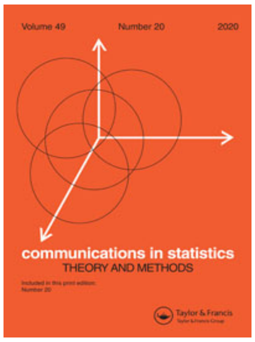

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

Goodness-of-fit tests for centralized Wishart processes

Gustav Alfelt, Taras Bodnar & Joanna Tyrcha

To cite this article: Gustav Alfelt, Taras Bodnar & Joanna Tyrcha (2020) Goodness-of-fit tests for centralized Wishart processes, Communications in Statistics - Theory and Methods, 49:20, 5060-5090, DOI: 10.1080/03610926.2019.1612917

To link to this article: https://doi.org/10.1080/03610926.2019.1612917

Published online: 22 May 2019.

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
2020, VOL. 49, NO. 20, 5060–5090
https://doi.org/10.1080/03610926.2019.1612917

Taylor & Francis
Taylor & Francis Group

Goodness-of-fit tests for centralized Wishart processes

Gustav Alfelt, Taras Bodnar, and Joanna Tyrcha

Department of Mathematics, Stockholm University, Stockholm, Sweden

ABSTRACT

In this paper we present several goodness-of-fit tests for the centralized Wishart process, a popular matrix-variate time series model used to capture the stochastic properties of realized covariance matrices. The new test procedures are based on the extended Bartlett decomposition derived from the properties of the Wishart distribution and allows to obtain sets of independently and standard normally distributed random variables under the null hypothesis. Several tests for normality and independence are then applied to these variables in order to support or to reject the underlying assumption of a centralized Wishart process. In order to investigate the influence of estimated parameters on the suggested testing procedures in the finite-sample case, a simulation study is conducted. Finally, the new test methods are applied to real data consisting of realized covariance matrices computed for the returns on six assets traded on the New York Stock Exchange.

ARTICLE HISTORY
Received 18 October 2018
Accepted 18 April 2019

KEYWORDS
Wishart autoregressive process; goodness-of-fit test; Bartlett decomposition; Wishart distribution; parameter uncertainty

JEL CLASSIFICATION
C12; C32; C52; C58

1. Introduction

The ability to model and to predict covariance matrices of asset returns is a crucial aspect in many financial applications, such as portfolio allocation, option pricing and risk management. A classical discrete time modeling approach is based on the multivariate GARCH-type models which were first introduced by Bollerslev et al. (1988) with the aim to capture the conditional heteroscedasticity present in the daily data. Further modifications of the multivariate GARCH processes were proposed by Engle and Kroner (1995); Engle (2002); Aielli (2013) with Bauwens et al. (2006) providing a review of these types of models. Another approach to model dynamics in the second moments of asset returns is based on the multivariate stochastic volatility models reviewed by Asai et al. (2006).

Due to the rapid development of the computer industry, it has recently become possible to use high-frequency financial data, such as one-minute returns, five-minutes returns, etc. to capture the dynamics in covariance matrices computed for daily frequency (see, e.g., Barndorff-Nielsen and Shephard (2004); Bibinger et al. (2014, 2017) and references therein). This has lead to the development of new approaches to model the stochastic behavior of daily (conditional) covariance matrices based on high frequency data, in which intraday returns are used to consistently estimate low frequency

CONTACT Gustav Alfelt gustava@math.su.se

covariance matrices. A number of such models are based on the assumption that the realized covariance matrices are driven by an underlying centralized Wishart process (Golosnoy et al. 2012; Noureldin et al. 2012).

Wishart processes as models for the multivariate stochastic volatility were introduced by Philipov and Glickman (2006); Gouriéroux et al. (2009). They have recently been intensively investigated in a number of studies, like Golosnoy et al. (2012); Noureldin et al. (2012); Jin and Maheu (2013); Gorgi et al. (2019); Opschoor et al. (2018); Yu et al. (2017); Anatolyev and Kobotaev (2018). This approach appears attractive since the conditional distribution of realized covariance matrices is modeled by the Wishart distribution which is a well-established model for almost surely positive definite matrices (see, e.g., Gupta and Nagar 2000), an inherent property of realized covariance matrices.

In order to investigate how well the model fits observed data, a measure of forecasting accuracy on out-of-sample data is commonly used. Golosnoy et al. (2012) evaluates the fit of the conditional autoregressive Wishart (CAW) model by computing vectors of standardized residuals which are tested for the presence of autocorrelation by applying a univariate Ljung-Box test. This is a common approach used in econometric literature which is based on the fact that the residuals in the misspecified model often are autocorrelated (Johnston and DiNardo 1997). However, this diagnostic procedure does not fully test the underlying assumption of a Wishart process, an assumption which must be fulfilled if the model is to be reliably used as, for example, a forecasting tool. Its main drawback is that it is based on the first two conditional moments of the Wishart process only and, consequently, it cannot detect deviations that are present in the moments of higher orders.

We contribute to the existent literature on the Wishart process by proposing new goodness-of-fit tests on this stochastic matrix-variate model. In the derivation we use the properties of the Wishart distribution to attain the extended Bartlett decomposition from which the test statistics are obtained. The size and power properties of the new test procedures are investigated by the means of a simulation study and they are then compared to those obtained by the application of standardized residuals as suggested in Golosnoy et al. (2012). An important difference between the new procedures and the existent approach is that while the method presented in Golosnoy et al. (2012) controls for the first two conditional moments of the Wishart process, the new procedures take the full distribution into consideration. Moreover, the new method is applicable to any model driven by a centralized Wishart process with known parameters, as well as in the large sample case when the model parameters are consistently estimated. As such it can be employed in a variety of areas, not restricted to cases of a particular model, such as the conditional autoregressive Wishart process.

The paper is structured as follows. Section 2 introduces a centralized Wishart process and presents its distributional properties. The extended Bartlett decomposition used in the derivation of the test statistics is given in Theorem 1. The goodness-of-fit tests are provided in Section 3, while their sizes and powers are investigated in Section 4 and compared to the results obtained for the existent approach. In Section 5 the new testing procedures are applied to real data consisting of six stocks traded on the New York Stock Exchange. Section 6 concludes. Proofs and complementary tables with the results of simulation study are moved to the Appendix.

2. Centralized Wishart process and its stochastic properties

Let $\{\mathbf{R}_t\}_{1\le t\le T}$ be a time series of symmetric positive definite $n\times n$ matrices and let $\mathcal{F}_{t-1}=\{\mathbf{R}_{t-1},\mathbf{R}_{t-2},\ldots\}$ denote a filtration determined by the observation matrices $\mathbf{R}_{t-1}$, $\mathbf{R}_{t-2},\ldots$. Then $\{\mathbf{R}_t\}_{1\le t\le T}$ follows a centralized Wishart process if

$$\mathbf{R}_t|\mathcal{F}_{t-1}\sim\mathcal{W}_n(\nu,\mathbf{S}_t) \tag{1}$$

where $\mathcal{W}_n$ denotes the $n\times n$ Wishart distribution, $\nu>n-1,\nu\in\mathbb{R}_+$ the scalar degrees of freedom[1] and $\mathbf{S}_t$ an $n\times n$ symmetric, positive definite scale matrix measurable with respect to $\mathcal{F}_{t-1}$ at time $t-1$. Further, let $\mathbf{S}_t^{\frac{1}{2}}$ be the lower-triangular Cholesky root of $\mathbf{S}_t$ and define $\mathbf{Q}_t=\mathbf{S}_t^{-\frac{1}{2}}\mathbf{R}_t(\mathbf{S}_t^{-\frac{1}{2}})'$. By Theorem 3.2.5 in Muirhead (1982) we then get

$$\mathbf{Q}_t|\mathcal{F}_{t-1}\sim\mathcal{W}_n(\nu,\mathbf{I}_n) \tag{2}$$

Note that the conditional distribution of $\mathbf{Q}_t$ given $\mathcal{F}_{t-1}$ does not depend on the σ–algebra $\mathcal{F}_{t-1}$ and, consequently, $\mathbf{Q}_t$ is unconditionally Wishart distributed.

Thus, the random matrix $\mathbf{Q}_t$ represents the underlying variability of the centralized Wishart process at time t once the scaling has been removed. In Theorem 1, we present an extension of Bartlett decomposition, presented in Bartlett (1934). While the decomposition in Bartlett (1934) is derived for integer valued degrees of freedom only, Theorem 1 allows for real valued degrees of freedom. This result will be later used in Section 3 to derive goodness-of-fit tests for a centralized Wishart process. The proof of Theorem 1 is similar to the proof of Theorem 3.2.14 in Muirhead (1982), and is given in the Appendix.[2]

Theorem 1. *(Extended Bartlett decomposition)*
Let $\mathbf{A}\sim\mathcal{W}_n(\nu,\mathbf{I}_n)$ where $\nu>n-1,\nu\in\mathbb{R}_+$ and define $\mathbf{A}=\mathbf{T}\mathbf{T}'$ where $\mathbf{T}=(t_{ij})_{i,j=1,\ldots,n}$ is a lower-triangular $n\times n$ matrix with positive diagonal elements. Then:

i. t_{ij}, $1\le j\le i\le n$ are mutually independent;
ii. $t_{ij}\sim\mathcal{N}(0,1)$ *(standard normal distribution) for* $1\le j< i\le n$;
iii. $t_{ii}^2\sim\Gamma(\frac{\nu-i+1}{2},2)$ *(gamma distribution with shape* $(\nu-i+1)/2$ *and scale 2) for* $i=1,\ldots,n$.

With the aid of Theorem 1 we can decompose $\mathbf{Q}_t$ in the following way:

$$\mathbf{Q}_t=\mathbf{U}_t\mathbf{U}_t' \tag{3}$$

where $\mathbf{U}_t$ is the lower-triangual Cholesky root whose squared diagonal elements $u_{ii,t}^2, i=1,\ldots,n$, are distributed as $\Gamma(\frac{\nu-i+1}{2},2)$, while all elements below the diagonal $u_{ij,t}, 1\le j<i\le n$, are standard normally distributed. Moreover, $u_{ij,t}, 1\le j\le i\le n$ are independent. Then define $e_{ij}=u_{ij}, 1\le j<i\le n$ and $e_{ii,t}=\Phi^{-1}(F_{\Gamma(\frac{\nu-i+1}{2},2)}(u_{ii,t}^2))$ for $i=1,\ldots,n$ where Φ denotes the cumulative distribution function (CDF) of the standard normal distribution and $F_{\Gamma(\frac{\nu-i+1}{2},2)}$ stands for the CDF of the $\Gamma(\frac{\nu-i+1}{2},2)$-distribution.

[1]Following the discussion on p. 87 in Muirhead (1982), we use the extended definition of the Wishart distribution allowing for any real valued degrees of freedom $\nu>n-1$.

[2]Simulation techniques based on this result have previously been considered in Ku and Bloomfield (2010) and Owen (2009). In this paper we supply a formal proof.

This integral transformation ensures that $e_{ii,t}, i = 1, \dots, n$, are standard normally distributed random variables (see e.g. Section 6.2.2 in Givens and Hoeting (2012)).

Now let

$$\mathbf{e}_t = (e_{11,t}, \dots, e_{n1,t}, e_{22,t}, \dots, e_{nn,t})' \tag{4}$$

for $1 \le t \le T$. Then, $\mathbf{e}_t \sim \mathcal{N}(\mathbf{0}_k, \mathbf{I}_k)$ (multivariate standard normal distribution) where $k = n(n+1)/2, \mathbf{0}_k$ is a $k \times 1$ zero vector and $\mathbf{I}_k$ a $k \times k$ identity matrix. Finally, using that the conditional distribution of $\mathbf{Q}_t$ does not depend on $\mathcal{F}_{t-1}$ and that in the definition of the vector $\mathbf{e}_t$ only the elements of matrix $\mathbf{Q}_t$ are used, we get that the residual vectors $\mathbf{e}_t, t = 1, \dots, T$ are mutually independently distributed. It is remarkable that this result holds independently of the equation used to model $\mathbf{S}_t$ and can be used to develop goodness-of-fit tests on a centralized Wishart processes which are presented in Section 3.

3. Goodness-of-fit tests

In this section we use the findings of Theorem 1 to derive several goodness-of-fit tests for centralized Wishart processes which in Section 4 are then applied to the special case of the conditionally autoregressive Wishart process considered in Golosnoy et al. (2012).

The null hypotheses of the considered tests are given by

$$\begin{aligned} H_0: &\ \mathbf{R}_t \ \text{ follows a centralized Wishart process with parameters } \nu \text{ and } \mathbf{S}_t \\ \text{against } H_1: &\ \mathbf{R}_t \ \text{ does not follow a centralized Wishart process with parameters } \nu \text{ and } \mathbf{S}_t. \end{aligned} \tag{5}$$

Note that the null hypothesis in (5) is also rejected when $\mathbf{R}_t$ still has a centralized Wishart process but with other parameters as specified under H_0. Hence, it also controls the validity of the model which is fitted to describe the dynamics in $\mathbf{S}_t$. Another possibility to reject H_0 is when the true data generating process deviates from the family of centralized Wishart processes.

In the derivation of the test statistics we employ the properties of centralized Wishart processes discussed in the previous section. Namely, the application of extended Bartlett decomposition with the integral transformation applied to the squared diagonal elements of the matrix $\mathbf{U}_t$ leads to the sequence of independent and multivariate standard normally distributed random vectors $\mathbf{e}_t$ as given in (4). Then the null hypothesis in (5) can be equivalently expressed in terms of the sequence $\mathbf{e}_t, t = 1, \dots, T$, where its distributional and time series properties are verified.

3.1. Noise matrix and its partitions

Using the transformation described after Theorem 1, we define the $k \times T$ *noise matrix* $\mathbf{E}$, again $k = n(n+1)/2$ being the number of components in the vector $\mathbf{e}_t$, as

$$\mathbf{E} = (\mathbf{e}_1, \dots, \mathbf{e}_T) = \{e_{ij,t}\}_{1 \le j \le i \le n, t=1,\dots,T} \tag{6}$$

which is the collection of the vectors given in (4). As a result, the matrix $\mathbf{E}$ consists of $Tn(n+1)/2$ independent and standard normally distributed random variables under the null hypothesis in (5). Hence goodness-of-fit tests evaluating these properties can be applied directly to $\mathbf{E}$ in order to test the validity of the Wishart process.

However, the independence and identical distribution of elements in $\mathbf{E}$ also allows for tests on any suitable partition of $\mathbf{E}$. Such a procedure can be informative about what parts of the matrix process that invalidates the null hypothesis, if any. Given that $\{\mathbf{R}_t\}_{1\leq t\leq T}$ is a discrete time process, it can be of interest to investigate if $\mathbf{R}_t$ at each t follows a centralized Wishart distribution, or if a sequence $\mathbf{R}_{t+1}, ..., \mathbf{R}_{t+\tau}$ follows a centralized Wishart process from time $t+1$ to $t+\tau$, for example during a selected week or month.

To this end, let $T = m\tau$ and consider the column-block partition $\mathbf{E}^B$ given as

$$\mathbf{E} = (\mathbf{E}_1, ..., \mathbf{E}_m) \quad \text{with} \quad \mathbf{E}_l = \left(\mathbf{e}_{l(\tau-1)+1}, ..., \mathbf{e}_{l\tau}\right) \tag{7}$$

Again note that following Theorem 1, the elements of $\mathbf{E}_l$ for $l = 1, ..., m$ are independently and standard normally distributed. Now let d_l be a goodness-of-fit test statistic calculated for $\mathbf{E}_l$ and define $\mathbf{d} = (d_1, ..., d_m)'$. Then the components of $\mathbf{d}$ are independently distributed and each of them is used to test the validity of a centralized Wishart process during the corresponding time period. Finally, one can base the decision of the validity of a centralized Wishart process for each submatrix $\mathbf{E}_l, l = 1, ..., m$ on the vector of p-values $\mathbf{p} = (p_1, ..., p_m)'$, where p_l is the p-value associated with d_l, instead of $\mathbf{d}$. In contrast to $\mathbf{d}$, the univariate marginal distributions of the vector $\mathbf{p}$ are uniform distributions on $[0, 1]$ and consequently they are independent of the dimensionality of matrices $\mathbf{E}_l, l = 1, ..., m$.

Similarly, appropriate goodness-of-fit tests can also be applied to each of the k component of the vectors $\mathbf{e}_t$ for $t = 1, ..., T$. To this end, consider the row partition $\mathbf{E}^C$ given as

$$\mathbf{E} = (\tilde{\mathbf{e}}_{11}, ..., \tilde{\mathbf{e}}_{n1}, \tilde{\mathbf{e}}_{22}, ..., \tilde{\mathbf{e}}_{nn})' \quad \text{with} \quad \tilde{\mathbf{e}}_{ij} = \left(e_{ij,1}, ..., e_{ij,T}\right)' \tag{8}$$

where $1 \leq j \leq i \leq n$. In this case, the aim is to investigate which components of the vector $\mathbf{e}_t$ are responsible for the violation of the null hypothesis in (5). Note that in general when several tests are performed simultaneously the multiplicity correction on the significance level should be kept in mind.

The construction of $\mathbf{E}$ together with appropriate partitions of $\mathbf{E}$ allows us to perform goodness-of-fit tests on any centralized Wishart process aiming to model observed series of symmetric positive definite matrices. When the null hypothesis of a centralized Wishart process is true, any given test statistic assessing independent standard normality, computed on the elements in $\mathbf{E}$ or on the elements in any matrix or vector of its partitions, will follow the corresponding null distribution.

3.2. Tests based on standard normality

Since the elements of $\mathbf{E}$ are independent and standard normally distributed under the null hypothesis in (5), a goodness-fit-test on a centralized Wishart process with known values of ν and $\mathbf{S}_t$ can be performed by testing if

$$\begin{aligned} H_0 &: e_{ij,t} \text{ is normally distributed} \\ \text{against } H_1 &: e_{ij,t} \text{ is not normally distributed} \end{aligned} \tag{9}$$

based on an independent sample $e_{ij,t}$ for $1 \le j \le i \le n$ and $t = 1, ..., T$.

The testing problem (9) is well studied in statistics with a number of existing approaches. For example, the test can be performed by using the Kolmogorov-Smirnov statistic (Massey 1951), the Anderson-Darling test (Anderson and Darling 1954), the Shapiro-Wilk approach (Shapiro and Wilk 1965), the Lilliefors test (Lilliefors 1967).

Similarly we can perform tests on the mean and the variance of the components of **E**. Namely, in using the t-test we will test if

$$H_0 : \mathbb{E}(e_{ij,1}) = 0 \quad \text{against} \quad H_1 : \mathbb{E}(e_{ij,1}) \neq 0 \tag{10}$$

The statistic of the t-test is given by

$$T_{mean} = \sqrt{kT}\frac{\bar{e}}{s_e} \tag{11}$$

where

$$\bar{e} = \frac{1}{kT}\sum_{j\le i}\sum_{t=1}^{T} e_{ij,t} \quad \text{and} \quad s_e^2 = \frac{1}{kT-1}\sum_{j\le i}\sum_{t=1}^{T}(e_{ij,t}-\bar{e})^2$$

are the sample mean and the sample variance obtained from $\{e_{ij,t}\}_{1\le j\le i\le n, t=1,...,T}$. Under the null hypothesis, it holds that T_{mean} is asymptotically standard normally distributed.

Finally, a test on the equality of the variance to one can be applied as well with the hypotheses given by

$$H_0 : \mathbb{V}ar(e_{ij,t}) = 1 \quad \text{against} \quad H_1 : \mathbb{V}ar(e_{ij,t}) \neq 1 \tag{12}$$

and the test statistic expressed as

$$T_{var} = (kT-1)s_e^2 \tag{13}$$

which has a χ^2_{kT-1}-distribution under H_0.

Correspondingly, the above tests can be applied to each matrix in the partition $\mathbf{E}^B$ and to each vector in the partition $\mathbf{E}^C$ in order to test for standard normality in different time periods and for different components of $\mathbf{e}_t$, respectively.

3.3. Tests based on autocorrelations

If a time series model is misspecified, then the residuals obtained from this model often appear to be autocorrelated (see, e.g., Tsay 2010). As a result, testing for the presence of autocorrelation in the residuals is a common way to validate the model in practice. Following this procedure and using the matrix **E**, in this section we construct several goodness-of-fit tests for a centralized Wishart process, designed to verify the presence of autocorrelated vectors in **E**. The testing hypotheses are given by

$$\begin{aligned} H_0 &: \text{No autocorrelation is present up to lag } L \\ \text{against } H_1 &: \text{Auto correlation is present up to lag } L \end{aligned} \tag{14}$$

This test is performed by employing the multivariate Ljung-Box test (see, Tsay 2010) whose asymptotic null distribution was derived by Hosking (1980). The test is applied to **E**, treating its columns as observation vectors.

Alternatively, a collection of univariate Ljung-Box tests can be applied to verify the presence of autocorrelations in each vector of the partition $\mathbf{E}^C$. This approach ignores the cross-sectional dependencies in the residuals, while reducing the degrees of freedom of the limiting χ^2-distribution and thus increasing the power of each individual test. Finally, the test for autocorrelation in residuals is usually performed not only to the original residuals but also to their corresponding squared values.

While the test on the autocorrelations in the residuals and their squared values check the properties of their first two moments, the goodness-of-fit tests of Section 3.2 monitor the whole distribution function. Consequently, the tests based on the Ljung-Box statistics seems to be a better choice if the fit of $\mathbf{S}_t$ is of interest, while the approach of Section 3.2 might be more useful to check if the Wishart distribution is an appropriate assumption for the standardized process $\mathbf{Q}_t$ as given in (3).

3.4. Specification of ν and $\mathbf{S}_t$

The tests in Sections 3.2 and 3.3 are designed under the assumption that both quantities ν and $\mathbf{S}_t$ are known precisely. The violation of this assumption could have a large effect on the performance of the suggested tests. The parameter ν is usually fitted in practice by the maximum likelihood estimator and, consequently, it is consistent and asymptotically normal distributed under some regularity conditions.

The situation of the specification of $\mathbf{S}_t$ is more involved. By its definition, the matrix $\mathbf{S}_t$ is measurable with respect to filtration $\mathcal{F}_{t-1}$. In practice, however, several approaches could be applied to model $\mathbf{S}_t$ with the conditional autoregressive Wishart (CAW) process considered in Golosnoy et al. (2012) and the HEAVY model introduced by Noureldin et al. (2012) to be the most popular ones. In both cases $\mathbf{S}_t$ is modeled by its previous realizations, the realizations of the process $\mathbf{R}_t$ up to time $t-1$, and some parameter matrices. Noureldin et al. (2012) used the maximum likelihood method to estimate the parameter matrices and proved that these estimators are consistent and asymptotically normally distributed. As a result, if the sample size is relatively large with respect to the process dimension, then the impact of the estimation error is expected to be small and it can be ignored in practice. On the other side, if the sample size is not large enough, then it has to be taken into account, which could be achieved via bootstrap (Davison and Hinkley 1997; Horowitz et al. 2003; Efron and Hastie 2016).

4. Finite sample performance

In order to investigate the size and power properties of the proposed test procedures under parameter uncertainty when the sample size T is finite, two simulation studies of the conditional autoregressive Wishart process of Golosnoy et al. (2012) are conducted in this section. The first study, found in Section 4.3, investigates the power when the autoregressive structure of a Wishart process is violated, i.e., the model equation for $\mathbf{S}_t$ is misspecified. The second study, found in Section 4.4, investigates the size and the power when both the autoregressive structure of the Wishart process as well as the assumption of the Wishart distribution are violated.

Table 1. Parameters and start values used in the simulation study.

Parameter	$\mathbf{A}_1$	$\mathbf{A}_2$	$\mathbf{B}_1$	$\mathbf{B}_2$	C	ν	$\mathbf{R}_0$	$\mathbf{S}_0$	$\mathbf{R}_{-1}$	$\mathbf{S}_{-1}$
Value	$\frac{1}{2}\mathbf{I}_n$	$\frac{2}{5}\mathbf{I}_n$	$\frac{1}{2}\mathbf{I}_n$	$\frac{2}{5}\mathbf{I}_n$	$\mathbf{I}_n$	20	$\mathbf{I}_n$	$\mathbf{I}_n$	$\mathbf{I}_n$	$\mathbf{I}_n$

4.1. Conditional autoregressive Wishart process

Given a filtration of past observations $\mathcal{F}_{t-1} = \{\mathbf{R}_{t-1}, \mathbf{R}_{t-2}, \ldots\}$ the conditionally autoregressive Wishart process of order (p, q) denoted by CAW(p, q) is specified as

$$\mathbf{R}_t|\mathcal{F}_{t-1} \sim \mathcal{W}_n(\nu, \mathbf{S}_t/\nu) \tag{15}$$

with

$$\mathbf{S}_t = \mathbf{CC}' + \sum_{i=1}^{p} \mathbf{B}_i\mathbf{S}_{t-i}\mathbf{B}_i' + \sum_{i=1}^{q} \mathbf{A}_i\mathbf{R}_{t-i}\mathbf{A}_i' \tag{16}$$

Golosnoy et al. (2012) derive the maximum likelihood estimators for the parameters of this model, while Noureldin et al. (2012) provided the conditions under which these estimators are consistent when the sample size T tends to infinity. In our simulation study we deal with the case of finite T with the aim to investigate the influence of parameter uncertainty on the performance of the suggested goodness-of-fit tests. We also compare the new procedures with the existent approach for model diagnostics which is proposed in Golosnoy et al. (2012) and is described in Section 4.2 as a benchmark procedure.

In the simulation study, we use the diagonal CAW(p, q) model (here denoted by DCAW(p, q)), as suggested in Golosnoy et al. (2012) of sizes $n = 2$, 4, 6. The DCAW model entails that each parameter matrix $\mathbf{A}_1, \ldots, \mathbf{A}_q, \mathbf{B}_1, \ldots \mathbf{B}_p$ in (15) and (16) is diagonal with n unknown parameters. The diagonal CAW model is chosen over the full CAW model in order to reduce parameter estimation time.

The simulations are performed by generating a series $\{\mathbf{R}_t\}_{1\le t\le T}$ of T matrices that are conditionally Wishart distributed according to (15) where $\mathbf{S}_t$ in (16) is computed with the starting values $\mathbf{R}_0, \mathbf{R}_{-1}, \mathbf{S}_0, \mathbf{S}_{-1}$ and parameters $\theta_{p,q} = \{\mathbf{A}_1, \ldots, \mathbf{A}_q, \mathbf{B}_1, \ldots \mathbf{B}_p, \mathbf{C}, \nu\}$ as given in Table 1. Then the parameters of the model are estimated by the maximum likelihood method in each simulation, and a scale matrix series $\{\hat{\mathbf{S}}_t\}_{1\le t\le T}$ is computed given the estimated parameters. Using both $\{\mathbf{R}_t\}_{1\le t\le T}$ and $\{\hat{\mathbf{S}}_t\}_{1\le t\le T}$ the noise matrix $\mathbf{E}$ is then computed according to (2)-(6). Finally, the proposed tests of Section 3 are applied first to $\mathbf{E}$, then to the matrices of the partition $\mathbf{E}^B$ and then to the vectors of the partition $\mathbf{E}^C$.

4.2. Benchmark procedure

As a benchmark method to the tests on the vector partition $\mathbf{E}^C$ proposed in Section 3, we consider the testing procedure presented in Golosnoy et al. (2012). Let

$$\mathbf{v}_t = Var[\mathbf{r}_t|\mathcal{F}_{t-1}]^{-\frac{1}{2}}(\mathbf{r}_t - E[\mathbf{r}_t|\mathcal{F}_{t-1}]) \tag{17}$$

where $\mathbf{r}_t = \text{vech}(\mathbf{R}_t)$ with vech(.) denote the **vech** operator which transforms a symmetric matrix to a vector (see, e.g., Harville (1997)). For the conditional mean vector and covariance matrix, Golosnoy et al. (2012) uses, in accordance with the Wishart distribution,

$$E[\mathbf{R}_t|\mathcal{F}_{t-1}] = \mathbf{S}_t \tag{18}$$

and

$$Cov\left[r_{ij,t}, r_{lm,t}|\mathcal{F}_{t-1}\right] = \frac{1}{\nu}\left(s_{il,t}s_{jm,t} + s_{im,t}s_{jl,t}\right) \tag{19}$$

where $\mathbf{S}_t = \{s_{ij,t}\}_{i,j=1,\dots,n}$ for $t = 1, \dots, T$.

As for the testing procedures of Section 3, we also define the corresponding noise matrix as

$$\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_T) = \{v_{ij,t}\}_{1\leq j\leq i\leq n, t=1,\dots,T} \tag{20}$$

with, for the k components of $\mathbf{v}_t$, the partition $\mathbf{V}^C$ given as

$$\mathbf{V} = (\tilde{\mathbf{v}}_{11}, \dots, \tilde{\mathbf{v}}_{n1}, \tilde{\mathbf{v}}_{22}, \dots, \tilde{\mathbf{v}}_{nn})' \quad \text{with} \quad \tilde{\mathbf{v}}_{ij} = \left(v_{ij,1}, \dots, v_{ij,T}\right)' \tag{21}$$

Thus, under the null hypothesis the elements of each vector $\tilde{\mathbf{v}}_{ij}$ will be serially uncorrelated. As such, the approach presented in Golosnoy et al. (2012) applies the univariate Ljung-Box test to each of these vectors as a model diagnostic.

Note that the benchmark procedure involves only the first two conditional moments of an autoregressive Wishart process. As such, we expect that it will be able to detect violations from the model assumptions related to this, like a misspecification of the conditional model for $\mathbf{S}_t$, while it might have some difficulties if the departures from the model assumptions are present in the dynamics of higher moments.

4.3. Violation of autoregressive structure

The model from which $\mathbf{R}_t, t = 1, \dots, T$ are drawn is a DCAW(2,2) model with parameters and starting values as listed in Table 1. In each simulation run, we consider several values for the model dimension $n \in \{2, 4, 6\}$ and generate 2100 random matrices. In order to remove the effect of the initial values, the first 100 observations are discarded, leaving the sample of size $T = 2000$ in each series.

The violation of the autoregressive structure in the equation for $\mathbf{S}_t$ is captured by fitting DCAW(0,0), DCAW(0,1), DCAW(1,1), DCAW(1,2), DCAW(2,1) models to the generated data by applying the maximum likelihood method.[3] We also fit DCAW(2,2) process to the generated data in order to study the size of the suggested tests whose nominal level is set to be equal to 5%.

The set of estimated parameters for a DCAW(p, q) model is denoted $\hat{\theta}_{p,q} = \{\hat{\mathbf{A}}_1, \dots, \hat{\mathbf{A}}_q, \hat{\mathbf{B}}_1, \dots \hat{\mathbf{B}}_p, \hat{\mathbf{C}}, \hat{\nu}\}$, and the associated conditional scale matrix series as $\{\hat{\mathbf{S}}_{t;p,q}\}$ for $t = 1, \dots, T$. We further use the notation $\mathbf{E}_{p,q}$ and $\mathbf{V}_{p,q}$ for the noise matrices $\mathbf{E}$ and $\mathbf{V}$ obtained from $\{\mathbf{R}_t\}, \{\hat{\mathbf{S}}_{t;p,q}\}, t = 1, \dots, T$ and the estimator $\hat{\nu}$. Finally, let $\mathbf{E}^B_{p,q}$ and $\mathbf{E}^C_{p,q}$

[3]The DCAW(1,0) model is not included in the simulation study, since it is not influenced by the observed data series $\{\mathbf{R}_t\}_{1\leq t\leq T}$ and the performance of the considered tests is similar to the ones obtained for the DCAW(0,0) model.

be the block and component partitions of $\mathbf{E}_{p,q}$ and $\mathbf{V}^C_{p,q}$ the component partition of $\mathbf{V}_{p,q}$. In $\mathbf{E}^B_{p,q}$, a block size of 20 columns is used.

The complete simulation procedure for each $n = 2$, 4, 6 is summarized as follows:

1. Simulate a matrix series of size 2100 using (15) and (16) given the starting values $\mathbf{S}_{-1}, \mathbf{S}_0, \mathbf{R}_{-1}$ and $\mathbf{R}_0$ and a parameter set $\theta_{2,2} = \{\mathbf{A}_1, \mathbf{A}_2, \mathbf{B}_1, \mathbf{B}_2, \mathbf{C}, \nu\}$ which are provided in Table 1. The first 100 values of the series is discarded in order to remove the effect of starting values. The resulting series is denoted by $\{\mathbf{R}_t\}_{1\le t\le T}$.
2. Estimate parameters sets $\hat{\theta}_{0,0}, \hat{\theta}_{0,1}$, $\hat{\theta}_{1,1}, \hat{\theta}_{1,2}, \hat{\theta}_{2,1}$ and $\hat{\theta}_{2,2}$ of the models DCAW(0,0), DCAW(0,1), DCAW(1,1), DCAW(1,2), DCAW(2,1), and DCAW(2,2) respectively.
3. Given the estimated parameters and the series $\{\mathbf{R}_t\}_{1\le t\le T}$, compute the corresponding scale matrix series $\{\hat{\mathbf{S}}_{t;p,q}\}_{1\le t\le T}$ corresponding to each of the models of different lag orders (p, q) as in the previous step.
4. For each order (p, q), calculate the noise matrices $\mathbf{E}_{p,q}$ and $\mathbf{V}_{p,q}$ as well as their partitions $\mathbf{E}^B_{p,q}$, $\mathbf{E}^C_{p,q}$, and $\mathbf{V}^C_{p,q}$.
5. Apply the following tests, and record their resulting p-values:
   - To $\mathbf{E}_{p,q}$: Anderson-Darling test on normality, Lilliefors test on normality, t-test for the equality of mean to 0, χ^2-test for the equality of variance 1, multivariate Ljung-Box test to the columns of $\mathbf{E}_{p,q}$ as well as to their the squared values;[4,5]
   - To each column block of partition $\mathbf{E}^B_{p,q}$: Anderson-Darling test on normality, Lilliefors test on normality, Shapiro-Wilks test on normality, t-test for the equality of mean to 0, χ^2-test for the equality of variance to 1;
   - To each vector of partition $\mathbf{E}^C_{p,q}$: Anderson-Darling test on normality, Lilliefors test on normality, Shapiro-Wilks test on normality, t-test for the equality of mean to 0, χ^2-test for the equality of variance 1, univariate Ljung-Box test to each vector as well as to their squares;
   - To each vector of the partition $\mathbf{V}^C_{p,q}$: univariate Ljung-Box test.
6. Repeat the steps (1)–(5) 1000 times and compute the relative number of rejections for each goodness-of-fit test with the nominal significance level of $\alpha = 5\%$.

The rejection rates in the case of $\mathbf{E}_{p,q}$ are reported in Table 2 for $n = 2$, 4, 6, while the corresponding results in the case of the component-wise partitions $\mathbf{E}^C_{p,q}$ and $\mathbf{V}^C_{p,q}$ are given in Table 3 for $n = 2$. Additional results for the component-wise partitions $\mathbf{E}^C_{p,q}$ and $\mathbf{V}^C_{p,q}$ for $n = 4$ and $n = 6$ as well as for the block partition $\mathbf{E}^B_{p,q}$ can be found in the Appendix (see, Tables A1–A4 and Table A5, respectively). In the case of the block partition the average rejection rates with respect to the blocks are provided. The Ljung-Box tests on squared residuals are denoted" Ljung-Box sq." in the tables. Note that while all tests asymptotically have the nominal rejection rate under the correct model, this does not necessarily hold in the finite-sample case.

[4]In all Ljung-Box tests the number of lags in the null hypothesis is set to 8, i.e., the closest integer to ln (2000).

[5]Since the number of elements in $\mathbf{E}_{p,q}$ extends 5000, the Shapiro-Wilks test for normality is not applied to this set, in line with Mahibbur Rahman and Govindarajulu (1997).

Table 2. Rejection rates of the goodness-of-fit tests based on $\mathbf{E}_{p,q}$ in the case of n = 2, 4, 6.

n = 2						
Test	DCAW(0,0)	DCAW(0,1)	DCAW(1,1)	DCAW(1,2)	DCAW(2,1)	DCAW(2,2)
Anderson-Darling	**0.86**	**0.21**	0.06	0.06	0.06	0.06
Lilliefors	**0.68**	**0.15**	0.05	0.04	0.05	0.05
Mean 0	0.00	0.00	0.00	0.00	0.00	0.00
Variance 1	0.00	0.00	0.00	0.00	0.00	0.00
Multivariate Ljung-Box	**1.00**	**1.00**	**0.35**	0.04	**0.35**	0.02
Multivariate Ljung-Box sq.	**1.00**	**0.55**	0.06	0.06	0.06	0.06
n = 4						
Test	DCAW(0,0)	DCAW(0,1)	DCAW(1,1)	DCAW(1,2)	DCAW(2,1)	DCAW(2,2)
Anderson-Darling	**0.95**	**0.27**	0.06	0.05	0.06	0.06
Lilliefors	**0.82**	**0.15**	0.04	0.04	0.04	0.04
Mean 0	0.00	0.00	0.00	0.00	0.00	0.00
Variance 1	0.00	0.00	0.00	0.00	0.00	0.00
Multivariate Ljung-Box	**1.00**	**1.00**	**0.36**	0.04	**0.35**	0.03
Multivariate Ljung-Box sq.	**1.00**	**0.70**	**0.07**	**0.07**	**0.07**	0.06
n = 6						
Test	DCAW(0,0)	DCAW(0,1)	DCAW(1,1)	DCAW(1,2)	DCAW(2,1)	DCAW(2,2)
Anderson-Darling	**1.00**	**0.49**	0.04	0.04	0.04	0.04
Lilliefors	**0.92**	**0.26**	0.02	0.03	0.03	0.03
Mean 0	0.00	0.00	0.00	0.00	0.00	0.00
Variance 1	0.00	0.00	0.00	0.00	0.00	0.00
Multivariate Ljung-Box	**1.00**	**1.00**	**0.38**	0.04	**0.38**	0.03
Multivariate Ljung-Box sq.	**1.00**	**0.83**	0.04	0.05	0.04	0.05

Data is simulated from a DCAW(2,2) model, and then estimated for DCAW model of order (0,0), (0,1), (1,1), (1,2), (2,1), (2,2). Rejection rates are reported for each test applied to each model fit. Rejection rates significantly larger than the nominal level of 5% are emphasized in bold.

Further, in order to determine if the empirical rejection rates obtained via simulations are larger than the nominal level $\alpha = 0.05$ with statistical significance, we apply a binomial test.[6] For a fixed model, sample size and test, let w denote the probability of a resulting p-value being below $\alpha = 0.05$. As such, for k independent simulations, the number of rejections will follow a binomial distribution with k trials and success probability w. We employ this to test the hypothesis $H_0 : w \leq \alpha$ against $H_1 : w > \alpha$, where the number of rejections will follow a binomial distribution with $k = 1000$ trials and success probability $w = 0.05$ under the null hypothesis. As such, each rejection rate based on a number of rejections that are equal to or exceeds the 99% quantile of this distribution is reported in bold, indicating that they are significantly larger (on the 1%-level) than the nominal level $\alpha = 0.05$. In the case of the block partition $\mathbf{E}^{B}_{p,q}$, the null distribution is binomial with trials equal to k times the number of blocks.

In Table 2 we observe that all tests on the noise matrix $\mathbf{E}_{p,q}$ possess rejection rates of around or below 0.05 for all dimensions $n = 2$, 4, and 6 under the true model DCAW(2,2), as expected. However, similar rejection rates are also present under the alternative model DCAW(1,2), indicating that the tests cannot distinguish between these two models. The specific reason for this have not been studied in detail, but it is likely that the second lag order, q, is able to capture the majority of the autoregressive structure, and that the DCAW(1,2) model thus is able to describe dynamics of the DCAW(2,2) model to a sufficiently large extent, for the given sample size and process

[6]Similar methods are presented in e.g. Oliveira and Ferreira (2010), Batsidis and Zografos (2013) and Batsidis et al. (2014). We would like to thank the anonymous referee for suggesting this type of approach.

Table 3. Rejection rates for the goodness-of-fit tests based on the component-wise partitions $\mathbf{E}^C_{p,q}$ and $\mathbf{V}^C_{p,q}$ in the case of n = 2.

	e_{11}	e_{12}	e_{22}		e_{11}	e_{12}	e_{22}
Set: $\mathbf{E}^C_{p,q}$. Test: Anderson-Darling.				Set: $\mathbf{E}^C_{p,q}$. Test: Variance 1.			
DCAW(0,0)	**0.74**	**0.20**	**0.69**	DCAW(0,0)	**0.15**	**0.15**	**0.21**
DCAW(0,1)	**0.16**	0.06	**0.18**	DCAW(0,1)	0.02	0.04	0.03
DCAW(1,1)	**0.07**	0.06	0.05	DCAW(1,1)	0.02	0.02	0.01
DCAW(1,2)	**0.07**	0.06	0.06	DCAW(1,2)	0.01	0.01	0.02
DCAW(2,1)	**0.07**	0.06	0.05	DCAW(2,1)	0.02	0.02	0.02
DCAW(2,2)	**0.07**	0.06	0.06	DCAW(2,2)	0.01	0.02	0.02
Set: $\mathbf{E}^C_{p,q}$. Test: Lilliefors.				Set: $\mathbf{E}^C_{p,q}$. Test: Ljung-Box.			
DCAW(0,0)	**0.57**	**0.12**	**0.53**	DCAW(0,0)	**1.00**	**1.00**	**1.00**
DCAW(0,1)	**0.11**	0.04	**0.12**	DCAW(0,1)	**1.00**	**1.00**	**1.00**
DCAW(1,1)	0.06	0.05	0.06	DCAW(1,1)	**0.28**	**0.32**	**0.27**
DCAW(1,2)	0.05	0.05	0.05	DCAW(1,2)	0.02	0.04	0.02
DCAW(2,1)	0.06	0.05	0.06	DCAW(2,1)	**0.28**	**0.31**	**0.27**
DCAW(2,2)	0.05	0.05	0.05	DCAW(2,2)	0.01	0.02	0.01
Set: $\mathbf{E}^C_{p,q}$. Test: Shapiro-Wilks.				Set: $\mathbf{E}^C_{p,q}$. Test: Ljung-Box sq.			
DCAW(0,0)	**0.84**	**0.28**	**0.79**	DCAW(0,0)	**1.00**	**1.00**	**1.00**
DCAW(0,1)	**0.24**	**0.07**	**0.24**	DCAW(0,1)	**0.40**	**0.48**	**0.36**
DCAW(1,1)	**0.07**	0.06	0.06	DCAW(1,1)	0.06	0.05	0.05
DCAW(1,2)	0.06	0.06	0.06	DCAW(1,2)	0.04	0.04	0.02
DCAW(2,1)	**0.07**	0.06	0.06	DCAW(2,1)	0.06	0.05	0.05
DCAW(2,2)	0.06	0.06	0.06	DCAW(2,2)	0.06	0.04	0.04
Set: $\mathbf{E}^C_{p,q}$. Test: Mean 0.				Set: $\mathbf{V}^C_{p,q}$. Test: Ljung-Box.	v_{11}	v_{12}	v_{22}
DCAW(0,0)	0.00	0.00	0.00	DCAW(0,0)	**1.00**	**1.00**	**1.00**
DCAW(0,1)	0.00	0.00	0.00	DCAW(0,1)	**1.00**	**1.00**	**1.00**
DCAW(1,1)	0.00	0.00	0.00	DCAW(1,1)	**0.29**	**0.31**	**0.28**
DCAW(1,2)	0.00	0.00	0.00	DCAW(1,2)	0.02	0.04	0.01
DCAW(2,1)	0.00	0.00	0.00	DCAW(2,1)	**0.29**	**0.31**	**0.27**
DCAW(2,2)	0.00	0.00	0.00	DCAW(2,2)	0.01	0.03	0.01

Data are simulated from a DCAW(2,2) model, and then estimated for DCAW models of order (0,0), (0,1), (1,1), (2,1), (1,2) and (2,2). Rejection rates are reported for each test applied to each model fit. Rejection rates significantly larger than the nominal level of 5% are emphasized in bold.

dimension. In the case of DCAW(1,1) and DCAW(2,1), the rejection rates of multivariate Ljung-Box test are around 0.3, while the other tests have rejection rates of around or below 0.05 as for the true model. For DCAW(0,1) and DCAW(0,0), both the normality tests and multivariate Ljung-Box tests have high rejection rates. The tests based on the mean and the variance has a rejection rate of zero for all models, indicating that the first two moments are well captured by each of the estimated models. In general, the Ljung-Box tests on squared residuals produce similar or lower rejection rates than the ordinary Ljung-Box test.

Table 3 shows similar results for the component-wise partitions $\mathbf{E}^C_{p,q}$ and $\mathbf{V}^C_{p,q}$ in case $n = 2$. Corresponding results for $n = 4$, 6 are found in Tables A1–A4 in the Appendix. The tests on components of DCAW(0,0) DCAW(0,1) exhibit rejection rates well above nominal level for a number of the considered tests. The univariate Ljung-Box test produces rejection rates of around 0.35 on components of DCAW(1,1) and DCAW(2,1), while rejection rates for components of DCAW(1,2) are around nominal level for all tests. Comparing the univariate Ljung-Box tests on the vector components of $\mathbf{E}^C_{p,q}$ derived in Section 3 and on the vector components of $\mathbf{V}^C_{p,q}$ derived in the benchmark method in Section 4.2, equivalent rejection rates are produced.

Finally, Table A5 in the Appendix show the rejection rates from the column block partition $\mathbf{E}^B_{p,q}$ with 20 columns in each block. Here high rejection rates for DCAW(0,1) and DCAW(0,0) is observed, mainly for the tests on mean and variance.

To summarize, concerning the tests on $\mathbf{E}_{p,q}$, the models DCAW(0,0) and DCAW(0,1) are rejected by the multivariate Ljung-Box test in 100% of the simulations, and rejected by the normality tests to some extent. The models DCAW(1,1) and DCAW(2,1) are rejected by the multivariate Ljung-Box in about 35% of the simulations. The DCAW(1,2) model does not exhibit a rejection rate above nominal levels for any of the tests. Regarding the univariate Ljung-Box method, the procedure proposed in Section 3 and the benchmark approach produce similar results. In general, a good performance by the tests for serial autocorrelation is not surprising, since the violations are present in the autoregressive structure in the introduced alternative models.

4.4. Violation of autoregressive structure and Wishart distribution

This study closely follows the study presented in Section 4.3 but a further violation of the true model is present in the data generating process. Given a filtration of past observations $\mathcal{F}_{t-1} = \{\mathbf{R}_{t-1}, \mathbf{R}_{t-2}, ...\}$ we define

$$\tilde{\mathbf{R}}_t = \frac{d-2}{\eta_t}\mathbf{R}_t \tag{22}$$

$$\mathbf{R}_t|\mathcal{F}_{t-1} \sim \mathcal{W}_n(\nu, \mathbf{S}_t/\nu) \tag{23}$$

$$\eta_t \sim \chi^2_d \tag{24}$$

$$\mathbf{S}_t = \mathbf{C}\mathbf{C}' + \sum_{i=1}^{p} \mathbf{B}_i \mathbf{S}_{t-i} \mathbf{B}'_i + \sum_{i=1}^{q} \mathbf{A}_i \tilde{\mathbf{R}}_{t-i} \mathbf{A}'_i \tag{25}$$

where $\mathbf{R}_t|\mathcal{F}_{t-1}$ and η_t are independent; the parameters and starting values are given again in Table 1. Thus, although $E[\tilde{\mathbf{R}}_t|\mathcal{F}_{t-1}] = E[\mathbf{R}_t|\mathcal{F}_{t-1}]$, $\tilde{\mathbf{R}}_t$ does not follow a centralized Wishart process, since the conditional distribution is not longer a Wishart distribution. In order to assess the rejection rates of the considered testing procedures based on this model, the steps 1-6 of the simulation study presented in Section 4.3 are repeated with the series $\{\mathbf{R}_t\}_{1\le t\le T}$ replaced by $\{\tilde{\mathbf{R}}_t\}_{1\le t\le T}$ for $d = 10$, 20, 30. Note that the case $d \to \infty$ produces the models presented in 4.3. Again, the procedure is repeated 1000 times and the rejection rates are computed for the noise matrix $\mathbf{E}_{p,q}$ (see, Table 4), the component-wise partitions $\mathbf{E}^C_{p,q}$ and $\mathbf{V}^C_{p,q}$ (see, Table 5 and Tables A6–A7 in the Appendix) as well as for the block partition[7] $\mathbf{E}^B_{p,q}$ (see, Table A8 in the Appendix) at nominal significance level $\alpha = 5\%$. Note that all of the considered models deviate from the process specified under H_0.

For $n = 2$ and each value $d = 10$, 20, 30, Table 4 displays rejection rates equal to 1 for all normality tests applied to $\mathbf{E}_{p,q}$ for each model DCAW(p, q). The multivariate Ljung-Box tests possess similar rejection rates to those presented in the simulation study of

[7]Again with column block size 20.

Table 4. Rejection rates of the goodness-of-fit tests based on $\mathbf{E}_{p,q}$ in the case of n = 2 and d = 10, 20, 30.

d = 10						
Test	DCAW(0,0)	DCAW(0,1)	DCAW(1,1)	DCAW(1,2)	DCAW(2,1)	DCAW(2,2)
Anderson-Darling	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Lilliefors	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Mean 0	0.00	0.00	0.00	0.00	0.00	0.00
Variance 1	0.00	0.00	0.00	0.00	0.00	0.00
Multivariate Ljung-Box	**1.00**	**1.00**	**0.32**	**0.07**	**0.32**	0.06
Multivariate Ljung-Box sq.	**1.00**	**0.96**	**0.11**	**0.10**	**0.11**	**0.10**
d = 20						
Test	DCAW(0,0)	DCAW(0,1)	DCAW(1,1)	DCAW(1,2)	DCAW(2,1)	DCAW(2,2)
Anderson-Darling	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Lilliefors	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Mean 0	0.00	0.00	0.00	0.00	0.00	0.00
Variance 1	0.00	0.00	0.00	0.00	0.00	0.00
Multivariate Ljung-Box	**1.00**	**1.00**	**0.33**	0.03	**0.33**	0.03
Multivariate Ljung-Box sq.	**1.00**	**0.90**	**0.07**	**0.07**	**0.07**	**0.07**
d = 30						
Test	DCAW(0,0)	DCAW(0,1)	DCAW(1,1)	DCAW(1,2)	DCAW(2,1)	DCAW(2,2)
Anderson-Darling	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Lilliefors	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Mean 0	0.00	0.00	0.00	0.00	0.00	0.00
Variance 1	0.00	0.00	0.00	0.00	0.00	0.00
Multivariate Ljung-Box	**1.00**	**1.00**	**0.34**	0.04	**0.33**	0.02
Multivariate Ljung-Box sq.	**1.00**	**0.83**	**0.07**	0.06	**0.07**	0.06

Data is simulated from (22), and then estimated for DCAW model of order (0,0), (0,1), (1,1), (1,2), (2,1), (2,2). Rejection rates are reported for each test applied to each model fit. Rejection rates significantly larger than the nominal level of 5% are emphasized in bold.

Section 4.3 for all models. Also, similarly to the study in the previous section, the tests on mean and variance produce low rejection rates.

Table 5 shows the rejection rates for the component-wise partitions $\mathbf{E}^C_{p,q}$ and $\mathbf{V}^C_{p,q}$ in the case of $d = 30$. The results for $d = 10$ and $d = 20$ are found in Tables A6 and A7 in the Appendix, respectively. The rejection rates are similar for each simulation study independently of $d = 10$, 20, 30, although they tend to drop with increasing d, regarding most tests. This is expected since larger values of d correspond to smaller violations to the assumption of the true DCAW(2,2) model. For each model, the tests on normality and variance produce high rejection rates regarding most of the components. Rejection rates for the mean tests and tests on the off-diagonal component e_{12} tend to be lower in general. However, rejection rates on the mean tend to be substantially higher than for the study in the previous subsection, see e.g. Table A6. The univariate Ljung-Box tests show similar patterns as those observed in the previous simulation study for both the tests proposed in Section 3 and the benchmark approach, with rejection rates close to 1 for DCAW(0,0) and DCAW(0,1); around 0.2–0.3 for DCAW(1,1) and DCAW(2,1); and rejection rates at nominal levels for DCAW(1,2) and DCAW(2,2). As such, given a correctly specified autoregressive structure, applying the benchmark method described in Section 4.2 will not detect violations to the distributional assumption, while applying the procedures presented in Section 3 will detect such violations.

Table A8 presents the rejection rates from the block partition $\mathbf{E}^B_{p,q}$ with 20 columns in each block. In general, each test has a rejection rate of 0.09 or above, for every model

Table 5. Rejection rates for the goodness-of-fit tests based on the component-wise partitions $\mathbf{E}^C_{p,q}$ and $\mathbf{V}^C_{p,q}$ in the case of n = 2 and d = 30.

	e_{11}	e_{12}	e_{22}		e_{11}	e_{12}	e_{22}
Set: $\mathbf{E}^C_{p,q}$. Test: Anderson-Darling.				Set: $\mathbf{E}^C_{p,q}$. Test: Variance 1.			
DCAW(0,0)	**1.00**	**0.77**	**1.00**	DCAW(0,0)	**1.00**	**1.00**	**0.85**
DCAW(0,1)	**1.00**	**0.33**	**1.00**	DCAW(0,1)	**1.00**	**1.00**	**0.98**
DCAW(1,1)	**1.00**	**0.17**	**1.00**	DCAW(1,1)	**1.00**	**1.00**	**1.00**
DCAW(1,2)	**1.00**	**0.18**	**1.00**	DCAW(1,2)	**1.00**	**1.00**	**1.00**
DCAW(2,1)	**1.00**	**0.17**	**1.00**	DCAW(2,1)	**1.00**	**1.00**	**1.00**
DCAW(2,2)	**1.00**	**0.18**	**1.00**	DCAW(2,2)	**1.00**	**1.00**	**1.00**
Set: $\mathbf{E}^C_{p,q}$. Test: Lilliefors.				Set: $\mathbf{E}^C_{p,q}$. Test: Ljung-Box.			
DCAW(0,0)	**1.00**	**0.52**	**1.00**	DCAW(0,0)	**1.00**	**1.00**	**1.00**
DCAW(0,1)	**1.00**	**0.18**	**0.99**	DCAW(0,1)	**1.00**	**1.00**	**1.00**
DCAW(1,1)	**0.96**	**0.10**	**0.96**	DCAW(1,1)	**0.27**	**0.33**	**0.25**
DCAW(1,2)	**0.96**	**0.11**	**0.95**	DCAW(1,2)	0.02	0.06	0.02
DCAW(2,1)	**0.96**	**0.09**	**0.96**	DCAW(2,1)	**0.27**	**0.33**	**0.25**
DCAW(2,2)	**0.95**	**0.12**	**0.96**	DCAW(2,2)	0.01	0.04	0.01
Set: $\mathbf{E}^C_{p,q}$. Test: Shapiro-Wilks.				Set: $\mathbf{E}^C_{p,q}$. Test: Ljung-Box sq.			
DCAW(0,0)	**1.00**	**0.87**	**1.00**	DCAW(0,0)	**1.00**	**1.00**	**1.00**
DCAW(0,1)	**1.00**	**0.53**	**1.00**	DCAW(0,1)	**0.53**	**0.61**	**0.66**
DCAW(1,1)	**1.00**	**0.32**	**1.00**	DCAW(1,1)	0.06	0.04	0.05
DCAW(1,2)	**1.00**	**0.32**	**1.00**	DCAW(1,2)	0.05	0.04	0.05
DCAW(2,1)	**1.00**	**0.32**	**1.00**	DCAW(2,1)	0.06	0.04	0.05
DCAW(2,2)	**1.00**	**0.32**	**1.00**	DCAW(2,2)	0.05	0.04	0.05
Set: $\mathbf{E}^C_{p,q}$. Test: Mean 0.				Set: $\mathbf{V}^C_{p,q}$. Test: Ljung-Box.	v_{11}	v_{12}	v_{22}
DCAW(0,0)	0.01	0.00	**0.50**	DCAW(0,0)	**1.00**	**1.00**	**1.00**
DCAW(0,1)	0.00	0.00	**0.22**	DCAW(0,1)	**1.00**	**1.00**	**1.00**
DCAW(1,1)	0.00	0.00	**0.07**	DCAW(1,1)	**0.24**	**0.30**	**0.23**
DCAW(1,2)	0.00	0.00	**0.08**	DCAW(1,2)	0.02	0.06	0.02
DCAW(2,1)	0.00	0.00	**0.07**	DCAW(2,1)	**0.24**	**0.30**	**0.23**
DCAW(2,2)	0.00	0.00	0.06	DCAW(2,2)	0.01	0.05	0.01

Data are simulated from (22), and then estimated for DCAW models of order (0,0), (0,1), (1,1), (2,1), (1,2) and (2,2). Rejection rates are reported for each test applied to each model fit. Rejection rates significantly larger than the nominal level of 5% are emphasized in bold.

and value of d. As mentioned previously, the rejection rates decrease as d increases and the violation of the assumption of a pure Wishart distribution becomes smaller.

To summarize, the normality tests on $\mathbf{E}_{p,q}$ presented in Section 3 are able to reject each of the tested models when the data are generated as in (22) with the values $d = 10$, 20, 30 in 100% of the simulations. The tests on the block and component partitions similarly produce rejection rates above nominal levels. This is in contrast to the benchmark method based on (20) presented in Golosnoy et al. (2012) which is not able to detect the violations in the DCAW(2,2) model.

5. Empirical application

5.1. Data

The goodness-of-fit testing procedures suggested in Section 3 are applied to evaluate the CAW model fitted to a series of realized covariance matrices for six stocks traded on the New York Stock Exchange: American Express Inc. (AXP), Citigroup (C), General Electric (GE), Home Depot Inc. (HD), International Business Machines (IBM) and JPMorgan Chase & Co. (JPM). This is the same data set as presented in Chiriac and

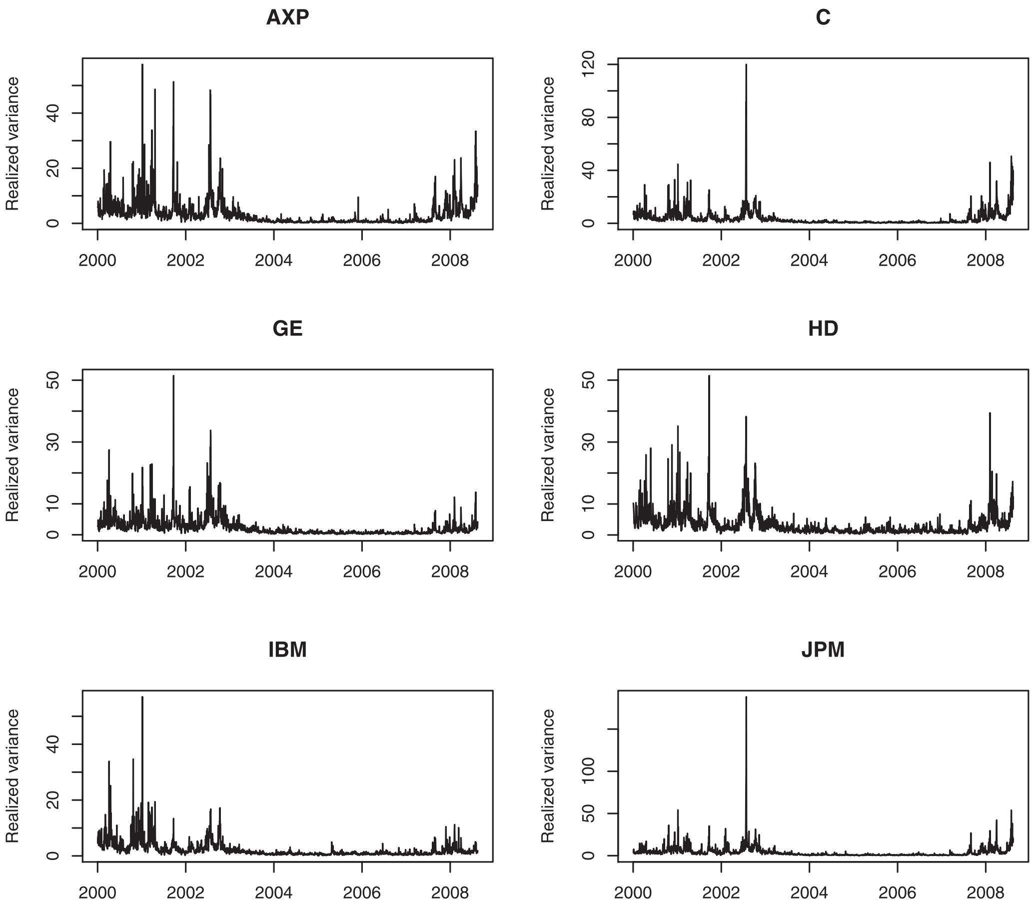

Figure 1. Realized variances of the six stocks traded on the New York Stock Exchange: American Express Inc. (AXP), Citigroup (C), General Electric (GE), Home Depot Inc. (HD), International Business Machines (IBM) and JPMorgan Chase & Co. (JPM) for the period from the 1st of January, 2000 to the 30th of July, 2008. The values are obtained as the corresponding diagonal elements $r_{ii,t}, i = 1, ..., 6$ of the realized covariance matrices $\mathbf{R}_t$ for $t = 1, ..., T$.

Voev (2011), that also describe how the realized covariance matrices have been computed. The series consists of daily realized covariance matrices in time period from the 1st of January, 2000 to the 30th of July, 2008 resulting in 2156 daily observations. The realized variance development of each stock over the considered time period is displayed in Figure 1. Higher volatilities are present during the period from 2000 to 2003 with a further increase observed at the end of the considered period.

5.2. Fitting the CAW models to data

Using the maximum likelihood method, CAW(p, q) models (as described by (15) and (16)) of orders $p, q \leq 2$ are fitted to the data described in Section 5.1. Since the considered CAW models employ up to two lagged values, the first two observations are used for $\mathbf{S}_{-1}, \mathbf{S}_0, \mathbf{R}_{-1}$ and $\mathbf{R}_0$. The remaining $T-2 = 2154$ data points are used to estimate the parameters. The estimating procedure is conducted first for the lowest lag order

Table 6. The p-values of the proposed goodness-of-fit tests on the validity of CAW(p,q) models with $p, q \leq 2$ based on $\mathbf{E}_{p,q}$.

Test	CAW(0,0)	CAW(0,1)	CAW(1,1)	CAW(1,2)	CAW(2,1)	CAW(2,2)
Anderson-Darling	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Lilliefors	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Mean 0	**0.00**	0.61	0.21	0.16	0.14	0.11
Variance 1	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Multivariate Ljung-Box	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Multivariate Ljung-Box sq.	**0.00**	**0.00**	**0.00**	**0.01**	**0.00**	**0.01**

The data consist of the realized covariance matrices calculated for the six stocks traded on the New York Stock Exchange: American Express Inc. (AXP), Citigroup (C), General Electric (GE), Home Depot Inc. (HD), International Business Machines (IBM) and JPMorgan Chase & Co. (JPM) for the period from the 1st of January, 2000 to the 30th of July, 2008. Bold values indicate that the null hypothesis of a CAW model is rejected at significance level of 5%.

model. Models with higher order lags then use the estimates of the lower order model as starting values in the numerical maximization of the likelihood function. Additionally, for each model, a set of several different starting values is considered in order to control for local maximums. The testing procedures used in Section 4 are then conducted.

Table 6 presents the p-values of the tests applied to the noise matrix $\mathbf{E}_{p,q}$ derived from each of the estimated CAW(p, q) models with $p, q \leq 2$.[8] For almost all of the considered tests with the exception of the t-test on mean, the null hypothesis of an CAW model is rejected at significance level $\alpha = 0.05$ (even 0.01).

Tables A9 and A10 in the Appendix show the p-values for the tests performed on the component-wise partitions $\mathbf{E}^C_{p,q}$ and $\mathbf{V}^C_{p,q}$. Here, the null hypothesis is rejected in most cases when the tests on normality are employed. However, as the model lag order increases, fewer tests reject the null hypothesis in general. Note that the univariate Ljung-Box test based on the partition $\mathbf{V}^C_{p,q}$ as in (21) and (17) rejects fewer components than the same test based on the partition $\mathbf{E}^C_{p,q}$. Such a result can be explained by the fact that the elements of the vectors in $\mathbf{E}^C_{p,q}$ are independent by construction, while the elements of the vectors in $\mathbf{V}^C_{p,q}$ are just uncorrelated. This could influence the power of the test suggested in Golosnoy et al. (2012).

Furthermore, Table A11 displays the rejection rates of the proposed goodness-of-fit tests applied to the block partition $\mathbf{E}^B_{p,q}$ with block size equal to 20 (resulting in 107 blocks), as described in Section 4. The tests for normality, namely Anderson-Darling test, Lilliefors test, and Shapiro-Wilk test, possess the average rejection rates of roughly 1 for all of the considered models, while the test on the mean, the test on the variance, and the test on autocorrelation are less powerful for the lag orders above (0,0). No rejection rate is however close to 5%, which would support the null hypothesis of each block having i.i.d standard normal elements.

Further, let $\mathbf{p} = (p_1, \ldots, p_{107})'$, where $p_l, l = 1, \ldots, 107$ is the p-value associated with a specific test applied to column block l. In accordance with the discussion in Section 3.1, the elements of $\mathbf{p}$ will be independent and will each follow a uniform distribution on $[0, 1]$ under the null hypothesis. In order to illustrate the model fit in different time periods, the vector $\mathbf{p}$ for p-values of the various tests on the block partition of the

[8]The lag order (1,0) is again excluded with the same motivation as in Section 4.

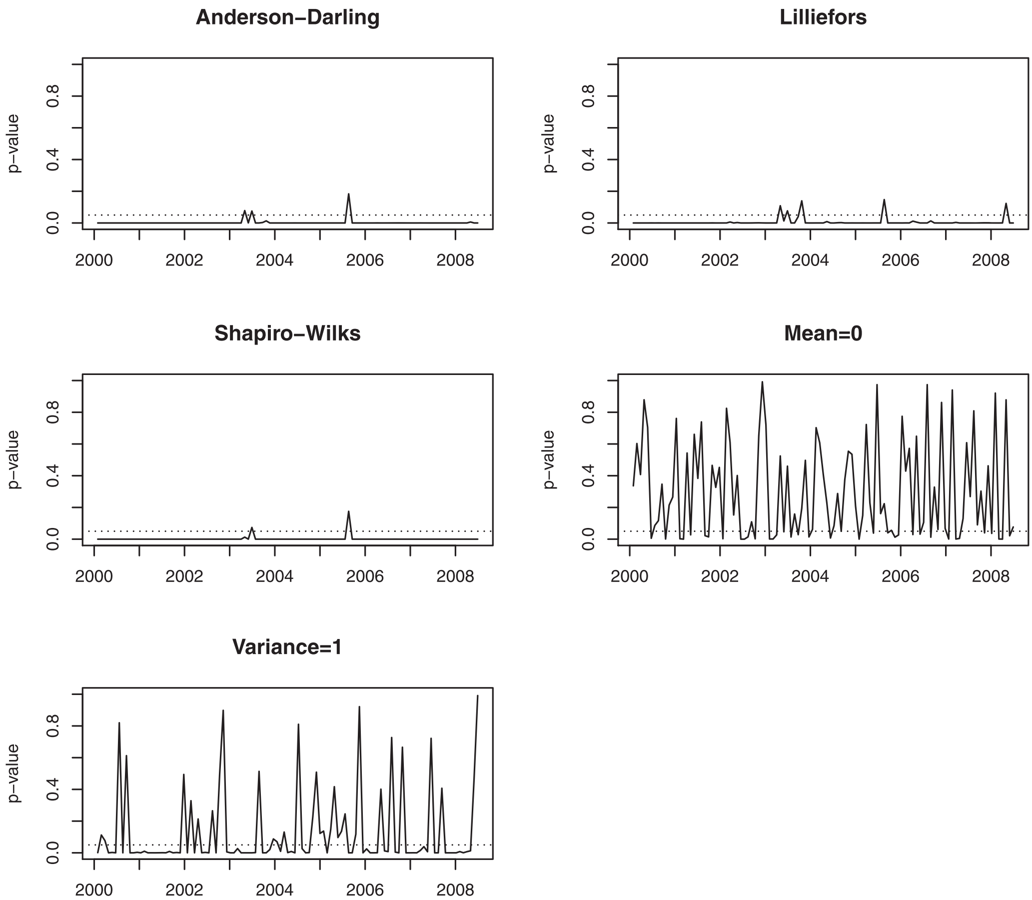

Figure 2. The p-values of the proposed goodness-of-fit tests on the validity of the CAW(2,2) model based on the column block partition $\mathbf{E}^{B}_{p,q}$ with the block size equal to 20, plotted over time. The data consist of the realized covariance matrices calculated for the six stocks traded on the New York Stock Exchange: American Express Inc. (AXP), Citigroup (C), General Electric (GE), Home Depot Inc. (HD), International Business Machines (IBM) and JPMorgan Chase & Co. (JPM) for the period from the 1st of January, 2000 to the 30th of July, 2008. The dotted line represents the nominal significance level of 5%.

CAW(2,2) model is plotted in Figure 2. The dotted line represents the nominal significance level of 5% and as such, a value below this line mean that the null hypohtesis for the corresponding block, which in turn represents 20 trading days,[9] is rejected at significance level 5%. The three tests for normality rejects basically every block, with a few exceptions which takes place mainly in the middle of the time period, namely around 2003–2006. According to Figure 1 this is a period where the realized stock variances are relatively low. The p-values from the test on the mean is relatively evenly spread over the time period, while the p-values obtained from the test on the variance tend to be somewhat clustered for low p-values.

[9]The last block represents 34 trading days.

In general the suggested testing procedures do not support the statement that the supplied data can be well fitted with a CAW model of the lag order (2,2) or lower.

6. Conclusion

Model diagnostics plays an important role when a time series model is fitted to real data. When the model is misspecified, then residuals calculated are usually autocorrelated. As a result, the classical way to validate the model's ability to fit real data is based on an autocorrelation test which is usually done by using the Ljung-Box test.

Recently, modeling and analyzing high-frequency data have become very popular topics in financial econometrics (see, e.g., Andersen et al. 2003; Hautsch 2011). Owning to developments of computer techniques, it is today possible to store and analyze huge data sets with the aim of improving the performance of holding portfolio. As a result, the usage of the realized covariance matrix in portfolio theory has become a popular topic in finance (Hautsch et al. 2015; Callot et al. 2017).

The Wishart autoregressive process has recently been introduced to capture the dynamics in realized covariance matrices (see, e.g., Gouriéroux et al. 2009; Golosnoy et al. 2012; Noureldin et al. 2012; Gorgi et al. 2019; Opschoor et al. 2018; Yu et al. 2017). In order to validate the fit of the model, the Ljung-Box test was considered in Golosnoy et al. 2012). However, this approach relies only on the first two conditional moments of the autoregressive Wishart process and, consequently, it cannot detect violations from the model in higher moments.

In this paper, we suggest an alternative procedure for validating the assumption of a centralized Wishart process which is based on the extended Barlett decomposition of a random matrix that has a Wishart distribution. As a result, several procedures for testing goodness-of-fit for centralized Wishart processes are derived. Since the model depends on unknown parameters which are estimated by employing the maximum likelihood method, the parameter uncertainty should be kept in mind when the testing procedures are applied to real data of small and moderate sizes. We investigate this point through simulations and find that the suggested goodness-of-fit tests are relatively robust to this issue. Finally, an application to a real data set consisting of six stock traded on the New York Stock Exchange is provided. Here, we found that the conditional autoregressive Wishart process does not provide a good fit to real data due to deviations from the model assumptions which are present in higher moments.

Regarding the test methods presented in this paper, one suggestion for future research include combining several of the suggested tests into one sequential test, for example in the spirit of Lin and Wu (2017). Another approach is to investigate the performance as process dimensionality increases from moderate to very high.

Moreover, our empirical finding motivates further research in the topic of modeling the dynamics in the realized covariance matrices which should take into account not only the conditionally heteroscedastic behavior of the scale matrix of the Wishart distribution, but also possess the ability to capture the dynamics in higher moments. Although this is a very important topic in financial engineering, the issue remain unresolved and require further methodological development and empirical investigation in the future.

Acknowledgments

The authors are grateful to the Editor, the Associated Editor and the anonymous referee for valuable comments and suggestions that helped to improve the paper.

References

Aielli, G. P. 2013. Dynamic conditional correlation: on properties and estimation. *Journal of Business & Economic Statistics* 31 (3):282–99. doi:10.1080/07350015.2013.771027.

Anatolyev, S., and N. Kobotaev. 2018. Modeling and forecasting realized covariance matrices with accounting for leverage. *Econometric Reviews* 37 (2):114–39. doi:10.1080/07474938.2015.1035165.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys. 2003. Modeling and forecasting realized volatility. *Econometrica* 71 (2):579–625. doi:10.1111/1468-0262.00418.

Anderson, T. W., and D. A. Darling. 1954. A test of goodness of fit. *Journal of the American Statistical Association* 49 (268):765–9. doi:10.1080/01621459.1954.10501232.

Asai, M., M. McAleer, and J. Yu. 2006. Multivariate stochastic volatility: a review. *Econometric Reviews* 25 (2–3):145–75. doi:10.1080/07474930600713564.

Barndorff-Nielsen, O. E., and N. Shephard. 2004. Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica* 72 (3):885–925. doi:10.1111/j.1468-0262.2004.00515.x.

Bartlett, M. S. 1934. On the theory of statistical regression. *Proceedings of the Royal Society of Edinburgh* 53:260–83. doi:10.1017/S0370164600015637.

Batsidis, A., N. Martin, L. Pardo, and K. Zografos. 2014. A necessary power divergence-type family of tests for testing elliptical symmetry. *Journal of Statistical Computation and Simulation* 84 (1):57–83. doi:10.1080/00949655.2012.694437.

Batsidis, A., and K. Zografos. 2013. A necessary test of fit of specific elliptical distributions based on an estimator of Song's measure. *Journal of Multivariate Analysis* 113:91–105. Special Issue on Multivariate Distribution Theory in Memory of Samuel Kotz. doi:10.1016/j.jmva.2011.09.006.

Bauwens, L., S. Laurent, and J. V. Rombouts. 2006. Multivariate GARCH models: a survey. *Journal of Applied Econometrics* 21 (1):79–109. doi:10.1002/jae.842.

Bibinger, M., N. Hautsch, P. Malec, and M. Reiss. 2017. Estimating the spot covariation of asset prices–statistical theory and empirical evidence. *Journal of Business & Economic Statistics* 8: 1–17doi:10.1080/07350015.2017.1356728.

Bibinger, M., N. Hautsch, P. Malec, and M. Reiß. 2014. Estimating the quadratic covariation matrix from noisy observations: Local method of moments and efficiency. *The Annals of Statistics* 42 (4):1312–46. doi:10.1214/14-AOS1224.

Bollerslev, T., R. F. Engle, and J. M. Wooldridge. 1988. A capital asset pricing model with time-varying covariances. *Journal of Political Economy* 96 (1):116–31. doi:10.1086/261527.

Callot, L. A., A. B. Kock, and M. C. Medeiros. 2017. Modeling and forecasting large realized covariance matrices and portfolio choice. *Journal of Applied Econometrics* 32 (1):140–58. doi:10.1002/jae.2512.

Chiriac, R., and V. Voev. 2011. Modelling and forecasting multivariate realized volatility. *Journal of Applied Econometrics* 26 (6):922–47. doi:10.1002/jae.1152.

Davison, A. C., and D. V. Hinkley. 1997. *Bootstrap methods and their application, volume 1*. Cambridge: Cambridge University Press.

Efron, B., and T. Hastie. 2016. *Computer age statistical inference: Algorithms, evidence, and data science. Institute of mathematical statistics monographs*. Cambridge: Cambridge University Press.

Engle, R. 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* 20 (3):339–50. doi:10.1198/073500102288618487.

Engle, R. F., and K. F. Kroner. 1995. Multivariate simultaneous generalized ARCH. *Econometric Theory* 11 (1):122–50. doi:10.1017/S0266466600009063.

Givens, G. H., and J. A. Hoeting. 2012. *Computational statistics*. Hoboken, NJ: John Wiley & Sons.

Golosnoy, V., B. Gribisch, and R. Liesenfeld. 2012. The conditional autoregressive Wishart model for multivariate stock market volatility. *Journal of Econometrics* 167 (1):211–23. doi:10.1016/j.jeconom.2011.11.004.

Gorgi, P., S. J. Koopman, P. R. Hansen, and P. Janus. 2019. Realized Wishart-GARCH: A score-driven multi-asset volatility model. *Journal of Financial Econometrics* 17 (1):1–32. doi:10.1093/jjfinec/nby007.

Gouriéroux, C., J. Jasiak, and R. Sufana. 2009. The Wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics* 150 (2):167–81. doi:10.1016/j.jeconom.2008.12.016.

Gupta, A. K., and D. K. Nagar. 2000. *Matrix variate distributions*. Boca Raton, FL: CRC Press.

Harville, D. 1997. *Matrix algebra from statistician's perspective*. New York: Springer.

Hautsch, N. 2011. *Econometrics of financial high-frequency data*. Berlin, Germany: Springer Science & Business Media.

Hautsch, N., L. M. Kyj, and P. Malec. 2015. Do high-frequency data improve high-dimensional portfolio allocations? *Journal of Applied Econometrics* 30 (2):263–90. doi:10.1002/jae.2361.

Horowitz, J. L., et al. 2003. The bootstrap in econometrics. *Statistical Science* 18 (2):211–8.

Hosking, J. R. 1980. The multivariate portmanteau statistic. *Journal of the American Statistical Association* 75 (371):602–8. doi:10.1080/01621459.1980.10477520.

Jin, X., and J. M. Maheu. 2013. Modeling realized covariances and returns. *Journal of Financial Econometrics* 11 (2):335–69. doi:10.1093/jjfinec/nbs022.

Johnston, J., and J. DiNardo. 1997. *Econometric methods*. New York: McGraw-Hill.

Ku, Y., and P. Bloomfield. 2010. Generating random Wishart matrices with fractional degrees of freedom in OX. Oxmetrics User Conferense.

Lilliefors, H. W. 1967. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association* 62 (318):399–402. doi:10.1080/01621459.1967.10482916.

Lin, J., and X. Wu. 2017. A sequential test for the specification of predictive densities. *The Econometrics Journal* 20 (2):190–220. doi:10.1111/ectj.12085.

Mahibbur Rahman, M., and Z. Govindarajulu. 1997. A modification of the test of Shapiro and Wilk for normality. *Journal of Applied Statistics* 24 (2):219–36. doi:10.1080/02664769723828.

Massey, F. J. 1951. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association* 46 (253):68–78. doi:10.1080/01621459.1951.10500769.

Muirhead, R. J. 1982. *Aspects of multivariate statistical theory*. New York: Wiley.

Noureldin, D., N. Shephard, and K. Sheppard. 2012. Multivariate high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics* 27 (6):907–33. doi:10.1002/jae.1260.

Oliveira, I. R. C. D., and D. F. Ferreira. 2010. Multivariate extension of chi-squared univariate normality test. *Journal of Statistical Computation and Simulation* 80 (5):513–26. doi:10.1080/00949650902731377.

Opschoor, A., P. Janus, A. Lucas, and D. V. Dijk. 2018. New heavy models for fat-tailed realized covariances and returns. *Journal of Business & Economic Statistics* 36 (4):643–57. doi:10.1080/07350015.2016.1245622.

Owen, A. 2009. Monte Carlo Theory, Methods and Examples.

Philipov, A., and M. E. Glickman. 2006. Multivariate stochastic volatility via Wishart processes. *Journal of Business & Economic Statistics* 24 (3):313–28. doi:10.1198/073500105000000306.
Shapiro, S. S., and M. B. Wilk. 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52 (3–4):591–611. doi:10.2307/2333709.
Tsay, R. S. 2010. *Analysis of financial time series*. Hoboken, NJ: John Wiley & Sons.
Yu, P. L., W. Li, and F. Ng. 2017. The generalized conditional autoregressive Wishart model for multivariate realized volatility. *Journal of Business & Economic Statistics* 35 :513–27. doi: 10.1080/07350015.2015.1096788.

Appendix

Proof of Theorem 1

The density function of A is given by

$$f(\mathbf{A}) = \frac{1}{2^{n\nu/2}\Gamma_n\left(\frac{\nu}{2}\right)} e^{tr\left(-\frac{1}{2}\mathbf{A}\right)} |\mathbf{A}|^{(\nu-n-1)/2} d\mathbf{A} \tag{26}$$

where $d\mathbf{A}$ is the volume element of $\mathbf{A}$. Following the proof of Theorem 3.2.14 in Muirhead (1982), we get

$$\begin{aligned} tr(\mathbf{A}) &= \sum_{j\leq i}^{n} t_{ij}^2 \\ |\mathbf{A}| &= \prod_{i=1}^{n} t_{ii}^2 \\ d\mathbf{A} &= 2^n \prod_{i=1}^{n} t_{ii}^{n+1-i} \bigwedge_{j\leq i}^{n} dt_{ij} \end{aligned}$$

Substituting these equalities into (26) leads to

$$f(\mathbf{A}) = \frac{1}{2^{n\nu/2}\Gamma_n\left(\frac{\nu}{2}\right)} e^{-\sum_{j\leq i}^{n} t_{ij}^2} 2^n \prod_{i=1}^{n} t_{ii}^{\nu-i} \bigwedge_{j\leq i}^{n} dt_{ij}$$

where

$$\begin{aligned} \Gamma_n\left(\frac{\nu}{2}\right) &= \pi^{n(n-1)/4} \prod_{i=1}^{n} \Gamma\left(\frac{\nu-i+1}{2}\right) \\ 2^{n\nu/2-n(n-1)/4-n} &= 2^{\sum_{i=1}^{n}\left(\frac{\nu-i+1}{2}-1\right)} \end{aligned}$$

Hence,

$$f(\mathbf{A}) = \prod_{j<i}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{t_{ij}^2}{2}} dt_{ij} \prod_{i=1}^{n} \frac{1}{2^{(\nu-i+1)/2-1}\Gamma\left(\frac{\nu-i+1}{2}\right)} e^{-\frac{t_{ii}^2}{2}} t_{ii}^{\nu-i} dt_{ii}$$

or, equivalently,

$$f(\mathbf{A}) = \prod_{j<i}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{t_{ij}^2}{2}} dt_{ij} \prod_{i=1}^{n} \frac{1}{2^{(\nu-i+1)/2}\Gamma\left(\frac{\nu-i+1}{2}\right)} e^{-\frac{t_{ii}^2}{2}} \left(t_{ii}^2\right)^{(\nu-i+1)/2-1} dt_{ii}^2$$

which is the joint density of independent random variables $t_{ij} \sim \mathcal{N}(0, 1)$ and $t_{ii}^2 \sim \Gamma(\frac{\nu-i+1}{2}, 2), 1 \leq j<i \leq n$. Finally, the support of shape-parameter of the univariate gamma distribution imposes the condition $\nu > n-1$.

Table A1. Rejection rates for the goodness-of-fit tests based on the component-wise partition $\mathbf{E}^C_{p,q}$ in the case of $n = 4$.

	e_{11}	e_{12}	e_{13}	e_{14}	e_{22}	e_{23}	e_{24}	e_{33}	e_{34}	e_{44}
Set: $\mathbf{E}^C_{p,q}$. Test: Anderson-Darling.										
DCAW(0,0)	**0.73**	**0.18**	**0.20**	**0.21**	**0.71**	**0.17**	**0.17**	**0.66**	**0.18**	**0.63**
DCAW(0,1)	**0.18**	0.06	**0.07**	0.06	**0.16**	0.05	0.05	**0.15**	0.06	**0.13**
DCAW(1,1)	0.05	0.05	0.05	0.06	0.04	0.06	0.04	0.05	0.06	0.05
DCAW(1,2)	0.05	0.05	0.05	0.05	0.04	0.05	0.04	0.05	0.05	0.05
DCAW(2,1)	0.05	0.05	0.05	0.06	0.04	0.06	0.04	0.05	0.06	0.05
DCAW(2,2)	0.06	0.05	0.05	0.05	0.04	0.05	0.04	0.05	0.05	0.05
Set: $\mathbf{E}^C_{p,q}$. Test: Lilliefors.										
DCAW(0,0)	**0.56**	**0.13**	**0.13**	**0.14**	**0.55**	**0.11**	**0.12**	**0.48**	**0.12**	**0.44**
DCAW(0,1)	**0.13**	0.05	0.05	0.05	**0.12**	0.04	0.04	**0.12**	0.04	**0.10**
DCAW(1,1)	0.05	0.04	0.06	0.04	0.03	0.05	0.05	0.05	0.05	0.05
DCAW(1,2)	0.05	0.04	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.05
DCAW(2,1)	0.05	0.04	0.06	0.04	0.03	0.05	0.05	0.05	0.04	0.05
DCAW(2,2)	0.05	0.04	0.06	0.05	0.04	0.04	0.04	0.05	0.04	0.05
Set: $\mathbf{E}^C_{p,q}$. Test: Shapiro-Wilks.										
DCAW(0,0)	**0.83**	**0.29**	**0.28**	**0.30**	**0.80**	**0.26**	**0.27**	**0.76**	**0.25**	**0.71**
DCAW(0,1)	**0.25**	**0.08**	**0.10**	**0.08**	**0.24**	**0.07**	**0.07**	**0.20**	**0.07**	**0.20**
DCAW(1,1)	0.05	0.05	0.06	0.06	0.05	0.05	0.05	0.06	0.05	0.05
DCAW(1,2)	0.05	0.04	0.06	**0.07**	0.05	0.05	0.05	0.05	0.05	0.05
DCAW(2,1)	0.06	0.05	0.06	**0.07**	0.04	0.05	0.05	0.06	0.05	0.05
DCAW(2,2)	0.05	0.04	0.05	0.06	0.05	0.05	0.05	0.05	0.05	0.05
Set: $\mathbf{E}^C_{p,q}$. Test: Mean 0.										
DCAW(0,0)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DCAW(0,1)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DCAW(1,1)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DCAW(1,2)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DCAW(2,1)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DCAW(2,2)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Data are simulated from a DCAW(2,2) model, and then estimated for DCAW models of order (0,0), (0,1), (1,1), (2,1), (1,2) and (2,2). Rejection rates are reported for each test applied to each model fit. Rejection rates significantly larger than the nominal level of 5% are emphasized in bold.

Table A2. Rejection rates for the goodness-of-fit tests based on the component-wise partitions $\mathbf{E}^C_{p,q}$ and $\mathbf{V}^C_{p,q}$ in the case of $n = 4$.

	e_{11}	e_{12}	e_{13}	e_{14}	e_{22}	e_{23}	e_{24}	e_{33}	e_{34}	e_{44}
Set: $\mathbf{E}^C_{p,q}$. Test: Variance 1.										
DCAW(0,0)	**0.34**	**0.30**	**0.32**	**0.33**	**0.20**	**0.18**	**0.19**	**0.29**	**0.31**	**0.58**
DCAW(0,1)	**0.08**	**0.07**	**0.08**	**0.07**	**0.07**	0.05	**0.07**	**0.08**	**0.07**	**0.12**
DCAW(1,1)	0.04	0.03	0.03	0.03	0.04	0.03	0.04	0.05	0.03	0.04
DCAW(1,2)	0.04	0.03	0.03	0.03	0.04	0.03	0.04	0.05	0.02	0.04
DCAW(2,1)	0.04	0.03	0.03	0.03	0.04	0.03	0.04	0.05	0.03	0.04
DCAW(2,2)	0.04	0.03	0.03	0.03	0.04	0.03	0.04	0.05	0.02	0.04
Set: $\mathbf{E}^C_{p,q}$. Test: Ljung-Box.										
DCAW(0,0)	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
DCAW(0,1)	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
DCAW(1,1)	**0.32**	**0.34**	**0.33**	**0.31**	**0.27**	**0.32**	**0.28**	**0.27**	**0.26**	**0.23**
DCAW(1,2)	0.04	0.04	0.03	0.04	0.03	0.05	0.03	0.04	0.03	0.03
DCAW(2,1)	**0.33**	**0.33**	**0.33**	**0.31**	**0.28**	**0.30**	**0.27**	**0.27**	**0.26**	**0.24**
DCAW(2,2)	0.02	0.03	0.02	0.02	0.02	0.03	0.02	0.02	0.03	0.02
Set: $\mathbf{E}^C_{p,q}$. Test: Ljung-Box sq.										
DCAW(0,0)	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
DCAW(0,1)	**0.41**	**0.46**	**0.47**	**0.46**	**0.38**	**0.42**	**0.44**	**0.33**	**0.38**	**0.31**
DCAW(1,1)	0.06	**0.07**	0.05	0.06	0.05	0.05	0.05	0.05	0.05	0.05
DCAW(1,2)	**0.07**	**0.07**	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.05
DCAW(2,1)	0.06	**0.07**	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
DCAW(2,2)	**0.07**	**0.07**	0.05	0.05	0.05	0.04	0.05	0.06	0.05	0.05
Set: $\mathbf{V}^C_{p,q}$. Test: Ljung-Box.	v_{11}	v_{12}	v_{13}	v_{14}	v_{22}	v_{23}	v_{24}	v_{33}	v_{34}	v_{44}
DCAW(0,0)	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
DCAW(0,1)	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
DCAW(1,1)	**0.33**	**0.34**	**0.32**	**0.30**	**0.31**	**0.32**	**0.31**	**0.30**	**0.30**	**0.28**
DCAW(1,2)	0.03	0.04	0.04	0.04	0.03	0.05	0.04	0.03	0.03	0.03
DCAW(2,1)	**0.34**	**0.33**	**0.31**	**0.30**	**0.31**	**0.31**	**0.31**	**0.31**	**0.29**	**0.29**
DCAW(2,2)	0.01	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.02

Data are simulated from a DCAW(2,2) model, and then estimated for DCAW models of order (0,0), (0,1), (1,1), (2,1), (1,2) and (2,2). Rejection rates are reported for each test applied to each model fit. Rejection rates significantly larger than the nominal level of 5% are emphasized in bold.

Table A3. Rejection rates for the goodness-of-fit tests based on the component-wise partition $\mathbf{E}^{C}_{p,q}$ in the case of $n = 6$.

	e_{11}	e_{12}	e_{13}	e_{14}	e_{15}	e_{16}	e_{22}	e_{23}	e_{24}	e_{25}	e_{26}	e_{33}	e_{34}	e_{35}	e_{36}	e_{44}	e_{45}	e_{46}	e_{55}	e_{56}	e_{66}
Set: $\mathbf{E}^{C}_{p,q}$ Test: Anderson-Darling.																					
DCAW(0,0)	**0.77**	**0.18**	**0.20**	**0.21**	**0.20**	**0.20**	**0.71**	**0.17**	**0.17**	**0.18**	**0.18**	**0.67**	**0.18**	**0.17**	**0.16**	**0.66**	**0.15**	**0.14**	**0.61**	**0.14**	**0.57**
DCAW(0,1)	**0.19**	**0.08**	**0.07**	0.06	**0.07**	**0.07**	**0.26**	**0.07**	0.06	0.06	**0.09**	**0.19**	0.06	0.05	0.06	**0.17**	0.06	**0.07**	**0.15**	0.06	**0.14**
DCAW(1,1)	0.04	0.04	0.04	0.06	0.05	0.06	0.05	0.05	0.06	0.04	0.06	0.04	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.06	0.04
DCAW(1,2)	0.05	0.04	0.04	0.05	0.05	0.05	0.06	0.05	0.05	0.04	0.05	0.04	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.06	0.04
DCAW(2,1)	0.04	0.04	0.04	0.06	0.05	0.06	0.05	0.05	0.06	0.04	0.06	0.04	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05
DCAW(2,2)	0.05	0.03	0.04	0.06	0.05	0.05	0.06	0.05	0.06	0.04	0.06	0.04	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.04
Set: $\mathbf{E}^{C}_{p,q}$ Test: Lilliefors.																					
DCAW(0,0)	**0.61**	**0.12**	**0.12**	**0.13**	**0.13**	**0.12**	**0.55**	**0.12**	**0.10**	**0.10**	**0.12**	**0.51**	**0.10**	**0.12**	**0.11**	**0.47**	**0.08**	**0.11**	**0.43**	**0.09**	**0.40**
DCAW(0,1)	**0.16**	**0.07**	0.04	0.05	0.05	0.05	**0.21**	0.05	0.05	0.05	**0.07**	**0.17**	0.05	0.04	0.05	**0.12**	0.06	0.05	**0.12**	0.04	**0.08**
DCAW(1,1)	0.04	0.03	0.03	0.05	0.06	0.05	0.04	0.04	0.05	0.04	0.05	0.04	0.04	0.04	0.06	0.04	0.05	0.05	0.04	0.04	0.05
DCAW(1,2)	0.04	0.04	0.03	0.05	0.05	0.05	0.05	0.04	0.06	0.04	0.05	0.03	0.04	0.04	0.05	0.05	0.05	0.04	0.05	0.05	0.04
DCAW(2,1)	0.04	0.03	0.03	0.05	0.06	0.05	0.04	0.04	0.05	0.04	0.05	0.04	0.04	0.03	0.06	0.04	0.05	0.05	0.04	0.04	0.04
DCAW(2,2)	0.04	0.04	0.04	0.05	0.05	0.06	0.04	0.05	0.05	0.04	0.05	0.03	0.04	0.04	0.05	0.04	0.05	0.04	0.05	0.05	0.04
Set: $\mathbf{E}^{C}_{p,q}$ Test: Shapiro-Wilks.																					
DCAW(0,0)	**0.85**	**0.29**	**0.30**	**0.27**	**0.29**	**0.29**	**0.81**	**0.26**	**0.25**	**0.25**	**0.26**	**0.77**	**0.27**	**0.23**	**0.25**	**0.75**	**0.21**	**0.24**	**0.70**	**0.22**	**0.68**
DCAW(0,1)	**0.26**	**0.11**	**0.10**	**0.08**	**0.07**	**0.09**	**0.32**	**0.09**	**0.09**	**0.09**	**0.10**	**0.26**	**0.09**	**0.09**	**0.09**	**0.24**	**0.08**	**0.08**	**0.21**	**0.07**	**0.18**
DCAW(1,1)	0.05	0.05	0.04	0.06	0.05	0.06	0.05	0.05	0.06	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.06	0.04	0.05	0.06
DCAW(1,2)	0.04	0.04	0.05	0.06	0.05	0.05	0.05	0.04	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.04	0.05	0.06
DCAW(2,1)	0.04	0.05	0.04	0.06	0.05	0.05	0.05	0.05	0.06	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.06	0.04	0.05	**0.07**
DCAW(2,2)	0.04	0.04	0.05	0.06	0.04	0.05	0.05	0.04	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.06	0.05	0.05	0.06
Set: $\mathbf{E}^{C}_{p,q}$ Test: Mean 0.																					
DCAW(0,0)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.01	0.00	0.01	0.01	0.01	0.01	0.02	**0.17**
DCAW(0,1)	0.02	0.00	0.00	0.00	0.00	0.00	**0.12**	0.02	0.01	0.02	0.02	**0.07**	0.00	0.01	0.01	0.01	0.00	0.00	0.03	0.00	**0.08**
DCAW(1,1)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DCAW(1,2)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DCAW(2,1)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DCAW(2,2)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Data are simulated from a DCAW(2,2) model, and then estimated for DCAW models of order (0,0), (0,1), (1,1), (2,1), (1,2) and (2,2). Rejection rates are reported for each test applied to each model fit. Rejection rates significantly larger than the nominal level of 5% are emphasized in bold.

Table A4. Rejection rates for the goodness-of-fit tests based on the component-wise partitions $\mathbf{E}^C_{p,q}$ and $\mathbf{V}^C_{p,q}$ in the case of $n = 6$.

	e_{11}	e_{12}	e_{13}	e_{14}	e_{15}	e_{16}	e_{22}	e_{23}	e_{24}	e_{25}	e_{26}	e_{33}	e_{34}	e_{35}	e_{36}	e_{44}	e_{45}	e_{46}	e_{55}	e_{56}	e_{66}
Set: $\mathbf{E}^C_{p,q}$ Test: Variance 1.																					
DCAW(0,0)	**0.53**	**0.51**	**0.49**	**0.48**	**0.49**	**0.49**	**0.28**	**0.22**	**0.22**	**0.26**	**0.23**	**0.20**	**0.20**	**0.22**	**0.21**	**0.36**	**0.40**	**0.36**	**0.65**	**0.67**	**0.87**
DCAW(0,1)	**0.12**	**0.12**	**0.14**	**0.10**	**0.11**	**0.13**	**0.20**	**0.09**	**0.08**	**0.08**	**0.07**	**0.14**	**0.08**	**0.09**	**0.08**	**0.18**	**0.13**	**0.12**	**0.22**	**0.21**	**0.33**
DCAW(1,1)	0.04	0.05	0.04	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.03
DCAW(1,2)	0.04	0.05	0.04	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.05	0.03	0.04	0.04	0.04	0.04
DCAW(2,1)	0.04	0.05	0.04	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.03	0.04
DCAW(2,2)	0.04	0.05	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.05	0.03	0.04	0.04	0.04	0.04
Set: $\mathbf{E}^C_{p,q}$ Test: Ljung-Box.																					
DCAW(0,0)	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
DCAW(0,1)	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
DCAW(1,1)	**0.31**	**0.34**	**0.32**	**0.32**	**0.32**	**0.31**	**0.30**	**0.32**	**0.29**	**0.31**	**0.32**	**0.25**	**0.25**	**0.26**	**0.27**	**0.24**	**0.23**	**0.24**	**0.20**	**0.23**	**0.18**
DCAW(1,2)	0.04	0.04	0.05	0.05	0.05	0.04	0.03	**0.08**	0.05	0.05	0.06	0.03	0.06	0.05	0.04	0.04	0.05	0.04	0.05	0.05	0.04
DCAW(2,1)	**0.32**	**0.34**	**0.31**	**0.32**	**0.32**	**0.31**	**0.30**	**0.32**	**0.29**	**0.30**	**0.31**	**0.25**	**0.24**	**0.25**	**0.26**	**0.25**	**0.23**	**0.24**	**0.20**	**0.23**	**0.18**
DCAW(2,2)	0.02	0.03	0.03	0.03	0.03	0.02	0.02	0.05	0.03	0.03	0.03	0.02	0.04	0.03	0.04	0.03	0.04	0.04	0.03	0.04	0.03
Set: $\mathbf{E}^C_{p,q}$ Test: Ljung-Box sq.																					
DCAW(0,0)	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**0.99**
DCAW(0,1)	**0.48**	**0.52**	**0.51**	**0.49**	**0.50**	**0.51**	**0.46**	**0.50**	**0.47**	**0.50**	**0.51**	**0.43**	**0.43**	**0.45**	**0.46**	**0.38**	**0.37**	**0.40**	**0.34**	**0.36**	**0.29**
DCAW(1,1)	0.05	0.05	0.04	0.05	0.05	0.05	0.06	0.06	0.05	0.04	0.06	0.06	0.05	0.05	0.06	0.06	0.05	0.05	0.06	0.06	0.06
DCAW(1,2)	0.05	0.06	0.05	0.06	0.05	0.06	0.05	0.06	0.05	0.04	0.05	0.05	0.05	0.05	0.06	0.05	0.05	0.05	0.05	0.05	0.05
DCAW(2,1)	0.05	0.05	0.03	0.05	0.05	0.05	0.06	0.06	0.05	0.04	0.06	0.05	0.05	0.05	0.06	0.06	0.05	0.05	0.06	0.06	0.06
DCAW(2,2)	0.05	0.06	0.04	0.05	0.05	0.05	0.06	0.05	0.04	0.04	0.05	0.06	0.05	0.06	0.06	0.05	0.05	0.04	0.05	0.05	0.05
	v_{11}	v_{12}	v_{13}	v_{14}	v_{15}	v_{16}	v_{22}	v_{23}	v_{24}	v_{25}	v_{26}	v_{33}	v_{34}	v_{35}	v_{36}	v_{44}	v_{45}	v_{46}	v_{55}	v_{56}	v_{66}
Set: $\mathbf{V}^C_{p,q}$ Test: Ljung-Box.																					
DCAW(0,0)	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
DCAW(0,1)	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
DCAW(1,1)	**0.33**	**0.34**	**0.32**	**0.31**	**0.32**	**0.31**	**0.33**	**0.34**	**0.31**	**0.33**	**0.34**	**0.30**	**0.29**	**0.30**	**0.28**	**0.32**	**0.31**	**0.29**	**0.28**	**0.26**	**0.29**
DCAW(1,2)	0.04	0.04	0.04	0.04	0.05	0.04	0.03	0.06	0.04	0.05	0.06	0.04	0.05	0.04	0.04	0.03	0.05	0.04	0.03	0.04	0.04
DCAW(2,1)	**0.34**	**0.34**	**0.32**	**0.30**	**0.31**	**0.31**	**0.34**	**0.33**	**0.30**	**0.33**	**0.34**	**0.30**	**0.28**	**0.29**	**0.27**	**0.32**	**0.30**	**0.28**	**0.28**	**0.26**	**0.29**
DCAW(2,2)	0.02	0.03	0.03	0.03	0.03	0.02	0.02	0.04	0.03	0.03	0.05	0.02	0.03	0.04	0.03	0.02	0.04	0.04	0.02	0.03	0.03

Data are simulated from a DCAW(2,2) model, and then estimated for DCAW models of order (0,0), (0,1), (1,1), (2,1), (1,2) and (2,2). Rejection rates are reported for each test applied to each model fit. Rejection rates significantly larger than the nominal level of 5% are emphasized in bold.

Table A5. Rejection rates for the goodness-of-fit tests based on the column block partition $\mathbf{E}^B_{p,q}$ in the case of $n = 2$, 4, 6.

n = 2						
Test	DCAW(0,0)	DCAW(0,1)	DCAW(1,1)	DCAW(1,2)	DCAW(2,1)	DCAW(2,2)
Anderson-Darling	**0.07**	0.05	0.05	0.05	0.05	0.05
Lilliefors	**0.06**	0.05	0.05	0.05	0.05	0.05
Shapiro-Wilks	**0.07**	0.05	0.05	0.05	0.05	0.05
Mean 0	**0.40**	**0.19**	0.05	0.05	0.05	0.05
Variance 1	**0.22**	**0.08**	0.05	0.05	0.05	0.05
n = 4						
Test	DCAW(0,0)	DCAW(0,1)	DCAW(1,1)	DCAW(1,2)	DCAW(2,1)	DCAW(2,2)
Anderson-Darling	**0.09**	0.05	0.05	0.05	0.05	0.05
Lilliefors	**0.08**	0.05	0.05	0.05	0.05	0.05
Shapiro-Wilks	**0.10**	**0.06**	0.05	0.05	0.05	0.05
Mean 0	**0.39**	**0.19**	0.05	0.05	0.05	0.05
Variance 1	**0.29**	**0.10**	0.05	0.05	0.05	0.05
n = 6						
Test	DCAW(0,0)	DCAW(0,1)	DCAW(1,1)	DCAW(1,2)	DCAW(2,1)	DCAW(2,2)
Anderson-Darling	**0.11**	**0.06**	0.05	0.05	0.05	0.05
Lilliefors	**0.09**	0.05	0.05	0.05	0.05	0.05
Shapiro-Wilks	**0.13**	**0.07**	0.05	0.05	0.05	0.05
Mean 0	**0.38**	**0.18**	0.05	0.05	0.05	0.05
Variance 1	**0.33**	**0.12**	0.05	0.05	0.05	0.05

Data are simulated from a DCAW(2,2) model, and then estimated for DCAW models of order (0,0), (0,1), (1,1), (2,1), (1,2) and (2,2). Rejection rates are reported for each test applied to each model fit. Rejection rates significantly larger than the nominal level of 5% are emphasized in bold.

Table A6. Rejection rates for the goodness-of-fit tests based on the component-wise partitions $\mathbf{E}^C_{p,q}$ and $\mathbf{V}^C_{p,q}$ in the case of $n = 2$ and $d = 10$.

	e_{11}	e_{12}	e_{22}		e_{11}	e_{12}	e_{22}
Set: $\mathbf{E}^C_{p,q}$. Test: Anderson-Darling.				Set: $\mathbf{E}^C_{p,q}$. Test: Variance 1.			
DCAW(0,0)	**1.00**	**1.00**	**1.00**	DCAW(0,0)	**1.00**	**1.00**	**1.00**
DCAW(0,1)	**1.00**	**1.00**	**1.00**	DCAW(0,1)	**1.00**	**1.00**	**1.00**
DCAW(1,1)	**1.00**	**0.97**	**1.00**	DCAW(1,1)	**1.00**	**1.00**	**1.00**
DCAW(1,2)	**1.00**	**0.97**	**1.00**	DCAW(1,2)	**1.00**	**1.00**	**1.00**
DCAW(2,1)	**1.00**	**0.97**	**1.00**	DCAW(2,1)	**1.00**	**1.00**	**1.00**
DCAW(2,2)	**1.00**	**0.97**	**1.00**	DCAW(2,2)	**1.00**	**1.00**	**1.00**
Set: $\mathbf{E}^C_{p,q}$. Test: Lilliefors.				Set: $\mathbf{E}^C_{p,q}$. Test: Ljung-Box.			
DCAW(0,0)	**1.00**	**1.00**	**1.00**	DCAW(0,0)	**1.00**	**1.00**	**1.00**
DCAW(0,1)	**1.00**	**0.94**	**1.00**	DCAW(0,1)	**1.00**	**1.00**	**1.00**
DCAW(1,1)	**1.00**	**0.80**	**1.00**	DCAW(1,1)	**0.24**	**0.35**	**0.22**
DCAW(1,2)	**1.00**	**0.80**	**1.00**	DCAW(1,2)	0.01	**0.12**	0.02
DCAW(2,1)	**1.00**	**0.81**	**1.00**	DCAW(2,1)	**0.24**	**0.35**	**0.23**
DCAW(2,2)	**1.00**	**0.80**	**1.00**	DCAW(2,2)	0.01	**0.10**	0.01
Set: $\mathbf{E}^C_{p,q}$. Test: Shapiro-Wilks.				Set: $\mathbf{E}^C_{p,q}$. Test: Ljung-Box sq.			
DCAW(0,0)	**1.00**	**1.00**	**1.00**	DCAW(0,0)	**1.00**	**1.00**	**1.00**
DCAW(0,1)	**1.00**	**1.00**	**1.00**	DCAW(0,1)	**0.67**	**0.77**	**0.84**
DCAW(1.1)	**1.00**	**0.99**	**1.00**	DCAW(1,1)	**0.08**	0.06	**0.07**
DCAW(1,2)	**1.00**	**0.99**	**1.00**	DCAW(1,2)	**0.07**	0.05	0.06
DCAW(2,1)	**1.00**	**0.99**	**1.00**	DCAW(2,1)	**0.08**	0.06	**0.07**
DCAW(2,2)	**1.00**	**0.99**	**1.00**	DCAW(2,2)	**0.07**	0.06	0.06
Set: $\mathbf{E}^C_{p,q}$. Test: Mean 0.				Set: $\mathbf{V}^C_{p,q}$. Test: Ljung-Box.	v_{11}	v_{12}	v_{22}
DCAW(0,0)	**1.00**	0.05	**1.00**	DCAW(0,0)	**1.00**	**1.00**	**1.00**
DCAW(0,1)	**1.00**	0.00	**1.00**	DCAW(0,1)	**1.00**	**1.00**	**1.00**
DCAW(1,1)	**1.00**	0.00	**1.00**	DCAW(1,1)	**0.15**	**0.23**	**0.14**
DCAW(1,2)	**1.00**	0.00	**1.00**	DCAW(1,2)	0.02	**0.11**	0.01
DCAW(2,1)	**1.00**	0.00	**1.00**	DCAW(2,1)	**0.16**	**0.23**	**0.14**
DCAW(2,2)	**1.00**	0.00	**1.00**	DCAW(2,2)	0.01	**0.10**	0.01

Data are simulated from (22), and then estimated for DCAW models of order (0,0), (0,1), (1,1), (2,1), (1,2) and (2,2). Rejection rates are reported for each test applied to each model fit. Rejection rates significantly larger than the nominal level of 5% are emphasized in bold.

Table A7. Rejection rates for the goodness-of-fit tests based on the component-wise partitions $\mathbf{E}^C_{p,q}$ and $\mathbf{V}^C_{p,q}$ in the case of $n=2$ and $d=20$.

	e_{11}	e_{12}	e_{22}		e_{11}	e_{12}	e_{22}
Set: $\mathbf{E}^C_{p,q}$. Test: Anderson-Darling.				Set: $\mathbf{E}^C_{p,q}$. Test: Variance 1.			
DCAW(0,0)	**1.00**	**0.94**	**1.00**	DCAW(0,0)	**1.00**	**1.00**	**0.98**
DCAW(0,1)	**1.00**	**0.66**	**1.00**	DCAW(0,1)	**1.00**	**1.00**	**1.00**
DCAW(1,1)	**1.00**	**0.40**	**1.00**	DCAW(1,1)	**1.00**	**1.00**	**1.00**
DCAW(1,2)	**1.00**	**0.41**	**1.00**	DCAW(1,2)	**1.00**	**1.00**	**1.00**
DCAW(2,1)	**1.00**	**0.39**	**1.00**	DCAW(2,1)	**1.00**	**1.00**	**1.00**
DCAW(2,2)	**1.00**	**0.41**	**1.00**	DCAW(2,2)	**1.00**	**1.00**	**1.00**
Set: $\mathbf{E}^C_{p,q}$. Test: Lilliefors.				Set: $\mathbf{E}^C_{p,q}$. Test: Ljung-Box.			
DCAW(0,0)	**1.00**	**0.79**	**1.00**	DCAW(0,0)	**1.00**	**1.00**	**1.00**
DCAW(0,1)	**1.00**	**0.36**	**1.00**	DCAW(0,1)	**1.00**	**1.00**	**1.00**
DCAW(1,1)	**1.00**	**0.19**	**1.00**	DCAW(1,1)	**0.26**	**0.34**	**0.25**
DCAW(1,2)	**1.00**	**0.18**	**1.00**	DCAW(1,2)	0.02	**0.08**	0.02
DCAW(2,1)	**1.00**	**0.19**	**1.00**	DCAW(2,1)	**0.27**	**0.33**	**0.25**
DCAW(2,2)	**1.00**	**0.18**	**1.00**	DCAW(2,2)	0.01	0.06	0.01
Set: $\mathbf{E}^C_{p,q}$. Test: Shapiro-Wilks.				Set: $\mathbf{E}^C_{p,q}$. Test: Ljung-Box sq.			
DCAW(0,0)	**1.00**	**0.98**	**1.00**	DCAW(0,0)	**1.00**	**1.00**	**1.00**
DCAW(0,1)	**1.00**	**0.82**	**1.00**	DCAW(0,1)	**0.58**	**0.66**	**0.75**
DCAW(1,1)	**1.00**	**0.60**	**1.00**	DCAW(1,1)	0.06	0.06	**0.07**
DCAW(1,2)	**1.00**	**0.60**	**1.00**	DCAW(1,2)	0.04	0.05	0.06
DCAW(2,1)	**1.00**	**0.60**	**1.00**	DCAW(2,1)	0.06	0.06	**0.07**
DCAW(2,2)	**1.00**	**0.60**	**1.00**	DCAW(2,2)	0.05	0.05	0.06
Set: $\mathbf{E}^C_{p,q}$. Test: Mean 0.				Set: $\mathbf{V}^C_{p,q}$. Test: Ljung-Box.	v_{11}	v_{12}	v_{22}
DCAW(0,0)	**0.30**	0.00	**0.97**	DCAW(0,0)	**1.00**	**1.00**	**1.00**
DCAW(0,1)	0.02	0.00	**0.98**	DCAW(0,1)	**1.00**	**1.00**	**1.00**
DCAW(1,1)	0.00	0.00	**0.95**	DCAW(1,1)	**0.23**	**0.27**	**0.22**
DCAW(1,2)	0.00	0.00	**0.93**	DCAW(1,2)	0.01	**0.08**	0.02
DCAW(2,1)	0.00	0.00	**0.95**	DCAW(2,1)	**0.23**	**0.27**	**0.22**
DCAW(2,2)	0.00	0.00	**0.93**	DCAW(2,2)	0.01	0.06	0.01

Data are simulated from (22), and then estimated for DCAW models of order (0,0), (0,1), (1,1), (2,1), (1,2) and (2,2). Rejection rates are reported for each test applied to each model fit. Rejection rates significantly larger than the nominal level of 5% are emphasized in bold.

Table A8. Rejection rates for the goodness-of-fit tests based on the column block partition $\mathbf{E}^B_{p,q}$ for the block size 20, $n=2$, and $d=10$, 20, 30.

d = 10						
Test	DCAW(0,0)	DCAW(0,1)	DCAW(1,1)	DCAW(1,2)	DCAW(2,1)	DCAW(2,2)
Anderson-Darling	**0.31**	**0.33**	**0.36**	**0.36**	**0.36**	**0.36**
Lilliefors	**0.24**	**0.26**	**0.28**	**0.28**	**0.28**	**0.28**
Shapiro-Wilks	**0.34**	**0.39**	**0.41**	**0.41**	**0.41**	**0.42**
Mean 0	**0.52**	**0.30**	**0.12**	**0.12**	**0.12**	**0.12**
Variance 1	**0.52**	**0.33**	**0.25**	**0.25**	**0.25**	**0.25**
d = 20						
Test	DCAW(0,0)	DCAW(0,1)	DCAW(1,1)	DCAW(1,2)	DCAW(2,1)	DCAW(2,2)
Anderson-Darling	**0.16**	**0.16**	**0.16**	**0.16**	**0.16**	**0.16**
Lilliefors	**0.13**	**0.12**	**0.12**	**0.12**	**0.12**	**0.12**
Shapiro-Wilks	**0.19**	**0.20**	**0.21**	**0.21**	**0.21**	**0.21**
Mean 0	**0.49**	**0.27**	**0.10**	**0.10**	**0.10**	**0.10**
Variance 1	**0.38**	**0.20**	**0.14**	**0.14**	**0.14**	**0.14**
d = 30						
Test	DCAW(0,0)	DCAW(0,1)	DCAW(1,1)	DCAW(1,2)	DCAW(2,1)	DCAW(2,2)
Anderson-Darling	**0.12**	**0.11**	**0.11**	**0.11**	**0.11**	**0.11**
Lilliefors	**0.10**	**0.09**	**0.09**	**0.09**	**0.09**	**0.09**
Shapiro-Wilks	**0.14**	**0.14**	**0.14**	**0.14**	**0.14**	**0.14**
Mean 0	**0.48**	**0.26**	**0.09**	**0.09**	**0.09**	**0.09**
Variance 1	**0.32**	**0.16**	**0.10**	**0.10**	**0.10**	**0.12**

Data are simulated from (22), and then estimated for DCAW models of order (0,0), (0,1), (1,1), (2,1), (1,2) and (2,2). Rejection rates are reported for each test applied to each model fit. Rejection rates significantly larger than the nominal level of 5% are emphasized in bold.

Table A9. The p-values of the proposed goodness-of-fit tests on the validity of CAW(p,q) models with $p, q \leq 2$ based on the component-wise partition $\mathbf{E}^C_{p,q}$.

	e_{11}	e_{12}	e_{13}	e_{14}	e_{15}	e_{16}	e_{22}	e_{23}	e_{24}	e_{25}	e_{26}	e_{33}	e_{34}	e_{35}	e_{36}	e_{44}	e_{45}	e_{46}	e_{55}	e_{56}	e_{66}
Set: $\mathbf{E}^C_{p,q}$ Test: Anderson-Darling.																					
CAW(0,0)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(0,1)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**	**0.00**	**0.01**	**0.00**
CAW(1,1)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.05**	**0.00**	**0.00**	**0.00**
CAW(1,2)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.02**	**0.00**	**0.01**	**0.00**
CAW(2,1)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**	**0.00**	**0.01**	**0.00**
CAW(2,2)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.03**	**0.00**	**0.01**	**0.00**
Set: $\mathbf{E}^C_{p,q}$ Test: Lilliefors.																					
CAW(0,0)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(0,1)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.10	**0.01**	**0.00**	**0.00**	**0.01**	**0.01**	**0.00**	**0.00**	**0.00**	0.18	**0.00**	0.12	**0.00**
CAW(1,1)	**0.00**	**0.00**	**0.00**	0.07	**0.02**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.03**	0.06	**0.00**	**0.00**	**0.01**	0.15	**0.00**	0.10	**0.00**
CAW(1,2)	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.04**	0.08	**0.00**	**0.00**	**0.05**	0.14	**0.00**	0.19	**0.00**
CAW(2,1)	**0.00**	**0.00**	**0.00**	**0.01**	**0.02**	**0.00**	**0.00**	**0.00**	**0.01**	**0.00**	**0.00**	**0.00**	**0.02**	**0.05**	**0.01**	**0.00**	0.06	0.11	**0.00**	0.06	**0.00**
CAW(2,2)	**0.00**	**0.00**	**0.00**	**0.03**	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**	**0.00**	**0.00**	**0.00**	0.08	0.13	**0.00**	**0.00**	**0.02**	0.10	**0.00**	0.12	**0.00**
Set: $\mathbf{E}^C_{p,q}$ Test: Shapiro-Wilks.																					
CAW(0,0)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(0,1)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(1,1)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(1,2)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(2,1)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(2,2)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Set: $\mathbf{E}^C_{p,q}$ Test: Mean 0.																					
CAW(0,0)	**0.00**	**0.04**	**0.01**	**0.05**	**0.00**	**0.03**	**0.00**	0.10	0.39	**0.00**	**0.00**	0.32	**0.00**	**0.00**	0.24	**0.00**	0.20	**0.00**	**0.00**	0.09	**0.00**
CAW(0,1)	**0.00**	0.51	0.43	0.78	0.66	0.93	**0.00**	**0.01**	0.20	0.44	**0.00**	0.28	0.16	0.06	0.13	**0.00**	0.25	0.39	**0.00**	0.90	**0.00**
CAW(1,1)	**0.00**	0.69	0.70	0.72	0.67	0.98	**0.00**	**0.04**	0.42	0.18	**0.00**	0.54	0.18	0.11	0.11	**0.00**	0.23	0.32	**0.00**	0.49	**0.00**
CAW(1,2)	**0.00**	0.48	0.62	0.67	0.46	0.94	**0.00**	**0.03**	0.31	0.25	**0.00**	0.43	0.27	0.07	0.08	**0.00**	0.30	0.41	**0.00**	0.58	**0.00**
CAW(2,1)	**0.00**	0.34	0.69	0.58	0.56	0.79	**0.00**	**0.03**	0.54	0.21	**0.00**	0.41	0.27	0.06	0.12	**0.00**	0.30	0.26	**0.00**	0.56	**0.00**
CAW(2,2)	**0.00**	0.40	0.57	0.50	0.45	0.78	**0.00**	**0.04**	0.34	0.23	**0.00**	0.47	0.17	0.10	0.10	**0.00**	0.28	0.31	**0.00**	0.62	**0.00**

The data consist of the realized covariance matrices calculated for the six stocks traded on the New York Stock Exchange: American Express Inc. (AXP), Citigroup (C), General Electric (GE), Home Depot Inc. (HD), International Business Machines (IBM) and JPMorgan Chase & Co. (JPM) for the period from the 1st of January, 2000 to the 30th of July, 2008. The p-values smaller than or equal to 0.05 are emphasized in bold.

Table A10. The p-values of the proposed goodness-of-fit tests on the validity of CAW(p,q) models with $p, q \leq 2$ based on the component-wise partitions $\mathbf{E}^C_{p,q}$ and $\mathbf{V}^C_{p,q}$.

	e11	*e12*	*e13*	*e14*	*e15*	*e16*	*e22*	*e23*	*e24*	*e25*	*e26*	*e33*	*e34*	*e35*	*e36*	*e44*	*e45*	*e46*	*e55*	*e56*	*e66*
Set: $\mathbf{E}^C_{p,q}$ Test: Variance 1.																					
CAW(0,0)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(0,1)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(1,1)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(1,2)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(2,1)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(2,2)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Set: $\mathbf{E}^C_{p,q}$ Test: Ljung-Box.																					
CAW(0,0)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(0,1)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(1,1)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(1,2)	**0.05**	**0.01**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(2,1)	0.45	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(2,2)	0.17	**0.04**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Set: $\mathbf{E}^C_{p,q}$ Test: Ljung-Box sq.																					
CAW(0,0)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(0,1)	**0.00**	**0.03**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.09	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.02**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(1,1)	**0.00**	0.30	0.93	0.52	0.77	0.05	**0.00**	0.65	0.06	0.95	0.07	**0.00**	0.35	**0.00**	**0.02**	0.78	0.24	0.33	0.14	0.89	0.07
CAW(1,2)	**0.00**	0.19	0.85	0.58	0.56	0.29	**0.00**	0.75	0.27	0.95	**0.02**	**0.01**	0.24	**0.00**	**0.01**	0.85	0.02	0.31	0.11	0.91	**0.02**
CAW(2,1)	**0.00**	0.39	0.68	0.75	0.56	0.08	**0.00**	0.86	0.09	0.97	**0.02**	**0.00**	0.39	**0.00**	**0.00**	0.76	**0.00**	0.15	0.23	**0.05**	**0.02**
CAW(2,2)	**0.00**	0.34	0.80	0.48	0.57	0.53	**0.00**	0.90	0.23	0.98	**0.04**	**0.01**	0.28	**0.00**	**0.00**	0.77	**0.00**	0.19	0.23	0.82	**0.02**
	v_{11}	v_{12}	v_{13}	v_{14}	v_{15}	v_{16}	v_{22}	v_{23}	v_{24}	v_{25}	v_{26}	v_{33}	v_{34}	v_{35}	v_{36}	v_{44}	v_{45}	v_{46}	v_{55}	v_{56}	v_{66}
Set: $\mathbf{V}^C_{p,q}$ Test: Ljung-Box.																					
CAW(0,0)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(0,1)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(1,1)	0.10	0.03	**0.00**	**0.00**	0.22	0.37	**0.00**	**0.00**	**0.01**	**0.01**	**0.05**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
CAW(1,2)	0.28	0.57	**0.00**	**0.00**	0.56	0.42	**0.00**	**0.00**	**0.00**	**0.05**	**0.04**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**	**0.00**	**0.00**	**0.02**	**0.00**
CAW(2,1)	0.61	**0.05**	**0.00**	**0.00**	**0.02**	0.13	0.08	**0.00**	**0.00**	0.23	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.02**	0.12	0.11	**0.00**	0.11	**0.00**
CAW(2,2)	0.48	0.22	**0.00**	**0.00**	0.27	0.50	**0.01**	**0.00**	**0.00**	0.21	0.13	**0.00**	**0.00**	**0.01**	**0.00**	**0.01**	**0.01**	0.22	**0.00**	**0.04**	**0.00**

The data consist of the realized covariance matrices calculated for the six stocks traded on the New York Stock Exchange: American Express Inc. (AXP), Citigroup (C), General Electric (GE), Home Depot Inc. (HD), International Business Machines (IBM) and JPMorgan Chase & Co. (JPM) for the period from the 1st of January, 2000 to the 30th of July, 2008. The p-values smaller than or equal to 0.05 are emphasized in bold.

Table A11. Average rejection rates of the proposed goodness-of-fit tests on the validity of CAW(p,q) models with $p, q \leq 2$ based on the column block partition $\mathbf{E}^{B}_{p,q}$ with the block size equal to 20.

Test	CAW(0,0)	CAW(0,1)	CAW(1,1)	CAW(1,2)	CAW(2,1)	CAW(2,2)
Anderson-Darling	**1.00**	**0.97**	**0.98**	**0.98**	**0.97**	**0.97**
Lilliefors	**1.00**	**0.94**	**0.95**	**0.96**	**0.96**	**0.95**
Shapiro-Wilks	**1.00**	**0.99**	**0.98**	**0.98**	**1.00**	**0.98**
Mean 0	**0.95**	**0.63**	**0.30**	**0.31**	**0.32**	**0.34**
Variance 1	**0.93**	**0.72**	**0.70**	**0.69**	**0.72**	**0.69**
Multivariate Ljung-Box	**0.95**	**0.52**	**0.55**	**0.59**	**0.56**	**0.58**
Multivariate Ljung-Box sq.	**0.86**	**0.34**	**0.37**	**0.36**	**0.36**	**0.38**

The data consist of the realized covariance matrices calculated for the six stocks traded on the New York Stock Exchange: American Express Inc. (AXP), Citigroup (C), General Electric (GE), Home Depot Inc. (HD), International Business Machines (IBM) and JPMorgan Chase & Co. (JPM) for the period from the 1st of January, 2000 to the 30th of July, 2008. Rejection rates significantly larger than 5% is emphazised in bold.