

# Dynamic Principal Component CAW Models for High-Dimensional Realized Covariance Matrices

**Bastian Gribisch\***

*Institute of Econometrics and Statistics, University of Cologne, Germany*

**Michael Stollenwerk**

*Institute of Econometrics and Statistics, University of Cologne, Germany*

(February 10, 2017)

## Abstract

We propose a new dynamic principal component CAW model (DPC-CAW) for time-series of high-dimensional realized covariance matrices of asset returns. The model performs a spectral decomposition of the scale matrix of a central Wishart distribution and assumes independent dynamics for the principal components' variances and the eigenvector processes. A three-step estimation procedure makes the model applicable to high-dimensional covariance matrices. We analyze the finite sample properties of the estimation approach and provide an empirical application to realized covariance matrices for 100 assets. The DPC-CAW model has particularly good forecasting properties and outperforms its competitors for realized covariance matrices.

*JEL classification:* C32, C58, G17;

*Keywords:* Realized volatility, Covariance matrix, Spectral Decomposition, Time-Series Models.

---

\*Corresponding Author. Institute of Econometrics and Statistics, University of Cologne, Universitätsstr. 22a, D-50937 Cologne, Germany. Tel.: +49(0)2214707711; Fax: +49(0)2214705074. *E-mail address:* bastian.gribisch@statistik.uni-koeln.de

# 1 Introduction

The modeling and forecasting of covariance matrices of asset returns is central to financial decision making since it provides a measurement of the risk involved in different investment allocations. It is specifically used in option pricing, risk management and portfolio allocation.

Traditionally multivariate GARCH (MGARCH) or multivariate stochastic volatility (MSV) models have been applied in order to estimate conditional covariance matrices from daily asset return vectors (see e.g. Bauwens et al., 2006, and Asai et al., 2006, for surveys). Nowadays the increasing availability of intraday asset return information enables the computation of consistent ex-post measures of daily (co)variation of asset prices, so-called realized (co)variances (see e.g. Andersen et al., 2003, and Barndorff-Nielsen and Shephard, 2004). These realized measures can then be modeled directly in order to obtain forecasts of the covariance matrix of asset returns. The literature provides broad evidence that models for realized covariance matrices provide more precise forecasts than MGARCH and MSV models (see e.g. Golosnoy et al., 2012, and the references therein). Pioneering approaches are found in Gouriéroux et al. (2009), Chiriac and Voev (2011), Bauer and Vorkink (2011), Noureldin et al. (2012) and Golosnoy et al. (2012).

The existing models have in common that applications to high-dimensional covariance matrices are complicated if not impossible and the empirical applications therefore do not exceed the 10-dimensional case. Realistic portfolios however typically consist of hundreds of assets which makes high-dimensional covariance matrix forecasting an important field of research. The development of models for high-dimensional applications is challenging, since the dimension of the object of interest is proportional to the square of the number of assets. This results in a huge number of model parameters and renders one-step maximum likelihood (ML) estimation virtually impossible (the so-called *curse of dimensionality*). An important task is therefore the development of multivariate volatility models which allow for feasible estimation approaches for high-dimensional applications.

Contributions on high-dimensional realized covariance modeling are sparse. Recently Bauwens et al. (2012) proposed the Realized DCC (Re-DCC) CAW model (see also Bauwens et al., 2014, and Bauwens et al., 2016, for applications and extensions), which resembles the DCC GARCH idea of Engle (2002) under the Conditional Autoregressive Wishart (CAW) setting of Golosnoy et al. (2012)

for realized covariance matrices. The model is applicable in high-dimensional settings via employing a three-step estimation procedure with correlation targeting, similar to the corresponding MGARCH model. Bauwens et al. (2012) provide an empirical application for 50 assets.

While the DCC idea builds on decomposing the conditional covariance matrix in variances and correlations, which are then estimated independently, an alternative strand of literature constructs orthogonal components via a spectral decomposition (SD) of the covariance matrix. The most prominent model here is the orthogonal GARCH (OGARCH) model of Alexander and Chibumba (1997) and Alexander (2001), where the estimation output can be readily interpreted in terms of (conditional) principal component analysis. Aielli and Caporin (2015) introduce additional flexibility via allowing for dynamic loading matrices. The resulting model is then called Dynamic Principal Component (DPC) GARCH model. Similar to the DCC approach, the framework assumes the presence of an auxiliary process generating orthonormal dynamic eigenvectors and allows for three-step estimation in order to be applicable in high-dimensional settings (the authors provide an application for up to 30 assets). This flexibility does not come without a cost: similar to the three-step approach for fitting DCC GARCH models, also the DPC three-step estimator suffers from inconsistency problems which are due to model misspecification and inconsistent targeting within the estimation steps.

In this paper we adapt the DPC-GARCH model of Aielli and Caporin (2015) to the modeling of high-dimensional realized covariance matrices. The model structure is based on the CAW framework of Golosnoy et al. (2012) assuming a conditional central Wishart distribution for the realized covariance matrix. This particular distributional assumption allows for a convenient Quasi Maximum Likelihood (QML) interpretation implying consistency of one-step estimation even if the Wishart assumption is violated. We focus on high-dimensional applications and present a scalar version of the resulting DPC-CAW model and its estimation via a three-step approach similar to Aielli and Caporin (2015). This focus builds on the common motivation of the DCC- and the DPC approaches: the simplification of parameter estimation in large-dimensional environments. The three-step approach suffers from similar inconsistency problems as the DCC GARCH, the Re-DCC CAW and the DPC-GARCH model. We therefore conduct an extensive simulation experiment which shows that biases are present but mainly concern the unconditional variances of lower order principal components

which are of minor relevance for covariance forecasting. An out-of-sample forecasting experiment for 100-dimensional realized covariance matrices finally shows that the DPC-CAW model has very good forecasting properties and outperforms its competitors including the Re-DCC approach of Bauwens et al. (2012). This finding is explained by the fact that the DPC model assumes independent variance dynamics for the component variances rather than the return variances. The results show that the DPC modeling approach induces more flexible covariance and correlation dynamics compared to the DCC model, which directly restricts the dynamics of the correlation process rather than the loading process.

The rest of the paper is organized as follows. In section 2 we briefly review the concept of realized covariance measures. Section 3 introduces the scalar DPC-CAW model including one-step and three-step ML estimation. Section 4 presents the results of a simulation experiment analyzing the bias and consistency of estimates obtained via the three-step approach. The empirical application to realized covariance matrices for 100 NYSE traded stocks including in-sample diagnostics and an extensive out-of-sample forecasting experiment is presented in Section 5. Section 6 concludes.

## 2 Realized Covariance Measures

Consider an  $n$ -dimensional vector of log-prices  $y(\tau)$ , where  $\tau \in \mathbb{R}_+$  represents continuous time. Assume that  $y(\tau)$  is a Brownian stochastic semimartingale with  $(n \times n)$  spot covariance matrix  $\Theta(\tau)$ . Without loss of generality restricting the trading day to the unit interval we obtain the 'true' integrated covariance matrix at day  $t$  as  $\Sigma_t = \int_{t-1}^t \Theta(\tau) d\tau$ .

Now assume that we observe  $m + 1$  uniformly spaced intraday log-prices. Then the  $j$ 'th intraday return vector on day  $t$  ( $t = 1, \dots, T$ ) is given by

$$r_{j,t} = y((t-1) + j/m) - y((t-1) + (j-1)/m), \quad j = 1, \dots, m, \quad t = 1, \dots, T. \quad (1)$$

Let the  $(n \times n)$  matrix  $R_t$  denote a realized measure, i.e. a nonparametric ex-post estimate of  $\Sigma_t$  exploiting high-frequency asset return information. A well-known example for  $R_t$  is the realized

covariance matrix, which is defined as

$$RC_t = \sum_{j=1}^m r_{j,t} r'_{j,t}. \quad (2)$$

In the absence of market microstructure noise and discontinuous price jumps it can be shown that  $RC_t$  is a consistent estimator of  $\Sigma_t$  as  $m \rightarrow \infty$  (see Barndorff-Nielsen and Shephard, 2004). If the observed intraday price data contains microstructure noise, jumps or non-synchronous trading one can employ one of several alternatives to the realized covariance matrix, such as the multivariate realized kernel of Barndorff-Nielsen et al. (2011).

### 3 The DPC-CAW Model

We model the time-evolution of  $n$ -dimensional stochastic positive-definite realized covariance measures  $\{R_t\}_{t=1}^T$ . Given the filtration  $\mathcal{F}_{t-1} = \{R_{t-1}, R_{t-2}, \dots\}$ ,  $R_t$  is assumed to follow a central Wishart distribution

$$R_t | \mathcal{F}_{t-1} \sim \mathcal{W}_n(\nu, S_t/\nu), \quad (3)$$

where  $\nu > n$  is the scalar degree of freedom, and  $S_t/\nu$  denotes the symmetric, positive definite  $n \times n$  scale matrix, such that

$$E[R_t | \mathcal{F}_{t-1}] = S_t. \quad (4)$$

We now adapt the covariance dynamics of the DPC-GARCH model of Aielli and Caporin (2015) to the direct modeling of realized covariance measures. Let

$$S_t = L_t D_t L_t' \quad (5)$$

denote the SD of the conditional mean of  $R_t$ , where the diagonal elements of  $D_t = \text{diag}(d_{1,t}, d_{2,t}, \dots, d_{n,t})$  are the eigenvalues of  $S_t$  and the columns of  $L_t$  are the associated orthonormal eigenvectors. We are interested in building a forecasting model for  $R_t$  where both the eigenvalues and the eigenvectors are allowed to vary persistently over time and which allows for convenient sequential estimation in

high-dimensional applications.

### 3.1 Eigenvector Driving Process

In order to obtain time-varying orthonormal eigenvectors  $L_t$ , we introduce a matrix-variate auxiliary process from which the loadings are obtained via computing the conditional SD. The auxiliary process is defined as a scalar BEKK recursion for realized covariance measures. Let

$$Q_t = (1 - a - b) S + a R_{t-1} + b Q_{t-1}, \tag{6}$$

$$Q_t = L_t G_t L_t'. \tag{7}$$

The scalars  $a$  and  $b$  and the intercept matrix  $S$  are parameters to be estimated.

We consider model (6)-(7) as the true data generating process (DGP) for the loading matrices. Note that we may generalize the scalar dynamics of Eq. (6) to full BEKK dynamics (see Aielli and Caporin, 2015, and Noureldin et al., 2014, for details). However, estimation of such a 'complete model' in high-dimensional settings is practically impossible since the number of autoregressive parameters is of order  $\mathcal{O}(n^2)$ . We therefore restrict the model to feasible scalar dynamics similar to the popular DCC-GARCH approach.

The spectral decomposition in Eq. (7) is not uniquely identified. Following Aielli and Caporin (2015) we therefore impose on all spectral decompositions within the model except Eq. (5)

**Assumption 1.** *The eigenvalues in a spectral decomposition are arranged in strictly decreasing order.*

The sign of each eigenvector is still unidentified. However within the model the eigenvector matrix appears only in quadratic form. Hence there is no need for imposing a sign restriction. The implicit assumption that the eigenvalues of  $Q_t$  are distinct holds almost surely and is thus mild.

In order to ensure that  $Q_t$  is always positive definite we furthermore impose

**Assumption 2.**  $0 \leq a, 0 \leq b, a + b < 1$ ;  $S$  and  $Q_0$  are positive definite.

We require an additional constraint on  $S$  in order to ensure a unique sequence of loadings. To see this intuitively, multiply Eq. (6) by some positive constant  $c$ . Given the data  $\{R_t\}_{t=1}^T$  and

provided assumption 2 still holds, this would produce the same eigenvector matrix series  $\{L_t\}_{t=1}^T$ .<sup>1</sup> Identification can be ensured by restricting the magnitude of the intercept matrix  $S$ . This is done implicitly in Assumption 3 of the eigenvalue driving process to be presented below.

### 3.2 Eigenvalue Driving Process

We employ  $n$  independent GARCH-type recursions in order to capture the dynamics of the diagonal elements of  $D_t$  in Eq. (5). Let

$$d_{i,t} = (1 - \alpha_i - \beta_i)\gamma_i + \alpha_i g_{i,t-1} + \beta_i d_{i,t-1}, \quad (8)$$

where  $g_{i,t} = e_i' L_t' R_t L_t e_i$  with  $e_i$  being an  $n \times 1$  vector of zeros with a 1 at the  $i$ 'th position. That is,  $g_{i,t}$  is the  $i$ 'th diagonal element of the random matrix  $L_t' R_t L_t$ . Generalizations of model (8) obtained by increasing the lag order or e.g. including HAR-type dynamics (see Corsi, 2009) are straightforward to implement.

Note that

$$E[L_t' R_t L_t | \mathcal{F}_{t-1}] = L_t' E[R_t | \mathcal{F}_{t-1}] L_t = L_t' L_t D_t L_t' L_t = D_t, \quad (9)$$

such that

$$E[g_{i,t} | \mathcal{F}_{t-1}] = E[e_i' L_t' R_t L_t e_i | \mathcal{F}_{t-1}] = e_i' E[L_t' R_t L_t | \mathcal{F}_{t-1}] e_i = e_i' D_t e_i = d_{i,t}. \quad (10)$$

Under the usual stationarity condition we then obtain

$$E[d_{i,t}] = \gamma_i. \quad (11)$$

We now employ the SD of the intercept matrix  $S = LDL'$ , where  $D = \text{diag}(\{d_i\}_{i=1}^n)$ , and set

$$\gamma_i = d_i, \quad i = 1, \dots, n. \quad (12)$$

---

<sup>1</sup>Note that  $cQ_t = cL_t G_t L_t' = L_t cG_t L_t' = L_t \tilde{G}_t L_t'$ .

That is  $\{d_i\}_{i=1}^n$  are the eigenvalues of the intercept matrix in the eigenvector driving recursion (see Eq. 6). The following assumption formalizes this idea and imposes stationarity conditions.

**Assumption 3.**  $\gamma_i = d_i$ ,  $0 \leq \alpha_i$ ,  $0 \leq \beta_i$ ,  $\alpha_i + \beta_i < 1$ ,  $0 < d_{i,0} \forall i$ .

Since all parameters are restricted to be positive this assumption also ensures that  $d_{i,t}$  is always positive and consequently  $S_t$  is always positive definite.

The targeting-like constraint of setting  $\gamma_i = d_i$  solves the problem of identifying a unique  $Q_t$  sequence since it implicitly imposes

$$\text{tr}(E[R_t]) = \text{tr}(S), \quad (13)$$

thus restricting the magnitude of the intercept matrix  $S$ . To see this consider

$$E[R_t] = E[E[R_t|\mathcal{F}_{t-1}]] = E[S_t], \quad (14)$$

such that together with the trace property  $\text{tr}(ABC) = \text{tr}(CAB) = \text{tr}(BCA)$  and orthonormality of  $L_t$  and  $L$  it holds that

$$\begin{aligned} \text{tr}(E[R_t]) &= E[\text{tr}(S_t)] = E[\text{tr}(L_t D_t L_t')] = E[\text{tr}(D_t L_t' L_t)] = E[\text{tr}(D_t)] \\ &= \text{tr}(D) = \text{tr}(DL'L) = \text{tr}(LDL') = \text{tr}(S). \end{aligned} \quad (15)$$

While this is not the only way to achieve identification of the eigenvector driving process it does entail an appealing interpretation of the model. Specifically if  $a = b = 0$ , the eigenvector driving process collapses to the constant matrix  $Q_t = S$ , such that  $L_t = L$ . The resulting specification resembles the popular orthogonal GARCH model of Alexander and Chibumba (1997) and Alexander (2001), such that the DPC-CAW is regarded as being a dynamic generalization of the OGARCH to the modeling of realized covariance measures.

Recall that according to Assumption 1 the diagonal elements of  $D$  are arranged in decreasing order, which implies that

$$E[d_{1,t}] > E[d_{2,t}] > \dots > E[d_{n,t}]. \quad (16)$$

This however does not imply that individual elements of  $d_t$  themselves are arranged in decreasing

order, since the random variables  $g_{i,t}$  are not bounded above. A situation where  $d_{i,t} < d_{i-1,t}$  happens particularly often in high dimensional applications where the elements of  $d$  are close to each other.

The conditional Wishart assumption for  $R_t$  in Eq. (3) results in a conditional Gamma distribution for  $g_{i,t}$ . Consider the following theorem of Rao (1965):

**Theorem 1.** *If an  $n \times n$  random matrix  $R$  has a central Wishart distribution with  $\nu$  degrees of freedom and scale matrix  $S$ , that is  $R \sim \mathcal{W}_n(S, \nu)$ , and  $X$  is a  $q \times n$  matrix of rank  $q$ , then:*

$$XRX' \sim \mathcal{W}_q(\nu, XSX').$$

Set  $X = e_i' L_t'$ , where  $e_i$  is defined as above,  $S = S_t/\nu$  and  $R = R_t$  to obtain  $XRX' = e_i' L_t' R_t L_t e_i = g_{i,t}$  and  $XSX' = e_i' L_t' \frac{S_t}{\nu} L_t e_i = \frac{1}{\nu} e_i' L_t' L_t D_t L_t L_t e_i = \frac{1}{\nu} e_i' D_t e_i = \frac{d_{i,t}}{\nu}$  such that

$$g_{i,t} | \mathcal{F}_{t-1} \sim \mathcal{W}_1\left(\nu, \frac{d_{i,t}}{\nu}\right). \quad (17)$$

Since the univariate Wishart resembles the Gamma density,  $g_{i,t}$  follows a conditional gamma distribution with shape parameter  $\nu/2$  and scale parameter  $2d_{i,t}/\nu$ :

$$g_{i,t} | \mathcal{F}_{t-1} \sim \text{Gamma}(\nu/2, 2d_{i,t}/\nu). \quad (18)$$

Equations (3) - (8) together with the Assumptions 1-3 constitute the scalar DPC-CAW model.

### 3.3 Maximum Likelihood Estimation

#### 3.3.1 One-Step Estimation

Low-dimensional applications (for, say, up to ten assets) allow for one-step estimation of the model parameters  $\theta = (\text{vech}(S)', a, b, \{\alpha_i, \beta_i\}_{i=1}^n, \nu)'$  of the DPC-CAW model. Estimation can then be

carried out by maximizing the log-likelihood function

$$\begin{aligned} \mathcal{L}(\theta) = & \sum_{t=1}^T \left[ \frac{n\nu}{2} \ln\left(\frac{\nu}{2}\right) - \frac{n(n-1)}{4} \ln(\pi) - \sum_{i=1}^n \ln \Gamma\left(\frac{\nu+1-i}{2}\right) + \left(\frac{\nu-n-1}{2}\right) \ln |R_t| \right. \\ & \left. - \frac{\nu}{2} \left[ \ln |S_t(\psi)| + \text{tr}(S_t(\psi)^{-1}R_t) \right] \right], \end{aligned} \quad (19)$$

where  $\psi$  summarizes the parameters for the  $Q_t$  and  $d_{i,t}$  recursions, such that  $\theta = (\psi, \nu)'$ . The parameter  $\nu$  can be treated as a nuisance parameter due to its irrelevance for the matrix forecast (see Eq. 4). In fact the first order conditions for the maximization of the log-likelihood with respect to  $\psi$  are proportional to  $\nu$ . Then

$$\hat{\psi} = \underset{\psi}{\text{argmax}} \mathcal{L}^*(\psi), \quad (20)$$

with

$$\mathcal{L}^*(\psi) = \sum_{t=1}^T -\frac{1}{2} \left[ \ln |S_t(\psi)| + \text{tr}(S_t(\psi)^{-1}R_t) \right]. \quad (21)$$

The score vector of observation  $t$  obtains as<sup>2</sup>

$$s_t(\psi) = \frac{1}{2} \left\{ \left[ (\text{vec}(R_t))' - (\text{vec}(S_t))' \right] (S_t^{-1} \otimes S_t^{-1}) \frac{\partial \text{vec}(S_t)}{\partial \psi} \right\}. \quad (22)$$

Assuming a correctly specified mean  $E[R_t|\mathcal{F}_{t-1}] = S_t$ ,  $s_t(\psi)$  is a martingale difference sequence since

$$E[s_t(\psi)|\mathcal{F}_{t-1}] = 0. \quad (23)$$

Consequently, as noted by Bauwens et al. (2012) and Noureldin et al. (2012), under usual regularity conditions (see e.g. Wooldridge, 1994)  $\hat{\psi}$  is consistent and asymptotically normal even if the Wishart assumption is violated, provided that the conditional mean is correctly specified. From the QL

---

<sup>2</sup>See Noureldin et al. (2012).

function in Eq. (19) we obtain the period- $t$  log-likelihood contribution

$$\begin{aligned}
\ell_t^* &= -\frac{1}{2} [\ln |S_t| + \text{tr} (S_t^{-1} R_t)] \\
&= -\frac{1}{2} [\ln |L_t D_t L_t'| + \text{tr} ((L_t D_t L_t')^{-1} R_t)] \\
&= -\frac{1}{2} [\ln |D_t| + \text{tr} (L_t D_t^{-1} L_t' R_t)] \\
&= -\frac{1}{2} \left[ \sum_{i=1}^n \ln (d_{i,t}) + \text{tr} (D_t^{-1} L_t' R_t L_t) \right] \\
&= -\frac{1}{2} \sum_{i=1}^n \left[ \ln (d_{i,t}) + \frac{g_{i,t}}{d_{i,t}} \right]. \tag{24}
\end{aligned}$$

In contrast to other CAW specifications no matrix inversions are required for the calculation of the QL function and therefore it is computed very quickly. Standard errors can be obtained by the well known sandwich formula e.g. provided in Bollerslev and Wooldridge (1992). However, initial investigation showed that the QL function is highly multi-modal, such that standard local gradient based optimization algorithms fail if the realized covariance measure comprises more than a few assets. We therefore recommend to use gradient-free methods, like MATLAB's *patternsearch*.

### 3.3.2 Three-Step Estimation

This paper focusses on high-dimensional forecasting of realized covariance measures due to its relevance for practical applications. Here the curse of dimensionality precludes one-step estimation of the DPC-CAW model. Aielli and Caporin (2015) propose a three-step estimation technique called the *DPC estimator*, which can be easily adapted to the CAW framework. The procedure works as follows:

1. Estimate  $S = LDL'$  via  $\hat{S} = T^{-1} \sum_{t=1}^T R_t$ ;
2. Conditional on step 1. estimate  $(a, b)'$  by fitting a scalar CAW model to the sequence of realized covariance measures, essentially assuming  $R_t | \mathcal{F}_{t-1} \sim \mathcal{W}_n(\nu, Q_t/\nu)$ , where  $Q_t$  is given by Eq. (6) and  $S \stackrel{!}{=} \hat{S}$ . Recover  $\{\hat{Q}_t\}_{t=1}^T$  to calculate  $\{\hat{g}_{i,t}\}_{t=1}^T$  for  $i = 1, \dots, n$ ;
3. Conditional on 1. and 2. estimate  $\{\alpha_i, \beta_i\}_{i=1}^n$  via univariate QML based on Eqs. (8) and (18)

separately  $\forall i$ . The  $i$ 'th log-likelihood is given by

$$\mathcal{L}_i(a, b) = \sum_{t=1}^T \left[ (\nu/2 - 1) \ln(g_{i,t}) - \ln(\Gamma(\nu/2)) - (\nu/2) \ln(2d_{i,t}/\nu) - 0.5\nu g_{i,t}/d_{i,t} \right]. \quad (25)$$

Analogous to the Wishart,  $\mathcal{L}_i(a, b)$  features a QML interpretation given the previously estimated  $\{\hat{g}_{i,t}\}_{t=1}^T$ .

Steps 1 and 2 estimate the parameters of the eigenvector driving process by fitting a scalar CAW model to the sequence of realized covariance measures directly and employing covariance targeting in order to alleviate the curse of dimensionality. The motivation for these two estimation steps stems from the finding that estimated  $\{Q_t\}$  sequences tend to show a particularly good fit to the conditional means of the realized covariance measures. This result is illustrated in Figure 1, which shows estimates of the individual  $Q_t$  elements obtained via one-step QML estimation of the DPC-CAW model to a 3-dimensional realized covariance subset of the data discussed in Section 4. The  $Q_t$ -dynamics closely follow the pattern of the realized (co)variance data.

Note that steps 1 and 2 result in biased and possibly inconsistent estimates of the parameters  $a$ ,  $b$  and  $S$  since the scalar CAW likelihood in step 2 is not correctly specified (the matrix  $Q_t$  is not the conditional mean of  $R_t$ ). Subsequently, conditional on Steps 1 and 2 the parameters of the eigenvalue driving processes are estimated. This last estimation step does not add to the possible inconsistency due to the QML interpretation of the according likelihoods given in Eq. (25). Notice that the intercept parameters  $\gamma_i$  were fixed in step 1, such that step 3 essentially corresponds to univariate GARCH estimation with variance targeting.

The quasi-likelihood functions in step 2 and 3 are smooth, hence standard gradient based optimization procedures can be applied. However, estimation of standard errors becomes complicated due to the aforementioned misspecification error (see also Aielli and Caporin, 2015).

The three-step approach is simple and intuitive but comes with the disadvantage of introducing bias and inconsistency in parameter estimation. Section 4 analyzes the properties of obtained estimates in an extensive simulation experiment. The results suggest that bias is present but acceptably small or of reduced impact, especially given the huge dimension of the estimation problem. In ad-

dition, the forecasting application of Section 5 shows that these issues do not negatively affect the out-of-sample performance, which stays in the focus of empirical applications.

The DPC estimator enables quick estimation of high-dimensional specifications. 3-step estimation of a 100-dimensional DPC-CAW specification with  $T = 2500$  takes at most 100 seconds using an Intel Core i7 2.60 GHz processor.

## 4 Simulation Experiment

We conduct an extensive simulation experiment in order to assess the finite sample properties of the DPC estimator. Since our focus lies on high-dimensional applications the cross-sectional size is set to  $n = 100$ .

The following parameter set up is used: The intercept matrix  $S = LDL'$  of the  $Q_t$  process is set equal to the average realized covariance matrix of the data employed in the empirical application of Section 5.1. We consider 9 distinct eigenvector recursion parameter set-ups, where the ARCH parameter  $a$  is set equal to 0.025, 0.05 or 0.1 and the GARCH parameter is chosen such that the persistence  $(a + b)$  equals 0.9, 0.95 or 0.99.

In order to achieve some variability in the eigenvalue recursion parameters they are drawn from uniform distributions according to

$$\alpha_i \sim U(0.22, 0.3), \quad \beta_i | \alpha_i \sim U(0.94 - \alpha_i, 0.99 - \alpha_i). \quad (26)$$

Consequently the persistence parameters  $(\alpha_i + \beta_i) \in [0.94, 0.99]$ . The degree of freedom parameter  $\nu$  is set to  $\nu = 100$ . This parameter set up is inspired by parameter estimates obtained in the empirical application of Section 5.1. The whole experiment covers 500 independent simulations for each of the four time series lengths  $T = 1000$ ,  $T = 2500$ ,  $T = 5000$ ,  $T = 10000$  and each of the 9 parameter constellations.

We focus on estimation steps 1 and 2 since the likelihood in estimation step 3 is correctly specified and QML-optimal conditional on steps 1 and 2.

**Estimation Step 1** Note that the symmetric  $100 \times 100$  parameter matrix  $S$  comprises 5050 distinct model parameters. We therefore restrict the analysis to the 100 eigenvalues  $d_i$ . These parameters are of particular importance since they determine the level of the eigenvalue recursions in estimation step 3. We only report results for  $T = 2500$  since the overall findings are similar over the different sample sizes, though showing slightly increasing biases and decreasing dispersion of estimation errors with increasing  $T$  indicating inconsistency.<sup>3</sup> Figure 2 shows the mean, maximum and minimum of percentage relative estimation errors for  $\hat{d}_i$ ,  $100 \cdot (\hat{d}_i - d_i)/d_i$  for  $i = 1, \dots, 100$ , where the  $d_i$ 's are sorted in descending order. Each subplot corresponds to one of the 9 parameter settings for  $(a, b)$ , where the setting  $(a = 0.025, a + b = 0.99)$  comes closest to our empirical findings of Section 5. We observe a considerable increase in relative biases with increasing persistence  $(a + b)$ . For  $a + b = 0.99$  and  $a = 0.1$  we obtain a relative bias of almost 100% for the smallest  $d_i$ . For  $a + b = 0.9$  instead, biases do not exceed 5% in absolute value. Overall, biases increase with the magnitude of the ARCH parameter  $a$ . For  $a = 0.05$  the largest bias is around 20% while for  $a = 0.1$  it almost reaches 100%. In general biases appear larger, the smaller the  $d_i$ . This trend is particularly obvious for the high ARCH, high persistence parameter set  $(a = 0.1, (a + b) = 0.99)$  which, however, does not appear to be relevant in practice (see the empirical results of Section 5). Note that the analysis is based on relative estimation errors and the total variation of 100 assets is mainly driven by the first one to three principal components, where biases are always less than 10% in absolute value. The  $d_i$  estimates are given in Table 1. We observe a sharp drop from the first to the second eigenvalue, as expected. The last twenty eigenvalues amount to less than 1% of the largest eigenvalue. The impact of the observed biases therefore appears minor.

**Estimation Step 2** Figure 3 reports Boxplots of relative estimation errors for  $a$ . Obtained biases range from 2.5% to 12% in absolute value and appear minor. We observe decreasing dispersion of estimation errors and increasing biases with increasing sample size for various parameter constellations.  $\hat{a}$  is upward biased for  $a = 0.025$  and downward biased for  $a = 0.05$  and  $a = 0.1$ . The largest biases are obtained for  $(a = 0.025, (a + b) = 0.9)$  and  $(a = 0.025, (a + b) = 0.95)$ , ranging from 9.8% to 12%. The smallest biases arise for  $a = 0.05$ , ranging from -2.6% to -5%. This medium ARCH

---

<sup>3</sup>Results are available upon request.

environment is close to our empirical results of Section 4.

Figure 4 depicts the distribution of relative estimation errors for  $b$ . Biases appear small with a maximum of 4.8% in absolute value for  $b = 0.875$  and  $(a + b) = 0.9$ . We again observe diverging biases for increasing sample sizes.

Summarizing the results, we find low to moderate biases in low ARCH / low persistence environments. In high ARCH / high persistence environments distortions occur which mainly affect the eigenvalues  $d_i$  for  $i > 3$ . These eigenvalues are rather low in absolute value. Hence biases are not expected to significantly affect the forecasting performance of the DPC-CAW model.

## 5 Empirical Application

### 5.1 Data

We apply the scalar DPC-CAW model introduced in Section 3 in order to capture the dynamics of 100-dimensional realized covariance measures. The data has been computed from one-minute intraday asset returns by the microstructure-noise and jump robust multivariate realized kernel method of Barndorff-Nielsen et al. (2011). The corresponding ticker symbols are shown in Table 2. Note that the choice of the particular type of realized measure is not an important issue here since the model can actually be fitted to any series of positive-definite realized covariance measures. The sample period starts at January 1, 2002, and ends on December 31, 2014, covering 3271 trading days.

Figure 5 depicts exemplary time-series plots of variance and covariance series and according sample autocorrelation functions for 4 stocks included in the data set. Descriptive statistics are provided in Table 3. The (co)variance processes are highly persistent, skewed to the right, leptokurtic and tend to move parallel to each other.

### 5.2 In-Sample Estimation Results

We start with analyzing the in-sample fit of the DPC-CAW model for various model-order settings using the BIC information criterion. The models are estimated with the 3-step estimation approach. We consider both, order selection for the eigenvector- and for the eigenvalue processes given in Eqs.

(6) and (8), jointly. For the eigenvalues we restrict the chosen order to be identical across the 100 assets. The BIC is computed at the value of the true (one-step) likelihood of Eq. (19), given the 3-step estimates. Table 4 shows the results. We find a clear indication for the standard (1,1) specification of the eigenvector recursion, which corresponds to the typically chosen DCC-GARCH specification for correlations. The distribution of BIC values over the various eigenvalue order-constellations is much more even and overall results in the preferred (3,4) model. For comparison we also report the BIC obtained for a standard HAR specification of the eigenvalue dynamics (see Corsi, 2009). The model boils down to a restricted autoregressive specification of order 20. The HAR model, although very popular in empirical applications, is not preferred in any case.

Table 5 reports a summary of the obtained estimates for the BIC-preferred (3,4)-(1,1) DPC-CAW specification. The estimated persistence for the eigenvector- and eigenvalue recursions is very high with  $(a + b) = 0.997$  and a median of  $\sum_{\ell=1}^p \alpha_{i,\ell} + \sum_{\ell=1}^q \beta_{i,\ell}$  of 0.978. This corresponds to the findings in Aielli and Caporin (2015) and resembles analogous results for scalar DCC-GARCH applications with intercept targeting.

The right panel of Figure 5 shows sample autocorrelation functions of standardized Pearson residuals from the DPC-CAW(3,4)-(1,1) model for exemplary variance and covariance series of four stocks included in the 100-dimensional data set. The results presented in the Figure are representative for the complete set of stocks. The residuals are obtained as standardized Martingale differences

$$\begin{aligned} e_t^* &= \left( \text{Cov}[\text{vech}(R_t) | \mathcal{F}_{t-1}] \right)^{-1/2} \text{vech} \left( R_t - E[R_t | \mathcal{F}_{t-1}] \right) \\ &= \left[ \frac{1}{\nu} L_n (I_{n^2} + K_{nn}) (S_t \otimes S_t) L_n' \right]^{-1/2} \text{vech} \left( R_t - S_t \right), \end{aligned}$$

where  $K_{nn}$  denotes the commutation matrix (see e.g. Lütkepohl, 2005) and  $L_n$  denotes the elimination matrix defined by  $\text{vech}(X) = L_n \text{vec}(X)$ . Under the null of correct model specification these residuals should be serially uncorrelated. The ACFs are depicted together with 95% Bartlett confidence bands for variance and covariance series separately and illustrate the overall good fit of the DPC-CAW approach. The model successfully reduces the serial dependence to a minimum. We however observe some remaining predictability in the residual series: 441 of the 5050 series do not

pass the Ljung-Box test for zero autocorrelation at the 1% level and 100 lags. The literature reports much worse fractions for applications of much lower dimension (see e.g. the model diagnostic results for the flexible CAW specifications in Golosnoy et al., 2012, for a 5-dimensional application). The diagnostics therefore imply a good fit to the complex dynamics of 5050 distinct variance and covariance series. Also note that we may interpret some remaining residual predictability as 'a price to pay' for the 3-step estimation approach which is however unavoidable for high-dimensional applications. The residual ACFs in Figure 5 show that remaining predictability is typically found in variance residuals. This may be related to the direct modeling of principal component variances rather than return variances.

### 5.3 Out-of-Sample Forecasting

We now compare the out-of-sample 1-period ahead forecasting performance of the DPC-CAW specification to alternative forecasting models proposed in the literature on realized covariance modeling. We consider two out-of-sample windows: The first window starts at January 1, 2008, and ends on December 31, 2011, covering the subprime crisis period. The window exhibits a particularly high volatility level and pronounced volatility peaks. The second window covers a period of low to moderate volatility from January 1, 2012, until December 31, 2014, representing normal stock market fluctuations (see the left panel in Figure 5 for exemplary time series plots). The models are re-estimated daily using a rolling window of the previous 1750 covariance measures, i.e. roughly 7 years of data. New forecasts are generated based on the updated parameter estimates.<sup>4</sup>

#### 5.3.1 Competing Models and Forecast Evaluation

The scalar Re-DCC model of Bauwens et al. (2012) represents the 'natural' competitor for the DPC-CAW approach. The Re-DCC model decouples correlations and variances which facilitates three-step estimation similar to the DPC estimator (see Bauwens et al., 2012, for details). The model assumes a conditional central Wishart distribution for the realized covariance measure and decomposes the

---

<sup>4</sup>Due to the computational complexity the computations were parallelized and performed using CHEOPS, a scientific High Performance Computer at the Regional Computing Center of the University of Cologne (RRZK) funded by the DFG.

scale matrix  $S_t$  into

$$S_t = V_t \rho_t V_t, \quad (27)$$

where  $V_t = \text{diag}(\sqrt{s_{11,t}}, \sqrt{s_{22,t}}, \dots, \sqrt{s_{nn,t}})$  and  $\rho_t$  is the correlation matrix implied by  $S_t$ . We consider GARCH( $p, q$ ) recursions for the conditional variances:

$$s_{ii,t} = \gamma_i + \sum_{k=1}^p \alpha_{k,i} r_{ii,t-k} + \sum_{l=1}^q \beta_{k,i} s_{ii,t-l}. \quad (28)$$

The correlation matrix  $\rho_t$  is parameterized as follows:

$$\rho_t = (1 - a - b) \bar{\rho} + a P_{t-1} + b \rho_{t-1}, \quad (29)$$

where  $P_t$  is the realized correlation matrix

$$P_t = \{\text{diag}(R_t)\}^{-1/2} R_t \{\text{diag}(R_t)\}^{-1/2}. \quad (30)$$

$\bar{\rho}$  is estimated by the sample mean of realized correlation measures (“correlation targeting”).

We also consider a constant conditional correlation CAW (CCC-CAW) model since it represents a restricted Re-DCC specification where  $a = b = 0$ . In a similar fashion we restrict the DPC-CAW model to  $a = b = 0$  in order to obtain the CAW-analogue to the OGARCH model (O-CAW). The set of competing models is completed by the  $DPC - CAW_{0f}$  model which is obtained by restricting the eigenvalue dynamics of the DPC-CAW model to  $\alpha_i = \alpha$  and  $\beta_i = \beta \forall i = 1, \dots, n$ . This particular model restriction turned out to be favorable in forecasting applications.

We furthermore analyze an exponentially weighted moving average (EWMA) specification, called RiskMetrics (see J.P. Morgan, 1996), which boils down to exponential smoothing of realized covariance matrices using a preset smoothing parameter  $\lambda$ . The forecast of the realized covariance matrix is then given by

$$E[R_t | \mathcal{F}_{t-1}] = (1 - \lambda) R_{t-1} + \lambda E[R_{t-1} | \mathcal{F}_{t-2}], \quad (31)$$

where  $\lambda$  is set to its typical value for daily data, i.e.  $\lambda = 0.94$ .

The first two columns of Table 6 provide an overview of all considered model specifications. The  $(p, q)$  column describes the number of GARCH lags in the conditional variance specification (Re-DCC and CCC-CAW) or in the eigenvalue recursion (DPC-CAW and O-CAW), respectively.

We now turn to the evaluation of the forecasting performance. Let  $L(\hat{X}, X)$  denote the Euclidean distance of the half-vectorization of the forecast error matrix given by

$$L(\hat{X}, X) = \text{vech}(\hat{X} - X)' \text{vech}(\hat{X} - X), \quad (32)$$

where  $\hat{X}$  represents a particular matrix forecast and  $X$  the according realization. We apply five different loss functions in order to evaluate the forecasting performance of the considered models:

- (i) MSE of predicted covariance matrix:  $L(\hat{R}_t, R_t)$ ;
- (ii) MSE of predicted variances:  $(\text{diag}(\hat{R}_t - R_t))' (\text{diag}(\hat{R}_t - R_t))$ ;
- (iii) MSE of predicted correlation matrix:  $L(\hat{\rho}_t, \rho_t)$ ;
- (iv) Variance of predicted global minimum variance portfolio (GMVP):  $V_{GMPV,t}$ ;
- (v) QLIKE:  $QLIKE_t = \ln |\hat{R}_t| + \text{vec}(\hat{R}_t^{-1} R_t)' \iota$ .

The model-specific forecast of the covariance matrix  $R_t$  is given by  $\hat{R}_t = E[R_t | \mathcal{F}_{t-1}]$  and accordingly  $\hat{\rho}_t = \{\text{diag}(\hat{R}_t)\}^{-1/2} \hat{R}_t \{\text{diag}(\hat{R}_t)\}^{-1/2}$ . We use the realized kernel estimate  $R_t$  as unbiased proxy for the true covariance matrix at period  $t$ .

Loss function (i) considers whole covariance matrix forecasts, while (ii) and (iii) focus on variances and correlations instead. These quantities are of particular interest since DCC frameworks model variance and correlation dynamics separately. Loss function (iv) considers economic losses via computing the forecast of the variance of the GMVP given by  $V_{GMPV,t} = \hat{w}' R_t \hat{w}$ , with  $\hat{w} = \hat{R}_t \iota / (\iota' \hat{R}_t^{-1} \iota)$ , where  $\iota$  is an  $n$ -dimensional vector of ones. See e.g Patton (2011) for a discussion of the properties of the QLIKE loss function (v), which is known to be robust to noisy volatility proxies.

We compute sample averages of the obtained losses over the respective forecasting windows and assess the significance of differences in losses via the model confidence set (MCS) approach of Hansen

et al. (2011). At a given confidence level  $(1 - \alpha)$  the MCS contains the single model or the set of models with the best forecasting performance. We select  $\alpha = 0.1$  as suggested by Hansen et al. (2011) and compute the confidence sets using the block bootstrap method with window lengths determined by the maximum number of significant parameters obtained by fitting an AR(p) process on the loss differences and 5,000 bootstrap replications.<sup>5</sup>

### 5.3.2 Forecasting Results

We start with discussing the forecasting results for the volatile market phase presented in Table 6. The DPC-CAW specifications generate the lowest losses for all considered loss functions. Except for the O-CAW model (QLIKE loss) and the CCC-CAW (variance loss) the 90% model confidence sets only include DPC- and Re-DCC CAW specifications. For the correlation loss and the economic GMVP loss function the MCS only includes the DPC approach. The most accurate variance forecasts are obtained for the Re-DCC model and the CCC-CAW approach with the DPC-CAW being still part of the MCS. The results confirm the presumption that the independent modeling of principal component variances with time-varying loadings offers a more precise description of the covariance and correlation dynamics, while the DCC- and CCC-CAW frameworks perform best in capturing pure variance dynamics by modeling them explicitly. The economic GMVP loss results then illustrate the importance of capturing correlation dynamics in the portfolio context. The DPC approach performs best in forecasting the global minimum portfolio variance.

The forecasting results for the calm market phase are given in Table 7. The findings are rather striking: the DPC-CAW model outperforms its competitors by a large margin and represents the only constituent of the 90% MCS across all considered loss functions.

We conclude that the DPC-CAW approach has particularly good forecasting properties and notably outperforms its Re-DCC competitors in forecasting asset return correlation and global minimum variance portfolio weights. This finding is of particular importance for portfolio management and is explained by the modeling of orthogonal component variances with time-varying loadings instead of pure scalar correlation dynamics in the DCC approach. The DPC model thereby offers more flexible

---

<sup>5</sup>We use the R package MCS of Bernardi and Catania (2015) to compute the model confidence sets.

correlation dynamics.

## 6 Conclusion

In this paper we propose a Dynamic Principal Component (DPC) CAW model for time-series of high-dimensional realized covariance measures of asset returns. The model represents an adaption of the DPC-GARCH model of Aielli and Caporin (2015) to the dynamics of realized covariance measures. A three-step estimation procedure similar to the DCC framework for asset returns makes the model applicable to high-dimensional covariance matrices.

We analyze the finite sample properties of the 3-step estimation approach in an extensive simulation experiment and provide an empirical application to realized covariance measures for 100 assets traded at the NYSE. The DPC-CAW model has particularly good forecasting properties and outperforms its competitors including DCC-CAW specifications for realized covariance measures.

## References

- [1] Aielli, G.P., Caporin, M., 2015. Dynamic principal components: a new class of multivariate GARCH models. Working Paper.
- [2] Alexander, C.O., 2001. Orthogonal GARCH. *Mastering Risk* 2, 21-38.
- [3] Alexander, C.O., Chibumba, A.M., 1997. Multivariate orthogonal factor GARCH. Mimeo: University of Sussex, UK.
- [4] Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realised volatility. *Econometrica* 71, 579–625.
- [5] Asai, M., McAleer, M., Yu, J., 2006. Multivariate stochastic volatility: a review. *Econometric Reviews* 25, 145–175.
- [6] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2011. Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics* 162: 149–169.
- [7] Barndorff-Nielsen, O.E., Shephard, N., 2004. Econometric analysis of realised covariation: high frequency covariance, regression and correlation in financial economics. *Econometrica* 72: 885–925.
- [8] Bauer, G.H., Vorkink, K., 2011. Forecasting multivariate realized stock market volatility. *Journal of Econometrics* 160, 93–101.
- [9] Bauwens, L., Braione, M., Storti, G., 2014. Forecasting comparison of long term component dynamic models for realized covariance matrices. CORE Working Paper.
- [10] Bauwens, L., Braione, M., Storti, G., 2016. A dynamic component model for forecasting high-dimensional realized covariance matrices. CORE Working Paper.
- [11] Bauwens, L., Laurent, S., Rombouts, J.V.K., 2006. Multivariate GARCH models: a survey. *Journal of Applied Econometrics* 21, 79–109.
- [12] Bauwens, L., Storti, G., and F. Violante. 2012. Dynamic conditional correlation models for realized covariance matrices. CORE Working paper.
- [13] Bernardi, M., Catania, L., 2015. The Model Confidence Set package for R. CEIS Working Paper.
- [14] Bollerslev, T., Wooldridge, J.M., 1992. Quasi maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews* 11:143–172.

- [15] Chiriac, R., Voev, V., 2011. Modelling and forecasting multivariate realized volatility. *Journal of Applied Econometrics* 26, 922–947.
- [16] Corsi, F., 2009. A simple approximative long-memory model of realized volatility. *Journal of Financial Econometrics*.
- [17] Engle, R.F., 2002. Dynamic conditional correlation: a simple class of multivariate GARCH models. *Journal of Business and Economic Statistics* 20: 339–350.
- [18] Golosnoy, V., Gribisch, B., Liesenfeld, R., 2012. The conditional autoregressive Wishart model for multivariate stock market volatility. *Journal of Econometrics* 167: 211–223.
- [19] Gouriéroux, C., Jasiak, J., Sufana, R., 2009. The Wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics* 150: 167–181.
- [20] Hansen, P.R., Lunde, A., Nason, J.M. 2011. The model confidence set. *Econometrica* 79: 453-497.
- [21] Lütkepohl, H., 2005. *New Introduction to Multiple Time Series Analysis*. Springer, Berlin.
- [22] Morgan, J.P., 1996. *RiskMetrics*. Technical Document, Fourth ed., New York.
- [23] Noureldin, D., Shephard, N., Sheppard, K., 2012. Multivariate High-Frequency-Based Volatility (HEAVY) Models. *Journal of Applied Econometrics* 27: 907–933.
- [24] Noureldin, D., Shephard, N., Sheppard, K., 2014. Multivariate rotated ARCH models. *Journal of Econometrics* 179: 16–30.
- [25] Patton, A.J., 2011. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* 16(1): 246–256.
- [26] Rao, C.R., 1965. *Linear statistical inference and its applications*. John Wiley & Sons.
- [27] Woolridge, J.M., 1994. Estimation and inference for dependent processes. *Handbook of econometrics* 4, 2639–2738.

125.16	16.39	10.98	9.26	7.99	6.47	6.42	5.57	5.31	5.31
5.22	4.84	4.80	3.97	3.83	3.79	3.69	3.43	3.35	3.22
3.10	3.00	2.95	2.83	2.74	2.71	2.66	2.61	2.55	2.48
2.37	2.28	2.27	2.19	2.17	2.13	2.11	2.07	2.01	1.94
1.93	1.87	1.85	1.82	1.79	1.77	1.72	1.69	1.64	1.62
1.59	1.57	1.54	1.49	1.47	1.47	1.42	1.41	1.37	1.34
1.32	1.29	1.29	1.28	1.25	1.25	1.24	1.23	1.23	1.20
1.19	1.18	1.12	1.12	1.06	1.05	1.05	1.03	1.02	1.00
0.99	0.98	0.97	0.93	0.91	0.88	0.85	0.83	0.83	0.82
0.81	0.81	0.79	0.72	0.68	0.65	0.54	0.51	0.47	0.45

Table 1: Sorted eigenvalues obtained from  $\hat{S} = T^{-1} \sum_{t=1}^T R_t$  for the data-set described in Section 5.

Symbol	Company	Symbol	Company
a	Agilent Technologies Inc.	gild	Gilead Sciences Inc.
aa	Alcoa Inc.	glw	Corning Incorporated
aapl	Apple Inc.	gps	Gap, Inc.
abt	Abbott Laboratories	gs	Goldman Sachs Group, Inc.
abx	Barrick Gold Corporation	hal	Halliburton Company
adbe	Adobe Systems Incorporated	hd	Home Depot, Inc.
adi	Analog Devices Inc.	hig	Hartford Financial Services Group, Inc.
adp	Automatic Data Processing	hon	Honeywell International Inc.
aig	American International Group Inc.	hpq	Hewlett-Packard Company
all	Allstate Corporation	ibm	International Business Machines Corporation
altr	Altera Corporation	intc	Intel Corporation
amat	Applied Materials Inc.	intu	Intuit Inc.
amd	Advanced Micro Devices Inc.	ip	Internation Paper Company
amgn	Amgen Inc.	jcp	J.C. Penney Company, Inc. Holding Company
amzn	Amazon.com, Inc.	jnj	Johnson & Johnson
apc	Anadarko Petroleum Corporation	jnpr	Juniper Networks, Inc.
axp	American Express Company	jpm	J P Morgan Chase & Co
ba	Boeing Company	klac	KLA-Tencor Corporation
bac	Bank of America Corporation	ko	Cocoa-Cola Company
bax	Baxter International Inc.	kr	Kroger Company
bbby	Bed Bath & Beyond Inc.	kss	Kohl's Corporation
bby	Best Buy Co., Inc.	lb	La Barge Inc.
bhi	Baker Hughes Incorporated	lltc	Linear Technology Corporation
bmy	Bristol-Myers Squibb Company	lly	Eli Lilly and Company
brcm	Broadcom Corporation	lmt	Lockheed Martin Corporation
c	Citigroup Inc.	low	Lowe's Companies, Inc.
cag	ConAgra, Inc.	luv	Southwest Airlines Company
cah	Cardinal Health Inc.	mas	Masco Corporation
cat	Caterpillar, Inc.	med	McDonald's Corporation
cbs	CBS Corporation new	mdt	Medtronic Inc.
cien	Ciena Corporation	met	MetLife, Inc.
cl	Colgate-Palmolive Company	mmc	Marsh & McLennan Companies, Inc.
cop	ConocoPhillips	mmm	3M Company
cost	Costco Wholesale Corporation new	mo	Altria Group
cscs	Cisco Systems, Inc.	mrk	Merck & Company, Inc. New
ctxs	Citrix Systems, Inc.	ms	Morgan Stanley Dean Witter & Co
cvs	CVS Caremark Corp.	msft	Microsoft Corporation
cvx	Chevron Corporation new	msi	Motorola Solutions, Inc.
dd	E.I. du Pont de Nemours and Company	mu	Micron Technology, Inc.
de	Deere & Company	nem	Newmont Mining Corporation
dis	Walt Disney Company	nke	Nike, Inc.
dow	Dow Chemical Company	ntap	NetApp, Inc.
duk	Duke Energy Corporation new	nvda	NVIDIA Corporation
ea	Electronic Arts Inc.	orcl	Oracle Corporation
ebay	Ebay Inc.	oxy	Occidental Petroleum Corporation
emc	EMC Corporation MA	payx	Paychex, Inc.
emr	Emerson Electric Company	pep	Pepsico, Inc.
f	Ford Motor Company DEL	pfe	Pfizer, Inc.
fitb	Fifth Third Bancorp	pg	Procter & Gamble Company
ge	General Electric Company	qcom	QUALCOMM Incorporated

Table 2: Data set of 100 stocks selected by liquidity from the S&P 500.

	Mean	Min.	Max.	Range	Std. dev.	Skewness	Kurtosis
Realized variances (100 time series)							
Min.	1.01	0.02	48.39	48.35	1.89	3.75	24.30
Median	3.29	0.10	117.00	116.88	5.04	8.90	146.03
Max.	12.51	0.35	7727.54	7727.50	151.31	43.58	2126.80
Realized covariances (4950 time series)							
Min.	0.20	-126.77	14.07	14.97	0.87	-1.09	33.86
Median	1.05	-3.32	63.42	68.12	2.61	10.08	169.49
Max.	3.90	-0.02	1262.30	1282.60	25.51	38.93	1851.31

Table 3: Descriptive statistics for the 5050 realized variance and covariance time series of the 100-dimensional data-set described in Section 5.

	Order of eigenvector process							
	(1,0)	(1,1)	(2,1)	(1,2)	(2,2)	(3,2)	(2,3)	(3,3)
(1,0)	-2.5515	-2.7671	-2.0818	-2.4645	-2.0505	-2.0298	-2.0460	-2.0998
(1,1)	-2.6289	-2.8342	-2.1779	-2.5363	-2.1531	-2.1402	-2.1502	-2.1892
(2,1)	-2.6287	-2.8340	-2.1778	-2.5361	-2.1531	-2.1404	-2.1503	-2.1890
(1,2)	-2.6304	-2.8361	-2.1793	-2.5377	-2.1546	-2.1418	-2.1517	-2.1905
(2,2)	-2.6303	-2.8360	-2.1793	-2.5376	-2.1546	-2.1420	-2.1518	-2.1904
(3,2)	-2.6301	-2.8358	-2.1791	-2.5374	-2.1544	-2.1418	-2.1516	-2.1902
(2,3)	-2.6311	-2.8371	-2.1801	-2.5384	-2.1553	-2.1428	-2.1525	-2.1911
(3,3)	-2.6309	-2.8370	-2.1799	-2.5382	-2.1552	-2.1426	-2.1524	-2.1909
(4,3)	-2.6308	-2.8368	-2.1798	-2.5380	-2.1550	-2.1425	-2.1522	-2.1908
(3,4)	-2.6314	<b>-2.8374</b>	-2.1801	-2.5384	-2.1554	-2.1428	-2.1526	-2.1911
(4,4)	-2.6313	-2.8372	-2.1800	-2.5383	-2.1553	-2.1427	-2.1524	-2.1910
(5,4)	-2.6312	-2.8371	-2.1798	-2.5381	-2.1551	-2.1426	-2.1523	-2.1908
(4,5)	-2.6313	-2.8373	-2.1800	-2.5382	-2.1553	-2.1428	-2.1525	-2.1910
(5,5)	-2.6312	-2.8372	-2.1799	-2.5381	-2.1552	-2.1426	-2.1524	-2.1908
HAR	-2.6280	-2.8341	-2.1773	-2.5357	-2.1525	-2.1397	-2.1497	-2.1883

Table 4: BIC information criteria for various lag-order constellations. BIC values:  $\times 10e7$ . Models are estimated using the 3-step estimation approach. The BIC is evaluated at the full (one-step) likelihood, see Eq. (19).

Eigenvalue Process								
	$\alpha_{i,1}$	$\alpha_{i,2}$	$\alpha_{i,3}$	$\beta_{i,1}$	$\beta_{i,2}$	$\beta_{i,3}$	$\beta_{i,4}$	$\sum_{\ell=1}^p \alpha_{i,\ell} + \sum_{\ell=1}^q \beta_{i,\ell}$
Median	0.311	0.074	0.000	0.132	0.068	0.135	0.135	0.978
Min.	0.025	0.000	0.000	0.000	0.000	0.000	0.000	0.947
Max.	0.492	0.180	0.116	0.519	0.348	0.373	0.378	0.987
Eigenvector Process								
	$a$	$b$	$a + b$					
	0.035	0.962	0.997					

Table 5: Summary of parameter estimates obtained by the DPC estimator for the 100-dimensional data-set described in Section 5 and the BIC selected model order (3,4)-(1,1).

Volatile Market: 01.01.2009 – 31.12.2011						
Model	(p,q)	Cov matrix	Var	Corr	GMVP $\times 10^2$	QLIKE
DPC-CAW	(1,1)	32289	8322	195.9	37.86	150.8
	(2,2)	32036	8296	195.5	37.74	150.4
	(3,3)	<b>31943</b>	8316	195.1	37.7	149.8
DPC-CAW <sub>of</sub>	(1,1)	32596	8511	193.5	37.76	147.0
	(2,2)	32262	8466	193.2	37.67	146.7
	(3,3)	32063	8450	<b>192.9</b>	<b>37.65</b>	<b>146.2</b>
Re-DCC-CAW	(1,1)	32346	8222	229.1	39.72	201.5
	(2,2)	32196	<b>8155</b>	228.8	39.59	200.4
	(3,3)	32335	8239	228.6	39.53	199.3
O-CAW	(1,1)	38834	10116	211.6	49.99	148.3
	(2,2)	38626	10112	211.2	49.97	148
	(3,3)	38528	10140	211	49.99	147.5
CCC-CAW	(1,1)	34359	8222	273.2	41.62	225.2
	(2,2)	34223	<b>8155</b>	273.2	41.54	224.7
	(3,3)	34371	8239	273.2	41.51	223.9
EWMA		37178	9823	204.9	38.9	162.9

Table 6: Mean daily forecasting losses for the period 01.01.2009 – 31.12.2011. The loss functions are defined in Section 5.3.1. The smallest value is shown in bold. Grey shaded values indicate that the 90% model confidence set includes the respective model.

Calm Market: 01.01.2012 – 31.12.2014						
Model	(p,q)	Cov matrix	Var	Corr	GMVP $\times 10^2$	QLIKE
DPC-CAW	(1,1)	1586	540.4	226.2	15.48	73.96
	(2,2)	1580	<b>539.7</b>	226.3	15.49	74.09
	(3,3)	<b>1578</b>	540.1	226.4	15.48	74.32
DPC-CAW <sub>of</sub>	(1,1)	1600	542.4	225.4	15.48	<b>66.76</b>
	(2,2)	1590	540.7	225.5	15.49	66.99
	(3,3)	1585	540.2	<b>225.4</b>	<b>15.48</b>	67.05
Re-DCC-CAW	(1,1)	1742	597.8	242.8	16.6	94.67
	(2,2)	1732	593.3	242.7	16.59	94.37
	(3,3)	1729	592.8	242.6	16.59	94.10
O-CAW	(1,1)	1756	609.4	243.9	20.08	110.16
	(2,2)	1750	608.2	244.2	20.08	110.43
	(3,3)	1749	607.6	244.5	20.08	110.59
CCC-CAW	(1,1)	1840	597.8	265.3	18.11	96.27
	(2,2)	1828	593.3	265.3	18.12	96.18
	(3,3)	1824	592.8	265.3	18.16	96.23
EWMA		1715	556.3	239.8	15.74	82.57

Table 7: Mean daily forecasting losses for the period 01.01.2012 – 31.12.2014. The loss functions are defined in Section 5.3.1. The smallest value is shown in bold. Grey shaded values indicate that the 90% model confidence set includes the respective model.

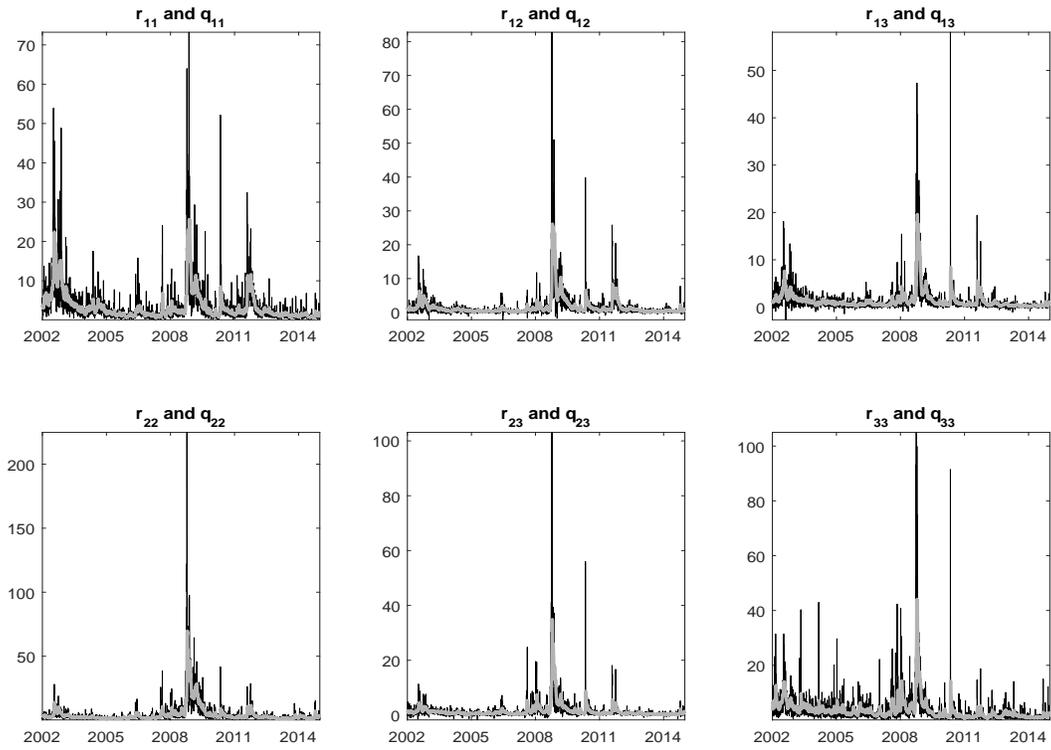


Figure 1: Black line: realized variances and covariances  $r_{i,j}$  of A ( $i = 1$ ), AA ( $i = 2$ ) and AAPL ( $i = 3$ ); Gray line: estimates of the individual  $Q_t$  elements obtained via one-step QML estimation of the DPC-CAW model as specified in Section 3 to the according set of 3-dimensional realized covariance matrices.

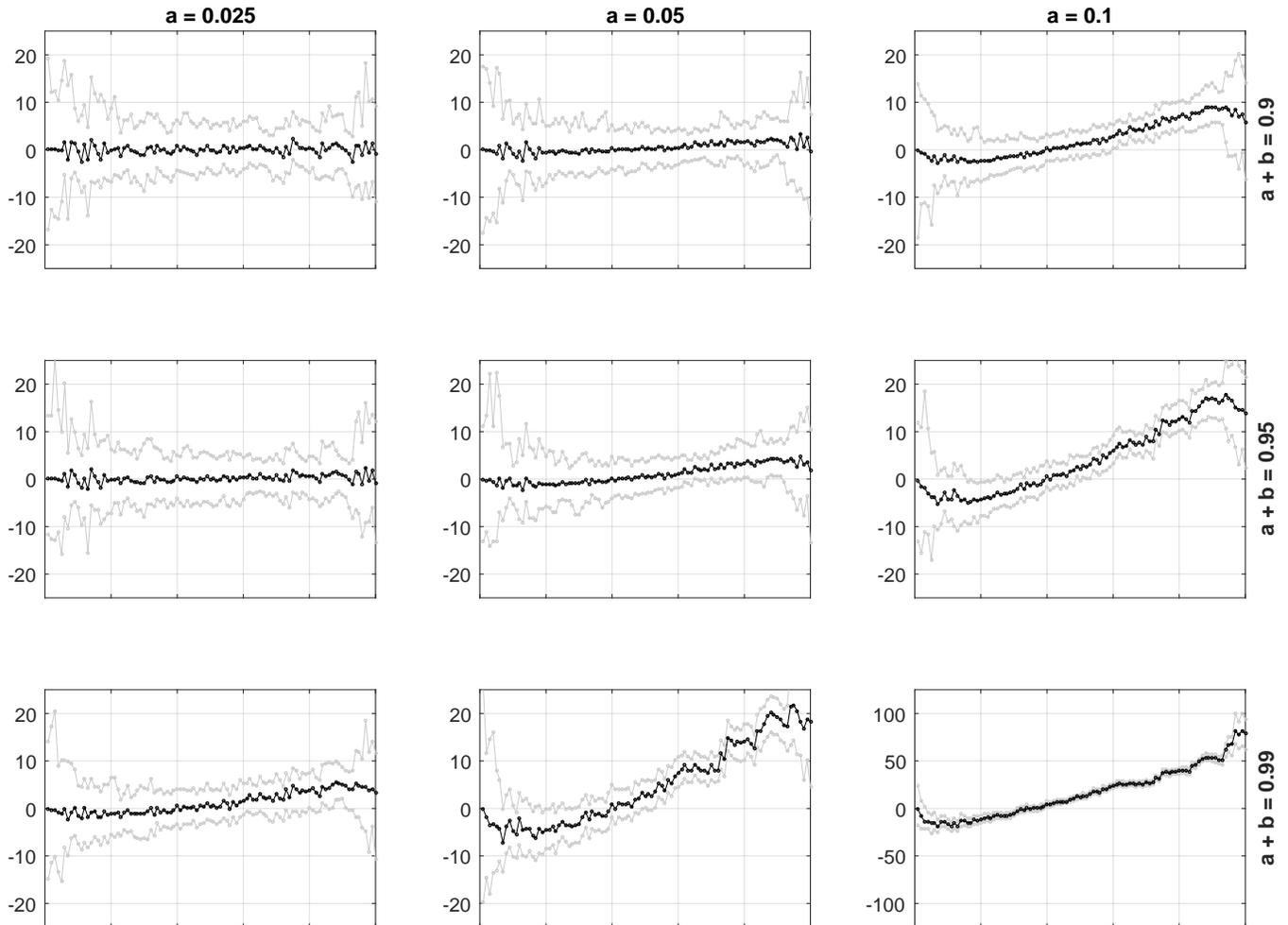


Figure 2: Black line: average relative estimation errors for  $d_i$ ; Gray line: maximum and minimum relative estimation errors for  $d_i$ . The results are obtained from the DPC estimator for the simulation experiment of Section 3 with  $T = 2500$ . The DGP parameter values are reported at the top of the panel for  $a$  and on the right side of the panel for  $(a+b)$ . Each line comprises 100 data points, one for each  $d_i$  in descending order with  $d_1$  being displayed on the left. Each data point is computed from the relative estimation errors in percent,  $100 \cdot (\hat{d}_i - d_i)/d_i$ .

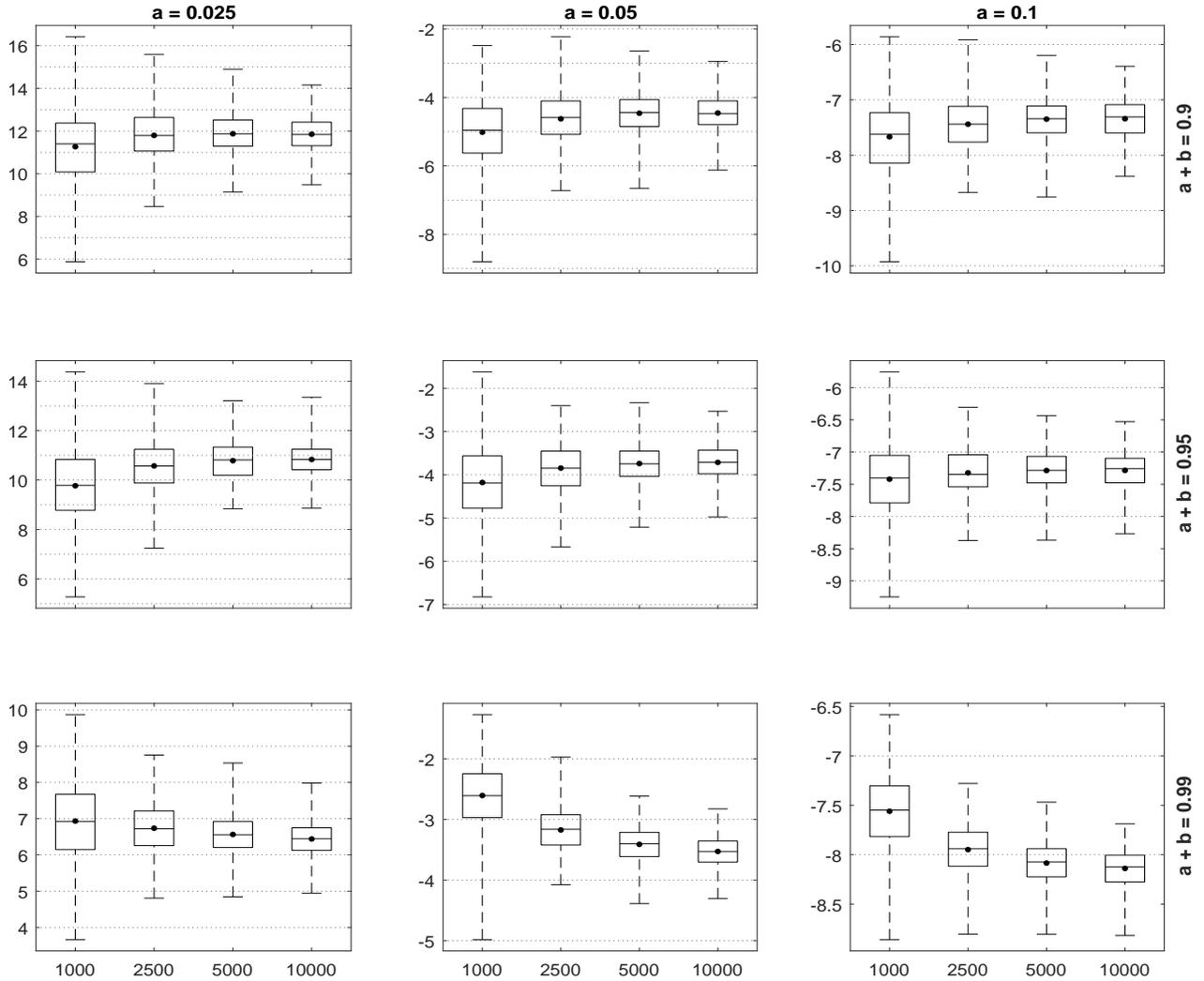


Figure 3: Boxplots of relative estimation errors  $100 \cdot (\hat{a} - a)/a$  obtained from the DPC estimator for the simulation experiment of Section 3. The DGP parameter values are reported at the top of the panel for  $a$  and on the right side of the panel for  $(a + b)$ . The first boxplot in each subplot comprises results for  $T = 1000$ , the second for  $T = 2500$ , the third for  $T = 5000$  and the fourth for  $T = 10000$ . The black dot denotes the mean.

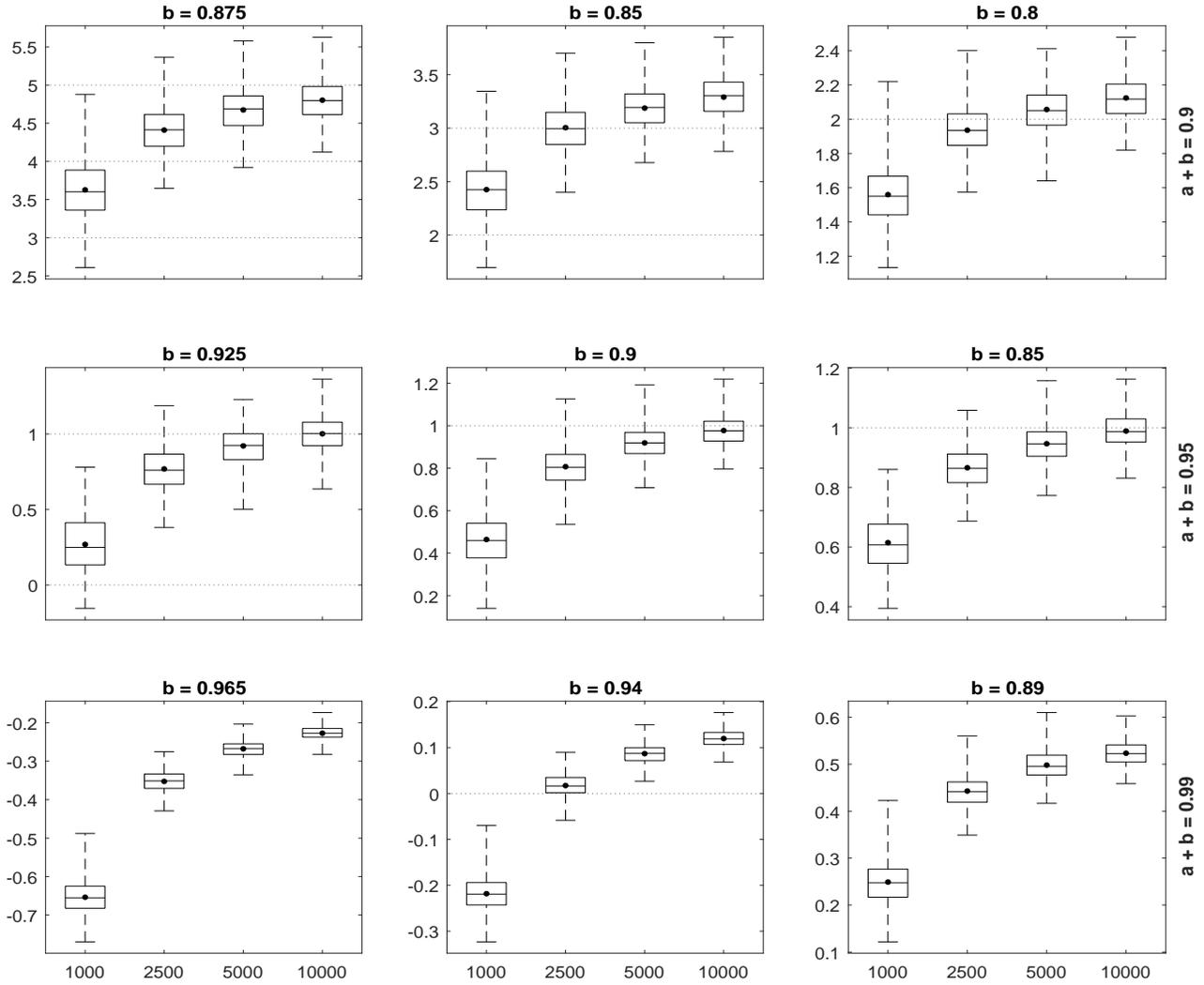


Figure 4: Boxplots of relative estimation errors  $100 \cdot (\hat{b} - b)/b$  obtained from the DPC estimator for the simulation experiment of Section 3. The DGP parameter values are reported at the top of each subplot for  $b$  and on the right side of the panel for  $(a+b)$ . The first boxplot in each subplot comprises results for  $T = 1000$ , the second for  $T = 2500$ , the third for  $T = 5000$  and the fourth for  $T = 10000$ . The black dot denotes the mean.

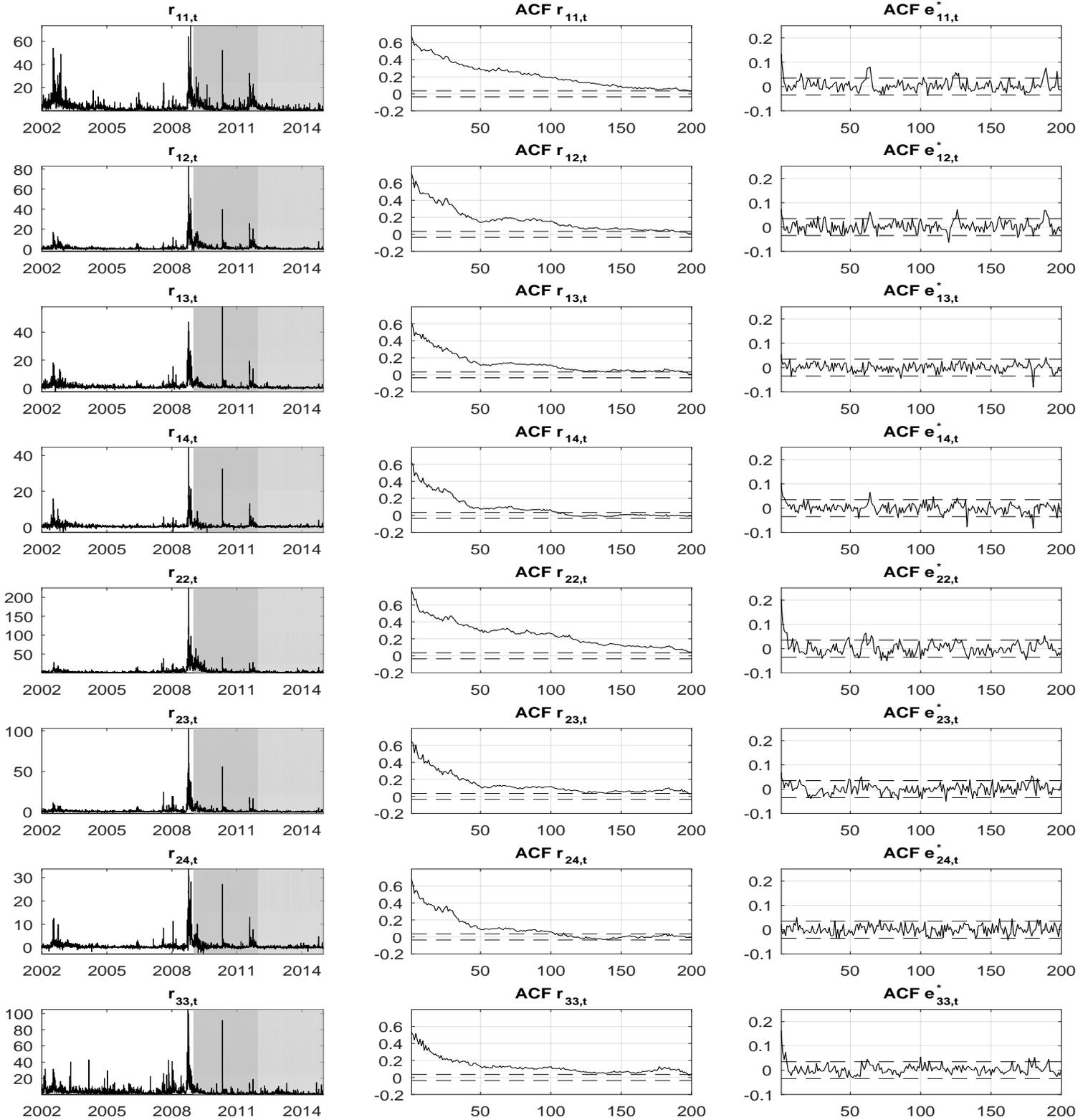


Figure 5: Realized (co)variance plots and sample autocorrelation functions (ACFs). Left panel: Sample of realized variances and covariances  $r_{i,j}$  of A ( $i = 1$ ), AA ( $i = 2$ ), AAPL ( $i = 3$ ) and ABT ( $i = 4$ ). Gray shaded areas indicate the periods covered by the forecasting experiment of Section 5.3. Middle panel: Sample ACFs of realized (co)variances together with 95% confidence bounds under the null of zero serial correlation. Right panel: Sample ACFs and according 95% confidence bounds of standardized Pearson residuals obtained from the BIC selected DPC-CAW(3,4)-(1,1) model estimated by the DPC estimator for the 100-dim. data-set illustrated in Section 5.