Thresholds, News Impact Surfaces and Dynamic Asymmetric Multivariate GARCH[*]

Massimiliano Caporin[†]
Department of Economic Sciences
University of Padova

Michael McAleer
Department of Quantitative Economics
Complutense University of Madrid

March 2009

Abstract: DAMGARCH is a new model that extends the VARMA-GARCH model of Ling and McAleer (2003) by introducing multiple thresholds and time-dependent structure in the asymmetry of the conditional variances. Analytical expressions for the news impact surface implied by the new model are also presented. DAMGARCH models the shocks affecting the conditional variances on the basis of an underlying multivariate distribution. It is possible to model explicitly asset-specific shocks and common innovations by partitioning the multivariate density support. This paper presents the model structure, describes the implementation issues, and provides the conditions for the existence of a unique stationary solution, and for consistency and asymptotic normality of the quasi-maximum likelihood estimators. The paper also presents an empirical example to highlight the usefulness of the new model.

Keywords: multivariate asymmetry, conditional variance, stationarity conditions, asymptotic theory, multivariate news impact curve.

JEL codes: C32, C51, C52

1. Introduction

Starting with the seminal work of Engle (1982) and Bollerslev (1986) for univariate models, and Bollerslev (1990) and Engle and Kroner (1995) for multivariate models, the modeling of conditional variances, covariances and correlations has attracted considerable interest in the risk and financial volatility literature. Several extensions and generalizations have been suggested for both the univariate and multivariate representations (see, for example, Bollerslev, Chou and Kroner (1992), Bollerslev, Engle and Nelson (1994), Li, Ling and McAleer (2002), McAleer (2005), and Bauwens, Laurent and Rombouts (2006)). The numerous proposed models have been applied to vastly different data sets including exchange rate forecasting, stock price volatility prediction, and market risk measurement through Value-at-Risk forecasts.

In comparison with the development of model specifications, the theoretical contributions have been limited. In fact, the conditions for the existence of a unique stationary and ergodic solution, and for the asymptotic theory of the parameter estimates have become available only for a subset of the proposed models (among others, see Bougerol and Picard (1992) and Ling and McAleer (2002a), (2002b) for univariate GARCH models, Comte and Lieberman (2003), and Ling and McAleer (2003) for multivariate GARCH – MGARCH henceforth – models)[1]. Furthermore, in the multivariate model case, the diagnostic checking of model adequacy is poorly covered in the literature, being restricted to some recent papers considering multivariate extensions of the well-known Ljung-Box test statistic (see Ling and Li, (1997) and Tsay (1998)).

One of the most important topics in the financial econometrics literature is the asymmetric behavior of conditional variances. The basic idea is that negative shocks have a different impact on the conditional variance evolution than do positive shocks of a similar magnitude. This issue was raised by Nelson (1990) in introducing the EGARCH model, and was also considered by Glosten, Jagannathan and Runkle (1992), Rabemananjara and Zakoian (1993) and Zakoian (1994) for the univariate case. For these models, some general results apply, including the conditions for stationarity and asymptotic theory for the quasi-maximum likelihood estimates (see Ling and McAleer, (2002a) and (2002b)). However, restricting attention to only a single asset may be too stringent, particularly if the primary goal is the measurement of the risk of an investment or a portfolio. In such cases, we could be interested in analyzing the effects of a shock on a set of assets, with a possible distinction between asset-specific shocks and market shocks.

---

[1] Note that the VARMA-GARCH model proposed by Ling and McAleer (2003) nests some other multivariate GARCH representations, including the CCC model of Bollerslev (1990). However, this class of models is non-nested with respect to the BEKK and Vech GARCH representations developed in Engle and Kroner (1995) (see also Caporin and McAleer, (2008)).

In addition to the possible mean effects, this paper focuses on the variance and covariance effects, monitoring an asset's conditional variance reaction to another asset's specific shock. For instance, the model we propose may be used to study the effects of an oil price shock on oil price volatility and on the volatilities of the stocks belonging to the auto sector, or the effects of a market shock (that could be represented by an unexpected macroeconomic shock) on the conditional variances of a set of stocks. Furthermore, in our modeling approach we will distinguish the effects of a shock's "sign" from these coming from the shock's "size". Note that the possible combination of sign and size may depend on the other asset's sign and size, with increased complexity according to the chosen multivariate framework. A related issue, the so-called 'leverage' effect (negative shocks should increase conditional variances while positive shocks should induce a reduction in the conditional variances) will not be addressed, given that our main focus is on conditional variance asymmetry[2].

Information on variance asymmetry could be useful for asset pricing, portfolio construction (given the relationship of such a shock-propagation mechanism with the asset correlations and their betas), and for market risk measurement (see, among others, Hafner and Herwartz (1998), Hansson and Hordahl (1998), and the references in Bauwens et al. (2006)). The structures needed to monitor, estimate and use the conditional variance asymmetries should be included in an appropriate multivariate model. A recent contribution in this direction was McAleer, Hoti and Chan (2009), who provided a multivariate generalization of the GJR model of Glosten et al. (1992). However, their approach is limited to a specific distinction between positive and negative shocks, and is based on an extension of the univariate analysis. The MGARCH literature includes several models with asymmetry, with interesting examples given in Kroner and Ng (1998) and De Goeij and Marquering (2004). In the cited papers, the asymmetry term enters either a Vech or BEKK representation (for definitions, see Engle and Kroner, (1995)). However, McAleer, Hoti and Chan (2009) seem to be the only authors to have dealt with asymmetry in the VARMA-GARCH model of Ling and McAleer (2003). Note that the VARMA-GARCH model nests the CCC model of Bollerslev (1990), so that the introduction of asymmetry in the VARMA-GARCH model of Ling and McAleer (2003) can be extended directly to the variance dynamics in the CCC model, and hence also the variance and covariance dynamics in the DCC model of Engle (2002) and the GARCC model of McAleer et al (2008).

---

[2] Note that the term 'leverage' is used by some authors to identify what we call 'asymmetry', that is, a different effect of negative and positive shocks of the same magnitude on the conditional variance (for an example, see Bauwens et al. (2006))

Despite the lack of theoretical contributions dealing with asymmetry and variance spillovers in the CCC and DCC models, the econometric literature includes several papers dealing with 'simple' GARCH(1,1) specifications and asymmetric effects in the correlations (see, among others, Cappiello, Engle and Sheppard (2006)). We do not follow this strand of literature as the effects of the variance misspecification on the correlation dynamics are not known, so that we prefer to generalise the variance dynamics. The model to be presented here could be extended following the GARCC model of McAleer et al. (2008) that allow for correlation dynamics and possibly also asymmetry in the correlations.

The purpose of this paper is to provide a general framework, in which both multivariate variance asymmetry and spillover effects are considered, to derive the conditions to ensure the existence of a unique stationary and ergodic solution, and to prove the consistency and asymptotic normality of the Quasi-Maximum Likelihood Estimator (QMLE) for the parameters of interest. In addition to the traditional asymmetric effect, we include time dependence in the asymmetric component of the variances, thereby extending the ideas of Caporin and McAleer (2006). We propose the Dynamic Asymmetric MGARCH (DAMGARCH) model that allows for time-varying asymmetry with spillover effects. The interactions between variances may depend both on a direct relation between the conditional variances (as in standard MGARCH models) and on spillover effects from the asymmetric component of the GARCH model. As DAMGARCH is a generalization of the DAGARCH model of Caporin and McAleer (2006), it inherits many of the properties of DAGARCH, namely the possibility of explaining asymmetry as well as persistence in asymmetry. DAMGARCH also represents a generalization of the VARMA-GARCH model of Ling and McAleer (2003). Therefore, it is non-nested with respect to the Vech class of models of Engle and Kroner (1995). Note that the DAMGARCH model generalizes the existing MGARCH models with asymmetry for the inclusion of spillovers in the asymmetry term, for the definition of asymmetry over a set of thresholds, and for the generalization of the indicator functions that drive the asymmetric effect.

We present a simple empirical analysis to compare a bivariate DAMGARCH model with basic CCC specifications where the conditional variances follow a standard GARCH(1,1) model, the asymmetric GJR-GARCH of Glosten et al. (1992), or the DAGARCH model of Caporin and McAleer (2006). The model proposed provides a higher likelihood and relevant insights into the asymmetric dynamics in the DAX and FTSE stock market indices.

Throughout the paper we use the following notation: ":" denotes horizontal matrix concatenation; $\rho(A)$ identifies the eigenvalue of matrix A with largest absolute value; *vec(A)* stacks the columns

of matrix A; *vecu(A)* stacks the columns of the lower triangular part of A below the main diagonal; *diag(a)* is a diagonal matrix with the vector a on the main diagonal; *dg(A)* is the vector containing the elements on the main diagonal of A; $\odot$ denotes the Hadamard matrix multiplication.

The remainder of the paper has the following structure. Section 2 defines the DAMGARCH model and considers three specific issues, namely the definition of thresholds (subsection 2.1), asymptotic properties of the model and of the QMLE (subsection 2.2), and estimation of DAMGARCH (subsection 2.3). In Section 3 we introduce the News Impact Surface and present a simulated example of the possible forms of the function, depending on the relations between the conditional variances. Section 4 presents an empirical analysis of two of stock market indices, comparing DAMGARCH with a set of CCC models. Section 5 gives some concluding comments.

2. DAMGARCH: Multivariate GARCH with Dynamic Asymmetry

In what follows, Y_t represents an *n*-dimensional vector of observable variables. For the moment, GARCH-in-mean effects are not considered, so it is assumed for simplicity that any mean component has been modeled adequately. The primary focus is on the mean residuals under the following equations:

$$Y_t = E\left[Y_t \mid I^{t-1}\right] + \varepsilon_t, \quad E\left[\varepsilon_t \mid I^{t-1}\right] = 0, \quad E\left[\varepsilon_t \varepsilon_t' \mid I^{t-1}\right] = \Sigma_t = D_t R_t D_t \tag{1}$$

in which I^{t-1} is the information set available at time $t-1$, $E\left[Y_t \mid I^{t-1}\right]$ is the conditional mean of Y_t, and ε_t is the *n*-dimensional mean residual vector at time t. The mean residuals have a conditionally time-dependent covariance matrix that can be decomposed into the contributions of the conditional variances and the conditional correlations[3]. Finally, D_t is a diagonal matrix of conditional volatilities, given by: $D_t = diag\left(\sigma_{1,t}, \sigma_{2,t}, ..., \sigma_{n,t}\right)$, and R_t is a (possibly time-dependent) correlation matrix. It is also assumed that the standardized and uncorrelated innovations, $\eta_t = \Gamma_t^{-1} D_t^{-1} \varepsilon_t$, are independent, with $\Gamma_t \Gamma_t' = R_t$. Note that, as distinct from standard practice, Γ_t, which is a full symmetric matrix, is not obtained by a Cholesky decomposition of the correlation

---

[3] It is implicitly assumed that the covariance dynamics are a by-product of the conditional variances and dynamic conditional correlations. Therefore, the Vech and BEKK representations (see Engle and Kroner (1995)) are not directly comparable with the model developed in this paper.

matrix. Differently, Γ_t comes from the eigendecomposition of the correlation matrix. In fact, $R_t = U_t \Delta_t U_t'$, where U_t is the matrix of eigenvectors and Δ_t is the diagonal matrix of eigenvalues. Using the fact that $U_t U_t' = I$, the identity matrix, Γ_t can be set to $\Gamma_t = U_t \Delta_t^{\frac{1}{2}} U_t'$. Finally, let $z_t = D_t^{-1} \varepsilon_t$ denote the standardized innovations, with R_t as the correlation matrix.

Define the vectors of conditional variances and squared innovations as $H_t = \left(\sigma_{1,t}^2, \sigma_{2,t}^2, \ldots \sigma_{n,t}^2 \right)' = dg\left(D_t D_t' \right)$ and $e_t = \left(\varepsilon_{1,t}^2, \varepsilon_{2,t}^2, \ldots \varepsilon_{n,t}^2 \right)'$, respectively. The following equations define the Dynamic Asymmetric MGARCH (hereafter DAMGARCH) model:

$$H_t = W + \sum_{i=1}^{s} B_i H_{t-i} + \sum_{m=1}^{r} \vec{G}_{t-m} , \tag{2}$$

where $\vec{G}_{t-m} = \sum_{j=1}^{l} \left\{ \left[A_{j,m} + \Psi_{j,m} G_{t-m-1} \right] I_j \left(\varepsilon_{t-m} \right) \left(\left(\varepsilon_{t-m} - \tilde{d}_j \right) \odot \left(\varepsilon_{t-m} - \tilde{d}_j \right) \right) \right\} \tag{3}$

and $G_{t-m} = \sum_{j=1}^{l} \left\{ \left[A_{j,m} + \Psi_{j,m} G_{t-m-1} \right] I_j \left(\varepsilon_{t-m} \right) \right\}, \tag{4}$

where, $B_i, i = 1, 2, \ldots, s$, $A_{j,m}$, $j = 1, 2, \ldots, l$, $m = 1, 2, \ldots, r$, $\Psi_{j,m}$, $j = 1, 2, \ldots, l$, $m = 1, 2, \ldots, r$, and G_t are n-dimensional square matrices, W and $\vec{G}_t$ are n-dimensional vectors, l is the number of subsets in which the support of the multivariate probability density function of ε_t has been partitioned (that is, there may be $l-1$ "threshold vectors"[4]). In addition, $I_j(\varepsilon_t)$ is a scalar (or a diagonal matrix) indicator function[5] (its structure will be further specified below) that verifies if the vector ε_t (each i component of ε_t) belongs to subset j of the joint support (of the marginal support), $\tilde{d}_j = 0$ or $\tilde{d}_j = \bar{d}_j$, where $\bar{d}_j$ is an n-dimensional vector that defines the upper (or lower) bounds of subset j

---

[4] The term "thresholds" is not appropriate when considering a multivariate density for which the threshold may be a vector with different components, as the marginal density may have different thresholds. In dealing with multivariate densities, reference will instead be made to a partition of the density support that defines some subsets.

[5] If a scalar, the function $I_j(\varepsilon_t)$ assumes the value 1 if the vector ε_t belongs to subset j, and 0 otherwise. If a diagonal matrix, each element in the diagonal of $I_j(\varepsilon_t)$ is an indicator function based on the marginal of the mean residuals, $I_j(\varepsilon_{i,t})$, and assumes the value 1 if the element $\varepsilon_{i,t}$ belongs to its specific subset j, and 0 otherwise.

(we will address below the structure of the subsets, the structure of $\bar{d}_j$ and its usefulness[6]). We highlight here that the vector $\tilde{d}_j = 0$ is a selection vector allowing a direct inclusion of the thresholds or bounds $\bar{d}_j$ in the GARCH equation. Note also that the vectors $\bar{d}_j$ characterize the thresholds used in the definition of the asymmetric components. These thresholds are not necessarily explicitly included in the GARCH equation (when $\tilde{d}_j = 0$), while they define the partitions of ε_t in all cases (and thus always enter in the functions $I_j(\varepsilon_t)$, as will be shown below). We call $\bar{d}_j$ the vectors of 'observed thresholds', and the definition will be motivated below. In the derivation of the asymptotic properties, we will use an alternative representation of the DAMGARCH model, which is given in Appendix A.1. The two terms $\vec{G}_{t-m}$ and G_{t-m} define the ARCH component of the model: the first is a function dependent on the thresholds and the second term, which in turn drives the dynamic asymmetry. Finally, we note that $\vec{G}_t$ in (3) is measurable with respect to the information set at time t, but it enters equation (2) with a lag of at least 1. Additional comments on the interpretation of these two elements are given below.

The indicator function, and therefore the number of thresholds (or number of subsets), can be defined not only on the $t-1$ innovation vectors, but also on a larger number of terms. In fact, it is possible to generalize $I_j(\varepsilon_t)$ to $I_j(\varepsilon_t, \varepsilon_{t-1}, ..., \varepsilon_{t-m})$. However, in this case, the number of thresholds (or subsets) may increase appreciably. Considering only the sign of the innovation, a single lag leads to $l = 2$, while the use of two lags leads to $l = 4$, with an exponential increase in the number of partitions.

Put differently, we can generalize equations (3) and (4) by increasing the number of lags for the terms G_t and $\vec{G}_t$. In this case, we can write:

$$G_{t-m} = \sum_{j=1}^{l} \left\{ \left[A_{j,m} + \Psi_{j,m}(L) G_t \right] I_j(\varepsilon_t) \right\},$$

$$\Psi_{j,m}(L) = \Psi_{j,m,1} L + \Psi_{j,m,2} L^2 . + ... + \Psi_{j,m,q} L^q,$$

(5)

---

[6] Note that the inclusion of $\bar{d}_j$ induces a continuous news impact surface, as will be shown below. The representation adopted here generalises the continuous news impact curve of Engle and Ng (1993) and Caporin and McAleer (2006) to the multivariate case.

with an obvious increase in the number of parameter matrices (similarly for $\vec{G}_t$). Section 2.3 considers the estimation problem and includes a discussion of the role of the number of parameters and feasible representations.

Note that equation (2) defines the dynamics of the conditional variances on the basis of (i) past conditional variances, and (ii) past squared innovations. While the first term represents the 'standard' GARCH component, the second term does not explicitly include the 'standard' ARCH component. In fact, the representation we choose can be recast with a slightly different structure, showing the asymmetric variance dynamic as an addition to the VARMA-GARCH structure of Ling and McAleer (2003). Actually, in a simple case, assuming $d_j=0$ $j=1,2$, $r=s=1$ (that is, where the thresholds do not directly enter in the ARCH component definition, which allows the omission of $\vec{G}_t$), we can write the ARCH coefficients of the model as follows:

$$
\begin{aligned}
H_t &= W + B_1 H_{t-1} + \left(\mathbf{A} + G_{t-1} \right) e_{t-1}, \\
G_t &= \sum_{j=1}^{l} \left\{ \left[A_j + \Psi_j G_{t-1} \right] I_j \left(\varepsilon_t \right) \right\},
\end{aligned}
\tag{6}
$$

so that we can rewrite the ARCH term as

$$
\mathbf{A} + G_t = \sum_{j=1}^{l} \left\{ \left[\mathbf{A} + A_j + \Psi_j G_{t-1} \right] I_j \left(\varepsilon_t \right) \right\}.
\tag{7}
$$

Equation (7) includes the traditional ARCH term $\mathbf{A}$, and additional ARCH matrices, the A_j $j = 1, 2..., l$ matrices, which are modifying the ARCH coefficients depending on the thresholds, and a dynamic asymmetric component G_{t-1}. Equation (7) also highlights that a sufficient condition for the identification of both the A_j (with $j = 1, 2..., l$) and the $\mathbf{A}$ matrices requires that at last one of the A_j $j = 1, 2..., l$ matrices must be set to zero. In this case, the matrices A_j $j = 1, 2..., l$ will define the differential effects on the conditional variances (ARCH component) of each subset with respect to a baseline subset. It may be considered a multivariate generalization of the GJR model where we have a standard ARCH coefficient and an additional coefficient that is added to the ARCH component only for negative shocks. In the DAMGARCH model, we have a baseline ARCH component referred, as an example, to small positive innovations and additional effects for negative small innovations, negative large innovations and positive large innovations. Furthermore, in

equation (3) we allowed for the direct dependence of the ARCH part from the thresholds in order to induce a continuous news impact curve, as in Engle and Ng (1993). Finally, the ARCH part of DAMGARCH can also be interpreted as the sum of two components: letting $r=1$ for simplicity, we may identify a standard ARCH part (threshold dependent and hence asymmetric) that is given by

$$\sum_{j=1}^{l} A_j I_j \left(\varepsilon_{t-1} \right) \left[\left(\varepsilon_{t-1} - \tilde{d}_j \right) \odot \left(\varepsilon_{t-1} - \tilde{d}_j \right) \right],$$

and a second term associated to the asymmetry persistence:

$$\sum_{j=1}^{l} \Psi_j G_{t-2} I_j \left(\varepsilon_{t-1} \right) \left[\left(\varepsilon_{t-1} - \tilde{d}_j \right) \odot \left(\varepsilon_{t-1} - \tilde{d}_j \right) \right].$$

This second term includes a component $\left(G_{t-2} \right)$ that carries the asymmetric effect up to time t-2 (depending on the information set at time t-2). The asymmetric behavior of time t conditional variances depends thus on past shocks (through the first term) but also on the asymmetry behavior in the previous period (through the second term). The pattern of G_t will capture the time-varying asymmetry effects over the conditional variances.

Put differently, we can define DAMGARCH as a MGARCH model, in which the time-varying ARCH coefficients, $A_j + \Psi_j G_{t-1}$, $j=1,2,...l$, depend on the partition to which time t-1 shock vector belongs, namely the A_j matrices, and on an autoregressive component that drives the persistence in the ARCH coefficients, as parameterized by the Ψ_j matrices. Caporin and McAleer (2006) provide a detailed discussion of the interpretation of the DAGARCH coefficients, which can be generalized directly to the DAMGARCH model.

A deeper discussion of the indicator function is required. We propose two alternative structures, which are defined over the multivariate density of the mean innovation vector, ε_t, and over the marginal densities of the univariate mean innovations, $\varepsilon_{i,t}$, respectively.

Consider the use of the multivariate density. In this case, define $S \subseteq \mathbb{R}^n$ as the support of the multivariate innovation density, so that:

$$I_j \left(\varepsilon_t \right) = \begin{cases} 1, & \varepsilon_t \in S_j \\ 0, & \text{otherwise,} \end{cases} \tag{8}$$

where S_j is a subset of S. Furthermore, we have

$$\bigcup_{j=1}^{l} S_j = S, \quad S_i \bigcap S_j = \emptyset, \quad i, j = 1, 2, ..., l, \quad i \neq j. \tag{9}$$

As an example, we may define the following three subsets of S:

$$
\begin{aligned}
S_1 &= \left\{ \varepsilon_t : \varepsilon_{i,t} < \overline{d}_L, i = 1,2...n \right\}, \\
S_3 &= \left\{ \varepsilon_t : \varepsilon_{i,t} > \overline{d}_U, i = 1,2...n \right\}, \\
S_2 &= S - S_1 - S_3.
\end{aligned}
\tag{10}
$$

In this example, assuming that $\overline{d}_L$ is a small negative number and $\overline{d}_U$ is a large positive number, the partition distinguishes extreme events from the remaining elements of S. The direct dependence of variances from thresholds (see equation 3) was introduced in order to induce continuity of the news impact to conditional variances. When the indicator functions are defined over the multivariate density support, continuity may not be simply achieved. In fact, thresholds may have a complex representation (see the examples in Appendix A.2). In these cases, we must set $\tilde{d}_j = 0$ and continuity may be obtained appropriately defining the A coefficient matrices associated to the partition. In the example of equation (10), we should set $A_1 = A_2 + A_a$ and $A_3 = A_2 + A_b$ where A_a and A_b are parameter matrices. Note this is similar to having a standard ARCH coefficient and two additional components associated with extreme events of either positive or negative sign and a baseline ARCH component A_2. Note that the thresholds in this case bracket vectors.

Put differently, we may define the thresholds over the marginal densities of the innovations. In this case, we may define the $I_j(\varepsilon_t)$ function as a diagonal matrix of dimension n, with $I_j(\varepsilon_{i,t})$ on the main diagonal. In turn, $I_j(\varepsilon_{i,t})$ is the indicator function for the inclusion of $\varepsilon_{i,t}$ in the j-th subset defined over the probability density support of $\varepsilon_{i,t}$. The $I_j(\varepsilon_{i,t})$ indicator function is the univariate counterpart of equation (8), namely:

$$\text{if } \tilde{d}_j = \overline{d}_j \qquad \begin{aligned} I_j\left(\varepsilon_{i,t}\right) &= \begin{cases} 1, & \varepsilon_{i,t} < \overline{d}_{i,j} \\ 0, & \text{otherwise} \end{cases}, \qquad j = 1, 2, \ldots, k \\[2mm] I_j\left(\varepsilon_{i,t}\right) &= \begin{cases} 1, & \varepsilon_{i,t} > \overline{d}_{i,j-1} \\ 0, & \text{otherwise} \end{cases}, \qquad j = k+1, \ldots, l \end{aligned} \qquad (11.a)$$

$$\text{if } \tilde{d}_j = 0 \qquad I_j\left(\varepsilon_{i,t}\right) = \begin{cases} 1, & \overline{d}_{i,j-1} < \varepsilon_{i,t} \le \overline{d}_{i,j} \\ 0, & \text{otherwise} \end{cases}, \qquad j = 1, 2, \ldots, l \qquad (11.b)$$

where the subset is expressed as a segment on the support of the probability density function of $\varepsilon_{i,t}$. Furthermore, for $j = 1$ (that is, the first subset), the condition in (11a,b) is $\varepsilon_{i,t} \le \overline{d}_{i,1}$, while for $j = l$ (that is, the last subset), the condition becomes $\varepsilon_{i,t} > \overline{d}_{i,l-1}$ with $\overline{d}_{i,1} < \ldots < \overline{d}_{i,k-1} < 0 < \overline{d}_{i,k+1} < \ldots < \overline{d}_{i,l-1}$, that is, the *k-th* threshold is equal to zero for all variables. The last assumption is imposed in order to simplify the model structure. Finally, the indicator function distinguishes positive and negative values in order to induce continuity in the news impact surface, which will be defined below. Note that by defining the thresholds over the marginal support we skip possible problems associated with the bracketing of vectors, given that we bracket single elements of the innovation vector. Furthermore, a single vector may have elements satisfying different functions $I_j(\varepsilon_t)$ without any constraint, adding flexibility to the model.

Note that the 'observed' threshold equivalence (apart from the zero threshold) over different components of the innovations was not imposed in (11), implying that we can have different 'observed' threshold values for each $\varepsilon_{i,t}$. In the following section, we will show that the indicator function defined in (8) may be more general than that defined in (11). Furthermore, we will show how to represent the indicator function that follows (11) in the form of (8) by defining the partitions of ε_t density function support appropriately.

If we follow (11) in defining the indicator functions, then the elements of $\overline{d}_j$ may be different over the variables and are defined accordingly to the structure of $I_j(\varepsilon_{i,t})$, namely $\overline{d}_j = \left\{\overline{d}_{j,1}, \ldots, \overline{d}_{j,l}\right\}$ for $j = 1, \ldots, k-1$, $\overline{d}_j = \left\{\overline{d}_{j-1,1}, \ldots, \overline{d}_{j-1,l}\right\}$ for $j = k+2, \ldots, l$, and $\overline{d}_j = 0_{n \times 1}$ for $j = k, k+1$. Under (8), the definition of $\overline{d}_j$ depends on the relations used to define the S_j subsets. In the example in (10), we have $\overline{d}_1 = \overline{d}_L i_n$, $\overline{d}_2 = 0_n$ and $\overline{d}_3 = \overline{d}_U i_n$, where i_n is an n-dimensional vector of ones and 0_n is an n-dimensional vector of zeros. However, note that the definition of the $\overline{d}_j$ elements may be more

complex under (8) than in (11), as will be shown below. Finally, note that the indicator function defined on the joint probability support can be represented in matrix form (instead of the scalar case previously used) by simply replacing one with an identity matrix.

The development of the DAMGARCH model is similar in spirit to Ling and McAleer (2003), McAleer et al. (2007) and McAleer et al. (2009). In fact, assuming a constant correlation matrix, and imposing the condition that $\vec{G}_{t-j} = A_j e_{t-j}$ (an n-dimensional square parameter matrix that is not influenced by asymmetric behavior) yields the VARMA-GARCH model of Ling and McAleer (2003). Moreover, the GARCC model proposed in McAleer et al. (2008) could be obtained assuming a time-dependent structure for the conditional correlation matrix, again under the restriction $\vec{G}_{t-j} = A_j e_{t-j}$.

A related approach was used in McAleer et al. (2009) for the introduction of asymmetric conditional variances in the MGARCH framework. In this case, the appropriate matrix is given by:

$$\vec{G}_{t-j} = \left[\mathbf{A} + A_1 I_1 \left(\varepsilon_{t-j} \right) \right] e_{t-j},$$

in addition, the matrix indicator function is defined as in (11), with a single threshold set to zero for all $\varepsilon_{i,t}$ and with the explicit inclusion of the ARCH parameter matrix. This model is also the multivariate counterpart of the GJR model of Glosten, Jagannathan and Runkle (1992) with constant correlations. Its representation using a partition over the support of ε_t includes 2^n subsets associated with all the possible combinations of positive and negative values in the elements of ε_t.

As the DAMGARCH model can nest all the previous cases, it follows that the CCC model of Bollerslev (1990), the DCC model of Engle (2002), and the varying conditional correlation (VCC) model of Tse and Tsui (2002), may be also considered as special cases of DAMGARCH. The CCC model is obtained by setting $\vec{G}_{t-j} = A_j e_{t-j}$, assuming that the matrices A_j and B_j are diagonal and that the model has a constant conditional correlation matrix.

The DAMGARCH model extends current multivariate representations of GARCH by introducing multiple thresholds and time-dependent asymmetry. However, DAMGARCH has a similar limitation of the standard multivariate representation, namely the problem of (high) dimensionality[7]. In order to resolve this problem, diagonal representations can be used, such as a separate univariate DA-GARCH model for each innovation variance. Diagonality implies that all the parameter

---

[7] See Table 1 for parameter number of several MGARCH models.

matrices are diagonal, while no restrictions are imposed on the thresholds, which could differ according to the variables involved. Furthermore, block structures could be considered, as in Billio et al. (2006). In that case, the parameter matrices could be partitioned and restricted on the basis of a particular asset classification.

2.1. Defining Thresholds and Model Specifications

As given in Caporin and McAleer (2006), the use of multiple thresholds with time-varying conditional variances may create problems in the definition of thresholds. In fact, if the thresholds are designed to identify the queues of the innovation density, they must be defined over the standardized innovations, as the thresholds should adapt to movements in the conditional variances. Consider a simple example in which a time series follows a GARCH(1,1) process, but without any mean dynamics. If we focus on the upper α-quantile of the mean distribution, this quantile is a function of the conditional variance and of the quantile of the standardized innovation density. Thus, in univariate representations, thresholds have to be defined over the standardized innovation, either by fixing a set of values or a set of percentiles a priori.

Continuing with this example, assume that the lowest threshold for mean innovations, ε_t, is fixed at $\bar{d}_L$, so that the indicator function for this case is $I(\varepsilon_t) = 1(\varepsilon_t < \bar{d}_L)$. The probability associated with this indicator function gives:

$$P\left(\varepsilon_t < \bar{d}_L\right) = P\left(z_t \sigma_t < \bar{d}_L\right) = P\left(z_t < \sigma_t^{-1} \bar{d}_L\right) = F_z\left(\sigma_t^{-1} \bar{d}_L\right) \tag{12}$$

where $F_z(.)$ is the cumulative density of the standardized innovations, z_t. We note that the probabilities are functions of the conditional variance. Therefore, fixing a value for $\bar{d}_L$ is not equivalent to defining a quantile on the mean innovation probability density function. A more appropriate thresholds formulation should consider fixing their value over the standardized innovation density (see Caporin and McAleer (2006) for further details).

A similar structure is needed for multivariate representations, as thresholds must then be defined over standardized innovations. However, a further difficulty arises with regard to the definition of thresholds according to the joint or marginal densities. The two approaches are equivalent if and only if there is zero correlation among the variables. For this reason, we believe that thresholds have to be defined over the standardized and uncorrelated innovations, that is, on the innovations defined as $\eta_t = \Gamma_t^{-1} D_t^{-1} \varepsilon_t$, where Γ_t is a symmetric matrix obtained from the spectral decomposition of the

R_t correlation matrix satisfying $\Gamma_t \Gamma_t^{'} = R_t$. Making a parallel with simultaneous equation systems, the shocks η_t may be compared with the structural shocks and, depending on their values, they affect the observed shocks, ε_t, and their variance dynamics.

Following this statement, we note that the observed thresholds $\bar{d}_j$ defined over the mean innovations ε_t will be time dependent: in fact we define them as $\bar{d}_{j,t} = D_t \Gamma_t d_j$, where d_j represents the vector of 'structural' thresholds defines with respect to η_t. Observed thresholds may be time dependent, given that they thresholds are a function of conditional standard deviations and conditional correlations, both of them possibly time dependent. Differently, the structural thresholds are assumed to be time independent

In the following, it is assumed that the thresholds are fixed over the probability density function of the η_t. The structural thresholds d_j can be fixed a priori or determined by a quantile relation, $d_j = F_i^{-1}(\alpha)$[8]. Furthermore, the term 'thresholds' will be used only with respect to the marginal densities, while the term 'support partitions' will be used with respect to the joint density. Note that the introduction of a threshold 'structure' on marginal or joint densities will be equivalent only in special cases, namely when the correlations are all equal to zero.

Thresholds and partitions can be defined as follows. Consider first the definition of thresholds over marginal densities. Assume that the thresholds are fixed over the components of η_t. Finally, define $F(.)$ as the joint cumulative density, and $F_i(.)$, $i = 1, 2, ..., n$, as the marginal cumulative densities of the η_t. It follows that:

$$I_j\left(\varepsilon_{i,t}\right) = \begin{cases} 1, & D_t \Gamma_t d_{j-1,i} < \varepsilon_{i,t} \le D_t \Gamma_t d_{j,i}, \\ 0, & \text{otherwise} \end{cases} \quad j = 1, 2, ..., l \tag{13}$$

where d_j is the vector of structural thresholds defined over the η_t innovations. Note that the condition in equation (13) is based on the elements of a time-dependent threshold vector, so that the indicator matrix function is given by $I_j(\varepsilon_t) = diag(I_j(\varepsilon_{1,t}), I_j(\varepsilon_{2,t}), ..., I_j(\varepsilon_{n,t}))$. Finally, for $j = 1$, the condition is $\varepsilon_{i,t} \le D_t \Gamma_t d_{1,i}$, while for $j = l$, the condition is $\varepsilon_{i,t} > D_t \Gamma_t d_{l,i}$. Equation (13) refers to the indicator functions as defined in (11.b). The previous comments and the discussion on threshold

---

[8] Note that the standardised innovations are also uncorrelated, so that the thresholds and the quantiles may be defined over either the marginal or the joint distribution function.

definition are valid also for the indicator functions of (11.a). The correlation matrix decomposition we used, differs from the standard Cholesky one. In fact, we preferred to consider a symmetric decomposition based on eigenvalues and eigenvectors in order to exclude the ordering effects induced by the Cholesky decomposition.

Put differently, the partition over the joint density of ε_t is defined as:

$$I_j\left(\varepsilon_t\right)=\begin{cases} 1, & \overline{d}_{j-1}<\varepsilon_t\leq\overline{d}_j \\ 0, & \text{otherwise} \end{cases}, \quad j=1,2,...,l \tag{14}$$

where the condition is satisfied if and only if the vector ε_t is included in the partition of the joint probability support. Specifically, equation (14) is equivalent to equation (7), as we can write the subset as:

$$S_j=\left\{\varepsilon_t:\overline{d}_{j-1,i}<\varepsilon_{i,t}\leq\overline{d}_{j,i}, \quad i=1,2,...,n\right\}, \quad j=1,2,...,l. \tag{15}$$

Note that equation (13) and(14) are not equivalent representations and is possible to move from one to the other when the correlations are all equal to zero and only in very special cases. Consider a bivariate example to illustrate the point. Assume that the correlation between the two variables is equal to zero. Then, the following figure represents a partition that can be obtained using either the marginal or the joint threshold definition (specifically, a single threshold that is set to zero):

[Insert Figure 1 here]

For the marginal threshold case, we have $l=2$, and a single threshold that is set equal to zero. For the joint partition, we have $l=4$, with each subset identifying a quadrant of the Cartesian plane. However, Figure 2 represents a support partition which is defined under the joint probability, but which cannot be obtained using the marginal threshold definition.

[Insert Figure 2 here]

This partition distinguishes between the cases where both variables are negative and the remaining combinations.

The fact that equations (13) and (14) are equivalent does not mean that the models defined over the joint or the marginal thresholds are also equivalent. In fact, the representation (14) over the joint support is associated with a more flexible model. In the case of the marginal thresholds, it follows that:

$$G_t = \left[A_1 + \Psi_1 G_{t-1} \right] I_1(\varepsilon_t) + \left[A_2 + \Psi_2 G_{t-1} \right] I_2(\varepsilon_t), \tag{16}$$

whereas over the joint support, it follows that:

$$G_t = \sum_{j=1}^{4} \left[\tilde{A}_j + \tilde{\Psi}_j G_{t-1} \right] \tilde{I}_j(\varepsilon_t). \tag{17}$$

The two representations are based on the same joint support partition. However, the second representation is more flexible since it allows different variance reactions for each of the four subsets of the Cartesian plane. The two equations are equivalent under the following parametric restrictions: let **x** be the first component and **y** the second, such that $j = 1$ identifies the subset with both components negative, $j = 4$ identifies the subset with both components positive, $j = 2$ defines the subset with positive **x** and negative **y,** and $j = 3$ defines the subset with negative **x** and positive **y.** It follows that (16) and (17) are equivalent if:

$$\tilde{A}_2 = \left[\left[\tilde{A}_4 \right]_{.,1} \quad \left[\tilde{A}_1 \right]_{.,2} \right], \qquad \tilde{\Psi}_2 = \left[\left[\tilde{\Psi}_4 \right]_{.,1} \quad \left[\tilde{\Psi}_1 \right]_{.,2} \right],$$

$$\tag{18}$$

$$\tilde{A}_3 = \left[\left[\tilde{A}_1 \right]_{.,1} \quad \left[\tilde{A}_4 \right]_{.,2} \right], \qquad \tilde{\Psi}_3 = \left[\left[\tilde{\Psi}_1 \right]_{.,1} \quad \left[\tilde{\Psi}_4 \right]_{.,2} \right],$$

where ".,1" denotes the first column of a matrix. Appendix A.2 includes two additional examples on partitions defined over the joint support.

When the correlations are not zero, the transformation of structural thresholds into observed thresholds may create non-contiguous (or not dense) sets in the support of ε_t that makes the bracketing of vectors as in equation (14) almost impossible. Differently, when thresholds are defined over marginal densities, the only effect of the correlations on the conditional variances is through the thresholds themselves. In fact, in the limiting case of diagonal specifications the conditional variances are close to be driven by univariate DAGARCH models, given that the only

link is in the thresholds (diagonal specifications exclude any spillover effect in the GARCH coefficients as well as in the ARCH and asymmetry terms). Finally, note that if the diagonal specification is coupled with uncorrelated standardized residuals η_t, then the DAMGARCH model collapses exactly on a collection of univariate DAGARCH models (now there are no links between the conditional variances and the standardized residuals).

Within the DAMGARCH model, the thresholds are not endogenous but must be fixed a priori on the basis of a distributional assumption for the structural residuals η_t. As an example, for the trivariate case, assuming multivariate normality and noting that η_t is a vector of independent innovations, we may define three threshold vectors:

$$
d_1 = \begin{bmatrix} z_\alpha \\ z_\alpha \\ z_\alpha \end{bmatrix} \quad
d_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad
d_3 = \begin{bmatrix} z_{1-\alpha} \\ z_{1-\alpha} \\ z_{1-\alpha} \end{bmatrix}
$$

where z_α identifies the α quantile of the univariate normal density. When the model has been estimated, the researcher may test the distributional assumptions, possibly update the beliefs, and re-estimate the DAMGARCH model. In addition, the thresholds may be defined on using the empirical densities of the η_t. In this case, an iterative estimation procedure should be use as we evidence in the estimation section.

2.2. Stationarity and Asymptotic Theory

In this paper, we focus on the variance model structure. The inclusion of ARMA mean components can be obtained using the results in McAleer et al. (2009). Furthermore, we assume a constant correlation matrix R, so that the extension to time-dependent correlations can be obtained as an extension of the results in McAleer et al. (2008). In the following, we provide the assumptions and the theorems stating the stationarity and the asymptotic properties of the DAMGARCH model. All the proofs are reported in Appendix A.4. The assumptions and the theorems are a direct extension of the results in Ling and McAleer (2003), and refer to the DAMGARCH model as defined in equations from (1) to (4), with one of the following restrictions: i) the model has no dynamic asymmetry effect (that is, the parameter matrices $\Psi_{1,1}...\Psi_{l,r}$ are all zero and thus the model collapses on a multiple threshold asymmetry specification which is a direct generalization of McAleer et al. (2009)); or, ii) the parameter matrices $A_{1,1}...A_{l,r}, \Psi_{1,1}...\Psi_{l,r}, B_1...B_s$ are all diagonal. The restrictions are needed to obtain the model structural properties as generalizations of the proofs

in Ling and McAleer (2003) and McAleer et al. (2009). The model parameter vector θ is defined as follows:

$$\theta = \begin{bmatrix} vec(W)' : vec(A_{1,1})' ... : vec(A_{l,r})' : vec(\Psi_{1,1})' ... : vec(\Psi_{l,r})' \\ : vec(B_1) ... : vec(B_s) : vecu(R) \end{bmatrix}'.$$

In addition, following McAleer et al. (2009) and McAleer et al. (2008), we assume that the parameter space Θ is a compact subspace of Euclidean space, such that θ is an interior point of Θ. Note that the thresholds are not included in the parameter vector.

Assumption 1: $E\left[Y_t \mid I^{t-1}\right] = 0$.

As a direct consequence of Assumption 1, the mean residuals are observable.

Assumption 2: The innovations $\eta_t = \Gamma^{-1} D_t^{-1} \varepsilon_t$ are independently and identically distributed. The structural thresholds are defined over the η_t. The structural thresholds are known.

As stated in Assumption 2, we assume the knowledge of structural thresholds which are fixed at quantiles of the underlying structural innovations, following the description in the previous section. The following additional assumptions are needed to derive the conditions to ensure the existence of a unique ergodic and stationary solution to the DAMGARCH model.

Assumption 3: R is a finite and positive definite symmetric matrix, with ones on the main diagonal and $\rho(R)$ having a positive lower bound over the parameter space Θ; all elements of B_i and $E\left[G_{t-1}^1 \tilde{z}_{t-j}\right]$ are non-negative $i=1,2,...s$, $j=1,2,...r$; W has elements with positive lower and upper bound over Θ; and all the roots of $\left| I - \sum_{i=1}^{r} E\left[G_{t-1}^1 \tilde{z}_{t-i}\right] L^i - \sum_{i=1}^{s} B_i L^i \right| = 0$ are outside the unit circle.

Assumption 4: $I - \sum_{i=1}^{r} E\left[G_{t-1}^1 \tilde{z}_{t-i}\right] L^i$ and $\sum_{i=1}^{s} B_i L^i$ are left coprime, and satisfy other identifiability conditions given in Jeantheau (1998) (the conditions are given in the proof to Theorem 3).

Theorem 1: *Under assumptions (1)-(3), the DAMGARCH model of equations (1)-(4) admits a unique second-order stationary solution, $\tilde{H}_t$, measurable with respect to the information set I^{t-1}, where I^{t-1} is a σ-field generated by the innovations $\tilde{z}_t$. The solution $\tilde{H}_t$ has the following causal expansion:*

$$\tilde{H}_t = W + \sum_{j=1}^{\infty} M' \left(\prod_{i=1}^{j} A_{t-i+1} \right) \xi_{t-j} \tag{19}$$

$$\underset{n\times(3nlr+ns)}{M'} = \left[\underset{n\times3nlr}{0} : I_n : \underset{n\times(s-1)}{0} \right] \tag{20}$$

$$A_t = \begin{bmatrix} \tilde{z}_t G_{t-1}^1 : \tilde{z}_t G_{t-1}^1 : \tilde{z}_t G_{t-1}^1 & \cdots & \tilde{z}_t G_{t-r}^1 : \tilde{z}_t G_{t-r}^1 : \tilde{z}_t G_{t-r}^1 & \tilde{z}_t B_1 & \cdots & \tilde{z}_t B_s \\ & 0_{(3nl-n)\times3nlr} & & & 0_{(3nl-n)\times ns} & \\ & I_{3nl(r-1)} & 0_{3nl(r-1)\times3nl} & & 0_{3nl(r-1)\times ns} & \\ G_{t-1}^1 : G_{t-1}^1 : G_{t-1}^1 & \cdots & G_{t-r}^1 : G_{t-r}^1 : G_{t-r}^1 & B_1 & \cdots & B_s \\ & 0_{n(s-1)\times3nlr} & & & I_{n(s-1)} & 0_{n(s-1)\times n} \end{bmatrix} \tag{21}$$

$$\xi_t = \left[(\tilde{z}_t W)', (i_{l-1} \otimes e_t)', G^{2,b'}, G_t^{2,c'}, 0_{3nl(r-1)\times1}, W', 0_{n(s-1)\times1} \right] \tag{22}$$

$$e_t = dg(\varepsilon_t \varepsilon_t') \qquad {}_1 e_j = dg(\bar{d}_j \bar{d}_j') \qquad {}_2 e_{t,j} = -2dg(\bar{d}_j \varepsilon_t') \tag{23}$$

$$\underset{(n\times nl)}{G_{t-m}^1} = \begin{bmatrix} \left(\left[A_{1,m} + \Psi_{1,m} G_{t-m-1} \right] I_1(\varepsilon_{t-m}) \right) : \left(\left[A_{2,m} + \Psi_{2,m} G_{t-m-1} \right] I_2(\varepsilon_{t-m}) \right) : \cdots \\ \cdots : \left(\left[A_{l,m} + \Psi_{l,m} G_{t-m-1} \right] I_l(\varepsilon_{t-m}) \right) \end{bmatrix} \tag{24}$$

$$\tilde{z}_t = diag\left(dg(z_t z_t') \right), \quad E[\tilde{z}_t] = I_n, \quad z_t = D_t^{-1} \varepsilon_t, \text{ and } D_t = diag\left(H_t^{1/2} \right) \tag{25}$$

where the quantity G_{t-1}^1 is defined in the alternative representation of the DAMGARCH model described in Appendix A.1. Hence, $\{Y_t, \tilde{H}_t\}$ are strictly stationary and ergodic.

The following theorem states the conditions to ensure the existence of moments for the DAMGARCH model.

Theorem 2: *Under assumptions (1) to (3), if $\rho\left(E\left[A_t^{\otimes b}\right]\right) < 1$, then the 2bth moments of Y_t are finite. Where b is a strictly positive integer, and $\otimes b$ denotes the Kroneker product of b matrices A_t defined in Theorem 1.*

We assume the coefficients are estimated by means of Quasi-Maximum Likelihood, following Bollerslev and Wooldridge (1992). A deeper discussion of the DAMGARCH estimation and the relevant implementation issues is included in section 2.3.

In order to prove the consistency of the QML estimates, we introduce the following assumption on logarithmic moments, as in Jeantheau (1998).

Assumption 5: For any $\theta \in \Theta$, we have $E_\theta\left[\log\left(|H_t|\right)\right] < 0$, where H_t is defined in (19).

The following two theorems define consistency and asymptotic normality of the quasi-maximum likelihood estimator for the DAMGARCH model parameters.

Theorem 3: *Define $\hat{\theta}$ as the quasi-maximum likelihood estimates of DAMGARCH. Under the conditions given by Jeantheau (1998) reported in Appendix A.4 and the theorems in Ling and McAleer (2003), we have $\hat{\theta} \xrightarrow{\ p\ } \theta$.*

Theorem 4: *Suppose that is Y_t generated by equations (1)-(4) satisfying assumptions (1)-(4). Given the consistency of the QMLE for DAMGARCH, under conditions 4.i), 4.ii) and 4.iii), we have $\sqrt{n}\left(\hat{\theta} - \theta\right) \xrightarrow{\ L\ } N\left(0, \Sigma_\theta^{-1}\Omega_\theta\Sigma_\theta^{-1}\right)$:*

4.i) $\dfrac{\partial^2 L}{\partial\theta\partial\theta'}$ exists and is continuous in an open and convex neighbor of θ;

4.ii) $n^{-1}\dfrac{\partial^2 L}{\partial\theta\partial\theta'}\bigg|_{\hat{\theta}}$ converges in probability to a finite non-singular covariance matrix

$\Sigma_\theta = E\left[n^{-1}\dfrac{\partial^2 L}{\partial\theta\partial\theta'}\right]$ *for any sequence $\hat{\theta}$ such that $\hat{\theta} \xrightarrow{\ p\ } \theta$;*

4.iii) $n^{-1} \dfrac{\partial L}{\partial \theta}\bigg|_{\hat{\theta}}$ *converges in law to a multivariate normal distribution* $N\left(0, \Omega_{\theta}\right)$, *with covariance*

matrix equal to $\Omega_{\theta} = \lim E\left[n^{-1} \dfrac{\partial L}{\partial \theta}\bigg|_{\hat{\theta}} \times \dfrac{\partial L}{\partial \theta'}\bigg|_{\hat{\theta}} \right]$.

2.3 Estimation

It was mentioned briefly in section 2.1 that the estimation of DAMGARCH could be considered through a quasi-maximum likelihood approach, following Bollerslev and Wooldridge (1992). This means that we can define an approximate likelihood function $L(\theta)$ that depends on the conditional covariance matrix, $L(\theta) = \sum_{t=1}^{T} l_t(\theta) = \sum_{t=1}^{T} l_t\left(\Sigma_t(\theta)\right)$. Traditionally, the approximate likelihood function is derived from a multivariate normal distribution. In the MGARCH literature, there also exists a two-step estimation approach that considers univariate estimation of the conditional variances and multivariate estimation of the correlation parameters, following Bollerslev (1990) and Engle (2002). It should be noted that the two-step approach cannot be used with DAMGARCH for two reasons: first, given the dependence of the conditional variance dynamics from the observed thresholds that, in turn, are defined over the conditional variances and correlations; second, by the inclusion of possible spillover effects across conditional variances in the traditional GARCH matrix. The two-step approach could be used only under the strong assumption of independence between the mean residuals, ε_t, and absence of spillovers in the GARCH component of the model.

Note that even when all the parameter matrices are diagonal (that is, when there are no spillovers between variables) but the correlations are not zero, the two-step approach cannot be used given that the observed thresholds still depend on the correlation matrix (the computation of thresholds requires knowledge of the correlation matrix).

The model complexity creates several implementation and numerical optimization problems, even with simplified and limited dimension systems. In order to reduce the computational burden, the following approximated estimation procedure is suggested. Note that full estimation can be considered at the cost of sensibly increasing computational time. Furthermore, numerical optimization problems could be reduced by implementing first-order derivatives, which will be considered in future extensions and applications of the current paper.

Recall that the number of variables is denoted by *n*. Thus, we suggest the following steps:

1) assume that the standardized and uncorrelated residuals are distributed according to a standard normal variables, fix the structural thresholds d_j for $j=1,2...l$ at theoretical quantiles of the normal distribution;

2) estimate a standard GARCH model on a univariate basis and save the conditional variances $_{GARCH}\sigma_{j,t}^2$, the standardized residuals $_{GARCH}\tilde{\eta}_{i,t} = \varepsilon_{i,t}\,_{GARCH}\sigma_{i,t}^{-1}$ for $i=1,2...n$, and the thresholds $_{GARCH}\bar{d}_{j,t}\left(\varepsilon_{i,t}\right) = \,_{GARCH}\sigma_{i,t}^{-1}d_j$ $j=1,2...l$;

3) estimate a univariate DAGARCH model (see Caporin and McAleer, (2006)) using the $_{GARCH}\sigma_{i,t}^{-1}d_j$ thresholds and save the conditional variances $_{DAGARCH}\sigma_{i,t}^2$ and the standardized residuals $_{DAGARCH}\tilde{\eta}_{i,t} = \varepsilon_{i,t}\,_{DAGARCH}\sigma_{i,t}^{-1}$ for $i=1,2...n$;

4) compute the unconditional correlation matrix (using the sample estimator) on the $_{DAGARCH}\tilde{\eta}_t = \left[\,_{DAGARCH}\tilde{\eta}_{1,t} : \,_{DAGARCH}\tilde{\eta}_{2,t} : ... _{DAGARCH}\tilde{\eta}_{n,t}\right]'$ series and save the correlation matrix R_n, the uncorrelated residuals $_R\eta_t = \left[\,_R\eta_{1,t} : \,_R\eta_{2,t} : ... _R\eta_{n,t}\right]'$ and the thresholds $_{DAMGARCH}\bar{d}_j = \Gamma_t D_t d_j$ $j=1,2...l$ (as defined in equation (11));

5) test using standard approaches the distribution assumption of step 1) and, if necessary, update the d_j thresholds; note that the threshold may be updated either modifying the distributional assumption or by computing them using the empirical model residuals.

If we assume that the model follows a diagonal specification in the GARCH conditional variance dynamics and all correlations estimated in step 4) are zero, then the previous steps allow the complete model estimation. The user just needs to validate the distributional assumptions in step 5), and, if needed, update the estimates of step 3) and 4). Some iterations of steps 3)-5) are needed if the thresholds are derived from empirical model residuals. Alternatively, the algorithm should proceed with the following steps. Note that Steps 1)-5) in this case are used to derive a reasonable vector of coefficient starting values, thereby reducing the computational time.

6) estimate the conditional variance parameters by fixing the correlation matrix and, using the thresholds defined in steps 4) and 5); then, save the conditional variances $_{DAMGARCH}\sigma_{i,t}^2$ and the standardized residuals $_{DAMGARCH}\tilde{\eta}_{i,t} = \varepsilon_{i,t}\,_{DAMGARCH}\sigma_{i,t}^{-1}$ for $i=1,2...n$;

7) compute the unconditional correlation matrix (using the sample estimator) on the

$$_{DAMGARCH}\tilde{\eta}_t = \left[_{DAMGARCH}\tilde{\eta}_{1,t} : _{DAMGARCH}\tilde{\eta}_{2,t} : \cdots _{DAMGARCH}\tilde{\eta}_{n,t} \right]' \text{ series and save the correlation}$$

matrix R_n, the uncorrelated residuals $_R\eta_t = \left[_R\eta_{1,t} : _R\eta_{2,t} : \cdots _R\eta_{n,t} \right]'$ and the thresholds

$$_{DAMGARCH}\bar{d}_j(\varepsilon_t) = \Gamma_t D_t d_j \quad j = 1,2...l \text{ (as defined in equation (11));}$$

8) test the distribution assumption of step 1) using standard approaches and, if necessary, update the d_j thresholds; note that the threshold may be updated either modifying the distributional assumption or by computing them using the empirical model residuals;

9) iterate steps 6) to 8) until convergence of the full model likelihood function (iterations are needed because we separate the estimation of conditional variance parameters from the estimation of correlations).

Given the parameter estimates, standard errors could be computed by numerical methods on the full system likelihood (that is, by the joint use of numerical gradient and Hessian computation in a Quasi-Maximum Likelihood approach, following Bollerslev and Wooldridge (1992)). Clearly, the proposed approach is suboptimal, but full systems estimation is likely to be viable only for small-dimensional systems. Notably, when full system estimation is considered, iterative estimation approaches should be in any case used if the initial assumption of innovation density is not supported by the data or if empirical thresholds are used.

The asymptotic properties of DAMGARCH are derived under an assumption of model identifiability and using a normal likelihood (that is, the QML standard approach). In the DAMGARCH model two different deviations from normality may be considered: leptokurtosis and asymmetry in the underlying shock density (different from the asymmetry in the conditional variances captured by the model). Given that the structural thresholds are based on the underlying innovations, model performance could be improved if the density used to determine thresholds is closer to the true density. For this purpose, steps 5) and 8) of the iterative estimation procedure allow the distributional assumption to be checked. Note that the asymptotic results are not affected by the misspecification of the likelihood function as they are derived within a QML framework.

At this point, a discussion of a feasible model structure and on the number of parameters is needed. The general model has a very high number of parameters: recall that n is the number of assets, s is the GARCH order, and r the ARCH order. Furthermore, $l-1$ is the number of thresholds (that is, we have l components in the asymmetric GARCH structure), and q is the order of the threshold function G_t. Therefore, the total number of parameters is: n for the conditional variance constants,

$n^2 \times s$ for the GARCH component, $n^2 \times (1+q) \times l$ for the threshold component, and $n \times (n-1)/2$ for the correlation matrix, namely $n + n^2 \times (s + l + l \times q) + n \times (n-1)/2$. Clearly, this is an intractable number of parameters, even for small dimensional systems. However, several restrictions could be considered: the use of diagonal parameter matrices (at the cost of excluding any spillover effects among the conditional variances, but allowing for an easier multi-step estimation procedure); introducing restrictions on the asymmetry dynamic (acting on the term G_t); fixing the number of thresholds at a small value, such as one ($l = 2$) for positive-negative or, as an example, to three ($l = 4$) for distinguishing among large and small positive (negative) values; or a combination of all of the above restrictions. Furthermore, we can expect that the standard GARCH orders should be small, possibly equal to 1; similarly, we may expect the threshold dynamics order to be small. Finally, note that if the model follows a pure ARCH dynamic (restricting s to zero), two-step estimation procedures are directly available.

Table 1 reports some examples, restricting to three the threshold number, imposing the standard GARCH orders, and fixing the asymmetry dynamics order to one. The number of DAMGARCH parameters is also compared with several alternative models. We show that the number of parameters in DAMGARCH is of order $O(n^2)$, namely the same order as the standard BEKK model, but lower than the order of the general Vech model, which is $O(n^4)$. Furthermore, the diagonal specification of DAMGARCH with common dynamics in the asymmetry has a parametric dimension that is comparable to that of the CCC model, but with additional interesting properties.

[Insert Table 1 here]

3. The News Impact Surface implied by DAMGARCH

Engle and Ng (1993) introduced the news impact curve, which is a useful tool for evaluating the effects of news on the conditional variances. The different reactions of the conditional variances to positive and negative shocks motivated the GJR and EGARCH representations of Glosten et al. (1992) and Nelson (1990), respectively. Both models permit a richer parameterization of the news impact curve as compared with the standard GARCH model. As an extension, Caporin and McAleer (2006) provided the news impact curve in the presence of multiple thresholds and dynamic asymmetry in conditional volatility.

This section provides a multivariate extension of the news impact curve for the DAMGARCH model. Without loss of generality, consider a simple model with two variables, structural shocks normally distributed, three thresholds set to zero, to the 5% and 95% quantiles, and all other orders restricted to one.

These values lead to the following DAMGARCH representation:

$$H_t = W + B_1 H_{t-1} + \vec{G}_{t-1},$$

(26)

$$\vec{G}_t = \sum_{j=1}^{4} \left\{ \left[A_j + \Psi_j G_{t-1} \right] I_j(\varepsilon_t) \left[(\varepsilon_t - \tilde{d}_j) \odot (\varepsilon_t - \tilde{d}_j) \right] \right\},$$

(27)

$$G_t = \sum_{j=1}^{4} \left\{ \left[A_j + \Psi_j G_{t-1} \right] I_j(\varepsilon_t) \right\}.$$

(28)

The parameter matrices have been set to the following specifications:

$$A_1 = \begin{bmatrix} 0.05 & 0 \\ 0.3 & 0.05 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0.05 & 0 \\ 0 & 0.05 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0.05 & 0 \\ 0 & 0.05 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0.05 & 0.3 \\ 0 & 0.05 \end{bmatrix}$$

(29)

$$W = \begin{bmatrix} 0.01 \\ 0.01 \end{bmatrix} \quad B_1 = \begin{bmatrix} 0.8 & 0.05 \\ 0.05 & 0.8 \end{bmatrix} \quad R = \begin{bmatrix} 1 & 0.7 \\ 0.7 & 1 \end{bmatrix} \quad \Psi = \begin{bmatrix} 0.10 & 0 \\ 0 & 0.10 \end{bmatrix}$$

Note that the first (second) asset's conditional variance depends on the second (first) asset's large positive (negative) shocks (see matrices A_1 and A_4). Furthermore, the two assets conditional variances are linked by a spillover effect (see matrix B_1). Finally, Ψ is constant over j.

Traditionally, the news impact curve represents the variance movements in response to an idiosyncratic shock, assuming that all past variances are evaluated at the unconditional variance implied by the model. For the simple GARCH(1,1) model, this implies:

$$NIC = \omega + \beta \bar{\sigma}^2 + \alpha \bar{\sigma}^2 z_t^2,$$

(30)

where z_t^2 represents the idiosyncratic component. In the DAMGARCH model, assuming that the correlations are constant, the news impact surface is given by:

$$NIS = W + BE[H] + \sum_{i=1}^{j} \left[\left(A_j + \Psi_j E[G] \right) diag \left(dg \left(\left[I_j \left(z_t \right) \left(z_t - \Gamma d_j \right) \left(z_t - \Gamma d_j \right)' \right] \right) \right) E[H] \right], \qquad (31)$$

where the derivation of the formula and the expectations are defined Appendix A.2. Note that the News Impact Surface is continuous by construction when thresholds are defined over the marginal density support as in (11.a). Differently, when thresholds are defined over the multivariate density support of η_t or when $\tilde{d}_j = 0$ the continuity of the News Impact Surface is not always guaranteed. As we have already argued, it may be induced by an appropriate definition of DAMGARCH parameter matrices (see section 2).

As an example, we report the News Impact Surfaces for the two asset example for two different cases: the first with the coefficients reported in (32), while the second is without any relation between the variances (i.e. all matrices are diagonal, excluding the correlation matrix).

[Insert Figures 3 to 6 here]

Note that when the assets are not correlated and there is no spillover or asymmetric behavior, the NIS collapses to the traditional News Impact Curve, as shown for the first asset in Figure 3 and for the second in Figure 4. When we introduce spillovers, the asymmetric component comes into play, modifying the NIS for the monitored assets depending on the shocks affecting the other asset (see Figures 5 and 6). If there is no spillover, the NIS for each asset collapses on the News Impact Curve of a univariate DAGARCH model.

4. Dynamic Asymmetric Effect: an empirical example

This section focuses on the estimation of the DAMGARCH model and its comparison with the simpler CCC-GARCH(1,1), CCC-GJR and CCC-DAGARCH models[9]. We consider the daily

---

[9] The conditional variances follows in these cases have the following specifications: CCC-GARCH(1,1) $\sigma_{i,t}^2 = \omega_i + \alpha_i \varepsilon_{t-1}^2 + \beta_i \sigma_{t-1}^2$; CCC-GJR $\sigma_{i,t}^2 = \omega_i + \alpha_i \varepsilon_{t-1}^2 + \gamma_i \varepsilon_{t-1}^2 I \left(\varepsilon_{t-1} < 0 \right) + \beta_i \sigma_{t-1}^2$; CCC-DAGARCH $\sigma_{i,t}^2 = \omega_i + \beta_i \sigma_{i,t-1}^2 + \sum_{j=1}^{4} \left(\overline{\gamma}_{i,j} + \overline{\phi}_{i,j} \gamma_{i,t-1} \right) I_{i,j} \left(z_{i,t-1} \right) \left(\varepsilon_{i,t-1} - \overline{d}_{i,j} \right)^2$ where $\overline{d}_{i,j} = d_{i,j} \sigma_{i,t-1}$, $z_{i,t} = \varepsilon_{i,t} \sigma_{i,t}^{-1}$, the thresholds $d_{i,j}$ are defined as the 5%, 50% and 95% quantiles of a standardized normal density, $\gamma_{i,t} = \sum_{j=1}^{4} \left(\overline{\gamma}_{i,j} + \overline{\phi}_{i,j} \gamma_{i,t-1} \right) I_{i,j} \left(z_{i,t} \right)$, $I_{i,j} \left(z_{i,t-1} \right) = 1 \left(z_{i,t-1} < d_{i,j} \right)$, $j = 1, 2$, and $I_{i,j} \left(z_{i,t-1} \right) = 1 \left(z_{i,t-1} > d_{i,j-1} \right)$, $j = 3, 4$.

closing levels of the DAX and FTSE 100 indices. The sample considered covers the period from 1998 to 2004 (1734 daily observations) and the data were downloaded from Datastream. The two markets are highly correlated and may show strong dependence in the extreme returns. Therefore, we may expect a NIS that is similar to that reported in the previous sections.

Table 5 reports the estimated coefficients of the DAMGARCH model, while Tables 2, 3 and 4 report the CCC-GARCH(1,1), the CCC-GJR and the CCC-DAGARCH estimates, respectively. Furthermore, Figures 7 and 8 report the conditional variances estimated by the three models for the FTSE index and the percentage differences between the CCC and DAMGARCH models. Figures 9 and 10 report the NISs for the DAX and the FTSE implied by the DAMGARCH model.

In this bivariate case we estimate a full DAMGARCH model with three thresholds, and all other orders set to 1. The thresholds were initially fixed at zero at the upper and lower 10% tails under a standardized normal distribution for the uncorrelated and variance standardized residuals. Following the estimation approach outlined in section 2, we verified the distribution of the empirical model residuals, which showed some deviation from normality (asymmetry for the FTSE returns and mild leptokurtosis for the DAX returns). Given this observation, we decided to fix the thresholds using empirical structural residuals and to iterate model estimation and definition of the thresholds until convergence of the likelihood function (the stopping rule was set to a change in the likelihood value lower than 1^{-4}). Convergence required only 4 iterations. Some descriptive statistics and the threshold values used in the estimation of DAMGARCH are included in Table 6. In addition, Figures 11 and 12 report the frequency histograms of the residuals. Graphical and descriptive analyses show the differences between the index innovations and the impact of the deviations from normality to the structural thresholds (and is more evident in FTSE).

Note that in the following, and in the Tables, the parameter matrices of DAMGARCH are matched with a subscript corresponding to the following partition of marginal innovations density:

1 - large negative values (below the lower threshold – 10% quantile);

2 - negative values;

3 - positive values;

4 - large positive values (above the upper threshold – 90% quantile).

Three thresholds were also used for the DAGARCH specifications and the coefficients subscript can be interpreted as for the DAMGARCH model.

The CCC models provide persistent conditional variance dynamics, a finding that is confirmed by the elevated values of the B matrix in the DAMGARCH model. The DAMGARCH model provides

significant coefficients in many parameter matrices, with the exception of the A matrix associated with positive innovations. Similar findings are observed in the CCC-DAGARCH estimates.

There is an evident interrelation between the two markets, in particular, for large positive shocks, and the correlation estimated by DAMGARCH is similar to that given by the CCC models.

Comparing the fitted conditional variances, we note some discrepancies, in particular, during periods of high volatility: the DAMGARCH peaks in the conditional variance seem, in some cases, to anticipate those produced by CCC models, an effect that may be due to an improved forecasting ability. Furthermore, the relative percent changes of CCC models with respect to the DAMGARCH ones reveals that, even if the patterns are very close (as we can observe in Figure 7), there are relevant differences, in some cases the variances are doubles while in other are halved. Notably, the differences are very high even comparing DAGARCH and DAMGARCH conditional variances.

Finally, the log-likelihood provided by the DAMGARCH model is much higher than that of the CCC models. Given there exists a nesting relationship between the models, empirical likelihood ratio tests can be used (in the following, p-values are derived presuming that the asymptotic density is the traditional one). These tests are in favor of the DAMGARCH model. The CCC-GJR model nests CCC-GARCH; in this case the test statistic has a value of 91.08, 2 degrees of freedom and a p-value of less than 1^{-10}. Differently, the likelihood comparison of CCC-GJR and DAMGARCH provides a test statistic equal to 82.15. The test has now 30 degrees of freedom: 2 in the B matrix (off diagonal coefficients are zero), 16 in the Ψ_i matrices (all coefficients restricted to zero), and 12 in the A_i matrices (8 for zero restrictions on the off-diagonal parameters plus 4 zero restrictions for diagonal elements in A_1 and A_4 – diagonal elements in A_2 and A_3 can be interpreted as the ARCH effect for negative and positive shocks). In this case, the P-value is less than 1^{-6} thereby supporting the inclusion of threshold asymmetry between variances. Furthermore, a comparison of the CCC-GJR and CCC-DAGARCH specifications is in favor of the former. It seems that the inclusion of multiple thresholds in the conditional variances is not useful, at least when we exclude any spillover across variances. Note that in the CCC-DAGARCH model we postulate that the thresholds do not depend on the correlation matrix. More interestingly, the comparison of CCC-GJR and DAMGARCH shows that the multivariate model with dynamic asymmetry, variance spillovers and thresholds correlation dependent provide relevant improvements with respect to traditional MGARCH models.

To conclude, the News Impact Surfaces reported in Figures 9 and 10 are similar to the example discussed in Section 3. They show that both the DAX and FTSE conditional variances depend on the other asset shocks, with a more pronounced effect in the case of DAX.

[Insert Figures 7 to 12 and Tables 2 to 6 here]

5. Concluding Remarks

This paper introduced a new MGARCH model, DAMGARCH, which generalized the VARMA-GARCH model of Ling and McAleer (2003) by introducing multivariate thresholds and time-dependent asymmetry in the ARCH component of the model. As a result, the proposed parameterization is able to explain variance asymmetry and threshold effects simultaneously with variance spillovers.

Furthermore, we provided the conditions for the existence of a unique stationary solution and, by generalizing the asymptotic theory in Ling and McAleer (2003), showed that the quasi-maximum likelihood estimators were consistent and asymptotically distributed as multivariate normal.

In addition, we presented the analytic form of the multivariate news impact curve, which was labelled the news impact surface, whose final purpose was a detailed graphical analysis of asymmetry and leverage effects.

In an illustrative empirical application, it was shown that the DAMGARCH model outperformed the standard CCC model in terms of the maximized log-likelihood values. An extended comparison of DAMGARCH and other more traditional MGARCH models is left for future research.

References

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

Billio, M., Caporin, M. and Gobbo, M., (2006), Flexible dynamic conditional correlation multivariate GARCH for asset allocation, *Applied Financial Economics Letters*, 2, 123-130.

Bollerslev, T., (1986), Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31, 307-327.

Bollerslev T., (1990), Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH approach, *Review of Economic and Statistics*, 72, 498-505.

Bollerslev T., Chou, R.Y. and Kroner, K. F., (1992), ARCH modeling in finance: a review of the theory and empirical evidence, *Journal of Econometrics*, 52, 5-59.

Bollerslev T., Engle, R. F. and Nelson, D. B., (1994), ARCH models. In R.F. Engle and D. McFadden, *Handbook of Econometrics*, Vol. 4, Elsevier, Amsterdam, pp. 2959-3038.

Bollerslev, T. and Wooldridge, J.M., (1992), Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances, *Econometric Reviews*, 11, 143-172.

Bougerol, P. And Picard, N., (1992), Stationarity of GARCH processes, *Journal of Econometrics*, 52, 115-127.

Caporin, M. And McAleer, M., (2006), Dynamic asymmetric GARCH, *Journal of Financial Econometrics*, 4, 385-412.

Caporin, M. And McAleer, M., (2008), Scalar BEKK and Indirect DCC, *Journal of Forecasting*,.27, 537-549

Cappiello L., Engle, R.F. And Sheppard, K., (2006), Asymmetric dynamics in the correlations of global equity and bond returns, *Journal of Financial Econometrics*, 4, 537-572.

Comte, F. and O. Lieberman, (2003), Asymptotic theory for multivariate GARCH processes, *Journal of Multivariate Analysis*, 84, 61-84

De Goeij, P. And Marquering, W., (2004), Modeling the conditional covariance between stock and bond returns: a multivariate GARCH approach, *Journal of Financial Econometrics*, 2, 531-564.

Engle, R.F., (1982), Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, 50, 987-1007.

Engle, R.F., (2002), Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models, *Journal of Business and Economic Statistics*, 20, 339-350.

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

Engle, R.F. and Ng, V., (1993), Measuring and testing the impact of news on volatility, *Journal of Finance*, 48, 1749-1778.

Glosten, L.R., Jagannathan, R. and Runkle, D.E., (1992), On the relation between the expected value and volatility of the nominal excess return on stocks, *Journal of Finance*, 46, 1779-1801.

Hafner, C., And Herwartz, H., (1998), Time-varying market price of risk in the CAPM approaches, empirical evidence and implications, *Finance*, 19, 93–112.

Hansson, B., And Hordahl, P., (1998), Testing the conditional CAPM using multivariate GARCH-M, *Applied Financial Economics*, 8, 377–388.

Kroner, F., And Ng, V., (1998), Modelling asymmetric comovements of asset returns, *Review of Financial Studies*, 11, 817–844.

Jeantheau, T., (1998), Strong consistency of estimators for multivariate ARCH models, *Econometric Theory*, 14, 70-86.

Li, W.K., Ling, S. and McAleer, M., (2002), Recent theoretical results for time series models with GARCH errors, *Journal of Economic Surveys*, 16, 245-269

Ling, S. and Li, W.K., (1997), Diagnostic checking of nonlinear multivariate time series with multivariate ARCH errors, *Journal of Time Series Analysis*, 18(5), 447-464.

Ling, S. and McAleer, M., (2002a), Stationarity and the existence of moments of a family of GARCH processes, *Journal of Econometrics*, 106, 109-117.

Ling, S. and McAleer, M., (2002b), Necessary and sufficient moment conditions for the GARCH(r,s,) and asymmetric power GARCH(r,s) models, *Econometric Theory*, 18, 722-729.

Ling, S. and McAleer, M., (2003), Asymptotic theory for a vector ARMA-GARCH model, *Econometric Theory*, 19, 278-308.

McAleer, M., (2005), Automated inference and learning in modeling financial volatility, *Econometric Theory*, 21, 232-261.

McAleer, M., Chan, F., Hoti S. and Lieberman, O., (2008), Generalized autoregressive conditional correlation, *Econometric Theory*, 24, 1554-1583.

McAleer, M., F. Chan and D. Marinova, (2007), An econometric analysis of asymmetric volatility: theory and application to patents, *Journal of Econometrics*, 139 (2), 259-284.

McAleer, M., Hoti S. and Chan, F., (2009), Structure and asymptotic theory for multivariate asymmetric conditional volatility, *Econometric Reviews*, 28 (5), 422-440.

Nelson, D.B., (1990), Conditional heteroskedasticity in asset pricing: a new approach, *Econometrica*, 59, 347-370.

Rabemananjara R. and Zakoian, J.M., (1993), Threshold ARCH models and asymmetries in volatility, *Journal of Applied Econometrics* 8, 31-49.

Tsay, R.S., (1998), Testing and modelling multivariate threshold models, *Journal of the American Statistical Association*, 93 (443), 1188-1202.

Tse, Y.K. and Tsui, A.K.C., (2002), A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations, *Journal of Business and Economic Statistics*, 20, 351-362.

Zakoian, J.M., (1994), Threshold heteroskedastic functions, *Journal of Economic Dynamics and Control*, 18, 931-955.

Appendix A.1 Alternative representation of the DAMGARCH model

Equations (2) - (4) can be represented in an alternative way as follows:

$$H_t = W + \sum_{i=1}^{s} B_i H_{t-i} + \sum_{m=1}^{r} G^1_{t-m} G^2_{t-m} \,,$$

$$G^1_{t-m} \atop {(n \times nl)} = \left[\begin{array}{l} \left(\left[A_{1,m} + \Psi_{1,m} G_{t-m-1} \right] I_1 \left(\varepsilon_{t-m} \right) \right) : \left(\left[A_{2,m} + \Psi_{2,m} G_{t-m-1} \right] I_2 \left(\varepsilon_{t-m} \right) \right) : \ldots \\ \ldots : \left(\left[A_{l,m} + \Psi_{l,m} G_{t-m-1} \right] I_l \left(\varepsilon_{t-m} \right) \right) \end{array} \right],$$

$$G^2_{t-m} \atop {(nl \times 1)} = \left[\begin{array}{l} dg \left(\left(\varepsilon_{t-m} - \tilde{d}_1 \right) \left(\varepsilon_{t-m} - \tilde{d}_1 \right)' \right)' : dg \left(\left(\varepsilon_{t-m} - \tilde{d}_2 \right) \left(\varepsilon_{t-m} - \tilde{d}_2 \right)' \right)' : \ldots \\ \ldots : dg \left(\left(\varepsilon_{t-m} - \tilde{d}_l \right) \left(\varepsilon_{t-m} - \tilde{d}_l \right)' \right)' \end{array} \right]',$$

$$G_{t-m} = G^1_{t-m} \left(i_l \otimes I_n \right).$$

where we used the equivalence $\left(\varepsilon_t - \tilde{d}_j \right) \odot \left(\varepsilon_t - \tilde{d}_j \right) = dg \left(\left(\varepsilon_t - \tilde{d}_j \right) \left(\varepsilon_t - \tilde{d}_j \right)' \right)$. Furthermore, using the following relations,

$$dg \left(\left(\varepsilon_t - \tilde{d}_j \right) \left(\varepsilon_t - \tilde{d}_j \right)' \right) = dg \left(\varepsilon_t \varepsilon_t' + \tilde{d}_j \tilde{d}_j' - \varepsilon_t \tilde{d}_j' - \tilde{d}_j \varepsilon_t' \right) = dg \left(\varepsilon_t \varepsilon_t' \right) + dg \left(\tilde{d}_j \tilde{d}_j' \right) - dg \left(\varepsilon_t \tilde{d}_j' \right) - dg \left(\tilde{d}_j \varepsilon_t' \right)$$

$$= dg \left(\varepsilon_t \varepsilon_t' \right) + dg \left(\tilde{d}_j \tilde{d}_j' \right) - 2 dg \left(\varepsilon_t \tilde{d}_j' \right) = e_t + {}_1 e_j + {}_2 e_{t,j},$$

we may rewrite G^2_{t-m} as

$$G^2_{t-m} = \left\{ \left(e_{t-m} + {}_1 e_1 + {}_2 e_{t-m,1} \right) : \left(e_{t-m} + {}_1 e_2 + {}_2 e_{t-m,2} \right) : \ldots : \left(e_{t-m} + {}_1 e_l + {}_2 e_{t-m,l} \right) \right\},$$

highlighting the fact that the component including the innovation contains an element which is time varying but constant over all partitions, a second element time invariant but changing across partitions and a cross term which is time varying and varying across partitions. Using this last result, the following representation can be derived:

$$H_t = W + \sum_{i=1}^{s} B_i H_{t-i} + \sum_{m=1}^{r} \left(G_{t-m}^1 G_{t-m}^{2,a} + G_{t-m}^1 G^{2,b} + G_{t-m}^1 G_{t-m}^{2,c} \right),$$

$$G_{t-m}^{2,a} = i_l \otimes e_{t-m}, \qquad G^{2,b} = \left\{ {}_1e_1 : {}_1e_2 : \dots : {}_1e_l \right\}, \qquad G_{t-m}^{2,c} = \left\{ {}_2e_{t-m,1} : {}_2e_{t-m,2} : \dots : {}_2e_{t-m,l} \right\}.$$

Appendix A.2 Possible partitions defined over the joint support

The flexibility of the partitions and of the models defined directly over the joint support accommodates particular representations, such as that depicted in Figure 3, which focuses on very extreme events. A natural question that may arise is the identification of common shocks or common components.

[Insert Figure A.1 here]

Finally, the partitions defined over the joint probability support may also accommodate non-linear relations between assets. A simple example is the distinction between extreme events of an elliptical multivariate distribution, as depicted in Figure A.1. This may be interesting for cases with constant correlations and thresholds defined over the standardised but correlated innovations.

[Insert Figure A.2 here]

Appendix A.3: Derivation of the News Impact Surfaces and of the unconditional estimates

Assume $s=1$ and $r=1$, and that the indicator function is defined over the marginal densities, so that $I_j(\varepsilon_t)$ is a diagonal matrix. The model representation is given by

$$H_t = W + B_1 H_{t-1} + \vec{G}_{t-1},$$

$$\vec{G}_{t-1} = \sum_{j=1}^{l} \left[A_j + \Psi_j G_{t-2} \right] I_j(\varepsilon_{t-1}) dg\left(\left(\varepsilon_{t-1} - \tilde{d}_j \right)\left(\varepsilon_{t-1} - \tilde{d}_j \right)' \right),$$

$$G_{t-1} = \sum_{j=1}^{l} \left[A_j + \Psi_j G_{t-2} \right] I_j(\varepsilon_{t-1}).$$

Lemma A.1

The focus of the indicator function refers to subset j: $I_j(\varepsilon_t)$. The following equality holds:

$$I_j(\varepsilon_t) = diag\left(I_j(\varepsilon_{1,t}), I_j(\varepsilon_{2,t}) ..., I_j(\varepsilon_{n,t})\right) = diag\left(I_j(z_{1,t}), I_j(z_{2,t}) ..., I_j(z_{n,t})\right) = I_j(z_t).$$

Proof:

Assume that the indicator functions are defined as in (11.a). Then, for a given j in $1,2,...,l$, and a given i in $1,2,...,k$,

$$I_j(\varepsilon_{i,t}) = I\left(\overline{d}_{i,j-1}(\varepsilon_{i,t}) < \varepsilon_{i,t} < \overline{d}_{i,j}(\varepsilon_{i,t})\right) = I\left(\left[D_t\Gamma_t d_{j-1}\right]_i < \varepsilon_{i,t} \leq \left[D_t\Gamma_t\overline{d}_j\right]_i\right)$$

$$= I\left(\left[D_t\Gamma_t d_{j-1}\right]_i < \left[D_t\Gamma_t\eta_t\right]_i \leq \left[D_t\Gamma_t d_j\right]_i\right) = I\left(\left[\Gamma_t d_{j-1}\right]_i < \left[\Gamma_t\eta_t\right]_i \leq \left[\Gamma_t d_j\right]_i\right)$$

$$= I\left(\left[\Gamma_t d_{j-1}\right]_i < z_{i,t} \leq \left[\Gamma_t d_j\right]_i\right) = I_j(z_{i,t}).$$

A similar proof can be derived when indicator functions follows (11.b). ∎

If the model has a unique stationary solution, we can write

$$E[H_t] = W + B_1 E[H_{t-1}] + E\left[\vec{G}_{t-1}\right].$$

As we are interested in the unconditional values, the expectations con be rewritten as follows:

$$E[H_t] = W + B_1 E[H_t] + E\left[\vec{G}_t\right],$$

by exploiting the structure of $\vec{G}_{t-1}$ and using Lemma A.1:

$$E\left[\vec{G}_t\right] = \sum_{j=1}^{l} E\left[\left(A_j + \Psi_j G_{t-1}\right) I_j\left(\varepsilon_t\right) dg\left(\left(\varepsilon_t - \tilde{d}_j\right)\left(\varepsilon_t - \tilde{d}_j\right)'\right)\right]$$

$$= \sum_{j=1}^{l}\left(A_j + \Psi_j E\left[G_{t-1}\right]\right) E\left[I_j\left(\varepsilon_t\right) dg\left(\left(\varepsilon_t - \tilde{d}_j\right)\left(\varepsilon_t - \tilde{d}_j\right)'\right)\right]$$

$$= \sum_{j=1}^{l}\left(A_j + \Psi_j E\left[G_t\right]\right) E\left[I_j\left(\varepsilon_t\right) dg\left(\left(\varepsilon_t - \tilde{d}_j\right)\left(\varepsilon_t - \tilde{d}_j\right)'\right)\right].$$

The unconditional value of the asymmetric term has to be numerically computed. The following equalities hold:

$$E\left[G_t\right] = \sum_{j=1}^{l}\left[A_j E\left[I_j\left(\varepsilon_t\right)\right] + \Psi_j E\left[G_{t-1} I_j\left(\varepsilon_t\right)\right]\right] = \sum_{j=1}^{l}\left[A_j E\left[I_j\left(z_t\right)\right] + \Psi_j E\left[G_{t-1}\right] E\left[I_j\left(z_t\right)\right]\right]$$

$$= \sum_{j=1}^{l}\left[A_j E\left[I_j\left(z_t\right)\right] + \Psi_j E\left[G_t\right] E\left[I_j\left(z_t\right)\right]\right] = \sum_{j=1}^{l}\left[A_j M_j + \Psi_j E\left[G_t\right] M_j\right],$$

where $M_j = E\left[I_j\left(z_t\right)\right]$ (which is a diagonal matrix). Note that this expectation can be numerically evaluated if the correlations are constant over time. Alternatively, we suggest using approximations and evaluate the quantity using the unconditional correlations implied by the correlation model. Given the values of M_j, we determine the unconditional value $E\left[G_t\right]$ by solving the following linear system

$$vec\left(E\left[G_t\right]\right) = vec\left(\sum_{j=1}^{l} A_j M_j\right) + vec\left(\sum_{j=1}^{l} \Psi_j E\left[G_t\right] M_j\right).$$

Next, consider then the following equality:

$$dg\left(\left(\varepsilon_{t-1} - \tilde{d}_j\right)\left(\varepsilon_{t-1} - \tilde{d}_j\right)'\right) = dg\left(D_{t-1}\left(z_{t-1} - \Gamma_{t-1} d_j\right)\left(z_{t-1} - \Gamma_{t-1} d_j\right)' D_{t-1}\right),$$

if $\tilde{d}_j = \bar{d}_j$ while if $\tilde{d}_j = 0$ we should consider

$$dg\left(\left(\varepsilon_{t-1} - \tilde{d}_j\right)\left(\varepsilon_{t-1} - \tilde{d}_j\right)'\right) = dg\left(D_{t-1} z_{t-1} z_{t-1}' D_{t-1}\right).$$

By focusing on the following expectation (when $\tilde{d}_j = \bar{d}_j$):

$$E\left[I_j(\varepsilon_t)dg\left((\varepsilon_t - \bar{d}_j)(\varepsilon_t - \bar{d}_j)'\right)\right] = E\left[I_j(z_t)dg\left(D_t(z_t - \Gamma_t d_j)(z_t - \Gamma_t d_j)' D_t\right)\right],$$

and using the fact that $I_j(z_t)$ is diagonal, we can write:

$$= E\left[dg\left(I_j(z_t)D_t(z_t - \Gamma_t d_j)(z_t - \Gamma_t d_j)' D_t\right)\right] = E\left[dg\left(D_t I_j(z_t)(\eta_t - \Gamma_t d_j)(\eta_t - \Gamma_t d_j)' D_t\right)\right].$$

This expression arises from the fact that the diagonal of the product within the internal parentheses is equivalent to the product of $I_j(\eta_t)$ with the dg(.) result given above. Furthermore, again using the fact that our interest is on the diagonal elements, we derive the following equalities:

$$E\left[dg\left(D_t I_j(z_t)(\eta_t - \Gamma_t d_j)(\eta_t - \Gamma_t d_j)' D_t\right)\right] =$$

$$= E\left[dg(D_t D_t) \odot dg\left(I_j(z_t)(\eta_t - \Gamma_t d_j)(\eta_t - \Gamma_t d_j)'\right)\right] =$$

$$= E\left[dg(D_t D_t)\right] \odot E\left[dg\left(I_j(z_t)(\eta_t - \Gamma_t d_j)(\eta_t - \Gamma_t d_j)'\right)\right] =$$

$$= E[H_t] \odot dg\left(E\left[I_j(z_t)(\eta_t - \Gamma d_j)(\eta_t - \Gamma d_j)'\right]\right).$$

Where we use the diagonality of D_t, and we replace the expectations with their unconditional values, namely $dg\left(E[D_t]E[D_t]\right) = dg\left(E[D_t D_t]\right) = E[H_t]$ and Γ comes from the decomposition of the unconditional correlation matrix. Note also, that the expectation of $E\left[dg(D_t D_t)\right]$ is independent from the expectation of $E\left[dg\left(I_j(z_t)(\eta_t - \Gamma_t d_j)(\eta_t - \Gamma_t d_j)'\right)\right]$ by the law of iterated expectations: the conditional standard deviations are a function of the information set at time t-1 while the innovations are referred to time t. Thus, defining

$$N_j = E\left[I_j(z_t)(z_t - \Gamma_t d_j)(z_t - \Gamma_t d_j)' \right],$$

it follows that:

$$E\left[dg\left(D_t I_j(z_t)(z_t - \Gamma_t d_j)(z_t - \Gamma_t d_j)' D_t \right) \right] = E[H_t] \odot dg(N_j) = E[H_t] \odot dg(N_j)$$
$$= dg(N_j) \odot E[H_t] = diag\left(dg(N_j) \right) E[H_t].$$

Note that when $\tilde{d}_j = 0$ the previous result is still valid but we must redefine N_j as follows:

$$\tilde{N}_j = E\left[I_j(z_t) z_t z_t' \right].$$

It should be emphasized that the unconditional expectation of the correlation matrix equals the correlation matrix if the matrix is constant, otherwise it has to be computed on the basis of a specified dynamic structure. Collecting the various results, we can then write

$$E\left[\vec{G}_{t-1} \right] = \sum_{j=1}^{l} (A_j + \Psi_j E[G_{t-2}]) E\left[I_j(z_{t-1}) dg\left((\varepsilon_{t-1} - \bar{d}_j)(\varepsilon_{t-1} - \bar{d}_j)' \right) \right]$$
$$= \sum_{j=1}^{l} \left[(A_j + \Psi_j E[G_t]) diag\left(dg(N_j) \right) E[H_t] \right]$$
$$E[H_t] = W + B_1 E[H_t] + \sum_{j=1}^{l} (A_j + \Psi_j E[G_t]) \left[diag\left(dg(N_j) \right) E[H_t] \right].$$

Solving with respect to the unconditional variances gives

$$E[H_t] = \left(I_n - B_1 - \sum_{j=1}^{l} \left[(A_j + \Psi_j E[G_t]) diag\left(dg(N_j) \right) \right] \right)^{-1} W.$$

Note that the unconditional value of the correlation matrix should be derived under the appropriate model that is used to define the dynamic conditional correlations, unless the conditional correlations are assumed to be constant. The unconditional variance of DAMGARCH is equivalent to

$$\Sigma = DRD,$$

and correlation targeting is imposed when the following equalities hold:

$$R^* = R,$$

$$W = \left(I_n - B_1 - \sum_{j=1}^{l} \left[\left(A_j + \Psi_j E \left[G_t^* \right] \right) diag \left(dg \left(N_j^* \right) \right) \right] \right) H^*,$$

where R^* and H^* refer to the corresponding sample estimators, N_j^* is evaluate using the decomposition of R^* and $E \left[G_t^* \right]$ depends on M_j^* which is also evaluated using the decomposition of R^*. Note also that the dynamic correlation model has also to be re-cast in a way ensuring correlation targeting. Note that this result includes as special cases the CCC, VARMA-GARCH, VARMA-AGARCH and DCC models.

Appendix A.4: Proofs of Theorems

Proof of Theorem 1.

Following Ling and McAleer (2003), we first rewrite DAMGARCH in the following form:

$$X_t = A_t X_{t-1} + \xi_t,$$

where

$$X_t = \left[G_t^{2,a\prime}, G^{2,b\prime}, G_t^{2,c\prime}, G_{t-1}^{2,a\prime}, G^{2,b\prime}, G_{t-1}^{2,c\prime}, ..., G_{t-r+1}^{2,a}{}', G^{2,b\prime}, G_{t-r+1}^{2,c}{}', H_t{}', H_{t-1}{}', ..., H_{t-s+1}{}' \right]',$$

which has dimension *(3nlr+ns)x1* and where the elements here included are defined in Appendix A.1. Note that the vector X_t contains threshold-dependent elements at time *t*, threshold dependent components, as well as the innovation (mean residuals) at time *t*. Using again the notation introduced in Appendix A.1, we consider the following representation of the variance dynamic:

$$H_t = W + \sum_{i=1}^{s} B_i H_{t-i} + \sum_{m=1}^{r} \left(G_{t-m}^1 G_{t-m}^{2,a} + G_{t-m}^1 G^{2,b} + G_{t-m}^1 G_{t-m}^{2,c} \right).$$

Multiplying by $\tilde{z}_t = diag\left(dg\left(z_t z_t' \right) \right)$ the equation for H_t yields:

$$e_t = \tilde{z}_t W + \sum_{i=1}^{s} \tilde{z}_t B_i H_{t-i} + \sum_{m=1}^{r} \left(\tilde{z}_t G_{t-m}^1 G_{t-m}^{2,a} + \tilde{z}_t G_{t-m}^1 G^{2,b} + \tilde{z}_t G_{t-m}^1 G_{t-m}^{2,c} \right),$$

given that $\tilde{z}_t H_t = e_t$. The previous equation implies a matrix A_t, with the following structure:

$$A_t = \begin{bmatrix} \tilde{z}_t G_{t-1}^1 : \tilde{z}_t G_{t-1}^1 : \tilde{z}_t G_{t-1}^1 & \cdots & \tilde{z}_t G_{t-r}^1 : \tilde{z}_t G_{t-r}^1 : \tilde{z}_t G_{t-r}^1 & \tilde{z}_t B_1 & \cdots & \tilde{z}_t B_s \\ & 0_{(3nl-n) \times 3nlr} & & & 0_{(3nl-n) \times ns} & \\ & I_{3nl(r-1)} & 0_{3nl(r-1) \times 3nl} & & 0_{3nl(r-1) \times ns} & \\ G_{t-1}^1 : G_{t-1}^1 : G_{t-1}^1 & \cdots & G_{t-r}^1 : G_{t-r}^1 : G_{t-r}^1 & B_1 & \cdots & B_s \\ & 0_{n(s-1) \times 3nlr} & & & I_{n(s-1)} & 0_{n(s-1) \times n} \end{bmatrix}.$$

Furthermore,

$$\xi_t = \left[\left(\tilde{z}_t W \right)', \left(i_{l-1} \otimes e_t \right)', G^{2,b\prime}, G_t^{2,c\prime}, 0_{3nl(r-1) \times 1}, W', 0_{n(s-1) \times 1} \right]$$

Given these quantities, and following Ling and McAleer (2003), we define the quantity

$$S_{m,t} = \xi_t + \sum_{j=1}^{m} \left(\prod_{i=1}^{j} A_{t-i+1} \right) \xi_{t-j}$$

where $m = 1,2,\dots$ Denote by $s_{m,t}$ the element of order k in the summation included in $S_{m,t}$, such that

$$E|s_{m,t}| = E \left| e_k' \xi_t + e_k' \left(\prod_{i=1}^{j} A_{t-i+1} \right) \xi_{t-j} \right|,$$

where e_k is a vector conformable with $S_{m,t}$ comprising zeros and with *1* in position *k*. Given that the matrices A_t are not independent, we should modify the proof of Ling and McAleer (2003). We first note that the following decomposition holds for each matrix A_t:

$$A_t = Z_t \tilde{A}_{t-1}$$

$$\tilde{A}_{t-1} = \begin{bmatrix} G_{t-1}^1 : G_{t-1}^1 : G_{t-1}^1 & \cdots & G_{t-r}^1 : G_{t-r}^1 : G_{t-r}^1 & B_1 & \cdots & B_s \\ & 0_{(3nl-n)\times 3nlr} & & & 0_{(3nl-n)\times ns} & \\ & I_{3nl(r-1)} & 0_{3nl(r-1)\times 3nl} & & 0_{3nl(r-1)\times ns} & \\ G_{t-1}^1 : G_{t-1}^1 : G_{t-1}^1 & \cdots & G_{t-r}^1 : G_{t-r}^1 : G_{t-r}^1 & B_1 & \cdots & B_s \\ & 0_{n(s-1)\times 3nlr} & & & I_{n(s-1)} & 0_{n(s-1)\times n} \end{bmatrix}$$

$$Z_t = \begin{bmatrix} \tilde{z}_t & 0_{n\times(3nlr+ns-n)} \\ 0_{(3nlr+ns-n)\times n} & I_{3nlr+ns-n} \end{bmatrix}$$

where $\tilde{A}_{t-1}$ depends on the information set at time *t-1* and Z_t depends on the information set at time *t*. Furthermore, consider a simple DAMGARCH model where the lags of the conditional variance dynamics are all restricted to be equal to one, which implies that $\tilde{A}_{t-1}$ depends only on the information at time *t-1*.and it is independent from Z_t.When increasing any lag length or order of the model, the following proof must be adapted. Thus, we have

$$E|s_{m,t}| = E\left|e_k'\left(\prod_{i=1}^{j} Z_{t-i+1}\tilde{A}_{t-i}\right)\xi_{t-j}\right| = e_k' E[Z_t]\left(\prod_{i=1}^{j-1} E\left[\tilde{A}_{t-i}Z_{t-i}\right]\right)E\left[\tilde{A}_{t-j}\xi_{t-j}\right] = e_k'\Delta_1 \bar{A}^j \Delta_2,$$

where we have used the previously introduced decomposition of A_t. If the model follows the general representation in equation (1)-(4) in the main text, the following equality is not satisfied, due to the inclusion in $\tilde{A}_{t-1}$ of the $G_{t-1}^1,...,G_{t-r}^1$ terms.

$$E\left[\prod_{i=1}^{j}\left(Z_{t-i+1}\tilde{A}_{t-i}\right)\xi_{t-j}\right] = E[Z_t]\left(\prod_{i=1}^{j-1} E\left[\tilde{A}_{t-i}Z_{t-i}\right]\right)E\left[\tilde{A}_{t-j}\xi_{t-j}\right]. \tag{A.3.1}$$

In fact, the terms $G_{t-1}^1,...,G_{t-r}^1$ depend on the past values of the innovations and are thus dependent on the past values of Z_t. However, under the restriction that the model has no dynamic asymmetry,

the terms $G_{t-1}^1, ..., G_{t-r}^1$ simplifies removing the dependence on past values of Z_t. In this last case, the expectations can be split given the dependence of $\tilde{A}_{t-1}$ on time *t-1* quantities only (when this is not the case, the expectation within the parentheses will involve additional terms), Δ_1 and Δ_2 are two matrices, and

$$\bar{A} = \begin{bmatrix} E\left[G_{t-1}^1 \tilde{z}_{t-1}\right] & ... & E\left[G_{t-r}^1\right] & B_1 & ... & B_s \\ & 0_{3nl \times (n+3nl)r} & & & 0_{3nl \times ns} & \\ & I_{(n+3nl)(r-1)} & 0_{(n+3nl)(r-1)\times(n+3nl)} & & 0_{(n+3nl)(r-1)\times ns} & \\ E\left[G_{t-1}^1 \tilde{z}_{t-1}\right] & ... & E\left[G_{t-r}^1\right] & B_1 & ... & B_s \\ & 0_{n(s-1)\times(n+3nl)r} & & & I_{n(s-1)} & 0_{n(s-1)\times n} \end{bmatrix}. \tag{A.3.2}$$

A similar result applies when the model parameter matrices are diagonal: by expanding the term G_{t-1}^1 in (A.3.1), and by using the diagonality of the parameter matrices and of $I_j\left(\varepsilon_{t-1}\right)$, we can show that the expectation in (A.3.2) is still valid. Unfortunately, such an approach cannot be used when the model has full parameter matrices.

When (A.3.1) is valid, it can be shown that Assumption 3 ensures the roots of the characteristic polynomial of $\bar{A}$ lie inside the unit circle, thereby proving the convergence of $\bar{A}^j$, and hence of the whole term. The remainder of the proof follows closely that in Ling and McAleer (2003) and in McAleer et al. (2009), also with respect to the proof of strict stationarity and ergodicity. ∎

Proof of Theorem 2.

Using Theorem 1 and the results reported in Appendix A.4, the proof follows by direct extension of the results in McAleer et al. (2009). ∎

Proof of Theorem 3.

Consistency is obtained by verifying the conditions given in Jeantheau (1998), namely

    i)   the parameter space Θ is compact;

    ii)   for any $\theta \in \Theta$, the model admits a unique strictly stationary and ergodic solution;

    iii)   there exists a deterministic constant *k*, such that $\forall t$ and $\forall \theta \in \Theta$, $\left|\Sigma_t\right| > k$;

iv)   model identifiability;

v) Σ_t is a continuous function of the parameter vector, θ;

vi) $E_\theta \left[\log |\Sigma_t| \right] < 0 \;\; \forall \theta \in \Theta$.

Note that the determinant of the conditional covariance matrix can be decomposed using equation (1) into $|\Sigma_t| = |D_t||R||D_t| = |D_t^2||R|$, where we have also used the assumption of a constant conditional correlation matrix. Furthermore, by Assumption 3, D_t is strictly positive, and there exists a constant k_1 such that $|D_t^2| > k_1 \;\; \forall t$. In addition, again using Assumption 3, there exists a second constant k_2 such that $|R| > k_2$. Then we can define a third constant $k = k_1 k_2$, such that $|\Sigma_t| > k \;\; \forall t$ and $\forall \theta \in \Theta$, where Θ is a compact subspace of an Euclidean space. This proves conditions i) and iii). Theorem 1 ensures the existence of a unique, strictly stationary and ergodic solution to DAMGARCH, verifying condition (ii). Assumption 4 deals with condition (iv), ensuring identifiability, while Assumption 5 imposes the log-moment condition, (vi). Finally, under Assumption 4, it is evident that the conditional variances are a continuous function of the parameter set, proving condition v). Condition (ii) refers to a unique strictly stationary and ergodic solution while Theorem 1 provides conditions for second-order stationary solution. However, using the results in Ling and McAleer (2003), Theorem 3.1, consistency can be proved under second-order stationary solutions.■

Proof of Theorem 4.

Using the previous results, the proof can be obtained by direct extension of the Theorems and Lemmas in Ling and McAleer (2003) and McAleer et al. (2009). ■

		ω	α	β
DAX	Coeff.	**0.074**	**0.143**	**0.833**
	100*St.dev.	0.429	0.338	0.467
FTSE	Coeff.	**0.009**	**0.087**	**0.908**
	100*St.dev.	0.010	0.039	0.042
Correlation		0.715		
Log-Likelihood -1708.484				

Table 2: CCC-GARCH estimates – bold values identify significant coefficients

		ω	α	γ	β
DAX	Coeff.	**0.023**	**0.022**	**0.097**	**0.919**
	100*St.dev.	0.040	0.070	0.364	0.080
FTSE	Coeff.	**0.009**	**0.003**	**0.104**	**0.935**
	100*St.dev.	0.008	0.079	0.322	0.140
Correlation		0.713			
Log-Likelihood -1662.922					

Table 3: CCC-GJR-GARCH estimates – bold values identify significant coefficients

		DAX	FTSE			DAX	FTSE
ω	Coeff.	**0.020**	**0.010**	Ψ_1 (DAX)	Coeff.	0.000	0.000
	100*St.dev.	0.023	0.009		100*St.dev.	0.074	0.160
B (DAX)	Coeff.	**0.930**	**0.930**	Ψ_2 (FTSE)	Coeff.	**0.064**	**0.175**
	100*St.dev.	0.093	0.049		100*St.dev.	0.404	0.784
A_1 (DAX)	Coeff.	0.000	0.000	Ψ_3 (FTSE)	Coeff.	0.000	0.000
	100*St.dev.	0.508	0.091		100*St.dev.	0.123	0.048
A_2 (DAX)	Coeff.	**0.106**	**0.103**	Ψ_4 (FTSE)	Coeff.	0.001	0.000
	100*St.dev.	0.112	0.096		100*St.dev.	0.268	0.326
A_3 (FTSE)	Coeff.	0.000	0.000	Correlation		0.714	
	100*St.dev.	0.204	0.088	Log-Likelihood -1660.82			
A_4 (DAX)	Coeff.	**0.150**	**0.084**				
	100*St.dev.	0.714	0.604				

Table 4: CCC-DAGARCH estimates – bold values identify significant coefficients

		DAX	FTSE			DAX	FTSE
ω	Coeff.	**0.015**	**0.009**	Ψ$_1$ (DAX)	Coeff.	0.000	0.000
	100*St.dev.	0.020	0.006		100*St.dev.	0.199	0.051
B (DAX)	Coeff.	**0.943**	0.000	Ψ$_1$ (FTSE)	Coeff.	**0.009**	0.003
	100*St.dev.	0.067	0.058		100*St.dev.	0.193	0.329
B (FTSE)	Coeff.	0.000	**0.928**	Ψ$_2$ (DAX)	Coeff.	**0.076**	**0.029**
	100*St.dev.	0.000	0.020		100*St.dev.	1.551	0.257
A$_1$ (DAX)	Coeff.	0.000	**0.024**	Ψ$_2$ (FTSE)	Coeff.	**0.273**	**0.185**
	100*St.dev.	0.075	0.401		100*St.dev.	0.503	0.716
A$_1$ (FTSE)	Coeff.	0.000	0.003	Ψ$_3$ (DAX)	Coeff.	0.000	**0.093**
	100*St.dev.	0.043	0.289		100*St.dev.	0.139	0.467
A$_2$ (DAX)	Coeff.	**0.063**	**0.011**	Ψ$_3$ (FTSE)	Coeff.	0.000	0.000
	100*St.dev.	0.004	0.058		100*St.dev.	0.067	0.116
A$_2$ (FTSE)	Coeff.	**0.008**	**0.059**	Ψ$_4$ (DAX)	Coeff.	0.000	**0.017**
	100*St.dev.	0.018	0.003		100*St.dev.	0.184	0.116
A$_3$ (DAX)	Coeff.	0.000	**0.003**	Ψ$_4$ (FTSE)	Coeff.	0.000	0.000
	100*St.dev.	0.052	0.097		100*St.dev.	0.180	0.204
A$_3$ (FTSE)	Coeff.	0.000	0.000	Corr	Coeff.	**0.718**	
	100*St.dev.	0.001	0.006		100*St.dev.	0.050	
A$_4$ (DAX)	Coeff.	**0.252**	**0.072**	Log-Likelihood 1622.425			
	100*St.dev.	0.530	0.546				
A$_4$ (FTSE)	Coeff.	**0.068**	**0.101**				
	100*St.dev.	0.003	0.390				

Table 5: DAMGARCH estimates – bold values identify significant coefficients

	DAX	FTSE
Mean	0.006	-0.013
Median	0.056	0.022
Maximum	5.466	3.589
Minimum	-3.517	-4.014
Std. Dev.	1.002	1.001
Skewness	-0.013	-0.165
Kurtosis	3.699	3.293
Correlation	0.004	
Jarque-Bera	35.331	14.083
Probability	0.000	0.001
10% quantile	-1.271	-1.248
90% quantile	1.278	1.233

Table 6: descriptive analysis of standardized and uncorrelated residuals (used for determining empirical structural innovations) and empirical quantiles used to define model thresholds – 10% (90%) quantile for the normal variable is -1.281 (1.281)

		2 (1)	3 (3)	4 (6)	5 (10)	10 (45)	20 (190)	100 (4950)	n $(n(n-1)/2)$
	DAMGARCH – $l=4$ – $s=r=q=1$								
	Assets number (number of correlations)	2 (1)	3 (3)	4 (6)	5 (10)	10 (45)	20 (190)	100 (4950)	n $(n(n-1)/2)$
DAMGARCH (including correlations)	Full	39	87	154	240	955	3810	95050	$n+(s+l+lq)n^2+\dfrac{n(n-1)}{2}$
	Diagonal	21	33	46	60	145	390	5950	$n+(s+l+lq)n+\dfrac{n(n-1)}{2}$
	Common Dynamic	27	60	106	165	655	2610	65050	$n+(s+l+1)n^2+\dfrac{n(n-1)}{2}$
	Diagonal and Common Dynamic	15	24	34	45	115	330	5650	$n+(s+l+1)n+\dfrac{n(n-1)}{2}$
	CCC (GARCH(s,r) and correlations)	9	15	22	30	85	270	5350	$n+(s+r)n+\dfrac{n(n-1)}{2}$
	DCC (GARCH(s,r) and correlations)	11	17	24	32	87	272	5352	$n+(s+r)n+\dfrac{n(n-1)}{2}+s_{DCC}+r_{DCC}$
	Diagonal BEKK(s,r)	7	12	18	25	75	250	5250	$\dfrac{n(n+1)}{2}+(s+r)n$
	Triangular BEKK(s,r)	9	18	30	45	165	630	15150	$(1+s+r)\dfrac{n(n+1)}{2}$
	BEKK(s,r)	11	24	42	65	255	1010	25050	$\dfrac{n(n+1)}{2}+(s+r)n^2$
	Diagonal Vech(s,r)	9	18	30	45	165	630	15150	$(1+s+r)\dfrac{n(n+1)}{2}$
	Vech(s,r)	21	78	210	465	6105	88410	$>5\times10^6$	$\dfrac{n(n+1)}{2}+(s+r)\left(\dfrac{n(n+1)}{2}\right)^2$

Table 1: Model dimension (s_{DCC} and r_{DCC} are the lag orders in the DCC model).
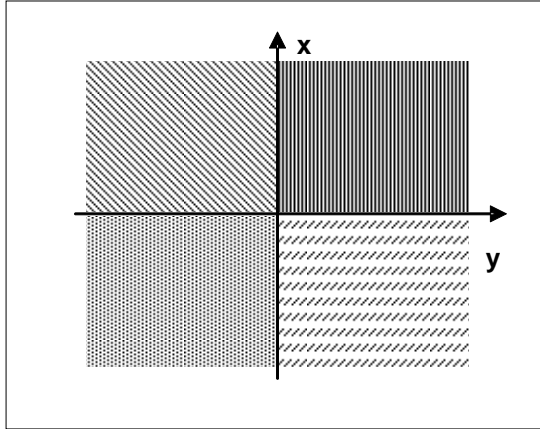
Figure 1: Multivariate GJR Representation
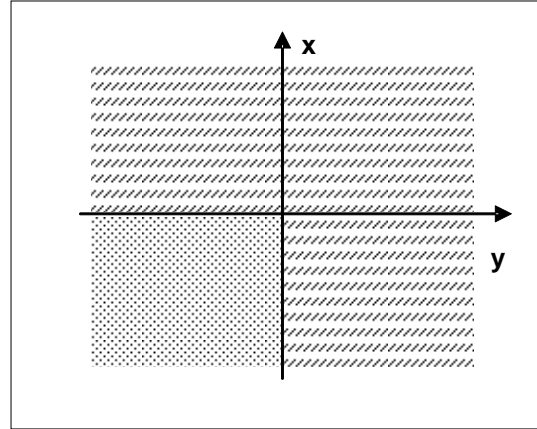
Figure 2: Partition Over the Joint Support
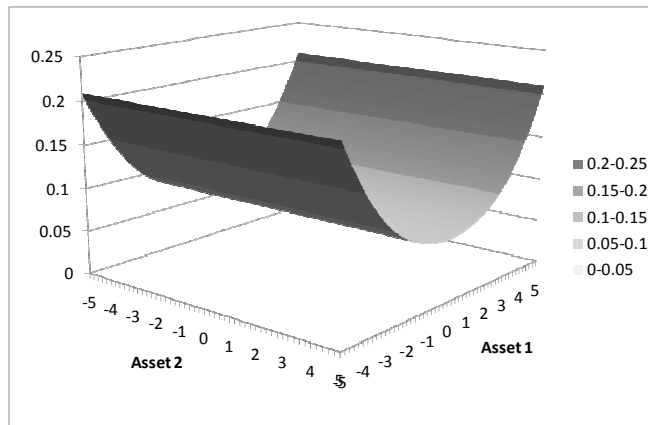
Figure 3: NIS of the first asset without spillovers
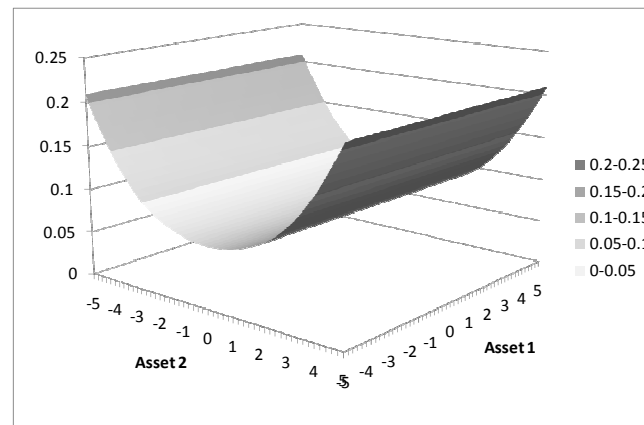
Figure 4: NIS of the second asset without spillovers

47
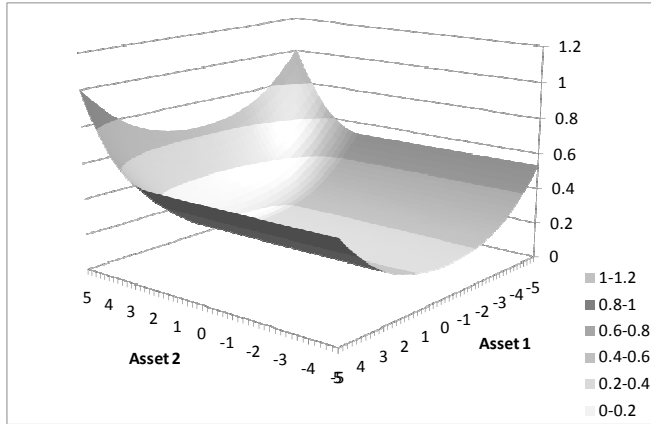
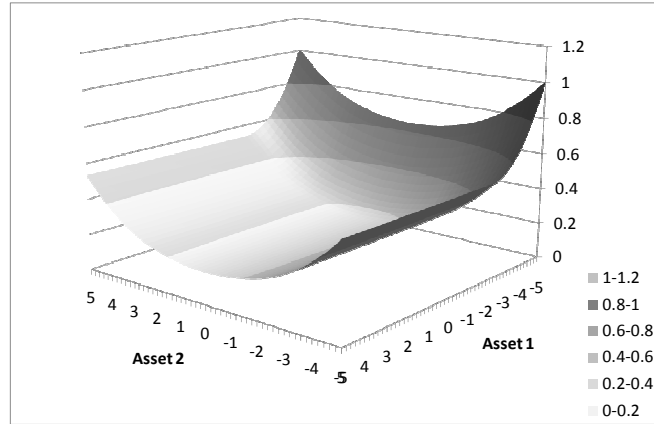Figure 5: NIS of the first asset with spillovers

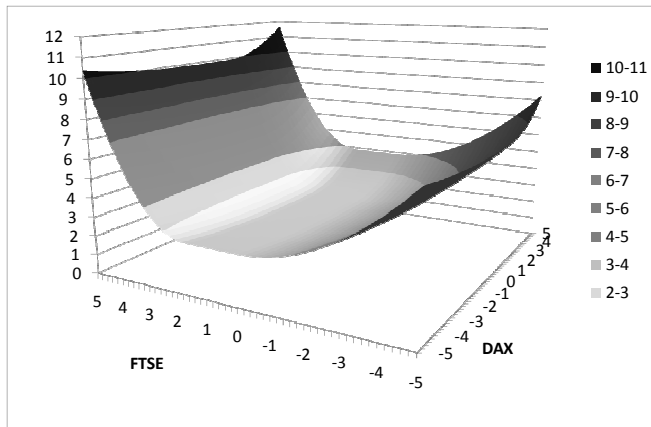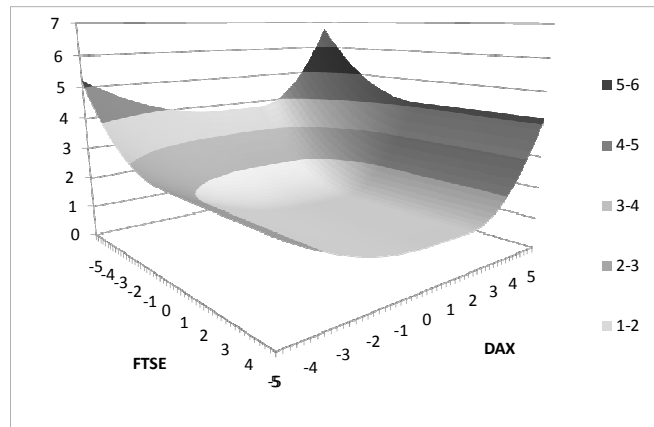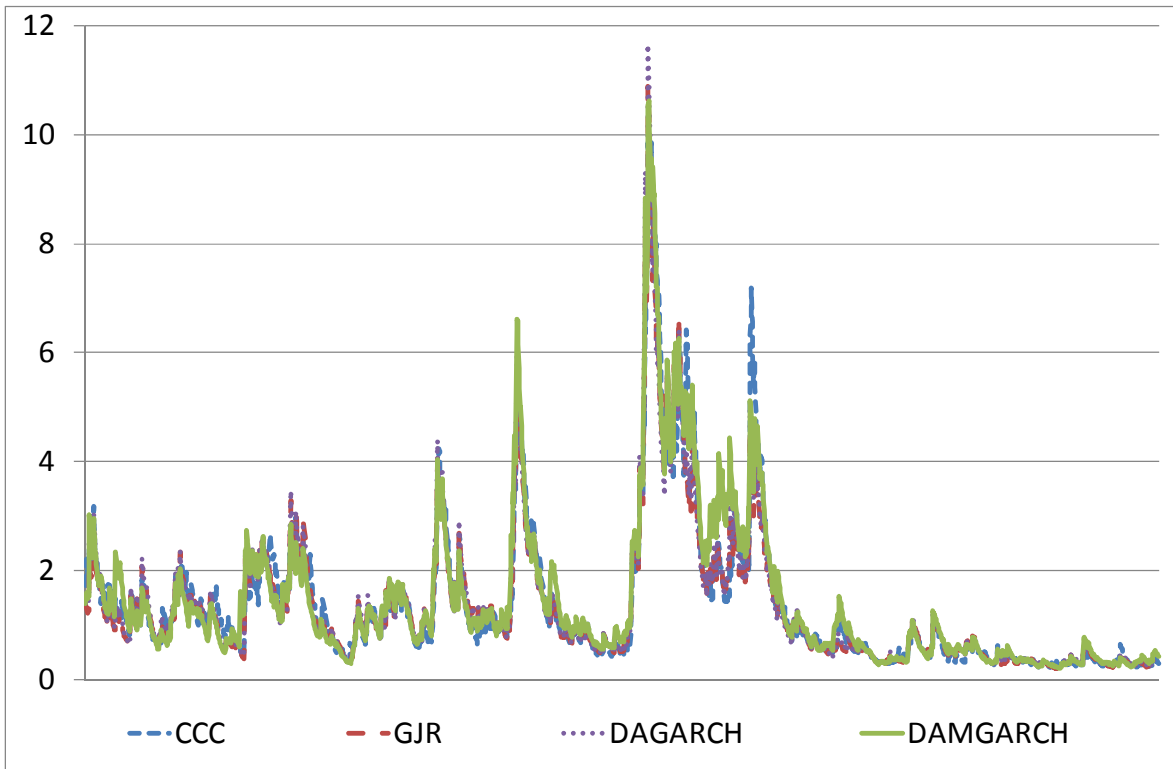Figure 6: NIS of the second asset with spillovers

Figure 9: DAX NIS

Figure 10: FTSE NIS

48

Figure 7: FTSE conditional variances given by CCC, GJR and DAMGARCH
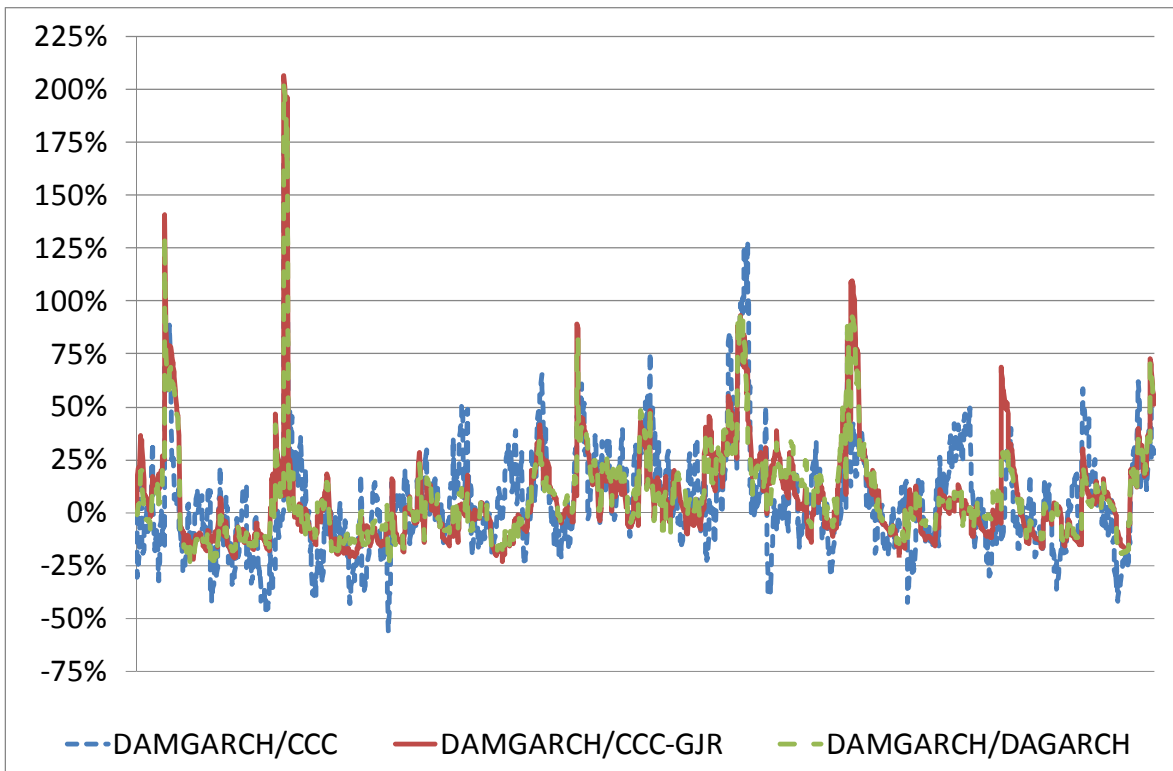
Figure 8: FTSE percentage difference between the conditional variances obtained from CCC and GJR with respect to the one provided by DAMGARCH
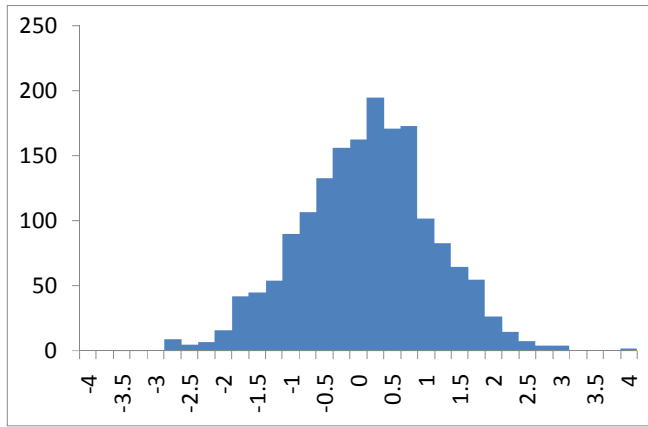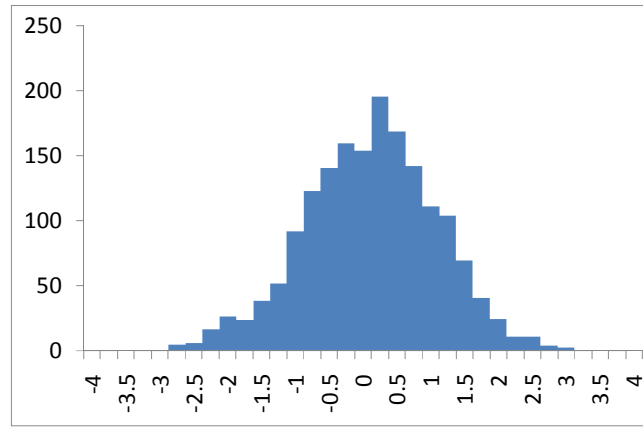
Figure 11: histogram of DAX structural innovations

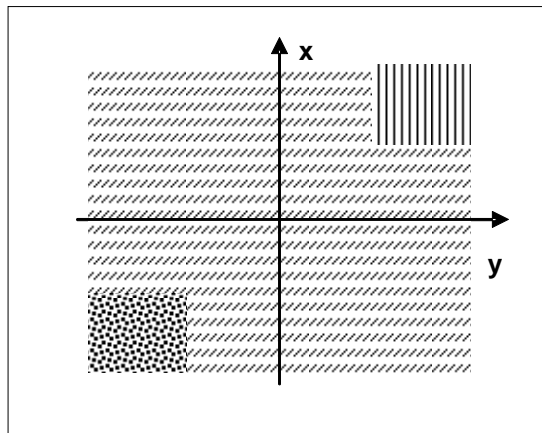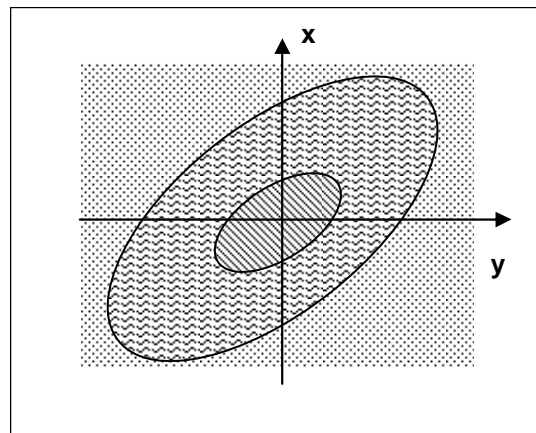Figure 12: histogram of FTSE structural innovations

Figure A.1: Extreme Events on a Bivariate Support

Figure A.2: A non-linear support partition