low conditional volatility, whereas NA is classified in the lowest unconditional volatility group during regime 2 and NI shows a similar behaviour during regime 3.

Another interesting aspect that stems from this analysis is that DJ, R1 and SP5, which consider different companies but with similar characteristics, always belong to the same group in each regime.

We recall that, when clustering within the regimes, we do not compare the series in the same subperiods, because they do not necessarily belong to the same regime at time t. The regime characterizes each financial series, in the sense that the change in regime happens in different days for different series. For example, the dominant markets, like the US market, change regimes before the other markets, which follow its behaviour with a spillover effect (see, for example, Theodossiou and Lee, 1993, and Gallo and Otranto, 2008). If we are interested to know how to classify the series at a certain date, we need to use the idea of time varying unconditional volatility.

4.3 Clustering the time varying unconditional volatilities

As said, the purpose of this third kind of classification is to evidence how the unconditional volatility classification changes along time, especially during some periods of particular interest, such as periods of economic crises and financial turmoil. The estimation of the ST-AMEM implies that at each time we have a different estimation of the unconditional volatility (2.6), so potentially it is possible to obtain T classifications, one for each date. Anyway, what is interesting is the dynamics of the clustering at particular time intervals, for example starting at the beginning of a crises, to evaluate how the different markets react to financial shocks and if there is a different timing in this reaction.

We focus on the period relative to the shock of September 11, which only impacted for a few days in the financial markets but with strong effects. We have selected the classifications relative to 7 September 2001, the last day in which the American markets were open before the terroristic attack of September 11, and 18 September 2001, the day after the date in which the financial markets were reopened.[3] In Figure 4 we show the levels

[3]We have used the day after to consider the spillover behavior of the first day in which the American

of unconditional volatilities, derived by ST-AMEM, in the two days, marking with white circles the data relative to September 7 and with black circles the data relative to September 18. It is evident that all the series show an increase in volatility, except NA and NI, which start from a high level already before September 11. The only index that does not change its volatility level is R2 (white and black circles are superposed). The lines join the series belonging to the same cluster (the thin lines are referred to September 7 and the bold lines to September 18). It is interesting to note that the series with low volatility on September 7, when R2 and SPT belonged to the same cluster, in September 18 are clearly separated, with large increase in volatility of SP4 and SPT. The series characterized by medium-high volatility in September 7 belong to two separate clusters: they contain series with an average unconditional volatility equal to 22.36 and 26.39 respectively. On September 18 they increase their volatility (except NI) and collapse in one cluster with average unconditional volatility level equal to 25. The only exception, on September 7, is CA, which increases its volatility and reaches the higher level of volatility of NA. On September 18, DA, which belonged to the highest level of volatility, as NA, further increases its unconditional volatility, which is very significantly different with respect to the other indices. In practice, this kind of analysis evidences not only the obvious increase of the volatility of almost all the indices, but also the size of these movements, which can be compared, and the changes in classification of the volatility indices after an abrupt and unexpected event.

4.4 Consistency of the classifications

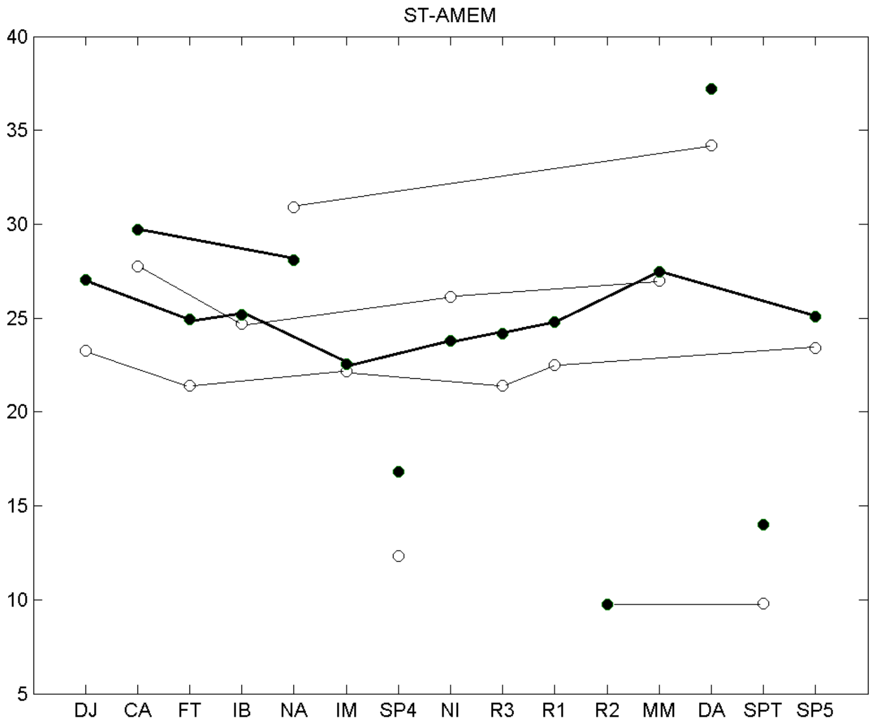
Given the three alternative ways to classify the time series in terms of unconditional volatilities, a natural question is if the different criteria are consistent, or, in other words, if the use of different models to estimate the unconditional volatility provides information which is coherent in the three different frameworks or not. We have interpreted the global unconditional volatility as a sort of average of unconditional volatilities within the regimes or the average of all the time varying volatilities; if this aspect is confirmed, we can conclude that the different criteria work coherently.

To reach this purpose, we simply compare each long time unconditional volatility with the corresponding average of unconditional volatilities within the three regimes and with the average of the T unconditional time varying volatilities. Whereas the last one could be calculated as a simple average of the T unconditional volatilities $u_{ST}(t)$ in (2.6), the average of the three unconditional volatilities within the regimes is more complicated because the duration and the frequency of each regime changes in the time interval considered. It is preferable to consider a weighted mean of the unconditional volatilities for each time t with weights given by the probabilities of each regime at time t. For this purpose, using the Hamilton (1990) filter and smoothing, we can obtain the smoothed probabilities $Pr(s_t|\mathbf{\Psi}_T)$ for each time t; the weighted average at time t of the unconditional volatilities within each regime is given by:

$$\bar{u}_{MS}(t) = \sum_{i=1}^{n} Pr(s_t = i|\mathbf{\Psi}_T)u_{MS}(i) \tag{4.2}$$

markets were re-open on the European countries, for effect of the different time zone.

Figure 4: Clustering time varying unconditional volatilities, based on ST-AMEM, on 7 September 2001 (white circles) and 18 September 2001 (black circles). The thin line joins the series belonging to the same cluster in the first date, the bold line joins the series belonging to the same cluster in the second date.



An average of the T unconditional volatilities $\bar{u}_{MS}(t)$ will provide an (alternative) estimation of the long time unconditional volatility of a certain series.

In Table 4 the three unconditional volatilities provided by each model and for each series are shown. It is possible to note the similar pattern, confirmed by a very high linear correlation (shown at the bottom of the table), more than 0.99 for each pair of models. The three models provide very similar values for the unconditional volatilities, as it is possible to note, in the last two columns of the table, the mean of the three values and the small standard deviation. In general, the AMEM-MS model provides an intermediate value of volatility between the one of the AMEM (the lowest) and the one of the ST-AMEM (the highest). Moreover, calculating a simple Wald statistics to evaluate the equality of a pair of unconditional volatilities, estimated by two different models, we notice that the AMEM-MS provides often the same unconditional volatility of the AMEM (pointed out by the superscript a) and/or the ST-AMEM (superscript c). The differences between the unconditional volatilities provided from the AMEM and the ST-AMEM (superscript b) are not significant only in five cases. It is also interesting to note that for each series the near equality of two unconditional volatilities provided by different models is verified.

This simple exercise supports the idea that the three models provide coherent information in terms of unconditional volatility, so the differences in the alternative clustering are

Table 4: Comparison of the long time unconditional volatilities obtained from three models.*

	Model				
Index	AMEM	AMEM-MS	AMEM-ST	Mean	St.dev.
DJ	12.24^{ab}	12.55^{ac}	12.74^{bc}	12.51	0.20
CA	14.53^{a}	15.03^{a}	15.55	15.04	0.41
FT	12.37^{ab}	12.78^{ac}	13.20^{bc}	12.78	0.34
IB	13.44^{a}	13.88^{a}	14.18	13.83	0.31
NA	14.60^{a}	15.30^{ac}	15.56^{c}	15.15	0.41
IM	9.99	10.77^{c}	11.07^{c}	10.61	0.46
SP4	10.30^{a}	10.77^{ac}	10.90^{c}	10.66	0.26
NI	14.50^{ab}	14.69^{ac}	14.86^{bc}	14.68	0.15
R3	11.31^{a}	11.68^{ac}	11.91^{c}	11.63	0.25
R1	11.66^{a}	12.04^{ac}	12.27^{c}	11.99	0.25
R2	9.73^{ab}	10.12^{ac}	9.73^{bc}	9.86	0.18
MM	13.39^{a}	14.03^{ac}	14.25^{c}	13.89	0.36
DA	16.64	17.68^{c}	17.98^{c}	17.43	0.57
SPT	10.62	11.38^{c}	11.60^{c}	11.20	0.42
SP5	12.56^{ab}	12.93^{ac}	13.11^{bc}	12.86	0.23
	Correlation coefficient between:				
	AMEM	AMEM-MS	0.994		
	AMEM	ST-AMEM	0.991		
	AMEM-MS	ST-AMEM	0.997		

*The superscripts a, b, c indicate that the difference between the long time unconditional volatility derived from AMEM and AMEM-MS, AMEM and ST-AMEM, AMEM-MS and ST-AMEM, respectively, is not significant using a Wald statistics at a nominal size equal to 0.01.

due to the different points of view and the different time spans: the full period; the periods of low, high and very high volatility; specific dates. Coherence is an important property if the three analyses are conducted in the same framework to guarantee the comparability of the results.

5 Concluding Remarks

The main purpose of this paper was to show how classifications could change, considering time series subject to changes in regime or, more in general, changes in the value of the parameters characterizing the data generating models. For this purpose we have used a model-based approach involving a clustering procedure based on a test statistic. The choice of the model could be considered subjective; in this framework we have adopted the AMEM class because of its recent success in the literature on volatility. In fact, it seems common opinion that models dealing directly with realized volatility are preferred to models which use the log-transformation to deal with data without constraints on the sign of the variable studied. Anyway the approach proposed could be easily adapted to other families of models, such as the GARCH one, with its extensions to the Markov Switching (see, for example, Dueker, 1997, and Klaassen, 2002) and the Smooth Transition (Gonzales-Rivera, 1998, Anderson et al., 1999, Lanne and Saikkonen, 2005) cases.

Similarly, this kind of approach could be used also in non financial applications, employing other popular models, belonging, for example, to the ARMA family. The reason why we have developed this idea in a financial framework is the basic role of the classification of assets in a certain time interval for financial traders; the support of statistics in this framework is fundamental and its importance in financial operations is increasing.

The clustering algorithm is performed estimating s univariate AMEM's (or AMEM-MS's or ST-AMEM's) and then comparing their coefficients. Maybe the multivariate modelling would be a more suitable and elegant framework for this kind of analysis, but at the cost of unfeasible models due to the large number of parameters to be estimated. This kind of problem is typical in the multivariate modelling for financial time series and remains an open problem; the solutions proposed are often based on hypotheses which are too strong (see Bauwens et al., 2006). For this reason and to provide a simple instrument for practical purposes we have preferred to deal with univariate models.

Future works could be devoted to identifying other characteristics of financial time series, together with unconditional volatility, to perform a characteristic-based clustering, in a similar way to the methodology proposed by Wang et al. (2006).

References

[1] Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P. (2000). Great realizations, *Risk* 13, 105-108.

[2] Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P. (2003). Modeling and forecasting realized volatility, *Econometrica* 71, 579-625.

[3] Andersen, T.G., Bollerslev, T., Diebold, F.X. (2010). Parametric and nonparametric volatility measurement, in Aït-Sahalia, Y., Hansen, L.P. (eds.): *Handbook of Financial Econometrics*, pages 67-138, North-Holland. Amsterdam.

[4] Anderson, H.M., Nam, K., Vahid, F (1999). Asymmetric nonlinear smooth transition GARCH models, in Rothmann, P. (ed.): *Nonlinear Time Series Analysis of Economic and Financial Data*, 191-207, Kluwer. Boston.

[5] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A, Shephard, N. (2008). Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise, *Econometrica* 76, 1481-1536.

[6] Bauwens, L., Laurent, S., Rombouts, J.V.K. (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics* 21, 79-109.

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

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

[9] Chan, K. S. and Tong, H., 1986. On estimating thresholds in autoregressive models. *Journal of Time Series Analysis*, 7, 178-190.

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

[11] Coretto, P., La Rocca, M., Storti, G., 2011. Group structured volatility. In Ingrassia, S., Rocci, R., Vichi, M. (Eds.) *New Perspectives in Statistical Modeling and Data Analysis*, in press. Springer.

[12] D'Urso P., Maharaj E.A. 2009 Autocorrelation-based fuzzy clustering of time series. Fuzzy Sets and Systems 160:3565-3589

[13] Dueker, M.J., 1997. Markov Switching in GARCH Processes and Mean-Reverting Stock-Market Volatility. Journal of Business and Economic Statistics, 15, 26–34.

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

[15] Engle, R.F., 2002a. 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.F., 2002b. New frontiers for ARCH models, Journal of Applied Econometrics, 17, 425-446.

[17] Engle, R.F., Gallo, G.M., 2006. A multiple indicators model for volatility using intra-daily data, Journal of Econometrics, 131, 3-27.

[18] Gallo, G.M., Otranto, E. (2008). Volatility Spillovers, Interdependence and Comovements: A Markov Switching Approach. Computational Statistics and Data Analysis 52, 3011-3026.

[19] Gallo, G.M., Otranto, E., 2011. Modeling realized volatility subject to changes of regime. Presented to ERCIM 2010, London, 10-12 December.

[20] Gonzales-Rivera, G., 1998. Smooth transition GARCH models. *Studies in Nonlinear Dynamics and Econometrics* 3, 161-178.

[21] Hamilton, J.D., 1990. Analysis of time series subject to changes in regime. Journal of Econometrics, 45, 39-70.

[22] Hubert, L. and Schultz, J., 1976. Quadratic assignment as a general data-analysis strategy. British Journal of Mathematics, Statistics and Psychology, 29, 190 241.

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

[24] Klaassen, F. 2002. Improving GARCH Volatility Forecasts with Regime-Switching GARCH. Empirical Economics, 27, 363–394.

[25] Lanne, M., Saikkonen, P., 2005. Nonlinear GARCH models for highly persistent volatility. *Econometrics Journal* 8,251-276.

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

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

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

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

[30] Mikosch, T. and Stărică, C. 2004. Nonstationarities in financial time series, the long-range depeendence, and the IGARCH effects. The Review of Economics and Statistics 86, 378-390.

[31] Otranto, E., 2008. Clustering heteroskedastic time series by model-based procedures. Comput Stat Data Anal, 52:4685-4698.

[32] Otranto, E., 2010. Identifying Financial Time Series with Similar Dynamic Conditional Correlation. Computational Statistics and Data Analysis 54, 1-15.

[33] Otranto, E., Gallo, G.M. (2002). A Nonparametric Bayesian Approach to Detect the Number of Regimes in Markov Switching Models. Econometric Reviews, 21, 477-496.

[34] Pattarin, F., Paterlini, S. and Minerva, T., 2004. Clustering financial time series: an application to mutual funds style analysis. Computational Statistics and Data Analysis, 47, 353-372.

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

[36] Planas, C., Depoutot, R., 2002. Controlling revisions in ARIMA-modelbased seasonal adjustment, Journal of Time Series Analysis 23, 193-213.

[37] Psaradakis, Z., Spagnolo, F. (2003). On the determination of the number of regimes in Markov-Switching autoregressive models, Journal of Time Series Analysis, 24, 237-252.

[38] Teräsvirta, T. 2009. Recent developments in GARCH modeling. In Andersen, T.G., Davis, R.A., Kreiß, J.-P., Mikosch, T. (Eds.) *Handbook of Financial Time Series*, 17-42. Springer-Verlag: Berlin Heidelberg

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

[40] Theodossiou, P. and Lee, U. (1993) Mean and Volatility Spillovers across Major National Stock Markets: Further Empirical Evidence. Journal of Financial Research, 16, 327-350, 1993.

[41] Wang, X., Smith, K. and Hyndman, R. (2006). Characteristic-Based Clustering for Time Series Data, *Data Mining and Knowledge Discovery*, 13, 335364.

[42] Zakoïan, J.M. (1994). Threshold Heteroskedastic Models, *Journal of Economic Dynamic and Control*, 18, 931-955.

Ultimi Contributi di Ricerca CRENoS

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

11/12 *Alessandro Fiori Maccioni,* "The risk neutral valuation paradox"
11/11 *Leonardo Becchetti, Vittorio Pelligra,* "Don't Be Ashamed to Say You Didn't Get Much: Redistributive Effects of Information Disclosure in Donations and Inequity-Aversion in Charitable Giving"
11/10 *Alessandra Colombelli, Marta Foddi, Raffaele Paci,* "The knowledge regions in the enlarged Europe"
11/09 *Emanuela Marrocu, Raffaele Paci, Stefano Usai,* "Proximity, Networks and Knowledge Production in Europe"
11/08 *Vittorio Pelligra,* "Empathy, Guilt-Aversion and Patterns of Reciprocity"
11/07 *Claudio Detotto, Edoardo Otranto,* "Cycles in Crime and Economy Revised"
11/06 *Juan Gabriel Brida, Marta Meleddu, Manuela Pulina,* "The attractiveness of a modern and contemporary art museum"
11/05 *Juan Gabriel Brida, Marta Meleddu, Manuela Pulina,* "An urban icon? The case of the Iceman Ötzi "
11/04 *Silvia Balia, Rinaldo Brau,* "A Country for Old Men? An Analysis of the Determinants of Long-Term Home Care in Europe"
11/03 *Luciano Mauro, Francesco Pigliaru,* "Social Capital, Institutions and Growth: Further Lessons from the Italian Regional Divide"
11/02 *Juan Gabriel Brida, Claudio Detotto, Manuela Pulina,* "How efficient is the Italian hospitality sector? A window DEA and truncated-Tobit analysis"
11/01 *Berardino Cesi, Dimitri Paolini,* "University choice, peer group and distance"
10/33 *Oliviero A. Carboni, Giuseppe Medda,* "A Neoclassical Growth Model with Public Spending"
10/32 *Vittorio Pelligra, Luca Stanca,* "To Give or Not To Give? Equity, Efficiency and Altruistic Behavior in a Survey-Based Experiment"
10/31 *Emanuela Marrocu, Raffaele Paci,* "Education or just Creativity: what matters most for economic performance?"
10/30 *Adriana Di Liberto, Stefano Usai,* TFP convergence across European regions: a comparative spatial dynamics analysis
10/29 *Oliviero A. Carboni, Heterogeneity in R&D Cooperation: An Empirical Investigation*
10/28 *Maurizio Conti, Giovanni Sulis,* "Human Capital, Employment Protection and Growth in Europe"
10/27 *Juan Gabriel Brida,* Manuela *Pulina*, *Eugenia Riaño, Sandra Zapata-Aguirre* "Investigating the behavior of embarking cruisers in a Caribbean homeport: a factor and a censured-Tobit analysis"
10/26 *Juan Gabriel Brida,* Manuela *Pulina*, *Eugenia Riaño,* "Visitors' experience in a modern art museum: a structural equation model"
10/25 *Gerardo Marletto, Cécile Silling,* "Distance matters – The environmental impact of regional and national supply chains of canned tomatoes"