Multivariate Markov Switching Dynamic Conditional Correlation GARCH representations for contagion analysis

M. Billio
Università Ca' Foscari Venezia

M. Caporin
Università Ca' Foscari Venezia

WORKING PAPER n. 05.02

May 2005

Multivariate Markov Switching Dynamic Conditional Correlation GARCH representations for contagion analysis*

Monica Billio[†]- Università Ca' Foscari Venezia - billio@unive.it
Massimiliano Caporin - Università Ca' Foscari Venezia - mcaporin@unive.it

May 31, 2006

Abstract

This paper provides an extension of the Dynamic Conditional Correlation model of Engle (2002) by allowing both the unconditional correlation and the parameters to be driven by an unobservable Markov chain. We provide the estimation algorithm and perform an empirical analysis of the contagion phenomenon in which our model is compared to the traditional CCC and DCC representations.

1 Introduction

Since the seminal work of Bollerslev (1990), multivariate GARCH models attracted considerable interest given their direct application in both financial and economic empirical researches. By now, they represent a fundamental tool for asset and risk management and are employed in most financial market analyses. They have been extended and updated following the enormous literature of the univariate GARCH models, trying to taken into account the empirical findings in a multivariate setting. However, a first order of problems came into play when considering large dimension multivariate GARCH models: we refer to the complexity of parameter estimation procedures, which directly derives from the high number of coefficients. A second set of problems concerns the asymptotic

*We aknowledge financial support from the Italian national research project on "The Euro and European financial market volatility: contagion, interdependence and volatility transmission" financed by the Italian Ministry of University and Research. Furthermore, we thank William De Pieri for research assistance and are grateful to Loriana Pelizzon, Claudio Pizzi, Domenico Sartore and the participants at the Forecasting Financial Markets 2004 conference and at the XLII Annual Meeting of the Italian Statistical Society for helpful comments. Usual disclaimer applies.

†Corresponding author - e-mail: billio@unive.it - address: Dipartimento di Scienze Economiche - Università Ca' Foscari di Venezia - San Giobbe, 873 - 30121 Venezia - Italy - phone: +39-041-234-9170 - fax: +39-041-234-9176.

properties of the quasi maximum likelihood estimators for this type of models, which is not yet theoretically derived. In fact, Jeantheau (1998) provides conditions for QMLE consistency using a pointwise convergence criterion whereas a uniform convergence is required.

The high number of parameter problems motivated for the search of a simple and easy-to-estimate GARCH representation, and, at the same time, prevented for a direct use of generalised representations when the number of variables increases. This paper belongs to this research area and provides an up-to-date review of current feasible multivariate GARCH representation.

In particular, we extend the DCC model of Engle (2002) introducing Markov switches in the unconditional correlation matrix and in the DCC parameters. From the methodological point of view, this is a natural extension of the model recently proposed by Pelletier (2004). Moreover, the interest for this extension comes from the opportunity to use these DCC GARCH representations for contagion analysis.

In fact, the 1990s were punctuated by a series of severe financial and currency crises: the 1992 European monetary system attacks, the 1994 Mexican peso collapse, the 1997 East Asian crises, the 1998 Russian collapse, the 1998 LTCM crisis, the 1999 Brazilian devaluation, and the 2000 technological crisis. One striking characteristic of several of these crises was how an initial country-specific shock was rapidly transmitted to markets of very different sizes and structures around the globe. This has prompted a surge of interest in "contagion". This is a relevant issue from the empirical and theoretical point of view. Volatility transmission and contagion are relevant at an international level by themselves and have important consequences for monetary policy, optimal asset allocation, risk measurement, capital adequacy, and asset pricing.

In this paper, contagion - as opposed to interdependence - conveys the idea that international propagation mechanisms are discontinuous and then a hidden Markov chain can be useful to describe this discontinuity. However, there is no agreement on this definition and many other definitions have been proposed. In the empirical applications, we show the presence of the loss of interdependence phenomenon, which supports the thesis of discontinuities in the volatility propagation mechanisms.

Section 2 reviews the currently available multivariate GARCH models that consider dynamic conditional correlations. In section 3 we introduce the Markov switching DCC model (MS-DCC). Section 4 shortly presents the contagion issue and section 5 describes the empirical application to a set of European stock market indices. Section 5 concludes.

2 Multivariate GARCH with Dynamic Correlations

The Vech-GARCH of Engle and Kroner (1995) is one of the more general representations; it is characterised by the following equations (the GARCH orders have been restricted to one for the sake of exposition, but higher orders can easily be handled):

$$\begin{aligned} Y_t &= f\left(I^{t-1};\theta\right)+\varepsilon_t, \qquad \varepsilon_t|I^{t-1}\sim iid\left(0,H_t\right) \\ Vech\left(H_t\right) &= C+AVech\left(\varepsilon_t\varepsilon_t'\right)+BVech\left(H_{t-1}\right) \end{aligned} \tag{1}$$

where Y_t is a k-dimensional vector of variables whose mean can be time dependent; I^{t-1} is the information set at time $t-1$; ε_t is the vector of residuals that are independently and identically distributed following a multivariate unspecified distribution with time-dependent variance-covariance matrix H_t; furthermore, H_t follows a multivariate GARCH(1,1)-type representation in which the parameter matrices are of dimension $n \times n$ with $n = k \times (k+1)/2$ and $Vech\left(X\right)$ stacks the lower triangular elements of X.

The constraints required to ensure the positivity of conditional variances and the positive semi-definiteness of the variance-covariance matrix create many problems both in the implementation and in the optimisation steps. For this reason, the largest part of the literature focuses on finding a representation that reduces the computational burden of a Vech-Multivariate GARCH. Among others, we cite the diagonal Vech and BEKK representations of Engle and Kroner (1995). However, the most used structure is the Constant Conditional Correlation, introduced by Bollerslev (1990).

The basic idea of the CCC-GARCH is in noting that the variance-covariance matrix can be represented as the product of a correlation matrix by a standard deviation matrix:

$$H_t = diag\left(\sigma_{1t},\sigma_{2t}...\sigma_{kt}\right)\ R\ diag\left(\sigma_{1t},\sigma_{2t}...\sigma_{kt}\right) = D_tRD_t \tag{2}$$

where R is a correlation matrix and D_t is a diagonal matrix of standard deviations. This simplification allows both the overcoming of most parameter restrictions (only the coefficients of the conditional variances have to be bounded using the standard univariate inequality) and of a simple two-step estimation strategy that considers univariate GARCH estimation in the first step, while in the second step the correlation matrix is estimated by its sample estimator. Therefore, this approach is feasible even for very large systems.

More recently, attention has been drawn to a direct modelisation of the correlation matrix, leaving aside the conditional variances. This new field was started by Engle (2002) who suggested generalising the CCC-GARCH model of Bollerslev (1990), allowing the conditional correlations to vary over time. Equation (2) is then replaced by

$$H_t = D_tR_tD_t \tag{3}$$

where the correlation matrix follows a time dependent relation as follows

$$\begin{aligned} R_t &= Q_t^{*-1} Q_t Q_t^{*-1} \\ Q_t &= (1-\alpha-\beta)\,\bar{Q} + \alpha \boldsymbol{\eta}_{t-1}\boldsymbol{\eta}_{t-1}' + \beta Q_{t-1} \\ Q_t^{*} &= diag\left(\sqrt{q_{11,t}}, \sqrt{q_{22,t}} \cdots \sqrt{q_{kk,t}}\right) \\ \bar{Q} &= \frac{1}{T}\sum_{i=1}^{T} \boldsymbol{\eta}_{t-1}\boldsymbol{\eta}_{t-1}' \end{aligned} \tag{4}$$

In this model, $\bar{Q}$ represents the unconditional correlations, Q_t^* guarantees that R_t is a correlation matrix ($q_{ii,t},\ i = 1, ..., k$ are the elements of the diagonal of Q_t) and $\boldsymbol{\eta}_t$ are the standardised residuals, $\boldsymbol{\eta}_t = D_t^{-1} E_t$ where $D_t = diag\left(\sigma_{1t}, \sigma_{2t} ... \sigma_{kt}\right)$. It is worth noting that the DCC just impose a GARCH-type structure on the conditional correlations and uses only two parameters to add a dynamic behaviour. A very similar approach was contemporaneously suggested by Tse and Tsui (2002), with the only difference that they use small sample correlation estimates (instead of the standardised residuals η_t) to avoid the further standardisation of Q_t^*.

It has also to be noted that the simplicity of the suggested approaches is coupled with a strong restriction: the dynamic of correlation is constant among all the variables. This constraint can be removed, as suggested by Engle (2002) estimating an unrestricted DCC (or Generalised DCC, let us call it GDCC), where

$$Q_t = (ii' - A - B) \odot \bar{Q} + A \odot \boldsymbol{\eta}_{t-1}\boldsymbol{\eta}_{t-1}' + B \odot Q_{t-1} \tag{5}$$

and A and B are full square matrices of dimension k, i is a vector of ones and $\odot$ is the Hadamard product (i.e. the element by element product). Conditions for positive definiteness of Q_t are provided in Ding and Engle (2001). Clearly, the unrestricted DCC model creates the well-known problem of the high number of parameters, a motivation that was at the base of the development of the CCC and DCC model classes. Therefore, we no longer consider this structure.

Now a lot of work is done on the modelisation of the correlation matrix. For example, Franses and Hafner (2003) propose a restricted parameterisation of the GDCC, suggest that $A = aa'$, where a is vector of dimension k and $B = \beta$ is a scalar and impose the positive definiteness by modifying the intercept term of the GDCC equation (let us call this model the Franses-Hafner DCC, FH-DCC). This model adds flexibility and the number of parameters decreases sensibly, but it loses the property of correlation targeting, since the unconditional mean of FH-DCC is not the unconditional correlation, as in the DCC model.

Other extensions of the DCC model have been suggested. Cappiello *et al.* (2003) propose to add a term in the DCC equation to take into account the asymmetry (Asymmetric DCC - ADCC). Their model is fairly general to include all previous cases, however the to impose the positive definiteness is a very hard computational problem.

Further extensions of DCC models are developed by Billio *et al.* (2004) and Billio and Caporin (2004). The first paper introduces the Flexible DCC model

(FDCC) where the GDCC of Engle (2002) is modified by the introduction of a constant and of a parameter matrix structure similar to the one used by Franses and Hafner (2003). The model is represented by the following equation

$$Q_t = \mathbf{c}\mathbf{c}' \odot \bar{Q} + \mathbf{a}\mathbf{a}' \odot \boldsymbol{\eta}_{t-1}\boldsymbol{\eta}_{t-1}' + \mathbf{b}\mathbf{b}' \odot Q_{t-1} \tag{6}$$

where $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are partitioned parameter vectors (e.g. $\mathbf{a} = [a_1 i\,(m_1)\,,\ a_2 i\,(m_2)\,,\ \cdots\,,\ a_w i\,(m_w)]'$, where w is the number of partitions and $i\,(m_j)$ is a row vector of one of dimension m_j with $\sum_j m_j = k$). The FDCC model loses the correlation targeting property as the FH-DCC but also reduces the number of parameters. Clearly, the number of parameters depends on the number of partitions imposed on the coefficient vectors, which can include several assets. Finally, the model provides a positive definite Q_t since it is composed by the sum of positive definite and semi-definite matrices.

Billio and Caporin (2004) generalise the FDCC to the Quadratic FDCC (QFDCC), which is similar to the model of Cappiello *et al.* (2003) and uses the structure of the FDCC. The QFDCC model is characterised by the following structure

$$Q_t = C\bar{Q}C' + A\boldsymbol{\eta}_{t-1}\boldsymbol{\eta}_{t-1}'A' + BQ_{t-1}B' \tag{7}$$

where A, B and C are partitioned parameter matrices. If the parameter matrices are diagonal the QFDCC model collapses on the FDCC one. The QFDCC model provides positive definite correlation matrices if the eigenvalues of $A+B$ are in modulus less than one and the matrix $C\bar{Q}C'$ is positive definite. This results derive from Engle and Kroner (1995) noting that the QFDCC is similar to a BEKK representation for the correlation matrix.

Finally, Chan *et al.* (2003) suggest a slightly different approach. They provide a general representation for a dynamic correlation model with stochastic coefficients that nests all the previous DCC-type models. The most important issue considered by Chan *et al.* (2003) pertains the asymptotic properties of their GARCC model: in fact they provide the regularity conditions under which the Quasi Maximum Likelihood estimator is consistent. Moreover, they show how the DCC of Engle (2002) is included as a particular case of their GARCC and argue that even the GDCC is a particular case of their representation.

Even if all these representations are useful to deal with high dimension problems and thus can be helpful for volatility transmission analysis (following King and Wadhwani (1990), King et al. (1994) and Ramchard and Susmel (1998)), they cannot account for discontinuities. In fact, the standard DCC model and its generalisations provide a feasible structure for the treatment of large systems, but they impose a fix unconditional correlation over the sample and for long samples this may be questionable. A change in the unconditional correlation levels can be hypothesize to explain some empirical findings and this change can be associated with a regime switch between two or more unconditional correlation levels.

The idea of adding Markov switching regimes to the correlation has already been considered by Pelletier (2004), who restricted his attention to a CCC model with regime switches (MS-CCC thereon). In particular, Pelletier suggested the following restricted representation for the correlation matrix

$$R_t = \Gamma\lambda(s_t) + I_k[1 - \lambda(s_t)] \tag{8}$$

where $\lambda(s_t)$ is a state-dependent variable assuming only positive values and $s_t = 1, 2, ..., S$. This representation is motivated by the computational problems related to a full estimation of S state specific correlation matrices, which clearly involve a very high number of parameters. However, this approach does not allow the correlations to change sign and this possibility cannot be excluded a priori.

In the next section we extend the DCC class by allowing both the unconditional correlation and the parameters to be driven by a latent Markov chain.

2.1 The estimation issue

The class of DCC models can be estimated with a two-step Quasi Maximum Likelihood approach, as demonstrated by Engle (2002). The full log-likelihood can be represented as

$$LogL(Y) = \frac{1}{T}\sum_{t=1}^{T} \log L(Y_t) = \frac{1}{T}\sum_{t=1}^{T}\left[-\frac{1}{2}\left(\log|H_t| + \varepsilon_t' H_t^{-1}\varepsilon_t\right)\right] \tag{9}$$

but recalling that $H_t = D_t R_t D_t$ and that $|D_t R_t D_t| = |D_t||R_t||D_t|$ we have

$$LogL(Y) = -\frac{1}{2T}\sum_{t=1}^{T}\left[2\log|D_t| + \log|R_t| + \varepsilon_t' D_t^{-1} R_t^{-1} D_t^{-1}\varepsilon_t\right] \tag{10}$$

Therefore, replacing in a first step the correlation matrix by an identity matrix, we can maximise only with respect to the parameters appearing in D_t. In a second stage, the estimation will then be performed conditionally on the estimate variances, i.e. on

$$LogL(Y|D) = -\frac{1}{2T}\sum_{t=1}^{T}\left[\log|R_t| + \eta_t' R_t^{-1}\eta_t\right] \tag{11}$$

Engle (2002) and Engle and Sheppard (2002) provide the asymptotic properties of the standard DCC model estimators, a result which is generalised by Chan et al. (2003).

3 A DCC model with Markov switching regimes: the MS-DCC model

To modelise a possible change in the unconditional correlations, we follow Pelletier (2004) and introduce a regime switch between two or more unconditional correlation levels. Moreover, we allow for changes in the sign of the switching correlations, a possibility which is excluded in Pelletier (2004) but cannot be excluded a priori.

Differently from Pelletier (2004), we consider the standard DCC model in the following reformulation:

$$Q_t = [1 - \alpha(s_t) - \beta(s_t)]\,\bar{Q}(s_t) + \alpha(s_t)\,\boldsymbol{\eta}_{t-1}\boldsymbol{\eta}'_{t-1} + \beta(s_t)\,Q_{t-1} \tag{12}$$

where both the unconditional correlation matrix and the parameters driving the system dynamics can be regime dependent. The Markov chain is governed by the following transition matrix

$$P = \{p_{ji}\ i, j = 1, ...S\} \tag{13}$$

where S is the number of regimes. In the following we refer to this model as the MS(S)-DCC.

As in Pelletier (2004), we restrict the regime dependent structure only to the correlations excluding any effect on variances. This restriction allows us to consider a two-step estimation procedure. In fact, a full Markov switching model will become highly unstable given the huge number of switching parameters.

Given the joint presence of regime switches and dynamic correlations the estimation of model (12) is very difficult being the Hamilton filter useless. In fact, since the matrix Q_t is not observed, the equation (12) should be modified into

$$Q_t^{ij} = \left[1 - \alpha_j - \beta_j\right]\bar{Q}_j + \alpha_j \boldsymbol{\eta}_{t-1}\boldsymbol{\eta}'_{t-1} + \beta_j Q_{t-1}^{li} \tag{14}$$

where the upperscript j,i and l refers to the state in t, $t-1$ and $t-2$, respectively; the dynamic structure of Q_t induce the dependence of the current regime to all the past regimes. Using equation (14) in a standard Hamilton filter, an S-fold increase of possible combinations is created at any new point in time. Therefore, some approximation is needed.

According to Kim (1994), we consider the following modified Hamilton filter:

[i] given the filtered probabilities $\Pr\left(s_{t-1} = i | I^{t-1}\right)$ as inputs, determine the joint probabilities:

$$\Pr\left(s_t = j, s_{t-1} = i | I^{t-1}\right) = \Pr\left(s_t = j | s_{t-1} = i\right)\Pr\left(s_{t-1} = i | I^{t-1}\right) \quad i, j = 1...S$$

[ii] evaluate the regime dependent likelihood:

$$Q_t^{ij} = \left[1 - \alpha_j - \beta_j\right] \bar{Q}_j + \alpha_j \boldsymbol{\eta}_{t-1} \boldsymbol{\eta}'_{t-1} + \beta_j Q_{t-1}^i \quad i, j = 1...S$$

$$\tilde{Q}_t^{ij} = diag\left(\sqrt{q_{11,t}^{ij}}, \sqrt{q_{22,t}^{ij}} \cdots \sqrt{q_{kk,t}^{ij}}\right)$$

$$R_t^{ij} = \left(\tilde{Q}_t^{ij}\right)^{-1} Q_t^{ij} \left(\tilde{Q}_t^{ij}\right)^{-1}$$

$$LogL_t(Y_t|D_t, s_t = j, s_{t-1} = i, I^{t-1}) = -\frac{1}{2T}\left(\log\left|R_t^{ij}\right| + \eta'_t \left(R_t^{ij}\right)^{-1} \eta_t\right)$$

[iii] evaluate the likelihood of observation t:

$$LogL_t(Y_t|D_t, I^{t-1}) = \sum_{j=1}^{S}\sum_{i=1}^{S} LogL_t(Y_t|D_t, s_t = j, s_{t-1} = i, I^{t-1}) \Pr\left(s_t = j, s_{t-1} = i|I^{t-1}\right)$$

$$LogL_t = LogL_{t-1} + LogL_t(Y_t|D_t, I^{t-1})$$

[iv] update the joint probabilities:

$$\Pr\left(s_t = j, s_{t-1} = i|I^t\right) = \frac{LogL_t(Y_t|D_t, s_t = j, s_{t-1} = i, I^{t-1}) \Pr\left(s_t = j, s_{t-1} = i|I^{t-1}\right)}{LogL_t(Y_t|D_t, I^{t-1})}$$

for $i, j = 1...S$;

[v] compute the filtered probabilities:

$$\Pr\left(s_t = j|I^t\right) = \sum_{i=1}^{S} \Pr\left(s_t = j, s_{t-1} = i|I^t\right) \quad j = 1...S$$

[vi] update the correlation matrix using the following approximation:

$$Q_t^j = \frac{\sum_{i=1}^{S} \Pr\left(s_t = j, s_{t-1} = i|I^t\right) Q_t^{ij}}{\Pr\left(s_t = j|I^t\right)}$$

[vii] iterate [i] to [vi] until the end of the sample.

Note that the last equation collapses an S^2-fold into an S-fold by an approximation based on the current state probabilities (the well-known Kim approximation). To initialise the filter, the regime probabilities could be set equal to the unconditional probabilities while Q_0^j can be obtained with a sample correlation estimator computed in two different subset of the full sample, possibly distinguishing before and during a crisis period. For this operation, some relevant information can be obtained by a preliminary analysis.

We should also specify a smoother for the hidden regimes. Given the approximation used in the filter, the smoother will include itself an implicit approximation as discussed by Kim and Nelson (1998). The algorithm requires as inputs the filtered probabilities obtained with the approximated filter and the transition probabilities. Since $\Pr\left(s_T = j|I^T\right)$ is known (it is obtained in the last iteration of the filter), it can be used to initialise the smoother, which is the following:

$$\Pr\left(s_t = j, s_{t+1} = m|I^T\right) = \frac{\Pr\left(s_{t+1} = m|I^T\right)\Pr\left(s_t = j|I^t\right)\Pr\left(s_{t+1} = m|s_t = j\right)}{\Pr\left(s_{t+1} = m|I^t\right)}$$

$$\Pr\left(s_t = j|I^T\right) = \sum_{m=1}^{S} \Pr\left(s_t = j, s_{t+1} = m|I^T\right) \tag{15}$$

for $j, m = 1...S$.

In large systems the MS-DCC model may have some convergence problems given the high number of parameters involved. In this case the parameterisation suggested by Pelletier (2004) can be used also in our framework. We should then rearrange equation (**??**), letting $Q_j = \Gamma\lambda\left(s_j\right) + I_k\left[1 - \lambda\left(s_j\right)\right]$where j is the regime in t. Thus, a single correlation matrix has to be estimated, sensibly reducing the number of parameters; however, as in Pelletier (2004), changes in the correlation sign are not allowed.

Finally, the MS structure can easily be extended to all the other DCC models reviewed in section 2.

4 Contagion definitions and literature

The last two decades have experienced a series of financial and currency crises, many of them carrying regional or even global consequences: the 1987 Wall Street crash, the 1992 European monetary system collapse, the 1994 Mexican pesos crisis, the 1997 "Asian Flu", the 1998 "Russian Cold", the 1999 Brazilian devaluation, the 2000 Internet bubble burst and the default crisis in Argentina of July 2001. Most of these crises hit emerging markets, which are more sensitive to shocks because of their underdeveloped financial markets and their large public deficits. The common feature shared by these events was that a country specific shock spreads rapidly to other markets of different sizes and structures all around the world.

It is still quite puzzling to justify the reason why a country reacts to a crisis affecting another country to which the former is not linked by any economic fundamentals. Many authors dealing with the topic of international propagation of shocks have referred to this circumstance as a contagion phenomenon, even if there is still no agreement on which factors lead to identify a contagion event precisely, and it is not yet clear how to define the contagion event itself.

Referring to the World Bank's classification, we can distinguish three definitions of contagion:

- *Broad definition*: contagion is identified with the general process of shock transmission across countries. The latter is supposed to work both in tranquil and crisis periods, and contagion is not only associated with negative shocks but also with positive spillover effects;

- *Restrictive definition*: this is probably the most controversial definition. Contagion is the propagation of shocks between two countries (or group of countries) in excess of what should be expected by fundamentals and considering the co-movements triggered by the common shocks. If we adopt this definition of contagion, we must be aware of what constitutes the underlying fundamentals. Otherwise, we are not able to appraise effectively whether excess co-movements have occurred and then whether contagion is displayed.

- *Very restrictive definition*: this is the one adopted by Forbes and Rigobon (2000). Contagion should be interpreted as the change in the transmission mechanisms that takes place during a turmoil period. For example, the latter can be inferred by a significant increase in the cross-market correlation. As we have said, this is the definition that will be used in this paper.

Many papers have focused on the question of contagion, testing for its existence with statistical methods. Their approaches vary with regard to the definition of contagion they choose as a starting point. As we have anticipated we will use the third definition.

Why do we concentrate on this aspect of contagion? Why is this definition of contagion important as is its exploration? Because, as observed by Forbes and Rigobon, the other definitions of contagion and relative approaches of analysis are unable to shed light on three main issues: international diversification, evaluation of the role and the potential effectiveness of international institutions and bail-out funds, and propagation mechanisms. Indeed, a critical assumption of investment strategies is that most economic disturbances are country specific. As a consequence, stock in different countries should be less correlated. However, if market correlation increases after a bad shock, this would undermine much of the rationale for international diversification.

The variety of empirical methods developed for the analysis of contagion has the aim of testing the stability of parameters in the sphere of a chosen econometric model. Evidence of parameter shifts is a signal of a change in the transmission mechanism, so according to the third definition there has been contagion. If, on the contrary, the parameters are constant, we should move to an interdependence case. Several methodologies have been used to statistically search for contagion in this way and others still have to be applied. Rigobon (2001) offers a good survey of these procedures, which are mainly based on OLS estimates (including GLS and FGLS), Principal Components, Probit models and correlation coefficient analysis.

However, the methodologies listed above carry some imperfections because the data often suffer from heteroskedasticity, endogenous and omitted variable problems. Some authors have tried to solve these problems in a similar way, although they have reached different conclusions in terms of contagion. In particular, Forbes and Rigobon (2002) developed a correlation analysis adjusting correlation coefficients only for heteroskedasticity under the assumption of no omitted variables or simultaneous equations problems. Corsetti et al. (2002) built up a model in which the specific shock of the country under crisis does not necessarily act as a global shock because this could bias the analysis in favour of interdependence instead of contagion. The authors therefore introduce more sophisticated assumptions about the ratio between the variance of the country-specific shock and the variance of the global factors weighted by factor loadings. Nevertheless, both these tests are still highly affected by the presence of omitted variables, the time zone and the windows used (see Billio and Pelizzon (2003)).

Working with Markov switching models and using regime probabilities (filtered or smoothed) to monitor the volatility transmission process is a relatively innovative approach, which has already suggested by Hassler (1995), Ang and Bekaert (1999) and Baele (2003), but which has not yet completely developed in contagion analysis (see however Pericoli and Sbracia (2003) for a review of empirical works considering Markov switching models, Edward and Susmel (2003), Gallo and Otranto (2004) and Billio *et al.* (2004)).

To combine a DCC-GARCH model with a Markov switching approach allows us to take into account several important aspects: first of all the heteroskedasticity of the data can be properly modelised; secondly, to consider dynamic correlation permits to analyse the dynamics of contagion; finally, to consider a latent Markov chain allows the endogenous definition of the crisis periods.

5 Detecting contagion with the MS-DCC model

This section presents an empirical application of the MS-DCC model. We consider a set of daily stock market indices: Standard & Poors 500 (S&P), FTSE100 (FTSE), EuroStoxx50 (EX), Nikkey225 (NK), Hang Seng (Hong Kong - HS), Straits (Singapore - STR) and KLSE (Malaysia - KLSE). The analysis is performed on the daily returns and the series run from January 2000 to December 2003. Data have been downloaded from Datastream.

Since holiday days are not common over the stock markets, we consider the following cut-off rule: we remove all common holidays while non-common holidays are replaced by a zero return. This approach has the advantage of not introducing spurious correlation in the data and to preserve all available points in time. Furthermore, to avoid any problem due to asynchronous trading we took the two-day moving average of the returns, as in Forbes and Rigobon (2002).

After this step, we filtered the idiosyncratic heteroskedasticity for each series by fitting univariate GARCH models with asymmetry (see Glosten *et al.* (1993)). The results are reported in Table 1. All indices evidence a relevant asymmetric effect: in all cases the γ parameter is highly significant[1].

On the variance-filtered series we estimated a set of static and dynamic correlation models: the CCC, the MS(2)-CCC, the DCC and the MS(2)-DCC; Table 2 reports the likelihoods of the four cases.

We also considered Markov switching models with three states but they have been excluded since the third states was highly unstable (its persistence probability resulted almost zero). It is important to note that the test for the presence of the Markov switches is not standard, due to the presence of nuisance parameters only under the alternative hypothesis. Thus, it is not possible to simply compare the likelihoods for determining the optimal state number (see Davies (1977, 1987) and Hansen (1992)). However, we can compare the DCC model with the CCC one and the MS(2)-DCC model with the MS(2)-CCC one by restricting the DCC coefficients to be zero (see Table 3 for the DCC parameter estimates). In that cases, the Likelihood Ratio (LR) test statistics (chi-squared with 2 degrees of freedom for which the 1% quantile is 9.21) are equal to 444.772 and 423.428, respectively. In both cases the null hypothesis of no dynamics in the correlations is strongly rejected giving a clear preference for the Dynamic Correlation models.

Finally, we verified the null hypothesis that the β parameter of the MS(2)-DCC model is stable over the two regimes. The associated Wald test statistics is equal to 214.96, highly rejecting the null hypothesis.

The estimation algorithms of the two Markov switching models have been initialised with the same starting values (for regime 0 the sample correlation matrix and for regime 1 no correlation) and reports at the optimum two very similar state dependent matrices[2]. Figure 1 reports a comparison of the smoothed probabilities of regime 1 obtained with the MS(2)-CCC and the MS(2)-DCC models. The two graphs provide similar patterns, both models identify almost the same regimes and switches among them (their concordance is 0,75). The estimation of the Markov chain affecting market movements is thus quite robust. However, even if the regimes classification is almost the same, the chain extraction with the DCC model is less precise. To this respect, it is important to remember that the filter algorithm for the MS-DCC model is approximated and this could explain this difference.

As evidenced in Table 2, it clearly emerges that the inclusion of Markov switches combined with the DCC dynamic evolution improves the likelihood. Transition probability matrices (reported in Table 4) evidence that the regimes are persistent and the DCC parameters are affected by the regime. Furthermore, Tables 5, 6 and 7 report the unconditional correlation matrices (for the CCC and DCC models) and the state dependent correlation matrices of the Markov switching models. The correlation matrices of Tables 6 and 7 can be associated

[1] In the GJR-GARCH model (Glosten *et al.*, 1993) the variance equation is

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \gamma\varepsilon_{t-1}^2 I(\varepsilon_{t-1}) + \beta\sigma_{t-1}^2$$

where $I(\varepsilon_{t-1})$ is an indicator function assuming the value 1 for negative values of ε_{t-1} and 0 otherwise.

[2] Moreover, the optimum seems not affected by the use of different starting values.

with a high correlation state (regime 0) and a lower correlation state (regime 1). Finally, the switches between the two regimes can be matched with the presence of the so-called "loss of interdependence" phenomenon, which happens when the link among the markets falls during the crisis. Whenever a turbulence starts in a market, the other markets immediately suffer the effects of the turbulence, but as soon as they identify the noisy signal and filter it out, the correlation with the turbulent area reduces. In that view, the low correlation state, regime 1, appear as the dominant regime for example after September 11, and again during the war operations in Afghanistan and Iraq. Indeed, even if this result cannot be accounted as contagion, correlation falls support the thesis of discontinuities in the propagation mechanisms and it has been evidenced by several authors (see, for example, Billio and Pelizzon (2003)).

It is also important to note that the MS-DCC α parameter associated to regime 1 (lower correlation regime) is larger than in regime 0. This parameter represents the loading of market shocks into the correlation dynamics and an increase in the α associated to low correlations can be interpreted as an increase in the reply to market movements. This is not surprising: noting that the low correlation is associated with high volatile periods we can expect that correlation dynamic dependence on market movements increases while the unconditional correlation falls.

6 Conclusions

In this paper we introduce a generalisation of the DCC model of Engle (2002) by allowing for Markov switches in both the parameter and the unconditional correlation. Following Kim (1994), we develop a modified Hamilton algorithm for model estimation and perform an empirical application on daily European and US stock market index returns. The proposed MS-DCC model is clearly preferred in terms of likelihood and allows an economic interpretation of the identified regimes. In particular, we point to the presence of the loss of interdependence phenomenon, which supports the thesis of discontinuities in the volatility propagation mechanisms.

References

[1] Ang A. and G. Bekaert (1999), International Asset Allocation with Time-Varying Correlations, NBER WP n. w7056.

[2] Baele L., 2003, Volatility Spillover Effects in European Equity Markets: Evidence from a Regime Switching Model, EIFC - Technology and Finance Working Papers 33, United Nations University, Institute for New Technologies, forthcoming *Journal of Financial and Quantitative Analysis.*

[3] Billio M. and M. Caporin, 2004, A generalised Dynamic Conditional Correlation model for portfolio risk evaluation, Proceedings of the Deloitte Conference in Risk Management 2004, Antwerp, 29-30 March 2004.

[4] Billio M., M. Caporin and M. Gobbo, 2004, Flexible Dynamic Conditional Correlation Multivariate GARCH models for asset allocation, Proceedings of the Forecasting Financial Market conference 2003, Paris, 4-6 June 2003 and Working Paper 04.03, GRETA, Venice.

[5] Billio M., M. Lo Duca and L. Pelizzon, 2004, Contagion Detection with Switching Regime Models: a Short and Long Run Analysis, mimeo GRETA, Venice.

[6] Billio M. and L. Pelizzon, 2003, Contagion and Interdependence in Stock Markets: Have they been misdiagnosed?, *Journal of Economics and Business*, 55, 5/6, 405-426.

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

[8] Cappiello L., R.F. Engle and K. Sheppard, 2003, Asymmetric dynamics in the correlations of global equity and bond returns, European Central Bank Working Paper 204.

[9] Chan F., S. Hoti and M. McAleer, 2003, Generalised Autoregressive Conditional Correlation, preprint Department of Economics, University of Western Australia.

[10] Corsetti G., M. Pericoli and M. Sbracia, 2002, Some contagion, some interdependence more pitfalls in tests of financial contagion, Discussion Paper n. 3310, CEPR.

[11] Davies R.B. (1977), Hypothesis testing when a nuisance parameter is present only under the alternative, Biometrika 64, 247-254.

[12] Davies R.B. (1987), Hypothesis testing when a nuisance parameter is present only under the alternative, Biometrika 74, 33-43.

[13] Ding Z. and R.F. Engle, 2001, Large scale conditional covariance matrix modeling estimation and testing, Working paper 01029, Department of Finance, NYU Stern School of Business.

[14] Edward S. and R. Susmel (2003), Interest-Rate Volatility in Emerging Markets, *Review of Economics and Statistics*, 85, 328-348.

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

[16] Engle R.F., 2002, Dynamic conditional correlation: a new simple calss of multivariate GARCH models, *Journal of Business and Economic Statistics*, 20, 339-350.

[17] Engle R.F. and K. Sheppard, 2001, Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH, University of California, San Diego, Discussion paper 2001-15, NBER Working Paper 8554.

[18] Forbes K. and R. Rigobon, 2000, Measuring contagion: conceptual and empirical issues, in S.Claessens and K. Forbes (Eds), *International Financial Contagion*, Kluwer Academic Publishers.

[19] Forbes K.J. and R. Rigobon, 2002, No contagion, only interdependence: measuring stock market comovements, *Journal of Finance*, 57/5, 2223-2261.

[20] Franses P.H. and C.M. Hafner, 2003, A Generalised Dynamic Conditional Correlation Model for Many Asset Returns, Econometric Institute Report EI 2003-18, Erasmus University Rotterdam.

[21] Gallo G.M. and E. Otranto (2004), Contagion and Interdependence in Financial Markets: a New Approach, FEDRA Working Papers.

[22] Glosten L., R. Jagannathan and D. Runkle (1993), On the relation between the expected value and volatility of nominal excess returns on stocks, *Journal of Finance*, 46, 1779-1801.

[23] Hansen B.E. (1992), The likelihood ratio test under non-standard conditions: Testing the Markov switching model of GNP, *Journal of Applied Econometrics*, 7, S61-S82. Also "Erratum", *Journal of Applied Econometrics* 1996, 11, 195-198.

[24] Hassler J., 1995, Regime Shifts and Volatility Spillover on International Stock Markets, IIES Seminar Paper n. 603.

[25] King M., E. Sentana and S. Wadhwani, 1994, Volatility and Links between National Stock markets, *Econometrica* 62/4, 901/933.

[26] King M. and S. Wadhwani, 1990, Transmission of volatility between stock markets, *Review of Financial Studies*, 3/1, 5-33

[27] Kim, C., 1994, Dynamic linear models with Markov switching, *Journal of Econometrics*, 60, 1-22

[28] Kim C. and C.R. Nelson, 1998, *State-Space models with regime switching*, MIT press.

[29] Pelletier D., 2004, Regime switching for dynamic correlations, forthcoming in *Journal of Econometrics.*

[30] Pericoli M. and M. Sbracia (2003), A Primer on Financial Contagion, *Journal of Economic Surveys*, 17(4), 571-608.

[31] Ramchand L. and R. Susmel, 1998, Volatility and Cross Correlation Across Major Stock markets, *Journal of Empirical Finance*, 5, 397-416.

[32] Rigobon R., 2001, Contagion: how to measure it?, NBER Working Paper No. w8118.

[33] Tse Y.K. and A.K.C. Tsui, 2002, A multivariate Generalised Autoregressive Conditional Heteroscedasticity model with time-varying correlations, *Journal of Business and Economic Statistics*, 20 (3), 351-362.

Table 1: GJR-GARCH estimates

Series	Parameters / Standard Error			
	ω	α	γ	β
S&P	0.01730	0.04662	0.15591	0.85110
	0.00013	0.00051	0.00074	0.00081
FTSE	0.12299	0.36178	0.18097	0.51987
	0.00095	0.00164	0.00181	0.00203
EX	0.21468	0.15853	0.18033	0.63704
	0.00318	0.00151	0.00128	0.00350
NK	0.22473	0.25267	0.06831	0.54264
	0.00507	0.00370	0.00147	0.00709
HS	0.15593	0.39825	0.17502	0.48429
	0.00143	0.00206	0.00190	0.00247
STR	0.01693	0.08890	0.09102	0.85219
	0.00011	0.00052	0.00053	0.00063
KLSE	0.02205	0.06315	0.13952	0.83608
	0.00019	0.00059	0.00071	0.00097

Table 2: LogLikelihoods of the fitted models

Model	Log-Likelihood
CCC	-2357.994
MS(2)-CCC	-2286.330
DCC	-2135.608
MS(2)-DCC	-2074.616

Table 3: DCC and MS(2)-DCC parameter estimates

DCC parameters		Coeff.	Std.Err.
DCC	α	0.186	0.012
	β	0.403	0.044
MS-DCC Regime 0	α	0.160	0.057
	β	0.379	0.030
MS-DCC Regime 1	α	0.237	0.026
	β	0.360	0.069

Table 4: Transition matrices of the MS(2)-CCC and MS(2)-DCC models

Tansition matrices		
MS-CCC	**0**	**1**
0	0.868 0.057	0.132
1	0.142	0.858 0.032
MS-DCC	**0**	**1**
0	0.744 0.031	0.256
1	0.142	0.827 0.025

Transition probabilities and standard errors

Table 5: Unconditional correlation matrix

Correlation Matrix (CCC and DCC)						
1	---	---	---	---	---	---
0.659	1	---	---	---	---	---
0.672	0.830	1	---	---	---	---
0.378	0.427	0.456	1	---	---	---
0.412	0.499	0.510	0.532	1	---	---
0.356	0.440	0.462	0.490	0.599	1	
0.102	0.085	0.147	0.225	0.272	0.333	1

Correlations have been estimated using the sample estimator

Table 6: State dependent correlation matrices of the MS-CCC model

	Correlation matrices - MS(2)-CCC						
	S&P500	FTSE100	EX50	NIK225	HS	STR	KLSE
	Regime 0						
S&P500	1	---	---	---	---	---	---
FTSE100	0.703 0.032	1	---	---	---	---	---
EX50	0.786 0.035	0.900 0.011	1	---	---	---	---
NIK225	0.433 0.050	0.553 0.067	0.535 0.061	1	---	---	---
HS	0.550 0.052	0.606 0.046	0.595 0.063	0.682 0.037	1	---	---
STR	0.501 0.064	0.561 0.067	0.540 0.045	0.652 0.070	0.761 0.052	1	---
KLSE	*0.049* *0.091*	*0.114* *0.073*	*0.146* *0.077*	0.276 0.081	0.163 0.083	0.329 0.110	1
	Regime 1						
S&P500	1	---	---	---	---	---	---
FTSE100	0.612 0.041	1	---	---	---	---	---
EX50	0.557 0.045	0.752 0.037	1	---	---	---	---
NIK225	0.318 0.072	0.277 0.075	0.368 0.072	1	---	---	---
HS	0.256 0.067	0.367 0.048	0.408 0.064	0.354 0.080	1	---	---
STR	0.196 0.073	0.294 0.058	0.372 0.049	0.294 0.081	0.415 0.050	1	---
KLSE	0.159 0.062	*0.029* *0.057*	0.132 0.060	*0.153* *0.096*	0.402 0.061	0.336 0.105	1

Estimated coefficients and standard errors - in italics we reports non-significant correlations at the 5% confidence level

Table 7: state dependent correlation matrices of the MS-DCC model

	Correlation matrices - MS(2)-DCC						
	S&P500	FTSE100	EX50	NIK225	HS	STR	KLSE
	Regime 0						
S&P500	1	---	---	---	---	---	---
FTSE100	0.699 0.127	1	---	---	---	---	---
EX50	0.819 0.074	0.917 0.062	1	---	---	---	---
NIK225	0.432 0.192	0.535 0.152	0.509 0.127	1	---	---	---
HS	0.579 0.100	0.604 0.121	0.592 0.182	0.729 0.098	1	---	---
STR	0.471 0.095	0.551 0.117	0.546 0.081	0.696 0.181	0.772 0.090	1	---
KLSE	0.381 0.181	0.482 0.232	0.467 0.177	0.635 0.210	0.582 0.180	0.780 0.276	1
	Regime 1						
S&P500	1	---	---	---	---	---	---
FTSE100	0.594 0.148	1	---	---	---	---	---
EX50	0.536 0.182	0.756 0.073	1	---	---	---	---
NIK225	0.329 0.166	0.354 0.174	0.423 0.104	1	---	---	---
HS	0.215 0.090	0.380 0.110	0.405 0.080	0.454 0.192	1	---	---
STR	0.272 0.123	0.376 0.128	0.418 0.067	0.386 0.108	0.497 0.163	1	---
KLSE	*-0.128* *0.079*	*-0.145* *0.085*	*-0.038* *0.040*	*-0.023* *0.032*	0.096 0.037	0.096 0.027	1

Estimated coefficients and standard errors - in italics we reports non-significant correlations at the 5% confidence level

Figure 1: smoothed probabilities of regime 0 for MS(2)-CCC and MS(2)-DCC models