CENTER FOR FINANCIAL STUDIES

No. 2006/09

Multivariate Normal Mixture GARCH

Markus Haas, Stefan Mittnik,
and Marc S. Paolella

Center for Financial Studies
an der Universität Frankfurt ▪ Mertonstraße 17-21 ▪ D-60325 Frankfurt am Main
Telefon: +49 (0)69 798-30050
Fax: +49 (0)69 798-30077 ▪ http://www.ifk-cfs.de ▪ E-Mail: info@ifk-cfs.de

Center for Financial Studies

The *Center for Financial Studies* is a nonprofit research organization, supported by an association of more than 120 banks, insurance companies, industrial corporations and public institutions. Established in 1968 and closely affiliated with the University of Frankfurt, it provides a strong link between the financial community and academia.

The CFS Working Paper Series presents the result of scientific research on selected topics in the field of money, banking and finance. The authors were either participants in the Center´s Research Fellow Program or members of one of the Center´s Research Projects.

If you would like to know more about the *Center for Financial Studies*, please let us know of your interest.

Prof. Dr. Jan Pieter Krahnen

Prof. Volker Wieland, Ph.D.

Center for Financial Studies
an der Universität Frankfurt ▪ Mertonstraße 17-21 ▪ D-60325 Frankfurt am Main
Telefon: +49 (0)69 798-30050
Fax: +49 (0)69 798-30077 ▪ http://www.ifk-cfs.de ▪ E-Mail: info@ifk-cfs.de

CFS Working Paper No. 2006/09

Multivariate Normal Mixture GARCH*

Markus Haas[1], Stefan Mittnik[2],
and Marc S. Paolella[3]

April 2006

Abstract:
We present a multivariate generalization of the mixed normal GARCH model proposed in Haas, Mittnik, and Paolella (2004a). Issues of parametrization and estimation are discussed. We derive conditions for covariance stationarity and the existence of the fourth moment, and provide expressions for the dynamic correlation structure of the process. These results are also applicable to the single-component multivariate GARCH(p, q) model and simplify the results existing in the literature. In an application to stock returns, we show that the disaggregation of the conditional (co)variance process generated by our model provides substantial intuition, and we highlight a number of findings with potential significance for portfolio selection and further financial applications, such as regime-dependent correlation structures and leverage effects.

JEL Classification: C32, C51, G10, G11

Keywords: Conditional Volatility, Regime-dependent Correlations, Leverage Effect, Multivariate GARCH, Second-order Dependence

* We thank participants of the 2005 NBER/NSF Satellite Workshop \Financial Risk and Time Series Analysis" in Munich for helpful comments and suggestions. The research of M. Haas und S. Mittnik was supported by the Deutsche Forschungsgemeinschaft (SFB 386). Part of the research of M. S. Paolella has been carried out within the National Centre of Competence in Research \Financial Valuation and Risk Management" (NCCR FINRISK), which is a research program supported by the Swiss National Science Foundation. Address correspondence to Stefan Mittnik, Chair of Financial Econometrics, Institute of Statistics, University of Munich, D-80799 Munich, Germany, or e-mail: finmetrics@stat.uni-muenchen.de.

1 Institute of Statistics, University of Munich

2 Institute of Statistics, University of Munich, Center for Financial Studies, Frankfurt, and Ifo Institute for Economic Research, Munich

3 Swiss Banking Institute, University of Zurich

Non–technical Summary

In this paper, we propose a multivariate generalization of the normal mixture GARCH model originally proposed in Haas, Mittnik, and Paolella (2004a, an earlier version has also been published as CFS Working Paper 2002/10). One of the most characteristic properties of this model is that it explicitly allows the evolution of risk inherent in a given financial position to depend on—unobservable—*states of the market*, such as, for example, bull and bear markets. This meets frequently expressed concerns about standard GARCH models, which are not able to capture state–dependent volatility dynamics.

As shown in Alexander and Lazar (2004, 2005), and Haas, Mittnik, and Paolella (2004a,b) for a considerable number of financial return series, the normal mixture GARCH model is well suited for modeling and forecasting the volatility of financial assets such as stocks and currencies, and consistently outperforms many competing approaches both in– and out–of–sample. However, while the existing literature on normal mixture GARCH models is confined to univariate processes, many applications in finance are inherently multivariate and require us to understand the dependence structure between assets. For example, in portfolio management, correlations between assets are often of predominant interest, because the size of the correlations determines the degree of risk reduction which can be achieved by efficient portfolio diversification. However, there is evidence that stock returns exhibit stronger dependence in bear markets, when volatility is high and market returns are decreasing. This issue is of considerable importance for portfolio selection and risk management, because it is in times of adverse market conditions that the benefits from diversification are most urgently needed. Models not taking into account the state–dependent correlation structure will thus tend to overstate the benefits of diversification in bear markets, and, consequently, they will underestimate the risk during such periods.

We discuss this and further implications of the mixture approach to multivariate GARCH models in the paper, and demonstrate their empirical relevance in an application to stock market returns. Moreover, we address issues of parametrization, estimation, and model selection, and we derive various relevant dynamic properties of the multivariate normal mixture GARCH process.

Nichttechnische Zusammenfassung

Die vorliegende Arbeit ist einer multivariaten Verallgemeinerung des sog. *Normal Mixture* GARCH Modells gewidmet, dessen univariate Variante von Haas, Mittnik und Paolella (2004a, siehe auch CFS Working Paper 2002/10) vorgeschlagen wurde. Dieses Modell unterscheidet sich von traditionellen GARCH–Ansätzen insbesondere dadurch, dass es eine Abhängigkeit der Risikoentwicklung von – typischerweise unbeobachtbaren – Marktzuständen explizit in Rechnung stellt. Dies wird durch die Beobachtung motiviert, dass das weit verbreitete GARCH Modell in seiner Standardvariante auch dann keine adäquate Beschreibung der Risikodynamik leistet, wenn die Normalverteilung durch flexiblere bedingte Verteilungen ersetzt wird. Zustandsabhängige Volatilitätsprozesse können etwa durch die variierende Dominanz heterogener Marktteilnehmer oder durch wechselnde Marktstimmungen ökonomisch zu erklären sein.

Anwendungen des *Normal Mixture* GARCH Modells auf zahlreiche Aktien– und Wechselkurszeitreihen (siehe z.B. Alexander und Lazar, 2004, 2005; und Haas, Mittnik und Paolella, 2004a,b) haben gezeigt, dass es sich zur Modellierung und Prognose des Volatilitätsprozesses der Renditen solcher Aktiva hervorragend eignet. Indes beschränken sich diese Analysen bisher auf die Untersuchung univariater Zeitreihen. Zahlreiche Probleme der Finanzwirtschaft erfordern jedoch zwingend eine multivariate Modellierung, mithin also eine Beschreibung der Abhängigkeitsstruktur zwischen den Renditen verschiedener Wertpapiere. Insbesondere für solche Analysen erweist sich der Mischungsansatz aber als besonders vielversprechend. So spielen etwa im Portfoliomanagement die Korrelationen zwischen einzelnen Wertpapierrenditen eine herausragende Rolle. Die Stärke der Korrelationen ist von entscheidender Bedeutung dafür, in welchem Ausmaß das Risiko eines effizienten Portfolios durch Diversifikation reduziert werden kann. Nun gibt es empirische Hinweise darauf, dass die Korrelationen etwa zwischen Aktien in Perioden, die durch starke Marktschwankungen und tendenziell fallende Kurse charakterisiert sind, stärker sind als in ruhigeren Perioden. Das bedeutet, dass die Vorteile der Diversifikation in genau jenen Perioden geringer sind, in denen ihr Nutzen am größten wäre. Modelle, die die Existenz unterschiedlicher Marktregime nicht berücksichtigen, werden daher dazu tendieren, die Korrelationen in den adversen Marktzuständen zu unterschätzen. Dies kann zu erheblichen Fehleinschätzungen des tatsächlichen Risikos während solcher Perioden führen.

Diese und weitere Implikationen des Mischungsansatzes im Kontext multivariater GARCH Modelle werden in der vorliegenden Arbeit diskutiert, und ihre Relevanz wird anhand einer empirischen Anwendung dokumentiert. Erörtert werden ferner Fragen der Parametrisierung und Schätzung des Modells, und einige relevante theoretische Eigenschaften werden hergeleitet.

1 Introduction

Since the publication of Engle's (1982) ARCH model and its generalization to GARCH by Bollerslev (1986), a considerable amount of research has been undertaken to develop models that adequately capture the volatility dynamics observed in financial return data at weekly, daily or higher frequencies. Within the GARCH class of models, the recently proposed family of normal mixture GARCH processes (Alexander and Lazar, 2004; Haas, Mittnik, and Paolella, 2004a,b) has been shown to be particularly well suited for analyzing and forecasting short–term financial volatility.[1] A finite mixture of a few normal distributions, say two or three, is capable of capturing the skewness and kurtosis detected in both conditional and unconditional return distributions, and can, when coupled with GARCH–type equations for the component variances, exhibit quite complex dynamics, as often observed in financial markets. For example, there may be components driven by nonstationary dynamics, while the overall process is still stationary. This corresponds to the observation that markets are stable most of the time, but, occasionally, subject to severe, short–lived fluctuations. Empirical results for several stock and exchange rate return series, as reported in Alexander and Lazar (2004, 2005), and Haas, Mittnik, and Paolella (2004a,b) show that the normal mixture GARCH process provides a plausible disaggregation of the conditional variance process, and that it performs well in out–of–sample density forecasting, which can be viewed as a rigorous check of model adequacy.

While the existing literature on normal mixture GARCH models is confined to univariate processes, many applications in finance are inherently multivariate and require us to understand the dependence structure between assets. For example, in applications to portfolio selection, correlations between assets are often of predominant interest. However, there is evidence that asset correlations are regime–dependent, in the sense that stock returns appear to exhibit stronger dependence during periods of high volatility, which are often associated with market downturns (see, for example, Patton, 2004). As stressed by Campbell, Koedijk, and Kofman (2002), the issue of regime–dependent correlations is of considerable interest for portfolio analysis, because it is in times of adverse market conditions that the benefits from diversification are most urgently needed.

In this paper, we generalize the normal mixture GARCH model proposed by Haas, Mittnik, and Paolella (2004a) to the multivariate setting. We will define the model in terms of the

[1] These models are generalizations of earlier proposed applications of normal mixture distributions in the GARCH context (see Vlaar and Palm, 1993; Palm and Vlaar, 1997; and Bauwens, Bos, and van Dijk, 1999). There is also some relationship with the models of Wong and Li (2001), and Cheung and Xu (2003), as well as with the Markov–switching (G)ARCH models of Cai (1994), Hamilton and Susmel (1994), Gray (1996), Dueker (1997), and Klaassen (2002). A detailed discussion of these models and their relationships is provided in Haas, Mittnik, and Paolella (2004a,b).

arguably most general multivariate GARCH specification, i.e., the vech model as defined by Bollerslev, Engle, and Wooldridge (1988). This model, without further restrictions, is not amendable for direct estimation, but it nests several more practicable specifications, such as the diagonal vech model, also proposed by Bollerslev, Engle, and Wooldridge (1988), and the BEKK model of Engle and Kroner (1995).[2]

For the multivariate normal mixture GARCH(p, q) model, we present conditions for covariance stationarity and the existence of the fourth unconditional moment, along with expressions for the autocorrelation matrices of the squared process. As the mixture model nests the single–component specification, these results are also applicable to the standard multivariate GARCH(p, q) model in vech form. For this model, our results improve upon the existing literature on this issue, both in terms of simplicity and interpretability, as will be discussed in Appendix D. Moreover, no results for asymmetric multivariate GARCH models, i.e., specifications with a leverage effect, exist in the literature so far.

In the most general specification of our model, we allow for leverage effects, i.e., an asymmetric reaction of variances and covariances to positive and negative shocks, as well as for asymmetry of the conditional mixture density. The second– and fourth–order moment structure for this general specification is detailed for the empirically most relevant GARCH(1,1) model.

The paper is organized as follows. In Section 2, we define the model and present results on its unconditional moments and its dynamic correlation structure. In Section 3, we provide an application to a bivariate stock return series. Section 4 concludes and identifies issues for further research. Technical details are gathered in a set of appendices.

2 The Model and its Properties

In this section, we define the multivariate normal mixture GARCH process, discuss estimation issues and present some theoretical properties.

2.1 Finite Mixtures of Multivariate Normal Distributions

An M–dimensional random vector X is said to have a k–component multivariate finite normal mixture distribution, or, in short, MNM(k), if its density is given by

$$f(x) = \sum_{j=1}^{k} \lambda_j \phi(x; \mu_j, H_j), \tag{1}$$

[2] See Bauwens, Laurent, and Rombouts (2006) for an overview over multivariate GARCH models.

where $\lambda_j > 0$, $j = 1, \ldots, k$, $\sum_j \lambda_j = 1$, are the mixing weights, and

$$\phi(x; \mu_j, H_j) = \frac{1}{(2\pi)^{M/2}\sqrt{|H_j|}} \exp\left\{-\frac{1}{2}(x-\mu_j)' H_j^{-1}(x-\mu_j)\right\}, \quad j = 1, \ldots, k, \tag{2}$$

are the component densities. The normal mixture random vector has finite moments of all orders, with expected value and covariance matrix given by (see, e.g., McLachlan and Peel, 2000)

$$\mathrm{E}(X) = \sum_{j=1}^{k} \lambda_j \mu_j, \tag{3}$$

and

$$\begin{aligned} \mathrm{Cov}(X) &= \sum_{j=1}^{k} \lambda_j (H_j + \mu_j \mu_j') - \left(\sum_{j=1}^{k} \lambda_j \mu_j\right)\left(\sum_{j=1}^{k} \lambda_j \mu_j\right)' \\ &= \sum_{j=1}^{k} \lambda_j H_j + \sum_{j=1}^{k} \lambda_j (\mu_j - \mathrm{E}(X))(\mu_j - \mathrm{E}(X))', \end{aligned} \tag{4}$$

respectively. We will also make use the third and fourth moments of a multivariate normal mixture distribution, which are given in Appendix B.

A question that naturally arises in the estimation of mixture distributions is identifiability. Obviously, a lack of identification always arises as a consequence of label switching, but this can be ruled out by restricting the parameter space such that no duplication appears, e.g., by imposing $\lambda_1 > \lambda_2 > \cdots > \lambda_k$. However, there is a more fundamental problem when the class of density functions to be mixed is linearly dependent (Yakowitz and Spragins, 1968). Fortunately, the class of multivariate finite normal mixtures is identifiable, as has been shown by Yakowitz and Spragins (1968), who generalized Teicher's (1963) results for univariate finite normal mixtures.

An issue which has not been satisfactorily resolved so far is the empirical determination of the number of mixture components, i.e., the choice of k in (1). It is well–known that standard test theory breaks down in this context (McLachlan and Peel, 2000). However, there is some evidence, that, at least for unconditional mixture models, the Bayesian information criterion of Schwarz (1978) provides a reasonably good indication for the number of components (see McLachlan and Peel, 2000, Ch. 6, for a survey and further references). According to Kass and Raftery (1995), a BIC difference of less than two corresponds to "not worth more than a bare mention", while differences between two and six imply positive evidence, differences between six and ten give rise to strong evidence, and differences greater than ten invoke very strong evidence.

2.2 Multivariate Normal Mixture GARCH Processes

The M–dimensional time series $\{\epsilon_t\}$ is said to be generated by a k–component multivariate normal mixture GARCH(p,q) process, or, in short, MNM(k)–GARCH(p,q), if its conditional distribution is a k–component multivariate normal mixture, denoted as

$$\epsilon_t|\Psi_{t-1} \sim \mathrm{MNM}(\lambda_1,\ldots,\lambda_k,\mu_1,\ldots,\mu_k,H_{1t},\ldots,H_{kt}), \tag{5}$$

where Ψ_t is the information set at time t. By imposing $\mu_k = -\sum_{j=1}^{k-1}(\lambda_j/\lambda_k)\mu_j$ on the mean of the kth component it is guaranteed that ϵ_t in (5) has zero mean. Furthermore, stack the $N := M(M+1)/2$ independent elements of the covariance matrices and the "squared" ϵ_t (i.e., $\epsilon_t\epsilon_t'$) in $h_{jt} := \mathrm{vech}(H_{jt})$, $j = 1,\ldots,k$, and $\eta_t := \mathrm{vech}(\epsilon_t\epsilon_t')$, respectively. Then, the component covariance matrices evolve according to

$$h_{jt} = A_{0j} + \sum_{i=1}^{q} A_{ij}\tilde{\eta}_{ij,t-i} + \sum_{i=1}^{p} B_{ij}h_{j,t-i}, \quad j = 1,\ldots,k, \tag{6}$$

where $\tilde{\eta}_{ij,t} = \mathrm{vech}[(\epsilon_t - \theta_{ij})(\epsilon_t - \theta_{ij})']$; θ_{ij}, $i = 1,\ldots,q$, and A_{0j} are columns of length M and N, respectively; and A_{ij}, $i = 1,\ldots,q$, and B_{ij}, $i = 1,\ldots,p$, are $N\times N$ matrices, $j = 1,\ldots,k$. The θ_{ij}'s are introduced in order to allow for the leverage effect in applications to stock market returns, i.e., the strong negative correlation between equity returns and future volatility. In the univariate GARCH literature, various specifications of the leverage effect exist. Our choice, i.e., incorporating the θ_{ij}'s in (6), can be viewed as a generalization of one of the earliest versions, namely Engle's (1990) asymmetric GARCH (AGARCH) model.[3] In the univariate framework, this model has been coupled with the normal mixture GARCH structure by Alexander and Lazar (2005). We will denote the asymmetric MNM(k)–GARCH(p,q) as MNM(k)–AGARCH(p,q). Moreover, in some applications, a symmetric conditional density will be appropriate, so that, in (5), $\mu_1 = \cdots = \mu_k = 0$. We will denote this restricted version as MNM$_S(k)$–(A)GARCH(p,q). An overview over the different model specifications is provided in Table 1.

To compactify the notation and facilitate the theoretical analysis of the model, note that, by (A.3) in Appendix A, $\mathrm{vech}(\epsilon_{t-i}\theta_{ij}' + \theta_{ij}\epsilon_{t-i}') = 2D_M^+\mathrm{vec}(\theta_{ij}\epsilon_{t-i}') = 2D_M^+(I_M\otimes\theta_{ij})\epsilon_{t-i}$. Then we rewrite (6) as

$$h_{jt} = \tilde{A}_{0j} + \sum_{i=1}^{q} A_{ij}\eta_{t-i} - \sum_{i=1}^{q}\Theta_{ij}\epsilon_{t-i} + \sum_{i=1}^{p} B_{ij}h_{j,t-i}, \quad j = 1,\ldots,k, \tag{7}$$

where $\tilde{A}_{0j} := A_{0j} + \sum_{i=1}^{q} A_{ij}\mathrm{vech}(\theta_{ij}\theta_{ij}')$, and $\Theta_{ij} := 2A_{ij}D_M^+(I_M\otimes\theta_{ij})$, $j = 1,\ldots,k$, $i = 1,\ldots,q$. Let $h_t := (h_{1t}',\ldots,h_{kt}')'$; $\tilde{A}_0 = (\tilde{A}_{01}',\ldots,\tilde{A}_{0k}')'$; $\Theta_i = (\Theta_{i1}',\ldots,\Theta_{ik}')'$, $A_i =$

[3] The acronym AGARCH is due to Engle and Ng (1993).

Table 1: Variants of MNM–GARCH models.

Model	Conditional Density	Leverage Effect
$\mathrm{MNM}_S(k)$–GARCH(p,q)	symmetric	no
$\mathrm{MNM}_S(k)$–AGARCH(p,q)	symmetric	yes
MNM(k)–GARCH(p,q)	possibly asymmetric	no
MNM(k)–AGARCH(p,q)	possibly asymmetric	yes

A symmetric conditional density is enforced by restricting the component means in (5) to zero, i.e., $\mu_1 = \cdots = \mu_k = 0$; while the absence of a leverage effect is imposed by restricting the θ_{ij}'s in (6) to zero, i.e., $\theta_{ij} = 0$, $j = 1, \ldots, k$, $i = 1, \ldots, q$.

$(A'_{i1}, \ldots, A'_{ik})'$, $i = 0, \ldots, q$; and $B_i = \bigoplus_{j=1}^{k} B_{ij}$, $i = 1, \ldots, p$, where $\bigoplus$ denotes the matrix direct sum. Using these definitions, we have

$$h_t = \tilde{A}_0 + \sum_{i=1}^{q} A_i \eta_{t-i} - \sum_{i=1}^{q} \Theta_i \epsilon_{t-i} + \sum_{i=1}^{p} B_i h_{t-i}. \tag{8}$$

For estimation purposes, the general formulation as given in (6) is not directly applicable, and parameter constraints are required in order to guarantee positive definiteness of all conditional covariances matrices. A particular restriction of the vech form of the multivariate GARCH process, which guarantees positive definiteness, is implied by the BEKK model of Engle and Kroner (1995) which specifies the covariance matrices as

$$H_{jt} = A^{\star}_{0j} A^{\star\prime}_{0j} + \sum_{\ell=1}^{L} \sum_{i=1}^{q} A^{\star}_{ij,\ell} (\epsilon_{t-i} - \theta_{ij})(\epsilon_{t-i} - \theta_{ij})' A^{\star\prime}_{ij,\ell} + \sum_{\ell=1}^{L} \sum_{i=1}^{p} B^{\star}_{ij,\ell} H_{j,t-i} B^{\star\prime}_{ij,\ell}, \quad j = 1, \ldots, k, \tag{9}$$

where $A^{\star}_{0j}$, $j = 1, \ldots, k$, are triangular matrices. As shown by Engle and Kroner (1995), each BEKK model implies a unique vech representation (the converse is not true), and, once a BEKK representation (9) is estimated, the matrices A_{ij} and B_{ij} of the vech model (6) can be recovered via

$$\begin{aligned} A_{ij} &= \sum_{\ell=1}^{L} D_M^{+} (A^{\star}_{ij,\ell} \otimes A^{\star}_{ij,\ell}) D_M, \quad i = 1, \ldots, q, \quad j = 1, \ldots, k, \\ B_{ij} &= \sum_{\ell=1}^{L} D_M^{+} (B^{\star}_{ij,\ell} \otimes B^{\star}_{ij,\ell}) D_M, \quad i = 1, \ldots, p, \quad j = 1, \ldots, k, \end{aligned} \tag{10}$$

where D_M and D_M^{+} denote the duplication matrix and its Moore–Penrose inverse, respectively, both of which we briefly review in Appendix A. Thus, all results derived for the vech model are also applicable to the BEKK model. In practical applications, $L = 1$ is the standard choice, as well as $p = q = 1$. For this specification, it follows from Proposition 2.1 of Engle and Kroner (1995) that the model is identified if the diagonal elements of $A^{\star}_{0j}$, as well as the top left elements of matrices $A^{\star}_{1j}$ and $B^{\star}_{1j}$, $j = 1, \ldots, k$, are restricted to be positive. We will thus impose these restrictions in the applications below. In addition, while, for $L = 1$, the BEKK model already involves fewer parameters than the unrestricted vech form, further simplifications can be obtained by assuming that $A^\star_{1j}$ and $B^\star_{1j}$, $j = 1, \ldots, k$, are diagonal matrices.

In the following discussion of the vech specification we will always assume that positive definite covariances matrices are guaranteed, without further specifying the constraints employed for achieving this.

2.3 Existence of Moments and Autocorrelation Structure

It is clear that, for practical purposes, the most important MNM(k)–AGARCH(p, q) process is the specification where $p = q = 1$, which is defined by (5) and

$$h_t = \tilde{A}_0 + A_1 \eta_{t-1} - \Theta_1 \epsilon_{t-1} + B_1 h_{t-1}. \tag{11}$$

For later reference, we summarize the dynamic properties of the process given by (5) and (11) in Proposition 1, while the corresponding results for the GARCH(p, q) specification, which require a considerable amount of additional notation, are developed in Appendix D.

We denote as $\rho(A)$ the largest eigenvalue in modulus of a square matrix A, i.e.,

$$\rho(A) := \max\{|z| : z \text{ is an eigenvalue of } A\}, \tag{12}$$

and define the vector of mixing weights $\lambda = (\lambda_1, \ldots, \lambda_k)'$. Following the classic papers of Engle (1982) and Bollerslev (1986), we assume for simplicity that the process starts indefinitely far in the past with finite fourth moments.

Proposition 1 *The MNM(k)–AGARCH(1,1) process given by (5) and (11) is covariance stationary if and only if $\rho(C_{11}) < 1$, where the $kN \times kN$ matrix C_{11} is defined by*

$$C_{11} = \lambda' \otimes A_1 + B_1. \tag{13}$$

In this case, the unconditional expectation of vector h_t is $E(h_t) = (I_{kN} - \lambda' \otimes A_1 - B_1)^{-1}[\tilde{A}_0 + A_1(\lambda' \otimes I_N)\tilde{\mu}]$, where $\tilde{\mu}$ is defined in Lemma 4 in Appendix B.1; and the unconditional expectation of η_t is $(\lambda' \otimes I_N)(E(h_t) + \tilde{\mu})$. Moreover, the unconditional fourth moment $E(\eta_t \eta_t')$ exists if and only if $\rho(C_{22}) < 1$, where C_{22} is the $(kN)^2 \times (kN)^2$ matrix given by

$$C_{22} = (A_1 \otimes A_1) G_M (I_N \otimes vec(\Lambda)' \otimes I_N)(K_{Nk} \otimes I_{kN}) + 2N_{kN}(B_1 \otimes \lambda' \otimes A_1) + B_1 \otimes B_1. \tag{14}$$

In (14), G_M is the $N^2 \times N^2$ matrix defined in (B.13) in Appendix B.2, $\Lambda = diag(\lambda_1 \ldots, \lambda_k)$, K_{mn} is the commutation matrix defined in Appendix A, and $N_n = (I_{n^2} + K_{nn})/2$. An expression for the fourth–moment matrix is given in Appendix C.1. If $\rho(C_{22}) < 1$ holds, the multidimensional autocovariance function of the squared process, $\Gamma_\tau := E(\eta_t \eta_{t-\tau}') - E(\eta_t)E(\eta_t)'$, is *given by*

$$\Gamma_\tau = (\lambda' \otimes I_N) C_{11}^{\tau-1} Q, \tag{15}$$

where Q is a constant matrix given in (C.21) in Appendix C.2.

The results of Proposition 1 are derived in Appendices B and C. From (15), the autocorrelation matrices, R_τ, can be calculated in the usual way. I.e., if D is a diagonal matrix with the square roots of $\mathrm{diag}(\Gamma_0)$ on its diagonal, where $\Gamma_0 := \mathrm{E}(\eta_t \eta_t') - \mathrm{E}(\eta_t)\mathrm{E}(\eta_t)'$, then

$$R_\tau = D^{-1} \Gamma_\tau D^{-1}. \tag{16}$$

Note that the term determining the rate of decay of Γ_τ is C_{11}^τ. Thus, under covariance stationarity, the largest eigenvalue in magnitude of the matrix C_{11} defined in (13) can be used as a measure for the persistence of shocks to volatility.

It may be worth pointing out that conditions (13) and (14), as well as the speed of decline of the autocorrelation function, do not depend on the "leverage terms" θ_{1j} in (6). Moreover, the stationarity condition $\rho(C_{11}) < 1$, where C_{11} is defined in (13), allows some components to be nonstationary, in the sense that the covariance stationarity condition for single–component multivariate GARCH(1,1) processes, i.e.,[4]

$$\rho(A_{1j} + B_{1j}) < 1, \tag{17}$$

is not satisfied for these components. Nevertheless, the overall process can still be stationary, as long as the corresponding mixture weights are sufficiently small. This parallels the situation in the univariate case (see Alexander and Lazar, 2004; and Haas, Mittnik, and Paolella, 2004a,b), and will be empirically illustrated in Section 3.

3 Application to Stock Market Returns

We investigate the bivariate time series of daily returns of the NASDAQ and the Dow Jones Industrial Average (DJIA) indices from January 1990 to December 1999, a sample of $T = 2516$ observations.[5] Continuously compounded percentage returns are considered, i.e., $r_{it} = 100 \times \log(P_{it}/P_{i,t-1})$, $i = 1, 2$, where P_{it} denotes the level of index i at time t. We let r_{1t} and r_{2t} denote the time–t return of the NASDAQ and the DJIA, respectively. As we want to concentrate our analysis on the volatility dynamics, a univariate linear AR(1) filter was applied to the series in order to remove (weak) low–order autocorrelation. Subsequently, all

[4] For this condition, see Bollerslev and Engle (1993), and Engle and Kroner (1995).

[5] The data were obtained from Datastream.

Table 2: Descriptive statistics of the filtered NASDAQ/DJIA returns.

	Covariance matrix		Correlation matrix		Skewness	Kurtosis
	NASDAQ	DJIA	NASDAQ	DJIA		
NASDAQ	1.229	0.721	1	0.728	–0.431	7.620
DJIA	0.721	0.798	0.728	1	–0.379	8.132

"Skewness" and "Kurtosis" refer to the standardized third and fourth moments, respectively. That is, Skewness $= m_3/m_2^{3/2}$, and Kurtosis $= m_4/m_2^2$, where m_i is the ith central moment (about the mean).

results are for the filtered version of the data. A few descriptive statistics of the filtered series are summarized in Table 2.

To make sure that all conditional covariance matrices are positive definite, we use the BEKK parametrization (9). Several versions of the general mixture GARCH model (5)–(6) with $p = q = 1$ have been estimated. Namely, the single–component model, which corresponds to $k = 1$ in (1), and which is just the standard Normal–GARCH process, has been estimated with and without imposing a symmetric reaction to negative and positive shocks. The first of these models, where $\theta_{11} = 0$ in (6), will be denoted by Normal–GARCH(1,1), and the second by Normal–AGARCH(1,1). Also, two–component models are considered with and without symmetric conditional mixture densities, i.e., with and without imposing $\mu_1 = \mu_2 = 0$ in (5), as well as with and without leverage effects. To refer to these different models, we will use the typology of Table 1.

Table (3) reports likelihood–based goodness–of–fit measures for the models and their rankings with respect to each of these criteria, i.e., the value of the maximized log–likelihood function, and the AIC and BIC criteria of Akaike (1973) and Schwarz (1978), respectively.

While it is not surprising that the Normal–GARCH model is the worst performer with respect to each of these criteria, several additional observations are worth mentioning. First, the normal mixture specifications allowing for asymmetric conditional densities, i.e., admitting nonzero component means in (5), are always favored against their symmetric counterparts. This is not the case when we consider the dynamic asymmetry, i.e., the asymmetric reaction of future variances to negative and positive shocks. The improvement in log–likelihood is much larger when passing from the symmetric MNM$_S$(2)–GARCH(1,1) to the MNM$_S$(2)–AGARCH(1,1) model (difference in log–likelihood: 25.3), than when passing from the asymmetric MNM(2)–GARCH(1,1) process to its AGARCH(1,1) counterpart (difference in log–likelihood: 14.5). As a consequence, the MNM(2)–GARCH(1,1) specification performs best overall according to the BIC. We note, however, that the difference in BIC for the latter two models is close to being insignificant according to the Kass and Raftery–recommendation

Table 3: Likelihood–based goodness of fit.

Distributional Model	K	L Value	L Rank	AIC Value	AIC Rank	BIC Value	BIC Rank
Normal–GARCH(1,1)	11	-5598.6	6	11219.3	6	11283.4	6
$\mathrm{MNM}_S(2)$–GARCH(1,1)	23	-5502.4	4	11050.8	4	11184.9	4
MNM(2)–GARCH(1,1)	25	-5478.8	3	11007.6	2	**11153.4**	**1**
Normal–AGARCH(1,1)	13	-5584.1	5	11194.1	5	11269.9	5
$\mathrm{MNM}_S(2)$–AGARCH(1,1)	27	-5477.1	2	11008.1	3	11165.6	3
MNM(2)–AGARCH(1,1)	29	$-$**5464.3**	**1**	**10986.5**	**1**	11155.6	2

The leftmost column states the type of volatility model fitted to the bivariate NASDAQ/DJIA returns. The column labeled K reports the number of parameters of a model; L is the log likelihood; $\text{AIC} = -2L + 2K$; and $\text{BIC} = -2L + K \log T$, where T is the number of observations. For each of the three criteria the criterion value and the ranking of the models are shown. Boldface entries indicate the best model for the particular criterion.

mentioned at the end of Section 2.1. Also, a closer inspection of the parameter estimates will reveal that the leverage effect may be an exclusive feature of the high–volatility component, so that the difference in the number of parameters between these models shrinks from four to two, which would reverse these conclusions.[6]

The maximum likelihood estimates are reported in Tables 4 and 5 for the models without and with leverage effect, i.e., dynamic asymmetry, respectively. Reported are the parameter matrices $A^\star_{0j}$, $A^\star_{1j}$, and $B^\star_{1j}$, $j = 1, 2$, of the BEKK representation (9), from which the parameter matrices of the vech representation, A_{1j} and B_{1j}, $j = 1, 2$, can be recovered via (10). In addition, we report the regime–specific persistence measures, i.e., the largest eigenvalues of the matrices $A_{1j} + B_{1j}$, $j = 1, 2$, where these matrices have been computed from the BEKK representation using (10), as well as the largest eigenvalues of the matrices C_{11} and C_{22} defined in Proposition 1, which provide information about the existence of the unconditional second and fourth moments, respectively. The two–component models have been ordered such that $\lambda_1 > \lambda_2$.

In discussing the parameter estimates, we first draw attention to a common characteristic of all mixture models fitted, whether they allow for asymmetry and/or leverage or not: All these models identify two components with distinctly different volatility dynamics. More precisely, the first component, i.e., the component with the larger mixing weight, is stationary in the sense that $\rho(A_{11} + B_{11}) < 1$, and it has less weight on the reaction parameters in A_{11} and

[6] Alternatively, a likelihood ratio test for $\theta_1 = \theta_2 = 0$ could be conducted. The associated test statistic, $LRT = 2 \times (5478.8 - 5464.3) = 29$, exceeds conventional critical values given by the asymptotically valid χ^2 distribution with four degrees of freedom, thus favoring the model with the leverage terms.

Table 4: MNM–GARCH(1,1) parameter estimates for NASDAQ/DJIA returns

	Normal–GARCH(1,1)	MNM$_S$(2)–GARCH(1,1)	MNM(2)–GARCH(1,1)
$A_{01}^{\star}$	$\begin{pmatrix} \underset{(0.019)}{0.128} & 0 \\ \underset{(0.019)}{0.045} & \underset{(0.014)}{0.030} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.028)}{0.024} & 0 \\ \underset{(0.016)}{-0.007} & \underset{(0.023)}{0.000} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.025)}{0.025} & 0 \\ \underset{(0.019)}{-0.006} & \underset{(0.043)}{0.000} \end{pmatrix}$
$A_{11}^{\star}$	$\begin{pmatrix} \underset{(0.034)}{0.373} & \underset{(0.035)}{-0.139} \\ \underset{(0.018)}{0.088} & \underset{(0.022)}{0.099} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.027)}{0.290} & \underset{(0.030)}{-0.144} \\ \underset{(0.011)}{0.063} & \underset{(0.018)}{0.060} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.027)}{0.264} & \underset{(0.030)}{-0.110} \\ \underset{(0.015)}{0.055} & \underset{(0.020)}{0.075} \end{pmatrix}$
$B_{11}^{\star}$	$\begin{pmatrix} \underset{(0.014)}{0.922} & \underset{(0.011)}{0.042} \\ \underset{(0.006)}{-0.029} & \underset{(0.006)}{1.005} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.006)}{0.954} & \underset{(0.005)}{0.027} \\ \underset{(0.001)}{-0.017} & \underset{(0.002)}{1.002} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.008)}{0.958} & \underset{(0.006)}{0.021} \\ \underset{(0.004)}{-0.016} & \underset{(0.003)}{1.000} \end{pmatrix}$
$\rho(A_{11}+B_{11})$	0.997	0.994	0.994
θ_{11}	–	–	–
λ_1	1	$\underset{(0.041)}{0.8270}$	$\underset{(0.033)}{0.836}$
μ_1	–	–	$\begin{pmatrix} \underset{(0.081)}{0.109}, \underset{(0.042)}{0.049} \end{pmatrix}'$
$A_{02}^{\star}$	–	$\begin{pmatrix} \underset{(0.103)}{0.554} & 0 \\ \underset{(0.113)}{0.328} & \underset{(0.059)}{0.150} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.082)}{0.448} & 0 \\ \underset{(0.089)}{0.398} & \underset{(0.187)}{0.000} \end{pmatrix}$
$A_{12}^{\star}$	–	$\begin{pmatrix} \underset{(0.146)}{0.731} & \underset{(0.165)}{0.012} \\ \underset{(0.111)}{0.220} & \underset{(0.124)}{0.345} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.143)}{0.753} & \underset{(0.144)}{-0.048} \\ \underset{(0.115)}{0.218} & \underset{(0.122)}{0.353} \end{pmatrix}$
$B_{12}^{\star}$	–	$\begin{pmatrix} \underset{(0.047)}{0.736} & \underset{(0.074)}{0.072} \\ \underset{(0.056)}{-0.083} & \underset{(0.053)}{0.973} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.075)}{0.829} & \underset{(0.074)}{-0.029} \\ \underset{(0.058)}{-0.038} & \underset{(0.061)}{0.916} \end{pmatrix}$
$\rho(A_{12}+B_{12})$	–	1.163	1.172
θ_{12}	–	–	–
λ_2	0	$\underset{(0.041)}{0.173}$	$\underset{(0.033)}{0.164}$
μ_2	–	–	$\begin{pmatrix} \underset{(0.108)}{-0.553}, \underset{(0.084)}{-0.248} \end{pmatrix}'$
$\rho(C_{11})$	0.997	0.995	0.996
$\rho(C_{22})$	0.994	0.994	0.994

Approximate standard errors are given in parentheses. Note that matrices $A_{0j}^{\star}$, $A_{1j}^{\star}$, and $B_{1j}^{\star}$, $j = 1, 2$, correspond to the BEKK representation (9) of the model, while matrices $A_{1j} + B_{1j}$, $j = 1, 2$, the maximal eigenvalues of which are reported, are associated with the vech representation (6). $\rho(C_{11})$ and $\rho(C_{22})$ denote the largest eigenvalues of the matrices C_{11} and C_{22}, defined in Proposition 1, which determine whether the unconditional second and fourth moments, respectively, exist.

Table 5: MNM–AGARCH(1,1) parameter estimates for NASDAQ/DJIA returns

	Normal–AGARCH(1,1)	MNM$_S$(2)–AGARCH(1,1)	MNM(2)–AGARCH(1,1)
$A^\star_{01}$	$\begin{pmatrix} \underset{(0.024)}{0.135} & 0 \\ \underset{(0.023)}{0.046} & \underset{(0.017)}{0.031} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.041)}{0.000} & 0 \\ \underset{(0.019)}{0.000} & \underset{(0.019)}{0.000} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.044)}{0.000} & 0 \\ \underset{(0.033)}{0.000} & \underset{(0.026)}{0.000} \end{pmatrix}$
$A^\star_{11}$	$\begin{pmatrix} \underset{(0.038)}{0.389} & \underset{(0.037)}{-0.135} \\ \underset{(0.020)}{0.094} & \underset{(0.023)}{0.108} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.028)}{0.288} & \underset{(0.030)}{-0.149} \\ \underset{(0.015)}{0.060} & \underset{(0.020)}{0.059} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.012)}{0.258} & \underset{(0.017)}{-0.114} \\ \underset{(0.013)}{0.052} & \underset{(0.018)}{0.068} \end{pmatrix}$
$B^\star_{11}$	$\begin{pmatrix} \underset{(0.017)}{0.911} & \underset{(0.014)}{0.042} \\ \underset{(0.008)}{-0.034} & \underset{(0.007)}{1.004} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.007)}{0.958} & \underset{(0.006)}{0.024} \\ \underset{(0.004)}{-0.015} & \underset{(0.003)}{1.001} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.002)}{0.963} & \underset{(0.002)}{0.017} \\ \underset{(0.002)}{-0.013} & \underset{(0.002)}{0.999} \end{pmatrix}$
$\rho(A_{11}+B_{11})$	0.996	0.998	0.996
θ_{11}	$\left(\underset{(0.062)}{0.305}, \underset{(0.079)}{0.243}\right)'$	$\left(\underset{(0.083)}{-0.113}, \underset{(0.097)}{-0.164}\right)'$	$\left(\underset{(0.076)}{-0.116}, \underset{(0.100)}{-0.153}\right)'$
λ_1	1	$\underset{(0.036)}{0.755}$	$\underset{(0.033)}{0.759}$
μ_1	–	–	$\left(\underset{(0.032)}{0.099}, \underset{(0.019)}{0.044}\right)'$
$A^\star_{02}$	–	$\begin{pmatrix} \underset{(0.128)}{0.132} & 0 \\ \underset{(0.080)}{-0.066} & \underset{(0.170)}{0.000} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.117)}{0.081} & 0 \\ \underset{(0.080)}{-0.088} & \underset{(0.210)}{0.000} \end{pmatrix}$
$A^\star_{12}$	–	$\begin{pmatrix} \underset{(0.094)}{0.635} & \underset{(0.097)}{0.027} \\ \underset{(0.078)}{0.193} & \underset{(0.074)}{0.312} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.059)}{0.603} & \underset{(0.071)}{-0.046} \\ \underset{(0.106)}{0.143} & \underset{(0.039)}{0.310} \end{pmatrix}$
$B^\star_{12}$	–	$\begin{pmatrix} \underset{(0.075)}{0.678} & \underset{(0.071)}{0.094} \\ \underset{(0.049)}{-0.121} & \underset{(0.039)}{0.989} \end{pmatrix}$	$\begin{pmatrix} \underset{(0.048)}{0.727} & \underset{(0.046)}{0.085} \\ \underset{(0.030)}{-0.091} & \underset{(0.027)}{0.981} \end{pmatrix}$
$\rho(A_{12}+B_{12})$	–	1.019	1.017
θ_{12}	–	$\left(\underset{(0.127)}{0.814}, \underset{(0.161)}{0.619}\right)'$	$\left(\underset{(0.114)}{0.878}, \underset{(0.138)}{0.636}\right)'$
λ_2	0	$\underset{(0.036)}{0.245}$	$\underset{(0.033)}{0.241}$
μ_2	–	–	$\left(\underset{(0.068)}{-0.310}, \underset{(0.055)}{-0.140}\right)'$
$\rho(C_{11})$	0.996	0.994	0.996
$\rho(C_{22})$	0.993	0.991	0.992

See the legend of Table 4 for explanations.

more weight on the persistence parameters in B_{11}, relative to the second component. The latter is nonstationary in the sense that $\rho(A_{12}+B_{12}) > 1$, and it has considerably more weight on the reaction and less on the persistence parameters. This implies that the high–volatility component reacts more strongly to shocks, but has a shorter memory. However, all estimated mixture models are stationary in the aggregate, because, for all models, the largest eigenvalue of the matrix C_{11}, defined in (13), is less than unity.

Also, if nonzero component means are allowed for, we observe that, both for the MNM(2)–GARCH(1,1) model in Table 4 and the MNM(2)–AGARCH(1,1) model in Table 5, the low–volatility component is associated with positive means, and the high–volatility component is associated with statistically significant negative means for both variables.

A similar finding holds for the leverage effects, i.e., the dynamic asymmetries in the GARCH structure, as reported in Table 5. For both mixture AGARCH models, a leverage effect seems to be present mainly in the high–volatility, bear market component. The leverage parameters in the first component, θ_{11}, are negative, and thus seem to indicate a “reverse” leverage effect, but they are also insignificant statistically. On the other hand, the leverage parameters of the nonstationary component, θ_{12}, are rather large, compared to those of the fitted Normal–AGARCH model, indicating a very strong negative relation between current returns and future volatility. Interestingly, this is in accordance with Figlewski and Wang (2000), who argue that the leverage effect is really a “down market effect” in the sense that, while there is a strong leverage effect associated with falling stock prices, there is a much weaker or nonexistent relation between positive stock returns and future volatility.

It is also interesting to note that the introduction of the leverage effects reduces the persistence measure of the high–volatility component somewhat, i.e., $\rho(A_{12} + B_{12})$ decreases.[7] However, at the same time, its mixing weight, λ_2, increases, so that the overall persistence of the model, as measured by $\rho(C_{11})$, remains approximately unchanged.

Another potentially relevant issue for financial applications is whether there are striking differences between the regime–specific correlation coefficients. Figure 1 displays the component–specific conditional correlations implied by the MNM(2)–GARCH(1,1) model.[8] The upper panel plots the conditional correlations in the positive–mean/low–volatility component, and the lower panel those in the negative–mean/high–volatility component. Most of the time, the correlation coefficient in the low–volatility regime is considerably smaller than that in the high–volatility regime, with the regimes' averages being 0.643 and 0.813, respectively.

[7] Admittedly, the interpretation of $\rho(A_{12} + B_{12})$ as a persistence measure is a little awkward when $\rho(A_{12} + B_{12}) > 1$.

[8] The results for the other mixture models are similar, and are available upon request.

It is clear that such a pattern can have significant implications for portfolio diversification, which will be a topic of future investigation. To tackle this task systematically, however, it may be more convenient to specify the dynamics in the correlation matrix directly, by using, for example, structures as proposed in Engle (2002), Tse and Tsui (2002), and, more recently, Pelletier (2006).

As the largest eigenvalue of the matrix C_{22}, reported in the bottom row of Tables 4 and 5, is below unity for all models, we can compute the theoretical auto– and cross–correlations implied by the fitted processes. As noted, for example, by He and Teräsvirta (2004), such calculations can help to assess whether a fitted model is capable of reproducing some of the dynamic properties of the data being investigated.

The auto– and cross–correlations are shown in Figures 2–9, along with their empirical counterparts. As expected, the mixture models mimic the empirical shapes much better than the single–component models, although the fit is not "optimal". A somewhat more surprising observation is the fact that the AGARCH specifications capture the observed auto- and cross-correlation structure less well than their GARCH counterparts. In particular, the AGARCH–implied auto– and cross–correlations tend to be somewhat smaller than the corresponding GARCH quantities. At first sight, and in view of Example 2 in Appendix C.2, this is somewhat surprising, as it is shown there, that, at least for the special case of the univariate QGARCH(1,1), the autocorrelations are *increasing* in θ^2. However, this result is true only if all other parameters of the model are held constant, and this is obviously not the case for the estimates reported in Tables 4 and 5. Nevertheless, these findings may indicate that, within the asymmetric GARCH structure adopted in (6), there is a trade–off between reproducing the correlation structure of the squares and capturing the asymmetric response of volatility to good and bad news. A possible consequence of this is to investigate other parameterizations of the leverage effect, such as that of Glosten, Jagannathan, and Runkle (1993) which has been used for multivariate GARCH models by Hansson and Hördahl (1998) and Kroner and Ng (1998).

4 Conclusions

In this paper, we have generalized the normal mixture GARCH model introduced in Haas, Mittnik, and Paolella (2004a) to the multivariate framework. For the vech representation of the multivariate GARCH process, conditions for covariance stationarity and the existence of the fourth moment were presented, along with expressions for the autocorrelation function of the squares.

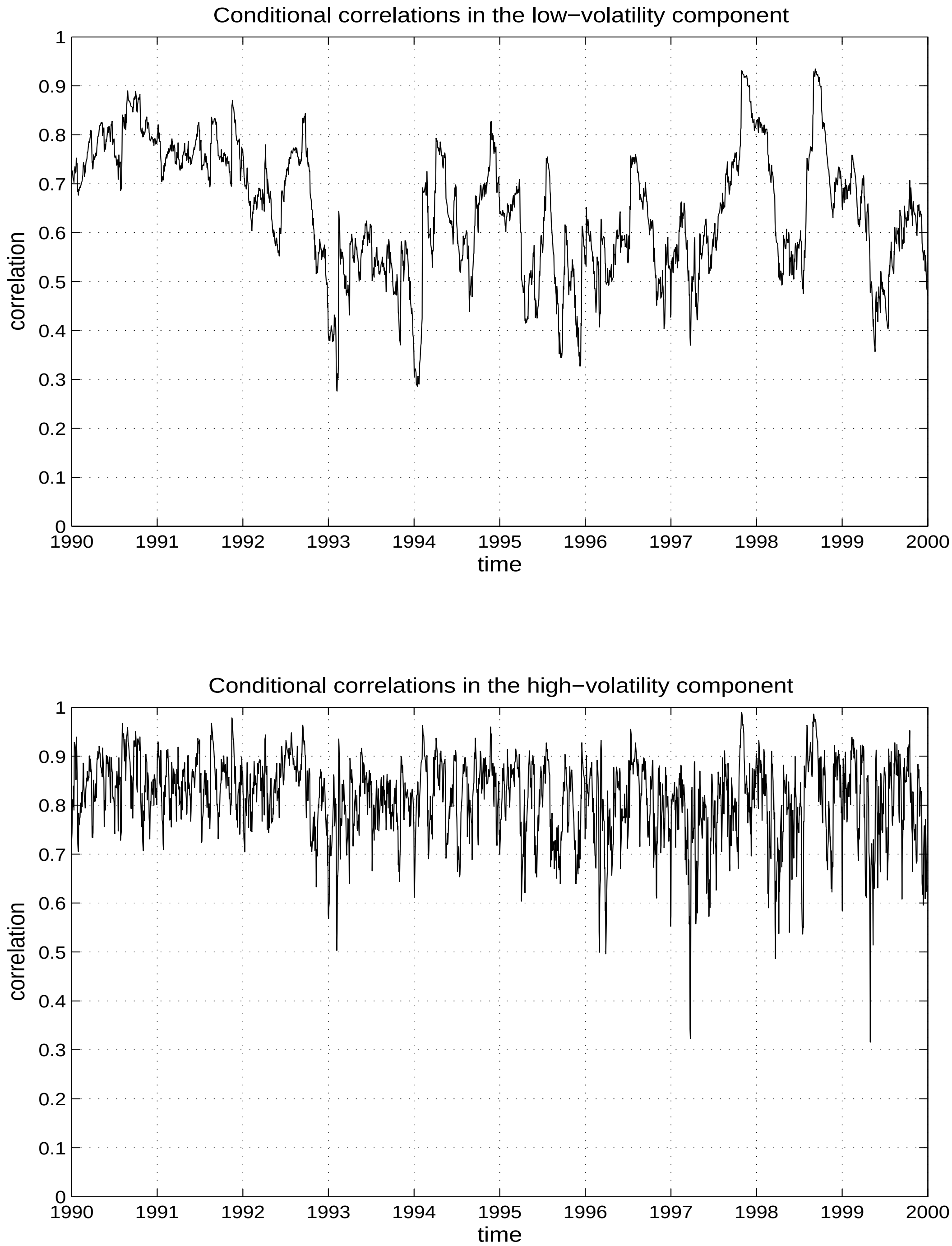

Figure 1: Shown are the implied component–specific correlations of the MNM(2)–GARCH(1,1) model, fitted to the NASDAQ/DJIA returns. The upper panel shows the conditional correlations in the low–volatility component, those in the high–volatility component are depicted in the lower panel.

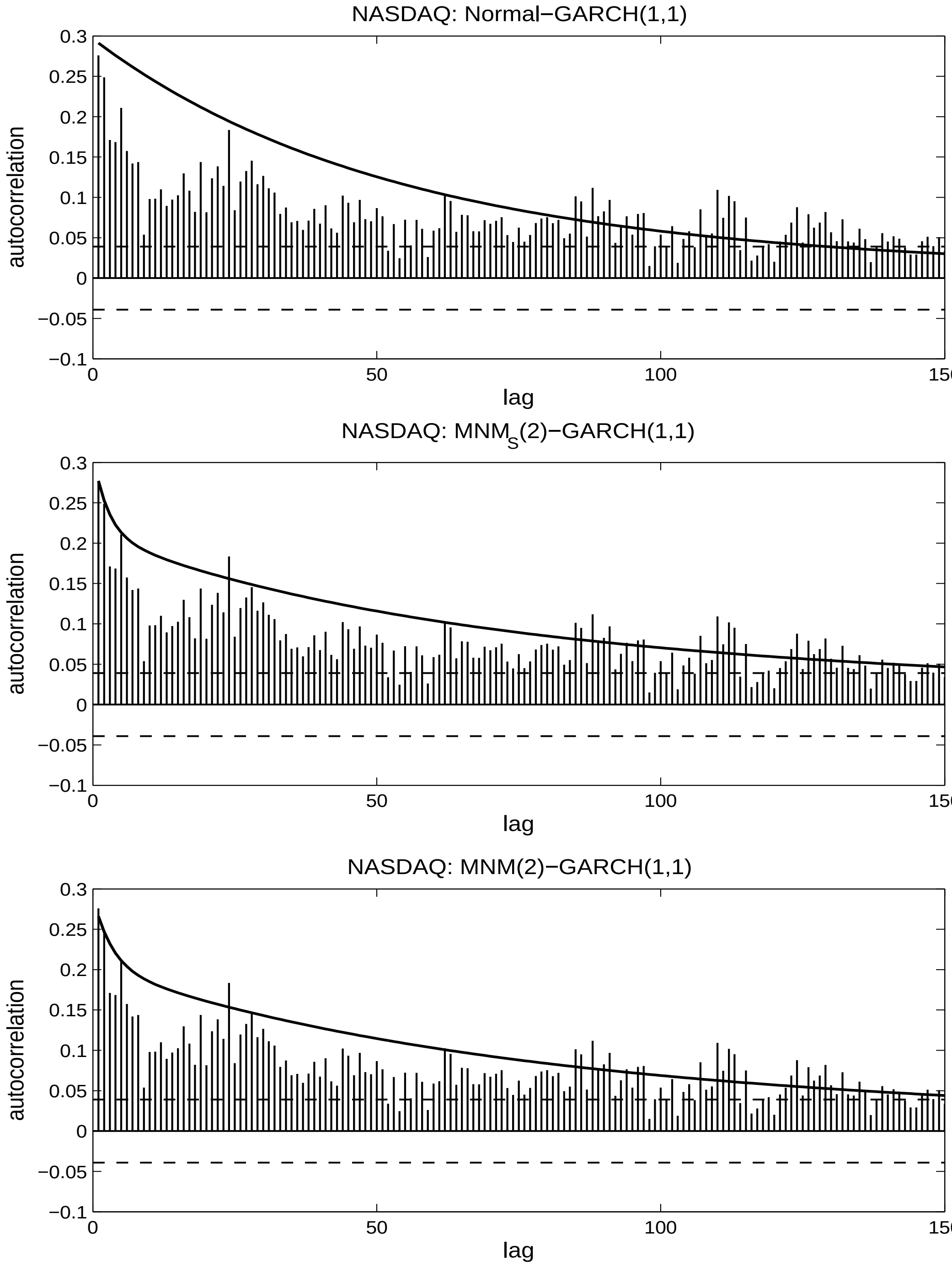

Figure 2: Shown are the empirical autocorrelations (vertical bars) of the squared (filtered) NASDAQ returns, as well as their theoretical counterparts (solid lines), as implied by the fitted Normal–GARCH(1,1) (top panel), MNM$_S$(2)–GARCH(1,1) (middle panel), and MNM(2)–GARCH(1,1) (bottom panel) models. The usual 95% asymptotic confidence intervals (dashed lines) associated with a white noise process with finite second moment are also included.

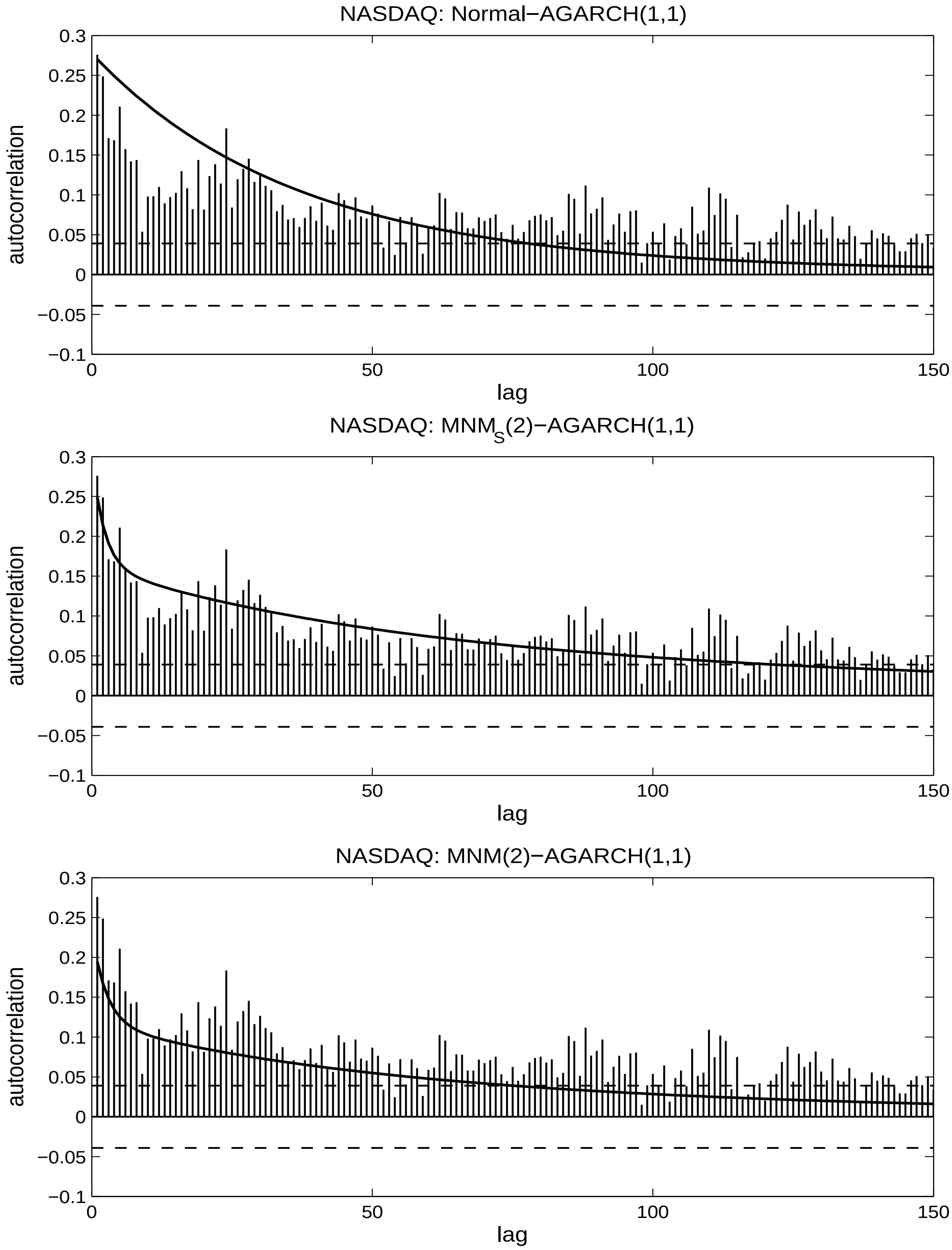

Figure 3: Shown are the empirical autocorrelations (vertical bars) of the squared (filtered) NASDAQ returns, as well as their theoretical counterparts (solid lines), as implied by the fitted Normal–AGARCH(1,1) (top panel), MNM$_S$(2)–AGARCH(1,1) (middle panel), and MNM(2)–AGARCH(1,1) (bottom panel) models. The usual 95% asymptotic confidence intervals (dashed lines) associated with a white noise process with finite second moment are also included.

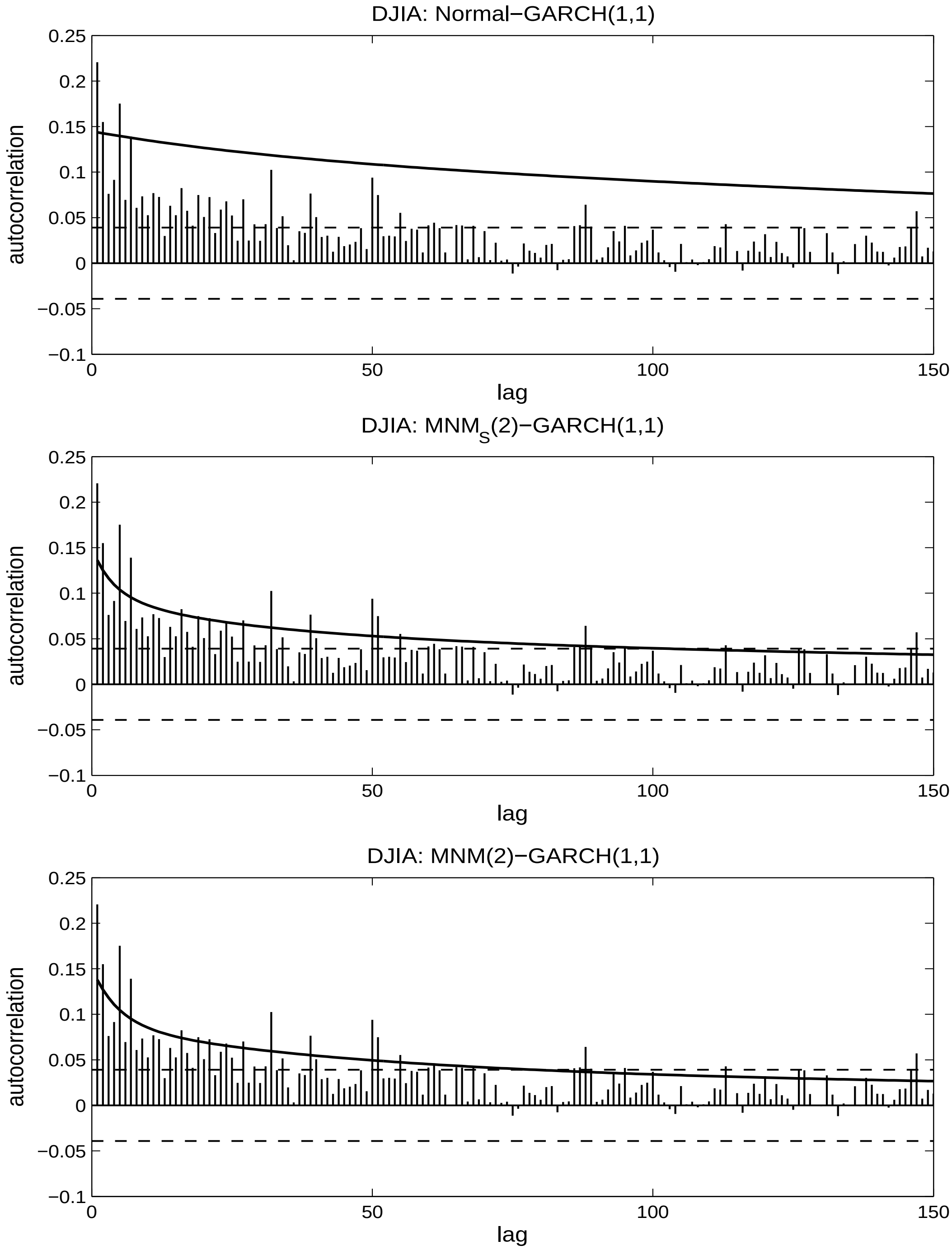

Figure 4: Shown are the empirical autocorrelations (vertical bars) of the squared (filtered) DJIA returns, as well as their theoretical counterparts (solid lines), as implied by the fitted Normal–GARCH(1,1) (top panel), MNM$_S$(2)–GARCH(1,1) (middle panel), and MNM(2)–GARCH(1,1) (bottom panel) models. The usual 95% asymptotic confidence intervals (dashed lines) associated with a white noise process with finite second moment are also included.

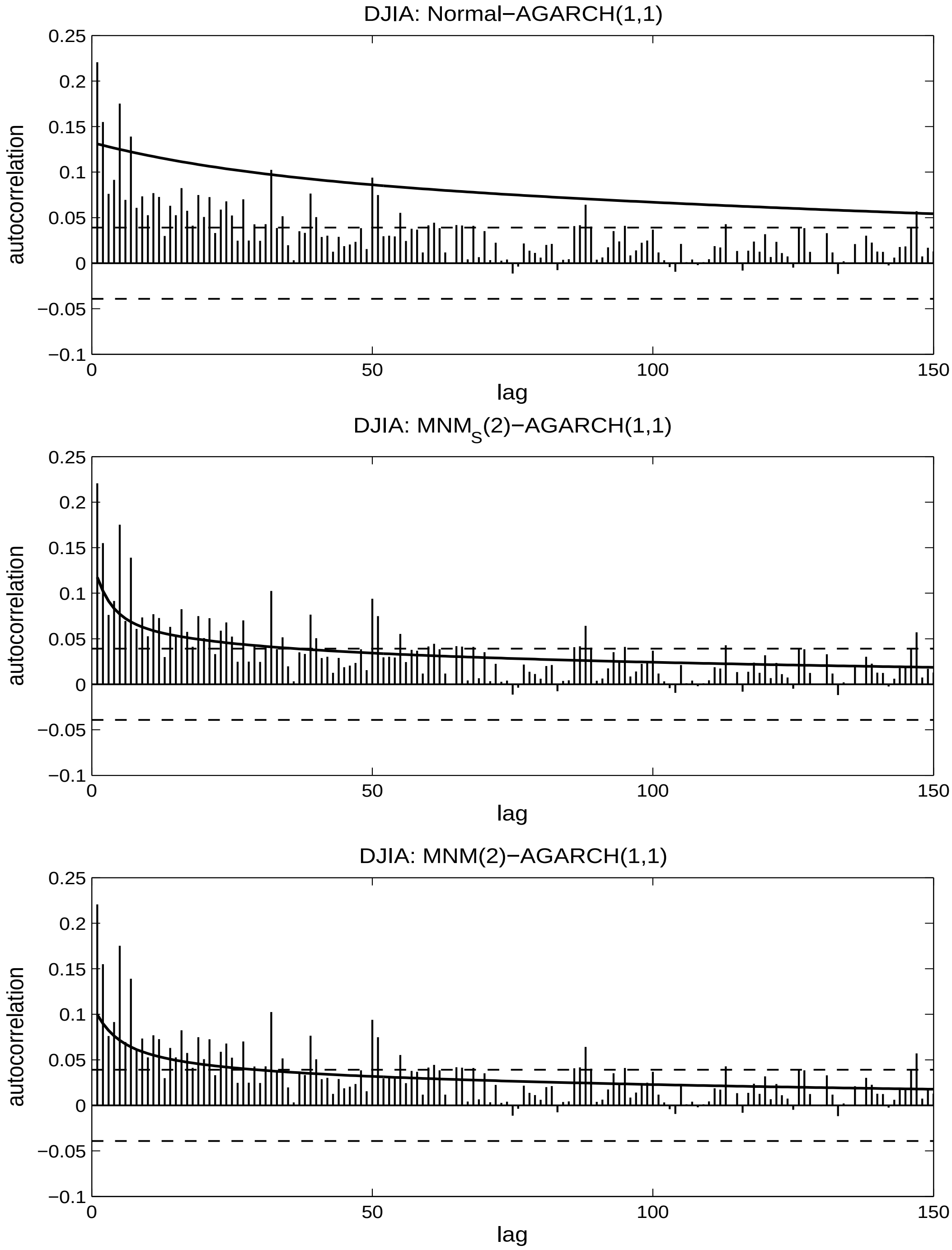

Figure 5: Shown are the empirical autocorrelations (vertical bars) of the squared (filtered) DJIA returns, as well as their theoretical counterparts (solid lines), as implied by the fitted Normal–AGARCH(1,1) (top panel), MNM$_S$(2)–AGARCH(1,1) (middle panel), and MNM(2)–AGARCH(1,1) (bottom panel) models. The usual 95% asymptotic confidence intervals (dashed lines) associated with a white noise process with finite second moment are also included.

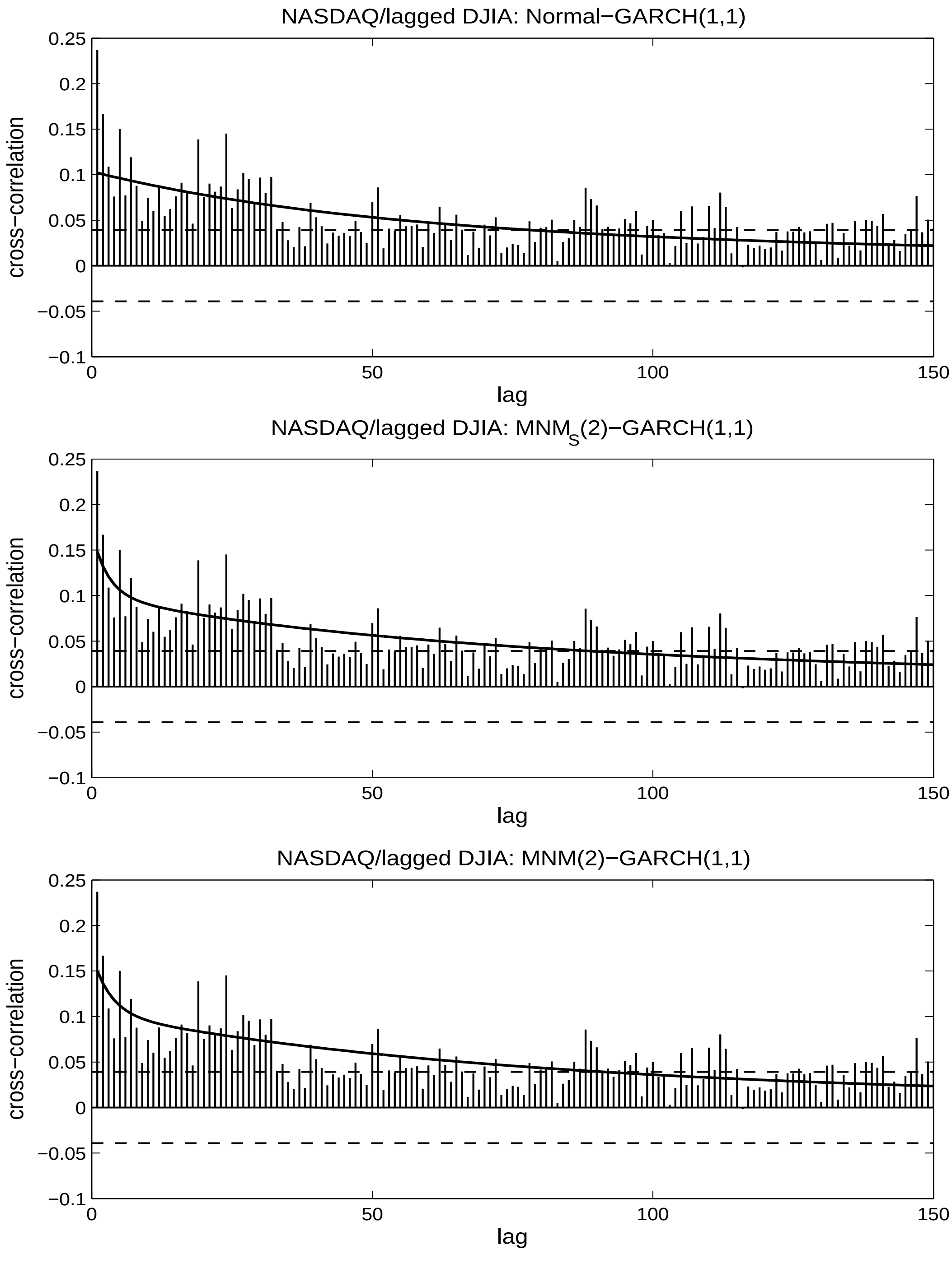

Figure 6: Shown are the empirical cross–correlations (vertical bars) of the (filtered) NASDAQ and DJIA returns, i.e., $\mathrm{Corr}(r_{1t}^2, r_{2,t-\tau}^2)$, $\tau = 1, \ldots, 150$, where r_{1t} and r_{2t} are the time–t returns of the NASDAQ and the DJIA, respectively. The solid lines represent the corresponding theoretical quantities, as implied by the fitted Normal–GARCH(1,1) (top panel), MNM$_S$(2)–GARCH(1,1) (middle panel), and MNM(2)–GARCH(1,1) (bottom panel) models. The usual 95% asymptotic confidence intervals (dashed lines) associated with a white noise process with finite second moment are also included.

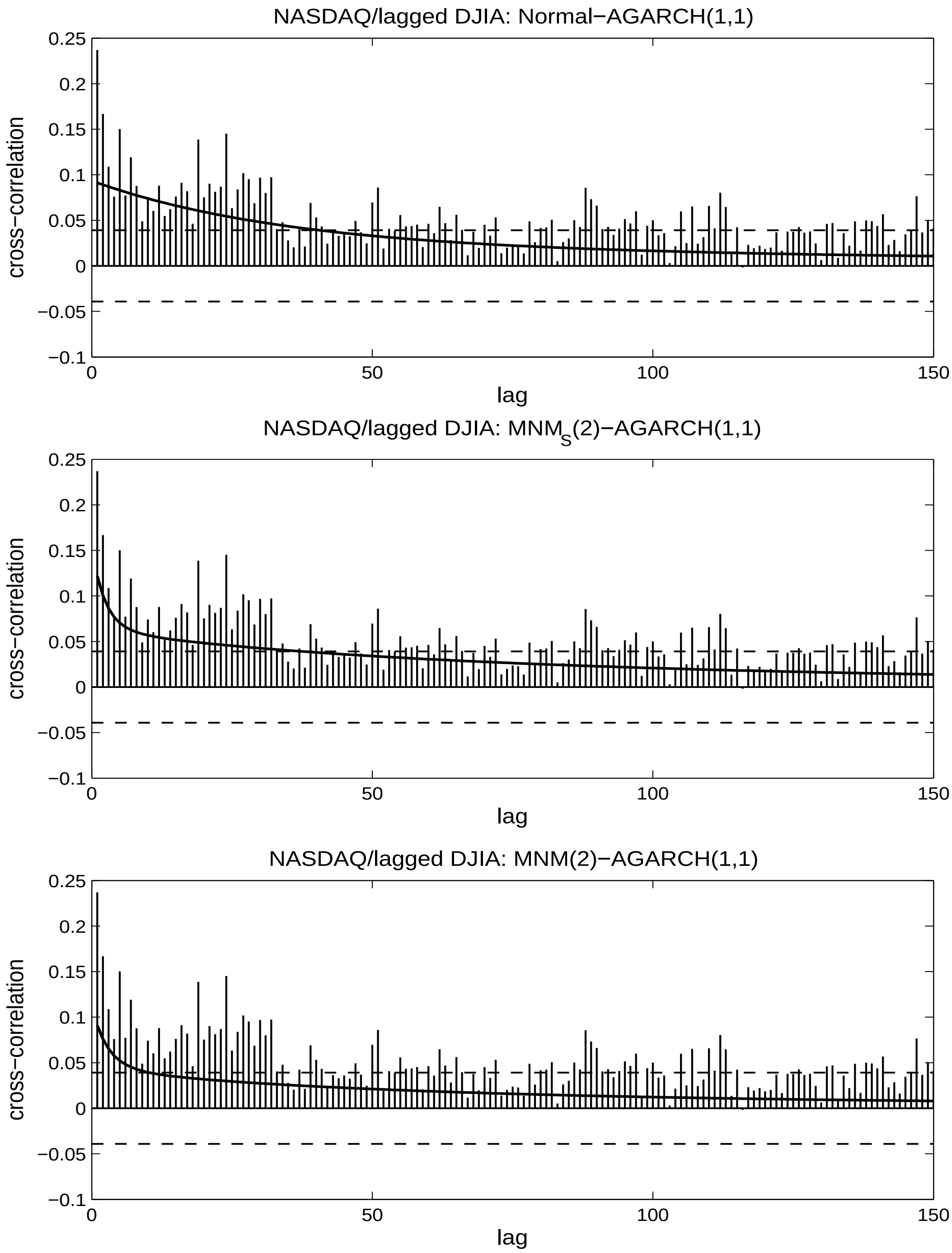

Figure 7: Shown are the empirical cross–correlations (vertical bars) of the (filtered) NASDAQ and DJIA returns, i.e., $\mathrm{Corr}(r_{1t}^2, r_{2,t-\tau}^2)$, $\tau = 1, \ldots, 150$, where r_{1t} and r_{2t} are the time–t returns of the NASDAQ and the DJIA, respectively. The solid lines represent the corresponding theoretical quantities, as implied by the fitted Normal–AGARCH(1,1) (top panel), $\mathrm{MNM}_S(2)$–AGARCH(1,1) (middle panel), and MNM(2)–AGARCH(1,1) (bottom panel) models. The usual 95% asymptotic confidence intervals (dashed lines) associated with a white noise process with finite second moment are also included.

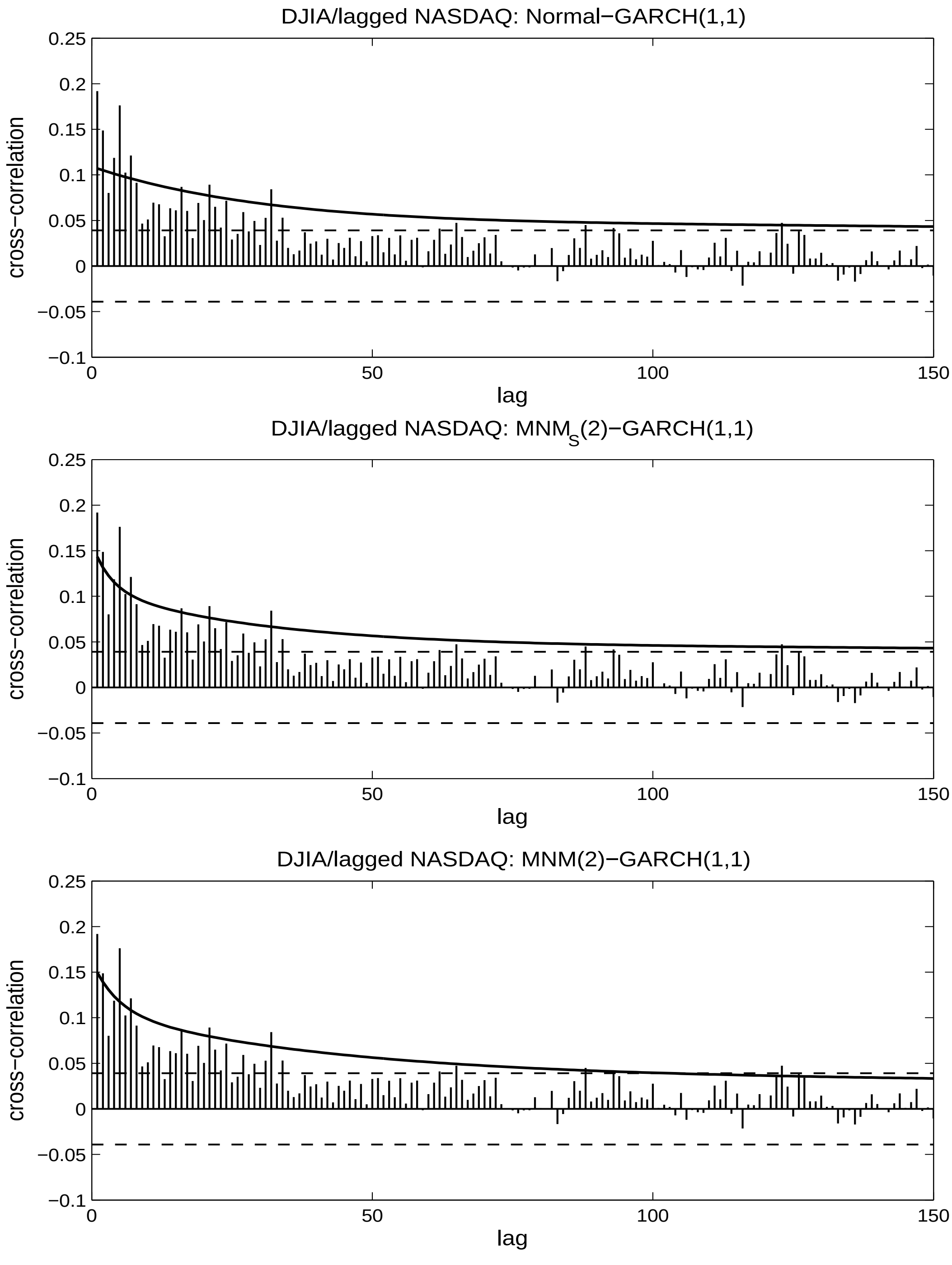

Figure 8: Shown are the empirical cross–correlations (vertical bars) of the (filtered) NASDAQ and DJIA returns, i.e., $\text{Corr}(r^2_{1,t-\tau}, r^2_{2t})$, $\tau = 1, \ldots, 150$, where r_{1t} and r_{2t} are the time–t returns of the NASDAQ and the DJIA, respectively. The solid lines represent the corresponding theoretical quantities, as implied by the fitted Normal–GARCH(1,1) (top panel), MNM$_S$(2)–GARCH(1,1) (middle panel), and MNM(2)–GARCH(1,1) (bottom panel) models. The usual 95% asymptotic confidence intervals (dashed lines) associated with a white noise process with finite second moment are also included.

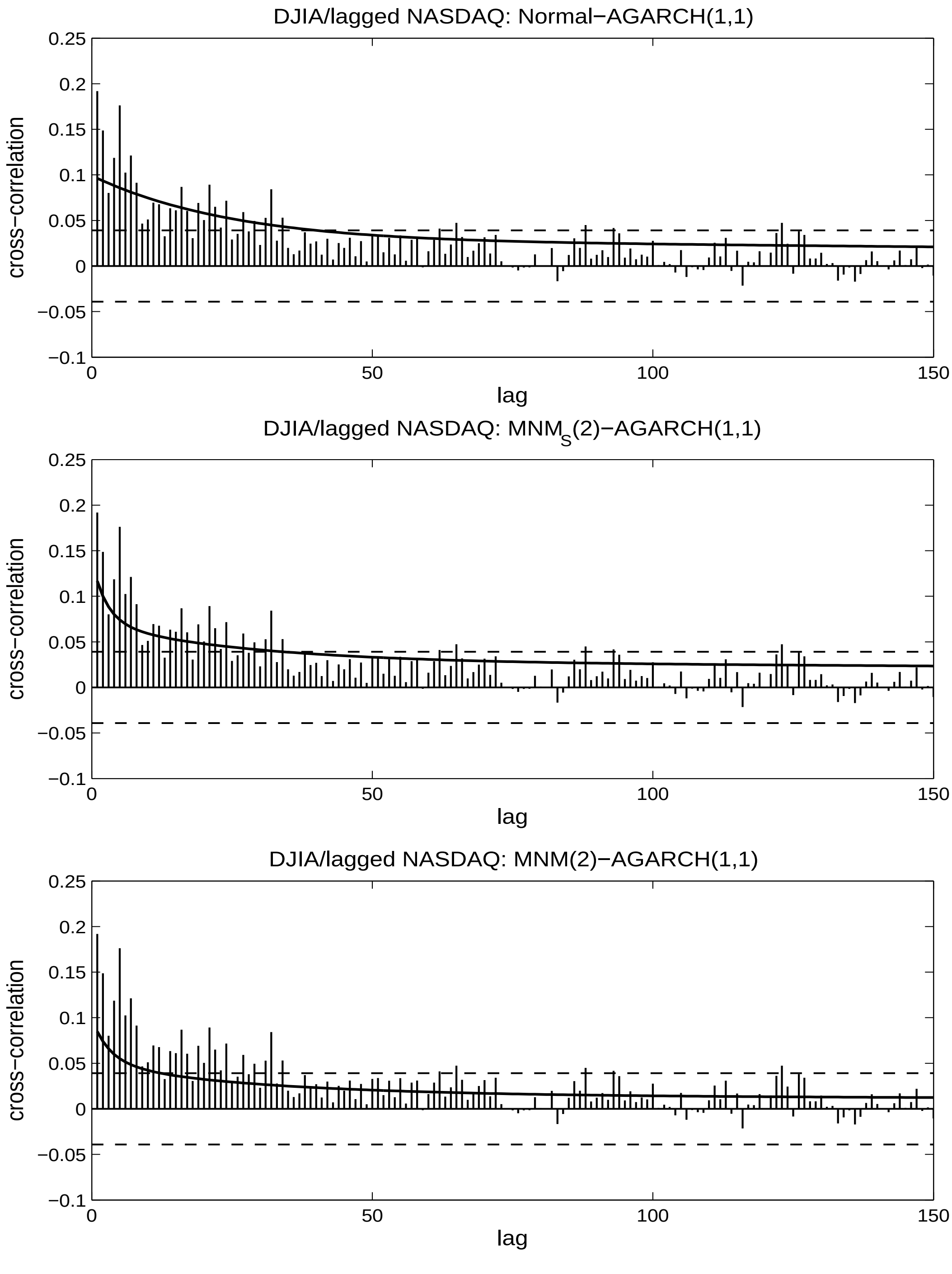

Figure 9: Shown are the empirical cross–correlations (vertical bars) of the (filtered) NASDAQ and DJIA returns, i.e., $\mathrm{Corr}(r_{1,t-\tau}^2, r_{2t}^2)$, $\tau = 1, \ldots, 150$, where r_{1t} and r_{2t} are the time–t returns of the NASDAQ and the DJIA, respectively. The solid lines represent the corresponding theoretical quantities, as implied by the fitted Normal–AGARCH(1,1) (top panel), MNM$_S$(2)–AGARCH(1,1) (middle panel), and MNM(2)–AGARCH(1,1) (bottom panel) models. The usual 95% asymptotic confidence intervals (dashed lines) associated with a white noise process with finite second moment are also included.

An application to daily returns of the NASDAQ and Dow Jones indices shows that the model captures interesting and relevant properties of the bivariate volatility process, such as regime–dependent leverage effects and conditional correlations. In view of these findings, it would be desirable to consider extensions of the model which allow for conditional forecasts of the next period's regime, which is not possible within the iid multinomial mixture approach adopted here.

A well–known disadvantage of the BEKK representation of the multivariate GARCH model is its large number of parameters, which renders estimation quite difficult when the dimension of the return series is larger than three or four. While this is true for standard GARCH models, this curse of dimensionality is even more burdensome in the mixture framework, as we have as many covariance matrices as mixture components. Thus, future research will concentrate on developing more parsimonious parameterizations for the component–specific covariance matrices. Factor structures as proposed in Alexander and Chibumba (1997), and Alexander (2001, 2002), as well as the dynamic conditional correlation models of Engle (2002), and Tse and Tsui (2002), are natural starting points to deal with this issue.

Another important topic of further research is the empirical comparison of the mixture GARCH process with other flexible multivariate GARCH models, such as those of Bauwens and Laurent (2005), and Aas, Haff, and Dimakos (2006), who employ a multivariate skewed t and the multivariate normal inverse Gaussian distribution, respectively.

Appendix

In the Appendix, we derive the conditions for the moments of the MNM–GARCH model. We also provide expressions for these moments and the autocorrelation structure of the process.

A Notation

To conveniently write down the unconditional moments of the multivariate normal mixture GARCH model, use of several patterned matrices is rather advantageous, and we define them here. A detailed discussion of (as well as explicit expressions for) these matrices can be found in Magnus (1988).[9] The first of these matrices is the *commutation matrix*, K_{mn}, which is the $mn \times mn$ matrix with the property that $K_{mn}\text{vec}(A) = \text{vec}(A')$ for every $m \times n$ matrix A. We will use the fact that the commutation matrix allows us to transform the vec of a Kronecker

[9] Chapter 1 of Magnus and Neudecker (1999) also provides useful information, as do the appendices in Hafner (2003) and Lütkepohl (2005).

product into the kronecker product of the vecs (Magnus, 1988, Theorem 3.6). More precisely, for an $m \times n$ matrix A and an $p \times q$ matrix B, it is true that

$$\mathrm{vec}(A \otimes B) = (I_n \otimes K_{qm} \otimes I_p)(\mathrm{vec}A \otimes \mathrm{vec}B). \tag{A.1}$$

The *elimination matrix*, L_n, is the $n(n+1)/2 \times n^2$ matrix that takes away the redundant elements of a symmetric $n \times n$ matrix, i.e., for every $n \times n$ matrix A, we have $L_n \mathrm{vec}(A) = \mathrm{vech}(A)$. In contrast, the *duplication matrix*, D_n, is the $n^2 \times n(n+1)/2$ matrix with the property that $D_n \mathrm{vech}(A) = \mathrm{vec}(A)$ for every symmetric $n \times n$ matrix A. Its Moore–Penrose inverse, D_n^+, is given by $D_n^+ = (D_n' D_n)^{-1} D_n'$ (Magnus, 1988, Theorem 4.1).

To compactify the expressions for the moments of our model, we will also made extensive use of the matrix $N_n = (I_{n^2} + K_{nn})/2$, which is discussed in Section 3.10 of Magnus (1988), and which has the property that, for every $n \times n$ matrix A,

$$2N_n \mathrm{vec}(A) = \mathrm{vec}(A + A'). \tag{A.2}$$

Note that the matrix D_n^+ has a similar property. Namely, because of $D_n^+ = L_n N_n$ (Magnus, 1988, p. 80), we have

$$2D_n^+ \mathrm{vec}(A) = \mathrm{vech}(A + A'). \tag{A.3}$$

B The Third and Fourth Moments of an Asymmetric Multivariate Normal Mixture Distribution

In this Appendix, we provide convenient expressions for the expectations of $\mathrm{vec}[\mathrm{vech}(xx')x']$ and $\mathrm{vec}[\mathrm{vech}(xx')\mathrm{vech}(xx')']$, when x has a multivariate normal mixture distribution with (possibly) nonzero means, as defined in (1) and (2). These expressions will be useful for computing the unconditional moments of the multivariate mixed normal GARCH process in Appendices C and D.

To derive the expressions given in this Appendix, we draw on results of Magnus and Neudecker (1979), Balestra and Holly (1990), and Hafner (2003). We state the central results as Lemmas 2–4 for the third, and Lemmas 5–8 for the fourth moment. Details of the derivations are presented only for the third moment, because those for the fourth moment are similar.[10]

B.1 The Third Moment

To find an expression for $\mathrm{vec}[\mathrm{vech}(xx')x']$, which is needed due to the inclusion of the leverage terms, we make use of a formula of Balestra and Holly (1990) which we state as Lemma 2.

[10] Detailed derivations are available on request from the authors.

Lemma 2 *(Balestra and Holly, 1990) For an M–dimensional random vector x, which is normally distributed with mean μ and covariance matrix H, we have*

$$E[(x \otimes x)x'] = vec(H)\mu' + 2N_M(\mu \otimes H) + (\mu \otimes \mu)\mu'. \tag{B.4}$$

We are interested in $\mathrm{E}\{\mathrm{vec}[\mathrm{vech}(xx')x']\}$ as a linear function in h, where $h = \mathrm{vech}(H)$. Such an expression is provided next.

Lemma 3 *For an M–dimensional random vector x, which is normally distributed with mean μ and covariance matrix H, we have*

$$E\{vec[vech(xx')x']\} = (I_M \otimes L_M)[\tilde{G}_M(\mu \otimes D_M)h + \mu \otimes \mu \otimes \mu], \tag{B.5}$$

where $h = vech(H)$, and

$$\tilde{G}_M = I_{M^3} + 2(I_M \otimes N_M)(K_{MM} \otimes I_M). \tag{B.6}$$

Proof. By Lemma 2, and using $\mathrm{vec}(ABC) = (C' \otimes A)\mathrm{vec}(B)$, we have

$$\begin{aligned} \mathrm{E}\{\mathrm{vec}[\mathrm{vech}(xx')x']\} &= \mathrm{E}\{\mathrm{vec}[L_M\mathrm{vec}(xx')x']\} = (I_M \otimes L_M)\mathrm{E}\{\mathrm{vec}[(x \otimes x)x']\} \\ &= (I_M \otimes L_M)\mathrm{vec}[\mathrm{vec}(H)\mu' + 2N_M(\mu \otimes H) + (\mu \otimes \mu)\mu']. \end{aligned}$$

Furthermore, $\mathrm{vec}[2N_M(\mu \otimes H)] = 2(I_M \otimes N_M)\mathrm{vec}(\mu \otimes H)$, and (A.1) implies that $\mathrm{vec}(\mu \otimes H) = (K_{MM} \otimes I_M)(\mu \otimes \mathrm{vec}(H))$. Finally, as $y \otimes x = \mathrm{vec}(xy')$ for vectors x and y, we have $\mu \otimes \mathrm{vec}(H) = \mathrm{vec}[\mathrm{vec}(H)\mu'] = \mathrm{vec}(D_M h\mu') = (\mu \otimes D_M)h$, and thus (B.5). ■

Next, we consider the case of a normal mixture distribution.

Lemma 4 *Assume that $x \sim MNM(\lambda_1, \ldots, \lambda_k, \mu_1, \ldots, \mu_k, H_1, \ldots, H_k)$. Let $\lambda = (\lambda_1, \ldots, \lambda_k)'$, $\Lambda = diag(\lambda)$; $h_j = vech(H_j)$, $j = 1, \ldots, k$; $h = (h_1', \ldots, h_k')'$; $\Upsilon = (\mu_1, \ldots, \mu_k)$; $\mu = vec(\Upsilon) = (\mu_1', \ldots, \mu_k')'$; $\tilde{\mu}_j = vech(\mu_j\mu_j')$, $j = 1, \ldots, k$; $\tilde{\Upsilon} = (\tilde{\mu}_1, \ldots, \tilde{\mu}_k)$; and $\tilde{\mu} = vec(\tilde{\Upsilon}) = (\tilde{\mu}_1', \ldots, \tilde{\mu}_k')'$. Then,*

$$\begin{aligned} &E\{vec[vech(xx')x']\} \qquad\qquad (B.7) \\ &\quad = (I_M \otimes L_M)\tilde{G}_M(\Upsilon\Lambda \otimes D_M)h + (I_M \otimes vec(\Lambda)' \otimes I_N)(K_{Mk} \otimes I_{kN})vec(\tilde{\mu}\mu'), \end{aligned}$$

where $N = M(M+1)/2$, and $\tilde{G}_M$ is defined in (B.6).

Proof. Lemma 4 follows from the fact that the third moment of the mixture is just the weighted average of the component–specific moments as given in (B.5), i.e., for x mixed normal as defined in Lemma 4, we have

$$\mathrm{E}\{\mathrm{vec}[\mathrm{vech}(xx')x']\} = (I_M \otimes L_M)\left\{\tilde{G}_M \sum_{j=1}^{k} \lambda_j(\mu_j \otimes D_M)h_j + \sum_{j=1}^{k} \lambda_j(\mu_j \otimes \mu_j \otimes \mu_j)\right\}. \tag{B.8}$$

Let e_j be the jth unit vector in $\mathbb{R}^k$. Then, for the first sum on the right–hand side of (B.8), we have that

$$\begin{aligned}\sum_{j=1}^{k}\lambda_j(\mu_j\otimes D_M)h_j &= \left\{\sum_{j=1}^{k}\lambda_j(e_j'\otimes\mu_j\otimes D_M)\right\}h=\left\{\left(\sum_{j=1}^{k}\lambda_j\mu_j e_j'\right)\otimes D_M\right\}h\\ &= (\Upsilon\Lambda\otimes D_M)h, \qquad\qquad (B.9)\end{aligned}$$

where, in the last equation of the first line in (B.9), we have used that $y'\otimes x = xy'$. For the second sum on the right–hand side of (B.8), we find

$$\begin{aligned}\sum_j\lambda_j(\mu_j\otimes\mu_j\otimes\mu_j) &= \sum_j\lambda_j\text{vec}[(\mu_j\otimes\mu_j)\mu_j'] = (I_M\otimes D_M)\sum_j\lambda_j\text{vec}(\tilde{\mu}_j\mu_j') \qquad (B.10)\\ &= (I_M\otimes D_M)\sum_j\lambda_j\text{vec}[(e_j'\otimes I_N)(\tilde{\mu}\mu')(e_j\otimes I_M)]\\ &= (I_M\otimes D_M)\sum_j\lambda_j(e_j'\otimes I_M\otimes e_j'\otimes I_N)\text{vec}(\tilde{\mu}\mu')\\ &= (I_M\otimes D_M)\sum_j\lambda_j(I_M\otimes e_j'\otimes e_j'\otimes I_N)(K_{Mk}\otimes I_{kN})\text{vec}(\tilde{\mu}\mu')\\ &= (I_M\otimes D_M)\sum_j\lambda_j(I_M\otimes\text{vec}(e_je_j')'\otimes I_N)(K_{Mk}\otimes I_{kN})\text{vec}(\tilde{\mu}\mu')\\ &= (I_M\otimes D_M)(I_M\otimes\text{vec}(\Lambda)'\otimes I_N)(K_{Mk}\otimes I_{kN})\text{vec}(\tilde{\mu}\mu'),\end{aligned}$$

where we have used the identity $(A\otimes b')K_{np} = b'\otimes A$ for $m\times n$ matrix A and $p\times 1$ vector b (Magnus, 1988, p. 36). Finally, because $(A\otimes B)(C\otimes D) = (AC)\otimes(BD)$ if AC and BD exist, we have $(I_M\otimes L_M)(I_M\otimes D_M) = (I_M\otimes L_MD_M)$, and, by Theorem 5.5 of Magnus (1988), $L_MD_M = I_N$, $N = M(M+1)/2$, so we get (B.7). ■

B.2 The Fourth Moment

For the fourth moment, we build on results of Magnus and Neudecker (1979) and Hafner (2003) which we state as Lemmas 5 and 6, respectively.

Lemma 5 *(Magnus and Neudecker, 1979, Theorem 4.3) For an M–dimensional random vector x, which is normally distributed with mean μ and covariance matrix H, we have*[11]

$$\begin{aligned}E[(x\otimes x)(x\otimes x)'] &= 2D_MD_M^+(H\otimes H)+vec(H)vec(H)' \qquad (B.11)\\ &\quad +2D_MD_M^+(H\otimes\mu\mu'+\mu\mu'\otimes H)\\ &\quad +vec(H)vec(\mu\mu')'+vec(\mu\mu')vec(H)'+vec(\mu\mu')vec(\mu\mu')'.\end{aligned}$$

[11] Magnus and Neudecker (1979) state the result in terms of the matrix N_n defined in Appendix A. In fact, by Theorem 4.2 of Magnus (1988), we have $N_n = D_nD_n^+$. Here, the representation in terms of $D_nD_n^+$ is preferable because this simplifies some of the expressions to be presented below.

For the result in Lemma 5 and generalizations, see also Magnus (1988, Ch. 10) and Ghazal and Neudecker (2000).

We are interested in E[vech(xx')vech(xx')$'$]. Using the identity vec(xx') $= x \otimes x$ and the definition of the elimination matrix L_M, this can be written as $L_M \mathrm{E}[(x \otimes x)(x \otimes x)']L'_M$, which is a simple transformation of (B.11). The case of a normal distribution with zero mean was considered by Hafner (2003).[12]

Lemma 6 *(Hafner, 2003, Theorem 1) For an M–dimensional normally distributed random vector x with zero mean and covariance matrix H, we have*

$$vec\{E[vech(xx')vech(xx')']\} = G_M vec(hh'), \tag{B.12}$$

where $h = vech(H)$, and

$$G_M = 2(L_M \otimes D_M^+)(I_M \otimes K_{MM} \otimes I_M)(D_M \otimes D_M) + I_{N^2}, \tag{B.13}$$

and $N := M(M+1)/2$ is the number of independent elements in H.

Our first step is to generalize (B.12) to the case of nonzero means, i.e., to consider the terms in the second and third line of (B.11).

Lemma 7 *For an M–dimensional normally distributed random vector x with mean μ and covariance matrix H, we have*

$$vec\{E[vech(xx')vech(xx')']\} = G_M vec(hh') + 2G_M N_N(\tilde{\mu} \otimes I_N)h + vec(\tilde{\mu}\tilde{\mu}'), \tag{B.14}$$

where G_M is defined in (B.13), $h = vech(H)$, $\tilde{\mu} = vech(\mu\mu')$, and $N = M(M+1)/2$.

The proof of Lemma 7 can be carried out along similar lines as the proof of Theorem 1 in Hafner (2003). The case of a multivariate normal mixture distribution is considered next. We make use of the notation introduced in Lemma 4.

Lemma 8 *Assume that $x \sim MNM(\lambda_1, \ldots, \lambda_k, \mu_1, \ldots, \mu_k, H_1, \ldots, H_k)$. Then,*

$$\begin{aligned} & vec\{E[vech(xx')vech(xx')']\} \\ & \quad = G_M(I_N \otimes vec(\Lambda)' \otimes I_N)(K_{Nk} \otimes I_{kN})vec(hh') + 2G_M N_N(\tilde{\Upsilon}\Lambda \otimes I_N)h \\ & \qquad + (I_N \otimes vec(\Lambda)' \otimes I_N)(K_{Nk} \otimes I_{kN})vec(\tilde{\mu}\tilde{\mu}'). \end{aligned} \tag{B.15}$$

[12] Actually, Hafner (2003) considered the more general class of spherical distributions which includes the normal as a special case.

Lemma 8 is obtained by combining the results of Lemma 7 with the fact that the fourth moment of the mixture is just the weighted average of the component–specific moments as given in (B.14), quite similar to equation (B.8) for the third moment, and by using arguments similar to those in the derivation of Lemma 4. For example, to show that

$$\sum_{j=1}^{k} \lambda_j \mathrm{vec}(h_j h_j') = (I_N \otimes \mathrm{vec}(\Lambda)' \otimes I_N)(K_{Nk} \otimes I_{Nk})\mathrm{vec}(hh'), \tag{B.16}$$

we essentially repeat the argument in (B.10).

C The Moments of the MNM(k)–AGARCH(1,1) Model

In this Appendix, we use the results of Appendix B to derive the unconditional second and fourth moments of the asymmetric multivariate mixed normal GARCH(1,1) model as given in equation (11), as well as the conditions for their existence. Using the results of Balestra and Holly (1990), higher–order moments could in principle also be derived, but the resulting expressions become unmanageable even for the central normal distribution, as the number of terms to be evaluated is explosive as the order increases. Thus, in view of the fact that such higher moments are of minor interest in applications, we concentrate on the second and the fourth moment.[13]

C.1 Moment Conditions

We will use the notation introduced in Section 2 and Lemmas 4 and 8. Also, as defined in (12), $\rho(A)$ denotes the largest eigenvalue in modulus of a square matrix A.

Define $W_t = (h_t', \mathrm{vec}(h_t h_t')')'$, and consider the expectation of W_t at time $t-2$, i.e., $\mathrm{E}(W_t|\Psi_{t-2})$. Clearly $\mathrm{E}(\eta_{t-1}|\Psi_{t-2}) = (\lambda' \otimes I_N)(h_{t-1} + \tilde{\mu})$, so that[14]

$$\mathrm{E}(h_t|\Psi_{t-2}) = \tilde{A}_0 + A_1(\lambda' \otimes I_N)\tilde{\mu} + (\lambda \otimes A_1 + B_1)h_{t-1}.$$

The conditional expectation of $\mathrm{vec}(h_t h_t')$ can be greatly simplified by extensively using the matrix N_n, and in particular its basic property (A.2). In addition, we will frequently use the identities $\mathrm{vec}(xy') = y \otimes x$ and $\mathrm{vec}(ABC) = (C' \otimes A)\mathrm{vec}(B)$. Thus,

$$\begin{aligned}
\mathrm{vec}(h_t h_t') \quad = \quad & \tilde{A}_0 \otimes \tilde{A}_0 + 2N_{kN}\mathrm{vec}[\tilde{A}_0(\eta_{t-1}'A_1' + h_{t-1}'B_1')] + 2N_{kN}\mathrm{vec}(A_1\eta_{t-1}h_{t-1}'B_1') \\
& + (A_1 \otimes A_1)\mathrm{vec}(\eta_{t-1}\eta_{t-1}') + (B_1 \otimes B_1)\mathrm{vec}(h_{t-1}h_{t-1}') \\
& + \mathrm{vec}(\Theta_1\epsilon_{t-1}\epsilon_{t-1}'\Theta_1') - 2N_{kN}\mathrm{vec}[(\tilde{A}_0 + A_1\eta_{t-1} + B_1 h_{t-1})\epsilon_{t-1}'\Theta_1'].
\end{aligned} \tag{C.17}$$

[13] For the univariate case, a condition for the existence of arbitrary integer even moments in given in Haas, Mittnik and Paolella (2004a).

[14] Note that $A_1(\lambda' \otimes I_N) = (1 \otimes A_1)(\lambda' \otimes I_N) = \lambda' \otimes A_1$.

Let us evaluate the conditional expectations of the components of (C.17). Observe that

$$\begin{aligned}
\mathrm{E}\{\mathrm{vec}(\tilde{A}_0\eta'_{t-1}A'_1)|\Psi_{t-2}\} &= \mathrm{vec}\{\tilde{A}_0(h'_{t-1}+\tilde{\mu}')(\lambda\otimes I_N)A'_1\} \\
&= [A_1(\lambda'\otimes I_N)\otimes\tilde{A}_0](h_{t-1}+\tilde{\mu}) \\
&= (\lambda'\otimes A_1\otimes\tilde{A}_0)(h_{t-1}+\tilde{\mu}),
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{E}[\mathrm{vec}(A_1\eta_{t-1}h'_{t-1}B'_1)|\Psi_{t-2}] &= \mathrm{vec}[A_1(\lambda'\otimes I_N)(h_{t-1}+\tilde{\mu})h'_{t-1}B'_1] \\
&= (B_1\otimes\lambda'\otimes A_1)\mathrm{vec}(h_{t-1}h'_{t-1}) + (B_1\otimes\lambda'\otimes A_1)\mathrm{vec}(\tilde{\mu}h'_{t-1}) \\
&= (B_1\otimes\lambda'\otimes A_1)\mathrm{vec}(h_{t-1}h'_{t-1}) + (B_1\otimes\lambda'\otimes A_1)(I_{kN}\otimes\tilde{\mu})h_{t-1} \\
&= (B_1\otimes\lambda'\otimes A_1)\mathrm{vec}(h_{t-1}h'_{t-1}) + [B_1\otimes(\lambda'\otimes A_1)\tilde{\mu}]h_{t-1}.
\end{aligned}$$

The expectation of $(A_1\otimes A_1)\mathrm{vec}(\eta_{t-1}\eta'_{t-1})$, given Ψ_{t-2}, follows from Lemma 8. It remains to consider those terms of (C.17) which involve ϵ_{t-1}. First, note that $\mathrm{E}(h_{t-1}\epsilon'_{t-1}|\Psi_{t-2}) = h_{t-1}\mathrm{E}(\epsilon'_{t-1}|\Psi_{t-2}) = 0$. Thus, we have two nonzero terms. The first is

$$\mathrm{E}[\mathrm{vec}(\Theta_1\epsilon_{t-1}\epsilon'_{t-1}\Theta'_1)|\Psi_{t-2}] = (\Theta_1\otimes\Theta_1)D_M(\lambda'\otimes I_N)(h_{t-1}+\tilde{\mu}),$$

and the second, using Lemma 4,

$$\begin{aligned}
\mathrm{E}[\mathrm{vec}(A_1\eta_{t-1}\epsilon'_{t-1}\Theta'_1)|\Psi_{t-2}] &= (\Theta_1\otimes A_1)\Big[(I_M\otimes L_M)\tilde{G}_M(\Upsilon\Lambda\otimes D_M)h_{t-1} \\
&\quad +(I_M\otimes\mathrm{vec}(\Lambda)'\otimes I_N)(K_{Mk}\otimes I_{kN})\mathrm{vec}(\tilde{\mu}\mu')\Big].
\end{aligned}$$

Next, define

$$d = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}, \quad C = \begin{pmatrix} C_{11} & 0_{kN\times kN} \\ C_{21} & C_{22} \end{pmatrix},$$

where

$$\begin{aligned}
d_1 &= \tilde{A}_0 + A_1(\lambda'\otimes I_N)\tilde{\mu} \\
d_2 &= \tilde{A}_0\otimes\tilde{A}_0 + 2N_{kN}(\lambda'\otimes A_1\otimes\tilde{A}_0)\tilde{\mu} + (A_1\otimes A_1)(I_N\otimes\mathrm{vec}(\Lambda)'\otimes I_N)(K_{Nk}\otimes I_{kN})\mathrm{vec}(\tilde{\mu}\tilde{\mu}') \\
&\quad +(\Theta_1\otimes\Theta_1)D_M(\lambda'\otimes I_N)\tilde{\mu} - 2N_{kN}(\Theta_1\otimes A_1)(I_M\otimes\mathrm{vec}(\Lambda)'\otimes I_N)(K_{Mk}\otimes I_{kN})\mathrm{vec}(\tilde{\mu}\mu'), \\
C_{11} &= \lambda'\otimes A_1 + B_1, \\
C_{21} &= 2N_{kN}(\lambda'\otimes A_1 + B_1)\otimes\tilde{A}_0 + 2N_{kN}[B_1\otimes(\lambda'\otimes A_1)\tilde{\mu}] + 2(A_1\otimes A_1)G_M N_N(\tilde{\Upsilon}\Lambda\otimes I_N) \\
&\quad +(\Theta_1\otimes\Theta_1)D_M(\lambda'\otimes I_N) - 2N_{kN}(\Theta_1\otimes A_1)(I_M\otimes L_M)\tilde{G}_M(\Upsilon\Lambda\otimes D_M), \\
C_{22} &= (A_1\otimes A_1)G_M(I_N\otimes\mathrm{vec}(\Lambda)'\otimes I_N)(K_{Nk}\otimes I_{kN}) + 2N_{kN}(B_1\otimes\lambda'\otimes A_1) + B_1\otimes B_1.
\end{aligned}$$

From the preceding analysis it is clear that

$$\mathrm{E}(W_t|\Psi_{t-2}) = d + CW_{t-1},$$

and, by iteration,

$$\mathrm{E}(W_t|\Psi_{t-\tau-1}) = \sum_{i=0}^{\tau-1} C^i d + C^\tau W_{t-\tau}. \tag{C.18}$$

From the block–triangular structure of C, we have, from (C.18), that

$$\mathrm{E}(h_t|\Psi_{t-\tau-1}) = \sum_{i=0}^{\tau-1} C_{11}^i d_1 + C_{11}^\tau h_{t-\tau}. \tag{C.19}$$

Thus, as we have assumed that the process starts indefinitely far in the past with finite fourth moments, the unconditional expectation $\mathrm{E}(h_t)$ exists and is given by the limit as $\tau \to \infty$, i.e.,

$$\mathrm{E}(h_t) = \lim_{\tau\to\infty} \mathrm{E}(h_t|\Psi_{t-\tau-1}) = \sum_{i=0}^{\infty} C_{11}^i d_1 = (I_{kN} - C_{11})^{-1} d_1$$

if and only if $\rho(C_{11}) < 1$, as stated in (13). By the same line of reasoning, $\mathrm{E}(W_t)$ exists and is given by $(I - C)^{-1}d$ if and only if $\mathrm{E}(h_t)$ exists and $\rho(C_{22}) < 1$, as claimed in (14).

Example 1 *Note that the expressions for the elements of d and C defined above simplify considerably if all mixture components have zero means, which may be appropriate when the (conditional) distribution of the returns under study exhibits leptokurtosis but no asymmetries. In particular, in this case the only extra term due to the leverage effects is $(\Theta_1 \otimes \Theta_1) D_M (\lambda' \otimes I_N)$ in the lower left block of C, i.e., C_{21}. Moreover, in the univariate, single–component case we get the QGARCH(1,1) model of Sentana (1995); and the unconditional fourth moment, in obvious notation, is*

$$E(\eta_t^2) = E(\epsilon_t^4) = \frac{3\alpha_0[\alpha_0(1+\alpha_1+\beta_1)+\theta_1^2]}{(1-\alpha_1-\beta_1)(1-3\alpha_1^2-2\alpha_1\beta_1-\beta_1^2)}, \tag{C.20}$$

which was given by Sentana (1995). This differs from the fourth moment of the standard GARCH(1,1) model of Bollerslev (1986) only by the extra θ_1^2 in the numerator of (C.20), which shows that the QGARCH(1,1) model has a greater fourth moment than its standard GARCH(1,1) counterpart. The variance, however, is $E(\epsilon_t^2) = \alpha_0/(1-\alpha_1-\beta_1)$, as in the standard GARCH(1,1), and the kurtosis is given by

$$\kappa = \frac{E(\epsilon_t^4)}{E^2(\epsilon_t^2)} = 3\frac{1-(\alpha_1+\beta_1)^2 + (1-\alpha_1-\beta_1)\theta^2/\alpha_0}{1-3\alpha_1^2-2\alpha_1\beta_1-\beta_1^2} = 3\frac{1-(\alpha_1+\beta_1)^2+\theta^2/E(\epsilon_t^2)}{1-3\alpha_1^2-2\alpha_1\beta_1-\beta_1^2},$$

which depends on the scale parameter α_0. Due to the factor $\theta^2/E(\epsilon_t^2)$, the unconditional kurtosis of the QGARCH(1,1) model exceeds that of the standard GARCH(1,1) process. However, as stressed by Carnero, Peña, and Ruiz (2004), in applications, θ^2 is usually small relative to $E(\epsilon_t^2)$, so that the difference is rather small.

C.2 Autocovariance Function of the Squares

To find the autocovariance matrices, i.e., $\Gamma_\tau = \mathrm{E}(\eta_t\eta'_{t-\tau}) - \mathrm{E}(\eta_t)\mathrm{E}(\eta_t)'$, we first note that (C.19) in Appendix C.1 implies

$$\mathrm{E}(h_t|\Psi_{t-\tau}) = \sum_{i=0}^{\tau-2} C_{11}^i d_i + C_{11}^{\tau-1} h_{t-\tau+1} = \mathrm{E}(h_t) + C_{11}^{\tau-1}[h_{t-\tau+1} - \mathrm{E}(h_t)].$$

Hence,

$$\begin{aligned}
\mathrm{E}(\eta_t\eta'_{t-\tau}) &= \mathrm{E}[\mathrm{E}(\eta_t|\Psi_{t-\tau})\eta'_{t-\tau}] \\
&= \mathrm{E}\{(\lambda' \otimes I_N)[\mathrm{E}(h_t|\Psi_{t-\tau}) + \tilde{\mu}]\eta'_{t-\tau}\} \\
&= (\lambda' \otimes I_N)\mathrm{E}\{[\mathrm{E}(h_t) + \tilde{\mu} + C_{11}^{\tau-1}(h_{t-\tau+1} - \mathrm{E}(h_t))]\eta'_{t-\tau}\} \\
&= \mathrm{E}(\eta_t)\mathrm{E}(\eta_t)' + (\lambda' \otimes I_N)C_{11}^{\tau-1}\mathrm{E}\left\{[\tilde{A}_0 + A_1\eta_{t-\tau} - \Theta_1\epsilon_{t-\tau} + B_1 h_{t-\tau} - \mathrm{E}(h_t)]\eta'_{t-\tau}\right\}.
\end{aligned}$$

Thus we have (15) with

$$Q = \mathrm{E}\left\{[\tilde{A}_0 + A_1\eta_t - \Theta_1\epsilon_t + B_1 h_t - \mathrm{E}(h_t)]\eta'_t\right\}. \tag{C.21}$$

Example 2 *For Sentana's (1995) univariate QGARCH(1,1) process considered in Example 1, tedious calculations show that the autocorrelation function of the squares is given by*

$$r_\tau = \begin{cases} \frac{2\alpha_0\alpha_1(1-\alpha_1\beta_1-\beta_1^2)+(3\alpha_1+\beta_1)(1-\alpha_1-\beta_1)\theta^2}{2\alpha_0(1-2\alpha_1\beta_1-\beta_1^2)+3(1-\alpha_1-\beta_1)\theta^2} & \tau = 1 \\ (\alpha_1+\beta_1)r_{\tau-1} & \tau > 1. \end{cases} \tag{C.22}$$

Thus, the decay pattern of the ACF is equal to that of the standard GARCH(1,1) process, as already noted by Sentana (1995). However, for given values of α_0, α_1, and β_1, the ACF of the QGARCH(1,1) process is always larger than that of the GARCH(1,1), and is increasing in θ^2: It is straightforward to see that $\partial r_\tau/\partial\theta^2 > 0$ is equivalent to $\alpha_1(1-\alpha_1\beta_1-\beta_1^2)(3\alpha_1+\beta_1)^{-1} < (1-2\alpha_1\beta_1-\beta_1^2)/3$, and simple manipulations reveal that this is equivalent to $3\alpha_1^2+2\alpha_1\beta_1+\beta_1^2 < 1$, which is just the condition for the existence of the fourth moment, and, thus, the ACF of ϵ_t^2.

D Moments of the MNM(k)–GARCH(p,q) process

In this Appendix, we indicate how the moments of the MNM(k)–GARCH(p,q) model may be computed for higher–order GARCH models, i.e., with p and/or q larger than 1. We keep the discussion short, because in most applications GARCH(1,1) rather than GARCH(p,q) will suffice, and the properties of the GARCH(1,1) case have been developed in detail in the preceding appendix. Moreover, in order to avoid clutter, we shall assume that all the components have zero means, i.e., in (5), $\mu_j = 0$, $j = 1, \ldots, k$, and that there are no leverage effects, i.e., in (6), $\theta_{ij} = 0$, $i = 1, \ldots, q$, $j = 1, \ldots, k$.

Recently, using the ARMA representation of a GARCH model, Zadrozny (2005) employed a state–space representation of the univariate GARCH(p, q) process to derive a condition for the existence of its fourth moment.[15] We use a similar approach to find a condition for the existence of the unconditional fourth–moment matrix of the multivariate mixed normal GARCH model. However, we use a different representation than Zadrozny (2005). Although the representation we use is less parsimonious, it is preferred in present context because, in addition to providing a condition and an expression for the fourth moment, it allows for the computation of the autocorrelation matrices of the process. Clearly, the results presented here also apply to the single–component case, i.e., the standard GARCH(p, q) model in vech form, the fourth–moment structure of which has been investigated by Hafner (2003). However, Hafner's (2003) analysis is based on the MA(∞) representation of the process, which makes the application of the results less convenient. A brief comparison of Hafner's (2003) analysis with our approach is provided at the end of this Appendix. A condition for the existence of the fourth moment in single–component multivariate GARCH(p, q) models has also been derived by Comte and Liebermann (2000).[16] Their condition involves a matrix which is composed of $2q$ terms, where q is the ARCH order. For $q = 1$, this matrix coincides with the matrix Z in Theorem 3 of Hafner (2003). However, Comte and Liebermann (2000) do not consider how to compute the autocovariances from their approach.

To write the model in VARMA form, define $\bar{h} = \mathrm{E}(\eta_t | \psi_{t-1}) = (\lambda' \otimes I_N) h_t$, and $u_t = \eta_t - \bar{h}_t$, so that $\{u_t\}$ is a white noise process (uncorrelated but not independent).[17] Then we can write the MNM–GARCH(p, q) process as a VARMA(r, v) model for h_t, i.e.,

$$h_t = A_0 + \sum_{i=1}^{r} C_i h_{t-i} + \sum_{i=1}^{v} A_i u_{t-i}, \tag{D.23}$$

where $r = \max\{p, q\}$, $v = \max\{q, 2\}$, $C_i = \lambda' \otimes A_i + B_i$, $A_i = 0$, for $i > q$, and $B_i = 0$, for $i > p$. To put the MNM–GARCH(p, q) model in VAR(1) form, we adopt a slightly modified form of the VAR(1) representation of a VARMA model discussed in Lütkepohl (2005, p. 426).

[15] Papers dealing with the fourth–moment structure of the univariate GARCH(p, q) model include Chen and An (1998), He and Teräsvirta (1999), Karanasos (1999), Ling (1999), Davidson (2002, Section 2.3), and Ling and McAleer (2002). There also exist results for other multivariate GARCH models than the vech model. For example, moment conditions for Jeantheau's (1998) generalization of Bollerslev's (1990) constant conditional correlation model are derived in Ling and McAleer (2003) and He and Teräsvirta (2004).

[16] Computation of the fourth moment in the bivariate case was also considered by Nijman and Sentana (1996).

[17] Recall that, in the present section, we assume zero component means and absence of leverage effects.

That is, we define

$$
\begin{aligned}
X_t &= \begin{pmatrix} h_t \\ \vdots \\ h_{t-r+1} \\ u_{t-1} \\ \vdots \\ u_{t-v+1} \end{pmatrix}, \quad \tilde{A}_0 = \begin{pmatrix} A_0 \\ 0_{N\{k(r-1)+(v-1)\}\times 1} \end{pmatrix}, \quad Z = \begin{pmatrix} A_1 \\ 0_{Nk(r-1)\times N} \\ I_N \\ 0_{N(v-2)\times N} \end{pmatrix}, \\
H &= \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}, \quad \text{where} \quad H_{11} = \begin{pmatrix} C_1 & \cdots & C_{r-1} & C_r \\ & I_{kN(r-1)} & & 0_{kN(r-1)\times kN} \end{pmatrix}, \\
H_{12} &= \begin{pmatrix} A_2 & \cdots & A_v \\ & 0_{kN(r-1)\times N(v-1)} & \end{pmatrix}, \quad H_{22} = \begin{pmatrix} 0_{N\times N(v-2)} & 0_{N\times N} \\ I_{N(v-2)} & 0_{N(v-2)\times N} \end{pmatrix}, \quad \text{(D.24)}
\end{aligned}
$$

and H_{21} is a $N(v-1)\times kNr$ matrix of zeros. Thus, X_t is of dimension $N(kr+v-1)$. Given the definitions in (D.24), we can write

$$X_t = \tilde{A}_0 + HX_{t-1} + Zu_{t-1}. \quad \text{(D.25)}$$

From (D.25), we can infer that the MNG–GARCH(p,q) process is stationary if $\rho(H_{11}) < 1$, or, equivalently, the roots of

$$\det(I_{kN} - \sum_{i=1}^{r} C_i z^i) = 0 \quad \text{(D.26)}$$

are outside the unit circle.

To find a condition for the existence of the fourth moment, i.e., of $\mathrm{E}(X_tX_t')$, define the matrices

$$\mathbb{I} := (I_{kN}, 0_{kN\times N\{k(r-1)+v-1\}}), \quad \text{(D.27)}$$

so that $h_th_t' = \mathbb{I}X_tX_t'\mathbb{I}'$, and

$$F_M := G_M(I_N \otimes \mathrm{vec}(\Lambda)' \otimes I_N)(K_{Nk} \otimes I_{kN}) - (\lambda' \otimes I_N \otimes \lambda' \otimes I_N). \quad \text{(D.28)}$$

Definition (D.28) is useful for calculating $\mathrm{E}(u_tu_t')$. In fact, as $\mathrm{E}(u_tu_t') = \mathrm{E}(\eta_t\eta_t') - \mathrm{E}(\bar{h}_t\bar{h}_t')$, and $\mathrm{E}(\bar{h}_t\bar{h}_t') = (\lambda' \otimes I_N)\mathrm{E}(h_th_t')(\lambda \otimes I_N)$, we have

$$\mathrm{E}[\mathrm{vec}(u_tu_t')] = L_M\mathrm{E}[\mathrm{vec}(h_th_t')].$$

Also note that $\mathrm{E}(h_tu_t') = \mathrm{E}[h_t\mathrm{E}(u_t'|\Psi_{t-1})] = 0$. Thus,

$$
\begin{aligned}
&\mathrm{E}[\mathrm{vec}(X_tX_t')|\Psi_{t-2}] \quad \text{(D.29)} \\
&= \tilde{A}_0 \otimes \tilde{A}_0 + (\tilde{A}_0 \otimes H + H \otimes \tilde{A}_0)X_{t-1} + [H \otimes H + (Z \otimes Z)F_M(\mathbb{I} \otimes \mathbb{I})]\mathrm{vec}(X_{t-1}X_{t-1}');
\end{aligned}
$$

and an argument quite similar to that of Appendix C shows that $\mathrm{E}(X_t X_t)$ exists if and only if $\rho(P) < 1$, where

$$P := H \otimes H + (Z \otimes Z) F_M (\mathbb{I} \otimes \mathbb{I}). \tag{D.30}$$

In case of existence, the unconditional moments $\mathrm{E}(X_t)$ and $\mathrm{E}(X_t X_t')$ can be computed by taking unconditional expectations and solving (D.25) and (D.29), respectively.

As mentioned above, representation (D.25) of the process—although less parsimonious than that used by Zadrozny (2005) in his analysis of the univariate GARCH(p,q) process—is useful for calculating the sequence of autocovariances of η_t, which is not dealt with in Zadrozny (2005). The reason is that, in present context, it is much more convenient to work with the VARMA representation of h_t rather than that of η_t, which involves determinantal terms and is quite difficult to handle. On the other hand,

$$\mathrm{E}(\eta_t \eta_{t-\tau}') = \mathrm{E}[(\bar{h}_t + u_t)(\bar{h}_{t-\tau} + u_{t-\tau})'] = \mathrm{E}(\bar{h}_t \bar{h}_{t-\tau}') + \mathrm{E}(\bar{h}_t u_{t-\tau}'), \tag{D.31}$$

as $\mathrm{E}(u_t u_{t-\tau}') = \mathrm{E}(u_t \bar{h}_{t-\tau}') = 0$. Thus, it is advantageous to explicitly model both h_t and u_t. The terms on the right–hand side of (D.31) can be extracted from

$$\begin{aligned}
\mathrm{E}(X_t X_{t-\tau}') &= \mathrm{E}\{\mathrm{E}(X_t|\Psi_{t-\tau-1}) X_{t-\tau}'\} \\
&= \mathrm{E}\left\{\left(\sum_{i=0}^{\tau-1} H^i \tilde{A}_0 + H^\tau X_{t-\tau}\right) X_{t-\tau}'\right\} \\
&= \mathrm{E}\{[(I - H^\tau)(I - H)^{-1} \tilde{A}_0 + H^\tau X_{t-\tau}] X_{t-\tau}'\} \\
&= \mathrm{E}(X_t)\mathrm{E}(X_t)' + H^\tau [\mathrm{E}(X_t X_t') - \mathrm{E}(X_t)\mathrm{E}(X_t)'].
\end{aligned} \tag{D.32}$$

For the first term on the right–hand side of (D.31), we have

$$\mathrm{E}(\bar{h}_t \bar{h}_{t-\tau}') = (\lambda' \otimes I_N)\mathrm{E}(h_t h_{t-\tau}')(\lambda \otimes I_N) = (\lambda' \otimes I_N)\mathbb{I}\mathrm{E}(X_t X_{t-\tau}')\mathbb{I}'(\lambda \otimes I_N), \tag{D.33}$$

and the second term is

$$\mathrm{E}(\bar{h}_t u_{t-\tau}') = (\lambda' \otimes I_N)\mathbb{I}\mathrm{E}(X_t X_{t-\tau+1}')\tilde{\mathbb{I}}', \tag{D.34}$$

where

$$\tilde{\mathbb{I}} = (0_{N \times kNr}, I_N, 0_{N \times N(v-2)}). \tag{D.35}$$

This completes the characterization of the fourth–moment structure of the multivariate mixed normal GARCH(p,q) process.

To compare with Hafner's (2003) method for the single–component multivariate GARCH(p,q), let us briefly sketch his argument when applied to the MNM(k)–GARCH(p,q) process. By inverting (D.23), we obtain the MA(∞) representation of h_t,

$$h_t = [I_{kN} - C(1)]^{-1} A_0 + [I_{kN} - C(L)]^{-1} A(L) u_t = \mathrm{E}(h_t) + \sum_{i=1}^{\infty} \Phi_i u_{t-i}, \tag{D.36}$$

where $C(z) = \sum_{i=1}^{r} C_i z^i$, L is the lag operator, $L^i y_t = y_{t-i}$, and the MA(∞) coefficient matrices, Φ_i, $i = 1, 2, \ldots$, can be calculated recursively in the usual way (see, e.g., Lütkepohl, 2005). Then, in case of existence of $\mathrm{E}(h_t h_t')$, or, equivalently, of $\mathrm{E}(u_t u_t')$,

$$\begin{aligned} \mathrm{vec}[\mathrm{E}(h_t h_t')] &= \mathrm{vec}[\mathrm{E}(h_t)\mathrm{E}(h_t)'] + \sum_{i=1}^{\infty}(\Phi_i \otimes \Phi_i)\mathrm{vec}[\mathrm{E}(u_t u_t')] && \text{(D.37)} \\ &= \mathrm{vec}[\mathrm{E}(h_t)\mathrm{E}(h_t)'] + \sum_{i=1}^{\infty}(\Phi_i \otimes \Phi_i)F_M \mathrm{vec}[\mathrm{E}(h_t h_t')]. && \text{(D.38)} \end{aligned}$$

Thus,

$$\mathrm{vec}[\mathrm{E}(h_t h_t')] = \left(I_{N^2 k^2} - \sum_{i=1}^{\infty}(\Phi_i \otimes \Phi_i)F_M \right)^{-1} \mathrm{vec}[\mathrm{E}(h_t)\mathrm{E}(h_t)']. \quad \text{(D.39)}$$

From Theorem 2 in Hafner (2003) we have that, under covariance stationarity, a condition for existence of $\mathrm{E}(h_t h_t')$ is $\rho(\Omega) < 1$, where $\Omega = \sum_{i=1}^{\infty}(\Phi_i \otimes \Phi_i)F_M$.

Note that Ω and P (as defined in (D.30)) are, in general, different matrices and do not have the same maximal eigenvalue. However, as expected, the conditions $\rho(P) < 1$ and $\rho(\Omega) < 1$ turn out to be equivalent, i.e., $\rho(\Omega) \lesseqgtr 1 \Leftrightarrow \rho(P) \lesseqgtr 1$. For example, in the GARCH(1,1) case, we can compute Ω explicitly as

$$\Omega = \sum_{i=1}^{\infty}(\Phi_i \otimes \Phi_i)F_M = \sum_{i=1}^{\infty}[(C_1^{i-1}A_1) \otimes (C_1^{i-1}A_1)]F_M = (I_{k^2 N^2} - C_1 \otimes C_1)^{-1}(A_1 \otimes A_1)F_M. \quad \text{(D.40)}$$

Thus, in the single–component, univariate case, where $F_M = 2$, $\rho(\Omega) = 2\alpha_1^2/(1-(\alpha_1+\beta_1)^2) < 1$ is equivalent to $3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2 < 1$ under stationarity, i.e., $\alpha_1 + \beta_1 < 1$. For $k > 1$ and/or $M > 1$, the equivalence is not obvious but can still be checked numerically, and an example is provided in Figure 10, where we consider the case $p = q = 1$, $k = 2$, and $M = 2$ (hence $N = 3$), with $\lambda = (0.75, 0.25)'$,

$$\begin{aligned} A_{11} &= \begin{pmatrix} A_{11,11} & 0.05 & 0.20 \\ 0.12 & 0.13 & 0.05 \\ 0.24 & 0.13 & 0.10 \end{pmatrix}, \quad A_{12} = \begin{pmatrix} 0.20 & 0.12 & 0.01 \\ 0.10 & 0.09 & 0.08 \\ 0.24 & 0.13 & 0.20 \end{pmatrix}, && \text{(D.41)} \\ B_{11} &= \begin{pmatrix} 0.23 & 0.16 & 0.30 \\ 0.29 & 0.14 & 0.05 \\ 0.03 & 0.12 & 0.13 \end{pmatrix}, \quad B_{12} = \begin{pmatrix} 0.32 & 0.04 & 0.03 \\ 0.02 & 0.04 & 0.18 \\ 0.11 & 0.05 & 0.25 \end{pmatrix}, \end{aligned}$$

and parameter $A_{11,11}$ varies from 0 to 0.25. Clearly, both $\rho(\Omega)$ and $\rho(P)$ are monotonically increasing in $A_{11,11}$, and they intersect exactly at $\rho(\Omega) = \rho(P) = 1$. We observe the same pattern if we let any other parameter vary or use different parameter matrices in (D.41).

However, while the conditions give rise to the same conclusions with respect to existence of $\mathrm{E}(\eta_t \eta_t')$, they have, relative to each other, several benefits and drawbacks. Clearly, an

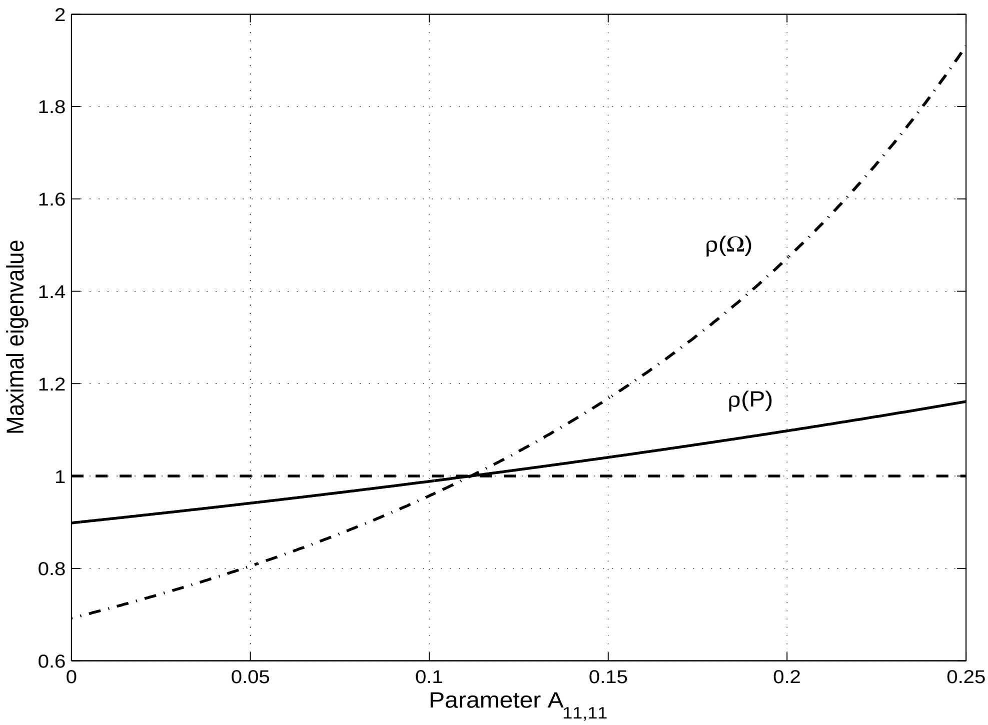

Figure 10: This figure illustrates the equivalence of the fourth–moment eigenvalue conditions based on matrices $\Omega = \sum_{i=1}^{\infty}(\Phi_i \otimes \Phi_i)F_M$, defined in (D.40), and P, defined in (D.30). Shown are, for the example in (D.41), the maximal eigenvalues, $\rho(P)$, solid, and $\rho(\Omega)$, dash–dotted, for values of the parameter $A_{11,11}$ increasing from 0 to 0.25.

advantage of Ω is that, in particular for high ARCH/GARCH orders, it is of a considerably lower dimension than P. On the other hand, computation of Ω requires the evaluation of the (infinite) sequence of MA coefficients Φ_i, $i = 1, 2, \ldots$, while the expression for P in (D.30) is more compact. Also, Hafner (2003: 35) argues that, if the fourth moment exists, the closeness of the maximum eigenvalue of Ω to unity may be considered as a measure for the degree of "persistence in kurtosis". However, this is questionable as, from (D.29), the appropriate measure of persistence in the fourth moment is $\rho(P)$. Thus, using Ω, the persistence in fourth moments is underestimated, as it is generally found that, for $\rho(\Omega) < 1$, we have $\rho(\Omega) < \rho(P) < 1$, as illustrated in Figure 10.

References

Aas, K., Haff, I. H., and Dimakos, X. K. (2006). Risk Estimation Using the Multivariate Normal Inverse Gaussian Distribution. *Journal of Risk*, 8:39–60.

Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. In Petrov, B. N. and Csaki, F., editors, *2nd International Symposium on Information Theory*, pages 267–281, Akademiai Kiado, Budapest.

Alexander, C. (2001). *Orthogonal GARCH*, volume 2 of *Mastering Risk*, pages 21–38. FT Prentice Hall, London.

Alexander, C. (2002). Principal Component Models for Generating Large GARCH Covariance Matrices. *Economic Notes*, 31:337–359.

Alexander, C. and Chibumba, A. M. (1997). Multivariate Orthogonal Factor GARCH. Mimeo, University of Sussex.

Alexander, C. and Lazar, E. (2004). Normal Mixture GARCH(1,1). Applications to Exchange Rate Modelling. Forthcoming in: *Journal of Applied Econometrics*.

Alexander, C. and Lazar, E. (2005). Asymmetries and Volatility Regimes in the European Equity Market. ICMA Centre Discussion Papers in Finance 2005–14, The Business School for Financial Markets at the University of Reading.

Balestra, P. and Holly, A. (1990). A General Kronecker Formula for the Moments of the Multivariate Normal Distribution. Cahiers de recherches économiques 9002, Département d'économétrie et d'économie politique, Université de Lausanne.

Bauwens, L., Bos, C. S., and van Dijk, H. K. (1999). Adaptive Polar Sampling with an Application to a Bayes Measure of Value-at-Risk. Tinbergen Institute Discussion Paper TI 99-082/4, Erasmus University Rotterdam.

Bauwens, L. and Laurent, S. (2005). A New Class of Multivariate Skew Densities, with Application to Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business & Economic Statistics*, 23:346–354.

Bauwens, L., Laurent, S., and Rombouts, J. V. K. (2006). Multivariate GARCH Models: A Survey. *Journal of Applied Econometrics*, 21:79–109.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31:307–327.

Bollerslev, T. (1990). Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model. *Review of Economics and Statistics*, 73:498–505.

Bollerslev, T. and Engle, R. F. (1993). Common Persistence in Conditional Variance. *Econometrica*, 61:167–186.

Bollerslev, T., Engle, R. F., and Wooldridge, J. M. (1988). A Capital Asset Pricing Model with Time-varying Covariances. *Journal of Political Economy*, 96:116–131.

Cai, J. (1994). A Markov Model of Switching–Regime ARCH. *Journal of Business and Economic Statistics*, 12:309–316.

Campbell, R., Koedijk, K., and Kofman, P. (2002). Increased Correlation in Bear Markets: A Downside Risk Perspective. *Financial Analysts Journal*, 58:87–94.

Carnero, M. A., Pena, D., and Ruiz, E. (2004). Persistence and Kurtosis in GARCH and Stochastic Volatility Models. *Journal of Financial Econometrics*, 2:319–342.

Chen, M. and An, H. Z. (1998). A Note on the Stationarity and the Existence of Moments of the GARCH Model. *Statistica Sinica*, 8:505–510.

Cheung, Y. M. and Xu, L. (2003). Dual Multivariate Auto-Regressive Modeling in State Space for Temporal Signal Separation. *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, 33:386–398.

Comte, F. and Liebermann, O. (2000). Second–Order Noncausality in Multivariate GARCH Processes. *Journal of Time Series Analysis*, 21:535–557.

Davidson, J. (2002). Establishing Conditions for the Functional Central Limit Theorem in Nonlinear and Semiparametric Time Series Processes. *Journal of Econometrics*, 106:243–269.

Dueker, M. J. (1997). Markov Switching in GARCH Processes and Mean–Reverting Stock–Market Volatility. *Journal of Business and Economic Statistics*, 15:26–34.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity With Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50:987–1008.

Engle, R. F. (1990). Stock Volatility and the Crash of '87: Discussion. *Review of Financial Studies*, 3:103–106.

Engle, R. F. (2002). Dynamic Conditional Correlation-a Simple Class of Multivariate GARCH Models. *Journal of Business and Economic Statistics*, 20:339–350.

Engle, R. F. and Kroner, K. F. (1995). Multivariate Simultaneous Generalized ARCH. *Econometric Theory*, 11:122–150.

Engle, R. F. and Ng, V. K. (1993). Measuring and Testing the Impact of News on Volatility. *Journal of Finance*, 48:1749–1778.

Figlewski, S. and Wang, X. (2000). Is the "Leverage Effect" a Leverage Effect? Working paper 00-037, Stern School of Business, New York University.

Ghazal, G. A. and Neudecker, H. (2000). On Second–order and Fourth–order Moments of Jointly Distributed Random Matrices: A Survey. *Linear Algebra and its Applications*, 321:61–93.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, 48:1779–1801.

Gray, S. F. (1996). Modeling the Conditional Distribution of Interest Rates as a Regime–Switching Process. *Journal of Financial Economics*, 42:27–62.

Haas, M., Mittnik, S., and Paolella, M. S. (2004a). Mixed Normal Conditional Heteroskedasticity. *Journal of Financial Econometrics*, 2:211–250.

Haas, M., Mittnik, S., and Paolella, M. S. (2004b). A New Approach to Markov–Switching GARCH Models. *Journal of Financial Econometrics*, 4:493–530.

Hafner, C. M. (2003). Fourth Moment Structure of Multivariate GARCH Models. *Journal of Financial Econometrics*, 1:26–54.

Hamilton, J. D. and Susmel, R. (1994). Autoregressive Conditional Heteroskedasticity and Changes in Regime. *Journal of Econometrics*, 64:307–333.

Hansson, B. and Hördahl, P. (1998). Testing the Conditional CAPM using Multivariate GARCH-M. *Applied Financial Economics*, 8:377–388.

He, C. and Teräsvirta, T. (1999). Fourth Moment Structure of the GARCH(p, q) Process. *Econometric Theory*, 15:824–846.

He, C. and Teräsvirta, T. (2004). An Extended Constant Correlation GARCH Model and its Fourth-Moment Structure. *Econometric Theory*, 20:904–926.

Jeantheau, T. (1998). Strong Consistency of Estimators for Multivariate ARCH Models. *Econometric Theory*, 14:70–86.

Karanasos, M. (1999). The Second Moment and the Autocovariance Function of the Squared Errors of the GARCH Model. *Journal of Econometrics*, 90:63–76.

Kass, R. E. and Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, 90:773–795.

Klaassen, F. (2002). Improving GARCH Volatility Forecasts with Regime–Switching GARCH. *Empirical Economics*, 27:363–394.

Kroner, K. F. and Ng, V. K. (1998). Modeling Asymmetric Comovements of Asset Returns. *Review of Financial Studies*, 11:817–844.

Ling, S. (1999). On the Probabilistic Properties of a Double Threshold ARMA Conditional Heteroskedastic Model. *Journal of Applied Probability*, 36:1–18.

Ling, S. and McAleer, M. (2002). Necessary and Sufficient Moment Conditions for the GARCH(r, s) and Asymmetric Power GARCH(r, s) Models. *Econometric Theory*, 18:722–729.

Ling, S. and McAleer, M. (2003). Asymptotic Theory for a Vector ARMA-GARCH Model. *Econometric Theory*, 19:280–310.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer–Verlag, Berlin.

Magnus, J. R. (1988). *Linear Structures*. Griffin, London.

Magnus, J. R. and Neudecker, H. (1979). The Commutation Matrix: Some Properties and Applications. *Annals of Statistics*, 7:381–394.

Magnus, J. R. and Neudecker, H. (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley, Chichester.

McLachlan, G. J. and Peel, D. (2000). *Finite Mixture Models*. John Wiley & Sons, New York.

Nijman, T. and Sentana, E. (1996). Marginalization and Contemporaneous Aggregation in Multivariate GARCH Processes. *Journal of Econometrics*, 71:71–87.

Palm, F. C. and Vlaar, P. J. G. (1997). Simple Diagnostic Procedures for Modeling Financial Time Series. *Allgemeines Statistisches Archiv*, 81:85–101.

Patton, A. J. (2004). On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation. *Journal of Financial Econometrics*, 2:130–168.

Pelletier, D. (2006). Regime Switching for Dynamic Correlations. *Journal of Econometrics*, 131:445–473.

Sentana, E. (1995). Quadratic ARCH Models. *Review of Economic Studies*, 62:639–661.

Teicher, H. (1963). Identifiability of Finite Mixtures. *The Annals of Mathematical Statistics*, 34:1265–1269.

Tse, Y. K. and Tsui, A. K. C. (2002). A Multivariate GARCH Model with Time-varying Correlations. *Journal of Business and Economic Statistics*, 20:351–362.

Vlaar, P. J. G. and Palm, F. C. (1993). The Message in Weekly Exchange Rates in the European Monetary System: Mean Reversion, Conditional Heteroscedasticity, and Jumps. *Journal of Business and Economic Statistics*, 11:351–360.

Wong, C. S. and Li, W. K. (2001). On a Mixture Autoregressive Conditional Heteroscedastic Model. *Journal of the American Statistical Association*, 96:982–985.

Yakowitz, S. J. and Spragins, J. D. (1968). On the Identifiability of Finite Mixtures. *The Annals of Mathematical Statistics*, 39:209–214.

Zadrozny, P. A. (2005). Necessary and Sufficient Restrictions for Existence of a Unique Fourth Moment of a Univariate GARCH(P, Q) Process. CESifo Working Paper 1505, Ifo Institute for Economic Research.

CFS Working Paper Series:

No.	Author(s)	Title
2006/09	**Markus Haas** **Stefan Mittnik** **Marc S. Paolella**	Multivariate Normal Mixture GARCH
2006/08	**Elena Carletti** **Philipp Hartmann** **Giancarlo Spagnolo**	Bank Mergers, Competition and Liquidity
2006/07	**Alexander Muermann** **Stephen H. Shore**	Strategic Trading and Manipulation with Spot Market Power
2006/06	**Jan Pieter Krahnen** **Frank A. Schmid** **Erik Theissen**	Investment Performance and Market Share: A Study of the German Mutual Fund Industry
2006/05	**Jan Pieter Krahnen**	Die Stabilität von Finanzmärkten: Wie kann die Wirtschaftspolitik Vertrauen schaffen?
2006/04	**Jan Pieter Krahnen** **Christian Wilde**	Risk Transfer with CDOs and Systemic Risk in Banking
2006/03	**Panos Parpas** **Berc Rustem** **Volker Wieland** **Stan Zakovic**	Mean Variance Optimization of Non–Linear Systems and Worst–case Analysis
2006/02	**Christoffer Carroll** **Miles S. Kimball**	Precautionary Saving and Precautionary Wealth
2006/01	**M. Hashem Pesaran**	Market Efficiency Today
2005/33	**Fulvio Corsi** **Uta Kretschmer** **Stefan Mittnik** **Christian Pigorsch**	The Volatility of Realized Volatility

Copies of working papers can be downloaded at http://www.ifk-cfs.de