Mathematical Statistics

Stockholm University

Singular conditional autoregressive Wishart model for realized covariance matrices

Gustav Alfelt, Taras Bodnar, Farrukh Javed, Joanna Tyrcha

Research Report 2020:8

ISSN 1650-0377

Postal address:
Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:
http://www.math.su.se

Mathematical Statistics
Stockholm University
Research Report **2020:8**,
http://www.math.su.se

Singular conditional autoregressive Wishart model for realized covariance matrices

Gustav Alfelt *[1], Taras Bodnar[1], Farrukh Javed[2], and Joanna Tyrcha[1]

[1] *Department of Mathematics, Stockholm University, SE-10691 Stockholm, Sweden*

[2] *Örebro University School of Business, Örebro University, SE-70281 Örebro, Sweden*

September 2020

Abstract

Realized covariance matrices are often constructed under the assumption that richness of intra-day return data is greater than the portfolio size, resulting in non-singular matrix measures. However, when for example the portfolio size is large, assets suffer from illiquidity issues, or market microstructure noise deters sampling on very high frequencies, this relation is not guaranteed. Under these common conditions, realized covariance matrices may obtain as singular by construction. Motivated by this situation, we introduce the Singular Conditional Autoregressive Wishart (SCAW) model to capture the temporal dynamics of time series of singular realized covariance matrices, extending the rich literature on econometric Wishart time series models to the singular case. This model is furthermore developed by covariance targeting adapted to matrices and a sectorwise BEKK-specification, allowing excellent scalability to large and extremely large portfolio sizes. Finally, the model is estimated to a 20 year long time series containing 50 stocks, and evaluated using out-of-sample forecast accuracy. It outperforms the benchmark Multivariate GARCH model with high statistical significance, and the sectorwise specification outperforms the baseline model, while using much fewer parameters.

Keywords: Time series matrix-variate model, Realized covariance matrix, High-dimensional data, Covariance targeting, Stock co-volatility.

JEL Classification: C32, C55, C58, G17.

*Corresponding Author: Gustav Alfelt. E-Mail: gustava@math.su.se

1 Introduction

The covariance matrix of asset returns plays a key role in several financial applications, such as portfolio allocation, risk management and option pricing. It is well-documented that this quantity changes over time, why describing and understanding its temporal dynamics is fundamental to financial decision making. A typical approach is to capture this evolution in discrete time by applying multivariate GARCH-type models, summarized in Bauwens et al. (2006), where the conditional covariance matrix is determined by past observations of daily returns. Another classic method is to use multivariate stochastic volatility-type models, reviewed in Asai et al. (2006), in where the covariance matrix process is assumed to be random.

During the last two decades, increased availability of asset price data on high frequencies has paved the way for numerous novel approaches in this area. Many of them are built upon the notion of realized covariance, where the daily return covariance matrix is estimated by a large number of intra-day returns, on for example five or ten minute intervals (see e.g. Andersen et al. (2003) and Barndorff-Nielsen and Shephard (2004)). Modeling the time series dynamics for realized covariance matrices in discrete time has become a large branch in the econometric literature. A popular approach is to assume the underlying stochastic process to be Wishart, a well-studied distribution that ensures positive-definiteness almost surely. For example, the Wishart Autoregressive (WAR) model introduced in Gouriéroux et al. (2009), assumes realized covariances are conditionally distributed as non-central Wishart, where the non-centrality parameter is described by historical realized covariance matrices. The High-Frequency-Based Volatility (HEAVY) model presented in Noureldin et al. (2012) and the Conditional Autoregressive Wishart (CAW) model introduced by Golosnoy et al. (2012) rely on the centralized Wishart distribution, where the scale matrix parameters are determined by past observations. A central Wishart distribution is also considered in Jin and Maheu (2012), but here the scale matrix is decomposed into either multiplicative components or additive components determined by sample means of historical realized covariances. The General Conditional Autoregressive Wishart (GCAW) model is proposed in Yu et al. (2017), parameterized with both a non-central parameter as in the WAR model and a scale matrix as in CAW model. In Anatolyev and Kobotaev (2018), the CAW model is extended to the Conditional Threshold Autoregressive Wishart (CTAW) model with the aim to include the effects of price asymmetry on future realized covariances. Goodness-of-fit tests for models driven by an underlying central Wishart distribution, such as the CAW model, is presented in Alfelt et al. (2019).

All of the models discussed above assume a realized covariance matrix that is positive definite. This can be ensured as long as the number, n, of intra-day returns used to compute the realized covariance matrix is larger than or equal to the number, m, of assets included into the portfolio. Regarding small and moderately sized portfolios or reasonably liquid assets, this relation can often be justified. However, in many applications it is of interest to consider portfolios of large dimensions, containing perhaps 50, 100 or even 1000 assets (see, e.g., Hautsch et al. (2015), Bodnar et al. (2020), Cai et al. (2020), Ding et al. (2020)). Furthermore, available data for the portfolio assets might be restricted, due to for example low liquidity resulting in a few price quotes per day. In addition, there might be limits to how high of an intra-day return sample frequency that is suitable, in presence of so-called market microstructure noise (see e.g. Aït-Sahalia and Yu (2009)). Any combination of these factors might result in a situation where $m > n$, and hence daily realized covariance matrices that are singular. Finally, Jacod and Podolskij (2013) derived an asymptotic test for inferring the rank of multivariate volatility processes.

This paper focuses on time series of singular realized covariance matrices, and extends the family of econometric autoregressive Wishart models to the singular case by introducing the Singular Conditional Autoregressive Wishart (SCAW) model to describe such time series. It is based on the assumption that realized covariance matrices follow a conditional singular Wishart distribution, described in, e.g., Srivastava (2003), Bodnar and Okhrin (2008), where the scale matrix parameter is determined by historical observations in an autoregressive fashion similar to the BEKK-specification of Engle and Kroner (1995), alike for example the CAW model in Golosnoy et al. (2012). This specification ensures positive definiteness and allows us to directly estimate the model parameter with the maximum likelihood method. Furthermore, parameter-based conditions for weak stationarity of the model are deduced.

Since the singular case is closely related to portfolios of large dimensions, a challenge in this setting is to capture the temporal dynamics of the time series, while simultaneously maintaining a parsimonious model that can be feasibly estimated. To deal with this scaling challenge, two novel approaches are introduced. The first one regards covariance targeting (see e.g. Pedersen and Rahbek (2014)), where the approach of Noureldin et al. (2014) is adapted to the matrix case. It concerns standardizing the time series by its unconditional mean, which allows implementing straightforward conditions on the model parameters such that positive definiteness is maintained also under a covariance targeting regime. This method circumvents estimating the large number of parameters present in the constant matrix of the BEKK-specification, greatly increasing estimation feasibility. The second approach utilizes the similarity of assets that belong to the same market sector. This specification assumes that the parameters governing temporal dynamics of the matrix time series are homogeneous for assets of the same sector. As such, the number of parameters does not depend on the number of portfolio assets, but rather of the number of market sectors these assets stem from. Combining these approaches results in a model that is well equipped for implementation on large or extremely large portfolios. In addition, an extension using the heterogeneous autoregressive (HAR) specification, adapted from Corsi (2009), is applied to the SCAW model. This approach considers long-time memory dependence by including historical realized covariance matrices on longer horizons, such as weekly or monthly.

In the empirical part of the paper the SCAW model with various specifications is estimated to a time series of 50 assets traded on the National Association of Securities Dealers Automated Quotations (NASDAQ) over 20 years. It is evaluated by several out-of-sample forecast measures and benchmarked against similarly specified Multivariate GARCH models. The results of the empirical study reveal that the SCAW model outperforms the benchmark model with high statistical significance for the vast majority of the forecasts measures. Moreover, it suggests that the presented sectorwise parameterization indeed can be useful, outperforming the original SCAW model both in and out-of-sample, despite having substantially fewer parameters. Finally, including the HAR-extension to this specification yields the lowest out-of-sample forecast accuracy for several measures, again by only employing a small number of parameters.

The rest of the paper is organized as follows. Section 2 introduces the SCAW model and presents its stochastic properties. In Section 3, covariance targeting, the sectorwise specification, and the HAR-extension are introduced. Section 4 governs the estimation procedure for the SCAW model with its various specifications. The empirical application is presented in Section 5, while Section 6 concludes. Proofs of the obtained theoretical results can be found in the Appendix.

2 Singular Conditional Autoregressive Wishart (SCAW) Model

Let $\mathbf{R}_t$ be an $m \times m$ realized covariance matrix, constructed using n intra-day return vectors recorded during day t. In addition, suppose that the number of intra-day return vectors used in the computation of $\mathbf{R}_t$ is less than the dimension of these vectors, such that $n < m$. As a result, $\mathbf{R}_t$ is a singular matrix by construction. Furthermore, let $\{\mathbf{R}_t\}$ be a discrete time series of such matrices, and let $\mathcal{F}_t$ denote the filtration associated with $\{\mathbf{R}_t\}$. Now, assume that conditioned on $\mathcal{F}_{t-1}$, $\mathbf{R}_t$ follows a singular Wishart distribution of dimension m. That is,

$$\mathbf{R}_t \mid \mathcal{F}_{t-1} \sim SW_m(n, \mathbf{S}_t/n), \tag{1}$$

where $SW_m(\nu, \mathbf{\Sigma})$ denote the singular Wishart distribution with degrees of freedom ν and symmetric, positive-definite scale matrix $\mathbf{\Sigma}$ of dimension $m \times m$. In addition, since $\mathbb{E}[\mathbf{R}_t \mid \mathcal{F}_{t-1}] = \mathbf{S}_t$, the matrix $\mathbf{S}_t$ is the conditional mean matrix of $\{\mathbf{R}_t\}$. Note that the singularity of $\mathbf{R}_t$ stems from the degrees of freedom n being lower than the matrix dimension m, while the conditional mean matrix, $\mathbf{S}_t$, is assumed to be non-singular.

Now, let $\mathbf{R}_t$ be partitioned as

$$\mathbf{R}_t = \begin{bmatrix} \mathbf{R}_{11,t} & \mathbf{R}_{12,t} \\ \mathbf{R}_{21,t} & \mathbf{R}_{22,t} \end{bmatrix}, \tag{2}$$

where $\mathbf{R}_{11,t}$ is an $n \times n$ non-singular matrix, $\mathbf{R}_{12,t}$ is $n \times (m-n)$, $\mathbf{R}_{21,t} = \mathbf{R}'_{12,t}$ and $\mathbf{R}_{22,t}$ is $(m-n) \times (m-n)$ with $\mathbf{R}_{22,t} = \mathbf{R}_{21,t}\mathbf{R}_{11,t}^{-1}\mathbf{R}_{12,t}$. That any singular, symmetric matrix can be partitioned this way is shown by, e.g., Harville (1997, Lemma 9.2.2). Consequently, in accordance with Srivastava (2003) regarding the singular Wishart distribution, the conditional density for $\mathbf{R}_t$ is given by

$$\begin{aligned} f(\mathbf{R}_t \mid \mathcal{F}_{t-1}) &= \frac{\pi^{n(n-m)/2}}{2^{mn/2}\Gamma_n(n/2)|\mathbf{S}_t/n|^{n/2}} |\mathbf{R}_{11,t}|^{(n-m-1)/2} \exp\left(-\frac{1}{2}\mathrm{tr}((\mathbf{S}_t/n)^{-1}\mathbf{R}_t)\right) \\ &= \frac{\pi^{n(n-m)/2} n^{m-n/2}}{2^{mn/2}\Gamma_n(n/2)|\mathbf{S}_t|^{n/2}} |\mathbf{R}_{11,t}|^{(n-m-1)/2} \exp\left(-\frac{n}{2}\mathrm{tr}(\mathbf{S}_t^{-1}\mathbf{R}_t)\right), \end{aligned} \tag{3}$$

where $\Gamma_n(\cdot)$ denotes the multivariate gamma function (see, e.g., Gupta and Nagar (2000)). In addition, the conditional conditional covariance matrix of $\mathbf{R}_t$ consists of the following elements

$$\mathrm{Cov}[r_{ij,t}, r_{kl,t} \mid \mathcal{F}_{t-1}] = \frac{1}{n}(s_{ik,t}s_{jl,t} + s_{il,t}s_{jk,t}), \tag{4}$$

for $i, j, k, l = 1, \ldots, m$, where $r_{ij,t}$ and $s_{ij,t}$ denotes the element on row i and column j of $\mathbf{R}_t$ and $\mathbf{S}_t$ respectively.

The conditional mean matrix $\mathbf{S}_t$, which is measurable by $\mathcal{F}_{t-1}$, captures the time series dynamics of singular realized covariance matrices $\{\mathbf{R}_t\}$. In the following it is assumed that

$$\mathbf{S}_t = \mathbf{C}\mathbf{C}' + \sum_{i=1}^{p} \mathbf{B}_i \mathbf{S}_{t-i} \mathbf{B}'_i + \sum_{j=1}^{q} \mathbf{A}_j \mathbf{R}_{t-j} \mathbf{A}'_j, \tag{5}$$

where we will denote the lag order of the model by (p, q) and $\mathbf{A}_j, \mathbf{B}_i, \mathbf{C}$ are $m \times m$ parameter matrices for $i = 1, \ldots, p$ and $j = 1, \ldots, q$ where $\mathbf{C}$ is lower-triangular with strictly positive diagonal elements. This form is similar to the BEKK specification of Engle and Kroner (1995) in the multivariate GARCH case, which is also adapted for the CAW model in Golosnoy et al. (2012). It ensures that $\mathbf{S}_t$ is symmetric and positive definite as long as the initial matrices $\mathbf{S}_0, \mathbf{S}_{-1}, \ldots, \mathbf{S}_{-p+1}$ are symmetric and positive semi-definite. It is notable that the conditional covariance matrix in GARCH-BEKK(p, q) coincides with the expression presented in (5) with $\mathbf{R}_{t-j} = \mathbf{x}_{t-j}\mathbf{x}'_{t-j}$, where $\mathbf{x}_t$ is the one day return vector of day t. As such, the proposed SCAW model (1) and (5) is a generalization of the GARCH-BEKK(p, q) process, where the GARCH-BEKK(p, q) model is a special case corresponding to $n = 1$. Furthermore, the specification of the SCAW(p, q) process is similar to the CAW(p, q) model suggested in Golosnoy et al. (2012) with the difference that the SCAW(p, q) process models singular realized covariance matrices, while Golosnoy et al. (2012) consider non-singular ones.

In the paper we will further consider different structures of the parameter matrices $\mathbf{A}_j, \mathbf{B}_i, \mathbf{C}$, $j = 1, \ldots, q$, $i = 1, \ldots, p$. Since large dimensional cases will generally be discussed, the matrices $\mathbf{A}_j, \mathbf{B}_i, \mathbf{C}$ needs to be specified parsimoniously such that estimation of the model remains feasible. If for example one allows $\mathbf{A}_j, \mathbf{B}_i$ to be general $m \times m$ matrices and $\mathbf{C}$ lower-triangular, the model (5) will consist of $m(m+1)/2 + (p+q)m^2$ parameters. With a large dimensional case, such as $m = 50$ and $p = q = 2$, this results in 11275 parameters, a formidable estimation exercise indeed.

2.1 Stochastic properties of the SCAW model

In this section we will present conditions under which the matrix-variate process $\{\mathbf{R}_t\}$ is weakly stationary. As with the CAW(p, q) model in Golosnoy et al. (2012), the stochastic properties of the SCAW(p, q) model defined in (1) and (5) are derived using the VARMA representation of the model. The proofs of the results presented in this section can be found in the Appendix.

Let vec$(\cdot)$ be the vectorization operator and let vech$(\cdot)$ be the half-vectorization operator. The symbol $\mathbf{D}_m$ denotes the duplication matrix, while $\mathbf{L}_m$ stands for the elimination matrix[1]. We define

$$\mathbf{r}_t \quad = \quad \text{vech}(\mathbf{R}_t), \quad \mathbf{s}_t = \text{vech}(\mathbf{S}_t), \quad \mathbf{c} = \text{vech}(\mathbf{C}\mathbf{C}'),$$

[1]The matrices $\mathbf{D}_m$ and $\mathbf{L}_m$ are defined as the matrices which satisfy the following equalities vec$(\mathbf{A}) = \mathbf{D}_m$vech$(\mathbf{A})$ and vech$(\mathbf{A}) = \mathbf{L}_mvec(\mathbf{A})$ for a symmetric matrix $\mathbf{A}$, respectively (see, e.g., Harville (1997)).

such that the vector representation of (5) is

$$\mathbf{s}_t \quad = \quad \mathbf{c} + \sum_{i=1}^{p} \boldsymbol{\mathcal{B}}_i \mathbf{s}_{t-i} + \sum_{j=1}^{q} \boldsymbol{\mathcal{A}}_j \mathbf{r}_{t-j}, \tag{6}$$

where $\boldsymbol{\mathcal{A}}_j$ and $\boldsymbol{\mathcal{B}}_i$ are $k \times k$ matrices, with $k = m(m+1)/2$ given by

$$\boldsymbol{\mathcal{A}}_j \quad = \quad \mathbf{L}_m(\mathbf{A}_j \otimes \mathbf{A}_j)\mathbf{D}_m, \quad \boldsymbol{\mathcal{B}}_i = \mathbf{L}_m(\mathbf{B}_i \otimes \mathbf{B}_i)\mathbf{D}_m,$$

where the symbol $\otimes$ denotes the Kroenecker product. Furthermore, $\mathbf{r}_t$ can be written as

$$\mathbf{r}_t = \mathbb{E}[\mathbf{r}_t \mid \mathbf{F}_{t-1}] + \mathbf{v}_t = \mathbf{s}_t + \mathbf{v}_t, \tag{7}$$

where $\mathbf{v}_t$ is a martingale difference sequence such that

$$\mathbb{E}[\mathbf{v}_t] = 0 \quad \text{and} \quad \mathbb{E}[\mathbf{v}_t \mathbf{v}_s'] = 0, \quad \forall s \neq t.$$

Plugging in (7) into (6), the SCAW(p, q) model can be written with the following VARMA$(\max(p, q), p)$ representation:

$$\mathbf{r}_t = \mathbf{c} + \sum_{i=1}^{\max(p,q)} (\boldsymbol{\mathcal{A}}_i + \boldsymbol{\mathcal{B}}_i)\mathbf{r}_{t-i} + \mathbf{v}_t - \sum_{j=1}^{p} \boldsymbol{\mathcal{B}}_j \mathbf{v}_{t-j}, \tag{8}$$

where $\boldsymbol{\mathcal{A}}_i = \boldsymbol{\mathcal{B}}_j = \mathbf{0}$ for $i > q, j > p$. From (8) the conditions for weak stationarity of $\mathbf{R}_t$ can be obtained. First, we derive a condition for the existence of the unconditional expectation of the SCAW(p, q) process, given by the following proposition.

Proposition 1. *The unconditional expectation of the SCAW(p,q) model is finite if and only if all eigenvalues of the matrix*

$$\boldsymbol{\Psi}_1 = \sum_{i=1}^{\max(p,q)} (\boldsymbol{\mathcal{A}}_i + \boldsymbol{\mathcal{B}}_i) \tag{9}$$

are less than 1 in modulus. In that case the unconditional expectation is given by

$$\mathbb{E}[\mathbf{r}_t] = \bar{\mathbf{r}} = \left(\mathbf{I}_k - \sum_{i=1}^{\max(p,q)} (\boldsymbol{\mathcal{A}}_i + \boldsymbol{\mathcal{B}}_i)\right)^{-1} \mathbf{c}. \tag{10}$$

Equation (8) can also be represented as an infinite vector moving average time series by (see, e.g., sections 11.3 and 11.4 in Lütkepohl (2005))

$$\mathbf{r}_t = \bar{\mathbf{r}} + \sum_{i=0}^{\infty} \mathbf{\Phi}_i \mathbf{v}_{t-i}, \quad \text{where} \tag{11}$$

$$\mathbf{\Phi}_i = -\boldsymbol{\mathcal{B}}_i + \sum_{j=1}^{i} (\boldsymbol{\mathcal{A}}_j + \boldsymbol{\mathcal{B}}_j)\mathbf{\Phi}_{i-j}, \quad i,j = 1,2,\ldots, \tag{12}$$

$$\mathbf{\Phi}_0 = \mathbf{I}_m. \tag{13}$$

Moreover, given that they exist, the autocovariances of $\mathbf{r}_t$ can then be expressed as

$$\mathbb{E}[(\mathbf{r}_t - \bar{\mathbf{r}})(\mathbf{r}_{t-\tau} - \bar{\mathbf{r}})'] = \sum_{i=0}^{\infty} \mathbf{\Phi}_{i+\tau}\, \mathbb{E}[\mathbf{v}_t \mathbf{v}_t'] \mathbf{\Phi}_i'. \tag{14}$$

Let

$$\mathbf{\Omega} = \frac{1}{n}(\mathbf{L}_m \otimes \mathbf{L}_m)[\mathbf{I}_{m^2} \otimes (\mathbf{I}_{m^2} + \mathbf{K}_{mm})](\mathbf{I}_m \otimes \mathbf{K}_{mm} \otimes \mathbf{I}_m)(\mathbf{D}_m \otimes \mathbf{D}_m), \tag{15}$$

where $\mathbf{K}_{mm}$ is the commutation matrix.[2] Then the following holds.

Proposition 2. *The unconditional second moment of the SCAW(p,q) model is finite if and only if all eigenvalues of the matrix*

$$\Psi_2 = \sum_{i=1}^{\infty} (\mathbf{\Phi}_i \otimes \mathbf{\Phi}_i)\mathbf{\Omega} \tag{16}$$

are less than 1 in modulus. In that case the second moment is given by

$$\text{vec}(\mathbb{E}[\mathbf{r}_t \mathbf{r}_t']) = (\mathbf{\Omega} + \mathbf{I}_{k^2}) \left(\mathbf{I}_{k^2} - \sum_{i=1}^{\infty} (\mathbf{\Phi}_{\boldsymbol{i}} \otimes \mathbf{\Phi}_i)\mathbf{\Omega}\right)^{-1} \text{vec}(\bar{\mathbf{r}}\bar{\mathbf{r}}'), \tag{17}$$

with $\bar{\mathbf{r}}$ given by (10).

[2]It is defined by the following equality $\mathbf{K}_{mm}\text{vec}(\mathbf{A}) = \text{vec}(\mathbf{A}')$ for any symmetric matrix $\mathbf{A}$ (see, Harville (1997).

Proposition 3. *Given that the unconditional second moments of the SCAW(p,q) model exist, the autocovariance matrix at lag τ is given by*

$$\mathrm{vec}(\mathbb{E}[(\mathbf{r}_t - \bar{\mathbf{r}})(\mathbf{r}_{t-\tau} - \bar{\mathbf{r}})']) = \sum_{i=0}^{\infty}(\mathbf{\Phi}_{i+\tau} \otimes \mathbf{\Phi}_i)\mathbf{\Omega} \left(\mathbf{I}_{k^2} - \sum_{i=1}^{\infty}(\mathbf{\Phi}_{\boldsymbol{i}} \otimes \mathbf{\Phi}_i)\mathbf{\Omega} \right)^{-1} \mathrm{vec}(\bar{\mathbf{r}}\bar{\mathbf{r}}').$$

As such, the process $\{\mathbf{R}_t\}$ under the SCAW(p, q) model defined by (1)-(5) is weakly stationary if and only if the eigenvalues of the matrix (16) are less than 1 in modulus.

3 Parameterization

As mentioned in Section 2, when the dimension of $\{\mathbf{R}_t\}$ grows large, it is important to parameterize $\mathbf{S}_t$ in (5) parsimoniously, in order to maintain feasible estimation. Simultaneously, the specification must be rich enough to capture the time series dynamics observed in data. This section discusses several parameterizations that can be applied to this end.

3.1 Covariance targeting

The constant term $\mathbf{CC}'$ in (5) consists of $m(m+1)/2$ parameters, rapidly increasing the estimation burden as the portfolio size grows. One approach to reduce the number of parameters is to consider *covariance targeting*, an extension of the idea of variance targeting (see Engle and Mezrich (1996)), where the constant term $\mathbf{CC}'$ is consistently estimated (see Pedersen and Rahbek (2014)) as follows. Let $\mathbf{R}_t = \mathbf{S}_t + \mathbf{V}_t$, where $\mathbf{V}_t$ is a martingale difference, s.t. $\mathbb{E}[\mathbf{V}_t] = 0$. Further denote the unconditional mean of $\{\mathbf{R}_t\}$ as $\mathbb{E}[\mathbf{R}_t] = \bar{\mathbf{S}}$. Then we can write (5) as

$$\begin{aligned}
\mathbf{S}_t &= \mathbf{CC}' + \sum_{i=1}^{p}\mathbf{B}_i\mathbf{S}_{t-i}\mathbf{B}_i' + \sum_{j=1}^{q}\mathbf{A}_j\mathbf{R}_{t-j}\mathbf{A}_j' \\
\mathbf{R}_t - \mathbf{V}_t &= \mathbf{CC}' + \sum_{i=1}^{p}\mathbf{B}_i(\mathbf{R}_{t-i} - \mathbf{V}_{t-i})\mathbf{B}_i' + \sum_{j=1}^{q}\mathbf{A}_j\mathbf{R}_{t-j}\mathbf{A}_j'.
\end{aligned}$$

Taking unconditional expectations we obtain

$$
\begin{aligned}
\mathbb{E}[\mathbf{R}_t] &= \mathbf{C}\mathbf{C}' + \sum_{i=1}^{p} \mathbf{B}_i \, \mathbb{E}[\mathbf{R}_t]\mathbf{B}_i' + \sum_{j=1}^{q} \mathbf{A}_j \, \mathbb{E}[\mathbf{R}_t]\mathbf{A}_j' \\
\bar{\mathbf{S}} &= \mathbf{C}\mathbf{C}' + \sum_{i=1}^{p} \mathbf{B}_i \bar{\mathbf{S}} \mathbf{B}_i' + \sum_{j=1}^{q} \mathbf{A}_j \bar{\mathbf{S}} \mathbf{A}_j'.
\end{aligned}
$$

such that

$$
\mathbf{C}\mathbf{C}' = \bar{\mathbf{S}} - \sum_{i=1}^{p} \mathbf{B}_i \bar{\mathbf{S}} \mathbf{B}_i' - \sum_{j=1}^{q} \mathbf{A}_j \bar{\mathbf{S}} \mathbf{A}_j'. \tag{18}
$$

The idea is then to replace $\mathbf{C}\mathbf{C}'$ in (5) by the expression (18), and to estimate $\bar{\mathbf{S}}$ by the sample mean of the process. This specification determines the constant term $\mathbf{C}\mathbf{C}'$ by the persistence parameters $\mathbf{A}_j$ and $\mathbf{B}_i$, such that $k = m(m+1)/2$ parameters less needs to be estimated in the model.

In order to ensure that the expression (18) is positive-definite, particular restrictions on the parameter matrices $\mathbf{A}_j$ and $\mathbf{B}_i$ must be imposed, and in general it is difficult to specify such conditions. One approach to circumvent this issue is considered in Noureldin et al. (2014) in the case of an ARCH model, where the original series of return vectors is rotated by its unconditional mean, to create a standardized series of returns particularly suitable to model with covariance targeting. In this paper we adapt this approach to singular realized covariance matrices in order to obtain a parsimonious parameterization. To this end, apply the eigenvalue decomposition to the unconditional expectation such that

$$
\bar{\mathbf{S}} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}',
$$

where $\mathbf{P}$ is a matrix with eigenvectors of $\bar{\mathbf{S}}$ as columns and $\mathbf{\Lambda}$ is a diagonal matrix with the eigenvalues of $\bar{\mathbf{S}}$ as diagonal entries. Note that although $\{\mathbf{R}_t\}$ is a series of singular matrices, its unconditional mean $\bar{\mathbf{S}}$ is non-singular and as such all the eigenvalues in $\mathbf{\Lambda}$ are positive. Further note that the symmetric square root of $\bar{\mathbf{S}}$ is $\bar{\mathbf{S}}^{1/2} = \mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{P}'$ and that $\bar{\mathbf{S}}^{-1/2} = \mathbf{P}\mathbf{\Lambda}^{-1/2}\mathbf{P}'$, since $\mathbf{P}$ is an orthogonal matrix.

Next we define the standardized realized covariance as

$$\mathbf{E}_t = \bar{\mathbf{S}}^{-1/2}\mathbf{R}_t(\bar{\mathbf{S}}^{-1/2})' = \mathbf{P}\boldsymbol{\Lambda}^{-1/2}\mathbf{P}'\mathbf{R}_t\mathbf{P}\boldsymbol{\Lambda}^{-1/2}\mathbf{P}'. \tag{19}$$

which has expected value

$$\mathbb{E}[\mathbf{E}_t] = \bar{\mathbf{S}}^{-1/2}\,\mathbb{E}[\mathbf{R}_t](\bar{\mathbf{S}}^{-1/2})' = \bar{\mathbf{S}}^{-1/2}\bar{\mathbf{S}}(\bar{\mathbf{S}}^{-1/2})' = \mathbf{I}_m.$$

Similarly, define $\mathbf{G}_t = \bar{\mathbf{S}}^{-1/2}\mathbf{S}_t(\bar{\mathbf{S}}^{-1/2})'$, such that

$$\mathbf{E}_t \sim SW_m(n, \mathbf{G}_t/n) \tag{20}$$

due to the affine transformation property of the singular Wishart distribution (see Theorem 2 of Bodnar et al. (2014)). As such, we will model $\mathbf{G}_t$ with an specification equivalent to (5) given by

$$\mathbf{G}_t = \tilde{\mathbf{C}}\tilde{\mathbf{C}}' + \sum_{i=1}^{p}\tilde{\mathbf{B}}_i\mathbf{G}_{t-i}\tilde{\mathbf{B}}_i' + \sum_{j=1}^{q}\tilde{\mathbf{A}}_j\mathbf{E}_{t-j}\tilde{\mathbf{A}}_j'.$$

Note that since $\mathbf{E}_t$ follows a conditional singular Wishart distribution and the specification of $\mathbf{G}_t$ is equivalent to that of $\mathbf{S}_t$, all results in Section (2.1) applies to the process $\{\mathbf{E}_t\}$ as well, with regards to parameters $\tilde{\mathbf{A}}_j$ and $\tilde{\mathbf{B}}_i$. Moreover, by applying the covariance targeting technique described in (18) with $\mathbb{E}[\mathbf{E}_t] = \mathbf{I}_m$ we get

$$\mathbf{G}_t = \left(\mathbf{I}_m - \sum_{i=1}^{p}\tilde{\mathbf{B}}_i\tilde{\mathbf{B}}_i' - \sum_{j=1}^{q}\tilde{\mathbf{A}}_j\tilde{\mathbf{A}}_j'\right) + \sum_{i=1}^{p}\tilde{\mathbf{B}}_i\mathbf{G}_{t-i}\tilde{\mathbf{B}}_i' + \sum_{j=1}^{q}\tilde{\mathbf{A}}_j\mathbf{E}_{t-j}\tilde{\mathbf{A}}_j'. \tag{21}$$

The restrictions on the persistence parameters $\tilde{\mathbf{A}}_j$ and $\tilde{\mathbf{B}}_i$ needed to ensure the positive definiteness of $\mathbf{G}_t$ are easily obtained for several parameterizations, as discussed below. In the following, the covariance targeting SCAW model described by (19), (20) and (21) will be referred to as $\mathrm{SCAW}_{\mathrm{CT}}(p,q)$.

Furthermore, since $\mathbf{S}_t = \bar{\mathbf{S}}^{1/2}\mathbf{G}_t(\bar{\mathbf{S}}^{1/2})'$ and $\mathbf{E}_t = \bar{\mathbf{S}}^{-1/2}\mathbf{R}_t(\bar{\mathbf{S}}^{-1/2})'$ the model

(21) for the standardized series $\{\mathbf{E}_t\}$ implies the following equalities

$$\begin{aligned}
\mathbf{A}_j &= \bar{\mathbf{S}}^{1/2}\tilde{\mathbf{A}}_j(\bar{\mathbf{S}}^{1/2})' && (22)\\
\mathbf{B}_i &= \bar{\mathbf{S}}^{1/2}\tilde{\mathbf{B}}_i(\bar{\mathbf{S}}^{1/2})' && (23)\\
\mathbf{C}\mathbf{C}' &= \bar{\mathbf{S}}^{1/2}\left(\mathbf{I}_m - \sum_{i=1}^{p}\tilde{\mathbf{B}}_i\tilde{\mathbf{B}}_i' - \sum_{j=1}^{q}\tilde{\mathbf{A}}_j\tilde{\mathbf{A}}_j'\right)(\bar{\mathbf{S}}^{1/2})' && (24)
\end{aligned}$$

for the parameterization in (5), modeling the non-standardized series $\{\mathbf{R}_t\}$.

Moreover, as discussed in Noureldin et al. (2014), there are seveal ways to parameterize the conditional mean, in this model described by (21): scalar and diagonal specification of $\tilde{\mathbf{A}}_j$ and $\tilde{\mathbf{B}}_i$, as well as specifications with common persistence or with orthogonal parameter matrices. In this presentation we will focus on the diagonal specification, such that $\tilde{\mathbf{A}}_j$ and $\tilde{\mathbf{B}}_i$ are all diagonal matrices, with the additional condition that the first element of each parameter matrix is positive, in order to ensure model identification. As such, the constant term in (21) will be positive-definite if and only if (see Engle and Kroner (1995))

$$\sum_{j=1}^{q}\tilde{a}_{j,ll}^2 + \sum_{i=1}^{p}\tilde{b}_{i,ll}^2 < 1, \quad l = 1, \ldots, m, \tag{25}$$

where $\tilde{a}_{j,ll}$ is the l:th diagonal element of $\tilde{\mathbf{A}}_j$ and $\tilde{b}_{i,ll}$ is the l:th diagonal element of $\tilde{\mathbf{B}}_i$. Conditions for the other specifications mentioned above can be obtained correspondingly. The diagonal parameterization of (21) results in $m(p+q)$ parameters, which is substantially lower than the $m(m+1)/2 + (p+q)m^2$ parameters in the original specification (5), particularly for large dimensional cases. In the example with $m = 50$ and $p = q = 2$ above, the diagonal model suggested thus results in 200 parameters, instead of the 11275 parameters in the original specification, making estimation much more feasible.

Hence, instead of modeling the series $\{\mathbf{R}_t\}$ with conditional means $\{\mathbf{S}_t\}$ directly, the above approach instead models the standardized realized covariances $\{\mathbf{E}_t\}$ with the conditional means $\{\mathbf{G}_t\}$. In turn, this model implies $\{\mathbf{S}_t\}$ to be specified as (5) with parameters obtained as (22)-(24). Note that while $\tilde{\mathbf{A}}_j$ and $\tilde{\mathbf{B}}_i$ are diagonal, the implied parameters for $\{\mathbf{S}_t\}$, $\mathbf{A}_j$ and $\mathbf{B}_i$, are in general not, since the transformations

(22)-(24) do not necessarily result in diagonal matrices. This does indeed suggest a rich dynamic for the original series of realized covariance matrices, as discussed in Noureldin et al. (2014) in the equivalent ARCH case. However, it does not mean that the specification $\{\mathbf{S}_t\}$ results in an entirely general BEKK model, since its parameters are constrained by the unconditional mean $\bar{\mathbf{S}}$.

3.2 Sectorwise parameterization

Prices for assets that belong to the same market sector tend to exhibit some level of similarity in price movements (see e.g. King (1966) and Chan et al. (2007)). To incorporate this feature, we introduce a model specification that assumes that covariance dynamics are homogeneous within market sectors. For this sectorwise parameterization, we define the diagonal elements of the parameter matrices $\tilde{\mathbf{A}}_j$ and $\tilde{\mathbf{B}}_i$, $j = 1, \dots, q$, $i = 1, \dots, p$, in (21) as $a_{ll,j} = a_{kk,j}$ and $b_{ll,i} = b_{kk,i}$ if asset l and asset k belong to the same market sector. The number of parameters for this specification is as such $s(p+q)$, where s denote the number of sectors that the considered assets belong to. Hence, the number of parameters for this approach is independent of process dimension m, which makes it an attractive modeling candidate when very large asset portfolios are considered.

3.3 HAR extension

To account for the high persistence in volatility processes, we also adapt the SCAW model with a heterogeneous autoregressive (HAR) extension, as proposed by Corsi (2009) in the univariate case and implemented by Golosnoy et al. (2012) in a matrix-variate version. Such an approach considers the long-memory dependence in daily volatility by including lagged realized covariances observed on longer horizons, like weekly and monthly. Consequently, for this specification, we define the conditional process mean $\mathbf{G}_t$ as

$$\mathbf{G}_t = \left(\mathbf{I}_m - \sum_{j=1}^{q} \tilde{\mathbf{A}}_j \tilde{\mathbf{A}}_j' - \tilde{\mathbf{D}}\tilde{\mathbf{D}}'\right) + \sum_{j=1}^{q} \tilde{\mathbf{A}}_j \mathbf{E}_{t-j} \tilde{\mathbf{A}}_j' + \tilde{\mathbf{D}}\mathbf{E}_{t-1}^{(h)}\tilde{\mathbf{D}}', \tag{26}$$

where $\mathbf{E}_t^{(h)}$ denote the standardized realized covariance, averaged of the last h trading days, up to trading day t. Further, we define $\tilde{\mathbf{D}}$ as a diagonal matrix with sectorwise parameterization as described above. As such, (26) can be specified in terms of (21), but we denote it with a separate parameter matrix $\tilde{\mathbf{D}}$ for ease of interpretation.

4 Estimation

Similar to Noureldin et al. (2014), we apply a two-step estimation procedure in order to obtain the parameter estimates of the considered model (21). Given a sample of the realized covariance process, $\{\mathbf{R}_t\}_{1\leq t\leq T}$, a method of moments approach is first used to estimate the unconditional mean of the process $\bar{\mathbf{S}}$, as

$$\hat{\bar{\mathbf{S}}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{R}_t. \tag{27}$$

The estimate $\hat{\bar{\mathbf{S}}}$ is then decomposed into estimates $\hat{\mathbf{P}}$ and $\hat{\mathbf{\Lambda}}$. From these estimates, a standardized series is obtained in correspondence to (19) as

$$\mathbf{E}_t = \hat{\mathbf{P}}\hat{\mathbf{\Lambda}}^{-1/2}\hat{\mathbf{P}}'\mathbf{R}_t\hat{\mathbf{P}}\hat{\mathbf{\Lambda}}^{-1/2}\hat{\mathbf{P}}',$$

consistent with the approach in Noureldin et al. (2014). In the second step, we estimate the diagonal parameter matrices $\tilde{\mathbf{A}}_j$ and $\tilde{\mathbf{B}}_i$, $i = 1,\ldots,p$, $j = 1\ldots,q$, in (21), by the maximum likelihood method. Similarly to Golosnoy et al. (2012), in order to ensure the positivity of the first diagonal element in each of the parameter matrices, the square roots of these values are estimated. To enforce the condition (25), the diagonal elements of $\tilde{\mathbf{A}}_j$ are specified according to the following function, for $l = 1, \dots, m$,

$$\tilde{a}_{ll,j} = \begin{cases} a^*_{ll,j} & \text{if} \quad s_l < 1 \\ \frac{a^*_{ll,j}(1-\epsilon)}{s_l} & \text{if} \quad s_l \geq 1 \end{cases} \tag{28}$$

$$\tilde{b}_{ll,i} = \begin{cases} b^*_{ll,i} & \text{if} \quad s_l < 1 \\ \frac{b^*_{ll,i}(1-\epsilon)}{s_l} & \text{if} \quad s_l \geq 1 \end{cases}, \tag{29}$$

where $s_l = \sum_{j=1}^{q} \tilde{a}^2_{j,ll} + \sum_{i=1}^{p} \tilde{b}^2_{i,ll}$ and ϵ is positive and close to zero. As such, we define the argument vector to the log-likelihood function as $\psi' = (\psi'_a, \psi'_b)$ with $\psi'_a = \left(\sqrt{a^*_{11,1}}, a^*_{22,1}, \dots, a^*_{ll,q}\right)$ and $\psi'_b = \left(\sqrt{b^*_{11,1}}, b^*_{22,1}, \dots, b^*_{ll,q}\right)$. Furthermore, since by (20), $\mathbf{E}_t$ follows a singular Wishart distribution, the log-likelihood obtains directly from the density (3) as

$$\begin{aligned} \mathcal{L}(\psi) &= \sum_{t=1}^{T} \left[c + \frac{n}{2} \ln |\mathbf{G}_t| + \frac{n-m-1}{2} \ln |\mathbf{E}_{11,t}| - \right. \\ &\left. - \frac{n-m-1}{2} \mathrm{tr}(\mathbf{G}_t^{-1} \mathbf{E}_t) \right], \end{aligned} \tag{30}$$

where

$$c = \frac{n(n-m)}{2} \ln(\pi) + \left(m - \frac{n}{2}\right) \ln n - \ln \Gamma_p \left(\frac{n}{2}\right). \tag{31}$$

Finding the vector ψ that maximizes the log-likelihood function (30) can then be done by applying numerical optimization techniques.

5 Empirical application

5.1 Data

The SCAW model presented in Section 2 is applied to analyze the daily realized covariance matrices of 50 assets traded at National Association of Securities Dealers Automated Quotations (NASDAQ) from mid 1997 to mid 2017. These assets are listed in Table 1 together with their associated market sector, following the NASDAQ sector classification (as e.g. Litimi et al. (2016), BenSaïda (2017)). Assets are selected such that the sample sector distribution is proportional to the sector distribution of assets traded at NASDAQ for the considered time period. The realized covariance matrix of these assets, for trading day t, is constructed as $\mathbf{R}_t = \sum_i^n \mathbf{x}_{t,i}\mathbf{x}_{t,i}'$, where $\mathbf{x}_{t,i}$ is the $m \times 1$ return vector obtained for the i:th 10-minute interval of day t between 09:30 and 16:00. In turn, this results in $n = 39$ return vectors, such that the rank of the $m \times m$ matrix $\mathbf{R}_t$ is 39, making it a singular matrix. The sample period starts 2nd of June 1997 and ends 15th of June 2017, resulting in about 20 years of data, and 4993 trading days. As such, the considered series covers two exceptionally volatile time periods: the so-called Dot-com bubble, which had its peak around the year 2000, and the global financial crisis of 2007-2008. Out of this time series sample, the first 90% of the trading days is used for estimating the models, while the remaining 10% of the sample is used to compute forecast accuracy.

Summary statistics of the realized variances (multiplied by 10^4 for easier reading) of the 50 considered assets are shown in Table 1. As is typical regarding empirical variances of asset returns, most of the series are considerably right skewed and leptokurtic. Table 2 summarizes statistics for the realized variance in each of the 12 market sectors in the sample. According to these statistics, the energy sector experiences the largest average variance, while assets in the financial sector are the most right skewed and leptokurtic.

5.2 Models

To study these data using the suggested SCAW model, the estimation and forecasting are performed using the various model specifications discussed in Section 3. The following SCAW-parameterizations are considered:

- $\text{SCAW}_{\text{CT}}(p, q)$: Parameter matrices $\tilde{\mathbf{A}}_j$ and $\tilde{\mathbf{B}}_i$, $j = 1, \ldots, q$, $i = 1, \ldots, p$ of (21) are diagonal.
- $\text{SCAW-SS}_{\text{CT}}(p, q)$: Parameter matrices $\tilde{\mathbf{A}}_j$ and $\tilde{\mathbf{B}}_i$, $j = 1, \ldots, q$, $i = 1, \ldots, p$ of (21) are diagonal. Further, $a_{ll,j} = a_{kk,j}$ and $b_{ll,i} = b_{kk,i}$ if asset l and asset k belongs to the same sector.
- $\text{SCAW-SS-HAR}_{\text{CT}}(q, h)$: Parameter matrices $\tilde{\mathbf{A}}_j$ and $\tilde{\mathbf{D}}$ of (26) are diagonal.

Further, $a_{ll,j} = a_{kk,j}$ and $d_{ll} = d_{kk}$ if asset l and asset k belongs to the same sector.

Similarly to Golosnoy et al. (2012), Multivariate GARCH models fitted to daily return data are used as forecast accuracy benchmarks to the parameterizations described above. To have comparable results, the benchmark models follow equivalent specifications and are consequently denoted $\text{MGARCH}_{\text{CT}}(p, q)$, $\text{MGARCH-SS}_{\text{CT}}(p, q)$ and $\text{MGARCH-SS-HAR}_{\text{CT}}(q, h)$. As discussed in Section 2, the Multivariate GARCH model with BEKK-specification can be thought of as a special case of SCAW model, with the number of intra-day returns (and matrix rank) $n = 1$.

5.3 Estimation

The first 90% of data consisting of 4494 trading days is used to estimate the models discussed in Section 5.2. The parameters of considered models are estimated as described in Section 4, where $\epsilon = 10^{-7}$ is used in Equations (28) and (29). Table 3 displays the number of parameters for each specification, and the maximum log-likelihood value (MLL) obtained from the optimization procedure. We also include two information criteria: Akaikes (AIC) and Bayesian (BIC). Within each set of model types, SCAW and MGARCH, the most favorable value in each column is highlighted in bold. With regards to the MLL-value and information criteria, $\text{SCAW-SS}_{\text{CT}}(2, 2)$ performs the best among the SCAW specifications. Note that the maximum log-likelihood value as well as the values of AIC and BIC computed for the SCAW and MGARCH models are not comparable, since these two models have probability distributions with the mass being concentrated on the spaces of different dimensions. In accordance with the partition (2), the dimension of the sample space subsets with probability mass is $mn - n(n-1)/2$. As such, with $m = 50$, the actual dimension of the sample space for the SCAW model is 1209, while it is 50 for the multivariate GARCH model. Hence the values of the density functions tend to be considerably smaller for the SCAW models, in general. Thus the values of all three performance measure can be compared only within the class of the two considered matrix-variate time series models.

Table 4 displays the estimates for the parameters estimates of the SCAW model with the most favourable MLL, AIC and BIC values, that is in the case of SCAW-$\text{SS}_{\text{CT}}(2,2)$ process. We observe that all parameters are statistically significant at 1% level. Moreover, the estimated elements of the matrices $\tilde{\mathbf{B}}_1$ and $\tilde{\mathbf{B}}_2$ are uniformly larger than the entries in the matrices $\tilde{\mathbf{A}}_1$ and $\tilde{\mathbf{A}}_2$, indicating larger effect of the previous values of the conditional mean matrices in comparison to the previous values of realized covariance matrices. This result is in line to the findings usually observed when a multivariate GARCH model is fitted to data.

5.4 Forecasting

The last 10% of the time series, 499 trading days, is used to compute out-of-sample forecast accuracy for the models discussed in Section 5.2. For each of the models, the ℓ-step-ahead forecast is computed recursively as

$$
\begin{aligned}
\mathbb{E}[\mathbf{R}_{t+\ell}|\mathcal{F}_t] &= \mathbb{E}[\mathbf{S}_{t+\ell}|\mathcal{F}_t] = \mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{P}'\,\mathbb{E}[\mathbf{G}_{t+\ell}|\mathcal{F}_t]\mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{P}', \quad \text{with} && (32)\\
\mathbb{E}[\mathbf{G}_{t+\ell}|\mathcal{F}_t] &= \left(\mathbf{I}_m - \sum_{i=1}^{p}\tilde{\mathbf{B}}_i\tilde{\mathbf{B}}_i' - \sum_{j=1}^{q}\tilde{\mathbf{A}}_j\tilde{\mathbf{A}}_j'\right) + \\
&\quad + \sum_{i=1}^{p}\tilde{\mathbf{B}}_i\,\mathbb{E}[\mathbf{G}_{t+\ell-i}|\mathcal{F}_t]\tilde{\mathbf{B}}_i' + \sum_{j=1}^{q}\tilde{\mathbf{A}}_j\,\mathbb{E}[\mathbf{E}_{t+\ell-j}|\mathcal{F}_t]\tilde{\mathbf{A}}_j', && (33)\\
\mathbb{E}[\mathbf{E}_{t+\ell-j}|\mathcal{F}_t] &= \mathbb{E}[\mathbf{G}_{t+\ell-j}|\mathcal{F}_t], && (34)
\end{aligned}
$$

where the parameter matrices are estimated as described in Section 4 and Section 5.3. The specification SCAW-SS-$\text{HAR}_{\text{CT}}(q,h)$ is computed similarly, employing that it can be represented in the form of (21).

The forecast accuracy of $\hat{\mathbf{R}}_{t+\ell} = \mathbb{E}[\mathbf{R}_{t+\ell}|\mathcal{F}_t]$ is evaluated using three measures. First, the average Frobenius norm of the ℓ-step-ahead forecast error is computed as

$$\text{FN}_\ell = \frac{1}{T_\ell}\sum_t ||\hat{\mathbf{R}}_{t+\ell} - \mathbf{R}_{t+\ell}||, \tag{35}$$

where T_ℓ is the sample-size for ℓ-step-ahead forecasts and $||\mathbf{M}||$ denote the Frobenius norm of the matrix $\mathbf{M}$. Further, in practice, one is often interested in applying covariance matrix forecasts in a portfolio setting. As such, we also compute mean squared error of the standard deviation of an equally weighted (EW) portfolio, a popular portfolio in financial literature (see, DeMiguel et al. (2009)), using the obtained covariance forecast and the realized covariance:

$$\mathrm{SD}_{\mathrm{EW},\ell} = \frac{1}{T_\ell}\sum_t \frac{1}{m^2}\left(\sqrt{\mathbf{1}'\hat{\mathbf{R}}_{t+\ell}\mathbf{1}} - \sqrt{\mathbf{1}'\mathbf{R}_{t+\ell}\mathbf{1}}\right)^2.$$

Another important portfolio is the global minimum variance (GMV) portfolio (cf., Frahm and Memmel (2010), Glombeck (2014), Bodnar et al. (2018), Bodnar et al. (2019), Ding et al. (2020)). This portfolio has the lowest risk of all possible portfolios of risky assets, and its weight vector is solely determined by the covariance matrix of asset returns. We employ the l-step ahead forecast of the covariance matrix $\hat{\mathbf{R}}_{t+\ell}$ produced by each model for the computation of the weights of the GMV portfolio expressed as

$$\hat{\mathbf{w}}_{t+\ell} = \hat{\mathbf{R}}_{t+\ell}^{-1}\mathbf{1}/(\mathbf{1}'\hat{\mathbf{R}}_{t+\ell}^{-1}\mathbf{1}). \quad (36)$$

As a performance measure of the constructed portfolio for different models, we use the standard deviation of the GMV portfolio of future time periods given by

$$\mathrm{SD}_{\mathrm{GMV},\ell} = \frac{1}{T_\ell}\sum_t \sqrt{\hat{\mathbf{w}}_{t+\ell}'\mathbf{R}_{t+\ell}\hat{\mathbf{w}}_{t+\ell}}.$$

The quantity $\mathrm{SD}_{\mathrm{GMV},\ell}$ measures the average forecasted standard deviation of the GMV portfolios in the out-of-sample period. Hence, $\mathrm{SD}_{\mathrm{EW},\ell}$ and $\mathrm{SD}_{\mathrm{GMV},\ell}$ illustrate the ability of each model to forecast different things. The former is a measure of the squared difference between the predicted standard deviation of the equally weighted portfolio, and the observed standard deviation of the equally weighted portfolio. The latter measures the predicted standard deviation of the GMV portfolio. For this measure, more accurate predictions will result in lower values, and the minimum value is obtained when inserting $\hat{\mathbf{R}}_{t+\ell} = \mathbf{R}_{t+\ell}$ into (36). An accurate prediction of this quantity has considerable economic value, since the standard deviation of the GMV portfolio is a key input value in many financial applications.

Finally, it is relevant to see if the difference in a forecast accuracy measure between the SCAW model and its benchmark is statistically significant. To this end, a two-sided paired t-test is applied to the sample of terms in FN_ℓ, $\text{SD}_{\text{EW},\ell}$ or $\text{SD}_{\text{GMV},\ell}$ for the SCAW model, and to the sample of terms for the same measure in the equivalent MGARCH specification. Significance level 0.01 is used for all the applied tests.

5.5 Results

Tables 5-7 summarizes the forecasts performance of the models discussed in Section 5.2, with forecast horizons $\ell = 1, 5, 10$. Within each set of model types, SCAW and MGARCH, the most favorable value in each column is highlighted in bold. A star next to a forecast value indicates that the model's forecast value is lower than the equivalent value of the other model type, with statistical significance at the 0.01 level.

Table 5 summarizes the forecast accuracy result in terms of FN_ℓ for the considered models. The model SCAW-$\text{SS}_{\text{CT}}(1, 2)$ performs best for each forecast horizon, and among all considered specifications. The models $\text{SCAW}_{\text{CT}}(0, 1)$, SCAW-$\text{SS}_{\text{CT}}(0, 1)$ and SCAW-$\text{SS}_{\text{CT}}(1, 2)$ outperform the benchmark model of equivalent specification with statisticance significance on the 0.01-level.

Table 6 summarizes the forecast accuracy result in terms of $\text{SD}_{\text{EW},\ell}$ for the considered models. With regards to this measures of portfolio forecast accuracy, SCAW-SS-$\text{HAR}_{\text{CT}}(1, 20)$ produce the lowest forecast error for each horizon. The SCAW models outperform the MGARCH counterpart for each specification and forecast horizon, and all of the differences are statistically significant at the 0.01 level.

Table 7 summarizes the forecast accuracy result in terms of $\text{SD}_{\text{GMV},\ell}$ for the considered models. With regards to this measures of portfolio forecast accuracy, SCAW-$\text{SS}_{\text{CT}}(1, 1)$ produce the lowest forecast accuracy for each horizon. The SCAW models outperform the MGARCH counterpart for each specification and forecast horizon. Furthermore, all of the differences are statistically significant at the 0.01 level, except in two cases for the longest forecast horizon of 10 days.

In general, it is noteworthy that for each forecast evaluation measure, a single model dominates for all considered forecast horizons, suggesting that these models are consistently superior regarding that measure. Three different models, SCAW-$\text{SS}_{\text{CT}}(1,2)$, SCAW-SS-$\text{HAR}_{\text{CT}}(1,20)$ and SCAW-$\text{SS}_{\text{CT}}(1,1)$ have the best performance for the three different measures FN_ℓ, $\text{SD}_{\text{EW},\ell}$ and $\text{SD}_{\text{GMV},\ell}$, respectively. This implies that different SCAW models may be suitable, depending on the application at hand.

Finally, it is notable that the specifications SCAW-$\text{SS}_{\text{CT}}(1,1)$, SCAW-$\text{SS}_{\text{CT}}(1,2)$, SCAW-$\text{SS}_{\text{CT}}(2,2)$ and SCAW-SS-$\text{HAR}_{\text{CT}}(1,20)$ outperform both $\text{SCAW}_{\text{CT}}(0,1)$ and $\text{SCAW}_{\text{CT}}(1,1)$ in terms of either AIC, BIC or one of the forecast measures FN_ℓ, $\text{SD}_{\text{EW},\ell}$ and $\text{SD}_{\text{GMV},\ell}$, $\ell = 1, 5, 10$, despite using a lower number of parameters (24, 36, 48 and 24 versus 50 and 100, respectively). This suggest that the sector-wise approach introduced in Section 3.2 indeed can be useful.

To summarize, the SCAW approach seems to perform well comparing to the benchmark, in terms of out-of-sample forecast accuracy for the time period studied. Among the various SCAW-specifications, the sectorwise parameterization, SCAW-$\text{SS}_{\text{CT}}(p,q)$, and the sectorwise specification with an HAR-extension, SCAW-SS-$\text{HAR}_{\text{CT}}(q,h)$, displays the most favourable results. The performance of these specifications are important, since the number of parameters of the sectorwise approach is independent of the number of assets m. As such, it is likely to be a feasible parameterization when very large asset portfolios are considered.

6 Conclusion

In this paper, we present the Singular Conditional Autoregressive Wishart (SCAW) model to capture the temporal dynamics for time series of singular realized covariance matrices. The model employs a BEKK-type specification, thus ensuring positive definitiveness, and allowing for straight forward estimation through the maximum likelihood method. Since the case of singular realized covariance is closely related to large portfolio dimensions, we also introduce methods to maintain parsimony in large dimensions. First, a covariance targeting approach adapted to the matrix case is presented. Second, we propose a sectorwise specification, utilizing asset homogeneity within market sectors. As an additional extension, the wellestablished HAR-approach is adapted to our model. These approaches results in a model well adapted for large or extremely large portfolio sizes.

The SCAW model is further estimated to 50 stocks in a time series covering 20 years, and moreover evaluated out-of-sample with Multivariate GARCH models of similar specifications as benchmark. This study reveals that the SCAW models indeed outperform the benchmark model in the vast majority of the forecast accuracy measures, with high statistical significance. Furthermore, it suggests that the sectorwise specification and HAR-extensions shows great promise, greatly improving both in-sample and out-of-sample performances in relation to the baseline model, while using only a fraction of the number of parameters. This is important, since again these specifications have good scaling properties with regards to portfolio size.

Future venues of research include extension of the SCAW model by for example the MIDAS-extension employed in Golosnoy et al. (2012), or an adaptation including the leverage effect in the spirit of Anatolyev and Kobotaev (2018).

Acknowledgement

For valuable discussions, the authors would like to thank seminar and conference participants at the Workshop on Financial Econometrics 2019 (Örebro University School of Business) and at the International Conference on Computational and Financial Econometrics 2019 (University of London).

Appendix

Lemma 1. *Given that* $\mathbb{E}[\mathbf{r}_t\mathbf{r}_t']$ *exists,*

$$\text{vec}(\mathbb{E}[\mathbf{r}_t\mathbf{r}_t']) = (\mathbf{\Omega} + \mathbf{I}_{k^2})\text{vec}(\mathbb{E}[\mathbf{s}_t\mathbf{s}_t']). \tag{37}$$

Proof. We have $\mathbf{r}_t = \text{vech}(\mathbf{R}_t) = \mathbf{L}_m\text{vec}(\mathbf{R}_t)$, and as such the conditional variance of $\mathbf{r}_t$ can be expressed as

$$\begin{aligned}\mathbb{V}[\mathbf{r}_t \mid \mathcal{F}_{t-1}] &= \mathbb{V}[\mathbf{L}_m \text{vec}(\mathbf{R}_t) \mid \mathcal{F}_{t-1}] \\ &= \mathbf{L}_m \, \mathbb{V}[\text{vec}(\mathbf{R}_t) \mid \mathcal{F}_{t-1}]\mathbf{L}_m'. \end{aligned} \tag{38}$$

Equation (4), expressing the conditional covariance of the singular Wishart distribution, can further be written as (see e.g. Muirhead (1982), p. 90)

$$\mathbb{V}[\text{vec}(\mathbf{R}_t) \mid \mathcal{F}_{t-1}] = \frac{1}{n}(\mathbf{I}_{m^2} + \mathbf{K}_{m,m})(\mathbf{S}_t \otimes \mathbf{S}_t), \tag{39}$$

where $\mathbf{K}_{m,m}$ is the commutation matrix. Using relationships between the Kroenecker product and the vec($\cdot$) operator, we obtain from (38) and (39) that

$$\begin{aligned}\text{vec}(\mathbb{V}[\mathbf{r}_t \mid \mathcal{F}_{t-1}]) &= \frac{1}{n}(\mathbf{L}_m \otimes \mathbf{L}_m)\text{vec}\left[(\mathbf{I}_{m^2} + \mathbf{K}_{m,m})(\mathbf{S}_t \otimes \mathbf{S}_t)\right] \\ &= \frac{1}{n}(\mathbf{L}_m \otimes \mathbf{L}_m)\left[\mathbf{I}_{m^2} \otimes (\mathbf{I}_{m^2} + \mathbf{K}_{m,m})\right]\text{vec}(\mathbf{S}_t \otimes \mathbf{S}_t) \\ &= \frac{1}{n}(\mathbf{L}_m \otimes \mathbf{L}_m)\left[\mathbf{I}_{m^2} \otimes (\mathbf{I}_{m^2} + \mathbf{K}_{m,m})\right] \times \\ &\quad \times(\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m)\left[\text{vec}(\mathbf{S}_t) \otimes \text{vec}(\mathbf{S}_t)\right] \\ &= \frac{1}{n}(\mathbf{L}_m \otimes \mathbf{L}_m)\left[\mathbf{I}_{m^2} \otimes (\mathbf{I}_{m^2} + \mathbf{K}_{m,m})\right] \times \\ &\quad \times(\mathbf{I}_m \otimes \mathbf{K}_{m,m} \otimes \mathbf{I}_m)(\mathbf{D}_m \otimes \mathbf{D}_m)\text{vec}(\mathbf{s}_t\mathbf{s}_t') \\ &= \mathbf{\Omega}\text{vec}(\mathbf{s}_t\mathbf{s}_t'), \end{aligned} \tag{40}$$

with $\mathbf{\Omega}$ defined as in (15). By the tower property we get $\mathbb{E}[\mathbf{r}_t\mathbf{r}_t'] = \mathbb{E}[\mathbb{V}[\mathbf{r}_t \mid \mathcal{F}_{t-1}]] + \mathbb{E}[\mathbf{s}_t\mathbf{s}_t']$. Vectorizing this expression and applying (40) yield

$$\begin{aligned}\text{vec}(\mathbb{E}[\mathbf{r}_t\mathbf{r}_t']) &= \text{vec}(\mathbb{E}[\mathbb{V}[\mathbf{r}_t \mid \mathcal{F}_{t-1}]]) + \text{vec}(\mathbb{E}[\mathbf{s}_t\mathbf{s}_t']) \\ &= \mathbb{E}[\text{vec}(\mathbb{V}[\mathbf{r}_t \mid \mathcal{F}_{t-1}])] + \text{vec}(\mathbb{E}[\mathbf{s}_t\mathbf{s}_t']) \\ &= \mathbf{\Omega}\text{vec}(\mathbb{E}[\mathbf{s}_t\mathbf{s}_t']) + \text{vec}(\mathbb{E}[\mathbf{s}_t\mathbf{s}_t']) \\ &= (\mathbf{\Omega} + \mathbf{I}_{k^2})\text{vec}(\mathbb{E}[\mathbf{s}_t\mathbf{s}_t']), \end{aligned}$$

completing the proof. □

Proof of Proposition 1. Taking the expectation of the VARMA representation in (8) yields

$$\mathbb{E}[\mathbf{r}_t] = \mathbf{c} + \sum_{i=1}^{\max(p,q)} (\boldsymbol{\mathcal{A}}_i + \boldsymbol{\mathcal{B}}_i)\, \mathbb{E}[\mathbf{r}_t],$$

since $\mathbb{E}[\mathbf{v}_t] = 0$. Inserting $\mathbb{E}[\mathbf{r}_t] = \bar{\mathbf{r}}$ leads to the solution expressed as

$$\bar{\mathbf{r}} = \left(\mathbf{I}_k - \sum_{i=1}^{max(p,q)} (\boldsymbol{\mathcal{A}}_i + \boldsymbol{\mathcal{B}}_i) \right)^{-1} \mathbf{c},$$

if and only if each eigenvalue of the matrix $\boldsymbol{\Psi}_1 = \sum_{i=1}^{\max(p,q)} (\boldsymbol{\mathcal{A}}_i + \boldsymbol{\mathcal{B}}_i)$ is less than 1 in modulus. □

Proof of Proposition 2. Inserting $\mathbb{E}[\mathbf{v}_t\mathbf{v}_t'] = \mathbb{E}[\mathbf{r}_t\mathbf{r}_t'] - \mathbb{E}[\mathbf{s}_t\mathbf{s}_t']$ into the autocovariance function (14) yields

$$\mathbb{E}[\mathbf{r}_t\mathbf{r}_t'] = \sum_{i=0}^{\infty} \boldsymbol{\Phi}_i \left(\mathbb{E}[\mathbf{r}_t\mathbf{r}_t'] - \mathbb{E}[\mathbf{s}_t\mathbf{s}_t'] \right) \boldsymbol{\Phi}_i' + \bar{\mathbf{r}}_t\bar{\mathbf{r}}_t{}'.$$

Vectorizing this expression and plugging in $\text{vec}(\mathbb{E}[\mathbf{r}_t\mathbf{r}_t']) = (\boldsymbol{\Omega} + \mathbf{I}_{k^2})\text{vec}(\mathbb{E}[\mathbf{s}_t\mathbf{s}_t'])$ from Lemma 1 further gives

$$\begin{aligned}
(\boldsymbol{\Omega} + \mathbf{I}_{k^2})\text{vec}(\mathbb{E}[\mathbf{s}_t\mathbf{s}_t']) &= \sum_{i=0}^{\infty} (\boldsymbol{\Phi}_i \otimes \boldsymbol{\Phi}_i)\boldsymbol{\Omega}\text{vec}(\mathbb{E}[\mathbf{s}_t\mathbf{s}_t']) + \text{vec}(\bar{\mathbf{r}}_t\bar{\mathbf{r}}_t{}') \\
\text{vec}(\bar{\mathbf{r}}_t\bar{\mathbf{r}}_t{}') &= \left[(\boldsymbol{\Omega} + \mathbf{I}_{k^2}) - \sum_{i=0}^{\infty} (\boldsymbol{\Phi}_i \otimes \boldsymbol{\Phi}_i)\boldsymbol{\Omega} \right] \text{vec}(\mathbb{E}[\mathbf{s}_t\mathbf{s}_t']) \\
\text{vec}(\bar{\mathbf{r}}_t\bar{\mathbf{r}}_t{}') &= \left[\mathbf{I}_{k^2} - \sum_{i=1}^{\infty} (\boldsymbol{\Phi}_i \otimes \boldsymbol{\Phi}_i)\boldsymbol{\Omega} \right] \text{vec}(\mathbb{E}[\mathbf{s}_t\mathbf{s}_t']), \qquad (41)
\end{aligned}$$

where the last equality is due to the fact that $\boldsymbol{\Phi}_0 = \mathbf{I}_k$. Equation (41) can be solved for $\text{vec}(\mathbb{E}[\mathbf{s}_t\mathbf{s}_t'])$ as

$$\text{vec}(\mathbb{E}[\mathbf{s}_t\mathbf{s}_t']) = \left[\mathbf{I}_{k^2} - \sum_{i=1}^{\infty} (\boldsymbol{\Phi}_i \otimes \boldsymbol{\Phi}_i)\boldsymbol{\Omega} \right]^{-1} \text{vec}(\bar{\mathbf{r}}_t\bar{\mathbf{r}}_t{}') \qquad (42)$$

if and only if the eigenvalues of the matrix $\boldsymbol{\Psi}_2 = \sum_{i=1}^{\infty} (\boldsymbol{\Phi}_i \otimes \boldsymbol{\Phi}_i)\boldsymbol{\Omega}$ are less than 1 in modulus. Inserting (42) into the expression of Lemma 1 then gives the result. □

Proof of Proposition 3. Vectorizing Equation (14) yields

$$\text{vec}(\mathbb{E}[(\mathbf{r}_t - \bar{\mathbf{r}})(\mathbf{r}_{t-\tau} - \bar{\mathbf{r}})']) = \sum_{i=0}^{\infty}(\mathbf{\Phi}_i \otimes \mathbf{\Phi}_{i+\tau})\text{vec}(\mathbb{E}[\mathbf{v}_t\mathbf{v}_t']). \tag{43}$$

Furthermore, the application of (7) and Lemma 1 leads to

$$\begin{aligned}
\text{vec}(\mathbb{E}[\mathbf{v}_t\mathbf{v}_t']) &= \text{vec}(\mathbb{E}[\mathbf{r}_t\mathbf{r}_t']) - \text{vec}(\mathbb{E}[\mathbf{s}_t\mathbf{s}_t']) \\
&= (\mathbf{I}_{k^2} - (\mathbf{\Omega} + \mathbf{I}_{k^2})^{-1})\text{vec}(\mathbb{E}[\mathbf{r}_t\mathbf{r}_t']),
\end{aligned}$$

and inserting this into (43) and applying Proposition 2 we get

$$\begin{aligned}
\text{vec}(\mathbb{E}[(\mathbf{r}_t - \bar{\mathbf{r}})(\mathbf{r}_{t-\tau} - \bar{\mathbf{r}})']) &= \sum_{i=0}^{\infty}(\mathbf{\Phi}_i \otimes \mathbf{\Phi}_{i+\tau})(\mathbf{I}_{k^2} - (\mathbf{\Omega} + \mathbf{I}_{k^2})^{-1}) \times \\
&\quad \times \text{vec}(\mathbb{E}[\mathbf{r}_t\mathbf{r}_t']) \\
&= \sum_{i=0}^{\infty}(\mathbf{\Phi}_i \otimes \mathbf{\Phi}_{i+\tau})(\mathbf{I}_{k^2} - (\mathbf{\Omega} + \mathbf{I}_{k^2})^{-1}) \times \\
&\quad \times (\mathbf{\Omega} + \mathbf{I}_{k^2}) \left(\mathbf{I}_{k^2} - \sum_{i=1}^{\infty}(\mathbf{\Phi}_{\boldsymbol{i}} \otimes \mathbf{\Phi}_i)\mathbf{\Omega}\right)^{-1} \text{vec}(\bar{\mathbf{r}}\bar{\mathbf{r}}') \\
&= \sum_{i=0}^{\infty}(\mathbf{\Phi}_i \otimes \mathbf{\Phi}_{i+\tau})\mathbf{\Omega} \times \\
&\quad \times \left(\mathbf{I}_{k^2} - \sum_{i=1}^{\infty}(\mathbf{\Phi}_{\boldsymbol{i}} \otimes \mathbf{\Phi}_i)\mathbf{\Omega}\right)^{-1} \text{vec}(\bar{\mathbf{r}}\bar{\mathbf{r}}'),
\end{aligned}$$

which completes the proof. □

References

Alfelt, G., Bodnar, T., and Tyrcha, J. (2019). Goodness-of-fit tests for centralized Wishart processes. *Communications in Statistics - Theory and Methods*, to appear.

Anatolyev, S. and Kobotaev, N. (2018). Modeling and forecasting realized covariance matrices with accounting for leverage. *Econometric Reviews*, 37(2):114–139.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and Forecasting Realized Volatility. *Econometrica*, 71(2):579–625.

Asai, M., McAleer, M., and Yu, J. (2006). Multivariate stochastic volatility: a review. *Econometric Reviews*, 25(2-3):145–175.

Aït-Sahalia, Y. and Yu, J. (2009). High frequency market microstructure noise estimates and liquidity measures. *Annals of Applied Statistics*, 3(1):422–457.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica*, 72(3):885–925.

Bauwens, L., Laurent, S., and Rombouts, J. V. (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics*, 21(1):79–109.

BenSaïda, A. (2017). Herding effect on idiosyncratic volatility in U.S. industries. *Finance Research Letters*, 23:121 – 132.

Bodnar, T., Dmytriv, S., Parolya, N., and Schmid, W. (2019). Tests for the weights of the global minimum variance portfolio in a high-dimensional setting. *IEEE Transactions on Signal Processing*, 67(17):4479–4493.

Bodnar, T., Mazur, S., and Okhrin, Y. (2014). Distribution of the product of singular Wishart Matrix and normal vector. *Theory of Probability and Mathematical Statistics*, 91:1–15.

Bodnar, T. and Okhrin, Y. (2008). Properties of the singular, inverse and generalized inverse partitioned Wishart distributions. *Journal of Multivariate Analysis*, 99(10):2389–2405.

Bodnar, T., Okhrin, Y., and Parolya, N. (2020). Optimal shrinkage-based portfolio selection in high dimensions. *Journal of Business & Economic Statistics*, under revision.

Bodnar, T., Parolya, N., and Schmid, W. (2018). Estimation of the global minimum variance portfolio in high dimensions. *European Journal of Operational Research*, 266(1):371–390.

Cai, T. T., Hu, J., Li, Y., and Zheng, X. (2020). High-dimensional minimum variance portfolio estimation based on high-frequency data. *Journal of Econometrics*, 214(2):482–494.

Chan, L. K., Lakonishok, J., and Swaminathan, B. (2007). Industry Classifications and Return Comovement. *Financial Analysts Journal*, 63(6):56–70.

Corsi, F. (2009). A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 7(2):174–196.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the $1/n$ portfolio strategy? *The Review of Financial Studies*, 22(5):1915–1953.

Ding, Y., Li, Y., and Zheng, X. (2020). High dimensional minimum variance portfolio estimation under statistical factor models. *Journal of Econometrics, Forthcoming.*

Engle, R. and Mezrich, J. (1996). GARCH for groups. *Risk*, 9:36–40.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory*, 11(1):122–150.

Frahm, G. and Memmel, C. (2010). Dominating estimators for minimum-variance portfolios. *Journal of Econometrics*, 159(2):289–302.

Glombeck, K. (2014). Statistical inference for high-dimensional global minimum variance portfolios. *Scandinavian Journal of Statistics*, 41(4):845–865.

Golosnoy, V., Gribisch, B., and Liesenfeld, R. (2012). The conditional autoregressive Wishart model for multivariate stock market volatility. *Journal of Econometrics*, 167(1):211–223.

Gouriéroux, C., Jasiak, J., and Sufana, R. (2009). The Wishart Autoregressive Process of Multivariate Stochastic Volatility. *Journal of Econometrics*, 150(2):167–181.

Gupta, A. K. and Nagar, D. K. (2000). *Matrix variate distributions.* CRC Press.

Harville, D. (1997). *Matrix Algebra from Statistician's Perspective.* Springer, New York.

Hautsch, N., Kyj, L. M., and Malec, P. (2015). Do high-frequency data improve high-dimensional portfolio allocations? *Journal of Applied Econometrics*, 30(2):263–290.

Jacod, J. and Podolskij, M. (2013). A test for the rank of the volatility process: the random perturbation approach. *The Annals of Statistics*, 41(5):2391–2427.

Jin, X. and Maheu, J. M. (2012). Modeling realized covariances and returns. *Journal of Financial Econometrics*, 11(2):335–369.

King, B. F. (1966). Market and Industry Factors in Stock Price Behavior. *The Journal of Business*, 39(1):139–190.

Litimi, H., BenSaïda, A., and Bouraoui, O. (2016). Herding and excessive risk in the American stock market: A sectoral analysis. *Research in International Business and Finance*, 38:6 – 21.

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

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory.* Wiley, New York.

Noureldin, D., Shephard, N., and Sheppard, K. (2012). Multivariate high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, 27(6):907–933.

Noureldin, D., Sheppard, K., and Shephard, N. (2014). Multivariate Rotated ARCH models. *Journal of Econometrics*, 179:16–30.

Pedersen, R. S. and Rahbek, A. (2014). Multivariate variance targeting in the BEKK-GARCH model. *The Econometrics Journal*, 17(1):24–55.

Srivastava, M. (2003). Singular Wishart and multivariate beta distributions. *The Annals of Statistics*, 31(5):1537–1560.

Yu, P. L., Li, W., and Ng, F. (2017). The Generalized Conditional Autoregressive Wishart Model for Multivariate Realized Volatility. *Journal of Business & Economic Statistics*, 35:513–527.

Asset symbol	Sector	Mean	Median	Std.dev.	Skewness	Kurtosis
AAPL	Technology	7.50	3.24	16.32	20.96	717.68
ACGL	Finance	5.88	1.22	17.43	9.47	148.73
ADSK	Technology	6.93	3.46	14.14	20.00	750.00
AEIS	Capital Goods	14.82	7.90	23.74	6.96	87.48
ALKS	Health Care	14.57	7.05	78.40	62.26	4198.73
ALXN	Health Care	16.73	5.57	42.68	22.04	889.67
AMAT	Technology	6.95	3.54	10.17	5.12	47.30
AMD	Technology	13.96	8.15	24.28	12.08	250.71
AMGN	Health Care	3.84	1.94	6.59	9.91	203.10
AMZN	Consumer Services	11.53	3.75	28.18	14.43	383.06
BIIB	Health Care	9.49	3.55	19.43	8.67	128.95
BLDP	Energy	20.62	12.33	33.13	10.30	206.96
LAMR	Consumer Services	7.25	2.76	15.82	10.33	184.48
MAR	Consumer Services	4.17	2.04	7.55	10.84	225.43
MAT	Consumer Non-Durables	4.76	2.26	15.05	34.39	1594.72
MNST	Consumer Non-Durables	27.05	4.60	182.58	59.89	3981.45
MSFT	Technology	2.83	1.49	4.40	6.44	70.66
MU	Technology	12.70	6.99	20.07	6.92	81.74
MVIS	Capital Goods	36.22	21.45	110.05	50.61	3130.91
MYL	Health Care	5.81	2.80	21.90	43.36	2467.91
NBIX	Health Care	21.72	8.95	43.53	9.62	180.42
NDSN	Capital Goods	7.17	3.56	11.57	6.76	89.07
NKTR	Health Care	16.57	8.66	36.92	25.14	1034.04
NTAP	Technology	13.93	4.37	36.38	20.80	804.59
NTRS	Finance	4.12	1.78	13.89	35.61	1827.34
NWBI	Finance	15.05	3.60	119.05	63.19	4280.99
ODP	Consumer Services	12.66	6.00	32.41	13.46	282.85
ONB	Finance	4.67	2.20	8.61	6.61	70.38
PAAS	Basic Industries	17.23	8.47	24.45	4.34	33.21
PAYX	Consumer Services	5.02	1.86	8.71	6.26	78.72
PDCO	Health Care	4.69	1.93	8.85	7.31	105.74
PDLI	Health Care	19.64	6.77	43.73	10.12	192.87
QCOM	Technology	6.79	2.80	12.15	5.28	44.64
RMBS	Technology	19.00	7.08	132.51	62.13	4185.96
RRD	Miscellaneous	5.31	2.56	43.34	47.42	2326.36
RYAAY	Transportation	16.24	3.73	63.71	18.54	476.35
SBUX	Consumer Services	5.04	2.53	7.72	5.68	54.07
SEIC	Finance	7.20	2.65	13.35	5.83	56.30
SIVB	Finance	9.23	3.37	29.89	21.26	649.90
SLGN	Consumer Durables	10.37	2.61	75.15	54.69	3459.49
TIVO	Miscellaneous	17.50	4.87	143.80	38.89	1665.27
TRMB	Capital Goods	12.52	4.18	27.89	12.68	311.98
TRMK	Finance	6.44	2.70	15.61	16.45	457.88
UBSI	Finance	6.15	2.82	12.74	16.75	545.36
WABC	Finance	4.52	2.31	8.81	14.04	374.78
WBA	Health Care	3.34	1.84	7.07	27.64	1237.02
VOD	Public Utilities	4.11	1.42	8.59	8.30	125.28
VRTX	Health Care	12.84	6.53	21.14	7.33	101.82
WTFC	Finance	7.07	2.49	18.93	19.20	627.96
ZION	Finance	6.87	2.29	19.76	12.72	252.23

Table 1: *Summary statistics for the realized variance (multiplied by* 10^4*) of the 50 assets considered.*

Sector	Nr.assets	Mean	Median	Std.dev.	Skewness	Kurtosis
Basic Industries	1	17.23	8.47	24.45	4.34	33.21
Capital Goods	4	17.68	7.61	59.31	82.02	9377.41
Consumer Durables	1	10.37	2.61	75.15	54.69	3459.49
Consumer Non-Durables	2	15.90	2.96	130.01	83.19	7769.35
Consumer Services	6	7.61	2.98	19.80	18.38	629.25
Energy	1	20.62	12.33	33.13	10.30	206.96
Finance	11	7.02	2.43	39.47	159.67	32358.94
Health Care	11	11.75	4.48	36.99	63.56	8048.78
Miscellaneous	2	11.41	3.44	106.37	49.79	2823.32
Public Utilities	1	4.11	1.42	8.59	8.30	125.28
Technology	9	10.06	4.23	48.09	146.27	26943.34
Transportation	1	16.24	3.73	63.71	18.54	476.35

Table 2: *Summary statistics for the realized variance (multiplied by* 10^4*) of the assets in each of the 12 market sectors considered.*

Model	Parameters	MLL	AIC	BIC
$\text{SCAW}_{\text{CT}}(0,1)$	50	-5517572	11035244	11035565
$\text{SCAW}_{\text{CT}}(1,1)$	100	-4364220	8728640	8729281
$\text{SCAW-SS}_{\text{CT}}(0,1)$	12	-5579755	11159534	11159611
$\text{SCAW-SS}_{\text{CT}}(1,1)$	24	-4400358	8800764	8800918
$\text{SCAW-SS}_{\text{CT}}(1,2)$	36	-4349293	8698658	8698889
$\text{SCAW-SS}_{\text{CT}}(2,2)$	48	**-4298311**	**8596718**	**8597026**
$\text{SCAW-SS-HAR}_{\text{CT}}(1,5)$	24	-4684222	9368492	9368646
$\text{SCAW-SS-HAR}_{\text{CT}}(1,20)$	24	-4608871	9217790	9217944
$\text{MGARCH}_{\text{CT}}(0,1)$	50	-131196	262492	262813
$\text{MGARCH}_{\text{CT}}(1,1)$	100	-116313	232826	233467
$\text{MGARCH-SS}_{\text{CT}}(0,1)$	12	-131448	262920	262997
$\text{MGARCH-SS}_{\text{CT}}(1,1)$	24	-117795	235638	235792
$\text{MGARCH-SS}_{\text{CT}}(1,2)$	36	-112502	225076	225307
$\text{MGARCH-SS}_{\text{CT}}(2,2)$	48	**-106728**	**213552**	**213860**
$\text{MGARCH-SS-HAR}_{\text{CT}}(1,5)$	24	-127951	255950	256104
$\text{MGARCH-SS-HAR}_{\text{CT}}(1,20)$	24	-122686	245420	245574

Table 3: *Summary of the estimation results for the considered models, with various values of* p, q *and* h*. The columns display, in order, the number of parameters, the maximum log-likelihood value (MLL), the Akaikes Information Criterion value (AIC) and the Bayesian Information Criterion value (BIC), for each specification. Within each type of model, SCAW and MGARCH, the most favourable value in each column is emphasized in bold.*

Sector	Nr. elements	$\tilde{\mathbf{A}}_1$	$\tilde{\mathbf{A}}_2$	$\tilde{\mathbf{B}}_1$	$\tilde{\mathbf{B}}_2$
Basic Industries	1	0.15***	0.17***	0.73***	0.65***
Capital Goods	4	0.22***	0.06***	0.61***	0.76***
Consumer Durables	1	0.24***	0.21***	0.65***	0.69***
Consumer Non-Durables	2	0.32***	0.22***	0.61***	0.69***
Consumer Services	6	0.27***	0.099***	0.63***	0.72***
Energy	1	0.36***	0.32***	0.59***	0.61***
Finance	11	0.28***	-0.15***	0.64***	0.7***
Health Care	11	0.25***	-0.14***	0.7***	0.66***
Miscellaneous	2	0.21***	0.35***	0.59***	0.68***
Public Utilities	1	0.13***	0.24***	0.65***	0.71***
Technology	9	0.19***	0.1***	0.74***	0.64***
Transportation	1	0.28***	0.26***	0.63***	0.67***

Table 4: *Estimates for the parameter matrices of SCAW-SS$_{CT}(2,2)$. Three stars next to the estimate indicates that the estimate is significant at the 0.01 level.*

Model		FN_ℓ		
	$\ell =$	1	5	10
$\mathrm{SCAW}_{\mathrm{CT}}(0,1)$		0.018	0.0173$*$	0.0175$*$
$\mathrm{SCAW}_{\mathrm{CT}}(1,1)$		0.0176	0.0177	0.0176
$\mathrm{SCAW\text{-}SS}_{\mathrm{CT}}(0,1)$		0.0176	0.0174$*$	0.0176$*$
$\mathrm{SCAW\text{-}SS}_{\mathrm{CT}}(1,1)$		0.0175	0.0175	0.0173
$\mathrm{SCAW\text{-}SS}_{\mathrm{CT}}(1,2)$		**0.0161**$*$	**0.0163**$*$	**0.0165**$*$
$\mathrm{SCAW\text{-}SS}_{\mathrm{CT}}(2,2)$		0.018	0.0182	0.0183
$\mathrm{SCAW\text{-}SS\text{-}HAR}_{\mathrm{CT}}(1,5)$		0.0181	0.018	0.0176
$\mathrm{SCAW\text{-}SS\text{-}HAR}_{\mathrm{CT}}(1,20)$		0.0182	0.0186	0.0188
$\mathrm{MGARCH}_{\mathrm{CT}}(0,1)$		0.0181	0.0184	0.0185
$\mathrm{MGARCH}_{\mathrm{CT}}(1,1)$		0.017	0.0174	0.0176
$\mathrm{MGARCH\text{-}SS}_{\mathrm{CT}}(0,1)$		0.018	0.0184	0.0185
$\mathrm{MGARCH\text{-}SS}_{\mathrm{CT}}(1,1)$		**0.0165**	**0.0169**	**0.0172**
$\mathrm{MGARCH\text{-}SS}_{\mathrm{CT}}(1,2)$		0.0175	0.0178	0.0179
$\mathrm{MGARCH\text{-}SS}_{\mathrm{CT}}(2,2)$		0.0177	0.0182	0.0184
$\mathrm{MGARCH\text{-}SS\text{-}HAR}_{\mathrm{CT}}(1,5)$		0.0178	0.0183	0.0185
$\mathrm{MGARCH\text{-}SS\text{-}HAR}_{\mathrm{CT}}(1,20)$		0.0175	0.018	0.0182

Table 5: *Summary of the forecast accuracy of FN_ℓ for the considered models, with various values of p, q, h and foreacst horizon ℓ. Within each type of model, SCAW and MGARCH, the most favourable value in each column is emphasized in bold. A star indicates that the model's forecast value is lower than the equivalent value of the other model type, with statistical significance at the* 0.01 *level.*

Model	$\ell =$	$SD_{EW,\ell}$ 1	5	10
$SCAW_{CT}(0,1)$		$2.10e^{-5}*$	$3.08e^{-5}*$	$3.19e^{-5}*$
$SCAW_{CT}(1,1)$		$1.26e^{-5}*$	$1.52e^{-5}*$	$1.64e^{-5}*$
$SCAW\text{-}SS_{CT}(0,1)$		$2.10e^{-5}*$	$3.17e^{-5}*$	$3.21e^{-5}*$
$SCAW\text{-}SS_{CT}(1,1)$		$1.30e^{-5}*$	$1.55e^{-5}*$	$1.70e^{-5}*$
$SCAW\text{-}SS_{CT}(1,2)$		$1.47e^{-5}*$	$1.60e^{-5}*$	$1.72e^{-5}*$
$SCAW\text{-}SS_{CT}(2,2)$		$1.64e^{-5}*$	$1.86e^{-5}*$	$2.03e^{-5}*$
$SCAW\text{-}SS\text{-}HAR_{CT}(1,5)$		$1.30e^{-5}*$	$1.94e^{-5}*$	$2.42e^{-5}*$
$SCAW\text{-}SS\text{-}HAR_{CT}(1,20)$		$\mathbf{1.19e^{-5}}*$	$\mathbf{1.44e^{-5}}*$	$\mathbf{1.55e^{-5}}*$
$MGARCH_{CT}(0,1)$		$5.25e^{-5}$	$5.37e^{-5}$	$5.36e^{-5}$
$MGARCH_{CT}(1,1)$		$3.53e^{-5}$	$3.75e^{-5}$	$3.89e^{-5}$
$MGARCH\text{-}SS_{CT}(0,1)$		$5.28e^{-5}$	$5.37e^{-5}$	$5.36e^{-5}$
$MGARCH\text{-}SS_{CT}(1,1)$		$3.30e^{-5}$	$3.53e^{-5}$	$3.72e^{-5}$
$MGARCH\text{-}SS_{CT}(1,2)$		$\mathbf{3.28e^{-5}}$	$\mathbf{3.44e^{-5}}$	$\mathbf{3.56e^{-5}}$
$MGARCH\text{-}SS_{CT}(2,2)$		$3.30e^{-5}$	$3.55e^{-5}$	$3.76e^{-5}$
$MGARCH\text{-}SS\text{-}HAR_{CT}(1,5)$		$4.95e^{-5}$	$5.28e^{-5}$	$5.36e^{-5}$
$MGARCH\text{-}SS\text{-}HAR_{CT}(1,20)$		$4.41e^{-5}$	$4.62e^{-5}$	$4.79e^{-5}$

Table 6: *Summary of the forecast accuracy of $SD_{EW,\ell}$ for the considered models, with various values of p, q, h and forecast horizon ℓ. Within each type of model, SCAW and MGARCH, the most favourable value in each column is emphasized in bold. A star indicates that the model's forecast value is lower than the equivalent value of the other model type, with statistical significance at the* 0.01 *level.*

Model	$\ell =$	$SD_{GMV,\ell}$ 1	5	10
$SCAW_{CT}(0,1)$		$6.10e^{-3}*$	$6.45e^{-3}*$	$6.51e^{-3}*$
$SCAW_{CT}(1,1)$		$5.75e^{-3}*$	$5.84e^{-3}*$	$5.93e^{-3}*$
$SCAW\text{-}SS_{CT}(0,1)$		$6.10e^{-3}*$	$6.48e^{-3}*$	$6.52e^{-3}$
$SCAW\text{-}SS_{CT}(1,1)$		$\mathbf{5.52e^{-3}}*$	$\mathbf{5.64e^{-3}}*$	$\mathbf{5.71e^{-3}}*$
$SCAW\text{-}SS_{CT}(1,2)$		$5.82e^{-3}*$	$5.90e^{-3}*$	$5.94e^{-3}*$
$SCAW\text{-}SS_{CT}(2,2)$		$5.62e^{-3}*$	$5.68e^{-3}*$	$5.71e^{-3}*$
$SCAW\text{-}SS\text{-}HAR_{CT}(1,5)$		$5.74e^{-3}*$	$5.91e^{-3}*$	$6.09e^{-3}*$
$SCAW\text{-}SS\text{-}HAR_{CT}(1,20)$		$5.72e^{-3}*$	$5.96e^{-3}*$	$6.20e^{-3}$
$MGARCH_{CT}(0,1)$		$6.54e^{-3}$	$6.67e^{-3}$	$6.68e^{-3}$
$MGARCH_{CT}(1,1)$		$6.35e^{-3}$	$6.54e^{-3}$	$6.55e^{-3}$
$MGARCH\text{-}SS_{CT}(0,1)$		$6.56e^{-3}$	$6.67e^{-3}$	$6.68e^{-3}$
$MGARCH\text{-}SS_{CT}(1,1)$		$\mathbf{6.04e^{-3}}$	$\mathbf{6.24e^{-3}}$	$\mathbf{6.29e^{-3}}$
$MGARCH\text{-}SS_{CT}(1,2)$		$6.23e^{-3}$	$6.44e^{-3}$	$6.47e^{-3}$
$MGARCH\text{-}SS_{CT}(2,2)$		$6.27e^{-3}$	$6.50e^{-3}$	$6.51e^{-3}$
$MGARCH\text{-}SS\text{-}HAR_{CT}(1,5)$		$6.48e^{-3}$	$6.65e^{-3}$	$6.67e^{-3}$
$MGARCH\text{-}SS\text{-}HAR_{CT}(1,20)$		$6.38e^{-3}$	$6.51e^{-3}$	$6.56e^{-3}$

Table 7: *Summary of the forecast accuracy of $SD_{GMV,\ell}$ for the considered models, with various values of p, q, h and forecast horizon ℓ. Within each type of model, SCAW and MGARCH, the most favourable value in each column is emphasized in bold. A star indicates that the model's forecast value is lower than the equivalent value of the other model type, with statistical significance at the* 0.01 *level.*