"Identification of structural multivariatie GARCH models"

Hafner, Christian ; Herwartz, Helmut ; Maxand, Simone

Abstract

Multivariate GARCH models are widely used to model volatility and correlation dynamics of financial time series. These models are typically silent about the transmission of implied orthogonalized shocks to vector returns. We propose a loss statistic to discriminate in a data-driven way between alternative structural assumptions about the transmission scheme. In its structural form, a four dimensional system comprising US and Latin American stock market returns points to a substantial volatility transmission from the US to the Latin American markets. The identified structural model improves the estimation of classical measures of portfolio risk,as well as corresponding variations.

Document type : *Document de travail (Working Paper)*

Référence bibliographique

Hafner, Christian ; Herwartz, Helmut ; Maxand, Simone. *Identification of structural multivariatie GARCH models.* CORE Discussion Paper ; 2018/20 (2018) 48 pages

2018/20

DP

Christian M. Hafner, Helmut Herwartz
and Simone Maxand

Identification of structural multivariate GARCH models

CORE
CENTER FOR
OPERATIONS RESEARCH
AND ECONOMETRICS

CORE
Voie du Roman Pays 34, L1.03.01
Tel (32 10) 47 43 04
Fax (32 10) 47 43 01
Email: immaq-library@uclouvain.be
https://uclouvain.be/en/research-institutes/immaq/core/discussion-papers.html

Identification of structural multivariate GARCH models

Christian M. Hafner* Helmut Herwartz† Simone Maxand‡

2nd May 2018

Abstract

Multivariate GARCH models are widely used to model volatility and correlation dynamics of financial time series. These models are typically silent about the transmission of implied orthogonalized shocks to vector returns. We propose a loss statistic to discriminate in a data-driven way between alternative structural assumptions about the transmission scheme. In its structural form, a four dimensional system comprising US and Latin American stock market returns points to a substantial volatility transmission from the US to the Latin American markets. The identified structural model improves the estimation of classical measures of portfolio risk, as well as corresponding variations.

Keywords: Structural innovations; identifying assumptions; MGARCH; portfolio risk; volatility transmission.
JEL Classification: C32, G15.

---

*Corresponding author, Institut of statistics, biostatistics, and actuarial sciences, and CORE, Université catholique de Louvain. e-mail `christian.hafner@uclouvain.be`

†Department of Economics, University of Goettingen, Germany. e-mail `hherwartz@uni-goettingen.de`. Much of the research documented in this paper has been done while the second author was visiting the Department of Economics and the Center of Environmental and Natural Resource Economics (CENRE) of the University of Chile, Santiago de Chile. The hospitality and productive research environment of both institutions are cordially acknowledged.

‡Department of Economics, University of Helsinki.

1 Introduction

Over the past decades, modeling volatility has been one of the most rapidly growing areas of research in empirical finance. This development has been spurred especially by the introduction of multivariate GARCH models (MGARCH), such as the so-called VEC model of Bollerslev et al. (1988), the BEKK model of Engle and Kroner (1995), or the DCC model of Engle (2002). These modeling frameworks generalize the classical univariate GARCH models of Engle (1982) and Bollerslev (1986), and allow us to deal with multi-dimensional problems such as portfolio allocation and optimization, portfolio risk evaluation, and asset pricing.

In the wake of a growing dependence across markets and countries, evidenced during the financial and economic crisis, 2007-2009, and the European debt crisis, concern has risen among central bankers, regulators, policy makers and portfolio managers about understanding the volatility linkages between countries, markets, and asset classes from a more fundamental perspective. In empirical assessments of contagion and transmission,[1] one may rely on either variance impulse response functions (e.g., Lin (1997), Hafner and Herwartz (2006a)), tests of specific parameter restrictions in MGARCH models (see, for example, Nakatani and Teräsvirta (2009), Billio et al. (2012), Woźniak (2015)), or tests for causality in variance, e.g. Hafner and Herwartz (2006b). Providing an alternative perspective, Diebold and Yilmaz (2009) suggest to model contemporaneous variance transmissions by aggregating the forecast error variance decomposition of an underlying vector autoregression (VAR) of realized variance measures into a 'spillover index' . For the purpose of measuring (co-)variance transmission in a timely manner at high frequency, Fengler and Herwartz (2018) extend this approach to a BEKK framework.

While conveying insightful information about the volatility and correlation dynamics, both realized volatility VAR and MGARCH models are limited in the sense that in most

[1] See, among others, Engle et al. (1990), Hamao et al. (1990), Forbes and Rigobon (2002), and Bali and Hovakimian (2009)

studies the underlying model of shock transmissions lacks identification in a strictly structural sense. For instance, Diebold and Yilmaz (2009) use the order-dependent Cholesky factorization and justify this choice by showing that alternative variable orderings obtain similar time profiles of aggregate volatility spillovers. Fengler and Herwartz (2018) motivate their model based on the economic plausibility of the estimated volatility spillover patterns. Indeed, identified structural models of realized volatility dynamics are scarce. As a notable example, Dimpfl and Jung (2012) exploit the chronological order of trading hours to estimate a structural VAR (SVAR) of realized volatilities. Similarly, MGARCH specifications often include an ad-hoc decomposition, mostly the symmetric square root, of the conditional covariance matrix.

Identification of structural MGARCH models has been addressed previously by, for example, van der Weide (2002), Rigobon (2003) and Weber (2010). The orthogonal GARCH model of van der Weide (2002) and Rigobon (2003) linearly combines orthogonal conditionally heteroskedastic shocks, where the weights of the linear combinations are time-invariant. The structural conditional correlation model of Weber (2010) builds upon a flexible form of vector-valued volatilities coupled with correlations of the underlying structural shocks which are constant or dynamic in the sense of Bollerslev (1990) and Engle (2002), respectively. It is worth mentioning that both identified MGARCH models - orthogonal GARCH and structural conditional correlation - come at the cost of imposing reduced form dynamic profiles that are more restrictive than those of the BEKK model. An example of using external information to identify an unrestricted BEKK model is Herwartz and Roestel (2016), who exploit a-priori assumptions on transient patterns of market dominance, and hence, volatility transmission. Under high frequency sampling, however, external economic information that could be employed for identification is generally scare and often not consensual.

In this paper, retaining the full flexibility of unrestricted BEKK models, we build upon recent advances in data based identification of SVARs to develop a new structural

MGARCH model. In particular, we exploit the uniqueness of non-Gaussian structural shocks for MGARCH identification (Moneta et al., 2013; Lanne et al., 2017; Gouriéroux et al., 2017).[2] The starting point is the definition of structural shocks as stochastically independent innovations, not merely orthogonal. This definition, which has been used before e.g. in Hafner and Herwartz (2006a), exploits the full information of the joint distribution of shocks, and not only their second order moment structure. For the identification of independent shocks we suggest a feasible loss (test) statistic that is a weighted sum of squared differences between empirical third and fourth order moments and their theoretical counterparts under the independence assumption. We show theoretically and via a simulation study that the minimization of the proposed test statistic consistently identifies the structural MGARCH model that has been estimated by quasi maximum likelihood (QML).

Intuitively, the idea is that under our assumptions, higher order moments, as opposed to second-order moments, are not invariant with respect to rotations of a decomposed conditional covariance matrix. The assumptions exclude the case of a multivariate normal distribution, under which no identification would be possible. In the univariate case, the conditional kurtosis of GARCH process only depends on the kurtosis of the i.i.d. innovations, which is a constant. In the multivariate case, however, the conditional kurtosis of portfolio returns is not constant under rotations of the i.i.d. innovation vector. This crucial information can be exploited empirically to identify structural innovations.

The identified structural model is particularly helpful in modelling the higher order moment structure of portfolio returns which is an important ingredient for modern risk management and regulatory purposes. For example, ESA (2016) prescribes the use of Cornish-Fisher expansions, based on empirical skewness and kurtosis, to estimate Value-at-Risk (VaR) measures for certain financial investment products. Consequently, these

[2]The recent literature on identifying SVARs comprises various data based approaches relying on unconditional heteroskedasticity, mixture distributions and iid non-Gaussian models. For an up-to-date textbook treatment of identification in SVARs the reader may consult Lütkepohl and Kilian (2017).

VaR measures will directly depend on the particular choice of the structural model. Furthermore, the structural model is also helpful in quantifying uncertainties inherent in conditional VaR and ES statistics. To further improve the modelling of tail events, we introduce a higher order risk measure by considering the difference between squared returns and conditional variances. While the conditional mean of this variable is zero, its conditional variance is directly linked to the conditional kurtosis of portfolio returns. Analysing the variation and tail risk of this variable provides important insights for risk management with respect to variations of classical risk measures over time.

The merits of our approach are discussed in detail based on an application to a four dimensional system of weekly stock returns of US and Latin American markets. In particular and opposite to an ad-hoc symmetric covariance decomposition, the estimation results suggest an active role of US markets in transmitting volatilities to Latin American markets. We show via simulations that VaR and expected shortfall (ES) measures are well approximated using the estimated structural model, as opposed to a symmetric model. We also show that this holds for quantities measuring the higher order type risk.

The next section introduces the structural MGARCH model and discusses the moment based identification criterion. In Section 3 we highlight how structural information can be beneficial for both types of risk analysis. A simulation study in Section 4 sheds light on the discriminatory strength of the identification criterion in finite samples. Section 5 provides an empirical illustration of identified volatility transmission patterns. The empirical analysis in Section 6 shows how structural information improves practical matters of monitoring portfolio risks. Section 7 concludes. Appendix A provides additional material on the marginal processes of simulated vector systems, and Appendix B on the QML BEKK model estimates.

2 Structural MGARCH and identification

Similar to modelling univariate conditionally heteroskedastic processes by means of GARCH models and its variants, MGARCH processes have been widely applied to quantify system covariances for speculative returns in a conditional manner. Let $\mathcal{F}_{t-1}$ denote the information set that is available at time $t-1$. For an N-dimensional vector valued system of speculative returns, denoted r_t, let

$$r_t = \mu_t + e_t, \tag{1}$$

where $\mu_t = E[r_t|\mathcal{F}_{t-1}] =: E_{t-1}[r_t]$, and e_t an error term with conditional mean equal to zero. In the analysis of daily (weekly) data one often assumes $\mu_t = 0$ (or $\mu_t = \mu$). Conditional on $\mathcal{F}_{t-1}$, MGARCH models specify the time varying covariance of e_t in a fully deterministic manner, i.e.

$$\text{Cov}[e_t|\mathcal{F}_{t-1}] = \text{Cov}_{t-1}[e_t] = H_t, \tag{2}$$

where the matrix process $\{H_t\}$ is symmetric and positive definite. Alternative approaches to modelling the conditional covariance H_t can be distinguished according to their flexibility of approximating the entire space of positive definite covariance matrices on the one hand, and the numerical tractability of their parameter space on the other hand, see Bauwens et al. (2006) for a review.

The so-called BEKK model (named after an early working paper by Baba, Engle, Kraft and Kroner) has become a widespread approach to formalize the conditional covariance of e_t (Engle and Kroner, 1995). The most flexible BEKK(p, q, K) model reads as

$$H_t = CC' + \sum_{k=1}^{K}\sum_{i=1}^{q} A'_{ki} e_{t-i} e'_{t-i} A_{ki} + \sum_{k=1}^{K}\sum_{i=1}^{p} B'_{ki} H_{t-i} B_{ki}\ , \tag{3}$$

where C is a lower triangular matrix and A_{ki} and B_{ki} are $N \times N$ parameter matrices.

Complementing theoretical results in Engle and Kroner (1995), Boussama et al. (2011) show that, given some regularity conditions, the MGARCH process $\{e_t\}$ is ergodic and strictly and weakly stationary if

$$\rho\left(\sum_{k=1}^{K}\sum_{i=1}^{q} A_{ki}\otimes A_{ki} + \sum_{k=1}^{K}\sum_{i=1}^{p} B_{ki}\otimes B_{ki}\right) < 1, \tag{4}$$

with $\rho(Z)$ denoting the spectral radius of a square matrix Z and $\otimes$ the Kronecker matrix product. Furthermore, Stelzer (2008) points out that, except for a few degenerate covariance processes, the space of BEKK models covers (approximately) the entire space of positive definite vec MGARCH processes of Bollerslev et al. (1988). While it is only mildly costly in terms of model flexibility, the BEKK model has the convenient feature to issue positive definite covariance paths under mild regularity and initial conditions.[3]

The vast majority of applications of the BEKK model use the parsimonious order specification BEKK(1,1,1). Omitting respective matrix indices for notational convenience, the BEKK(1,1,1) model is given by

$$H_t \quad = \quad CC' + A'e_{t-1}e'_{t-1}A + B'H_{t-1}B. \tag{5}$$

Accordingly, for $N = 2, 3, 4$ low order BEKK(1,1,1) models comprise 11, 24, and 42 parameters, respectively. Henceforth, the parameters of the BEKK(1,1,1) model are stacked into the parameter vector $\theta = (\text{vech}(C)', \text{vec}(A)', \text{vec}(B)')'$. To express the dependence of the conditional covariance matrix in (5) on θ, we will sometimes write $H_t(\theta)$.

To relate the conditional covariance in (2) and the MGARCH process in (1), one often considers an i.i.d. vector valued innovation process $\{\xi_t\}$ with mean zero and unit covariance as the source of stochastic variation of e_t (and, hence, r_t). This relation is often formalized by means of ad-hoc decompositions of H_t such as, for example, the symmetric

[3]Selected applications of the BEKK model are provided by Chan et al. (1992); Bekaert and Harvey (1995); Baele (2005); Fountas and Karanasos (2007); Hassan and Malik (2007).

eigenvalue decomposition or Cholesky factors. Denoting by $H_t^{1/2}$ the matrix square root of H_t obtained by eigenvalue decomposition, the MGARCH process reads as

$$e_t|\mathcal{F}_{t-1} \;=\; H_t^{1/2}\xi_t,\ \xi_t \sim iid(0, I_N). \tag{6}$$

Viewing the expression in (6) as a structural scheme, the j-th column of $H_t^{1/2}$ formalizes how single orthogonalized shocks ξ_{jt} in ξ_t affect the returns (or their reduced form residuals) collected in r_t (e_t). Similarly, the i-th row of $H_t^{1/2}$ unravels the contribution of each shock in ξ_t to uncertainty/volatility received by a single market r_{it} (e_{it}). Importantly, the eigenvalue decomposition implies for each market a symmetry of cross market volatility reception and transmission.[4] On a priori grounds one might argue that this implication lacks economic justification in many contexts. For instance, it appears intuitive to expect a wedge between patterns of market specific volatility transmission and reception if the considered markets differ considerably in terms of economic importance and functioning or market valuation. An analyst might at least warrant a more flexible model framework which could also nest as special cases particular a-priori schemes.

Generalizing the exposition in (6), the identification problem in MGARCH models can be made more explicit in terms of a structural transmission scheme as

$$e_t|\mathcal{F}_{t-1} \;=\; D_t\xi_t, \tag{7}$$

where $\mathrm{Cov}[e_t|\mathcal{F}_{t-1}] = D_tD_t' = H_t$. As a main consequence this generalization implies that the decomposition is not unique without further assumptions. For instance, under the assumption of conditional normality all possible covariance decompositions of the form

[4] The symmetric square root $H_t^{1/2}$ of H_t is obtained as $H_t^{1/2} = \Gamma_t\Lambda_t^{1/2}\Gamma_t'$, where the eigenvectors of H_t are the columns of Γ_t, and the diagonal matrix Λ_t has the eigenvalues of H_t along its diagonal. Opposite to choosing the symmetric decomposition factors $H_t^{1/2}$, a-priori opting for a (triangular) Cholesky factor of H_t renders the analysis conditional on the presumed ordering of variables. Opting for a lower triangular Cholesky factor, shocks ξ_{1t} would contribute to all system returns, while r_{1t} (or e_{1t}) would not receive any effects of shocks other than ξ_{1t}.

$H_t = D_t D_t'$ are observationally equivalent. Hence, in this case the determination of a particular decomposition scheme, i.e. of a structural representation of the MGARCH model, has to rely on external non-data based information. Unlike a-priori choices, the structural approach followed in this paper processes data-based information to determine the most convenient specification of D_t. We next describe the identification problem and our approach to solving it.

2.1 The identification problem

Following (7), the identification problem can be stated more explicitly by formalizing the space of structural covariance decompositions as

$$H_t = H_t^{1/2} H_t^{1/2} = H_t^{1/2} R_\delta R_\delta' H_t^{1/2} = D_t D_t', \tag{8}$$

where R_δ is a rotation matrix such that $R_\delta R_\delta' = I_N$, and $D_t = H_t^{1/2} R_\delta$.[5] When choosing $R_\delta = I_N$, for instance, the structural model in (7) corresponds to the symmetric model in (6). More specific, R_δ is parameterized as the product of distinct forms of Givens rotation matrices where the elements of δ, denoted δ_i, $0 \leq \delta_i < \pi$, are rotation angles. For a model of dimension N, δ comprises $N(N-1)/2$ rotation angles. For instance, in the case of $N = 3$,

$$R_\delta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\delta_1) & -\sin(\delta_1) \\ 0 & \sin(\delta_1) & \cos(\delta_1) \end{bmatrix} \times \begin{bmatrix} \cos(\delta_2) & 0 & -\sin(\delta_2) \\ 0 & 1 & 0 \\ \sin(\delta_2) & 0 & \cos(\delta_2) \end{bmatrix} \times \begin{bmatrix} \cos(\delta_3) & -\sin(\delta_3) & 0 \\ \sin(\delta_3) & \cos(\delta_3) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{9}$$

Apparently, the matrices D_t depend on both the reduced form MGARCH parameters $\theta \in \Theta$ and the rotation angles δ, i.e., $D_t(\theta, \delta) = H_t(\theta)^{1/2} R_\delta$. Consequently, model implied

[5] The supposition of $H_t^{1/2}$ as baseline factor of H_t is without loss of generality. With a respective adaptation of the rotation matrix other factors (e.g., Cholesky factors) might be used as well.

structural shocks read as $\xi_t(\theta, \delta) = D_t^{-1}(\theta, \delta)e_t$. While the vector of coefficients θ can be uniquely determined by means of QML estimation, δ lacks identification without additional information. Put differently, the specification of D_t focuses on the most appropriate choice of R_δ for given (estimates of) θ and $\{H_t\}$.

The selection of D_t (i.e., of R_δ) is similar to the identification problem in structural VAR models. Main examples of SVAR identification strategies which rely on external information are the narrative approach (see, e.g., Romer and Romer, 2010), the sign restrictions approach (see, e.g., Faust, 1998; Uhlig, 2005), and a-priori choices of specific models (e.g. the Cholesky decomposition by Sims, 1980). Given the high-frequency nature of financial return data, one may favor data-based external information for identification rather than economic (or a-priori) assumptions which are likely to lack sufficient motivation when analysing speculative returns sampled at medium or high frequency (daily or weekly, say). The recent literature on identification in SVARs has shown that the data-based identification of structural relations offers unique solutions if structural shocks are non-Gaussian and independently distributed (Lanne et al., 2017; Gouriéroux et al., 2017). In empirical applications of GARCH models the supposition of Gaussian GARCH residuals is regularly confirmed as overly restrictive and models incorporating leptokurtic innovations have been put forward (e.g, Bollerslev, 1990). This naturally allows to draw upon the branch of identification techniques for non-Gaussian systems, and to exploit the particular structure of third and fourth order co-moments that is implied by independently distributed non-Gaussian model innovations.

2.2 Moment based identification

The main theoretical motivation for our approach is the following result of Lancaster (1954), which is a characterization of the multivariate normal distribution: If ξ_t is a vector of independent standardized random variables with finite cumulants, and the non-trivial transformation $\xi_t^* = R_\delta \xi_t$ gives a vector of independent standardized random variables, then each component of ξ_t^* is normal (and hence so is each component of ξ_t). Thus, under the requirement of independence, the only situation where non-identifiability occurs is the case of a normal distribution. In other words, structural innovations are identifiable if the innovation vector is not normally distributed.

Under non-normality, the core idea to identify the structural MGARCH model is to start with an a-priori decomposition $H_t(\theta)^{1/2}$, obtaining standardized reduced form errors $e_t^{std} := H_t(\theta)^{-1/2}e_t$. In a second step, this factor is subjected to systematic rotations R_δ to obtain a specific (rotated) matrix $D_t(\theta,\delta) = H_t(\theta)^{1/2}R_\delta$ that implies innovations which are best in matching a set of co-moment conditions that apply to independent shocks.

Before characterizing the identification technique, we need to impose some assumptions on the vector of innovations $\xi_t(\theta,\delta) = D_t^{-1}(\theta,\delta)e_t$. In the following, we distinguish between the vector of independent structural error terms which is related to rotation angle δ_0 by $\xi_t := \xi_t^{\delta_0} = \xi_t(\theta,\delta_0)$ and the general vector of structural shocks $\xi_t^\delta = \xi_t(\theta,\delta)$ which depends on rotation angle δ and is not necessarily independent. The following assumptions are imposed on ξ_t.

Assumption 1 ξ_t, $t = 1,\ldots,T$, *is an N-dimensional vector with the following properties:*

(i) At most one of the components of the random vector ξ_t has a normal distribution.

(ii) The components ξ_{it} are mutually independent with $E[\xi_{it}] = 0$ and $Var[\xi_t] = I_N$.

(iii) For some $\epsilon > 0$, $E|\xi_{kt}|^{6+\epsilon} < \infty$, $k = 1,\ldots,N$.

In the following, we discuss moment properties of independent shocks in order to propose a diagnostic statistic in Section 2.3 to identify them based on higher order moments. We first introduce some notation and define respectively the marginal skewness and kurtosis of innovations, $m_i^{(3)} := E[\xi_{it}^3]$ and $m_i^{(4)} := \mathrm{E}[\xi_{it}^4]$, $i = 1,\ldots,N$. Furthermore, let $m^{(g)} := (m_1^{(g)},\ldots,m_N^{(g)})'$ denote N-dimensional moment vectors for $g \in \mathbb{N}$. In particular, the skewness vector corresponds to $m^{(3)}$, the kurtosis vector to $m^{(4)}$, and the second order moment $m^{(2)}$ is unity by definition. The operator diag(a) stacks an n-dimensional vector a into a diagonal $(n \times n)$ matrix.

Let us now define the matrices of third and fourth order cross-products of ξ_t as

$$\Phi_t := (\xi_t \otimes \xi_t)\xi_t' \quad \text{and} \quad \Psi_t := \xi_t\xi_t' \otimes \xi_t\xi_t'$$

and corresponding expectations,

$$\begin{aligned} \Phi := E[\Phi_t] &= \operatorname{diag}(m^{(3)})L_N \\ \Psi = E[\Psi_t] &= 2D_N D_N^+ + \operatorname{vec}(I_N)\operatorname{vec}(I_N)' + \operatorname{diag}(\operatorname{vec}(\operatorname{diag}(m^{(4)} - 3\iota_N))) \end{aligned}$$

where L_N is the unique $N^2 \times N$ matrix defined by the property $\operatorname{diag}(A) = L'\operatorname{vec}(A)$ for any $N \times N$ matrix A, D_N is the duplication matrix, D_N^+ its generalized inverse, and $\iota_N := (1, 1, \ldots, 1)'$ an N-dimensional vector of ones. A derivation of these expressions is straightforward following the lines of Proposition 5.3 of Hafner and Rombouts (2007). An elementwise characterization of the matrix Ψ was derived in Fengler and Herwartz (2018).

Furthermore, define $\boldsymbol{\phi}_t$ as the vector containing the non-redundant elements of vec(Φ_t), except for the terms ξ_{it}^3, $i = 1, \ldots, N$. That is, $\boldsymbol{\phi}_t$ contains all cross-products of the type $\xi_{it}^2\xi_{jt}$ and $\xi_{it}\xi_{jt}\xi_{kt}$, so that $\boldsymbol{\phi}_t$ is of dimension $q_\phi := N(N-1)(N+4)/6$. Similarly, define the vector $\boldsymbol{\psi}_t$ as the vector containing the unique elements of vec(Ψ_t), except for the terms ξ_{it}^4, $i = 1, \ldots, N$. That is, $\boldsymbol{\psi}_t$ contains all cross-products of the type $\xi_{it}^3\xi_{jt}$, $\xi_{it}^2\xi_{jt}\xi_{kt}$, $\xi_{it}^2\xi_{jt}^2$ and $\xi_{it}\xi_{jt}\xi_{kt}\xi_{lt}$, and it can easily be checked that $\boldsymbol{\psi}_t$ is of dimension $q_\psi := N(N-1)(N^2+7N+18)/24$.

Finally, let $\boldsymbol{\phi} := \mathrm{E}[\boldsymbol{\phi}_t]$ and $\boldsymbol{\psi} := \mathrm{E}[\boldsymbol{\psi}_t]$ be the vectors of expectations of third- and fourth-order cross-products of innovations. Table 1 gives the dimensions of $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$ as a function of N for $N = 2, 3, 4$. Note that $q_\phi = O(N^3)$ and $q_\psi = O(N^4)$.

N	q_ϕ	q_ψ
2	2	3
3	7	12
4	16	31

Table 1: *Number of third and fourth order cross-products as a function of* N*. The functions are* $q_\phi = N(N-1)(N+4)/6$ *and* $q_\psi = N(N-1)(N^2+7N+18)/24$.

The variance-covariance matrix of $\boldsymbol{\phi}_t$ can be obtained as

$$\text{vech}\left(\text{Var}(\boldsymbol{\phi}_t)\right) = C_1 \begin{pmatrix} 0 \\ 1 \\ \text{vecl}(m^{(3)}m^{(3)\prime}) \\ m^{(4)} \end{pmatrix}$$

where $\text{vecl}(\cdot)$ denotes the operator that stacks the lower triangular part of a matrix, excluding the diagonal, into a column vector, and C_1 is a $q_\phi(q_\phi+1)/2 \times N(N+1)/2 + 2$-dimensional binary selection matrix, that is, each row contains exactly one entry of unity, and zeros elsewhere. For example, in the bivariate case ($N = 2$), we have $\boldsymbol{\phi}_t = (\xi_{1t}^2\xi_{2t}, \xi_{1t}\xi_{2t}^2)'$ and

$$\text{Var}(\boldsymbol{\phi}_t) = \begin{pmatrix} m_1^{(4)} & m_1^{(3)}m_2^{(3)} \\ m_1^{(3)}m_2^{(3)} & m_2^{(4)} \end{pmatrix}$$

such that C_1 has dimension 3×5 and is given by

$$C_1 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Similarly, the variance-covariance matrix of $\boldsymbol{\psi}_t$ is given by

$$\text{vech}\left(\text{Var}(\boldsymbol{\psi}_t)\right) = C_2 \begin{pmatrix} 0 \\ 1 \\ m^{(4)} \\ \text{vecl}(m^{(4)}m^{(4)\prime}) - \iota_{N(N-1)/2} \\ m^{(6)} \end{pmatrix}$$

where C_2 is a $q_\psi(q_\psi+1)/2 \times (2(N+1)+N(N-1)/2)$-dimensional binary selection matrix.

To construct the covariance-matrix of the vector $(\boldsymbol{\phi}_t', \boldsymbol{\psi}_t')'$, it remains to calculate the covariance between $\boldsymbol{\phi}_t$ and $\boldsymbol{\psi}_t$, which is obtained as

$$\text{vec}\left(\text{Cov}(\boldsymbol{\phi}_t, \boldsymbol{\psi}_t)\right) = C_3 \begin{pmatrix} 0 \\ m^{(3)} \\ m^{(5)} \\ \text{vec}(m^{(3)}m^{(4)\prime}) \end{pmatrix},$$

where C_3 is a $q_\phi q_\psi \times (N+1)^2$ binary selection matrix.

We now stack both vectors into the q-dimensional vector $S_t = (\boldsymbol{\phi}_t',\, \boldsymbol{\psi}_t')'$ (e.g. $q = 47$ for $N = 4$), and $q = q_\phi + q_\psi$. All variances and covariances of $\boldsymbol{\phi}_t$ and $\boldsymbol{\psi}_t$ are then used to construct $\Sigma := \text{Var}((\boldsymbol{\phi}_t', \boldsymbol{\psi}_t')')$. Furthermore, $\hat{\boldsymbol{\phi}}_t$ and $\hat{\boldsymbol{\psi}}_t$ ($\hat{S}_t$) define the respective estimators obtained from $\hat{\xi}_t = \xi_t(\hat{\theta}, \delta_0)$. We discuss consistency of these estimators in the following Section 2.3.

2.3 A diagnostic and its asymptotic properties

The co-moment structure detailed above holds for the case where the independent structural shocks in ξ_t can be distinguished from the space of dependent distributions obtaining non-trivial relations among third and fourth order relations. Noticing that alternative choices of δ in (8) obtain shocks ξ_t^δ with distinct joint distributions, we suggest to select the vector of rotation angles in a way such that implied empirical third and fourth order co-moments of ξ_t^δ are best in line with the moments of ξ_t. For the vector $\mathrm{E}[S_t^\delta]$ of stacked theoretical third and fourth order moments of ξ_t^δ, let $\bar{S}_T^\delta = \frac{1}{T}\sum_{t=1}^T S_t^\delta = \left(\bar{\boldsymbol{\phi}}_T^{\delta\prime}, \bar{\boldsymbol{\psi}}_T^{\delta\prime}\right)'$ denote the empirical counterpart. To discriminate between distinct choices of the rotation angle δ we consider the statistic

$$\lambda_\delta^T = T(\bar{S}_T^\delta - E[S_t])'\Sigma_\delta^{-1}(\bar{S}_T^\delta - E[S_t]), \tag{10}$$

where the covariance matrix Σ_δ collects the second order (co-)moments of the elements in $S_t^\delta = \left(\boldsymbol{\phi}_t^{\delta\prime}, \boldsymbol{\psi}_t^{\delta\prime}\right)'$. The entries in Σ_δ are calculated under the assumption of independence with respect to rotation angle δ. Similarly, the expectation vector $E[S_t]$ comprises $E[\boldsymbol{\phi}_t]$ and $E[\boldsymbol{\psi}_t]$ defined under independence.[6]

If independence is violated, and under non-sphericity, the third and fourth order moment conditions provided in Section 2.2 are (partly) violated such that λ_δ^T in (10) is expected to diverge for increasing T. Hence, minimizing λ_δ^T provides a particular decomposition of H_t which implies weakest dependence among corresponding orthogonalized shocks. Using the statistic in (10), the identified structural model reads as

$$\widehat{D}_t = D_t(\theta, \hat{\delta}), \text{with } \hat{\delta} = \operatorname{argmin}_\delta\{\lambda_\delta |\, \xi_t = D_t(\theta, \delta)^{-1}e_t\}. \tag{11}$$

Regarding the identification scheme in (10) and (11) we make the following two remarks. First, while the asymptotics for the statistic in (10) hold for the true independent shocks ξ_t, consistency of QML MGARCH estimation $\hat{\theta}$ coupled with continuous mapping arguments motivate that the result remains valid even for estimated (and rotated) MGARCH innovation estimates. A rigorous proof of these considerations is provided below. Second,

[6]For the class of spherical distributions, of which the standard normal is a special case, the above moment conditions hold as well and they are invariant with respect to orthogonal rotations. However, the standard normal is the only member of this class with independent components, so that spherical distributions are excluded by Assumption $1(i) - (ii)$. In practice, observing a statistic λ_δ^T that is a non-trivial function of δ is an indicator of non-sphericity.

the covariance matrix Σ is characterized by a sparse structure, mostly comprising elements of zero or unity, see the results of Section 2.2. Only a few higher order moments have to be estimated. Accordingly, the asymptotic distribution of λ_δ^T in (10) requires finiteness of marginal moments of the order of $6+\epsilon$, $\epsilon > 0$ (see e.g. White, 1984, p. 119).[7]

This section focuses on the theoretical derivation of the asymptotic properties of the estimator $\widehat{D}_t$, or rather $\hat{\delta}$. The decomposition $D_t(\theta, \delta)$ depends on both estimated coefficients in θ and δ. In empirical practice the parametric specification of the MGARCH model is typically unknown such that model parameters θ are subjected to QML estimation. Accordingly, a feasible evaluation of (10) relies on estimated rather than true innovation estimates $\hat{\xi}_t^\delta = D_t(\hat{\theta}, \delta)^{-1} e_t$.

For the derivation of asymptotic properties in this section we proceed in three steps. Firstly, keeping the rotation angles δ_0 fixed, we derive consistency of the shocks $\hat{\xi}_t = D_t(\hat{\theta}, \delta_0)^{-1} e_t$ based on the QML estimator $\hat{\theta}$ which implies consistency of estimates $\widehat{S}_t = (\hat{\boldsymbol{\psi}}_t', \hat{\boldsymbol{\phi}}_t')'$. Secondly, we show the asymptotic properties of $\lambda_{\delta_0}^T$ and $\hat{\lambda}_{\delta_0}^T$ (based on $\hat{\theta}$). Thirdly, we study the properties of the statistic λ_δ^T to finally establish consistency of $\hat{\delta}$.

To implement QML estimation of the parameters, the assumption of conditional normality of ξ_t, and hence $e_t \sim N(0, H_t)$, is commonly adopted. Jeantheau (2000) initiated the derivation of asymptotic properties of QML estimators in MGARCH models. In particular, he provides regularity conditions, such as ergodicity and compactness to establish consistency of QML estimates. Comte and Lieberman (2003) show consistency and asymptotic normality of the QML estimator of the BEKK model assuming existence of, respectively, second and eighth order moments of the MGARCH process. Weakening these requirements, Hafner and Preminger (2009) show consistency of QML estimators under the condition of finite second order moments of the (non-Gaussian) innovations ξ_t, and establish asymptotic normality of QML estimators given sixth order moments of the MGARCH process.

[7] In the case where cross products of the form $\xi_{it}^3 \xi_{jt}$ are removed from $\boldsymbol{\psi}_t$, S_t and $\bar{S}_T$ accordingly, the asymptotics of (10) can be established with the weaker assumption of finite moments of order $4+\epsilon$, $\epsilon > 0$.

Consistency of the QML estimator, i.e. $\hat{\theta} \xrightarrow{p} \theta$, and application of a continuous mapping theorem imply that $H_t(\hat{\theta}) - H_t(\theta) \xrightarrow{p} 0$ for $T \to \infty$. Consequently, the estimated decomposition $D_t(\hat{\theta}, \delta_0)$ converges to the true decomposition, $D_t(\hat{\theta}, \delta_0) - D_t(\theta, \delta_0) \xrightarrow{p} 0, \quad T \to \infty$. Furthermore, for $T \to \infty$,

$$\hat{\xi}_t - \xi_t = D_t(\hat{\theta}, \delta_0)^{-1} e_t - D_t(\theta, \delta_0)^{-1} e_t \xrightarrow{p} 0 \quad \text{and} \quad \hat{\xi}_t^2 - \xi_t^2 \xrightarrow{p} 0. \tag{12}$$

Convergence of the estimated structural innovations implies convergence of the vectors of third and fourth order cross products:

$$\hat{\boldsymbol{\psi}}_t - \boldsymbol{\psi}_t \xrightarrow{p} 0 \quad \text{and} \quad \hat{\boldsymbol{\phi}}_t - \boldsymbol{\phi}_t \xrightarrow{p} 0, \qquad T \to \infty. \tag{13}$$

In the following, we establish consistency of the empirical third and fourth order moments $\bar{\boldsymbol{\phi}} = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{\phi}_t$ and $\bar{\boldsymbol{\psi}} = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{\psi}_t$. Under Assumption 1, the empirical moments converge in probability by the application of a weak law of large numbers. For instance,

$$\frac{1}{T} \sum_{t=1}^{T} \xi_{ti}^2 \xi_{tj}^2 \xrightarrow{p} E(\xi_i^2 \xi_j^2) = 1, \quad T \to \infty. \tag{14}$$

Similarly, the remaining third and fourth order empirical moments defined in Section 2.2 converge to their expectation. For the vector of co-moments it follows that

$$\bar{\boldsymbol{\phi}} \xrightarrow{p} E(\boldsymbol{\phi}) \quad \text{and} \quad \bar{\boldsymbol{\psi}} \xrightarrow{p} E(\boldsymbol{\psi}) \qquad \text{for } T \to \infty. \tag{15}$$

Consistency of the QML estimator of the MGARCH model implies that the structural model determination by means of the statistic in (10) also applies asymptotically to estimated vectors of orthogonalized shocks $\hat{\xi}_t$. Combining the result from (15) with the convergence of the estimated moments stated in (13) establishes consistency of the estimated empirical moments, i.e.

$$\hat{\bar{\boldsymbol{\phi}}} \xrightarrow{p} E(\boldsymbol{\phi}) \quad \text{and} \quad \hat{\bar{\boldsymbol{\psi}}} \xrightarrow{p} E(\boldsymbol{\psi}) \qquad \text{for } T \to \infty. \tag{16}$$

While we have derived convergence under the assumption of $\delta = \delta_0$, the results hold for any rotation angle $\delta \in [0, 2\pi]$. Consequently, for the statistic in (10) it follows that for any rotation angle δ,

$$|\lambda_\delta^T - \lambda_\delta| \xrightarrow{p} 0, \qquad T \to \infty. \tag{17}$$

The limit λ_δ measures the distance between the expected third and fourth moments under rotation angle δ and δ_0.

Under the assumption of independence, i.e. $S_T = S_T^{\delta_0}$, the centered vector of empirical third and fourth order moments converges to a multivariate normal distribution, i.e., for $T \to \infty$,

$$\sqrt{T}(\bar{S}_T - E[S_t]) \xrightarrow{d} \mathcal{N}(0, \Sigma) \tag{18}$$

with covariance matrix Σ. This result follows from a multivariate central limit theorem and is, for instance, stated in Davidson (1994) in its general form. Accordingly,

$$T(\bar{S}_T - E[S_t])'\Sigma^{-1}(\bar{S}_T - E[S_t]) \xrightarrow{d} \chi^2(q), \quad T \to \infty. \tag{19}$$

With the result of (16) and equivalent statements for the other moments the estimated counterpart $\overline{\widehat{S}}$ has the same asymptotic distribution as given in (19). Similarly, the covariance matrix Σ can be substituted by its estimated version $\widehat{\Sigma}$ so that

$$T(\overline{\widehat{S}}_T - E[S_t])'\widehat{\Sigma}^{-1}(\overline{\widehat{S}}_T - E[S_t]) \xrightarrow{d} \chi^2(q), \quad T \to \infty. \tag{20}$$

Based on results from the independent component analysis literature, the decomposition $H_t = D_t D_t'$ is unique for independent non-Gaussian components $\xi_{k,t}$ (Comon, 1994; Lancaster, 1954). In the following, we derive that minimizing the statistic λ_δ with respect to the rotation angles δ provides a consistent estimator of this unique decomposition D_t.

For the standardized version of the vector of reduced form errors $e_t^{std} = H_t(\theta)^{-1/2} e_t$, we are interested in determining the rotation angles δ so that the components of $\xi_t^\delta = R_\delta^{-1} e_t^{std}$ are independent. Recall that $\bar{S}_T^\delta$ contains the empirical counterparts of $\mathrm{E}(S_t)$ dependent on the rotation angles δ. Thus, the statistic λ_δ contains the sum of squared elements of the weighted difference between $\bar{S}_T^\delta$ and $\mathrm{E}[S_t]$, i.e. $\Sigma_\delta^{-1/2}(\bar{S}_T^\delta - \mathrm{E}[S_t])$. This converges in probability to a zero vector under independence, i.e. for $\xi_t = R_{\delta_0}^{-1} e_t^{std}$.

Given this background, the statistic λ_δ minimizes the sum of squared standardized elements of the matrices of third and fourth order co-moments, except for the marginal skewness and kurtosis. Alternatively, one could maximize the information contained in the marginal skewness and kurtosis. This technique is used in a classical independent component analysis algorithm, the so-called joint approximate diagonalization of eigenmatrices (JADE). Following, for instance, Miettinen et al. (2015), the described identification technique can be formulated as the counterpart to standard ICA methodologies. While we are interested in minimizing the standardized cross-moments, JADE maximizes the standardized marginal fourth order moments. We obtain from Miettinen et al. (2015) that the minimum of $\lambda_{\delta_0} = 0$, i.e. at $\xi_t = R_{\delta_0}^{-1} e_t^{std}$, is unique up to permutation, sign-changes, and scaling. Thus, for independent ξ_t the criterion λ_δ is minimal, i.e. the minimum of λ_δ is obtained at rotation angle δ_0.

It remains to show that the consideration of alternative rotation angles δ leads to a consistent estimator of δ_0, i.e. $\hat{\delta} \to \delta_0$ for $T \to \infty$. For this, we follow the argumentation of Matteson and Tsay (2017). From the foregoing considerations we know that $\lambda_{\hat{\delta}} \geq \lambda_{\delta_0}$ and $\lambda_{\delta_0}^T \geq \lambda_{\hat{\delta}}^T$. Consequently, $\lambda_{\delta_0}^T - \lambda_{\delta_0} \geq \lambda_{\hat{\delta}}^T - \lambda_{\delta_0} \geq \lambda_{\hat{\delta}}^T - \lambda_{\hat{\delta}}$. Using (17) it follows

$$\begin{aligned}|\lambda_{\hat{\delta}}^T - \lambda_{\delta_0}| &\leq \max(|\lambda_{\delta_0}^T - \lambda_{\delta_0}|, |\lambda_{\hat{\delta}}^T - \lambda_{\hat{\delta}}|) \\ &\leq \sup_\delta |\lambda_\delta^T - \lambda_\delta| \xrightarrow{p} 0, \qquad T \to \infty.\end{aligned}$$

By continuity it follows that $R_{\hat{\delta}} \xrightarrow{p} R_{\delta_0}$ and, following Matteson and Tsay (2017), $\hat{\delta} \xrightarrow{p}$

δ_0 for $T \to \infty$.

3 Portfolio risk analysis

MGARCH models have become a widespread tool for the risk analysis of investment portfolios. Since these models provide a reduced form evaluation of time varying covariances, alternative structural specifications do not materialize in the assessment of portfolio variances. Moreover, alternative structural MGARCH models cannot be further discriminated under conditional normality, since this distribution is fully characterized by its first two moments. However, in a non-Gaussian framework, alternative structural models exhibit distinct higher order characteristics, and identification can be achieved by defining structural innovations as independent.

In this section we argue that higher order characteristics are important for risk management in two respects:[8] First, in the context of approximations of tail risk measures such as Value-at-Risk, and second, for measures of variability of squared portfolio returns around their conditional expectation. The latter risk type will be shown to be directly related to the conditional kurtosis of portfolio returns and will therefore be termed 'kurtosis risk' while the former is about approximating 'first order' risks. In the following we show the importance of conditional fourth order moments for both types of (portfolio) risk. Henceforth, let w denote an N-dimensional vector of portfolio weights. Accordingly, the stochastic part of portfolio returns reads as $\tau_t = w'e_t$, and the conditional portfolio

[8] The importance of higher order moments for risk measures has also been recognized by regulatory authorities. For example, the European supervisory authorities EBA, EIOPA and ESMA have deviced regulatory technical standards for packaged retail and insurance-based investment products (see ESA (2016)) that include the use of skewness and kurtosis via Cornish-Fisher expansions to evaluate market risk measures.

variance is

$$\begin{aligned}\sigma_t^2 &= \mathrm{Var}[\tau_t|\mathcal{F}_{t-1}] = \mathrm{E}[\tau_t^2|\mathcal{F}_{t-1}] \\ &= \mathrm{tr}[w'H_t w] = \mathrm{tr}[\widetilde{W}H_t], \end{aligned} \tag{21}$$

where $\widetilde{W} = ww'$. The following outline of risk concepts is conditional on a specific choice of portfolio weights w. Next, we (i) describe our simulation based approach to first order risk assessment, (ii) derive conditional fourth order moments of portfolio returns, and (iii) outline the measurement of kurtosis risk.

3.1 First order risk

The main measures for evaluating market risk, recommended by the Basel committee of Banking Supervision, see e.g. Chapter 2 of McNeil et al. (2nd edition, 2016), are the Value-at-Risk (VaR) and the expected shortfall (ES). In a non-Gaussian context VaR and ES may depend on higher order moments such that the choice between a structural or symmetric model for the decomposition of H_t is not innocuous for first order risk assessment.

In the empirical part of this work, we demonstrate the relevance of the model choice for higher order moments by simulating risk measures for alternative models. It will turn out that, for our data, risk measures based on an asymmetric structural model tend to give a better approximation of actual risks in comparison with a symmetric model. To be precise, at each point in time t, conditional on the information set $\mathcal{F}_{t-1}$, we simulate a very large number (i.e. 10^6) of portfolio returns τ_t using independent bootstrap draws from model implied innovations ξ_t, and then obtain the α-quantile of their empirical distribution, which will give an approximation of the conditional VaR at time t.[9] Formally,

$$\mathrm{VaR}_\alpha^\bullet(\tau_t|\mathcal{F}_{t-1}) = -F_{\bullet,\alpha}^{-1}(\tau_t|\mathcal{F}_{t-1}), \tag{22}$$

[9]We provide a more detailed description of the resampling scheme in Section 6.1.

where F is short for the (portfolio) return distribution function, '$\bullet$' indicates alternative model choices (D_t vs. $H_t^{1/2}$), and α is a nominal probability which we set alternatively to $\alpha = .010, .025, .050, .100, .250$.[10]

In the same vein, we can simulate the conditional ES, i.e.,

$$\mathrm{ES}^{\bullet}_{\alpha}(\tau_t|\mathcal{F}_{t-1}) = -\mathrm{E}_{\bullet}\left[\tau_t|\left(\tau_t < -\mathrm{VaR}^{\bullet}_{\alpha}(\tau_t|\mathcal{F}_{t-1})\right), \mathcal{F}_{t-1}\right], \tag{23}$$

where the expectation on the right hand side is taken with respect to the simulated distribution of portfolio returns.[11] To highlight the role of tail events for the determination of ES, one might also consider the expected excess shortfall, i.e.

$$\mathrm{EES}^{\bullet}_{\alpha}(\tau_t|\mathcal{F}_{t-1}) = \mathrm{ES}^{\bullet}_{\alpha}(\tau_t|\mathcal{F}_{t-1}) - \mathrm{VaR}^{\bullet}_{\alpha}(\tau_t|\mathcal{F}_{t-1}). \tag{24}$$

In analysing VaR exceedances, however, we do not only consider their average but also address in how far the structural MGARCH models are useful to manage their distribution. For this purpose, we extract from the simulated return distributions the interquartile range of VaR exceedances which provides a statistical tool that can be contrasted against empirical patterns of exceedances. Specifically, one would expect that, on average, about 50% of all VaR exceedances fall within the model implied interquartile range. Formally, the interquartile range is given by the interval

$$\left[\underline{\mathrm{ES}}^{\bullet}_{\alpha}(\tau_t|\mathcal{F}_{t-1}), \overline{\mathrm{ES}}^{\bullet}_{\alpha}(\tau_t|\mathcal{F}_{t-1})\right], \tag{25}$$

[10]See, e.g., Angelidis and Degiannakis (2007) for an overview of GARCH based VaR modelling.

[11]See, e.g., McNeil and Frey (2000) or Zhu and Galbraith (2011) for two examples for GARCH based ES modelling.

where the bounds are implicitly defined by the two equations

$$\begin{aligned}\text{Prob}\,[\tau_t > \underline{\text{ES}}^\bullet_\alpha(\tau_t|\mathcal{F}_{t-1})|\tau_t < -\text{VaR}^\bullet_\alpha(\tau_t|\mathcal{F}_{t-1}),\mathcal{F}_{t-1}] &= 0.25\\ \text{and}\;\; \text{Prob}\left[\tau_t < \overline{\text{ES}}^\bullet_\alpha(\tau_t|\mathcal{F}_{t-1})|\tau_t < -\text{VaR}^\bullet_\alpha(\tau_t|\mathcal{F}_{t-1}),\mathcal{F}_{t-1}\right] &= 0.75.\end{aligned}$$

3.2 Fourth order moments of portfolio returns

In the univariate case, the conditional kurtosis of a GARCH process with i.i.d. innovations is a constant, given by the kurtosis of the innovations. In the multivariate case however, as we will see in the following, the conditional kurtosis of portfolio returns is no longer constant and, in particular, is not invariant with respect to orthogonal rotations of the innovation vector. Using result 11, p.98 of Lütkepohl (1996), the conditional fourth order moment of portfolio returns is given by

$$\begin{aligned}E[\tau_t^4|\mathcal{F}_{t-1}] &= E[\text{tr}[\widetilde{W}(e_te_t')\widetilde{W}(e_te_t')]|\mathcal{F}_{t-1}]\\ &= E[\text{vec}(\widetilde{W})'(D_t\xi_t\xi_t'D_t' \otimes D_t\xi_t\xi_t'D_t')\text{vec}(\widetilde{W})],\end{aligned} \tag{26}$$

where e_t has been replaced by its structural representation which is known conditional on $\mathcal{F}_{t-1}$. Noticing that the Kronecker product in (26) involves identical matrices and using the results 7, 7.2(6) and 8(a) (Lütkepohl, 1996, p.97), one obtains

$$\begin{aligned}E[\tau_t^4|\mathcal{F}_{t-1}] &= E[\text{vec}(\widetilde{W})'(D_t \otimes D_t)\text{vec}(\xi_t\xi_t')\text{vec}(\xi_t\xi_t')'(D_t' \otimes D_t')\text{vec}(\widetilde{W})]\\ &= \text{vec}(\widetilde{W})'(D_t \otimes D_t)E[\text{vec}(\xi_t\xi_t')\text{vec}(\xi_t\xi_t')'](D_t' \otimes D_t')\text{vec}(\widetilde{W})]\\ &= \text{vec}(\widetilde{W})'(D_t \otimes D_t)\Psi(D_t' \otimes D_t')\text{vec}(\widetilde{W})] \;=:\; \varsigma_t.\end{aligned} \tag{27}$$

In (27) the $N^2 \times N^2$ matrix Ψ, defined as

$$\Psi := E[\text{vec}(\xi_t\xi_t')\text{vec}(\xi_t\xi_t')'], \tag{28}$$

collects the fourth order moments of the structural shocks ξ_t.

Structural estimates $\varsigma_t^{(0)} = \varsigma_t(D_t)$ allow for a comparison with the corresponding statistics retrieved from the symmetric decomposition $H_t^{1/2}$, i.e.

$$\begin{aligned} \varsigma_t^{(1)} &= \varsigma_t(H_t^{1/2}) \\ &= \text{vec}(\widetilde{W})'(H_t^{1/2} \otimes H_t^{1/2})\Psi(H_t^{1/2} \otimes H_t^{1/2})\text{vec}(\widetilde{W}). \end{aligned} \tag{29}$$

Since both statistics $\varsigma_t^{(0)}$ and $\varsigma_t^{(1)}$ evaluate fourth order properties, it is convenient to focus the analysis on estimated time varying 'kurtosis' statistics which could be defined as $\kappa_t^{(0)} = \varsigma_t^{(0)}/\sigma_t^4$ and $\kappa_t^{(1)} = \varsigma_t^{(1)}/\sigma_t^4$. To illustrate the difference between $\kappa_t^{(0)}$ and $\kappa_t^{(1)}$, we provide a numerical example. Set $N = 2$, $H_t = I_2$, $w = (0.5, 0.5)'$, and Ψ is obtained by assuming independent standardized Student-t distributed ξ_t with $\nu = 5$, 10 and 15 degrees of freedom.[12] The rotation matrix R is defined as

$$R_\delta = \begin{pmatrix} \cos\delta & -\sin\delta \\ \sin\delta & \cos\delta \end{pmatrix}.$$

Figure 1 shows the kurtosis $\kappa_t^{(0)}$ as a function of δ. The kurtosis $\kappa_t^{(1)}$ using the symmetric decomposition is obtained for $\delta = 0$. Obviously, the kurtosis variation is substantial if the degrees of freedom parameter is not too large. Therefore, identifying structural innovations using the λ_δ statistic of the previous section will have a non-negligible impact on the conditional kurtosis of portfolio returns.

3.3 Kurtosis risk

While measures of volatility, VaR and ES convey important information for risk management, it is important to notice that such statistics are - even conditionally - subject to inherent estimation uncertainty. In an MGARCH context, in particular, the vari-

[12]These distributions will also be considered in the simulation study in Section 4.

ation of squared returns contributes to estimation uncertainty of conditional variances, and exposes VaR and ES evaluations to conditional uncertainty. In statistical terms, the fluctuation of squared portfolio returns around the conditional variances depends on the conditional kurtosis of portfolio returns. Accordingly, we refer to such patterns of higher order risk as kurtosis risk. Recalling the results of Section 3.2, note that kurtosis risk is not invariant with respect to the specification of the structural MGARCH model.

Let us define the difference between squared portfolio returns and conditional variances as

$$m_t := \tau_t^2 - \sigma_t^2, \tag{30}$$

which is a zero mean random variable (conditional on $\mathcal{F}_{t-1}$ and unconditionally) that describes surprises to portfolio risk. Such surprises can be expected to be small (large) if $\mathrm{Var}[m_t|\mathcal{F}_{t-1}]$ is small (large). Using the definition of ς_t in (27), the assessment of variations of risk involves a quantification of conditional fourth order moments of τ_t,

$$v_t := \mathrm{Var}[m_t|\mathcal{F}_{t-1}] = E[\tau_t^4|\mathcal{F}_{t-1}] - \sigma_t^4 = \varsigma_t - \sigma_t^4. \tag{31}$$

A standardized measure is given by

$$\tilde{v}_t := \frac{v_t}{\sigma_t^4} = \kappa_t - 1, \tag{32}$$

which directly links v_t to the conditional kurtosis κ_t. Unlike the portfolio variance σ_t^2, the risk statistic in (32) depends on the specification of the structural model. Hence, it is interesting to contrast rival structural specifications with regard to their scope in quantifying time varying patterns of kurtosis risk. Accordingly, the empirical counterpart of $\tilde{v}_t$ is given by

$$\widetilde{m}_t^2 = \frac{m_t^2}{\sigma_t^4}. \tag{33}$$

While the empirical process $\{\widetilde{m}_t^2\}$ is conditionally observable, it is important to notice that the simulation of processes $\tilde{v}_t^{(0)} = \tilde{v}_t(D_t)$ and $\tilde{v}_t^{(1)} = \tilde{v}_t(H_t^{1/2})$ depends on the structural assumptions.

Apart from modelling mean profiles of $\widetilde{m}_t^2$, we evaluate model accuracies in capturing the (tail) event that $\widetilde{m}_t^2$ exceeds a prespecified critical quantile of its (simulated) distribution. This is analogous to Value-at-risk analysis of first order risk. In the present context, however, the event describes a positive excess violation of an upper quantile of $\widetilde{m}_t^2$, as opposed to a lower quantile of the return distribution. Formally, we define the conditional 'kurtosis-at-risk' (KaR) as

$$\mathrm{KaR}_\gamma^\bullet(\widetilde{m}_t^2|\mathcal{F}_{t-1}) = G_{\bullet,\gamma}^{-1}(\widetilde{m}_t^2|\mathcal{F}_{t-1}), \tag{34}$$

where $G(\cdot)$ is the distribution function of $\widetilde{m}_t^2$, '$\bullet$' indicates alternative model choices (D_t vs. $H_t^{1/2}$), and γ is a nominal probability level.

Similar to expected shortfall analysis of first order risk, we also consider the expected violation of the threshold given that an event of threshold violation has occurred. Formally, we define the conditional 'expected kurtosis shortfall' (EKS) as

$$\mathrm{EKS}_\gamma^\bullet(\widetilde{m}_t^2|\mathcal{F}_{t-1}) = \mathrm{E}_\bullet\left[\widetilde{m}_t^2|\left(\widetilde{m}_t^2 > \mathrm{KaR}_\gamma^\bullet(\widetilde{m}_t^2|\mathcal{F}_{t-1})\right), \mathcal{F}_{t-1}\right]. \tag{35}$$

Furthermore, we also consider the distribution of risk exceedances, and in particular the interquartile range, analogous to our analysis of first order risk. For kurtosis risk, the interquartile range is defined as the interval

$$\left[\underline{\mathrm{EKS}}_\gamma^\bullet(\widetilde{m}_t^2|\mathcal{F}_{t-1}), \overline{\mathrm{EKS}}_\gamma^\bullet(\widetilde{m}_t^2|\mathcal{F}_{t-1})\right], \tag{36}$$

where the bounds are implicitly determined by the following probability statements

$$\begin{aligned}
\text{Prob}\left[\widetilde{m}_t^2 > \underline{\text{EKS}}_\alpha^\bullet(\widetilde{m}_t^2|\mathcal{F}_{t-1})|\left(\widetilde{m}_t^2 > \text{KaR}_\gamma^\bullet(\widetilde{m}_t^2|\mathcal{F}_{t-1})\right), \mathcal{F}_{t-1}\right] &= 0.25 \\
\text{and Prob}\left[\widetilde{m}_t^2 < \overline{\text{EKS}}_\gamma^\bullet(\widetilde{m}_t^2|\mathcal{F}_{t-1})|\left(\widetilde{m}_t^2 > \text{KaR}_\gamma^\bullet(\widetilde{m}_t^2|\mathcal{F}_{t-1})\right), \mathcal{F}_{t-1}\right] &= 0.75.
\end{aligned}$$

4 Simulation study

The purpose of the simulation study is to uncover the scope of distance measures as defined in (10) for identifying the structural parameters in D_t. Owing to consistency of QML estimation, one may assume that this potential is only mildly affected by the distinction of true covariances $\{H_t\}$ and their estimated counterparts $\{\widehat{H}_t\}$. Therefore we disgard QML estimation steps in the simulations, and evaluate model selection outcomes under the assumption that $\{H_t\}_{t=1}^T$ is known to the analyst.

4.1 Simulation design

Simulated covariance dynamics accord with a BEKK(1,1,1) model (see (5)),[13]

$$C = \begin{pmatrix} 4.00 & 0.00 & 0.00 \\ 14.5 & 2.00 & 0.00 \\ 25.0 & -8.50 & 2.50 \end{pmatrix} /1000,\; A = \begin{pmatrix} .14 & .05 & .05 \\ -.05 & .14 & .05 \\ -.03 & .05 & .14 \end{pmatrix},\; B = \begin{pmatrix} .96 & -.06 & .02 \\ .04 & .96 & .02 \\ .04 & .02 & .96 \end{pmatrix}.$$

While the dynamic formalization uniquely determines $\{H_t\}_{t=1}^T$, a core aspect of the Monte Carlo study is the transmission of latent iid structural innovations (ξ_t) to observable reduced form disturbances (e_t). To generate (returns) e_t from the structural model, the

[13] The marginal processes implied by the simulated BEKK model are shortly described in the Appendix A. As it turns out, marginal processes show typical characteristics of empirical univariate GARCH processes in terms of covariance stationarity, persistence and news response.

transmission matrix D_t is determined by means of a rotation of $H_t^{1/2}$ as

$$D_t = H_t^{1/2} R_\delta, \tag{37}$$

where R_δ is in the form of (9) and rotation angles are $\delta = (\delta_1, \delta_2, \delta_3)' = (.10,\ .25,\ .40)\pi$. We generate MGARCH return data from standardized Student-t innovations ξ_t with 5, 10, 15, 30 and 100 degrees of freedom. Benchmark results for the unidentified Gaussian model are also provided. Distinguished sample sizes are $T + 100 = 1100$, 2100, 4100 and 8100, before discarding the first 100 draws to immunize the analysis from initialization effects. Each experiment is performed 10000 times.

To analyse the discriminatory strength of the moment based criterion in (10), the simulated data $\{e_t, H_t\}_{t=1}^T$ are subjected to a structural analysis presuming rival specifications of the decomposition in (8). On the one hand, a candidate decomposition matrix is chosen in accordance with the true model, $D_t^{(0)} = D_t$. One expects smallest loss statistics λ if an analyst has access to the true decomposition scheme. In addition, seven alternative (and false) decomposition schemes are evaluated in terms of the identification criterion. Owing to consistency of the moment estimators entering the covariance Σ in (18) all these choices should obtain (at least asymptotically) loss statistics in excess of their counterparts obtained from the true structural model. The following decomposition matrices are used

$$\begin{aligned}
D_t^{(0)} &= H_t^{1/2} R_\delta, \ \text{ true model, clockwise rotation of } H_t^{1/2}, \\
D_t^{(1)} &= H_t^{1/2}, \text{ unrotated model, eigenvalue decomposition,} \\
D_t^{(\overline{q})} &= H_t^{1/2} R_{\overline{q}\delta}, \overline{q} = 1.010,\ 1.025,\ 1.050 \text{ `excess' clockwise rotations,} \\
D_t^{(\underline{q})} &= H_t^{1/2} R_{\underline{q}\delta}, \underline{q} = (\overline{q})^{-1} \text{ `insufficient' clockwise rotations.}
\end{aligned}$$

To facilitate the discussion of simulation outcomes the loss statistics in (10) are denoted

$\lambda(D_t^{(q)})$ in this Section. Loss statistics $\lambda(D_t^{(q)})$ rely only on vectors $\boldsymbol{\psi}_t$.[14]

4.2 Simulation results

4.2.1 Documentation

In terms of frequency estimates Table 1 documents how often we obtained

(i) smaller loss statistics for the true model ($D_t^{(0)}$) in comparison with the unrotated symmetric covariance decomposition ($D_t^{(1)} = H_t^{1/2}$)

(ii) how often loss statistics $\lambda(D_t^{(0)})$ are smaller than statistics derived from both, over- and underrotated decompositions ($\overline{q} = 1.010, 1.025, 1.050$; $(\overline{q})^{-1}$).

Before taking a more detailed look at simulation outcomes, it is instructive to point out that we might expect two groups of basic outcomes from the experiments. First, noticing the uniqueness of non-Gaussian independent shocks, we expect the frequencies of observing $\lambda(D_t^{(0)}) < \lambda(H_t^{1/2})$ to increase with the sample size and to shrink with the number of Student-t degrees of freedom. Asymptotically, we expect a unit frequency of having $\lambda(D_t^{(0)})$ smaller than $\lambda(H_t^{1/2})$ unless samples originate from the Gaussian model. Second, with the presumption that the true rotation ($\lambda(D_t^{(0)})$) corresponds to a minimum of the loss statistic, frequencies of preferring the true rotation within a 'symmetric' space of over- and underrotations ($\lambda(D_t^{(\underline{q})}), \lambda(D_t^{(\overline{q})})$) should increase with the rotation bounds, i.e. with $\bar{q}$.

[14] Undocumented results from further simulations show that the identification outcomes are very similar if model selection is based jointly on third and fourth order comoments, i.e. on $(\boldsymbol{\phi}_t', \boldsymbol{\psi}_t')'$. Moreover, it is worth to recall that a consistent determination of $\mathrm{Cov}[\boldsymbol{\psi}_t]$ requires finiteness of marginal moments of ξ_{it} up to order $6 + \epsilon$. In some simulation experiments (e.g. when drawing data from standardized Student-t distributions with 5 degrees of freedom) these moment conditions are violated. Removing the six fourth order moments of the form $\xi_{it}^3 \xi_{jt}$ from $\boldsymbol{\psi}_t$ obtains a dependence diagnostic which only requires the existence of moments up to order $4 + \epsilon$. Relying on a six dimensional vector of comoments (i.e. using moments from $\xi_{it}^2 \xi_{jt}^2$ or $\xi_{it}^2 \xi_{jt} \xi_{kt}$) obtains simulation results which are very close to those documented in Table 1 of this study. Detailed results on unreported simulation outcomes are available from the authors upon request.

4.2.2 Discussion

Both of the expectations outlined before are confirmed by the Monte Carlo results. Contrasting loss measures from the true decomposition against counterparts obtained from the symmetric decomposition is supportive for the true model with frequencies close to unity under various scenarios with either sizeable deviations from the Gaussian model ($\nu = 5$, $T = 1000$) or sufficiently rich sample information. Having a standardized Student-t distribution with 30 degrees of freedom, samples as large as $T = 8000$ are informative to rule out the symmetric model against the true structural model in 92% of all Monte Carlo replications. In the Gaussian case (i.e. $\nu \to \infty$), the outcomes of model selection are purely random, with selection frequencies of 50% for both alternatives.

Comparing the loss statistic for the true model with those derived from over- and underrotations shows that minimizing the loss statistic obtains a local minimum in the neighbourhood of the true structural model (see also the exposition in Figure 1). With $T = 1000$ observations and Student-t shocks with 15 degrees of freedom, the true decomposition obtains a loss measure smaller than those from both bounds of false rotations with $q = 1.050$ in almost 20% of all Monte Carlo experiments. Pointing to the consistency of the minimization of the loss statistic, the respective success frequencies increase to almost 30%, 40% and 60% if the sample size increases to $T = 2000, 4000$ and $T = 8000$, respectively.

5 Empirical illustration

In this Section we illustrate the merits of the independence based structural MGARCH approach by analysing a four dimensional system of the US and three Latin American stock market returns. In this context one might regard US stock markets as important issuers of information. Hence, profiles of volatility transmission that are implied by the often used symmetric covariance matrix decomposition might be considered critical for such a system. Next, we briefly introduce the data and provide QML estimation results from the BEKK(1,1,1) model. After QML estimation least dependent MGARCH innovations are extracted by means of the moment based estimator described in Section 2.2, and we discuss implications of the estimated structural model ($\{D_t\}_{t=1}^T$) and an ad-hoc counterpart of symmetric covariance decompositions ($\{H_t^{1/2}\}_{t=1}^T$).

5.1 Data

We study data of daily nominal local-currency stock market indices ranging from January 1992 to November 2007 and taken from Diebold and Yilmaz (2009). The considered system comprises weekly real returns from the US market (r_{1t}), and three Latin American stock markets namely Argentina (r_{2t}), Brazil (r_{3t}) and Chile (r_{4t}). Following Diebold and Yilmaz (2009) weekly returns are changes in log prices, Friday-to-Friday (or Thursday when Friday is not available) subsequently converted from nominal to real terms by means of consumer price indices from the IMF's International Financial Statistics.[15] The sample size is $T = 829$. We consider time invariant conditional return expectations, $\mu_t = \mu$, and subject centered real returns to QML MGARCH estimation.[16] The analysed real return series are displayed in Figure 2.

[15]Diebold and Yilmaz (2009) take the data from Datastream and Global Financial Data. We use already processed data provided along with corresponding RATS code on `https://estima.com/procedures/dieboldyilmaz_ej2009.zip`. For a more detailed description of the data transformations see Diebold and Yilmaz (2009).

[16]Average log returns are 0.11E-02 (US), -0.21E-03 (Argentina), 0.23E-02 (Brazil) and 0.11E-02 (Chile).

5.2 QML estimation and rotation

QML estimates $\hat{\theta}$ (with t-ratios in parentheses) for the BEKK(1,1,1) model are documented in Appendix B.[17] As also indicated by suitable Wald type tests,[18] the empirical model differs significantly from the diagonal BEKK model, such that cross equation (i.e. cross market) effects are important to describe conditional (co)variance dynamics. Figure 3 shows the MGARCH assessment of conditional standard deviations of the analysed vector returns. The multivariate variance assessment captures suitably clustering patterns of return variation and highlights both common and market specific phases of excess stock market risks.

Minimizing squared deviations of third and fourth order cross moments of structural shocks from theoretical counterparts obtains estimated structural MGARCH matrices as

$$D_t = H_t^{1/2} R_{\hat{\delta}}, \tag{38}$$

where

$$R_{\hat{\delta}} = \begin{pmatrix} 0.863 & -0.134 & -0.324 & -0.364 \\ 0.348 & 0.805 & 0.466 & 0.113 \\ 0.142 & -0.508 & 0.823 & -0.210 \\ 0.338 & -0.274 & 0.002 & 0.900 \end{pmatrix}.$$

Comparing the moment statistics in (10) obtained from a symmetric decomposition ($\delta = 0$, $R_\delta = I_N$) and the model after rotation is supportive for the asymmetric structural specification. Samples of estimated orthogonal shocks $\{H_t^{-1/2} e_t\}_{t=1}^T$ and $\{R_{\hat{\delta}} H_t^{-1/2} e_t\}_{t=1}^T$

[17] To estimate the four dimensional BEKK model we use a modified version of the module 'arch_mg.src' that is comprised in *JMulTi* (Lütkepohl and Krätzig 2005, `http://www.jmulti.de/`). The modifications consist mainly in extending the number of alternative initializations of the iterative estimation procedure, and the selection of initial values from preestimates of bivariate submodels. The estimated parameters are checked to correspond to a maximum of the log-likelihood function. Given that the multivariate GARCH innovations ξ_t are not multivariate Gaussian distributed, the QML covariance matrix is more reliable for diagnosing parameter significance than corresponding ML quantities.

[18] Subjecting all off-diagonal parameters of the BEKK matrices A and B to a joint significance test obtains a QML χ^2-statistic of 121.74 which is significant at conventional levels according to an asymptotic $\chi^2(24)$-distribution.

obtain test statistics of $\lambda_{\delta=0} = 126.6$ and $\lambda_{\hat{\delta}} = 39.73$, respectively. While the former is clearly significant at any conventional level, the latter obtains an associated p-value of 0.76 from a supposed χ^2-distribution with 47 degrees of freedom. Moreover, the data based $\lambda_{\hat{\delta}}$ statistic obtains a p-value of 0.802 when subjecting the search for infimum statistics to resampling with 1000 replications. Hence, estimated structural shocks $\{R_{\hat{\delta}} H_t^{-1/2} e_t\}_{t=1}^T$ can be considered independent with conventional significance. To unravel if the statistically identified MGARCH model is economically reasonable, we next illustrate model implied patterns of volatility transmission and reception among US and Latin American stock markets.

5.3 Volatility transmission

Time varying elements of estimated covariance decompositions are displayed in Figure 4. To underpin the potential merits of the structural model, single panels of Figure 4 jointly display respective elements of $\widehat{H}_t^{1/2}$ and $\widehat{D}_t = \widehat{H}_t^{1/2} R_{\hat{\delta}}$. For the interpretation of the estimation results it is important to notice that the US market is ordered first within the dynamic system. By construction, volatility transmission and reception patterns retrieved from $H_t^{1/2}$ are symmetric. Along the diagonal panels of Figure 4 we see how the MGARCH models quantify the effects of (structural) shocks materialising at their 'own' market. Confirming a-priori intuition, the two alternative model specifications almost agree in their assessment of US market 'own' effects.[19]

A-priori one would expect the Latin American markets to be more affected by volatility transmission from the US than vice versa. The estimated structural model $\{\widehat{D}_t\}$ implies that all Latin American markets are more affected by innovations in the US markets than under model symmetry ($\{\widehat{H}_t^{1/2}\}$). Likewise, the US market is less affected by Latin

[19] Any candidate decomposition of H_t is 'unique' up to the sign and ordering of its columns. We have ordered the columns of decomposition matrices such that in each column the respective diagonal element is the largest. By implication of this ordering, structural shocks have strongest impacts on the market of their origin (Lütkepohl and Netšunajev, 2014).

American markets under the estimated structural model.

As a reflection of relative market capitalization, among the Latin American markets the Brazilian market is least affected by the US and shows almost a coincidence of 'own' effects under symmetric and asymmetric volatility transmission. Specifically, in the more flexible asymmetric model and in comparison with implications of model symmetry, the Brazilian market becomes a more active transmitter of volatility towards the Argentinian market which is the smallest in terms of market capitalization within the considered system.[20]

In the time dimension, volatility spillovers operating from the US to the Latin American markets are of particular strength during the period 10/03/1997 to 09/03/1999, which roughly corresponds to the great economic recession in Argentina (1998-2002), and to the depreciation of the Brazilian real in 1999 (Samba effect).

6 Modelling and controlling portfolio risks

To further highlight that the merits of the identified model go beyond the minimization of the statistical criterion in (10), we discuss in this section if the distinction of alternative MGARCH model structures (i.e. contrasting $H_t^{1/2}$ vs. D_t) is useful for an active management of portfolio risks. Henceforth, we refer to the alternative structural models as the symmetric model ($H_t^{1/2}$) and the identified or asymmetric model (D_t). For both types of risk analysis described in Section 3 we choose a variety of five alternative nominal coverage levels, namely $\alpha = 0.010, 0.025, 0.050, 0.100, 0.250$ (first order risk modelling) and $\gamma = 0.750, 0.900, 0.950, 0.975, 0.990$ (kurtosis risk modelling).[21] Next we describe briefly

[20] Worldbank data indicate for 2000 a market capitalization of US firms of 15108 trillion USD. In relative terms the market capitalization of Argentina, Brazil and Chile has been, respectively, 0.3%, 1.5% and 0.4%. Hence, relative to the Brazilian market, Argentinian and Chilean market capitalizations have been 20.25% and 26.7%. Source: *https://data.worldbank.org/indicator/CM.MKT.LCAP.CD?locations=AR-BR-CL-US*, retrieved on Dec, 20th, 2017.

[21] For assessing first order risks one typically considers small nominal levels ($\alpha = 0.01, 0.025$, say). Since our interest is also in the scope of alternative model specifications to capture the conditional return distribution more generally, we consider larger nominal levels ($\alpha = 0.10, 0.25$) which correspond to more

alternative portfolios and the simulation based determination of portfolio returns and their distributional properties. Subsequently, we provide empirical outcomes for alternative risk assessments within the four dimensional system of stock markets introduced in Section 5. For both types of risk modelling the analysis comprises (i) a graphical illustration of risks based on the equal weight portfolio, (ii) a listing of applied loss functions, and (iii) a comparative discussion of empirical loss statistics.

6.1 Portfolios and the simulation of return distributions

We consider (i) six stylized portfolios, denoted P_1 to P_6 and (ii) 1000 portfolios with random structure of positive weights ($w_i > 0, \sum_{i=1}^4 w_i = 1$).[22] Throughout, the simulation of portfolio returns and their higher order properties is conditional on the processes of estimated covariances H_t and their alternative decompositions $\{D_t\}$ and $\{H_t^{1/2}\}$. Our aim is to model the specific distribution of portfolio returns $\tau_t|\mathcal{F}_{t-1}$ which we can suitably contrast with empirical outcomes. Accordingly, performance statistics (loss measures) are obtained from averaging over samples of empirical portfolio returns $\{\tau_t\}_{t=1}^T$, whereas the conditional return features are determined at each time instance t over the number of generated bootstrap samples.

Noticing that our interest is in the behaviour of (powers of) returns in the tail of the respective distributions and in modelling conditional probabilities or interquartile ranges of portfolio returns, all simulations rely on $B = 10^6$ bootstrap samples. Bootstrap

frequent risky states.

[22]The six stylized portfolios are the equal weight portfolio ($w_i = 1/4, \mathrm{P}_1$), minimum variance portfolios (P_2), and portfolios where one market enters with weight $w_i = 0.5$ and the remaining markets are equally weighted ($w_1 = 0.5$, P_3; $w_2 = 0.5$, P_4; $w_3 = 0.5$, P_5; $w_4 = 0.5$, P_6). Time varying weight vectors of the minimum variance portfolios are determined as

$$w_t = H_t^{-1}\mathbf{1}/c,$$

where $c = \mathbf{1}'H_t^{-1}\mathbf{1}$ and $\mathbf{1}$ is a four dimensional vector of ones. Opposite to both the remaining stylized portfolios and simulated portfolio structures, portfolio weights of minimum variance portfolios can be smaller than zero. Simulated portfolio weights obtain as $w_i = \tilde{w}_i/(\sum_{i=1}^4 \tilde{w}_i)$, where $\tilde{w}_i$, $i = 1, \ldots, 4$, are drawn from the uniform distribution.

samples $\{\xi_t^*\}_{t=1}^T$ are drawn from estimated innovations $\{\hat{\xi}_t\}_{t=1}^T$ originating alternatively from the symmetric ($\hat{\xi}_t = \widehat{H}_t^{-1/2} e_t$) or the asymmetric model ($\hat{\xi}_t = \widehat{D}_t^{-1} e_t$). Assuming cross equation independence, we compose bootstrap vectors ξ_t^* by drawing its elements ξ_{it}^* with replacement from the marginal distributions, i.e. from $\{\hat{\xi}_{it}\}_{t=1}^T$.[23] After their generation, bootstrap vectors $\{\xi_t^*\}_{t=1}^T$ are used to obtain samples of bootstrap returns based alternatively on the symmetric ($e_t^* = H_t^{1/2}\xi_t^*$) or the asymmetric model ($e_t^* = D_t\xi_t^*$). Let b, $b = 1, 2, \ldots, B$, be an index to distinguish single bootstrap draws, where $B = 10^6$. Then, bootstrap samples of portfolio returns read explicitly as

$$\{\{\tau_{t,b}^* = w' e_{t,b}^*\}_{t=1}^T\}_{b=1}^B. \tag{39}$$

In the following we omit the replication index b for notational convenience. It is important, however, to keep in mind that bootstrap portfolio returns are dependent on the distinction between the symmetric and asymmetric MGARCH model. In each time instance, the samples in (39) are used to determine risk statistics that have been introduced in Section 3. In particular, we determine VaR, ES, KaR, the interquartile ranges for conditional exceedance statistics (see (25) and (36) in Section 3), and model the probability that shortfall returns are less than the negative expected shortfall.

6.2 Model performance in first order risk analysis

6.2.1 Loss functions

Let $I\{\}$ denote an indicator function. To rank alternative model specifications we consider the following loss functions that apply to portfolio returns falling short of the conditional VaR (i.e. $\tau_t < -\mathrm{VaR}_{\alpha,t}^{\bullet}$):

[23] Noticing that the symmetric model is not developed under an assumption of cross equation independence one may expect some mismatch of theoretical and simulated moments when sampling conditional on $\{H_t^{1/2}\}_{t=1}^T$.

1. Empirical coverage of conditional VaR estimates

$$\mathcal{L}_1 = \frac{1}{T}\sum_{t=1}^{T} I\left\{\tau_t < -\mathrm{VaR}^{\bullet}_{\alpha}(\tau_t|\mathcal{F}_{t-1})\right\}$$

2. Mean excess shortfall[24]

$$\mathcal{L}_2 = \frac{1}{T}\sum_{t=1}^{T} (|\tau_t| - \mathrm{VaR}^{\bullet}_{\alpha}(\tau_t|\mathcal{F}_{t-1}))I\left\{\tau_t < -\mathrm{VaR}^{\bullet}_{\alpha}(\tau_t|\mathcal{F}_{t-1})\right\}$$

3. Empirical coverage of the simulated interquartile ranges

$$\mathcal{L}_3 = \frac{1}{T}\sum_{t=1}^{T} I\left\{\tau_t \in \left[\underline{\mathrm{ES}}^{\bullet}_{\alpha}(\tau_t|\mathcal{F}_{t-1}), \overline{\mathrm{ES}}^{\bullet}_{\alpha}(\tau_t|\mathcal{F}_{t-1})\right]\right\}$$

4. Assessing the probability that shortfall returns are less than the $\mathrm{ES}^{\bullet}_t$

$$\mathcal{L}_4 = \frac{1}{T}\sum_{t=1}^{T}\left|\mathrm{Prob}^{\bullet}_t\left[\tau_t < -\mathrm{ES}^{\bullet}_{\alpha}(\tau_t|\mathcal{F}_{t-1})\right] - I\left\{\tau_t < -\mathrm{ES}^{\bullet}_{\alpha}(\tau_t|\mathcal{F}_{t-1})\right\}\right|$$

While the first loss functions is common for VaR assessments, $\mathcal{L}_2$ provides an empirical counterpart of EES as defined in (24) and is, except for the adjustment by means of $\mathrm{VaR}^{\bullet}_{\alpha}(\tau_t|\mathcal{F}_{t-1}$, in full analogy to common assessments of ES statistics. Providing conceptually distinct loss statistics, the interquartile coverage $\mathcal{L}_3$ and the probability assessment $\mathcal{L}_4$ have not yet been considered in the related literature. Both, however, are particularly informative for a risk model's scope to describe distributional properties of unfavourable states of portfolio performance. By construction, empirical outcomes for $\mathcal{L}_3$ are the more favourable the closer these estimates are to a nominal coverage of 50%. The fourth loss function consists of comparing the absolute distance between model implied probabilities

[24]The informative content of mean (excess) shortfall statistics might suffer from single outlying return observations. Unreported results determined for robust median shortfall statistics are qualitatively very similar to those documented for $\mathcal{L}_2$.

and binary empirical outcomes. Hence, risk monitoring is the more effective the smaller is $\mathcal{L}_4$.

6.2.2 Time varying first order risks for the equal weight portfolio

Figure 5 displays empirical equal weight portfolio returns (first line) joint with VaR estimates at levels $\alpha = 0.025$ (second line) and $\alpha = 0.01$ (third line). Corresponding estimates of expected excess shortfall (EES) are shown in the fourth ($\alpha = 0.025$) and fifth panel ($\alpha = 0.01$) of Figure 5. Eyeball inspection of Figure 5 reveals that all dynamic risk assessments cope suitably with variation clusters which can be seen for the process of portfolio returns. Concerning the VaR estimates, both the symmetric and the asymmetric MGARCH model issue very similar risk estimates at the nominal level $\alpha = 0.025$ and slightly distinct statistics for $\alpha = 0.01$. For both levels the VaRs obtained from the asymmetric model are somewhat more conservative, i.e. larger in absolute value. As it becomes apparent from the comparison of model implied EES statistics, the symmetric and the asymmetric MGARCH specification differ in particular with respect to the assignment of probabilities to tail events of (very) small portfolio returns. The EES statistics issued from the asymmetric model are throughout more conservative, i.e., larger in absolute value.

6.2.3 Empirical loss statistics

Core summary statistics for modelling equal weight portfolio risk are documented in the two leftmost columns of Table 2. In terms of the loss statistics $\mathcal{L}_1$ (coverage) and $\mathcal{L}_2$ (mean excess shortfall) and taking a joint perspective over all considered nominal coverage levels, the performance of both alternative MGARCH variants is similar. For 4 out of 5 nominal coverage levels, however, the asymmetric model obtains more favourable outcomes for $\mathcal{L}_3$ and $\mathcal{L}_4$. For instance, in case of the nominal level of $\alpha = 0.01$ ($\alpha = 0.025$) the empirical coverage of interquartile ranges for shortfall returns ($\mathcal{L}_3$) are 0.125 and 0.40 (0.682 and

0.476) for the symmetric and asymmetric model, respectively.

Summarizing the results for six stylized portfolios (see the left hand side panel of Table 3), we find that the asymmetric model improves accuracy of risk assessments with empirical VaR levels coming closer to their nominal counterparts if the nominal level is large (i.e. $\alpha \geq 0.1$) such that the number of critical events increases.

The coverage of interquartile ranges of shortfall returns ($\mathcal{L}_3$) or their probabilities of being less than the expected shortfall ($\mathcal{L}_4$) provides strongest support for the asymmetric model. Nine out of ten respective average loss statistics documented in Table 3 are in favour of the asymmetric model. For instance, with nominal level of $\alpha = 0.01$ and over six portfolios the total counts of shortfall returns are 49 and 34 for the symmetric and the asymmetric model, respectively. Out of these, 13 and 19 shortfall returns are covered by the respective interquartile ranges obtaining empirical coverage frequencies of 26.7% and 55.8%. While the latter cannot be distinguished statistically from the nominal 50% coverage, the former violates this level with 5% significance. On average and conditional on the six stylized portfolios, the empirical coverage of interquartile ranges determined by means of the asymmetric model is closer to the nominal 50% coverage for each choice of α except for $\alpha = 0.05$. In this case we see a slight 'lead' of using the symmetric model.

As documented in the lower panel of Table 3, summary statistics for 1000 portfolios with randomized weight structure largely confirm average results documented for the six stylized portfolios. Modelling probabilities of observing a return below the expected shortfall ($\mathcal{L}_4$) obtains uniformly smaller average approximation errors when the analysis conditions on the asymmetric model. In particular for small nominal coverage levels $\alpha = 0.01, 0.025$ the empirical coverage of interquartile ranges for shortfall returns is more accurate for diagnostics developed from the asymmetric model.

6.3 Model performance in kurtosis risk analysis

6.3.1 Loss functions

To rank alternative model specifications in kurtosis risk assessment we consider two loss functions that apply to time instances when variations of portfolio risks exceed their simulated thresholds (i.e. $\widetilde{m}_t^2 > \mathrm{KaR}^\bullet_{\gamma,t}$):

1. Empirical frequency of excess variations of portfolio risks ($\mathcal{M}_1$)

$$\mathcal{M}_1 = \frac{1}{T}\sum_{t=1}^{T} I\left\{\widetilde{m}_t^2 > \mathrm{KaR}^\bullet_\gamma(\widetilde{m}_t^2|\mathcal{F}_{t-1})\right\}$$

2. Coverage of such events offered by the simulated interquartile ranges ($\mathcal{M}_2$).

$$\mathcal{M}_2 = \frac{1}{T}\sum_{t=1}^{T} I\left\{\widetilde{m}_t^2 \in \left[\underline{\mathrm{EKS}}^\bullet_\gamma(\widetilde{m}_t^2|\mathcal{F}_{t-1}), \overline{\mathrm{EKS}}^\bullet_\gamma(\widetilde{m}_t^2|\mathcal{F}_{t-1})\right]\right\}$$

6.4 Time varying variations of portfolio risks

For the case of equal weight portfolios Figure 6 shows the empirical processes of portfolio standard deviations σ_t, squared deviations from the conditional variance $\widetilde{m}_t^2$ (see (33)) and expected (scaled) variations of portfolio risks ($\tilde{v}_t^{(1)}$ and $\tilde{v}_t^{(0)}$, see (32)).

The empirical profile of $\widetilde{m}_t^2$ exhibits a couple of strong outliers which coincide with sizeable new information entering the conditional variance process.[25] Noticing that the likelihood of outlying observations in empirical portfolio returns increases with the kurtosis of the return distribution, it is intuitive to observe that both displayed mean profiles $\tilde{v}_t^{(0)}$ and $\tilde{v}_t^{(1)}$ imply time varying kurtosis levels in excess of 3. Throughout, model implied kurtosis (minus 1!) statistics are markedly larger for the asymmetric model ($\tilde{v}_t^{(0)} \approx 3.2$ on average (with standard deviation of 0.07)) in comparison with statistics obtained from

[25]The strong dispersion of the empirical distribution of $\widetilde{m}_t^2$ is also reflected in the fact that, on average, 8.08%, 6.03% and 3.25% of all observations characterizing equal weight portfolios are above thresholds of 3, 5 and 10, respectively.

the symmetric model ($\tilde{v}_t^{(1)} \approx 2.5$ (0.06)). Seeing large and frequent outlying observations for $\widetilde{m}_t^2$, we may expect from the kurtosis differential that the asymmetric model has some lead in managing tail events.

6.4.1 Results for stylized portfolios

Complementing the graphical displays in Figure 6, Table 4 summarizes the performance of alternative approaches to quantify kurtosis risk of equal weight portfolio returns. Except for the least and most conservative nominal levels, the empirical frequencies of excessive statistics $\widetilde{m}_t^2$ ($\mathcal{M}_1$) are throughout closer to the nominal counterparts when conditioning the analysis on the asymmetric model. The symmetric model provides risk thresholds which are more conservative in comparison with the asymmetric model. Using the simulated interquartile ranges for conditional interval 'prediction' ($\mathcal{M}_2$), we find that the quantities determined by means of the asymmetric model are more trustworthy for several nominal levels γ. In particular, for the most conservative nominal level ($\gamma = 0.99$) we have that both model variants yield five violations of the respective quantile. None of these violations is covered by the interquartile range determined under model symmetry, while the asymmetric specification obtains intervals capturing 3 out of 5 critical events.

Core results discussed for the equal weight portfolio also hold for the remaining stylized portfolios. Summary loss statistics for all stylized portfolios are displayed in the left hand side panel of Table 5. With a few exceptions, the realized mean excesses over risk thresholds are smaller for the asymmetric model. Empirical frequencies of kurtosis risks in excess of the specified thresholds are closer to the nominal counterparts for the asymmetric models. Moreover, the coverage of the interquartile range of simulated excessive risks is generally closer to the nominal 50% coverage for the asymmetric specification. While extreme events characterized by setting $\gamma = 0.99$ are rare for a given portfolio, aggregating the number of respective threshold violations over all six portfolios obtains that empirical statistics $\widetilde{m}_t^2$ exceed the thresholds implied by the asymmetric and symmetric model in

35 and 36 cases, respectively. From these instances of quite strong surprises to risk (in total) 48.6% and 33.3% are covered by the model specific interquartile ranges. While the former statistic is in line with the nominal coverage, the latter differs from its nominal reference with 5% significance.

6.4.2 Results for randomized portfolio compositions

Modelling results for kurtosis risk analysis applied to 1000 portfolios with randomized weight structures are shown in the right hand side panel of Table 5. In terms of threshold exceedances ($\mathcal{M}_1$) both MGARCH variants perform somehow conservative, however, for nominal levels $\gamma = 0.75, 0.90, 0.95, 0.975$ the empirical frequencies of excess risks are closer to their nominal counterpart for the asymmetric model. Noticing the huge number of simulated cases ($T = 829$ times $(1 - \hat{\gamma})$ times 1000 simulated portfolios) performance differentials between the two model alternatives are in excess of two times the respective standard error for the nominal levels $\gamma = 0.75, 0.90, 0.95$ and $\gamma = 0.975$. At levels of $\gamma = 0.975$ and $\gamma = 0.990$ interquartile ranges of excessive risks derived from the symmetric model cover their empirical counterparts with frequencies of 38% and 31%, respectively. Using the asymmetric model for this purpose obtains interquartile ranges with improved precision, covering respective events with frequencies of 42.6% and 55.8%, respectively.

7 Conclusions

Being widely used in empirical practice for various aspects of portfolio monitoring and management, multivariate GARCH (MGARCH) models often lack an interpretation in a structural sense beyond a-priori considerations. Building upon recent advances in the data based identification of structural VARs, we exploit the uniqueness of independent non-Gaussian innovations in MGARCH processes to determine covariance decompositions in a data based manner.

We show that the structural model is asymptotically well identified if the structural analysis follows consistent QML estimation. We provide simulation based evidence pointing at consistency of the proposed identification scheme and sensitivity to stronger deviations from the unidentified conditionally Gaussian MGARCH model.

Our empirical analysis provides structural insights into volatility transmission and reception characterizing a four dimensional system of US and Latin American stock markets (Argentina, Brazil, Chile). The devised structural model obtains volatility transmission patterns which are better justified in economic terms in comparison with corresponding profiles retrieved from a symmetric ad-hoc covariance decomposition. Moreover, the identified structural model turns out preferable when it comes to an active management of first order (conditional VaR, expected shortfall) and higher order (conditional kurtosis) risk patterns inherent in portfolio returns.

Throughout, the analysis in this work relies on the supposition of parametric stability of the MGARCH model specification. While stability of reduced form parameters might be tested with conventional QML methods, the detection of changing time profiles of volatility transmission (conditional on estimated covariances) appears as a promising direction for future research. In this respect, the moment based diagnostic suggested in this work might provide an interesting starting point for sensitivity analysis by means of recursive or adaptive modelling techniques (Foster and Nelson, 1996; Härdle et al., 2003; Golosnoy et al., 2012; Ibragimov and Müller, 2010).

References

Angelidis, T. and S. Degiannakis (2007): "Backtesting VaR models: A two-stage procedure," *Journal of Risk Model Validation*, 1, 27–48.

Baele, L. (2005): "Volatility spillover effects in European equity markets," *Journal of Financial and Quantitative Analysis*, 40, 373–401.

Bali, T. G. and A. Hovakimian (2009): "Volatility Spreads and Expected Stock Returns," *Management Science*, 55, 1797–1812.

Bauwens, L., S. Laurent, and J. V. K. Rombouts (2006): "Multivariate GARCH models: A survey," *Journal of Applied Econometrics*, 31, 79–109.

Bekaert, G. and C. R. Harvey (1995): "Time-varying world market integration," *The Journal of Finance*, 50, 403–444.

Billio, M., M. Getmansky, A. W. Lo, and Pelizzon (2012): "Econometric measures of connectedness and systemic risk in the finance and insurance sectors," *Journal of Financial Economics*, 104, 535–559.

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

——— (1990): "Modeling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model," *Review of Economics and Statistics*, 72, 498–505.

Bollerslev, T., R. F. Engle, and J. M. Wooldridge (1988): "A capital asset pricing model with time varying covariances," *Journal of Political Economy*, 96, 116–131.

Boussama, F., F. Fuchs, and R. Stelzer (2011): "Stationarity and geometric ergodicity of BEKK multivariate GARCH models," *Stochastic Processes and their Applications*, 121, 2331–2360.

Chan, K. C., G. A. Karolyi, and R. Stulz (1992): "Global financial markets and the risk premium on US equity," *Journal of Financial Economics*, 32, 137–167.

Comon, P. (1994): "Independent component analysis, A new concept?" *Signal Processing*, 36, 287–314.

Comte, F. and O. Lieberman (2003): "Asymptotic Theory for Multivariate GARCH processes," *Journal of Multivariate Analysis*, 84, 61–84.

Davidson, J. (1994): *Stochastic Limit Theory*, Oxford University Press.

Diebold, F. X. and K. Yilmaz (2009): "Measuring Financial Asset Return and Volatility Spillovers, With Application to Global Equity Markets," *The Economic Journal*, 119, 158–171.

Dimpfl, T. and R. C. Jung (2012): "Financial market spillovers around the globe," *Applied Financial Economics*, 22, 45–57.

Engle, R. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1007.

Engle, R. F. (2002): "Dynamical Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroscedastic Models," *Journal of Business and Economic Statistics*, 20, 339–350.

Engle, R. F., T. Ito, and W. L. Lin (1990): "Meteor Showers or Heat Waves? Heteroskedastic Intra- Daily Volatility in the Foreign Exchange Market," *Econometrica*, 58, 525–542.

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

ESA (2016): "Final draft regulatory technical standards," Working paper, Joint Committee of the European Supervisory Authorities ESMA, EBA, and EIOPA.

Faust, J. (1998): "The robustness of identified VAR conclusions about money," *Carnegie-Rochester Conference Series on Public Policy*, 49, 207–244.

Fengler, M. R. and H. Herwartz (2018): "Measuring spot variance spillovers when (co)variances are time-varying – the case of multivariate GARCH models," *Oxford Bulletin of Economics and Statistics*, 80, 135?159.

Forbes, K. J. and R. Rigobon (2002): "No contagion, only interdependence: Measuring stock market comovements," *Journal of Finance*, 57, 2223–2261.

Foster, D. and D. B. Nelson (1996): "Continuous Record Asymptotics for Rolling Sample Variance Estimators," *Econometrica*, 64, 139–174.

Fountas, S. and M. Karanasos (2007): "Inflation, output growth, and nominal and real uncertainty: empirical evidence for the G7," *Journal of International Money and finance*, 26, 229–250.

Golosnoy, V., I. Okhrin, and W. Schmid (2012): "Statistical Surveillance of Volatility Forecasting Models," *Journal of Financial Econometrics*, 10, 513–543.

Gouriéroux, C., A. Monfort, and J. Renne (2017): "Statistical inference for independent component analysis: Application to structural VAR models," *Journal of Econometrics*, 196, 111–126.

Hafner, C. M. and H. Herwartz (2006a): "Volatility impulse responses for multivariate GARCH models: An exchange rate illustration," *Journal of International Money and Finance*, 25, 719–740.

——— (2006b): "A Lagrange multipler test for causality in variance," *Economics Letters*, 93, 137–141.

Hafner, C. M. and A. Preminger (2009): "On asymptotic theory for multivariate GARCH models," *Journal of Multivariate Analysis*, 100, 2044–2054.

Hafner, C. M. and J. V. Rombouts (2007): "Semiparametric multivariate volatility models," *Econometric Theory*, 23, 251?280.

Hamao, Y., R. Masulis, and V. K. Ng (1990): "Correlations in Price Changes and Volatility across International Stock Markets," *Review of Financial Studies*, 3, 281–307.

Härdle, W., H. Herwartz, and V. Spokoiny (2003): "Time Inhomogenous Multiple Volatility Modelling," *Journal Financial Econometrics*, 1, 55–95.

Hassan, S. A. and F. Malik (2007): "Multivariate GARCH modeling of sector volatility transmission," *The Quarterly Review of Economics and Finance*, 47, 470–480.

Herwartz, H. and J. Roestel (2016): "Market dominance targeting - A flexible structural framework to assess transmission in financial markets," Research paper, University of Göttingen.

Ibragimov, R. and U. K. Müller (2010): "t-Statistic Based Correlation and Heterogeneity Robust Inference," *Journal of Business and Eocnomics Statistics*, 28, 453–468.

Jeantheau, T. (2000): "Strong Consistency of Estimators for Multivariate ARCH Models," *Econometric Theory*, 14, 70–86.

Lancaster, H. O. (1954): "Traces and Cumulants of Quadratic Forms in Normal Variables," *Journal of the Royal Statistical Society. Series B (Methodological)*, 16, 247–254.

Lanne, M., M. Meitz, and P. Saikkonen (2017): "Identification and estimation of non-Gaussian structural vector autoregressions," *Journal of Econometrics*, 196, 288–304.

Lin, W.-L. (1997): "Impulse response function for conditional volatility in GARCH models," *Journal of Business & Economic Statistics*, 15, 15–25.

Lütkepohl, H. (1996): *Handbook of Matrices*, John Wiley & Sons, Ltd., Chichester.

Lütkepohl, H. and L. Kilian (2017): *Structural Vector Autoregressive Analysis*, Cambridge: Cambridge University Press.

Lütkepohl, H. and A. Netšunajev (2014): "Structural Vector Autoregressions with Smooth Transition in Variances: The Interaction Between U.S. Monetary Policy and the Stock Market," DIW Discussion Paper 1388.

Matteson, D. S. and R. S. Tsay (2017): "Independent Component Analysis via Distance Covariance," *Journal of the American Statistical Association*, 112, 623–637.

McNeil, A., R. Frey, and P. Embrechts (2nd edition, 2016): *Quantitative Risk Management*, Princeton, NJ: Princeton University Press.

McNeil, A. J. and R. Frey (2000): "Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach," *Journal of Empirical Finance*, 7, 271–300.

Miettinen, J., S. Taskinen, K. Nordhausen, and H. Oja (2015): "Fourth moments and independent component analysis," *Statist. Sci.*, 30, 372–390.

Moneta, A., D. Entner, P. O. Hoyer, and A. Coad (2013): "Causal Inference by independent component analysis: Theory and applications," 75, 705–730.

Nakatani, T. and T. Teräsvirta (2009): "Testing for volatility interactions in the Constant Conditional Correlation GARCH model," *Econometrics Journal*, 12, 147–163.

Rigobon, R. (2003): "Identification through heteroskedasticity," *The Review of Economics and Statistics*, 85, pp. 777–792.

Romer, C. D. and D. H. Romer (2010): "The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks," *The American Economic Review*, 100, 763–801.

Sims, C. A. (1980): "Macroeconomics and reality," *Econometrica*, 48, pp. 1–48.

Stelzer, R. (2008): "On the Relation between the Vec and BEKK Multivariate GARCH Models," *Econometric Theory*, 24, 1131–1136.

Uhlig, H. (2005): "What are the effects of monetary policy on output? Results from an agnostic identification procedure," *Journal of Monetary Economics*, 52, 381–419.

van der Weide, R. (2002): "GO-GARCH: A multivariate generalized orthogonal GARCH model," *Journal of Applied Econometrics*, 17, 549–564.

Weber, E. (2010): "Structural Conditional Correlation," *Journal of Financial Econometrics*, 8, 392–407.

White, H. (1984): *Asymptotic theory for econometricians*, Economic Theory, Econometrics and Mathematical Economics Series, Academic Press.

Woźniak, T. (2015): "Testing causality between two vectors in multivariate GARCH models," *International Journal of Forecasting*, 31, 876–894.

Zhu, D. and J. W. Galbraith (2011): "Modeling and forecasting expected shortfall with the generalized asymmetric Student-t and asymmetric exponential power distributions," *Journal of Empirical Finance*, 18, 765–778.