



UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Scienze Economiche “Marco Fanno”

THRESHOLDS, NEWS IMPACT SURFACES AND
DYNAMIC ASYMMETRIC MULTIVARIATE GARCH

MASSIMILIANO CAPORIN
Università di Padova

MICHAEL McALEER
University of Western Australia

January 2008

“MARCO FANNO” WORKING PAPER N.64

Thresholds, News Impact Surfaces and Dynamic Asymmetric Multivariate GARCH

Massimiliano Caporin*
Dipartimento di Scienze Economiche “Marco Fanno”
Università degli Studi di Padova

Michael McAleer
School of Economics and Commerce
University of Western Australia

December 2007

Abstract: DAMGARCH 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. 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 provides analytical expressions for the news impact surface implied by DAMGARCH and an empirical example.

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

JEL codes: C32, C51, C52

* Corresponding author: Massimiliano Caporin, Dipartimento di Scienze Economiche “Marco Fanno”, Università degli Studi di Padova, Via Del Santo, 33 - 35123 Padova – Italy – tel. +39-049-827-4258 – fax. +39-049-827-4211– email: massimiliano.caporin@unipd.it

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, Chou and Kroner (1992), Bollerslev, Engle and Nelson (1994), Li, Ling and McAleer (2003), 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 Pircard (1996) and Ling and McAleer (2002a, b) for univariate GARCH models, and Ling and McAleer (2003) for multivariate GARCH 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).

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 (1991) in introducing the EGARCH model, and was also considered by Glosten, Jagannathan and Runkle (1993), 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, b). However, restricting attention to only a single asset may be too strict, particularly if the primary goal is the measurement of the risk of an investment or a portfolio.

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).

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.

In addition to the possible mean effects, the paper focuses on the variance and covariance effects, monitoring an asset's conditional variance reaction to another asset's specific shock. Alternatively, 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 is an unexpected macroeconomic shock) on all the conditional variances might be considered. Furthermore, the effects of a shock will be distinguished based on its sign and 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 is the so-called 'leverage' effect, that is, negative shocks should increase conditional variances while positive shocks should induce a reduction in the conditional variances. Some multivariate GARCH models with asymmetry do not include leverage effects².

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. An early contribution in this direction was Hoti, Chan and McAleer (2002) who provided a multivariate generalization of the GJR model. 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 multivariate GARCH 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, the work of Hoti, Chan and McAleer (2002) seem to be the only one to deal 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 dynamics in the DCC model of Engle (2002)).

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 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 ideas in McAleer et al. (2006) 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 Multivariate GARCH (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 multivariate GARCH 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 multivariate GARCH 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 a basic CCC specification. The model proposed provides a higher likelihood and relevant insights into the asymmetric dynamics in the DAX and FTSE stock market indices.

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 a set of stock market indices in an asset allocation framework, and compares DAMGARCH with the CCC model. 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:

$$Y_t = E[Y_t | I^{t-1}] + \varepsilon_t,$$

$$E[\varepsilon_t | I^{t-1}] = 0, \tag{1}$$

$$E[\varepsilon_t \varepsilon_t' | I^{t-1}] = \Sigma_t = D_t R_t D_t$$

in which I^{t-1} is the information set available to time $t-1$, $E[Y_t | I^{t-1}]$ is the conditional mean of Y_t , and ε_t is the n -dimensional mean residual vector at time t . Furthermore, the innovations have a conditional mean of zero, and a conditionally time-dependent covariance matrix that can be decomposed into the contributions of the conditional variances and the conditional correlations³. Finally, D_t is a diagonal matrix of conditional volatilities, given by:

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

$$D_t = \text{diag}(\sigma_{1,t}, \sigma_{2,t}, \dots, \sigma_{n,t}),$$

in which $\text{diag}(a)$ is a matrix operator that creates a diagonal matrix from the elements of the vector a , 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 independent⁴, 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 matrix. Differently, Γ_t comes from an eigenvalue-eigenvector decomposition 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 $\Gamma_t = U_t \Delta_t^{1/2} U_t'$. 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 $H_t = (\sigma_{1,t}^2, \sigma_{2,t}^2, \dots, \sigma_{n,t}^2)'$ and $e_t = (\varepsilon_{1,t}^2, \varepsilon_{2,t}^2, \dots, \varepsilon_{n,t}^2)'$, respectively. The following equations define the Dynamic Asymmetric Multivariate GARCH (hereafter DAMGARCH) model:

$$H_t = W + \sum_{i=1}^s B_i H_{t-i} + \sum_{j=1}^l \tilde{G}_{t-j}, \quad (2)$$

$$\tilde{G}_t = \sum_{j=1}^l \left\{ \left[A_j + \Psi_j G_{t-1} \right] I_j(\varepsilon_t) \left[(\varepsilon_t - \tilde{d}_j) \square (\varepsilon_t - \tilde{d}_j) \right] \right\}^5 \quad (3)$$

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

where $B_i, i=1,2,\dots,s$, $A_j, j=1,2,\dots,l$, $\Psi_j, j=1,2,\dots,l$, and G_t are n -dimensional square matrices, and W and \tilde{G}_t are an n -dimensional vectors⁶. Furthermore, l is the number of subsets in which the

⁴ Independence and zero correlation are equivalent under a multivariate normal distribution, but not otherwise.

⁵ Note that the following equality holds: $(\varepsilon_t - \tilde{d}_j) \square (\varepsilon_t - \tilde{d}_j) = dg \left((\varepsilon_t - \tilde{d}_j) (\varepsilon_t - \tilde{d}_j)' \right)$.

support of the multivariate probability density function of ε_t has been partitioned (that is, there may be $l-1$ “threshold vectors”⁷). Finally, $I_j(\varepsilon_t)$ is a scalar (or a diagonal matrix) indicator function⁸ 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 (we will address below the structure of the subsets, the structure of \bar{d}_j and its usefulness), and \square denotes the Hadamard matrix product (that is, the element by element product). Note 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 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 $I_j(\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-m})$. 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 \bar{G}_t . In this case, we can write:

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.

⁷ 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.

⁸ 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. The structure of the indicator function will be further specified below.

$$G_t = \sum_{j=1}^l \left\{ \left[A_j + \Psi_j(L) G_t \right] I_j(\varepsilon_t) \right\}, \quad (5)$$

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

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. 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.

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). In fact, in a simple case, assuming $d_j=0$ $j=1,2$ (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 coefficient matrix \mathbf{A} explicitly as follows:

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

so that we can rewrite the ARCH term as

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

Equation (7) includes the traditional ARCH term, but also highlights that a sufficient condition for the identification of both the A_j and the \mathbf{A} (with $j = 1, 2, \dots, l$) matrices requires that at last one of the A_j $j = 1, 2, \dots, l$ matrices must be set to zero. In this case, the matrices A_j $j = 1, 2, \dots, l$ will define the differential effects of each partition on the GARCH structure, such that, in the representation we

propose, they define the global partition effect on the conditional variance evolution. 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 $r=l$ for simplicity, we may identify a standard ARCH part (threshold dependent and hence asymmetric) that is given by

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

and a second term associated to the asymmetry persistence:

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

This second term includes a component (G_{t-2}) 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 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 multivariate GARCH model, in which the time-varying ARCH coefficients, $A_j + \Psi_j G_{t-1}, j=1, 2, \dots, l$, 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:

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

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

$$\bigcup_{j=1}^l S_j = S, \quad S_i \cap S_j = \emptyset, \quad i, j = 1, 2, \dots, l, \quad i \neq j. \quad (9)$$

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

$$\begin{aligned} S_1 &= \{\varepsilon_t : \varepsilon_{i,t} < \bar{d}_L, i = 1, 2, \dots, n\}, \\ S_3 &= \{\varepsilon_t : \varepsilon_{i,t} > \bar{d}_U, i = 1, 2, \dots, n\}, \\ S_2 &= S - S_1 - S_3. \end{aligned} \quad (10)$$

In this example, assuming that \bar{d}_L is a small negative number and \bar{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 $A_3 = A_2 + A_b$. Note this is similar to having a standard ARCH coefficient and two additional components associated to extreme events of either positive or negative sign.

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:

$$I_j(\varepsilon_{i,t}) = \begin{cases} 1, & \varepsilon_{i,t} < \bar{d}_{i,j} \\ 0, & \text{otherwise} \end{cases}, \quad j = 1, 2, \dots, k \quad (11.a)$$

$$I_j(\varepsilon_{i,t}) = \begin{cases} 1, & \varepsilon_{i,t} > \bar{d}_{i,j-1} \\ 0, & \text{otherwise} \end{cases}, \quad j = k+1, \dots, l$$

if $\tilde{d}_j = \bar{d}_j$, and if $\tilde{d}_j = 0$:

$$I_j(\varepsilon_{i,t}) = \begin{cases} 1, & \bar{d}_{i,j-1} < \varepsilon_{i,t} \leq \bar{d}_{i,j} \\ 0, & \text{otherwise} \end{cases}, \quad j = 1, 2, \dots, l \quad (11.b)$$

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 $\varepsilon_{i,t} \leq \bar{d}_{i,1}$, while for $j = l$ (that is, the last subset), the condition becomes $\varepsilon_{i,t} > \bar{d}_{i,l-1}$; $\bar{d}_{i,1} < \dots < \bar{d}_{i,k-1} < 0 < \bar{d}_{i,k+1} < \dots < \bar{d}_{i,l-1}$, 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 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) is more general than that defined in (11), which represents a special case. Furthermore, we will show that any representation of the indicator function that follows (11) can be re-cast in the form of (8) by defining the partitions of ε_i density function support appropriately.

If we follow (11) in defining the indicator functions, then the elements of \bar{d}_j may be different over the variables and are defined accordingly to the structure of $I_j(\varepsilon_{i,t})$, namely $\bar{d}_j = \{\bar{d}_{j,1}, \dots, \bar{d}_{j,l}\}$ for

$j=1,\dots,k-1$, $\bar{d}_j = \{\bar{d}_{j-1,1}, \dots, \bar{d}_{j-1,l}\}$ for $j=k+2, \dots, l$, and $\bar{d}_j = 0_{n \times 1}$ for $j=k, k+1$. Under (8), the definition of \bar{d}_j depends on the relations used to define the S_j subsets. In the example in (10), we have $\bar{d}_1 = \bar{d}_L i_n$, $\bar{d}_2 = 0_n$ and $\bar{d}_3 = \bar{d}_U i_n$, 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 \bar{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 Hoti, Chan and McAleer (2002), Ling and McAleer (2003), and McAleer, Chan and Marinova (2002). In fact, assuming a constant correlation matrix, and imposing the condition that $\bar{G}_{t-j} = A_j e_{t-j}$ (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. (2006) could be obtained assuming a time-dependent structure for the conditional correlation matrix, again under the restriction $\bar{G}_{t-j} = A_j e_{t-j}$.

A related development was used in Hoti, Chan and McAleer (2002) for the introduction of asymmetric conditional variances in the multivariate GARCH framework. In this case, the appropriate matrix is given by:

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

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. In the full DAMGARCH representation of equations (1), (2), (3) and (4), the number of GARCH parameters is equal to $n + n^2(s + 2l)$ (excluding the parameters in the correlation matrix).

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 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, Caporin and Gobbo (2006). In that case, the parameter matrices could be partitioned and restricted on the basis of a particular asset classification.

Finally, we note that the DAMGARCH model has an implicit Self Exciting Threshold VAR representation (see Tsay (1998), for a definition of threshold VAR models) under the restriction $\vec{d}_j = 0$. The representation can be derived by generalizing the results in Caporin and McAleer (2006).

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:

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

where $F_z(\cdot)$ 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. The correct formulation should consider fixing a threshold 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, and a parametric restriction is imposed (this will be clarified below). For this reason, thresholds should 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 eigenvalue-eigenvector 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.

The observed thresholds defined over the mean innovations will be time independent for ε_t but not for η_t , as the following relations show:

$$P(\varepsilon_t < \bar{d}_j) = P(D_t \Gamma_t \eta_t < \bar{d}_j) = P(\eta_t < \Gamma_t^{-1} D_t^{-1} \bar{d}_j) = F_\eta(\Gamma_t^{-1} D_t^{-1} \bar{d}_j), \quad (13)$$

where \bar{d}_j is the vector defining the j -th thresholds, and $F_{\eta}(\cdot)$ is the multivariate cumulative density of the η_t innovations. The last equality shows that these probabilities are time varying since they are a function of the time-varying conditional variances and correlations. As a result, the observed thresholds defined over ε_t induce time-dependent thresholds over η_t . Given that our interest is in the identification of extreme events in the underlying innovations, η_t , the threshold time-dependence make their identification problematic. We therefore suggest reversing the threshold definition making ε_t thresholds time dependent. We can define a set of fixed thresholds over η_t , and refer to them as d_j . These ‘structural’ thresholds are such that:

$$P(\varepsilon_t < D_t \Gamma_t d_j) = P(D_t \Gamma_t \eta_t < D_t \Gamma_t d_j) = P(\eta_t < d_j) = F(d_j). \quad (14)$$

Therefore, the ‘observed’ thresholds equal $\bar{d}_j = \Gamma_t D_t d_j$, and they will be time dependent, by construction. Equation (14) ensures the following relation:

$$P(\eta_t < d_j) = \prod_{i=1}^n P(\eta_{i,t} < d_{i,j}) = \prod_{i=1}^n F_i(d_{i,j}), \quad j = 1, 2, \dots, l, \quad (15)$$

which arises from the independence among the standardized and uncorrelated innovations.

In the following, it is assumed that the thresholds are fixed over the probability density function of the η_t . In this case, we can either fix a priori the threshold values or a set of quantiles. 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.

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 $F(\cdot)$ as the joint cumulative density, and $F_i(\cdot)$, $i = 1, 2, \dots, n$, as the marginal cumulative densities of the η_t . It follows that:

$$I_j(\varepsilon_{i,t}) = \begin{cases} 1, & \Gamma_t D_t d_{j-1,i} < \varepsilon_{i,t} \leq \Gamma_t D_t d_{j,i}, \\ 0, & \text{otherwise} \end{cases}, \quad j = 1, 2, \dots, l \quad (16)$$

where d_j is the vector of structural thresholds defined over the η_t innovations. The structural thresholds d_j can be fixed a priori or determined by a quantile relation, $d_j = F_i^{-1}(\alpha)$ ⁹. Note that the condition in equation (16) is based on the elements of a time-dependent threshold vector, so that the indicator matrix function is given by $I_j(\varepsilon_t) = \text{diag}(I_j(\varepsilon_{1,t}), I_j(\varepsilon_{2,t}), \dots, I_j(\varepsilon_{n,t}))$. Finally, for $j=1$, the condition is $\varepsilon_{i,t} \leq \Gamma_i D_i d_{1,i}$, while for $j=l$, the condition is $\varepsilon_{i,t} > \Gamma_i D_i d_{l,i}$. Equation (16) refers to the indicator functions as defined in (11.b). The previous comments and the discussion on threshold definition are valid also for the indicator functions of (11.a).

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

$$I_j(\varepsilon_t) = \begin{cases} 1, & \bar{d}_{j-1} < \varepsilon_t \leq \bar{d}_j \\ 0, & \text{otherwise} \end{cases}, \quad j=1, 2, \dots, l \quad (17)$$

where the condition is satisfied if and only if the vector ε_t is included in the partition of the joint probability support.¹⁰ Specifically, equation (17) is equivalent to equation (7), as we can write the subset as:

$$S_j = \{\varepsilon_t : \bar{d}_{j-1,i} < \varepsilon_{i,t} \leq \bar{d}_{j,i}, \quad i=1, 2, \dots, n\}, \quad j=1, 2, \dots, l. \quad (18)$$

Note that equation (16) is always included in (17), while the reverse is not always true. Consider a bivariate example to illustrate the point. 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]

⁹ 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.

¹⁰ Note that, given (14), equation (17) is equivalent to $I_j(\eta_t) = \begin{cases} 1, & d_{j-1} < \eta_t \leq d_j \\ 0, & \text{otherwise} \end{cases}, \quad j=1, 2, \dots, l.$

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 (16) and (17) are equivalent does not mean that the models defined over the joint or the marginal thresholds are also equivalent. In fact, the representation (17) over the joint support is associated with a more flexible model. In the case of the marginal thresholds, it follows that:

$$G_t = [A_1 + \Psi_1 G_{t-1}] I_1(\varepsilon_t) + [A_2 + \Psi_2 G_{t-1}] I_2(\varepsilon_t), \quad (19)$$

whereas over the joint support, it follows that:

$$G_t = \sum_{j=1}^4 [\tilde{A}_j + \tilde{\Psi}_j G_{t-1}] \tilde{I}_j(\varepsilon_t). \quad (20)$$

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 \mathbf{x} be the first component and \mathbf{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 \mathbf{x} and negative \mathbf{y} , and $j = 3$ defines the subset with negative \mathbf{x} and positive \mathbf{y} . It follows that (19) and (20) are equivalent if:

$$\begin{aligned}
\tilde{A}_2 &= \begin{bmatrix} [\tilde{A}_4]_{.,1} & [\tilde{A}_1]_{.,2} \end{bmatrix}, & \tilde{\Psi}_2 &= \begin{bmatrix} [\tilde{\Psi}_4]_{.,1} & [\tilde{\Psi}_1]_{.,2} \end{bmatrix}, \\
\tilde{A}_3 &= \begin{bmatrix} [\tilde{A}_1]_{.,1} & [\tilde{A}_4]_{.,2} \end{bmatrix}, & \tilde{\Psi}_3 &= \begin{bmatrix} [\tilde{\Psi}_1]_{.,1} & [\tilde{\Psi}_4]_{.,2} \end{bmatrix},
\end{aligned} \tag{21}$$

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

Finally, we show that when there is a single common threshold set to zero, the ‘observed’ and ‘structural’ thresholds are equivalent and time independent. 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.

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 Hoti, Chan and McAleer (2002). 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. (2006). In the following, we provide the assumptions and the theorems stating the stationarity and the asymptotic properties of the DAMGARCH model. The assumptions and the theorems are a direct extension of the results in Ling and McAleer (2003).

Assumption 1: $E[Y_t | I^{t-1}] = 0$.

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

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

Under Assumption 2, the indicator functions defining the support partition depend on the structural thresholds. 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 positive definite symmetric matrix, with a positive lower bound over the parameter space Θ , all elements of B_i and $E[G_{t-1}^1 \tilde{z}_{t-i}]$ are non-negative; W has elements with positive lower and upper bound over Θ ; and all the roots of $\left| I - \sum_{i=1}^r E[G_{t-1}^1 \tilde{z}_{t-i}] L^i - \sum_{i=1}^s B_i L^i \right| = 0$ are outside the unit circle, where G_{t-1}^1 and \tilde{z}_{t-i} is defined below.

Assumption 4: $I - \sum_{i=1}^r E[G_{t-1}^1 \tilde{z}_{t-i}] L^i$ 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), (2), (3) and (4), the DAMGARCH model 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 $\tilde{\eta}_t$. The solution \tilde{H}_t has the following causal expansion:

$$\tilde{H}_t = W + \sum_{j=1}^{\infty} M' \left(\prod_{i=1}^j A_{t-i+1} \right) \xi_{t-j} \quad (22)$$

$$M' = \begin{bmatrix} \mathbf{0} & : & I_n & : & \mathbf{0} \\ n \times m & & n \times (r \times n) & & n \times ((s-1) \times m) \end{bmatrix} \quad (23)$$

$$A_t = \begin{bmatrix} \tilde{z}_t G_{t-1}^1 & \dots & \tilde{z}_t G_{t-r}^1 & \tilde{z}_t B_1 & \dots & \tilde{z}_t B_s \\ & 0_{3nl \times (n+3nl)r} & & & 0_{3nl \times ns} & \\ & I_{(n+3nl)(r-1)} & 0_{(n+3nl)(r-1) \times (n+3nl)} & & 0_{(n+3nl)(r-1) \times ns} & \\ G_{t-1}^1 & \dots & G_{t-r}^1 & B_1 & \dots & B_s \\ & 0_{n(s-1) \times (n+3nl)r} & & & I_{n(s-1)} & 0_{n(s-1) \times n} \end{bmatrix} \quad (24)$$

$$\xi_t = \begin{bmatrix} \tilde{z}_t W : (I_l \otimes \tilde{z}_t) \{ {}_1 e_{t,j} \}_{j=1}^l : (I_l \otimes \tilde{z}_t) \{ {}_2 e_{t,j} \}_{j=1}^l : (I_l \otimes \tilde{z}_t) \{ {}_3 e_{t,j} \}_{j=1}^l : \\ 0_{(n+3nl)(r-1) \times 1} : W : 0_{n(s-1) \times 1} \end{bmatrix} \quad (25)$$

$$\begin{aligned} e_t &= dg(\varepsilon_t \varepsilon_t') & {}_1 e_{t,j} &= dg(\bar{d}_j \bar{d}_j') \\ {}_2 e_{t,j} &= -dg(\varepsilon_t \bar{d}_j') & {}_3 e_{t,j} &= -dg(\bar{d}_j \varepsilon_t') \end{aligned} \quad (26)$$

$$G_t^1 = \left[\left([A_1 + \Psi_1 G_{t-1}] I_1(\varepsilon_t) \right) : \left([A_2 + \Psi_2 G_{t-1}] I_2(\varepsilon_t) \right) \dots \left([A_l + \Psi_l G_{t-1}] I_l(\varepsilon_t) \right) \right] \left[(i_l \otimes I_n) : I_l \otimes (i_3' \otimes I_n) \right]_{(n \times (n+3nl))} \quad (27)$$

$$\tilde{z}_t = dg(z_t z_t'), \quad E[\tilde{z}_t] = I_n, \quad z_t = D_t^{-1} \varepsilon_t, \quad \text{and} \quad D_t = \text{diag}(H_t^{1/2}) \quad (28)$$

and : denotes matrix horizontal concatenation.

Proof: See Appendix A.4. ■

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

Theorem 2: Under assumptions (1) to (4), if $\rho(E[A_t^{\otimes b}]) < 1$, then the 2bth moments of Y_t are finite; b is a strictly positive integer, and $\otimes b$ denotes the Kroneker product of b matrices A_t defined in Theorem 1.

Proof: Using Theorem 1 and the results reported in Appendix A.4, the proof follows by direct extension of the results in McAleer et al. (2006), page 32. ■

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 $E_{\theta_0} \left[\log \left(\left| \tilde{\Sigma}_t \right| \right) \right] < 0$

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) and reported below, we have $\hat{\theta} \xrightarrow{p} \theta$.

Proof: 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 $\forall \theta \in \Theta$, $|\Sigma_t| > k$;
- iv) model identifiability;
- v) Σ_t is a continuous function of the parameter vector, θ_0 ;
- vi) $E_{\theta} \left[\log |\Sigma_t| \right] < 0 \quad \forall \theta \in \Theta$.

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 $|D_t^2| > k_1 \quad \forall t$. 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 $|\Sigma_t| > k \quad \forall t$ 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). ■

Theorem 4: Given the consistency of the QMLE for DAMGARCH, under conditions 4.i), 4.ii) and 4.iii), we have $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{L} N(0, \Sigma^{-1} \Omega \Sigma^{-1})$:

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

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

$\Sigma_0 = E \left[n^{-1} \frac{\partial^2 L}{\partial \theta \theta'} \Big|_{\theta_0} \right]$ for any sequence $\hat{\theta}$ such that $\hat{\theta} \xrightarrow{p} \theta$;

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

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

Proof: Using the previous results, the proof can be obtained by direct extension of the results in Ling and McAleer (2003). ■

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 $L(\theta_0)$ that depends on the conditional

covariance matrix, $L(\theta_0) = \sum_{t=1}^T l_t(\theta_0) = \sum_{t=1}^T l_t(\Sigma_t(\theta_0))$. Traditionally, the approximate likelihood

function is derived from a multivariate normal distribution. In the multivariate GARCH 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 (1991) and Engle (2002). It should be noted that the two-step approach cannot be used

with DAMGARCH, given the dependence of the conditional variance dynamics from the observed thresholds that, in turn, are defined over the conditional variances and correlations. Furthermore, the distribution of the structural innovation, η_t , is not known. The two-step approach is available only in the limiting case of a single threshold that is set to zero for all the variables. Note that when all the parameter matrices are diagonal (that is, when there are no spillovers between variables), the two-step approach cannot be used given that the thresholds still depend on the correlation matrix. The two-step approach could be used only under the strong assumption of independence between the mean residuals, ε_t .

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 iterative approximated estimation procedure is suggested. Note that full estimation can be considered at the cost of increasing computational time and reducing the reliability of the estimates. 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, and compute the thresholds d_j for $j=1,2\dots l$;
- 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 $i=1,2\dots n$, and the thresholds ${}_{GARCH}\bar{d}_{j,t}(\varepsilon_{i,t}) = {}_{GARCH}\sigma_{i,t}^{-1}d_j$ $j=1,2\dots l$;
- 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 $i=1,2\dots n$;
- 4) compute the unconditional correlation matrix (using the sample estimator) on the ${}_{DAGARCH}\tilde{\eta}_t = [{}_{DAGARCH}\tilde{\eta}_{1,t} \cdot {}_{DAGARCH}\tilde{\eta}_{2,t} \cdot \dots \cdot {}_{DAGARCH}\tilde{\eta}_{n,t}]'$ series and save the correlation matrix R_n , the uncorrelated residuals ${}_R\eta_t = [{}_R\eta_{1,t} \cdot {}_R\eta_{2,t} \cdot \dots \cdot {}_R\eta_{n,t}]'$ and the thresholds ${}_{DAMGARCH}\bar{d}_j = \Gamma_t D_t d_j$ $j=1,2\dots l$ (as defined in equation (11));

5) test the distribution assumption of step 1) and, if necessary, update the $\mathbf{d}_j(\eta_t)$ thresholds.

If we assume that the model follows a diagonal specification in the conditional variance dynamics, we can iterate steps 3) to 5) until convergence, using in step 3) the thresholds defined in steps 4) and 5), that is, the thresholds ${}_{DAMGARCH}d_j(\varepsilon_t) = \Gamma^{-1}D_t^{-1}\bar{d}_j(\eta_t) \quad j=1,2\dots l$ with the updated $\mathbf{d}_j(\eta_t)$ component. Alternatively, the algorithm could include also the following steps.

- 6) estimate the conditional variance parameters by fixing the correlation matrix and, using the thresholds defined in steps 4) and 5), save the conditional variances ${}_{DAMGARCH}\sigma_{i,t}^2$ and the standardized residuals ${}_{DAMGARCH}\tilde{\eta}_{i,t} = \varepsilon_{i,t} / {}_{DAMGARCH}\sigma_{i,t}$ for $i=1,2\dots n$;
- 7) compute the unconditional correlation matrix (using the sample estimator) on the ${}_{DAMGARCH}\tilde{\eta}_t = [{}_{DAMGARCH}\tilde{\eta}_{1,t} : {}_{DAMGARCH}\tilde{\eta}_{2,t} : \dots : {}_{DAMGARCH}\tilde{\eta}_{n,t}]'$ series and save the correlation matrix R_n , the uncorrelated residuals ${}_R\eta_t = [{}_R\eta_{1,t} : {}_R\eta_{2,t} : \dots : {}_R\eta_{n,t}]'$ and the thresholds ${}_{DAMGARCH}d_j(\varepsilon_t) = \Gamma^{-1}D_t^{-1}\bar{d}_j(\eta_t) \quad j=1,2\dots l$ (as defined in equation (11));
- 8) test the distribution assumption of step 1) and, if necessary, update the $\bar{d}_j(\eta_{i,t})$ thresholds;
- 9) iterate steps 6) to 8) until convergence.

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 (1990)). Clearly, the proposed approach is suboptimal, but full systems estimation is likely to be viable only for small-dimensional systems.

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. (1993) and Nelson (1991), 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:

$$H_t = W + B_1 H_{t-1} + \vec{G}_{t-1}, \quad (29)$$

$$\vec{G}_t = \sum_{j=1}^4 \left\{ [A_j + \Psi_j G_{t-1}] I_j(\varepsilon_t) [(\varepsilon_t - \tilde{d}_j) \square (\varepsilon_t - \tilde{d}_j)] \right\}, \quad (30)$$

$$G_t = \sum_{j=1}^4 \left\{ [A_j + \Psi_j G_{t-1}] I_j(\varepsilon_t) \right\}. \quad (31)$$

The parameter matrices have been set to the following specifications:

$$A_1 = \begin{bmatrix} 0.15 & 0 \\ 0.20 & 0.10 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0.15 & 0 \\ 0.10 & 0.10 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0.15 & 0 \\ 0.10 & 0.10 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0.15 & 0 \\ 0.20 & 0.10 \end{bmatrix} \quad (32)$$

$$W = \begin{bmatrix} 0.01 \\ 0.01 \end{bmatrix} \quad B_1 = \begin{bmatrix} 0.85 & 0 \\ 0.07 & 0.78 \end{bmatrix} \quad R = \begin{bmatrix} 1 & 0.7 \\ 0.7 & 1 \end{bmatrix}$$

Note that the first asset's conditional variance does not depend on the second asset's conditional variance, but the two are correlated. Put differently, the second asset strongly depends on shocks arising from the first asset, both in the GARCH and the asymmetric terms.

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:

$$NIC = \omega + \beta \bar{\sigma}^2 + \alpha \bar{\sigma}^2 z_t^2, \quad (33)$$

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:

$$NIS = W + BE[H] + \sum_{i=1}^j \left[(A_i + \Psi_i E[G]) \text{diag} \left(dg \left(\Gamma \left[I_i(\eta_t)(\eta_t - d_i)(\eta_t - d_i)' \right] \Gamma' \right) \right) E[H] \right], \quad (34)$$

where the expectations are defined in the appendix. 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 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 with the correlation set to zero.

[Insert Figures 3 to 7 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 5. Put differently, even if the shocks are not correlated but if there are asymmetric spillover effects, the NIS depends on the shocks of both assets with ‘symmetric’ behavior. When we also introduce the correlation, the NIS changes in both cases. For the first asset, the correlation introduces dependence on the other asset’s innovations, while for the second asset the correlation increases the NIS level and introduces asymmetric behavior.

4. Dynamic Asymmetric Effect: an empirical example

This section focuses on the estimation of the DAMGARCH model and its comparison with a simpler CCC model. We consider the daily closing levels of the DAX and FTSE 100 indices. The

sample considered covers the period from 1998 to 2004 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 an NIS that is similar to that reported in the previous sections.

Table 3 reports the estimated coefficients of the DAMGARCH model, while Table 2 reports the CCC estimates. Furthermore, Figures 8 and 9 report the conditional variances estimated by the two models, while Figures 10, 11, 12 and 13 report the NISs for the DAX and the FTSE, together with their isometric representations, respectively.

For the bivariate system, a full DAMGARCH model with three thresholds, and all other orders set to 1, was estimated. The thresholds were fixed at zero at the upper and lower 5% tails under a standardized normal distribution for the uncorrelated and variance standardized residuals.

The thresholds define a partition on the bivariate probability support, which is comprised of 16 subsets, 4 of which are finite. In the following, the matrices are matched with a subscript corresponding to the following order:

- 1 - large negative values (below the lower threshold);
- 2 - negative values;
- 3 - positive values;
- 4 - large positive values (above the upper threshold).

The CCC model provides 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 a significant coefficient in most parameter matrices, with the exception of the A and Ψ matrices associated with negative values.

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 model.

Comparing the fitted conditional variances, we note some discrepancies, in particular, during periods of high volatility: the DAMGARCH peaks in the conditional variance seem to anticipate those produced by CCC, an effect that may be due to an improved forecasting ability. This empirical result needs further investigation, which is left for future research.

Finally, the log-likelihood provided by the DAMGARCH model is much higher than that of the CCC model. Given there exists a nesting relationship between the two models, a simple likelihood ratio test provides evidence in favor of the DAMGARCH model. The test statistic is 175.716 for 32 restrictions, such that the associated p-value is smaller than $1e-20$. This analysis is, however, not conclusive but needs further investigation, which is left for future research.

[Insert Figures 8 to 12 and Tables 2 and 3 here]

5. Concluding Remarks

This paper introduced a new multivariate GARCH model, DAMGARCH, which generalizes 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), we 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 multivariate GARCH models is left for future research.

Acknowledgements

We would like to thank Monica Billio, Gabriele Fiorentini, Mark Paoletta, Paolo Paruolo, Domenico Sartore and the participants of the Second Italian Congress of Econometrics and Empirical Economics 2007 (Rimini, Italy) and of the Workshop in Computational and Financial Econometrics 2007 (Geneva, Switzerland) for helpful comments and suggestions. The first author acknowledges financial support from the Italian Ministry of University and Research project PRIN2006 “Econometric analysis of interdependence, stabilisation and contagion in real and financial markets”. The second author wishes to acknowledge the financial support of the Australian Research Council.

References

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

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

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

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

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

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

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

De Goeij, P. and W. Marquering, (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 K.F. Kroner, (1995), Multivariate simultaneous generalized ARCH, *Econometric Theory*, 11, 122-150.

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

Glosten, L.R., R. Jagannathan, and D.E. Runkle, (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 H. Herwartz, (1998), Time-varying market price of risk in the CAPM. approaches, empirical evidence and implications, *Finance*, 19, 93–112.

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

Hoti, S., F. Chan and M. McAleer, (2002), Structure and asymptotic theory for multivariate asymmetric volatility: empirical evidence for country risk ratings, Invited paper presented to the Australasian Meeting of the Econometric Society, Brisbane, Australia, July 2002.

Kroner, F., and V. Ng, (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.

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

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

Ling, S. and M. McAleer, (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 M. McAleer, (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., F. Chan, S. Hoti and O. Lieberman, (2006), Generalized autoregressive conditional correlation, unpublished paper, School of Economics and Commerce, University of Western Australia.

McAleer, M., F. Chan and D. Marinova, (2002), An econometric analysis of asymmetric volatility: theory and application to patents, Invited paper presented to the Australasian Meeting of the Econometric Society, Brisbane, Australia, July 2002; Journal of Econometrics, forthcoming.

Nelson, D.B., (1991), Conditional heteroskedasticity in asset pricing: a new approach, *Econometrica*, 59, 347-370

Rabemananjara R. and J.M. Zakoian, (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 A.K.C. Tsui, (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 using the following:

$$H_t = W + \sum_{i=1}^s B_i H_{t-i} + \sum_{j=1}^r G^1_{t-j} G^2_{t-j} \quad (\text{A.1.1})$$

$$G^1_{(n \times n)} = \left[\left([A_1 + \Psi_1 G_{t-1}] I_1(\varepsilon_t) \right) : \left([A_2 + \Psi_2 G_{t-1}] I_2(\varepsilon_t) \right) \dots : \left([A_l + \Psi_l G_{t-1}] I_l(\varepsilon_t) \right) \right] \quad (\text{A.1.2})$$

$$G^2_{(n \times 1)} = \left[dg \left((\varepsilon_t - \tilde{d}_1)(\varepsilon_t - \tilde{d}_1)' \right)' : dg \left((\varepsilon_t - \tilde{d}_2)(\varepsilon_t - \tilde{d}_2)' \right)' \dots : dg \left((\varepsilon_t - \tilde{d}_l)(\varepsilon_t - \tilde{d}_l)' \right)' \right]' \quad (\text{A.1.3})$$

$$G_t = G^1_t (i_t \otimes I_n) \quad (\text{A.1.4})$$

where : denotes matrix horizontal concatenation and $dg(A)$ is a vector comprising the diagonal elements of A.

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

$$\begin{aligned} H_t &= W + B_1 H_{t-1} + \vec{G}_{t-1}, \\ \vec{G}_{t-1} &= \sum_{j=1}^l [A_j + \Psi_j G_{t-2}] I_j(\varepsilon_{t-1}) \text{dg} \left((\varepsilon_{t-1} - \tilde{d}_j) (\varepsilon_{t-1} - \tilde{d}_j)' \right), \\ G_{t-1} &= \sum_{j=1}^l [A_j + \Psi_j G_{t-2}] I_j(\varepsilon_{t-1}), \end{aligned}$$

where $\text{dg}(A)$ is the vector given by the diagonal of a matrix A .

Lemma A.1

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

$$I_j(\varepsilon_t) = \text{diag} \left(I_j(\varepsilon_{1,t}), I_j(\varepsilon_{2,t}), \dots, I_j(\varepsilon_{n,t}) \right) = \text{diag} \left(I_j(\eta_{1,t}), I_j(\eta_{2,t}), \dots, I_j(\eta_{n,t}) \right) = I_j(\eta_t).$$

Proof:

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

$$\begin{aligned} I_j(\varepsilon_{i,t}) &= I \left(\bar{d}_{i,j-1}(\varepsilon_{i,t}) < \varepsilon_{i,t} < \bar{d}_{i,j}(\varepsilon_{i,t}) \right) = I \left([D_t \Gamma_t \bar{d}_{j-1}]_i < \varepsilon_{i,t} \leq [D_t \Gamma_t \bar{d}_j]_i \right) \\ &= I \left([D_t \Gamma_t \bar{d}_{j-1}]_i < [D_t \Gamma_t \eta_t]_i \leq [D_t \Gamma_t \bar{d}_j]_i \right) \\ &= I \left(d_{i,j-1} < \eta_{i,t} \leq d_{i,j} \right) = I_j(\eta_{i,t}). \end{aligned}$$

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

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

$$E[H_t] = W + B_1 E[H_{t-1}] + E[\tilde{G}_{t-1}]$$

As we are interested in the unconditional values, we can write:

$$E[H] = W + B_1 E[H] + E[\tilde{G}_{t-1}]$$

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

$$\begin{aligned} E[\tilde{G}_{t-1}] &= \sum_{j=1}^l \left[A_j E \left[I_j(\varepsilon_{t-1}) dg \left((\varepsilon_{t-1} - \tilde{d}_j)(\varepsilon_{t-1} - \tilde{d}_j)' \right) \right] + \right. \\ &+ \left. \Psi_j E \left[G_{t-2} I_j(\varepsilon_{t-1}) dg \left((\varepsilon_{t-1} - \tilde{d}_j)(\varepsilon_{t-1} - \tilde{d}_j)' \right) \right] \right] \\ &= \sum_{j=1}^l \left[A_j E \left[I_j(\eta_{t-1}) dg \left((\varepsilon_{t-1} - \tilde{d}_j)(\varepsilon_{t-1} - \tilde{d}_j)' \right) \right] + \right. \\ &+ \left. \Psi_j E[G_{t-2}] E \left[I_j(\eta_{t-1}) dg \left((\varepsilon_{t-1} - \tilde{d}_j)(\varepsilon_{t-1} - \tilde{d}_j)' \right) \right] \right] \end{aligned}$$

The unconditional value of the asymmetric term is given by

$$\begin{aligned} E[G_t] &= \sum_{j=1}^l \left[A_j E[I_j(\varepsilon_{t-1})] + \Psi_j E[G_{t-2} I_j(\varepsilon_{t-1})] \right] \\ &= \sum_{j=1}^l \left[A_j E[I_j(\eta_t)] + \Psi_j E[G_t] E[I_j(\eta_t)] \right] \\ &= \left(I - \sum_{j=1}^l \Psi_j E[I_j(\eta_t)] \right)^{-1} \left(\sum_{j=1}^l A_j E[I_j(\eta_t)] \right) \\ &= \left(I - \sum_{j=1}^l \Psi_j M_j \right)^{-1} \left(\sum_{j=1}^l A_j M_j \right), \end{aligned}$$

which is a square matrix, with $M_j = E[I_j(\eta_t)]$. Consider then the following equality:

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

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

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

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

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

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

$$\begin{aligned} &= E\left[dg\left(I_j(\eta_t)D_t\Gamma_t(\eta_t - d_j)(\eta_t - d_j)'\Gamma_t'D_t\right)\right] \\ &= E\left[dg\left(D_t\Gamma_t I_j(\eta_t)(\eta_t - d_j)(\eta_t - d_j)'\Gamma_t'D_t\right)\right]. \end{aligned}$$

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 $dg(\cdot)$ result given above. Using the convention that

$$\begin{aligned} &E\left[dg\left(D_t\Gamma_t I_j(\eta_t)(\eta_t - d_j)(\eta_t - d_j)'\Gamma_t'D_t\right)\right] = \\ &= dg\left(E[D_t]E[\Gamma_t]E\left[I_j(\eta_t)(\eta_t - d_j)(\eta_t - d_j)'\right]E[\Gamma_t]'E[D_t]\right) \end{aligned}$$

which generalizes the result that the expectation of a matrix is the matrix of the expectations, using the law of iterated expectations under the assumption that the conditional variances and conditional correlations depend on the information set at time $t-1$.

Thus, defining

$$N_j = E \left[I_j (\eta_t) (\eta_t - d_j) (\eta_t - d_j)' \right]$$

it follows that:

$$\begin{aligned} & E \left[dg \left(D_t \Gamma I_j (\eta_t) (\eta_t - d_j) (\eta_t - d_j)' \Gamma' D_t \right) \right] = \\ & = dg \left(E[D_t] E[\Gamma] N_j E[\Gamma]' E[D_t] \right) = E[H_t] \square dg \left(E[\Gamma] N_j E[\Gamma]' \right) = \\ & = dg \left(E[\Gamma] N_j E[\Gamma]' \right) \square E[H] = \text{diag} \left(dg \left(\Gamma N_j \Gamma' \right) \right) E[H], \end{aligned}$$

using the diagonality of D_t , and replacing the expectations with their unconditional values, namely $E[\Gamma_t] = \Gamma$ and $dg(E[D_t]E[D_t]) = dg(E[D_t D_t]) = E[H_t] = E[H]$. Note that when $\tilde{d}_j = 0$ the previous result is still valid but we must redefine N_j as follows:

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

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

$$\begin{aligned} E[\tilde{G}_{t-1}] &= \sum_{j=1}^l \left[A_j E \left[I_j (\eta_{t-1}) dg \left((\varepsilon_{t-1} - \bar{d}_j) (\varepsilon_{t-1} - \bar{d}_j)' \right) \right] + \Psi_j E[G_{t-2}] E \left[I_j (\eta_{t-1}) dg \left((\varepsilon_{t-1} - \bar{d}_j) (\varepsilon_{t-1} - \bar{d}_j)' \right) \right] \right] \\ &= \sum_{j=1}^l \left[A_j \text{diag} \left(dg \left(\Gamma N_j \Gamma' \right) \right) E[H] + \Psi_j E[G_t] \text{diag} \left(dg \left(\Gamma N_j \Gamma' \right) \right) E[H] \right] \end{aligned}$$

$$E[H] = W + B_1 E[H] + \sum_{j=1}^l \left[A_j \left[\text{diag} \left(dg \left(\Gamma N_j \Gamma' \right) \right) E[H] \right] + \Psi_j E[G_t] \left[\text{diag} \left(dg \left(\Gamma N_j \Gamma' \right) \right) E[H] \right] \right]$$

Rearranging the results and substituting the expression of $E[G_t]$ gives

$$\begin{aligned}
E[H] &= \left(I_n - B_1 - \sum_{j=1}^l \left[(A_j + \Psi_j E[G_t]) \text{diag} \left(dg \left(\Gamma N_j \Gamma' \right) \right) \right] \right)^{-1} W \\
&= \left(I_n - B_1 - \sum_{j=1}^l \left[\left(A_j + \Psi_j \left(I - \sum_{j=1}^l \Psi_j M_j \right)^{-1} \left(\sum_{j=1}^l A_j M_j \right) \right) \text{diag} \left(dg \left(\Gamma N_j \Gamma' \right) \right) \right] \right)^{-1} W
\end{aligned}$$

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:

$$\begin{aligned}
R^* &= R, \\
W &= \left(I_n - B_1 - \sum_{j=1}^l \left[\left(A_j + \Psi_j \left(I - \sum_{j=1}^l \Psi_j M_j \right)^{-1} \left(\sum_{j=1}^l A_j M_j \right) \right) \text{diag} \left(dg \left(\Gamma^* N_j \Gamma^{*'} \right) \right) \right] \right)^{-1} H^*
\end{aligned}$$

where starred quantities refer to the corresponding sample estimators. Note that this result includes as special cases the CCC, VARMA-GARCH, VARMA-AGARCH and DCC models.

Appendix A.4 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

$$X_t = \begin{bmatrix} e_t, \{ {}_1e_{t,j} \}_{j=1}^l, \{ {}_2e_{t,j} \}_{j=1}^l, \{ {}_3e_{t,j} \}_{j=1}^l, e_{t-1}, \{ {}_1e_{t-1,j} \}_{j=1}^l, \{ {}_2e_{t-1,j} \}_{j=1}^l, \{ {}_3e_{t-1,j} \}_{j=1}^l, \dots \\ e_{t-r+1}, \{ {}_1e_{t-r+1,j} \}_{j=1}^l, \{ {}_2e_{t-r+1,j} \}_{j=1}^l, \{ {}_3e_{t-r+1,j} \}_{j=1}^l, H_t, H_{t-1}, \dots, H_{t-s+1} \end{bmatrix}'$$

which has dimension $((n+3nl)r+ns) \times l$, and

$$\begin{aligned} e_{t,j} &= dg \left((\varepsilon_t - \tilde{d}_j)(\varepsilon_t - \tilde{d}_j)' \right) \\ &= dg \left(\varepsilon_t \varepsilon_t' + \tilde{d}_j \tilde{d}_j' - \varepsilon_t \tilde{d}_j' - \tilde{d}_j \varepsilon_t' \right) \\ &= dg \left(\varepsilon_t \varepsilon_t' \right) + dg \left(\tilde{d}_j \tilde{d}_j' \right) - dg \left(\varepsilon_t \tilde{d}_j' \right) - dg \left(\tilde{d}_j \varepsilon_t' \right) \\ &= e_t + {}_1e_{t,j} + {}_2e_{t,j} + {}_3e_{t,j} \end{aligned}$$

Furthermore, $\{ {}_1e_{t,j} \}_{j=1}^l$ denotes the vector of dimension nl containing the values of ${}_1e_{t,j}$ for all the l partitions.

Note that we have included the threshold-dependent innovation (in deviation from the threshold) at time t and the innovation (mean residuals) at time t . For the simple (1,1) model, we have

$$\begin{aligned} H_t &= W + B_1 H_{t-1} + \sum_{j=1}^l [A_j + \Psi_j G_{t-2}] I_j (\varepsilon_{t-1}) (e_{t-1} + {}_1e_{t-1,j} + {}_2e_{t-1,j}), \\ G_t &= \sum_{j=1}^l [A_j + \Psi_j G_{t-1}] I_j (\varepsilon_t), \end{aligned}$$

which implies a matrix A_t , with the following structure:

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

where

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

Furthermore,

$$G_t^1 = \left[\left([A_1 + \Psi_1 G_{t-1}] I_1(\varepsilon_t) \right) : \left([A_2 + \Psi_2 G_{t-1}] I_2(\varepsilon_t) \right) \dots \left([A_l + \Psi_l G_{t-1}] I_l(\varepsilon_t) \right) \right] \left[(i_l \otimes I_n) : I_l \otimes (i_3' \otimes I_n) \right]$$

$(n \times (n+3nl))$

and

$$\xi_t = \left[\begin{array}{c} \tilde{z}_t W : (I_l \otimes \tilde{z}_t) \left\{ {}_1 e_{t,j} \right\}_{j=1}^l : (I_l \otimes \tilde{z}_t) \left\{ {}_2 e_{t,j} \right\}_{j=1}^l : (I_l \otimes \tilde{z}_t) \left\{ {}_3 e_{t,j} \right\}_{j=1}^l : \\ \mathbf{0}_{(n+3nl)(r-1) \times 1} : W : \mathbf{0}_{n(s-1) \times 1} \end{array} \right]$$

where $(I_l \otimes \tilde{z}_t) \left\{ {}_1 e_{t,j} \right\}_{j=1}^l$ indicates that the first matrix product multiplies the vector $\left\{ {}_1 e_{t,j} \right\}_{j=1}^l$ and “:” identifies matrix vertical concatenation.

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

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

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

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

where e_k is a vector conformable with $S_{n,t}$ comprising zeros and with 1 in position k . Note that the following decomposition holds for the matrix A_t :

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

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 on information at time $t-1$. When increasing any lag length, the following proof must be adapted. Thus, we have

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

where we have used the previously introduced decomposition of A_t , so 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

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

Then, 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).

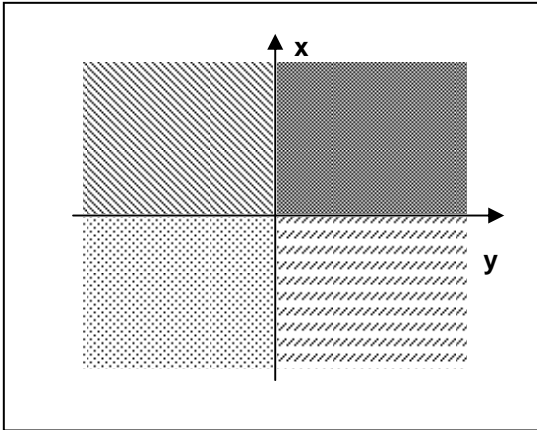


Figure 1: Multivariate GJR Representation

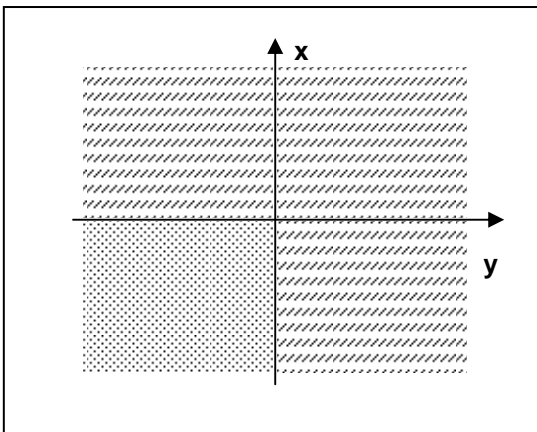


Figure 2: Partition Defined Over the Joint Support

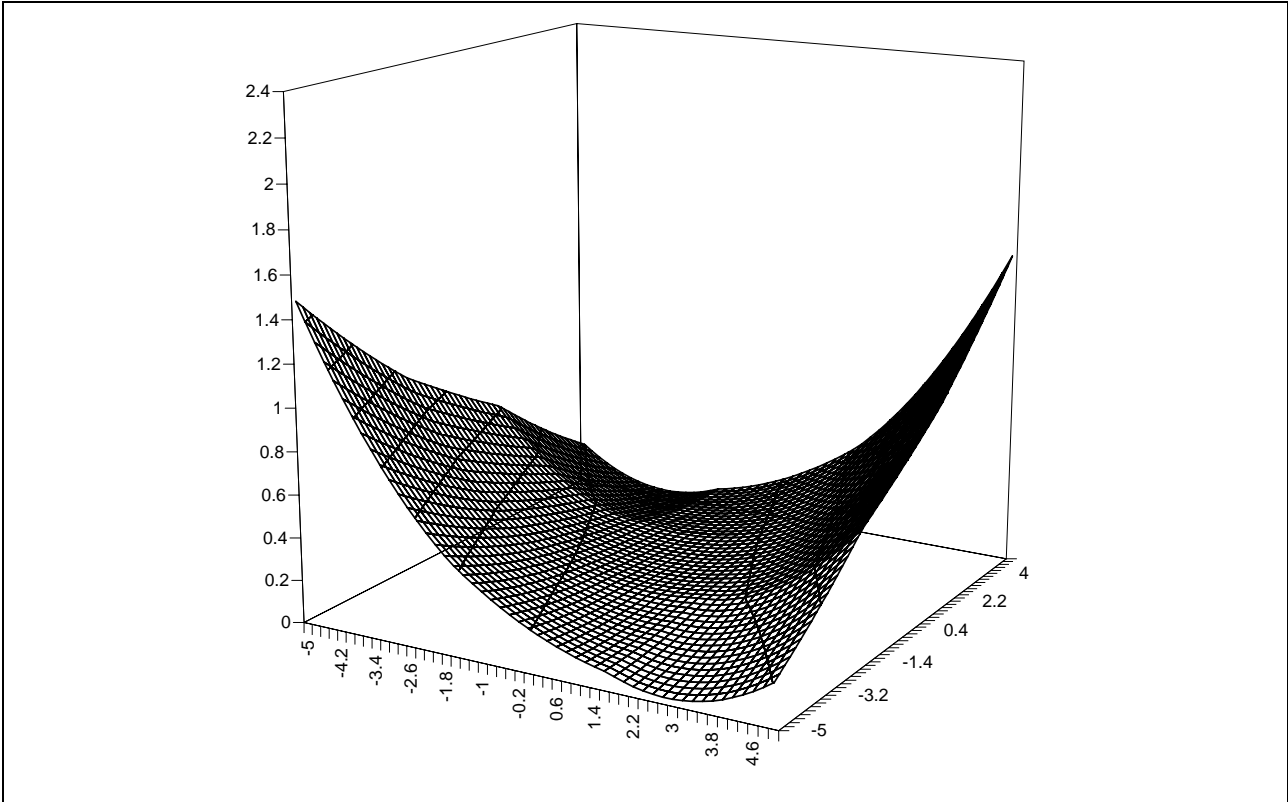


Figure 3: NIS of the first asset when the correlation is 0.7

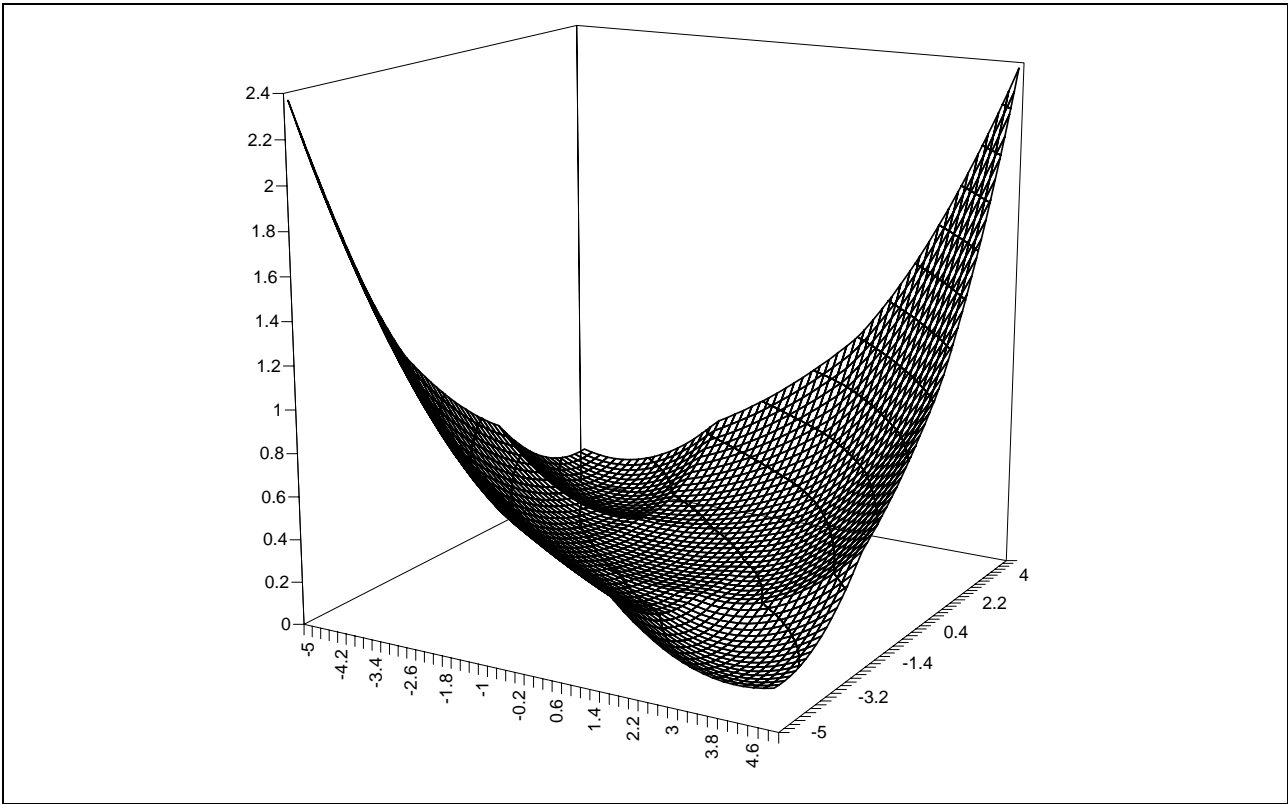


Figure 4: NIS of the second asset when the correlation is 0.7

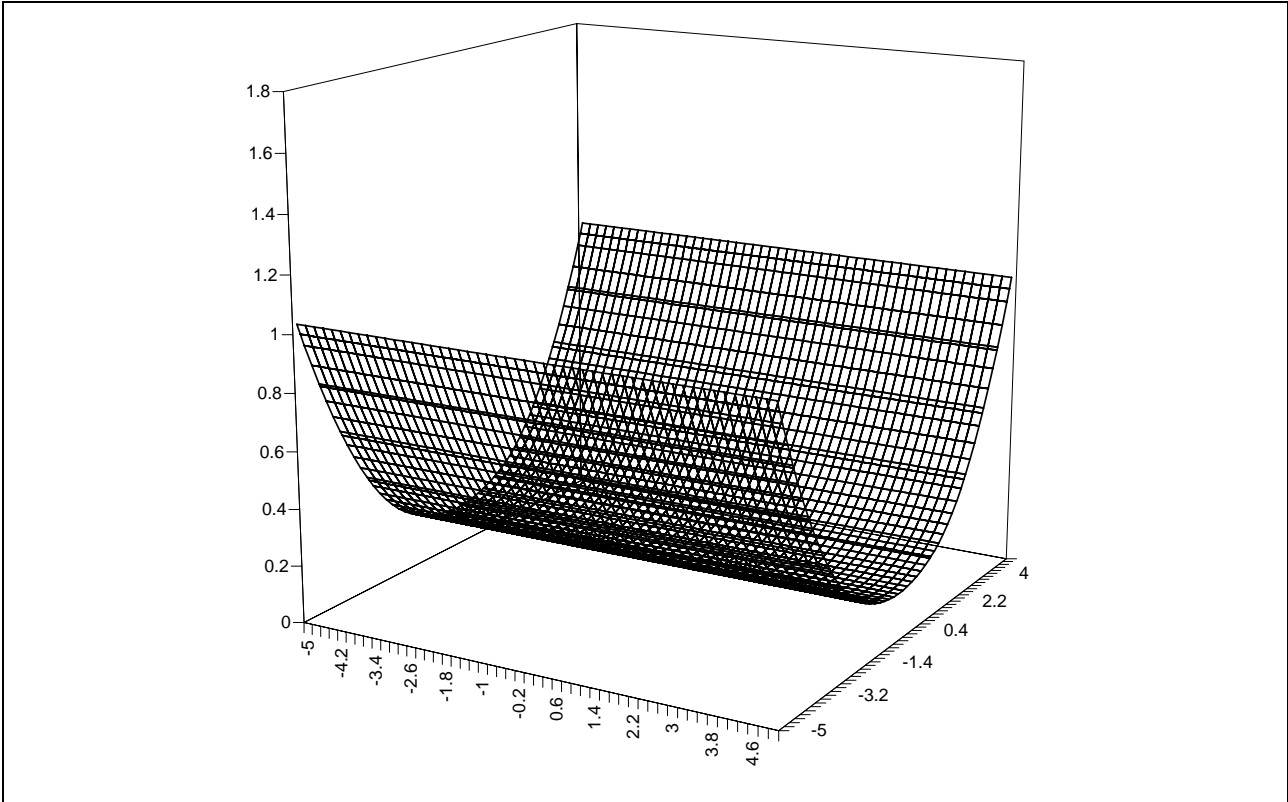


Figure 5: NIS of the first asset when the correlation is 0

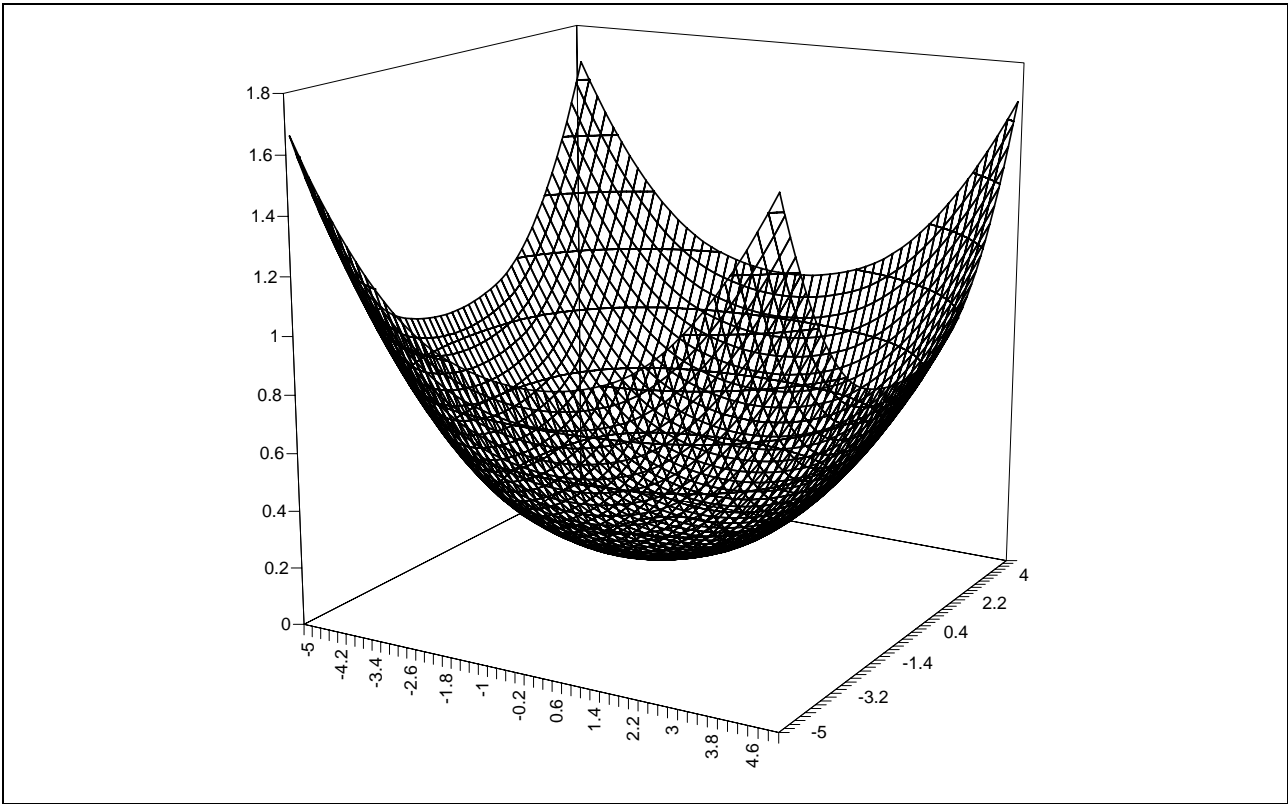


Figure 6: NIS of the second asset when the correlation is 0

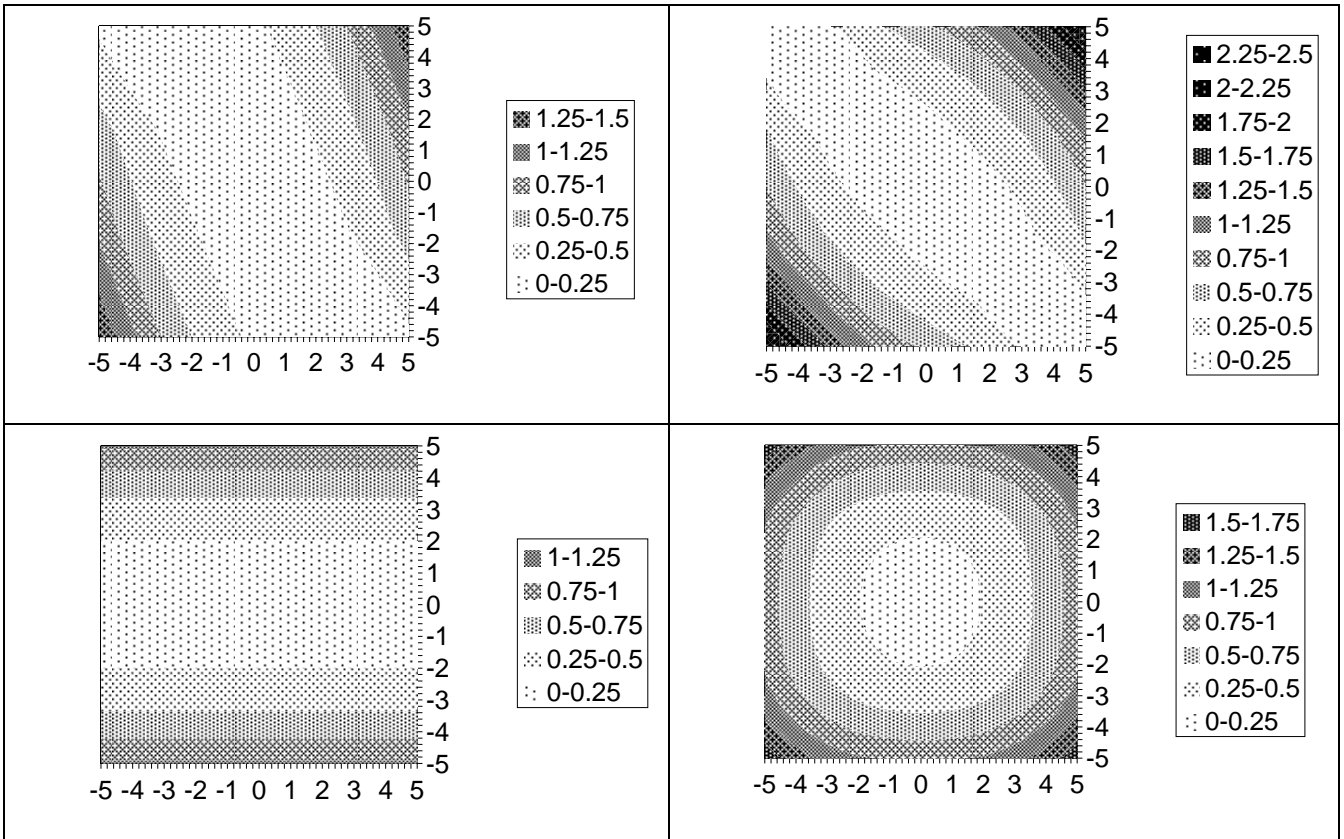


Figure 7: isometric representations of the NIS reported in Figures 3 to 6.

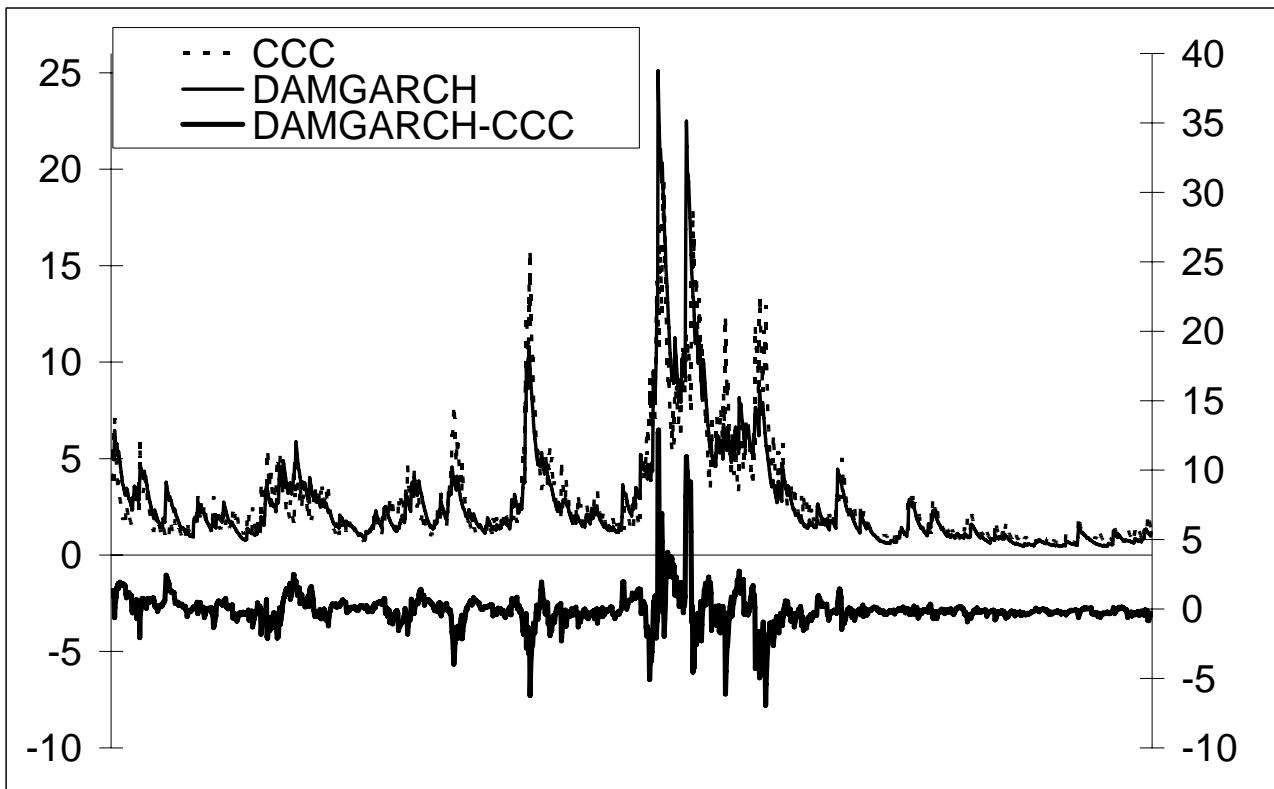


Figure 8: DAX conditional variances given by CCC and DAMGARCH, together with their difference

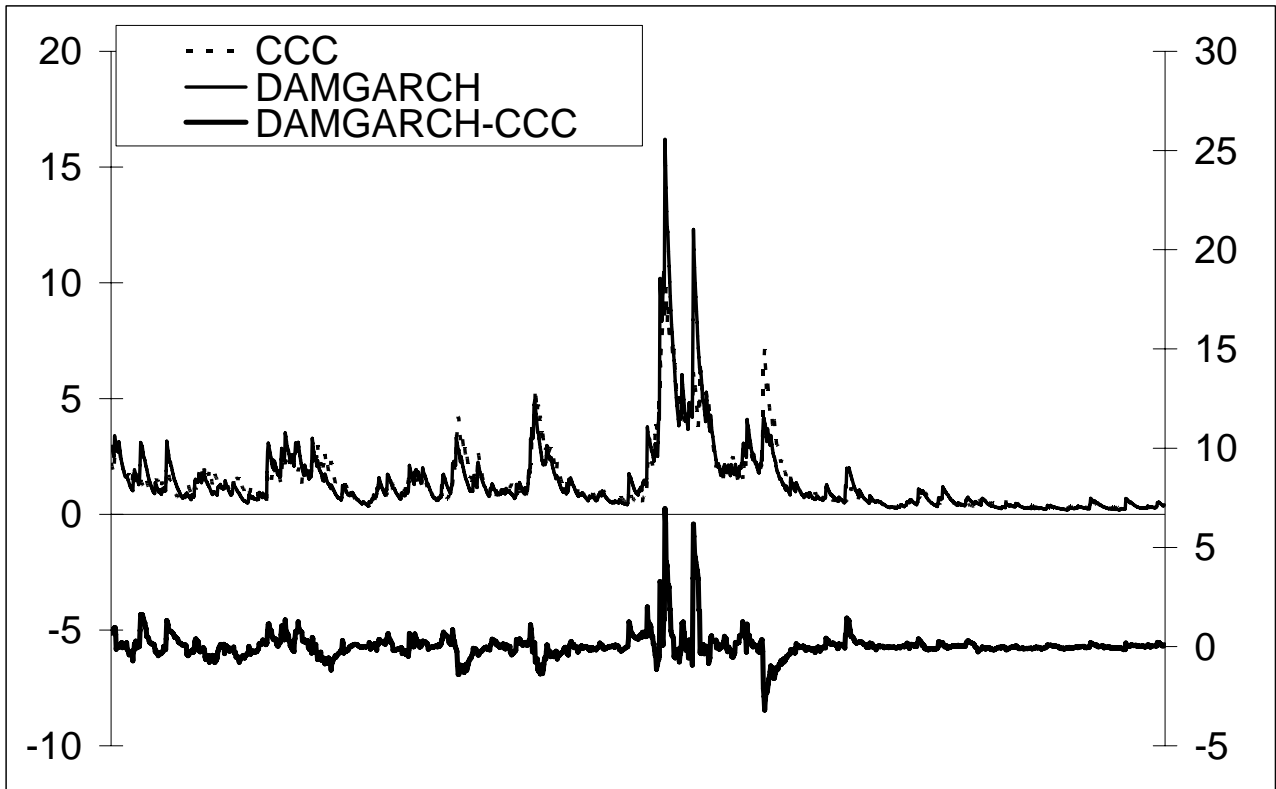


Figure 9: FTSE conditional variances given by CCC and DAMGARCH, together with their difference

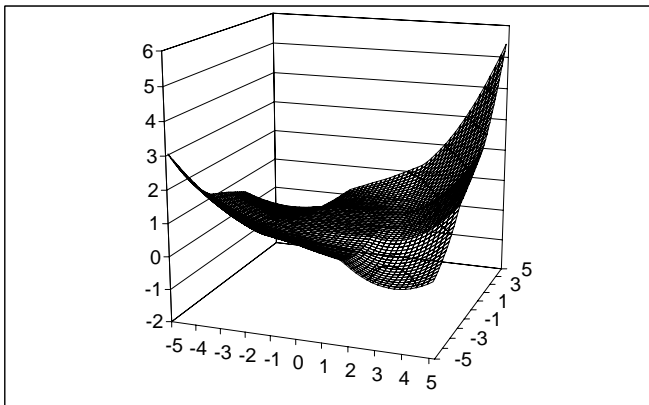


Figure 10: DAX NIS

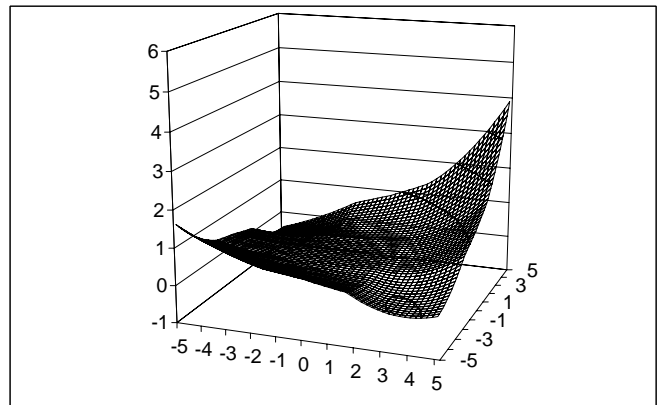


Figure 11: FTSE NIS

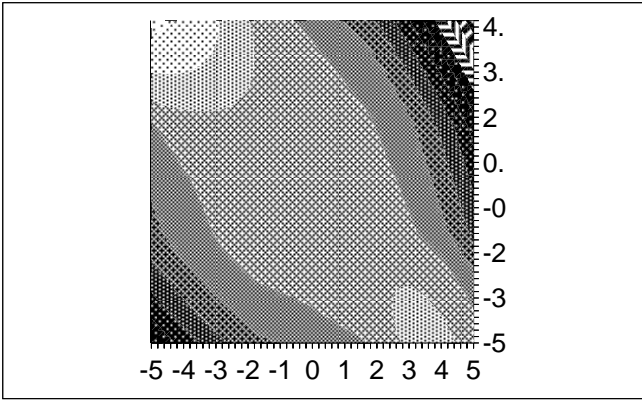


Figure 12: DAX NIS

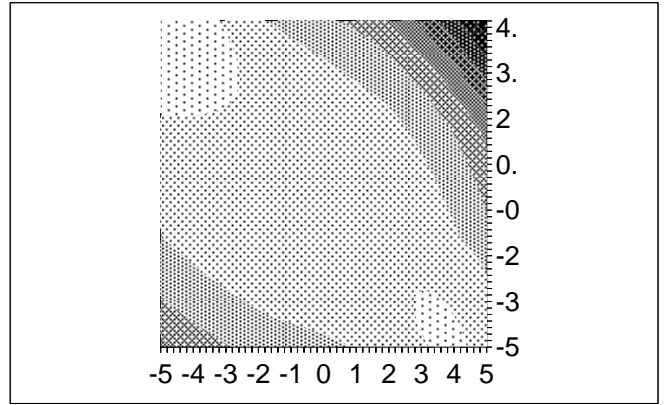


Figure 13: FTSE NIS

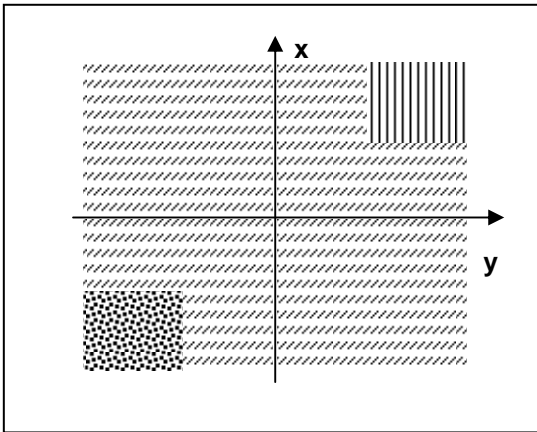


Figure A.1: Extreme Events on a Bivariate Support

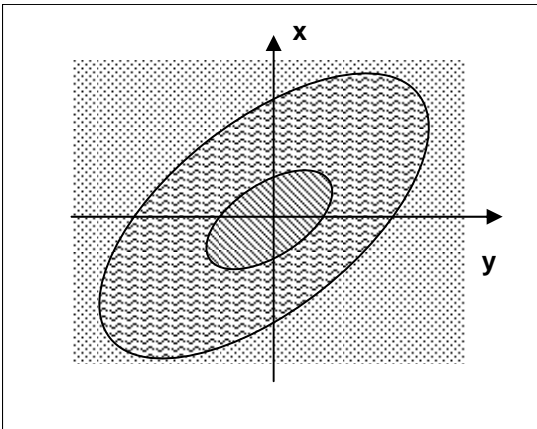


Figure A.2: A non-linear support partition

		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

		DAX	FTSE			DAX	FTSE
ω	Coeff.	0.018	0.010	Ψ_1 (DAX)	Coeff.	0.000	0.046
	St.dev.	1.8E-04	1.6E-04		St.dev.	2.1E-03	6.3E-03
B (DAX)	Coeff.	0.940	0.000	Ψ_1 (FTSE)	Coeff.	0.013	0.032
	St.dev.	3.3E-05	2.1E-03		St.dev.	1.3E-03	2.2E-03
B (FTSE)	Coeff.	0.000	0.926	Ψ_2 (DAX)	Coeff.	0.000	0.000
	St.dev.	6.2E-04	1.9E-06		St.dev.	3.0E-03	1.4E-03
A ₁ (DAX)	Coeff.	0.210	0.295	Ψ_2 (FTSE)	Coeff.	0.000	0.000
	St.dev.	3.6E-03	7.5E-03		St.dev.	1.5E-03	1.5E-03
A ₁ (FTSE)	Coeff.	0.056	0.222	Ψ_3 (DAX)	Coeff.	0.000	0.001
	St.dev.	3.9E-03	4.8E-03		St.dev.	1.6E-03	1.5E-03
A ₂ (DAX)	Coeff.	0.000	0.000	Ψ_3 (FTSE)	Coeff.	0.157	0.599
	St.dev.	4.8E-04	1.4E-03		St.dev.	4.6E-03	9.5E-03
A ₂ (FTSE)	Coeff.	0.000	0.007	Ψ_4 (DAX)	Coeff.	0.000	0.083
	St.dev.	4.7E-04	6.3E-04		St.dev.	2.0E-03	6.2E-03
A ₃ (DAX)	Coeff.	0.061	0.000	Ψ_4 (FTSE)	Coeff.	0.043	0.000
	St.dev.	4.4E-04	1.1E-03		St.dev.	9.7E-03	1.0E-03
A ₃ (FTSE)	Coeff.	0.004	0.034	Corr	Coeff.	0.730	
	St.dev.	4.7E-04	3.1E-04		St.dev.	3.7E-04	
A ₄ (DAX)	Coeff.	0.494	0.164	Log-Likelihood -1620.626			
	St.dev.	1.2E-02	2.8E-03				
A ₄ (FTSE)	Coeff.	0.247	0.175				
	St.dev.	8.8E-03	2.9E-03				

Table 3: DAMGARCH estimates

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

Table 1: Model dimension