Dynamic Conditional Correlation Models with Asymmetric Multivariate Laplace Innovations

Giovanni Urga*
(*Cass Business School, London, UK, Bergamo University, Italy*)

Juan-Pablo Cajigas
(Investor Bank, London, UK)

Alexios Ghalanos
(*Cass Business School, London, UK*)

First version: 28 April 2007
This version: 2 October 2011

Abstract

In this paper we propose to estimate multivariate GARCH processes and a class of dynamic conditional correlation (DCC) models assuming that the n-dimensional returns series follow the Asymmetric Multivariate Laplace (AML) distribution. The AML distribution is able to capture asymmetry and leptokurtosis which characterise returns from financial assets. Under general conditions, it preserves desirable properties such as finiteness of moments and stability under geometric summation. We prove that the maximum likelihood estimator provides consistent estimates for a variety of DCC models when AML distribution is assumed for standardised residuals. We also prove strict stationary of DCC models. The empirical validity of the proposed framework is tested by fitting 21 FTSE All-World stock indices and 12 bond return indices and evaluate its in-sample performance via alternative risk management measures.We provide clear evidence that in our data set the asymmetric generalised (AGDCC) model with AML distribution overwhelmingly outperforms the variety of DCC models that assume normality of innovations.

J.E.L. Classification: C51, C52, G1.

Keywords: Multivariate GARCH, Dynamic Conditional Correlations, Asymmetric Multivariate Laplace Distribution, Maximum Likelihood Estimation Method, Risk Management Measures.

*Corresponding author: g.urga@city.ac.uk. Centre for Econometric Analysis, Faculty of Finance, Cass Business School, 106, Bunhill Row, London, EC1Y 8TZ (U.K.) Tel. 00.44.(0)20.70408698; Fax. 00.44.(0)20.70408881; http://www.cass.city.ac.uk/faculty/g.urga http://www.cass.city.ac.uk/cea/index.html

1 Introduction

This paper proposes a multivariate time-varying framework for modelling and forecasting cross-market correlations, where innovations are assumed to follow the Asymmetric Multivariate Laplace (AML) distribution introduced by Kotz et al. (2003). A good understanding of the dynamic properties of cross-market correlation (or dependence across markets) is vital to assess the level of integration between international markets, both for investment purposes and for increasing the capacity to produce reliable forecasts (Engle, 2009). Modelling the dynamics of volatilities of returns from financial assets has been one of the work horses in the development of financial econometrics over the last years (Bollerslev, 2001; Engle, 2001). Nonetheless, until recently, most of the advances in the field have been developed essentially in univariate cases. Bollerslev (2009) provides "an encyclopedic type reference guide to the long list of ARCH acronyms that have been used in the literature". The growth in techniques for modelling the dynamics of covariances and correlations has lagged considerably behind that in modelling time-varying volatility. One of the main reasons for this uneven expansion is the problem posed by the so-called “curse of dimensionality”, which is related to the estimation of unrestricted multivariate GARCH (MGARCH) models in high-dimensional settings. Bauwens et al. (2006) and Silvennoinen and Teräsvirta (2009a) provide comprehensive surveys on MGARCH models.

Among the alternative MGARCH specifications, the dynamic conditional correlation (DCC) model proposed by Engle (2002) and Tse and Tsui (2002) has proved particularly suitable in providing a parsimonious, flexible and feasible model that significantly reduces the “curse of dimensionality”. In this framework, the dynamic variance-covariance matrix of conditional returns is specified as a function of univariate variances and linear correlations. When the model is estimated by maximum likelihood, the DCC approach allows the log-likelihood function to be split into two parts, the first is used to estimate the parameters of the univariate volatilities while the second the correlations. By using this two-step estimation technique, large systems can be consistently estimated with limited computational costs and without imposing too many restrictions. Hafner et al. (2005), Silvennoinen and Terasvirta (2009b), Engle and Colacito (2006), Pelletier (2006), Palandri (2009), and Audrino and Trojani (2011) propose interesting developments of this framework.

A vital assumption of the DCC model is that standardized residuals are normally distributed. The assumption of Gaussian innovations does not allow to embody the modelling of skewness and excess kurtosis, which are two important features of financial data. There are two strands of literature that drop the normality assumption. The first uses nonparametric and semiparametric methods of investigations. For the univariate case, the main contributions relate to Engle and González-Rivera (1991), Drost and Klaassen (1997), González-Rivera (1997), González-Rivera and Drost (1999). There are also a few papers studying the multivariate case, such Hafner et al.(2005), Long and Ullah (2005) and Hafner and Rombouts (2007). However, in this case the estimation remains feasible only when a small number of assets is considered. The second strand of contributions uses thick-tailed distributions to improve efficiency, as, for the univariate case, in Bollerslev (1987), Baillie and Bollerslev (1989), Nelson (1991). Fiorentini et al. (2003) propose to relax normality using the multivariate Student-t distribution; Bauwens and Laurent (2005) adopt a multivariate skewed-Student-t distribution to fit a DCC (1,1) model using a small number of assets. However, because of the presence of the degrees of freedom, the use of multivariate Student-t distributions invalidates the two stage approach, which is a crucial feature of the DCC approach. Thus, both nonparametric/semiparametric methods and thick-tailed distributions do not overcome the limits of the standard assumption of normally distributed returns.

Our paper is in the spirit of Mencia and Sentana (2005, 2009, 2010) who use a Generalized Hyperbolic (GH) distribution in a model where the variance matrix dynamics follow a conditionally heteroskedastic single factor model and the conditional variance of the factor obeys a univariate GQARCH (1,1) process. This framework allows for flexible tail modelling, but at the cost of limiting the inclusion of rich dynamics for the conditional variance matrix due to "curse of dimensionality". For the case of highly parameterised specification, such as for instance the asymmetric generalised DCC model (AGDCC), the estimation using the GH distribution is extremely difficult. Note that both the AML and the normal distributions are nested into the GH as detailed in Kotz et al. (2003) and Kozubowski et al. (2010).

In this paper we relax the assumption of Gaussian distributed returns for DCC models by allowing for AML distributed returns, able to capture leptokurtosis and asymmetry. In the univariate context, the Laplace or double-exponential distribution has been widely used in modelling financial data. See for instance Madan and Seneta (1990), Madan et al. (1998), Linden (2001), Heyde and Kou (2004), Komunjer (2005). To the best of our knowledge, this is the first paper that proposes the use of the AML distribution to model financial returns in a MGARCH setting. This multivariate distribution has desirable properties such as additivity and finiteness of moments, and a density function with a closed-form that makes the maximum likelihood estimation method easy to implement. Thus, the AML distribution is more suited than stable Paretian distributions for modelling financial data[1]. The AML distribution belongs to the subclass of geometric stable distributions, a characteristic that in the case of the AML distribution can be used to model linear combinations of random variables with univariate symmetric Laplace distributions. Further, the AML distribution is a location-scale mixture of normals and therefore enjoys the flexibility of mixtures of normals. This feature is extremely important as it allows to use this distribution in the computation, for instance, of the parametric-VaR of portfolios of financial assets, characteristic that was thought exclusive of the Pareto-stable distribution and in particular of its most widely used limiting case, such as the normal distribution[2]. The empirical applications in this paper indicate that the AML distribution fits the data better than the multivariate Gaussian counterpart. Thus, the distinguishing features of the AML distribution overcome many drawbacks of the existing literature and allow for more reliable inference.

The remainder of the paper is organized as follows. In Section 2, we present the theoretical

[1] Pareto stable distributions allow for skewness and excess kurtosis, but they have infinite second moments, which prevent the traditional estimation of, for example, market risk using variances of returns and lower-order dependence using correlation measures between returns.

[2] The property that linear combination of multivariate AML distributiions are AML is going to be useful "only" for the one-day ahead VaR computation if one uses a GARCH model for the variance. For more than one-day ahead one would need to compute the VaR through simulations. Basel II ask 10-day ahead VaR computations. We wish to thank Denis Pelletier for bringing this point to our attention.

framework of DCC models with AML distributed standardised residuals. In Section 3, we introduce the minimum contrast estimator of Pfanzgal (1969) and its maximum likelihood extension. In Section 4, we introduce the AGDCC models with AML innovations, providing in Section 4.1 condition for consistency of the MLE for MGARCH models when an AML distribution is assumed for standardised residuals. In Section 5, we provide the proof of consistency of the MLE parameter estimates in the AGDCC (and its nested models). We also prove strict stationary of DCC models. In Section 6, we report the results from an empirical application using a sample of 21 FTSE All-World stock indices and 13 bond return indices. We also report estimates of in-sample risk management measures to evaluate, by means of a backtesting analysis, the performance of the normal versus the AML distribution in the estimation of the variance-covariance matrices. Section 7 concludes.

NOTATION. We use "$\circ$" to denote the Hadamard product, "$\|\cdot\|$" the Euclidean norm of a vector, "$\rightarrow$" the ordinary limit, "$\overset{D}{=}$"equality in distribution "$\overset{a.s.}{=}$" almost sure equality, "$\overset{a.s.}{\rightarrow}$" almost sure convergence; "$\xrightarrow{d}$" convergence in distribution, and "$\xrightarrow{p}$" convergence in probability.

2 Dynamic Conditional Correlation Models with Asymmetric Multivariate Laplace Distribution

Consider the $n-$dimensional zero-mean return process $r_t \in \mathbb{R}^n$, $t = 1, ..., T$

$$r_t = H_t^{1/2}(\theta)\,\varepsilon_t, \tag{1}$$

$$H_t = var(r_t|\Omega_{t-1})$$

where θ is the set of parameters to estimate, Ω_{t-1} is the information set at time $t-1$, and ε_t is an i.i.d. process with unit variance. In the DCC setting, H_t is modelled directly as a function of dynamic univariate variances and dynamic linear correlations

$$H_t = D_t R_t D_t \tag{2}$$

where $D_t \in \mathbb{R}^{n\times n}$ is a diagonal matrix with elements $\sqrt{h_{it}}, i = 1, ..., n,\ t = 1, .., T$, and R_t is defined as

$$R_t = (Q_t^*)^{-1} Q_t (Q_t^*)^{-1} \tag{3}$$

where Q_t^* is a diagonal matrix of the form

$$Q_t^* = (Diag\ Q_t)^{1/2} \tag{4}$$

and

$$Q_t = \left(1 - \sum_{i=1}^{q} a_i - \sum_{j=1}^{p} b_j\right) \overline{Q} + \sum_{i=1}^{q} a_i \varepsilon_{t-i} \varepsilon'_{t-i} + \sum_{j=1}^{p} b_j Q_{t-j}. \tag{5}$$

Here, $\overline{Q} \in \mathbb{R}^{n\times n}$ is the unconditional variance-covariance matrix of ε_t, $R_t \in \mathbb{R}^{n\times n}$ is a conditional correlation matrix, and α_i and β_j are positive-scalar parameters satisfying the constraints $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$, where q corresponds to the number of autoregressive shock lags, p corresponds to the number of persistence lags. Specification (3) ensures that R_t is a valid correlation matrix, while (2) and (5), in addition to the condition of stationarity, ensure that H_t is positive definite. The decomposition in (2) and (5) is particularly appealing because it allows for a two step estimation procedure that makes feasible the estimation of high-dimensional processes.

Specification (5) can be enriched by allowing for asymmetries in conditional correlations as well as for asset-specific correlations, as proposed by Cappiello et al. (2006). We refer to this general model as the Asymmetric Generalised Dynamic Conditional Correlation (AGDCC henceforth) (p,q,s) model, where in addition to q and p defined above, s corresponds to the number of asymmetric shock lags. Thus, specification (5) can be augmented as follows:

$$\begin{aligned} Q_t \ &= \ \overline{Q} - \sum_{i=1}^{q} A_i A'_i \circ \overline{Q} - \sum_{j=1}^{p} B_j B'_j \circ \overline{Q} - \sum_{k=1}^{s} G_k G'_k \circ \overline{N} \\ &+ \sum_{i=1}^{q} A_i A'_i \circ \varepsilon_{t-i} \varepsilon'_{t-i} + \sum_{j=1}^{p} B_j B'_j \circ Q_{t-j} + \sum_{k=1}^{s} G_k G'_k \circ \eta_{t-k} \eta'_{t-k} \end{aligned} \tag{6}$$

A, B, and $G \in \mathbb{R}^{n\times n}$ are diagonal parameter matrices with elements a_{ii}, b_{ii} and g_{ii} respectively; $\eta_t = I[\varepsilon_t < 0] \circ \varepsilon_\tau$ with $I[.]$ an indicator function taking value 1 if the argument is true and 0 otherwise, and $\bar{N} = E(\eta_t \eta_t')$. Q_t is positive-definite if

$$\overline{Q} - \sum_{i=1}^{q} A_i A_i' \circ \overline{Q} - \sum_{j=1}^{p} B_j B_j' \circ \overline{Q} - \sum_{k=1}^{s} G_k G_k' \circ \overline{N} \tag{7}$$

is positive definite, which for the scalar ADCC simplifies to the constraints $a^2 + b^2 + \delta g^2 < 1$, where δ is the maximum eigenvalue $\left[\overline{Q}^{-1/2}\overline{NQ}^{-1/2}\right]$ (see Cappiello et al. 2006, p. 544)

The AGDCC(p, q, s) model nests the following specifications:

- DCC(p, q) when $G = 0$ and A, B are scalars.

- ADCC (p, q, s) when A, B and G are scalars.

- GDCC (p, q) when $G = 0$.

For the ADCC and the GDCC models, equation (6) is respectively:

$$\begin{aligned} Q_t^{ADCC(p,q,r)} &= \left(1 - \sum_{i=1}^{q} a_i - \sum_{j=1}^{p} b_j\right)\overline{Q} - \left(\sum_{k=1}^{s} g_k\right)\overline{N} \\ &\quad + \sum_{i=1}^{q} a_i \left(\varepsilon_{t-i}\varepsilon_{t-i}'\right) + \sum_{j=1}^{p} b_j Q_{t-j} + \sum_{k=1}^{s} g_k \left(\eta_{t-k}\eta_{t-k}'\right), \end{aligned} \tag{8}$$

$$\begin{aligned} Q_t^{GDCC(p,q)} &= \overline{Q} - \sum_{i=1}^{q} A_i A_i' \circ \overline{Q} - \sum_{j=1}^{p} B_j B_j' \circ \overline{Q} + \sum_{i=1}^{q} A_i A_i' \circ \varepsilon_{t-i}\varepsilon_{t-i}' \\ &\quad + \sum_{j=1}^{p} B_j B_j' \circ Q_{t-j}. \end{aligned} \tag{9}$$

All models can be estimated in two stages. In the first stage, univariate volatilities are estimated by assuming zero correlations. In the second stage, correlations are estimated once standardised residuals are obtained. In the paper, we refer to the process in the first stage as AGDCC*, and to the process in the second stage as AGDCC$^\diamondsuit$.

Asymmetric Multivariate Laplace Distribution. An important assumption of the DCC model is that standardized residuals are normally distributed. Under the assumption of normality, Quasi Maximum Likelihood Estimator (QMLE) can be employed to get feasible and consistent though inefficient DCC coefficients of conditional correlations (Bollerslev and Wooldridge, 1992). Nevertheless, where time-varying volatilities are estimated by assuming a normal-GARCH process for the innovations, even for correctly specified models, statistically significant levels of skewness and excess kurtosis can still be found. This feature has important implications not only for the econometric properties of parameter estimates, but also for the use of these models in financial applications such as portfolio allocation, VaR and Expected Shortfall analyses.

In this paper, we propose to estimate conditional correlations assuming that the innovations follow the asymmetric Laplace distribution proposed by Kozubowski and Podgorski (2001) as a subclass of geometric stable distributions. In particular, we adopt the multivariate generalisation of the asymmetric Laplace laws in Kotz et al. (2003).

In the geometric stable model, the return $r_{f(p)}$ is considered to be the sum of smaller returns $r^{(i)}$ over the period of time $f(p)$ which is a stopping time random variable with geometric probability function $P(f(p) = j) = p(1-p)^{j-1}$, $j = 1, 2, ...$ The geometric stable distribution can be approximated to a normalised geometric stable model sum when the p parameter of the stopping time function $f(p)$ approaches zero. More formally, the random vector r has a geometric stable distribution in $\mathbb{R}^n$ if and only if, as $p \to 0$

$$a(p) \sum_{i=1}^{f(p)} \left(\boldsymbol{\kappa}(p) + \mathbf{r}^{(i)}\right) \xrightarrow{d} r, \tag{10}$$

where $\left\{ \{\mathbf{r}^{(d)} = \left(r_1^{(d)}, ..., r_n^{(d)}\right)', d \geq 1 \right\}$ is a sequence of i.i.d. random vectors in $\Re^n$ independent of $f(p)$, $a(p) > 0$, $\boldsymbol{\kappa}(p) \in \mathbb{R}^n$. Kozubowski and Podgorski (2001) show that when each vector in r has mean m_i, $i = 1, ...n$, a variance σ_{ij}, $i = 1, ...n$, $j = 1, ...n$, and for $a(p) = \sqrt{p}$ and $\boldsymbol{\kappa}(p) = m\left(\sqrt{p} - 1\right)$, the random variable r defined in (10) has an AML distribution with the characteristic function

$$\Psi(\mathbf{t}) = \frac{1}{1 + \frac{1}{2}\mathbf{t}'\mathbf{H}\mathbf{t} - \mathbf{i}\mathbf{t}'\mathbf{m}} \tag{11}$$

where $\mathbf{t} \in \mathbb{R}^n$, and $\mathbf{H} \in \mathbb{R}^{n \times n}$ is a positive-definite matrix.

The density function of the $n-$variate AML distribution allowing for time dependency in H_t and r_t is given by

$$f(\mathbf{r}) = \frac{2 \exp\left(r_t' H_t^{-1} m\right)}{(2\pi)^{n/2} \left|H_t\right|^{1/2}} \left(\frac{r_t' H_t^{-1} r_t'}{2 + m' H_t^{-1} m}\right)^{v/2} K_v \left(\sqrt{(2 + m' H_t^{-1} m)(r_t' H_t^{-1} r_t')}\right) \tag{12}$$

where $v = (2 - n)/2$ and $K_v(u)$ is the modified Bessel function of the third kind defined by $K_v(u) = \frac{(u/2)^v \Gamma(1/2)}{\Gamma(v+1/2)} \int_1^\infty e^{-ut}(t^2 - 1)^{v-1/2} dt$, $u > 0, v \geq -1/2$. The vector m is the location parameter and the matrix H is the scale parameter of this distribution. Note that $m \equiv Hb$ where $b \in \mathbb{R}^n$.

A very important characteristic of the AML distribution is that it is unimodal with mode equal to zero. Thus, the m parameter does not only determine the mean of the distribution, but also its level of asymmetry. When $m = 0$, i.e. $b = 0$, the distribution is symmetric and it collapses, as it can clearly be seen in equation (11), to the elliptical case (see discussion in Johnson and Kotz, 1972), which is the distribution used in all models with zero mean residuals.

As shown in Kotz et al. (2003), AML distributions can also be obtained as a limiting case of the GH distribution, introduced by Barndorff-Nielsen (1977). These are location-scale mixtures of normal distributions, i.e. if we assume that $\mathbf{w}$ has a GH distribution in $\mathbb{R}^n$ then

$$\mathbf{r} \stackrel{D}{=} \boldsymbol{\mu} + m\boldsymbol{\xi} + \boldsymbol{\xi}^{1/2}\mathbf{Z} \tag{13}$$

where $\mathbf{Z} \sim N_n(0, H)$, $\boldsymbol{\mu} \in \mathbb{R}^n$, and $\boldsymbol{\xi}$ is a generalised inverse Gaussian (GIG) variable with parameters ν, γ, and δ, i.e. $\boldsymbol{\xi} \sim GIG(\nu, \gamma, \delta)$, where ν and γ are shape parameters, and δ scale parameter. AML distributions appear when $\mu = 0$ and when $\boldsymbol{\xi}$ is not $GIG(\nu, \gamma, \delta)$ but standard exponential, i.e. $\boldsymbol{\xi} \sim EXP(1)$. Note that the limiting case $GIG(1, 0, 2)$ is equivalent to $EXP(1)$, just as the multivariate normal is a limiting case of the GH distribution (see Kotz et al. 2003, and Kozubowski et al. 2010).

The representation of the AML distribution as a location-scale mixture of normal distributions is given by

$$\mathbf{r} \overset{D}{=} m\boldsymbol{\xi} + \boldsymbol{\xi}^{1/2} Z \tag{14}$$

where in this case $\boldsymbol{\xi} \sim EXP(1)$. From this it can easily be seen that $E(\mathbf{r}) = m$ and $Var(\mathbf{r}) = H + mm'$. This is of particular importance for the estimation of the MGARCH model. Contrary to the Gaussian case, the variance of a random variable with AML distribution does not coincide with the scale parameter of the distribution. Note that $Var(\mathbf{r}) = H$ only when the distribution is elliptical, i.e. when $m = 0$.

In contrast with the majority of GH distributions, the AML distribution in the special case $m = 0$ is stable, just as the normal. This condition implies an important property necessary for the modelling of financial portfolios known as the additivity property, which is basically the concept that a linear combination of independent random variables with stability index α is also stable with the same parameter α (see Khindanova et al. 2001).

Pareto stable distributions are stable under random summation. Formally, the random variable $\mathbf{r}$ is said to be Pareto stable if for any $a_i > 0$, $i = 1, ..., d$, there exist a constant $c = d^{1/\alpha}$ and $u_d \in \mathbb{R}^n$ for any $d \geq 2$ such that

$$a_1\mathbf{r}^{(1)} + ... + a_d\mathbf{r}^{(d)} \overset{D}{=} c\mathbf{r} + u_d \tag{15}$$

where $\mathbf{r}^{(1)}, ..., \mathbf{r}^{(d)}$ are independent copies of $\mathbf{r}$. In an alike way Laplace laws are stable, but under geometric summation instead of random summation. To be able to preserve stability we have to constrain the normalising constants $a(p)$ and $\boldsymbol{\kappa}(p)$ in (10) to

$$a(p) = \sqrt[\alpha]{p}, \quad \boldsymbol{\kappa}(p) = 0 \tag{16}$$

The first condition implies that for the case of the AML distribution $\alpha = 2$. This is the same α value of the normal distribution which is the only Pareto-stable distribution with a finite second moment. The second condition $\kappa(p) = 0$ implies $m = 0$.

In the next section, we introduce the consistent minimum contrast estimator and its extension to maximum likelihood estimator to estimate multivariate ARCH models under normality of innovations.

3 Consistency of minimum contrast estimates and its maximum likelihood extension.

Let us assume that $\{r_t, t = 1, ..., T\}$ is a multivariate ARCH process

$$r_t = \Upsilon_\theta(\underline{r}_{t-1}) + \Delta_\theta(\underline{r}_{t-1})\varepsilon_t, \tag{17}$$

where $\underline{r}_{t-1} = (r_{t-1}, r_{t-2}, ...)'$, Υ_θ is a measurable function from $(\mathbb{R}^d)^n \to \mathbb{R}^d$, Δ_θ is a measurable function from $(\mathbb{R}^d)^n \otimes \mathbb{R}^d$, θ is a vector of parameters that belong to a parameter space Θ, and $\varepsilon_t \sim i.i.d.(0, R)$, where R is a square matrix.

The conditional covariance matrix of the error $\Delta_\theta(\underline{r}_{t-1})\varepsilon_t$ is defined as

$$H_t(\theta) = \Delta_\theta(\underline{r}_{t-1})R\Delta_\theta(\underline{r}_{t-1})'.$$

In what follows we will refer to θ_0 as to the true (population) value of θ.

Consider now the following assumptions:

P1. Θ is a compact (closed and bounded) and convex parameter space.

P2. The stochastic process $\{r_t, t = 1, ..., T\}$ is strictly stationary and ergodic.

P3. The function $F(\theta_0, \theta) = E_{\theta_0}(f(\underline{r_1}, \theta))$, has a unique finite minimum at θ_0, where f is a real valued measurable function continuous in θ.

P4. $\forall \theta \in \Theta, E_{\theta_0}(\inf(f_*(\theta, \rho), 0)) > -\infty$, where $f_*(\theta, \rho) = \inf\{f(\underline{r_t}, \theta'), \theta' \in B(\theta, \rho)\}$ where $B(\theta, \rho)$ is a ball of center θ and radius ρ.

Assumption P1 requires the knowledge of bounds of the true parameter value. As discussed in Newey and McFadden (1994), this assumption can be relaxed if the instability of $F_T(\underline{r_t}, \theta)$, contrast process such that $F_T(\underline{r_t}, \theta) = T^{-1}\sum_{t=1}^{T} f(\underline{r_t}, \theta)$, when θ is unbounded is not extreme. Assumption P3 secures *identification*, i.e., the existence of a unique minimum of $F(\theta_0, \theta)$ when this is evaluated at the true parameter value. Assumption P2 in conjunction with Assumption P4 secure uniform convergence of $F_T(\underline{r_t}, \theta)$ to $F(\theta_0, \theta)$, condition that is also required for consistency (Newey and McFadden, 1994).

Consistency of the MLE estimator under Assumptions P1-P4 has been proved by Pfanzgal (1969) as a special case of a more general class of estimators known as "minimum contrast estimators". We report the result here in the following lemma:

Lemma 1 *Let* $(\Omega, \digamma, P)$ *be a probability space,* $\theta \in \Theta$ *and* $F_T(\underline{r_t}, \theta)$ *be a contrast process such that* $F_T(\underline{r_t}, \theta) = T^{-1} \sum_{t=1}^{T} f(\underline{r_t}, \theta)$*. Under assumptions P1-P4, the minimum contrast estimator* $\widehat{\theta} \overset{a.s}{\rightarrow} \theta_0$.

4 Estimating AGDCC Models with AML Distributed Standardised Residuals

We turn now to the estimation of DCC models employing AML distributions. The likelihood function $L_T^{AML}(\theta)$, with θ the set of parameters to estimate, assuming a AML distribution for the conditional returns is proportional to

$$
\begin{aligned}
L_T^{AML}(\theta) \quad \propto \quad &= \sum_{t=1}^{T} \Bigg\{ r_t' H_t^{-1} m - \frac{1}{2} \ln |H_t| + \\
&\frac{v}{2} \left(\ln(r_t' H_t^{-1} r_t) - \ln(2 + m' H_t^{-1} m) \right) \\
&+ \ln \left[K_v \left(\sqrt{(2 + m' H_t^{-1} m)(r_t' H_t^{-1} r_t)} \right) \right] \Bigg\}.
\end{aligned} \tag{18}
$$

From $H_t \equiv D_t R_t D_t$, equation (18) can be written as

$$
\begin{aligned}
L_T^{AML}(\theta) \quad = \quad & \sum_{t=1}^{T} \Bigg\{ r_t' (D_t R_t D_t)^{-1} m - \frac{1}{2} \ln |(D_t R_t D_t)| \\
&+ \frac{v}{2} \left(\ln(r_t' D_t^{-1} R_t^{-1} D_t^{-1} r_t) - \ln(2 + m'(D_t R_t D_t)^{-1} m) \right) \\
&+ \ln \left[K_v \left(\sqrt{(2 + m'(D_t R_t D_t)^{-1} m)(r_t' D_t^{-1} R_t^{-1} D_t^{-1} r_t)} \right) \right] \Bigg\}
\end{aligned} \tag{19}
$$

The m parameter cannot be estimated in the first step because of zero mean properties of data. Thus, assume $R_t = I_n$ and $m = 0$ and let us denote with ζ the set of parameters in the matrix of variances D_t. The first stage likelihood function is

$$
\begin{aligned}
L_T^{AML}(\boldsymbol{\zeta}) \quad = \quad & \sum_{t=1}^{T} \Bigg\{ -\frac{1}{2} \ln \left| D_t^2 \right| + \frac{v}{2} \left(\ln(r_t' D_t^{-2} r_t) - \ln(2) \right) \\
&+ \ln \left[K_v \left(\sqrt{2(r_t' D_t^{-2} r_t)} \right) \right] \Bigg\}
\end{aligned} \tag{20}
$$

Contrary to the normal case, $L_T^{AML}(\zeta)$ cannot be expressed as the sum of n-log-likelihood functions, i.e. the parameters in ζ have to be estimated maximizing one single log-likelihood function. This, however, does allow to continue to use the two-step estimation technique although it does extend the computing time for estimation, unless we use normal distribution for the first stage.

Defining $\varepsilon_t = r_t' D_t^{-1}$ and $\varepsilon_t^* = m' D_t^{-1}$, the second-stage log-likelihood is given by

$$\begin{aligned} L_T^{AML}(\boldsymbol{\varphi} \mid \widehat{\boldsymbol{\zeta}}) \;=\; & \sum_{t=1}^{T} \Big\{ \varepsilon_t R_t^{-1} (\boldsymbol{\varepsilon}_t^*)' - \frac{1}{2} \ln |R_t| + \\ & \frac{v}{2} \left(\ln(\varepsilon_t R_t^{-1} \varepsilon_t') - \ln(2 + \varepsilon_t^* R_t^{-1} (\boldsymbol{\varepsilon}_t^*)') \right) \\ & + \ln K_v \left(\sqrt{(2 + \varepsilon_t^* R_t^{-1} (\boldsymbol{\varepsilon}_t^*)')(\varepsilon_t R_t^{-1} \varepsilon_t')} \right) \Big\} \end{aligned} \tag{21}$$

where $\boldsymbol{\varphi}$ is the set of parameters in R_t.

The computation of (20) and (21) is feasible but it may be computationally demanding. An alternative is the numerical solution suggested in Kotz et al. (2003). In this case, maximum likelihood estimation method yields consistent and asymptotically efficient parameter estimates when the assumed distribution is correctly specified. The use of a flexible distribution like the AML is in this regard very important.

From the minimum contrast estimator of Pfanzgal provided in Lemma 1, we adapt the framework of Jeantheau (1998) to the case of AML innovations and in Section 4.1 below, we provide condition for consistency of the MLE for MGARCH models when an AML distribution is assumed for standardised residuals. In Section 5, we provide the proof of consistency of the MLE parameter estimates in the AGDCC and its nested models.

4.1 Consistent MLE under AML innovations

We now prove that $\widehat{\theta}$ is consistent for the case of AML-MLE.

Consider the negative log-density of the AML distribution,

$$\begin{aligned} f(\underline{r}_t, \theta) &= \frac{1}{2}\log(\det H_t(\theta)) - (r_t' H_t^{-1}(\theta) m) \\ &-\frac{\nu}{2}\left[\log(r_t' H_t^{-1}(\theta) r_t) - \log(2 + m' H_t^{-1}(\theta) m)\right] \\ &- \log\left(K_\nu\left(\sqrt{(2 + m' H_t^{-1}(\theta) m)(r_t' H_t^{-1}(\theta) r_t)}\right)\right) \end{aligned} \tag{22}$$

where $K_\nu(\cdot)$ is the modified Bessel function of the second kind (also called modified Bessel function of the third kind. See Bowman, 1958, and Relton 1965).

We now adapt Lemma 1 to the MLE case in a similar spirit to Jeantheau (1998). This requires specialising Assumptions P1-P4 as follows:

Assumption 1. Θ is a compact, convex parameter space, where $\theta \in \Theta$.

Assumption 2. The process defined by (17) is strictly stationary and ergodic.

Assumption 3. There exists a constant $[c > 0]$ such that $\det(H_t(\theta)) \geq c$ for all t, and $\theta \in \Theta$.

Assumption 4. $\forall \theta \in \Theta, E_{\theta_0}\left(|\log(\det(H_t(\theta))|\right) < \infty$, where θ_0 denotes the true parameter value.

Assumption 5. The function $H_t(\theta)$ is such that $\forall \theta \in \Theta, \forall \theta_0 \in \Theta, H_t(\theta) = H_t(\theta_0) \Rightarrow \theta = \theta_0$.

Assumption 6. The function $H_t(\theta)$ is a continuous function of the parameter θ.

Assumption 7. $H_t(\theta)$ is positive definite all $\forall \theta \in \Theta$.

In what follows, we show how for the specific case where the log-density of r_t is given by (22), the four assumptions listed by Pfanzgal (1969) can be replaced by Assumptions 1 to 7.

The following lemma shows under what conditions P4 can be replaced by Assumption 3:

Lemma 2 *Let r_t be defined by $f(\underline{r}_t, \theta)$ as in (22). Then, Assumptions 3, 4 and 7 imply that P4 holds.*

Proof. By Assumption 7, $H_t^{-1}(\theta)$ will be positive definite and by Assumption 3 $F_T(\underline{r_t}, \theta) > \log(c)$. Also, in view of that the highest eigenvalue $H_t(\theta)$ is bounded from above, the log-likelihood has a lower bound. Given that the probability that all elements in r_t are zero is zero we have,

$$K_\nu\left(\sqrt{(2+m'H_t^{-1}(\theta)m)(r_t{}'H_t^{-1}(\theta)r_t)}\right)<\infty \tag{23}$$

■

Note that Assumption P4 is necessary to ensure that the minimum of the log-density function does exist. Assumption 3 is much more easy to verify and Lemma 2 provides the connection between the two.

We need the following Lemma to show that P3 (identification condition) holds for the case when the AML distribution is assumed.

Lemma 3 *If Assumptions 1-5 hold, and if $f(\underline{r}_t,\theta)$ is the density function of ε_t, then $F_T(\underline{r_t},\theta)\overset{a.s.}{\to}F(\theta_0,\theta)$ and $F(\theta_0,\theta)$ has a unique optimum at θ_0.*

Proof. The first part of the Lemma follows from the strict stationarity and ergodicity of the process $\{r_t\}$ (Assumption 2) and the ergodic theorem. To prove that $F_T(\theta_0,\theta)$ has a unique finite optimum in θ_0, first by the concavity of the logarithm function and by Jensen's Inequality we have that

$$E_{\theta_0}\left(\log\left(\frac{L\left(H_t(\theta),m(\theta);r_t\right)}{L\left(H_t(\theta_0),m(\theta_0);r_t\right)}\right)\right)\le\log E_{\theta_0}\left(\left(\frac{L\left(H_t(\theta),m(\theta);r_t\right)}{L\left(H_t(\theta_0),m(\theta_0);r_t\right)}\right)\right) \tag{24}$$

where $L\left(H_t(\theta_0),m(\theta_0);r_t\right)$ is the likelihood function evaluated at the true parameter value, and $L\left(H_t(\theta),m(\theta);r_t\right)$ is the likelihood function evaluated at any other parameter value in the compact (Assumption 1) parameter space Θ. From

$$\begin{aligned}E_{\theta_0}\left(\frac{L\left(H_t(\theta),m(\theta);r_t\right)}{L\left(H_t(\theta_0),m(\theta_0);r_t\right)}\right) &= \int\left(\frac{L\left(H_t(\theta),m(\theta);r_t\right)}{L\left(H_t(\theta_0),m(\theta_0);r_t\right)}\right)L\left(H_t(\theta_0),m(\theta_0);r_t\right)dr\\ &= \int L\left(H_t(\theta),m(\theta);r_t\right)dr=1,\end{aligned} \tag{25}$$

we can rewrite equation (24) as

$$E_{\theta_0}\left(\log\left(\frac{L\left(H_t(\theta),m(\theta);r_t\right)}{L\left(H_t(\theta_0),m(\theta_0);r_t\right)}\right)\right)\le 0 \tag{26}$$

and

$$E_{\theta_0}\left(\log L(H_t(\theta), m(\theta); r_t)\right) \leq E_{\theta_0}\left(\log L\left(H_t(\theta_0), m(\theta_0); r_t\right)\right). \tag{27}$$

The equality holds if and only if $H_t(\theta) = H_t(\theta_0)$ and $m(\theta) = m(\theta_0)$. Therefore, from Assumption 5, we have $F(\theta_0, \theta) = F(\theta_0, \theta_0)$ if and only if $\theta = \theta_0$. ■

The following theorem provides the proof of consistency of parameter estimates using the AML-MLE

Theorem 1 *Under Assumptions 1-7, the AML-Maximum Likelihood Estimator (AML-MLE) for the multivariate heteroskedastic model (17) is consistent,* $\widehat{\theta} \xrightarrow{p} \theta_0$.

Proof. The MLE is a minimum contrast estimator when the maximization is taken over the negative-log-likelihood function. The consistency of a minimum contrast estimator was proven by Pfanzgal (1969) under Assumptions P1-P4 in Lemma 1. Assumptions 1-2 replace Assumptions P1-P2. For a process $\{r_t, t = 1, ..., T\}$ following the dynamics of process (17) and parameters estimated by *AML-MLE,* we have that Assumption P4 holds if Assumption 3, 4 and 7 hold (Lemma 2), and that Assumption P3 holds if Assumptions 3-5 hold (Lemma 3) hold. ■

5 Consistency of MLE for AGDCC Models with AML Innovations

In the previous section, we provided the conditions for consistency of the AML-MLE for a general multivariate ARCH models. We now apply this result to the specific case of the AGDCC model; we verify under what conditions Assumptions 1-7 hold in the case of this particular multivariate ARCH model. As already mentioned in Section 2, we refer to the process defined by the first-step estimation as AGDCC*, and to the process defined by the second-step estimation as AGDCC$^\Diamond$.

5.1 *AGDCC** model

In this section, we apply the result of consistency derived in Section 4.1 to the particular case when the multivariate ARCH process is of the type AGDCC*.

5.1.1 Stationarity

In order to evaluate under what conditions Assumption 2 (strict stationarity) is valid, and given that for this process $R = I_n$, we write (3) as

$$r_{it} = \sqrt{h_{it}}\varepsilon_{it}, \qquad i = 1, ..., n \tag{28}$$

$$h_{it} = w_i + \sum_{j=1}^{q} \alpha_{ij} r_{it-j}^2 + \sum_{k=1}^{p} \beta_{ik} h_{it-k} \tag{29}$$

We can rewrite equation (2) as

$$H_t = \begin{bmatrix} h_{1,t} & \dots & 0 \\ \vdots & \ddots & 0 \\ 0 & \dots & h_{n,t} \end{bmatrix} \tag{30}$$

Let us define $diagH_t' = [h_{1t}, ..., h_{nt}]'$ as the vector containing the diagonal elements of H_t. Following specification (29) we can write

$$diagH_t = \begin{pmatrix} w_i \\ \vdots \\ w_n \end{pmatrix} + \sum_{j=1}^{q} N_j \begin{pmatrix} r_{1,t-j}^2 \\ \vdots \\ r_{n,t-j}^2 \end{pmatrix} + \sum_{k=1}^{p} M_k diagH_{t-k} \tag{31}$$

where N_j and $M_k \in \mathbb{R}^{n \times n}$, and where we assume that all coefficients are positive. Strict stationarity is stated in the following lemma:

Lemma 4 *Let us assume that* θ_0 *is such that* $\det(I_n - \sum_{i=1}^{n} (N_i - M_j) \lambda^i)$ *has its roots outside the unit circle. Then, the CCC model of Bollerslev (1990) has a unique, strictly stationary and ergodic solution.*

Proof. See Jeantheau (1998, pp.73-76). ■

This result is very useful as it shows how for the CCC model strict stationarity follows from covariance stationarity. Covariance stationarity can be easily verified in this case by checking the roots in $\det(I_n - \sum_{i=1}^{n} (N_i - M_j) \lambda^i)$.

5.1.2 Identification

The next step is to verify conditions for identification. To do this it is convenient first to rewrite equation (29) as

$$P(L)\begin{pmatrix} h_{1,t}(\theta) \\ \vdots \\ h_{n,t}(\theta) \end{pmatrix} = \begin{pmatrix} w_i \\ \vdots \\ w_n \end{pmatrix} + Q(L)\begin{pmatrix} r_{1,t}^2 \\ \vdots \\ r_{n,t}^2 \end{pmatrix} \tag{32}$$

where L is the backshift operator, and P and Q are two matrices with polynomial coefficients such that $P(L) = I_n - \sum_{i=1}^{p} \beta_i L^i$ and $Q(L) = \sum_{i=1}^{q} \alpha_i L^i$ where I_n is the identity matrix. For multivariate GARCH processes, the condition of identification is intimately related to the so-called *minimal* condition in multivariate time series analysis. To provide the results, let us consider the following definition first:

Definition.

(a) A polynomial matrix $M(L)$ with degree d_{ij} , i.e. $M_{ij}(L) = \sum_{l=0}^{d_{ij}} a_{ij,l} L^l$, is column reduced if and only if $\det(a_{ij,d_j}) \neq 0$. We define also $d_j(M) = \sup_i d_{ij}$

(b) Denote by $M * P$ the set of matrices with polynomial coefficients. A square matrix $M(L) \in M * P$ is unimodular if and only if its determinant is independent of L and non zero.

(c) Let $A, B \in M * P$ such that $\det(A) \neq 0$ and $\det(B) \neq 0$. The matrix $D \in M * P$ is called the greatest common left divisor of A and B if and only if every left divisor of D is also a left divisor of A and B, and if and only if left divisor of A and B is also a left divisor of D.

(d) Two matrices $A, B \in M * P$ are coprime if and only if any of their greatest common left divisor is unimodular.

Then, a *minimal* multivariate GARCH process is defined as follows

Lemma 5 *The multivariate GARCH (p,q) specification given in (32) is minimal if*

1. $P(0) = I_d$ and $Q(0) = 0$
2. $det(P) \neq 0$ and $det(Q) \neq 0$

3. P and Q are coprime.

4. $\forall j, 1 \leq j \leq d, d_j(P) = d_j \leq p$ and $d_j(Q) = d_j \leq q$.

5. P or Q is column reduced.

This is not the case for MGARCH models because the greatest common left divisor is not unique for polynomial matrices. This is the reason way the notion of "column-reduced" matrix must be introduced in Definition **(d)**.

We now introduce two additional assumptions:

Assumption 8. There exist two strictly positive constants c_1 and c_2 such that all the w_i elements in (29) are greater than $c_1^{1/n}$ and $\det(R) \geq c_2$

Assumption 9. The formulation at θ_0 for the model (28)-(29) is *minimal.*

Assumption 8 identifies primitive conditions in relation to the existence of a positive bound for the determinant of the conditional variance-covariance matrix $H_t(\theta)$ when this is defined as in (29), and it also allows us to verify more easily conditions stated in Assumption 3. Assumption 9 is related to the identification condition stated in Assumption 5.

The following theorem establishes consistency of the $AGDCC^*$ model estimated by AML-MLE.

Theorem 2 *Under Assumptions 1-9, the use of the AML-MLE for the $AGDCC^*$ model defined by (28) and (29) provides consistent parameter estimates.*

Proof. In Lemma 4, we showed that the $AGDCC^*$ process in strictly stationary and ergodic when it is covariance-stationary. This satisfies Assumption 2. By definition α_i and β_i are positive, therefore by Assumption 8 we have $\det(D_t) \geq c^{1/2}$ and $\det(H_t) \geq c_1 c_2 > 0$, satisfying Assumption 3. Because (29) is weakly stationary, we know that $E_0(h_{i,t}) < +\infty$. By Jensen's inequality, we have that $E_0(\log(h_{i,t})) < +\infty$. This entails that $E_0(\log(\det(H_t))) < \infty$.Further, given that from Assumption 8 $E_0(\log(\det(H_t))) > -\infty$, thus $E_0(|\log(\det(H_t))|) < +\infty$, satisfying Assumption 4. Jeantheau (1998, Proposition 3.4, p. 79) proved that under a weakly stationary solution for the process and under Assumption 8, H_t in the CCC model is given such that $\forall\theta \in \Theta, \forall\theta_0 \in \Theta, H_t(\theta) = H_t(\theta_0) \Rightarrow \theta = \theta_0$, satisfying Assumption 5. Finally, Assumption 6 holds. ■

Some remarks regarding Theorem 2 follow. For a positive-definite matrix A, we have $|A| > 0$. Secondly, Assumption 9 is necessary in order to satisfy identification, avoiding the case where two (or more) representations of (32) are equivalent. Two $VARMA$ representations are equivalent if $P(L)^{-1}Q(L)$ results in the same operator $\Psi(L)$ (Dufour and Pelletier, 2008). We need the *minimal* condition to avoid that the elements of $P(L)$ and $Q(L)$ cancel out when we take $P(L)^{-1}Q(L)$ in (32).

5.2 $\boldsymbol{AGDCC^{\diamond}}$ Model

So far we have proved consistency of parameter estimates involved in the first-stage process $AGDCC^{*}$, involving the conditional variances. This is an important step given that parameters estimated in the second-stage are inconsistent if the residuals in the first-stage are standardised by inconsistent conditional variances. In this section, we proceed with the analysis of consistency of parameters estimated in the second-step estimation. We first derive the conditions for the strict stationarity of the AGDCC$^{\diamond}$ process. Then, we introduce the concept of *top-Lyapunov exponent* and how it helps to define the existence of a strict-stationarity solution for the AGDCC$^{\diamond}$ process. In addition, we generalise this result for the models nested in the AGDCC$^{\diamond}$ process, namely the GDCC$^{\diamond}$ (p, q), ADCC$^{\diamond}$ (p, q, s), and DCC$^{\diamond}(p, q)$ processes reported in Appendix A. Further, we modify Lemma 5 to account for the new dynamics introduced. Finally, a theorem provides consistency for the entire process.

First, we verify the main conditions required for consistency, i.e. strict-stationarity and identification. Equations (1)-(3) for the case of the AGDCC$^{\diamond}$ process can be written as follows

$$r_t = g(D_t, R_t, \varepsilon_t) \tag{33}$$

$$R_t = \eta(Q_t) \tag{34}$$

where g and η are measurable functions.

5.2.1 Stationarity

In order to prove the strict stationarity of AGDCC$^{\lozenge}$, we make use of that for any process $Z_t = f(\Gamma_t)$, where $f(\cdot)$ is a measure preserving function, if the process Γ_t is strictly stationary then Z_t is also strictly stationary. This entails that we need to prove the strict-stationarity of equation (6) in order to obtain the strict stationary of the AGDCC$^{\lozenge}$ process.

Equation (6) can be written as

$$U_t = FU_{t-1} + C_t, \qquad t = 1, ..., T \tag{35}$$

where

$$U_t = \begin{bmatrix} Q_t \\ . \\ . \\ . \\ Q_{t-p-1} \end{bmatrix} \tag{36}$$

$$F = \begin{bmatrix} B_1B_1' & B_2B_2' & . & . & . & B_{p-1}B_{p-1}' & B_pB_p' \\ [I_d] & [0] & . & . & . & [0] & [0] \\ [0] & [I_d] & . & . & . & [0] & [0] \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ [0] & [0] & . & . & . & [I_d] & [0] \end{bmatrix} \tag{37}$$

$$C_t = \begin{bmatrix} \Big\{ \overline{Q} - \sum_{i=1}^q A_iA_i' \circ \overline{Q} - \sum_{j=1}^p B_jB_j' \circ \overline{Q} - \sum_{k=1}^r G_kG_k' \circ \overline{N} \\ + \sum_{i=1}^q A_iA_i' \circ \varepsilon_{t-i}\varepsilon_{t-i}' + \sum_{k=1}^r G_kG_k' \circ \eta_{t-k}\eta_{t-k}' \Big\} \\ 0 \\ . \\ . \\ 0 \end{bmatrix} \tag{38}$$

where the sup-products $(B_iB_i' \circ Q_{t-1})$, $i = 1, ..., p$ in FU_{t-1} are Hadamard. Note that $U_t \in \mathbb{R}^n, B \in \mathbb{R}^{n \times n}$, and that $C_t \in \mathbb{R}^{n \times n}$. This is a valid representation in the sense that a stationary solution is independent of the future at any given time.

The necessary and sufficient conditions for strict stationarity of generalised autoregressive processes of the form $U_t = F_tU_{t-1} + C_t$, were derived by Bougerol and Picard (1992). They define the strict stationarity of stochastic recurrence equations in terms of a metric measure of the autoregressive component F_t called the Lyapunov exponent: for the process $U_t = F_t U_{t-1} + C_t$, the top Lyapunov exponent, when $E\left(\log^+ \|F_1\|\right) < \infty$, is defined by

$$\gamma = \inf(n^{-1} E[\log \|F_1 ... F_t\|], n \in N) \tag{39}$$

where $\|\cdot\|$ is the operator norm on $\mathbb{R}^n$. In our case the process for U_t is more simple (the coefficient F is a constant) and the top Lyapunov exponent can be defined as $\gamma = \log \rho(F)$ where $\rho(M) = \max_{1 \leq i \leq n} |\lambda_i|$ is the spectral radius of the matrix $M \in \mathbb{R}^{n \times n}$ and $\lambda_1, ..., \lambda_n$ are the eigenvalues of M.

Bougerol and Picard (1992) provide the strictly stationary solution of stochastic recurrence equations in terms of their top Lyapunov exponent. For the convenience of the reader, we formulate the argument in Lemma 6 below:

Lemma 6 *Let us suppose that the stochastic recurrence equation $U_t = F_t U_{t-1} + C_t$, $t = 1, ..., T$, with an F_t i.i.d. coefficient is irreducible and that $E\left(\log^+ \|F_1\|\right) < \infty$ and $E[\log^+ \|C_1\|] < \infty$. Then U_t has a non-anticipative strictly stationary solution if and only if the top Lyapunov exponent γ is strictly negative.*

Proof. See Bougerol and Picard (1992). ■

The process considered by Bougerol and Picard (1992) is more general than the process in (35), where the coefficient matrix F is constant, rather than F_t as an i.i.d. sequence. For a F constant matrix of a specific form, we have the following corollary to Lemma 6

Corollary 1 *Consider the autoregressive process $U_t = FU_{t-1} + C_t, \quad t = 1, ..., T$ given by*

$$\begin{bmatrix} u_t \\ . \\ . \\ . \\ u_{t-p-1} \end{bmatrix} = \begin{bmatrix} f_1 & f_2 & . & . & . & f_{p-1} & f_p \\ 1 & 0 & . & . & . & 0 & 0 \\ 0 & 1 & . & . & . & 0 & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & 1 & 0 \end{bmatrix} \begin{bmatrix} u_{t-1} \\ . \\ . \\ . \\ u_{t-p} \end{bmatrix} + \begin{bmatrix} c_t \\ 0 \\ . \\ . \\ 0 \end{bmatrix} \tag{40}$$

If $E[\log^+ \|C_1\|] < \infty$, $E(U_t^2) < \infty$ and U_t has a non-anticipative weakly stationary solution, then this solution is also strictly stationary.

Proof. If U_t has a weakly stationary solution then $f_i > 0, \forall i, i = 1, ..., p$, $\sum_{i=1}^{p} f_i < 1$ and the spectral radius of F is $\rho(F) < 1$. For a constant matrix F, we know that $\gamma = \log \rho(F)$ and therefore $\gamma < 0$. If we add the condition $E(U_t^2) < \infty$ then the result follows. ■

Corollary 1 shows that as in the case of the process AGDCC* covariance stationarity is enough to secure strict stationarity. For the AGDCC$^\diamondsuit$ process the connection relies on the behaviour of the Lyapunov exponent associated to the process; when the roots in $1-\sum_{i=1}^{p} f_i$ are outside the unit circle, the Lyapunov exponent is strictly negative.

We now use the results of Corollary 1 to provide the conditions for the strict stationary of the AGDCC$^\diamondsuit$ model.

Lemma 7 *Consider the matrix* $\Pi = I_n - \sum_{i=1}^{q} A_i A_i' - \sum_{j=1}^{p} B_j B_j' - \sum_{k=1}^{r} G_k G_k'$. *The process* r_t *that follows the AGDCC*$^\diamondsuit(p, q, s)$ *model is strictly stationary if* Π *is a positive-definite matrix.*

Proof. If Π is a positive-definite matrix then the AGDCC$^\diamondsuit(p, q, s)$ process is weak stationary. The matrix F in (37) is constant and of the form (40), and also $E(|Q_t|^2) < \infty$ almost surely because $E\left(|\varepsilon_t|^2\right) < \infty$ and $E(|\eta_t|^2) < \infty$. From Corollary 1, Q_t is strictly stationary and thus r_t is strictly stationary. ■

This result provides the conditions for the strict stationarity of the AGDCC$^\diamondsuit$ process, the most general form of the DCC models. In Lemma 10 in Appendix A, we report the extension of Lemma 7 to the nested DCC, ADCC, and GDCC models.

We turn now to the last step of our analysis to prove the identification of the parameter estimates present in the AGDCC$^\diamondsuit$ process.

5.2.2 Identification

Let us rewrite (6) as

$$\begin{aligned} P(L) \circ Q_t &= \overline{Q} - \sum_{i=1}^{q} A_i A_i' \circ \overline{Q} - \sum_{j=1}^{p} B_j B_j' \circ \overline{Q} \\ &\quad - \sum_{k=1}^{s} G_k G_k' \circ \overline{N} + Q(L) \circ \varepsilon_t \varepsilon_t' + S(L) \circ \eta_t \eta_t' \end{aligned} \tag{41}$$

where $S(L) = \sum_{k=1}^{s} G_k G_k' L^k$. We have now three matrices with polynomial coefficients, and the equivalence of representations explained in the previous section can take place in the two pairs $P(L)^{-1}Q(L)$ and $P(L)^{-1}S(L)$.

In order to accommodate the new matrix S in (41), we extend the definition of *minimal specification* presented in Lemma 5 as follows:

Lemma 8 *The multivariate GARCH (p,q) specification given in (41) is minimal if*

1. $P(0) = I_d$, $Q(0) = 0$ and $S(0) = 0$
2. $det(P) \neq 0$, $det(Q) \neq 0$ and $det(S) \neq 0$
3. P and Q are coprime or(and) P and S are coprime
4. $\forall j, 1 \leq j \leq d, d_j(P) = d_j \leq p$, $d_j(Q) = d_j \leq q$, and $d_j(S) = d_j \leq r$.
5. P or Q is column reduced, and P or S is column reduced

Lemma 9 *Let (P_1, Q_1) define a minimal formulation of a multivariate GARCH (p,q) model, such that there exists a weakly stationary solution denoted ε_t; then, if ε_t is also a solution of another model written with (P_2, Q_2), there exists j, such that $d_j(P_2) > d_j(P_1)$ or $d_j(Q_2) > d_j(Q_1)$.*

Proof. See Jeantheau (1998, p.79) ■

Lemma 9 (Proposition 3.3 in Jeantheau, 1998) states that if two processes have the same stationary solution, then the supremes of the degrees of the matrices with polynomial coefficients P_1 and P_2 or Q_1 and Q_2 cannot be both equals. Lemma 9 is useful to prove the validity Theorem 3 below, which establishes consistency of the $AGDCC^{\diamond}$ model estimated by AML-MLE.

Theorem 3 *Under the assumption that the formulation at θ_0 for the model (6) is minimal, the AML-MLE for the AGDCC model defined by (6) is consistent, $\hat{\theta} \xrightarrow{p} \theta_0$.*

Proof. In Lemma 7, we proved the strict stationarity of the process, satisfying Assumption 2. If the process is strictly stationary and all parameters are greater than zero then

$\det(Q_t) \geq c$ where $c > 0$, satisfying Assumption 3. For Assumption 4, we can apply the same argument as in Proof of Theorem 2: defining $Q_{i,t}$ as the diagonal element of Q_t, covariance stationary implies $E_0(Q_{i,t}) < \infty$, Jensen's inequality yields $E_0(\log(Q_{i,t})) < +\infty$, and given that $\det(Q_t) \geq c$ we get $E_0(|\log(Q_{i,t})|) < \infty$. To verify Assumption 5, consider first that for

$$Q_{t,0} = Q_t\left(\theta_0\right)$$

we have that (6) is

$$\begin{aligned} Q_{t,0} &= \overline{Q} - \sum_{i=1}^{q} A_{i,0}A_{i,0}' \circ \overline{Q} - \sum_{j=1}^{p} B_{j,0}B_{j,0}' \circ \overline{Q} - \sum_{k=1}^{s} G_{k,0}G_{k,0}' \circ \overline{N} \\ &\quad + \sum_{i=1}^{q} A_{i,0}A_{i,0}' \circ \varepsilon_{t-i}\varepsilon_{t-i}' + \sum_{j=1}^{p} B_{j,0}B_{j,0}' \circ Q_{t-j,0} + \sum_{k=1}^{s} G_{k,0}G_{k,0}' \circ \eta_{t-k}\eta_{t-k}' \end{aligned} \tag{42}$$

If $Q_t = Q_{t,0}$ then

$$\begin{aligned} 0 &= \sum_{i=1}^{q} M_i \circ \varepsilon_{t-i}\varepsilon_{t-i}' + \sum_{j=1}^{p} M_{q+j} \circ Q_{t-j} + \sum_{k=1}^{s} M_{q+p+k} \circ \eta_{t-k}\eta_{t-k}' \\ &\quad - \sum_{i=1}^{q} M_i \circ \overline{Q} - \sum_{j=1}^{p} M_{q+j} \circ \overline{Q} - \sum_{k=1}^{s} M_{q+p+k} \circ \overline{N} \end{aligned} \tag{43}$$

where $M_i = A_{i,0}A_{i,0}' - A_iA_i'$, $M_{q+j} = B_{j,0}B_{j,0}' - B_jB_j'$, and $M_{q+p+k} = G_{k,0}G_{k,0}' - G_kG_k'$. We must prove that all terms are equal to 0. First, (43) yields

$$M_1 \circ \varepsilon_{t-1}\varepsilon_{t-1}' = U \tag{44}$$

where U is a $\mathcal{F}_{t-2}$-measurable matrix. Lemma 3.1 in Jeantheau (1998, page 78) implies that both M_1 and U are equal to 0. From $M_1 = 0$, we have

$$\begin{aligned} M_{q+1} \circ Q_{t-j} &= -\sum_{i=1}^{q} M_i \circ \varepsilon_{t-i}\varepsilon_{t-i}' - \sum_{k=1}^{s} M_{q+p+k} \circ \eta_{t-k}\eta_{t-k}' \\ &\quad + \sum_{i=1}^{q} M_i \circ \overline{Q} + \sum_{j=1}^{p} M_{q+j} \circ \overline{Q} + \sum_{k=1}^{s} M_{q+p+k} \circ \overline{N}. \end{aligned} \tag{45}$$

Now, when M_{q+1} is different from zero we have that if P is column reduced then $M_{q+1}\det(\beta_{d_j}) \neq 0$ because of Definition **(a)**. From Lemma 9, we have that the left term of (45) must have at least one column j with $d_j(P)$ lags, but this contradicts (45) as the right term (45) which has only $d_j(P) - 1$ lags. Therefore, it must be $M_{q+1} = 0$. The same same demonstration holds if Q is column reduced. The proof is complete by iterating the same demonstration for $M_2, M_{q+2}, M_3, ..., M_{q+p}, M_{q+p+1}, ..., M_{q+p+r}$ and show that all terms are equal to zero. ■

Note that Theorem 3 provides weak consistent results which is good enough for our purposes. However, under mildly stronger assumptions strong consistency can also be proved (Pfanzgal, 1969).

6 Empirical Application

The application reported in this section is intended to provide evidence about the appropriateness of the use of the AML-DCC type specifications compared to the normal-DCC models. We focus on specification tests for the distribution of standardised residuals and on the features of parameter estimates.

We consider shares indices of 21 countries listed in the FTSE All-World Indices and bond indices of 13 countries constructed by Datastream. We refer the interested reader to Cappiello et al. (2006) for a detailed description of the data[3]. The frequency is weekly and spans over the period 08/01/1987-07/02/2002 (785 observations). The 21 countries of the share indices are: Australia, Austria, Belgium, Canada, Denmark, France, Germany, Hong Kong, Ireland, Italy, Japan, Mexico, Netherlands, New Zealand, Norway, Singapore, Spain, Sweden, Switzerland, United Kingdom, and the United States. The 13 countries of the bond indices are Austria, Belgium, Canada, Denmark, France, Germany, Ireland, Japan, Netherlands, Sweden, Switzerland, and the United Kingdom. Weekly returns for bonds and shares were calculated through log differences using Friday to Friday closing prices

$$r_{it} = \log\left(\frac{P_{it}}{P_{it-1}}\right), \quad i = 1, ...n \tag{46}$$

where P_{it} is the price of assets i at time t. We estimate the four models described above: AGDDC (1,1,1), GDCC(1,1), ADCC(1,1,1), and the DCC(1,1). Table 1 reports the parameter estimates of the joint GJR-GARCH (1,1) processes for the univariate volatilities which

[3] We wish to thank Kevin Sheppard for providing us with the dataset.

were calculated to include a mean intercept. The asymmetry parameter was only significant in 4 out of the 21 equity indices, and only 1 of the 13 bond indices, while the average persistence[4] was 0.95 for both groups. In the case of Japan Fixed Income, it would appear that a GARCH model might not be appropriate, which a visual inspection of the autocorrelation plot of the square and absolute residuals confirmed.

[**Insert Table 1 here**]

To evaluate the multivariate distributions we implemented the visual diagnostic proposed in Kawakatsu (2006). The idea is based on the fact that if $r_t \sim N(0, H_t)$, then $r_t H_t^{-1} r_t'$ has a $\chi^2(n)$ distribution. Although we do not know the distribution of $r_t H_t^{-1} r_t'$ when $r_t \sim AML(m, H_t)$, we generated an empirical distribution in order to perform the comparison. We found that when comparing the quantile-quantile plots of the AML and Normal models against the empirical distribution, the former provided for a better fit. To reinforce our findings about the inconvenience of the assumption of multivariate normality, we also performed the omnibus test of Doornik and Hansen (2008). Multivariate normality was overwhelmingly rejected for the raw and standardized data after fitting the normal DCC, ADCC, GDCC, and AGDCC models. All $p-$values are ≈ 0, and thus we do not report them in the paper, but they are available upon request. Before estimating the models for the conditional correlation, we evaluated the constancy of correlation performing the LM test of Tse (2000). We overwhelmingly reject the null of constant correlation with a p-value $= 0.000$, and in this case too we do not report the results, available upon request. Estimates for the AGDCC models for the multivariate normal and Laplace distributions are reported in Tables 2 - 3, respectively[5].

[**Insert Tables 2-3 here**]

Panel A of Table 4 reports the results by log-likelihood and the BIC and AIC information criteria, respectively.

[4] In the GJR-Normal model persistence is equal to $\alpha + \beta + 0.5 * \gamma$.

[5] Estimates for the DCC, ADCC, GDCC are not reported but available upon request.

[Insert Table 4 here]

It is clear that the increase in likelihood from the scalar DCC to the AGDCC under both distributions is significant at the 1% level indicating that the more generalized model with asymmetric effects favors the data, while the increased likelihood across comparable models under the AML distribution is significant at the 1% level suggesting a better fit to the data once tail dependence is taken into account.

In Table 5, we report estimates of in-sample risk management measures such Value-at-Risk (VaR), Tail Risk (Wong, 2010, TR), and Expected Shortfall (ES). Though we have considered alternative portfolios, in this paper we only report results for a portfolio composed of 75% stocks and 25% bonds. We then proceed to evaluate the results by means of a backtesting analysis.We compare the performance from the use of both Normal and AML distributions in the estimation process of the variance-covariance matrices. We estimate two conditional variance matrices H_t using an AGDCC (1,1,1) model with first stage normal and second stage normal and a second AGDCC (1,1,1) model with first stage normal and second stage AML. For the univariate volatilities specification we selected a GJR-GARCH(1,1) specification.

In the estimation of the risk quantiles, we employed the normal and asymmetric Laplace. The backtesting results of our exercise for the risk measures are reported in Table 5. Note that all backtesting analyses are based on observations from 201 to 785 in order to be able to apply the Filtered Historical Simulation (FHS), as proposed by Barone-Adesi et al. (1999) (see also Christoffersen 2009), we needed to exclude the first 200 weekly log-returns. Therefore, to allow a fair comparison between alternative methodologies, we carried out all the tests on the same sample. The unconditional (UC) test supports the AGDCC-AML model at any quantile. In terms of clustering of VaR violations (IND), there is evidence that the best option is again the AGDCC-AML at both 5% and 1% quantiles. When we consider the conditional (CC) test, where the number and clustering of VaR failures are jointly considered in a single test, the result is again in favour of the AGDCC-AML model for every quantile. Table 5 also reports the backtesting results for the TR measure. In this case too AGDCC-MVN is dominated by the AGDCC-AML models. Since TR is a parameteric test, the FHS cannot be used. Finally, in the last rows of Table 5, the results for the ES backtest are reported. There is mixed evidence, with AGDCC-AML model at 1% while AGDCC-MVN is chosen at 5% quantile.

[Insert Table 5 here]

Figure 1 reports the graphs of a geographical aggregation in four groups (Europe, Europe non EMU, Americas, and Australasia) of time varying correlations from the AGDCC-ML model, showing the common spikes around the 1987 and 1998 crashes. Finally, Figures 2-4 provide a visual inspection of the time varying correlations in the equity and fixed income across selected countries of the four groups.

[Insert Figures 1-4 here]

7 Conclusions

In this paper we proposed a multivariate asymmetric generalised dynamic conditional correlation GARCH model, where returns are assumed to follow the asymmetric multivariate Laplace distribution. This multivariate distribution is able to capture leptokurtosis and asymmetry, it has desirable properties such as additivity and finiteness of moments, and a density function with a closed-form that makes the maximum likelihood estimation method easy to implement. We proved that maximum likelihood estimation provides consistent estimate of parameters of AGDCC-AML models. We also proved the strict stationarity of the AGDCC model and its nested versions. The empirical validity of the AGDCC-AML model is tested by fitting 21 FTSE All-World stock indices and 13 bond return indices of Cappiello et al. (2006). We also report estimates of in-sample risk management measures (VaR, TR, and ES) for a portfolio composed of 75% stocks and 25% bonds. We evaluated the results by means of a backtesting analysis, by comparing the performance from the use of both normal and AML distributions in the estimation process of the variance-covariance matrices.

Overall, we provide clear evidence that the AML distribution overwhelmingly outperforms the case in which we assume normality of innovations.

There are several extensions to the results in this paper worth considering. The empirical validity of AGDCC-AML model may be tested using data at higher frequency. Further, to estimate the risk quantiles, alternative (to normal and AML) distributions, such as the Student-t, the generalised Pareto, the asymmetric power (Komunjer, 2007), can also be employed. In addition, it would interesting to explore whether the results in this paper may be extended to the broader class of GH distributions. Another important issue is to explore the ex-ante forecasting advantage in using a dynamic model instead of unconditional measures, which call for the need of simulation based approaches. These interesting developments are beyond the scope of the present paper, but they are the object of ongoing research.

Acknowledgment

We wish to thank participants in the international conference "*Common Features in London*" (Cass Business School, London, 16-17 December 2004), in particular Robert Engle, Neil Shephard, Kevin Sheppard; the "*1st Italian Congress of Econometrics and Empirical Economics*" (Universita' Ca' Foscari, Venice, 24-25 January 2005); the "*Simulation Based and Finite Sample Inference in Finance II*" Conference (Québec City, 29-30 April 2005), in particular our discussant Tongshu Ma; the "*ESRC Econometric Study Group Conference*" (Bristol, 14-16 July, 2005); the CEA@Cass/ESRC seminar on "*The Use of Copulas and DCC Models to Measure Dependence in Finance*" (Cass Business School, 17 March 2006), in particular Valentina Corradi, Andrew Patton, and Fabio Trojani; the Finance Seminar (Tilburg, 20 March 2006) in particular Ralph S.J. Koijen; the Bag Lunch Seminar (Finance Department, Kellogg Management School), in particular Torben Andersen, Ravi Jagannathan, Ernst Schaumburg; the CEA@Cass/ESRC seminar on "*Using Thick Distributions, DCC and Copula Models*" (Cass Business School, London,19 May 2006), in particular Enrique Sentana and Javier Mencia; the "*4th OxMetrics User Conference*" (Cass Business School, London, 15-16 September 2006), in particular Andrew Harvey, David Hendry and Sebastien Laurent; the "*2007 North American Summer Meeting of the Econometric Society*".(Duke University, USA, 21-24 June 2007); the "*2007 Far Eastern Meeting of the Econometric Society*" (Institute of Economics, Academia Sinica, Taipei, Taiwan, 11-13 July 2007); the "*2009 Summer School on Econometrics and Finance*" (Wang Yanan Institute for Studies in Economics, Xiamen University, China, 7-11 June 2009) for useful discussions and comments which helped to progress with this paper. Thanks also to Annamaria Bianchi, Simona Boffelli, Arturo Leccadito, Denis Pelletier, Elisabetta Pellini, Fabio Trojani and Radu Tunaru for detailed comments on previous versions of the manuscript, and to Chihwa Kao for long discussions on the identification and consistency issue. Kevin Sheppard and Lorenzo Cappiello provided the dataset and useful insights to estimate the various models. Very special thanks to Lorenzo Trapani for having carefully read and checked various versions of the manuscript, providing useful technical insights: without his help this revision could not have been completed. We wish to thank the Editor, Ron Gallant, and four anonymous referees, for very useful comments and suggestions that greatly help to improve the paper. However, the usual disclaimer applies. We wish to acknowledge financial support from Centre for Econometric Analysis at Cass and INQUIRE UK 2004. This article represents the views of the authors and not of INQUIRE UK.

References

Audrino, F., Trojani, F., 2011. A general multivariate threshhold GARCH model with dynamic conditional correlations. Journal of Business & Economic Statistics 29, 138-149

Baillie, R. T., Bollerslev, T., 1989. The message in daily exchange rates: a conditional variance tale. Journal of Business & Economic Statistics 7, 297-305.

Barndorff-Nielsen, O., 1977. Exponentially decreasing distributions for the logarithm of particle size. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences 353, 401-419.

Bauwens, L., Laurent, S., 2005. A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models. Journal of Business and Economic Statistics 23, 346-354.

Bauwens, L., Laurent,S., Rombouts, J., 2006. Multivariate GARCH models: a survey. Journal of Applied Econometrics 21, 79-109.

Barone-Adesi, G., Giannopoulos, K. and Vosper, L. ,1999. VaR without correlations for non-linear portfolios. Journal of Futures Markets 19, 583-602.

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

Bollerslev, T.,1987. A conditionally heteroskedastic time series model for speculative prices and rates of return. Review of Economics and Statistics 69, 542-547.

Bollerslev, T., 1990. Modelling the coherence in short-run nominal exchange rates: a multivariate GARCH. Review of Economics and Statistics 72, 498-505.

Bollerslev, T., 2001. Financial econometrics: past developments and future challenges. Journal of Econometrics 100, 41-51.

Bollerslev, T. (2009), Glossary of ARCH (GARCH), in Volatility and Time Series: Essays in Honour of Robert F. Engle (eds. Tim Bollerslev, Jeffrey R. Russell and Mark Watson), Oxford University Press, Oxford, UK.

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

Bougerol, P., and Picard, N., 1992. Stationarity of GARCH processes and some nonnegative time series. Journal of Econometrics 52, 115-127.

Boussama F. (1998). Ergodicity, Mixing and Estimation in GARCH models. Ph.D. Dissertation, University of Paris 7.

Bowman, F. C. (1958). Introduction to Bessel Functions. Dover Publications, New York.

Cappiello, L., Engle, R., Sheppard, K., 2006. Asymmetric dynamics in the correlations of global equity and bond returns. Jourrnal of Financial Econometrics 25, 537-572.

Christoffersen, P., 2009. Value–at–Risk models, in Handbook of Financial Time Series, (eds. T.G. Anderson, R.A. Davis, J.P. Kreiβ, T. Mikosch), 201-229. Springer-Verlag Berlin

Heidelberg.

Doornik, J. A., Hansen, H., 2008. An omnibus test for univariate and multivariate normality. Oxford Bulletin of Economcis and Statistics 70, 927-939.

Drost, C.D., Klaassen, C.A.J., 1997. Efficient estimation in semiparametric GARCH models. Journal of Econometrics 81, 193-221.

Dufour, J-M, and Pelletier, D., 2008. Practical methods for modelling weak VARMA processes: identification, estimation and specification with a macroeconomic application. Discussion Paper, McGill University (Department of Economics), CIREQ and CIRANO.

Engle, R., 2001. Financial econometrics – A new discipline with new methods. Journal of Econometrics 100, 53-56.

Engle, R., 2002. Dynamic conditional correlation - A simple class of multivariate GARCH models. Journal of Business & Economic Statistics 20, 339- 350.

Engle, R., 2009. Anticipating Correlations. A New Paradigm for Risk Management. Princeton University Press, Princeton and Oxford .

Engle, R., Colacito, R., 2006. Testing and valuing dynamic correlations for asset allocation. Journal of Business & Economic Statistics 2, 238-253.

Engle, R., Gonzalez-Rivera, G., 1991. Semiparametric ARCH models. Journal of Business & Economic Statistics 9, 345-359.

Engle, R., Sheppard, K., 2001. Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. NBER Working Paper No. W8554

Fiorentini, G., Sentana, E., Calzolari, G., 2003. Maximum likelihood estimation and inference in multivariate conditionally heteroskedastic dynamic regression models with Student-t innovations. Journal of Business & Economic Statistics 21, 532-546.

Gonzalez-Rivera, G. , 1997. A note on adaptation in ARCH models. Econometric Reviews 16, 55-68.

Gonzalez-Rivera, G., Drost, C.D., 1999. Efficiency comparisons of maximum- likelihood-based estimators in GARCH models. Journal of Econometrics 93, 93-111.

Hafner, C. M., van Dijk, D., Franses, H. P., 2005. Semi-parametric modelling of correlation dynamics, in Dek Terrell, Thomas B. Fomby (ed.) Econometric Analysis of Financial and Economic Time Series (Advances in Econometrics, Volume 20), Emerald Group Publishing Limited, pp.59-103.

Hafner, C. M., Rombouts, J. V. K., 2007. Semiparametric multivariate volatility models. Econometric Theory 23, 251–280.

Heyde, C. C., Kou, S. G., 2004. On the controversy over tailweight of distributions. Operations Research Letters 32, 399-408.

Johnson, N. L., Kotz, S., 1972. Distributions in Statistics: Continuous Multivariate Distributions, Wiley, New York.

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

Kawakatsu, H., 2006. Matrix exponential GARCH. Journal of Econometrics 134, 95-128.

Khindanova, I., Rachev, S., Schwartz, E., 2001. Stable modeling of Value at Risk. Mathematical and Computer Modeling 34, 1223-1259.

Komunjer, I., 2005. Quasi-maximum likelihood estimation for conditional quantiles. Journal of Econometrics 128, 137-164.

Komunjer, I., 2007. Asymmetric power distribution: theory and applications to risk measurement. Journal of Applied Econometrics 22, 891-921.

Kozubowski, T. J., Podgorski, K. and Rychlik, I.,2010. Multivariate generalized Laplace distributions and related random fields. Preprint 2010:47, Department of Mathematical Sciences, University of Gothenburg.

Kotz, S., Kozubowski, T. J., Podgorski, K., 2003. The Laplace Distribution and Generalizations. A Revisit with Applications to Communications, Economics, Engineering, and Finance, Birkhauser, Berlin.

Kozubowski, T. J., Podgorski, K., 2001. Asymmetric Laplace laws and modeling financial data. Mathematical and Computer Modelling 34, 1003-1021.

Linden, M., 2001. A model for stock return distribution. International Journal of Finance and Economics 6, 159-169.

Long, X., Ullah, A., 2005. Nonparametric and semiparametric multivariate GARCH model. Mimeo, Department of Economics, University of California, Riverside.

Madan, D. B., Seneta, E., 1990. The variance gamma model for share market returns. Journal of Business 63, 511-524.

Madan, D. B., Carr, P. P., Chang, E. C., 1998. The variance gamma process and option pricing. European Finance Review 2, 79-105.

Mencia, F. J., Sentana, E., 2005. Estimation and testing of dynamic models with generalised hyperbolic innovations. CEMFI working paper No. 0411/CEPR Discussion Paper No. 5177.

Mencia, F. J., Sentana, E., 2009. Multivariate location-scale mixtures of normals and mean-variance-skewness portfolio allocation. Journal of Econometrics 153, 105-121.

Mencia, F. J., Sentana, E., 2010. Distributional tests in multivariate dynamic models with normal and student t innovations. CEMFI Working Paper 0804 (revised June 2010). Forthcoming in Review of Economics and Statistics.

Nelson, D.V., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59, 347-370.

Newey, W., McFadden, D., 1994. Large sample estimation and hypothesis testing," in Handbook of Econometrics (Vol. 4), eds. R. Engle and D. McFadden, New York: Elsevier Science, 2113–2245.

Palandri, A., 2009. Sequential conditional correlations: inference and evaluation. Journal of Econometrics 153, 122-132.

Pfanzgal, J., 1969. On the measurability and consistency of minimum contrast estimates. Metrika 14, 249-272.

Pelletier, D., 2006. Regime switching for dynamic correlations. Journal of Econometrics 131, 445-473.

Relton, F. E., 1965. Applied Bessel Functions. Dover Publications, New York.

Silvennoinen, A., Terasvirta, T., 2009a. Multivariate GARCH models, in Handbook of Financial Time Series, (eds. T.G. Anderson, R.A. Davis, J.P. Kreiβ, T. Mikosch), pp.

201-229. Springer-Verlag Berlin Heidelberg.

Silvennoinen, A., Terasvirta, T., 2009b. Modeling multivariate autoregressive conditional heteroskedasticity with the double smooth transition conditional correlation GARCH model. Journal of Financial Econometrics 7, 373-411.

Tse, I. K., 2000. A tests for constant correlations in a multivariate GARCH model. Journal of Econometrics 98, 107-127.

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

Wong, W.K., 2010. Backtesting value-at-risk based on tail losses. Journal of Empirical Finance 17, 526-538.

Appendices

Appendix A (Extension of Lemma 7 to the GDCC, ADCC, DCC models)

Lemma 10 *The GDCC$^{\diamond}$ (p,q), ADCC$^{\diamond}$ (p,q,s), and DCC$^{\diamond}(p,q)$ processes are strictly stationary if they are weakly stationary.*

Proof. For the GDCC$^{\diamond}$ (p,q) process we have that the vector in (38) is substituted by

$$C_t = \begin{bmatrix} \overline{Q} - \sum_{i=1}^{q} A_i A_i' \circ \overline{Q} - \sum_{j=1}^{p} B_j B_j' \circ \overline{Q} + \sum_{i=1}^{s} A_i A_i' \circ \varepsilon_{t-i}\varepsilon_{t-i}' \\ 0 \\ . \\ . \\ 0 \end{bmatrix} \tag{47}$$

The matrix Π is in this case given by $\left(I_d - \sum_{i=1}^{q} A_i A_i' - \sum_{j=1}^{p} B_j B_j'\right)$. As in Lemma 7, $E[\log^+ \|C_1\|] < \infty$ as $E\left(\varepsilon_t\varepsilon_t'\right) < \infty$.The conditions for $E\left(\log^+ \|F\|\right) < \infty$ and $\gamma < 0$ are the same as for the AGDCC$^{\diamond}$ (p,q,s) model. For the ADCC$^{\diamond}$ (p,q,s) process, relations (37) and (38) are

$$F = \begin{bmatrix} b_1 & b_2 & . & . & . & b_{p-1} & b_p \\ 1 & 0 & . & . & . & 0 & 0 \\ 0 & 1 & . & . & . & 0 & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & 1 & 0 \end{bmatrix} \tag{48}$$

$$C_t = \begin{bmatrix} \left\{\left(1 - \sum_{i=1}^{q} a_i - \sum_{j=1}^{p} b_j\right) \overline{Q} - \left(\sum_{k=1}^{s} g_k\right) \overline{N} \right. \\ \left. + \sum_{i=1}^{q} a_i \left(\varepsilon_{t-i}\varepsilon_{t-i}'\right) + \sum_{k=1}^{s} g_k \left(\eta_{t-k}\eta_{t-k}'\right)\right\} \\ 0 \\ . \\ . \\ 0 \end{bmatrix} \tag{49}$$

For the DCC$^{\diamond}$ (p,q) process, F is given as in (48) and the C_t vector is

$$C_t = \begin{bmatrix} \left(1 - \sum_{i=1}^{q} a_i - \sum_{j=1}^{p} b_j\right) \overline{Q} + \sum_{i=1}^{q} a_i \left(\xi_{t-i}\xi_{t-i}'\right) \\ 0 \\ . \\ . \\ 0 \end{bmatrix} \tag{50}$$

In these cases Π is not a matrix but a scalar. For the ADCC$^{\lozenge}$ (p,q,r) process $\Pi = 1 - \sum_{i=1}^{q} a_i - \sum_{j=1}^{p} b_j - \sum_{k=1}^{s} g_k$ and for the DCC$^{\lozenge}$ (p,q) process $\Pi = 1 - \sum_{i=1}^{q} a_i - \sum_{j=1}^{p} b_j$. The conditions for weak stationarity are respectively $\sum_{i=1}^{q} a_i + \sum_{j=1}^{p} b_j + \delta \sum_{k=1}^{s} g_k < 1$ and $\sum_{i=1}^{q} a_i + \sum_{j=1}^{p} b_j < 1$. Again $E[\log^{+} \|C_1\|] < \infty$ as $E\left(\varepsilon_t \varepsilon_t'\right) < \infty$ and $E(\eta_t \eta_t') < \infty$. The result follows from Corollary 1. ■

Appendix B (The AGDCC model with positive definite constraints for the intercept).

The AGDCC model of Cappiello et at (2006) models the multivariate dynamics as

$$Q_t = \Omega + A'\varepsilon_{t-1}\varepsilon_{t-1}'A + B'Q_{t-1}B + G'\eta_{t-1}\eta_{t-1}'G,$$

where Ω is the variance targeting intercept introduced in order to reduce the dimensionality of the problem. The intercept equation is given by $\Omega = (\bar{Q} - A'\bar{Q}A - B'\bar{Q}B - G'\bar{N}G)$ which for the diagonal version simplifies to $\Omega = \bar{Q} \circ (ii' - a'a - b'b) - \bar{N} \circ g'g$. Unlike the scalar case, checking for positive definiteness of Q_t would require checking the eigenvalues of the intercept matrix. In the context of optimization, we could explicitly (direct method) include a nonlinear constraint of positive eigenvalues, though this would make the problem highly nonlinear and would not be able to handle the case of complex values. A popular method is to implicitly constrain the optimization (indirect method) by checking for positive and real eigenvalues in the main optimization and then 'kick' the optimizer when these conditions are not met. This method, related to the exterior penalty function method, has numerous well documented problems, including creating a high degree of non-linearity and possibly non-smoothness and local optima. Additionally, for boundary cases this also creates instability in the evaluation of numerical derivatives (particularly in relation to standard error calculation) as the so called 'kick' creates a discontinuity in the function. In one-stage optimization one would usually make use of the gradient and hessian at the optimal solution returned by the solver. However, for the case of the 2-stage DCC (and variants) we cannot do this and must usually resort to numerical derivatives, perturbing the first stage GARCH parameters to obtain the second stage partitioned standard error matrix for the DCC parameters. As a result, it is not unfathomable that the numerical evaluation of the scores based on first stage results lead to an unstable and singular matrix which cannot be inverted. In this context, we propose an alternative method based on a well known property of positive definite matrices.

Definition A1. *A Matrix M is positive definite if and only if there is a positive definite matrix $B > 0$, with $B^2 = M$.*

The matrix B is called the "square root" of M. This matrix B is unique, but only under the assumption $B > 0$. In terms of the optimization problem, we can include the following constraint to ensure the positive definiteness of the intercept Ω, $B^2 - \Omega = 0$. If Ω has a "square root" then it is positive semi-definite. One therefore models the lower triangular part of B which creates an added $N(N-1)/2+N$ parameters in the optimization problem. Since modelling these additional parameters and also imposing a constraint given these parameters is equivalent to optimizing $N(N-1)/2+N$ parameters, then we might as well just forget about variance targeting and model the intercept directly, as CC' with diagonal elements of C restricted to be positive in order to ensure uniqueness (Billio et at, 2006), hence requiring $N(N-1)/2+N$ parameters without any additional constraints.

Table 1: Parameter estimates and p-values from the first stage univariate GARCH.

	Equity				Fixed Income			
	ω	α	γ	β	ω	α	γ	β
Australia	0.00003	0.0831	0.0188	**0.8639**				
	[0.626]	*[0.207]*	*[0.802]*	*[0.000]*				
Austria	0.00003	**0.1098**	0.0416	**0.8286**	0.00002	**0.0630**	0.0000	**0.8700**
	[0.156]	*[0.017]*	*[0.431]*	*[0.000]*	*[0.083]*	*[0.030]*	*[0.999]*	*[0.000]*
Belgium	0.00001	**0.0629**	0.0046	**0.9200**	0.00001	**0.0509**	0.0123	**0.8955**
	[0.283]	*[0.008]*	*[0.894]*	*[0.000]*	*[0.157]*	*[0.015]*	*[0.559]*	*[0.000]*
Canada	0.00004	0.0746	0.1858	**0.7703**	0.00001	0.0349	0.0811	**0.8521**
	[0.188]	*[0.080]*	*[0.253]*	*[0.000]*	*[0.098]*	*[0.285]*	*[0.067]*	*[0.000]*
Denmark	0.00004	0.1018	0.0248	**0.8289**	0.00001	**0.0405**	0.0340	**0.9126**
	[0.297]	*[0.016]*	*[0.716]*	*[0.000]*	*[0.194]*	*[0.028]*	*[0.178]*	*[0.000]*
France	0.00003	0.0492	0.0491	**0.8850**	0.00001	**0.0554**	0.0263	**0.8788**
	[0.091]	*[0.088]*	*[0.240]*	*[0.000]*	*[0.111]*	*[0.038]*	*[0.329]*	*[0.000]*
Germany	0.00000	**0.0633**	0.0000	**0.9341**	0.00001	**0.0558**	0.0000	**0.8857**
	[0.551]	*[0.003]*	*[1.000]*	*[0.000]*	*[0.090]*	*[0.038]*	*[1.000]*	*[0.000]*
HK	**0.00029**	0.0810	0.3068	**0.5965**				
	[0.049]	*[0.063]*	*[0.099]*	*[0.000]*				
Ireland	0.00002	0.0856	0.0000	**0.9040**	0.00000	**0.0435**	0.0039	**0.9448**
	[0.471]	*[0.270]*	*[1.000]*	*[0.000]*	*[0.182]*	*[0.005]*	*[0.816]*	*[0.000]*
Italy	**0.00012**	**0.0895**	0.0092	**0.7982**				
	[0.047]	*[0.031]*	*[0.796]*	*[0.000]*				
Japan	0.00003	0.0359	**0.1292**	**0.8779**	0.00003	0.1068	0.0104	0.7906
	[0.286]	*[0.113]*	*[0.003]*	*[0.000]*	*[0.182]*	*[0.832]*	*[0.997]*	*[0.208]*
Mexico	0.00029	0.0860	0.1660	**0.7287**				
	[0.139]	*[0.070]*	*[0.204]*	*[0.000]*				
Netherlands	**0.00006**	**0.1199**	0.1316	**0.7169**	0.00002	0.0563	0.0000	**0.8671**
	[0.025]	*[0.045]*	*[0.319]*	*[0.000]*	*[0.317]*	*[0.062]*	*[1.000]*	*[0.000]*
NZ	0.00005	0.0768	0.0000	**0.8800**				
	[0.539]	*[0.345]*	*[1.000]*	*[0.000]*				
Norway	0.00004	0.1028	0.0000	**0.8628**				
	[0.714]	*[0.370]*	*[1.000]*	*[0.000]*				
Singapore	**0.00012**	**0.0690**	**0.1865**	**0.7668**				
	[0.000]	*[0.046]*	*[0.003]*	*[0.000]*				
Spain	0.00008	**0.1196**	0.0646	**0.7780**				
	[0.402]	*[0.033]*	*[0.580]*	*[0.000]*				
Sweden	0.00006	0.0755	0.0770	**0.8365**	0.00000	0.0249	0.0359	**0.9408**
	[0.065]	*[0.055]*	*[0.199]*	*[0.000]*	*[0.335]*	*[0.326]*	*[0.140]*	*[0.000]*
Switzerland	0.00001	0.0066	0.0000	**0.9819**	0.00000	0.0000	0.0000	**0.9999**
	[0.376]	*[0.738]*	*[1.000]*	*[0.000]*	*[0.951]*	*[1.000]*	*[1.000]*	*[0.000]*
UK	0.00000	0.0392	**0.0000**	**0.9533**	0.00000	**0.0222**	0.0000	**0.9744**
	[0.076]	*[1.000]*	*[0.036]*	*[0.000]*	*[0.353]*	*[0.006]*	*[1.000]*	*[0.000]*
USA	0.00003	0.0000	0.2427	**0.8110**	0.00000	0.0000	0.0539	**0.9299**
	[0.430]	*[0.125]*	*[1.000]*	*[0.000]*	*[0.457]*	*[1.000]*	*[0.189]*	*[0.000]*

Notes to Table 1: Values in bold denote 5% significance level.
The GJR-GARCH model: $\sigma^2_{t-1} = \omega + \beta\sigma^2_{t-1} + \alpha\varepsilon^2_{t-1} + \gamma\varepsilon^2_{t-1}I_{t-1}$, where $I_{t-1} = \begin{cases} 0, \varepsilon_{t-1} \geqslant 0 \\ 1, \text{otherwise} \end{cases}$

Table 2: Parameter estimates and p-values from the Asymmetric Generalized DCC (AGDCC) Model with Multivariate Normal Distribution.

	Equity			Fixed Income		
	a	**b**	**g**	**a**	**b**	**g**
Australia	0.0175	**0.9684**	**- 0.0495**			
	[0.662]	*[0.000]*	*[0.015]*			
Austria	0.0296	**0.9670**	- 0.0882	0.0332	**0.9786**	- 0.0189
	[0.253]	*[0.000]*	*[0.204]*	*[0.168]*	*[0.000]*	*[0.119]*
Belgium	**0.0659**	**0.9576**	**- 0.2016**	0.0527	**0.9767**	- 0.0336
	[0.030]	*[0.000]*	*[0.001]*	*[0.190]*	*[0.000]*	*[0.099]*
Canada	- 0.0108	**0.9712**	**- 0.2229**	0.0528	**0.9430**	**- 0.1273**
	[0.768]	*[0.000]*	*[0.000]*	*[0.113]*	*[0.000]*	*[0.033]*
Denmark	0.0318	**0.9719**	**- 0.0961**	0.1025	**0.9773**	- 0.0674
	[0.488]	*[0.000]*	*[0.041]*	*[0.077]*	*[0.000]*	*[0.091]*
France	0.0804	**0.9529**	**- 0.1779**	0.0854	**0.9768**	**- 0.0541**
	[0.066]	*[0.000]*	*[0.000]*	*[0.115]*	*[0.000]*	*[0.039]*
Germany	0.0122	**0.9663**	**- 0.0483**	0.0685	**0.9779**	- 0.0421
	[0.070]	*[0.000]*	*[0.000]*	*[0.205]*	*[0.000]*	*[0.123]*
HK	- 0.0082	**0.9889**	- 0.0150			
	[0.551]	*[0.000]*	*[0.206]*			
Ireland	**0.0796**	**0.9550**	- 0.0517	0.0615	**0.9747**	- 0.0313
	[0.010]	*[0.000]*	*[0.580]*	*[0.075]*	*[0.000]*	*[0.140]*
Italy	0.0758	**0.9720**	**- 0.2416**			
	[0.262]	*[0.000]*	*[0.000]*			
Japan	0.0496	**0.9871**	- 0.0688	0.1117	**0.9959**	- 0.0640
	[0.400]	*[0.000]*	*[0.193]*	*[0.201]*	*[0.000]*	*[0.135]*
Mexico	- 0.0005	**0.9709**	- 0.1226			
	[0.991]	*[0.000]*	*[0.197]*			
Netherlands	**0.0209**	**0.9676**	**- 0.0537**	0.0574	**0.9771**	- 0.0337
	[0.043]	*[0.000]*	*[0.000]*	*[0.150]*	*[0.000]*	*[0.257]*
NZ	- 0.0004	**0.9861**	- 0.0261			
	[0.991]	*[0.000]*	*[0.104]*			
Norway	0.0219	**0.9643**	- 0.0596			
	[0.714]	*[0.000]*	*[0.317]*			
Singapore	- 0.0111	**0.9940**	- 0.1274			
	[0.884]	*[0.000]*	*[0.093]*			
Spain	**0.1315**	**0.9430**	**- 0.2382**			
	[0.013]	*[0.000]*	*[0.000]*			
Sweden	0.0085	**0.9539**	**- 0.0995**	**0.0543**	**0.9778**	**- 0.0818**
	[0.685]	*[0.000]*	*[0.000]*	*[0.000]*	*[0.000]*	*[0.000]*
Switzerland	**0.0162**	**0.9493**	**- 0.0539**	0.0588	**0.9803**	**- 0.0301**
	[0.032]	*[0.000]*	*[0.000]*	*[0.137]*	*[0.000]*	*[0.035]*
UK	**0.1017**	**0.9468**	- 0.1124	**0.0508**	**0.9765**	- 0.0100
	[0.002]	*[0.000]*	*[0.092]*	*[0.000]*	*[0.000]*	*[0.788]*
USA	0.0117	**0.9625**	**- 0.1553**	0.0433	**0.9761**	- 0.1229
	[0.539]	*[0.000]*	*[0.000]*	*[0.208]*	*[0.000]*	*[0.264]*

Notes to Table 2: Values in bold denote 5% significance level.
The AGDCC Model: $Q_t = \Omega + aa' \circ \varepsilon_{t-1}\varepsilon'_{t-1} + gg' \circ \eta_{t-1}\eta'_{t-1} + bb' \circ Q_{t-1}$, where $\circ$ denotes the Hadamard operator.

Table 3: Parameter estimates and p-values from the Asymmetric Generalized DCC (AGDCC) Model with Asymmetric Multivariate Laplace Distribution.

	Equity			**Fixed Income**		
	a	**b**	**g**	**a**	**b**	**g**
Australia	0.0257	**0.9701**	- 0.0502			
	[0.608]	*[0.000]*	*[0.383]*			
Austria	0.0314	**0.9723**	- 0.0822	0.0325	**0.9801**	- 0.0172
	[0.445]	*[0.000]*	*[0.406]*	*[0.161]*	*[0.000]*	*[0.176]*
Belgium	0.0636	**0.9605**	**- 0.1963**	0.0528	**0.9790**	- 0.0327
	[0.290]	*[0.000]*	*[0.005]*	*[0.181]*	*[0.000]*	*[0.245]*
Canada	- 0.0043	**0.9676**	**- 0.2244**	0.0394	**0.9471**	- 0.1280
	[0.905]	*[0.000]*	*[0.000]*	*[0.479]*	*[0.000]*	*[0.184]*
Denmark	0.0339	**0.9766**	- 0.0886	0.0818	**0.9800**	- 0.0541
	[0.499]	*[0.000]*	*[0.325]*	*[0.100]*	*[0.000]*	*[0.151]*
France	**0.0836**	**0.9561**	**- 0.1651**	0.0836	**0.9788**	- 0.0523
	[0.048]	*[0.000]*	*[0.000]*	*[0.112]*	*[0.000]*	*[0.121]*
Germany	0.0126	**0.9698**	**- 0.0489**	0.0689	**0.9805**	- 0.0428
	[0.172]	*[0.000]*	*[0.002]*	*[0.200]*	*[0.000]*	*[0.097]*
HK	- 0.0071	**0.9844**	- 0.0193			
	[0.677]	*[0.000]*	*[0.530]*			
Ireland	**0.0738**	**0.9535**	- 0.0588	0.0588	**0.9779**	- 0.0308
	[0.013]	*[0.000]*	*[0.668]*	*[0.143]*	*[0.000]*	*[0.235]*
Italy	0.0778	**0.9698**	**- 0.2385**			
	[0.188]	*[0.000]*	*[0.003]*			
Japan	0.0499	**0.9889**	- 0.0714	0.1072	**0.9965**	- 0.0644
	[0.394]	*[0.000]*	*[0.456]*	*[0.304]*	*[0.000]*	*[0.066]*
Mexico	- 0.0035	**0.9698**	- 0.1185			
	[0.926]	*[0.000]*	*[0.054]*			
Netherlands	**0.0269**	**0.9674**	**- 0.0517**	0.0573	**0.9796**	- 0.0344
	[0.001]	*[0.000]*	*[0.036]*	*[0.127]*	*[0.000]*	*[0.192]*
NZ	0.0070	**0.9883**	- 0.0356			
	[0.832]	*[0.000]*	*[0.313]*			
Norway	0.0251	**0.9646**	- 0.0594			
	[0.475]	*[0.000]*	*[0.527]*			
Singapore	- 0.0080	**0.9950**	- 0.1298			
	[0.919]	*[0.000]*	*[0.371]*			
Spain	0.1279	**0.9345**	**- 0.2401**			
	[0.060]	*[0.000]*	*[0.026]*			
Sweden	0.0116	**0.9494**	**- 0.1121**	**0.0564**	**0.9794**	**- 0.0815**
	[0.667]	*[0.000]*	*[0.000]*	*[0.000]*	*[0.000]*	*[0.000]*
Switzerland	**0.0172**	**0.9574**	**- 0.0491**	0.0611	**0.9829**	**- 0.0293**
	[0.007]	*[0.000]*	*[0.051]*	*[0.137]*	*[0.000]*	*[0.033]*
UK	**0.1112**	**0.9489**	- 0.1146	**0.0615**	**0.9796**	- 0.0154
	[0.000]	*[0.000]*	*[0.315]*	*[0.004]*	*[0.000]*	*[0.691]*
USA	0.0127	**0.9599**	**- 0.1568**	0.0395	**0.9755**	- 0.1321
	[0.494]	*[0.000]*	*[0.000]*	*[0.175]*	*[0.000]*	*[0.310]*

Notes to Table 3: Values in bold denote 5% significance level.
The AGDCC Model: $Q_t = \Omega + aa' \circ \varepsilon_{t-1}\varepsilon'_{t-1} + gg' \circ \eta_{t-1}\eta'_{t-1} + bb' \circ Q_{t-1}$, where $\circ$ denotes the Hadamard operator.

Table 4: Log-Likelihood, BIC and AIC of Asymmetric Generalized DCC (AGDCC) and nested models with Multivariate Normal and Asymmetric Multivariate Laplace distributions.

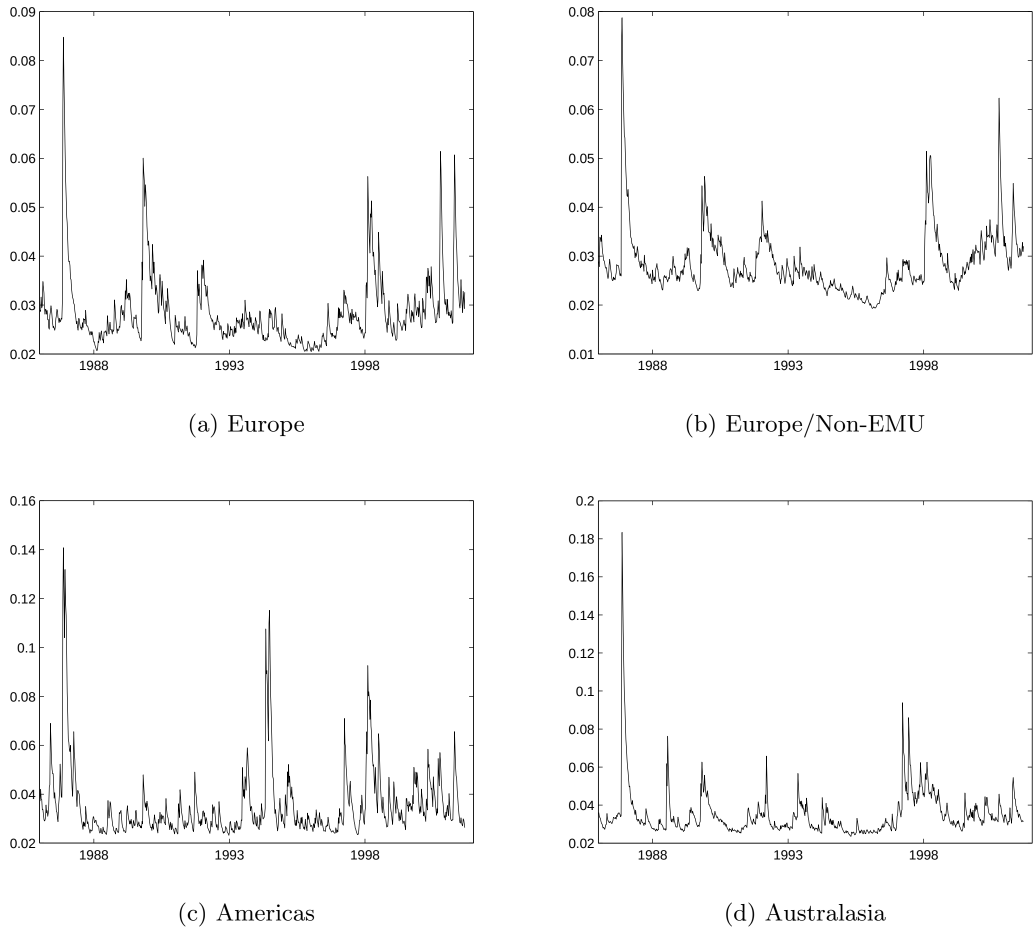
Model	**Parameters**	**MVN**	**AML**
	Panel A (Log-Likelihood)		
DCC	561+136+2	78,285.37	78,391.44
ADCC	561+136+3	78,310.61	78,490.62
GDCC	561+136+68	78,378.47	78,546.09
AGDCC	561+136+102	78,471.79	78,599.19
	Panel B (BIC)		
DCC	561+136+2	-5.60652	-5.61448
ADCC	561+136+3	-5.60803	-5.62154
GDCC	561+136+68	-5.58827	-5.60085
AGDCC	561+136+102	-5.58228	-5.59183
	Panel C (AIC)		
DCC	561+136+2	-5.82131	-5.88215
ADCC	561+136+3	-5.82312	-5.88959
GDCC	561+136+68	-5.82334	-5.89375
AGDCC	561+136+102	-5.82779	-5.89774

Notes to Table 4: The nested models are the scalar DCC, asymmetric scalar DCC, generalized DCC and asymmetric generalized DCC. The likelihoods can be directly compared between the two distributions as they are nested in the Generalized Hyperbolic Distribution (see Kotz et al.(2003) and Kozubowski et al.(2010)). The Bayesian Information Coefficient (BIC) is calculated as $(-2LL)/N + mlog_e(N)/N$, and the Akaike Information Coefficient (AIC) as $(-2LL)/N + 2m/N$ where N is the number of observations and m the number of estimated parameters. The 'Parameters' column reports the estimated number of parameters as : 'AGDCC Intercept(lower triangular)' + 'GJR (1st Stage)'+'AGDCC (2nd Stage)', respectively.

Table 5: Value at Risk Tests for Asymmetric Generalized DCC (AGDCC) Model with Multivariate Normal and Asymmetric Multivariate Laplace distributions.

Model	**GJR-N - AGDCC-MVN**		**GJR-N - AGDCC-AML**	
Test/Quantile	**5%**	**1%**	**5%**	**1%**
UC	0.3059	**0.5749**	**0.4239**	0.1610
FHS		*0.5053*	*0.6487*	
IND	0.1124	0.0012	**0.1398**	**0.8972**
FHS			*0.2390*	*0.8524*
CC	0.1252	0.0022	**0.2174**	**0.2495**
FHS			*0.4179*	*0.2784*
TR	0.0387	0.2168	**0.9999**	**0.9985**
ES	**0.5000**	0.4640	0.4810	**0.7670**
FHS	*0.5800*			*0.6790*

Notes to Table 5: The table reports the insample p-values for the Unconditional (UC), Independence (IND) and Conditional (CC) Value at Risk tests, and tests of Tail Risk (TR) and Expected shortfall (ES) for the 5% and 1% quantiles under the 2 distributions. A Filtered Historical Simulation (FHS) is also displayed for each model/test with the highest parametric p-value (in bold).



(a) Europe

(b) Europe/Non-EMU

(c) Americas

(d) Australasia

Figure 1: Plots of the weekly average Equity volatilities across the 4 major regions from the Asymmetric Generalized DCC model with Asymmetric Multivariate Laplace distribution

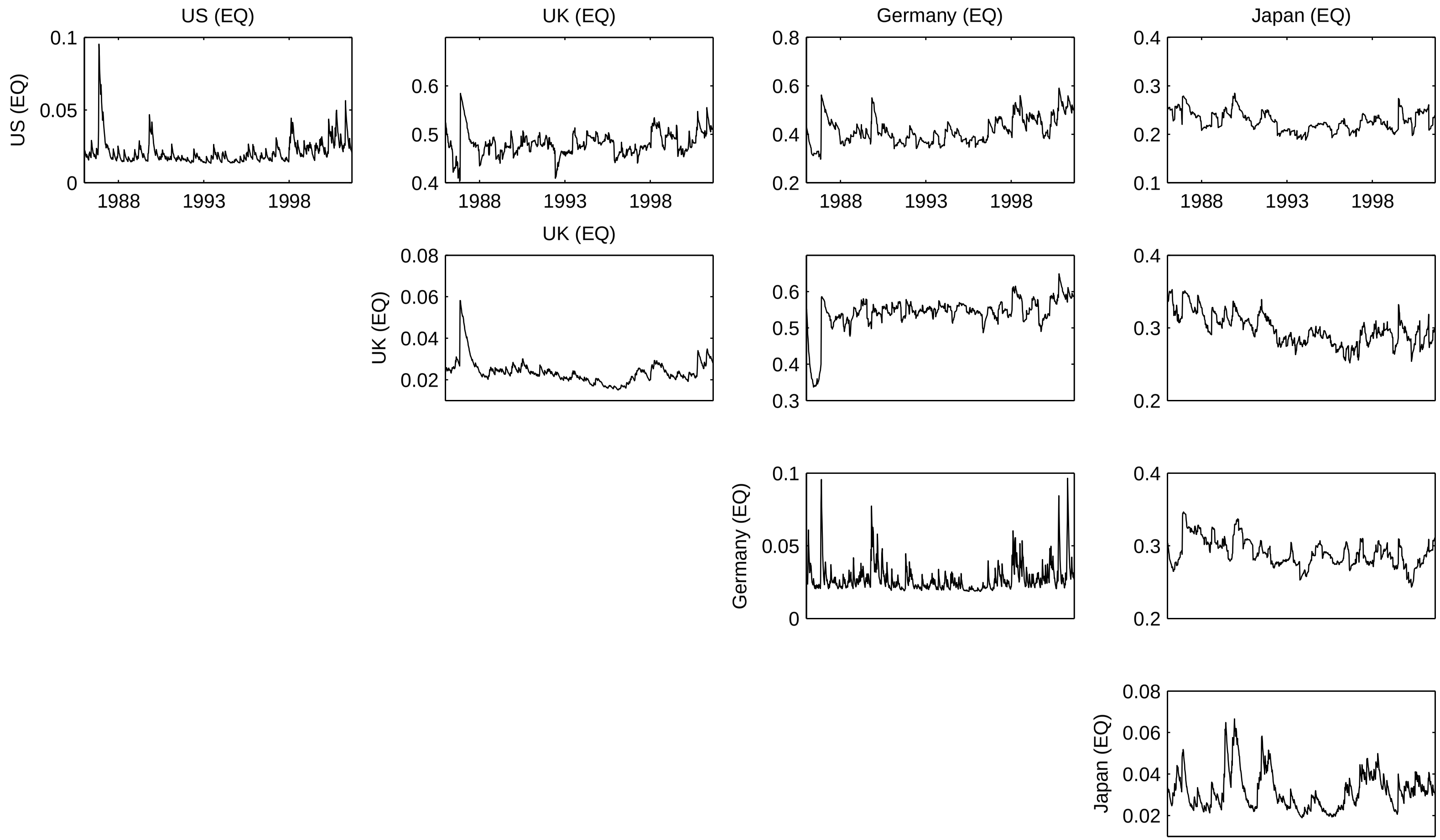


Figure 2: Plots of selected time-varying weekly Equity (EQ) correlations of the Asymmetric Generalized DCC model with Asymmetric Multivariate Laplace distribution. The diagonal plots are the time-varying weekly volatilities.

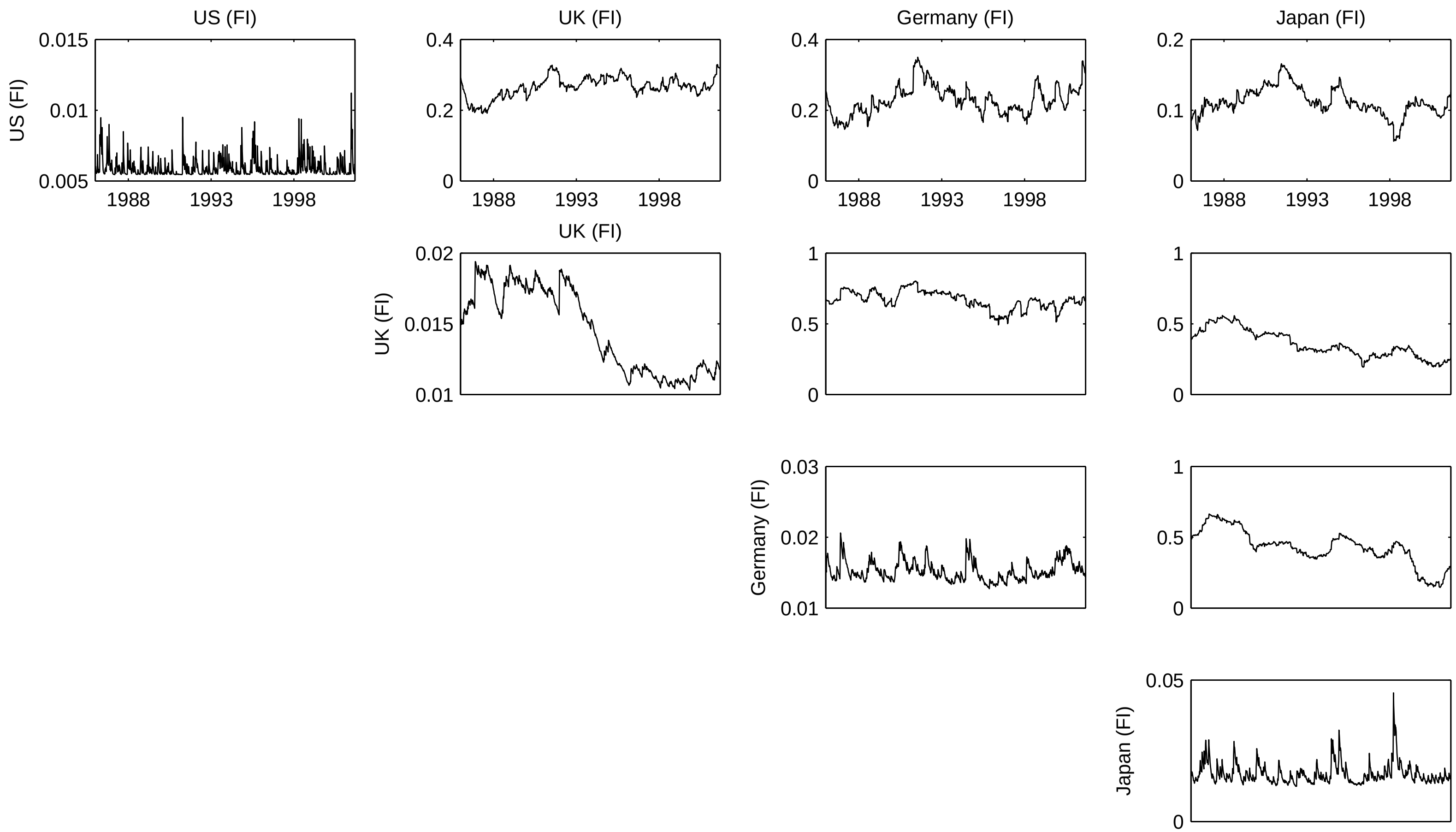


Figure 3: Plots of selected time-varying weekly Fixed Income (FI) correlations of the Asymmetric Generalized DCC model with Asymmetric Multivariate Laplace distribution. The diagonal plots are the time-varying weekly volatilities.

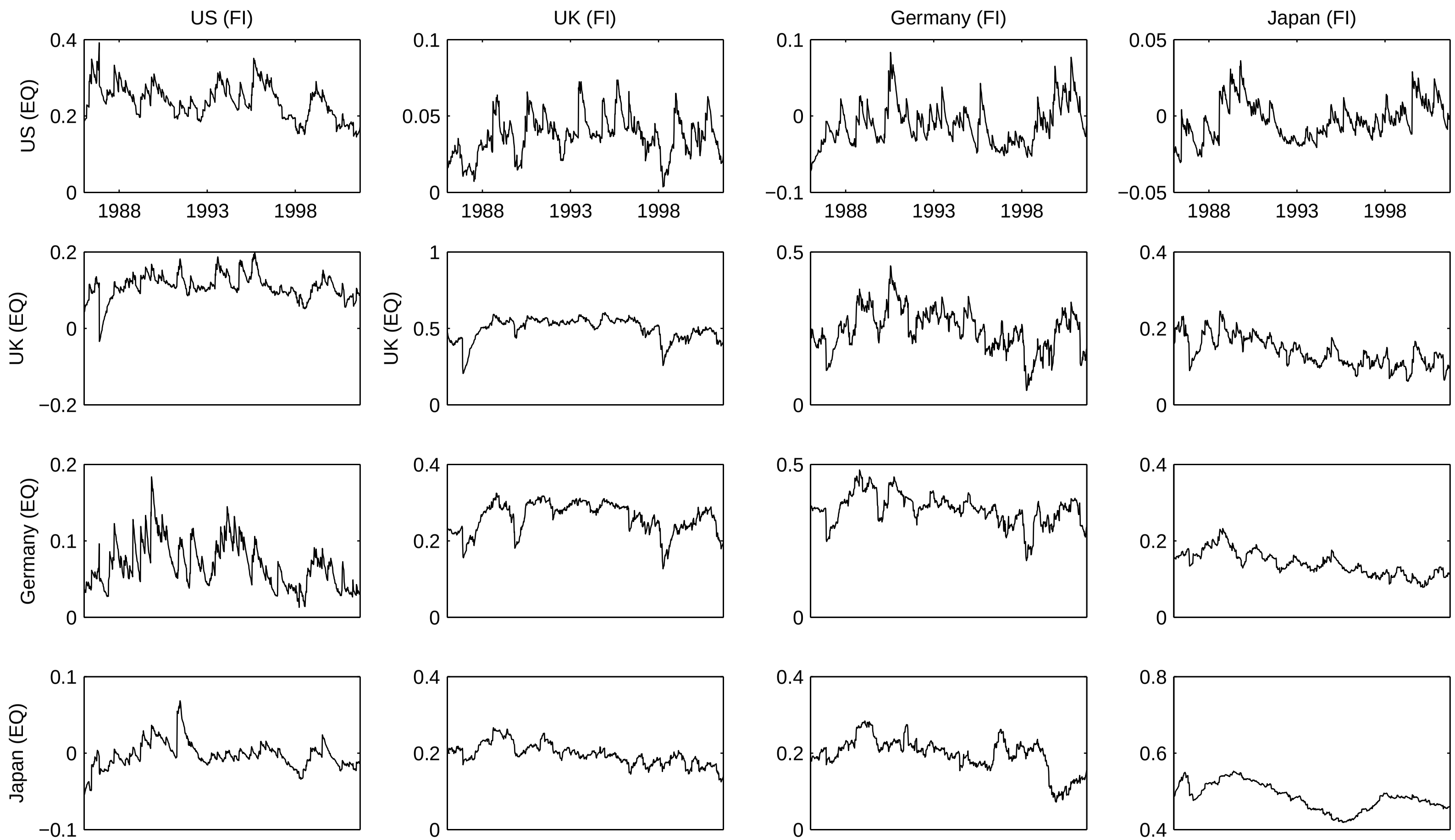


Figure 4: Plots of selected time-varying weekly Equity (EQ) against Fixed Income (FI) correlations of the Asymmetric Generalized DCC model with Asymmetric Multivariate Laplace distribution.