

Hartkopf, Jan Patrick

Article — Published Version

Composite forecasting of vast-dimensional realized covariance matrices using factor state-space models

Empirical Economics

Provided in Cooperation with:

Springer Nature

Suggested Citation: Hartkopf, Jan Patrick (2022) : Composite forecasting of vast-dimensional realized covariance matrices using factor state-space models, Empirical Economics, ISSN 1435-8921, Springer, Berlin, Heidelberg, Vol. 64, Iss. 1, pp. 393-436, <https://doi.org/10.1007/s00181-022-02245-1>

This Version is available at:

<https://hdl.handle.net/10419/309887>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Composite forecasting of vast-dimensional realized covariance matrices using factor state-space models

Jan Patrick Hartkopf¹

Received: 22 May 2021 / Accepted: 6 April 2022 / Published online: 5 May 2022
© The Author(s) 2022, corrected publication 2022

Abstract

We propose a dynamic factor state-space model for the prediction of high-dimensional realized covariance matrices of asset returns. Using a block LDL decomposition of the joint covariance matrix of assets and factors, we express the realized covariance matrix of the individual assets similar to an approximate factor model. We model the individual parts, i.e., the factor and residual covariances as well as the factor loadings, independently via a tractable state-space approach. This results in closed-form Matrix- F predictive densities for the distinct covariance elements and Student's t predictive densities for the factor loadings. In an out-of-sample forecasting and portfolio selection exercise we compare the performance of the proposed factor model under different specifications of the residual dynamics. These includes block diagonal residuals based on the GICS sector classifications and strict diagonality assumptions as well as combinations of both using linear shrinkage. We find that the proposed model performs very well in an empirical application to realized covariance matrices for 225 NYSE-traded stocks using the well-known Fama–French factors and sector-specific factors represented by exchange traded funds.

Keywords Factor model · Realized covariance · State-space model · Composite prediction

JEL Classification C32 · C38 · C51 · C58 · G17

1 Introduction

Modeling and forecasting covariance matrices of financial assets are essential in various fields like option pricing, risk management and portfolio allocation. Numerous approaches have been studied in the past decades, mostly based on daily asset return

✉ Jan Patrick Hartkopf
jan.hartkopf@gmail.com

¹ Cologne Graduate School, University of Cologne, Universitätsstr. 22a, 50937 Cologne, Germany

data. Due to the increasing availability of intraday asset return data it is nowadays possible to compute accurate nonparametric ex-post estimates of daily integrated covariance matrices of asset returns, called realized covariance, which are known to incorporate much more information about daily conditional covariances than daily return series, and come with the advantage that time-series models can directly be applied.¹

However, this involves two key challenges. Firstly, the need to assure positive definiteness of the covariance matrix forecasts, and secondly, to cope with the parameter proliferation arising through the fact that the number of objects to be modeled is proportional to the square of the number of assets. These aspects become even more pronounced in vast-dimensional applications that are considered in this paper.

Nevertheless, the direct time-series modeling of realized covariance matrices has arisen as an important new strand of literature, pioneered by the studies of [Gourieroux et al. \(2009\)](#), [Noureldin et al. \(2012\)](#), [Golosnoy et al. \(2012\)](#), and [Jin and Maheu \(2013\)](#), who focused on capturing the time-series behavior of realized covariance matrices through matrix-variate distributions like the Wishart and inverse-Wishart. In addition [Bauer and Vorkink \(2011\)](#) and [Chiriac and Voev \(2011\)](#) focused on decompositions of realized covariances to apply standard time-series models for the prediction of future realizations. More recent contributions examine additional aspects in modeling realized covariances by using more complex distributions like mixtures of inverse-Wisharts or the Matrix- F distribution that allow for fat tails in the covariance ([Jin and Maheu 2016](#); [Opschoor et al. 2017](#); [Opschoor and Lucas 2019](#); [Vasallo et al. 2021](#); [Zhou et al. 2019](#)). As further example, the Riesz distribution allows for incorporating liquidity differences in the intraday returns and reflecting tail heterogeneity ([Gribisch and Hartkopf 2022](#); [Blasques et al. 2021](#)). The benefits of LASSO (least absolute shrinkage and selection operator) techniques on the predictive accuracy are examined in [Callot et al. \(2017\)](#), the incorporation of (multiplicative) long-term components is studied by, e.g., [Bauwens et al. \(2017\)](#). However, most of the proposed models are limited to low-dimensional applications, say with 30 assets or less.

In this paper we contribute to another branch in realized covariance modeling, i.e., the application of factor structures for the assets' covariance matrices to deal with empirically more realistic scenarios in increasing dimensions. Prominent examples of the factor approach are the models proposed by [Tao et al. \(2011\)](#), [Asai and McAleer \(2015\)](#), [Jin et al. \(2019\)](#), [Sheppard and Xu \(2019\)](#), [Gribisch et al. \(2020\)](#) and [Brito et al. \(2018\)](#). While the first three approaches extract implicit factors from realized covariance data via an eigenanalysis and therefore build dynamic time-series models, the remaining three rely on economically motivated observed risk factors, like in the [Fama and French \(1993\)](#) three-factor model.

The use of observed risk factors is particularly appealing as it allows for time-variation in both the factor loadings and the residual components which yields flexibility gains when modeling the complex structures driving the evolution of realized covariances. We follow this route and propose a flexible factor state-space model for the prediction of realized covariance matrices of asset returns based on observed

¹ For a description of the concept of realized covariance matrices see, for example, [Andersen et al. \(2003\)](#), [Barndorff-Nielsen and Shephard \(2004\)](#), [Park and Linton \(2012\)](#), and [Lunde et al. \(2016\)](#).

risk factors which is capable of handling dynamic systems of vast dimensions (say 200 and more).

Our method relies on a factor decomposition of the realized asset covariance matrix from a joint construction of the realized measure for the factors and assets. After observing the covariances of both the assets and factors, we construct realized factor loadings and realized residual components from standard matrix decompositions resulting in a time-series for each of the individual parts. We then propose independent state-space models applied to the time-series of the factors, factor loadings and residual components akin to an approximate factor model taking the ‘true’ integrated factor and residual covariance matrices and integrated loadings as a latent state variables which are observed through their noisy realized counterparts.

For the factor and residual covariance matrices we combine independent Wishart measurement densities with parsimoniously parameterized matrix-variate Beta transitions in the vein of Windle and Carvalho (2014). Besides yielding positive definite covariance forecasts by construction, this specification has the advantage of being tractable in a sense that the predictive distribution of the data, and hence, the likelihood function is available in closed form. As a further contribution, we derive additional useful properties of the covariance model at hand. For the factor loadings we rely on well-established models from the TVP-VAR literature with stochastic volatility estimated by maximum likelihood (ML) techniques based on Kalman-like filtering.

Although a separate modeling approach has certain drawbacks as it ignores data-imposed dependencies in the measurements, it comes with the huge advantage of reducing model complexity allowing to efficiently handle vast-dimensional portfolios. In contrast to other state-space approaches that require high computational effort, our approach enables parameter estimation and prediction on a minute time scale. Based on the individual models, predictions for each component are readily made completely in parallel and are combined afterward to obtain a composite forecast of the full asset covariance matrix. This makes the model particularly appealing for practitioners.

In order to investigate whether the loss of information arising from the composite nature of our proposed model has any influence on its predictive accuracy we later include the Factor HEAVY (high-frequency-based volatility) model of Shepard and Xu (2019) in the set of competing models. The Factor HEAVY relies on a GARCH-type modeling of the observed factors’ and assets’ covariances based on a joint Wishart distribution combined with a strict factor structure (i.e., a sparse diagonal residual component). The diagonal residual assumption is crucial for their model as it enables a neat factorization into (conditionally) independent low-dimensional components allowing for straightforward estimation by quasi-maximum likelihood (QML). However, it imposes severe restrictions on the residual component which can be unrealistic in vast-dimensional applications. Our modeling approach allows to relax this assumption without effort.

The proposed composite factor state-space (CFSS)² model relies on the same factor decomposition as used in Brito et al. (2018). They combine the LASSO approach of Callot et al. (2017) for the factor and residual covariances with linear heterogeneous

² The name composite factor state-space model arises from the composite nature of how the predictions for the full asset covariance matrix are obtained.

autoregressive (HAR, Corsi 2009) processes for the factor loadings. Their approach has two major drawbacks in comparison with the CFSS model. First, since their model specification is prone to over-parameterization, they have to restrict the dynamics of the residual component to be driven by past residual variances only, neglecting possibly important information when forecasting the assets' covariance matrices. Second, and maybe more important, they operate with transformed time-series for the covariance parts based on the matrix-logarithm and make use of the matrix exponential function to ensure positive definiteness of the predictions which induces a bias to the forecasts. This bias might be neglectable in short-term forecasts; however, it gets more pronounced over longer horizons.

In the empirical part, we apply the CFSS model to daily covariance matrices for the returns of 225 NYSE traded stocks. In an extensive out-of-sample analysis we compare several model specifications based on varying sets of observed risk factors and different residual structures, as well as different restrictions imposed to the dynamics of the realized covariance components. We illustrate the predictive performance of our approach relative to several competing models. As performance measures we consider the accuracy of the (co)variance predictions and the ability to produce predictions for the global-minimum variance portfolio under different side restrictions as well as the ability to produce predictions for the mean-variance portfolio using momentum signals. Our out-of-sample results show that the CFSS model performs favorable in almost all dimensions relative to its competitors.

The remainder of the paper is organized as follows: Sect. 2 introduces the proposed CFSS model. The maximum likelihood estimation and parameter restrictions and the construction of composite forecasts are discussed in Sect. 3. Section 4 describes the data set. Section 5 discusses implementation details and presents the out-of-sample results. Finally, Sect. 6 concludes. Additional derivations for the model (parts) at hand are provided in Appendix, additional in- and out-of-sample results are deferred to Supplementary Appendix.

2 The model

2.1 Factor decomposition of realized covariances

Consider a positive definite, symmetric consistent estimate C_t of the joint latent integrated covariance matrix Σ_t , observed for a panel of p logarithmic asset prices and q observed factor prices at the trading days $t = 1, \dots, T$. Let C_t be partitioned as

$$C_t = \begin{pmatrix} C_t^f & C_t^{rf} \\ C_t^{fr} & C_t^r \end{pmatrix}, \quad (1)$$

where C_t^f constitutes the $q \times q$ realized covariance matrix for the factors, C_t^r denotes the $p \times p$ realized covariance matrix for the assets and $C_t^{fr} = (C_t^{rf})'$ denotes the $p \times q$ matrix of the realized covariances between the factors and assets.

In the present paper our particular interest is to predict the realized covariance matrix of the asset returns C_t^r . To this end, we exploit a factor decomposition relating the assets' realized covariance to the risk factors via

$$C_t^r = B_t C_t^f B_t' + C_t^e, \tag{2}$$

where the matrices B_t ($p \times q$) and C_t^e ($p \times p$) are uniquely found from a block LDL factorization of the block matrix representation in Eq. (1) as the regression coefficient matrix and the resulting Schur complement of C_t^f in C_t , i.e.,

$$B_t = C_t^{fr} (C_t^f)^{-1}, \quad C_t^e = C_t^r - C_t^{fr} (C_t^f)^{-1} C_t^{rf}. \tag{3}$$

The matrices B_t and C_t^e can be interpreted as realized factor loadings and as realized residual covariance matrix approximating the integrated loadings \mathcal{B}_t and residual components Σ_t^e , respectively.

Since we observe the complete covariance C_t at each point in time, we can use Eq. (3) to determine distinct time-series for C_t^f , C_t^e and B_t , such that these can be modeled and predicted separately. Afterward, they can be recomposed to obtain a forecast for the major quantity of interest, the full covariance of the assets C_t^r . Specifications for the dynamics of the individual parts are presented in the following section.

2.2 Model equations

We take the integrated components Σ_t^f , Σ_t^e and \mathcal{B}_t as latent state variables observed through their noisy realized counterparts and model them by a state-space approach with independent measurement densities $f(C_t^f | \Sigma_t^f)$, $f(C_t^e | \Sigma_t^e)$, $f(B_t | \mathcal{B}_t)$, and corresponding transition equations.

Before specifying the respective measurements for the individual parts, we make two additional assumptions on B_t and C_t^e . First, we assume row-wise independence of the betas, i.e., $\text{Cov}(b_{it}, b_{jt}) = 0 \forall i \neq j$, and $i, j = 1, \dots, p$, where b_{it} is a vector containing the i th row's entries of B_t . This assumption is comparably mild as, e.g., Brito et al. (2018) assume element-wise independence for the entries in B_t . Second, we impose the following block structure for the residual covariance

$$C_t^e = \begin{pmatrix} C_{1t}^e & * & * \\ * & \ddots & * \\ * & * & C_{St}^e \end{pmatrix}, \tag{4}$$

where the S diagonal blocks are of size $p_i \times p_i, i = 1, \dots, S$. The asterisk '*' indicates residual covariances which are economically insignificant and thus are ignored.³ The structure in Eq. (4) is supported by empirical evidence in many studies when the number of factors is chosen large enough, see, e.g., Fan et al. (2016), Ait-Sahalia and Xiu (2017), Brito et al. (2018), Gribisch et al. (2020). Expecting the significant

³ Economic significance is explained in more detail in Sect. 4.

entries to cluster in blocks around the diagonal might appear arbitrary. However, by rearranging the rows and columns in C_t^r such a block structure can almost always be achieved (at least approximately). As it is now common in the literature to model Σ_t^e as a sparse matrix (e.g., Sheppard and Xu 2019 assume Σ_t^e to be strict diagonal and only use residual variance information in a one-factor setting), we impose a block-diagonal structure and assume that the integrated residual covariance Σ_t^e is sufficiently measured by the marginals $f(C_{1t}^e | \Sigma_{1t}^e), \dots, f(C_{St}^e | \Sigma_{St}^e)$.

Next, we specify the measurement distributions for the different components C_t^f , $\{C_{it}^e\}_{i=1}^S$, and $\{b_{it}\}_{i=1}^P$, which read as follows

$$C_t^f | \Sigma_t^f \sim \mathcal{W}_q(n^f, \Sigma_t^f/n^f), \quad C_{it}^e | \Sigma_{it}^e \sim \mathcal{W}_{p_i}(n_i^e, \Sigma_{it}^e/n_i^e), \quad b_{it} | \beta_{it}, \Sigma_{it}^b \sim \mathcal{N}_q(\beta_{it}, \Sigma_{it}^b), \tag{5}$$

where $\mathcal{W}_m(d, s)$ denotes a m -dimensional Wishart distribution with $d > m - 1$ degrees of freedom and scale matrix s , and $\mathcal{N}_m(a, b)$ denotes a m -dimensional Normal distribution with mean a and covariance b . The latter choice is motivated by the fact that similar (marginal) distributions result from the assumption of a joint conditional Wishart distribution for C_t in (1), which can be seen as a natural candidate for the measurement (c.f. Gribisch et al. 2020, and see Philipov and Glickman 2006; Golosnoy et al. 2012; Noureldin et al. 2012; Bauwens et al. 2016, for applications of the Wishart to realized covariances).

In order to complete the composite factor state space (CFSS) model, we specify the transition densities for the latent time-varying integrated factor and residual covariances, Σ_t^f and $\{\Sigma_{it}^e\}$, as well as the integrated factor loadings $\{\beta_{it}\}$ and their time-varying covariance matrices $\{\Sigma_{it}^b\}$. For C_t^f and the distinct blocks $\{C_{it}^e\}$ we adopt the model of Windle and Carvalho (2014), which combines the measurements in Eq. (5) with a generalized Matrix-Beta type- I transition for the integrated precision, i.e., the inverse of the integrated covariances Σ_t^f and $\{\Sigma_{it}^e\}$, respectively, implying a shifted Matrix- F transition for Σ_t^f and $\{\Sigma_{it}^e\}$ themselves (see ‘‘Appendix A.2’’ for a derivation). In the vein of Windle and Carvalho (2014) we refer to this model as Uhlig Extension (UE) as it depicts a generalization of the process originally proposed in Uhlig (1994, 1997).

The UE model comes with three major advantages. First, irrespective of the dimension of the underlying covariance parts, it is parameterized parsimoniously with only three parameters. Parsimony has proven to be particularly advantageous in predicting high-dimensional covariance matrices (see, e.g., Bauwens et al. 2016). Second, it yields positive definite (p.d.) forecasts for C_t^f and C_{it}^e by construction. Hence, there is no need to resort to any transformation like the Cholesky decomposition that depends on the ordering of assets in C_t , or the matrix logarithmic transformation (Bauer and Vorkink 2011), which may induce severe biases of the predictions \widehat{C}_{t+1}^r through the exponential re-transformation. Third, and most important, for the UE model we can exploit the conjugacy between the Wishart and the Matrix-Beta distribution to obtain closed-form filtering formulas for tracking the latent states of the system and allowing the model parameters to be estimated by maximum likelihood in one step. (Details are given in the following section.)

In the UE model for the factor covariance matrix, the $q \times q$ symmetric, p.d. integrated precision $\Omega_t^f = (\Sigma_t^f)^{-1}$ evolves through

$$\Omega_t^f = \mathcal{U}(\Omega_{t-1}^f)' \Psi_t^f \mathcal{U}(\Omega_{t-1}^f) / \lambda^f, \quad \Psi_t^f \sim \mathcal{B}_q^I \left(\frac{k^f}{2}, \frac{n^f}{2} \right), \tag{6}$$

with conditional expectation given by $E[\Omega_t^f | \Omega_{t-1}^f] = \Omega_{t-1}^f k^f / [\lambda^f (k^f + n^f)]$ and initial condition $\Omega_1^f \sim \mathcal{W}_q(k^f, (n^f S_0^f)^{-1} / \lambda^f)$. $\mathcal{U}(\cdot)$ denotes the upper Cholesky factor, Ψ_t^f are $q \times q$ symmetric, p.d. iid Matrix-Beta type-*I* shocks to Ω_t^f .⁴ Besides the symmetric scale matrix S_0^f of the initial distribution, the UE model has three parameters, the degrees of freedom (d.o.f.) $n^f, k^f > q - 1$ and the smoothing parameter $\lambda^f > 0$, controlling the dynamics of the Ω_t^f process.⁵

Similarly to the integrated factor precision, we define $\Omega_{it}^e = (\Sigma_{it}^e)^{-1}$ for the $i = 1, \dots, S$ distinct $p_i \times p_i$ symmetric, p.d. blocks of Σ_{it}^e , and assume

$$\Omega_{it}^e = \mathcal{U}(\Omega_{it-1}^e)' \Psi_{it}^e \mathcal{U}(\Omega_{it-1}^e) / \lambda_i^e, \quad \Psi_{it}^e \sim \mathcal{B}_{p_i}^I \left(\frac{k_i^e}{2}, \frac{n_i^e}{2} \right), \tag{7}$$

with initial conditions $\Omega_{i1}^e \sim \mathcal{W}_{p_i}(k_i^e, (n_i^e S_{i0}^e)^{-1} / \lambda_i^e)$, where $n_i^e, k_i^e > p_i - 1, \lambda_i^e > 0$ and S_{i0}^e is the p.d. $p_i \times p_i$ initial scale matrix.

For the factor loadings $\{b_{it}\}$ we use a local-level version of the time-varying parameter VAR with stochastic volatility proposed in Moura and Noriller (2019). This model assumes the latent β s to follow heteroscedastic random walks and uses a Wishart process for modeling the volatility of the system. The latent $q \times 1$ vectors of factor loadings $b_{it}, i = 1, \dots, p$, evolve as driftless random walk processes of the form

$$\beta_{it} = \beta_{it-1} + \eta_{it}, \quad \eta_{it} \sim \mathcal{N}(0, \Sigma_{it}^b / \sigma_i), \tag{8}$$

where the transition is allowed to be affected by multivariate stochastic volatility shocks via Σ_{it}^b , scaled by $1/\sigma_i$. Similar to the UE model, the evolution of $\Omega_{it}^b = (\Sigma_{it}^b)^{-1}$ is specified by the following multiplicative transition

$$\Omega_{it}^b = \mathcal{U}(\Omega_{it-1}^b)' \Psi_{it}^b \mathcal{U}(\Omega_{it-1}^b) / \lambda_i^b, \quad \Psi_{it}^b \sim \mathcal{B}_q^I \left(\frac{k_i^b + 1}{2}, \frac{1}{2} \right), \tag{9}$$

with fixed initial condition. Here, Ψ_{it}^b are $q \times q$ iid singular Beta type-*I* shocks to Ω_{it}^b , with single d.o.f. parameter k_i^b governing the time-variation in the precision, given a predetermined value of $0 < \lambda_i^b < 1$. Again, exploiting the conjugacy between the Normal, Wishart and Matrix-Beta distributions, analytical expressions for the filtering

⁴ See, e.g., Gupta and Nagar (2000) for useful theorems about the matrix-variate Beta distribution.

⁵ In the literature the smoothing parameter λ^f is typically not estimated, but tied to the d.o.f. n^f and k^f through different restrictions which are directly imposed during estimation of the latter. We discuss these restrictions in Sect. 3.2.

distribution of the loadings and the likelihood function can be found, enabling fast and straightforward estimation of the unknown model parameters. It is worth noting that for $k_i^b \rightarrow \infty$ the nonlinear state-space model in Eqs. (5), (8) and (9) collapses to a homoscedastic version (Kim 2014; Moura and Noriller 2019). For $1/\sigma_i \rightarrow 0$ the latent factor loadings are time-invariant.

3 Estimation and forecasting

3.1 Predictive distributions and likelihood estimation

This section describes the ML estimation of the CFSS model. The unknown model parameters are estimated by independently maximizing the distinct elements in the conditional log-likelihood

$$\log L^r = \log L^f(\theta^f; C_{1:T}^f) + \sum_{i=1}^S \log L_i^e(\theta_i^e; C_{i1:T}^e) + \sum_{j=1}^p \log L_j^b(\theta_j^b; b_{j1:T}), \quad (10)$$

where L^f , L_i^e and L_j^b denote the likelihood contributions of the factors, residual blocks and factor loadings, with parameters $\theta^f = (n^f, k^f, \lambda^f)'$, $\theta_i^e = (n_i^e, k_i^e, \lambda_i^e)'$ and $\theta_j^b = (k_j^b, \lambda_j^b, \sigma_j)'$, respectively.⁶ The notation $A_{s:\tau}$ is used to denote the collection $\{A_s, \dots, A_\tau\}$.

The measurement distributions in Eq. (5) combined with the corresponding transition equations for Ω_t^f , $\{\Omega_{it}^e\}$, $\{\beta_{it}\}$ and $\{\Omega_{it}^b\}$ in Eqs. (6) to (9) constitute independent nonlinear state-space models for the realized quantities C_t^f , $\{C_{it}^e\}$ and $\{b_{jt}\}$. By exploiting the conjugacy between Normal, Wishart and Matrix-Beta distributions, Windle and Carvalho (2014) and Moura and Noriller (2019) derived several useful propositions for the UE model and the TVP-VAR, including the forward filtering distributions for the latent precision matrices and factor loadings, which turn out to be Wishart and multivariate- t , respectively. In addition, Windle and Carvalho (2014) propose a backward sampling scheme to efficiently draw the latent sequences $\{\Omega_t^f\}$ ($\{\Omega_{it}^e\}$, $i = 1, \dots, S$) in a single sweep within a Bayesian MCMC sampler. In the following we focus on their results enabling fast and easy to implement ML estimation of the model parameters, and the prediction of future realizations of C_t^f in Eq. (2).

Let $C_{1:t}^f$ capture the information on the realized factor covariance matrices from period 1 up to t . With the transition in Eq. (6) and the corresponding initial condition, the predictive distribution for C_{t+1}^f given past information is of Matrix- F type, i.e.,

⁶ The likelihood maximization is carried out using the MATLAB built-in numerical optimization routine `fmincon` under several restrictions. Code for the implementation of the estimation procedure is made accessible in the supplementary material.

$C_{t+1}^f | C_{1:t}^f \sim \mathcal{F}_q(n^f, k^f, \lambda^f S_t^f)$, with density given by

$$f(C_{t+1}^f | C_{1:t}^f) = \frac{\Gamma_q(\frac{n^f+k^f}{2})}{\Gamma_q(\frac{n^f}{2})\Gamma_q(\frac{k^f}{2})} \frac{|C_{t+1}^f|^{(n^f-q-1)/2} |\lambda^f S_t^f|^{k^f/2}}{|\lambda^f S_t^f + C_{t+1}^f|^{(n^f+k^f)/2}}, \tag{11}$$

with d.o.f. n^f, k^f and scale $\lambda^f S_t^f$, where $S_t^f = \lambda^f S_{t-1}^f + C_t^f$ (see Proposition 3 in Windle and Carvalho 2014). The d.o.f. k^f are known as fat-tail parameters with low (high) values indicating fatter (thinner) tails. For $k^f \rightarrow \infty$ the predictive distribution degenerates to the Wishart with n^f d.o.f. (see Opschoor et al. 2017). The conditional first moment that is used to generate one-step forecasts for C_{t+1}^f is given by⁷

$$E[C_{t+1}^f | C_{1:t}^f] = \frac{\lambda^f n^f}{k^f - q - 1} S_t^f, \tag{12}$$

where solving the recursion for S_t^f yields

$$S_t^f = \sum_{i=0}^{t-1} (\lambda^f)^i C_{t-i}^f + (\lambda^f)^t S_0^f, \tag{13}$$

implying that the forecast of C_{t+1}^f will be a scaled, geometrically weighted sum of the previous observations. From Eq. (13) one can see that λ^f controls the degree of smoothing of the observations. This is crucial when forming estimates and one-step ahead predictions (cf. Windle and Carvalho 2014).⁸

Based on Eq. (11) the likelihood function for the factor part obtains as $L^f = f(C_1^f; S_0^f) \prod_{t=2}^T f(C_t^f | C_{1:t-1}^f)$.⁹ Analogous results hold for the predictive distributions $f(C_{it+1}^e | C_{i1:t}^e)$ and likelihoods L_i^e of the distinct residual covariance blocks, and left out here for space saving reasons.

⁷ A derivation for the conditional second moment, i.e., $\text{Cov}[\text{vec}(C_{t+1}^f)]$, is given in ‘‘Appendix A.1’’.

⁸ In Sect. 3.2 we propose several restrictions on the smoothing parameter, directly relating it to the shape of the predictive distribution.

⁹ The likelihood function is based on S_0^f , which could already be challenging to estimate in small to medium dimensional applications. To circumvent this problem, we apply the approach proposed by Windle and Carvalho (2014), i.e., we set aside the first T' observations and use the set $\{S_{T'}^f, C_{T':T}^f\}$, with $S_{T'}^f = \sum_{i=0}^{T'-1} (\lambda^f)^i C_{T'-i}^f$ to estimate the remaining model parameters.

Now, let $b_{i1:t}$ capture the information on the realized factor loadings from period 1 up to t , for $i = 1, \dots, p$. The predictive density of b_{it+1} is that of a scaled multivariate t distribution (see Moura and Noriller 2019, Corollary 1)¹⁰

$$f(b_{it+1} | b_{i1:t}) = \frac{\Gamma\left(\frac{k_i^b+1}{2}\right)}{\pi^{q/2}\Gamma\left(\frac{k_i^b-q+1}{2}\right)} |P_{it+1}|^{1/2} [1 + (b_{it+1} - \bar{\beta}_{it+1})' P_{it+1} (b_{it+1} - \bar{\beta}_{it+1})]^{-\frac{k_i^b+1}{2}}, \tag{14}$$

with $k_i^b - q + 1$ d.o.f., mean $\bar{\beta}_{it+1}$, and scale $(k_i^b - q + 1)^{-1} P_{it+1}^{-1}$, where

$$\bar{\beta}_{it+1} = \frac{N_{it}}{(N_{it} + 1)} (\bar{\beta}_{it} + b_{it} N_{it}^{-1}), \tag{15}$$

$$P_{it+1} = \frac{N_{it}}{k_i^b (N_{it} + 1)} (S_{it}^b)^{-1}, \tag{16}$$

$$S_{it+1}^b = \lambda_i^b S_{it}^b + \frac{\lambda_i^b N_{it}}{(k_i^b + 1)(N_{it} + 1)} (b_{it} - \bar{\beta}_{it})(b_{it} - \bar{\beta}_{it})', \tag{17}$$

$$N_{it+1} = \frac{\sigma_i \lambda_i^b (N_{it} + 1)}{\sigma_i + \lambda_i^b (N_{it} + 1)}. \tag{18}$$

The initial conditions are fixed parameters.¹¹

The predictive expectation of the factor loadings is given by

$$E[b_{it+1} | b_{i1:t}] = \bar{\beta}_{it+1}. \tag{19}$$

Again, the likelihood function L_i^b is obtained as the product of the predictive densities in Eq. (14).

3.2 Parameter restrictions

Our proposed CFSS model as defined by Eqs. (5) to (9) includes $3 \times (1 + S + p)$ parameters, whereof the majority relates to the covariance dynamics of the factor loadings. In order to further increase the parsimony of our approach we utilize commonly imposed restrictions to the d.o.f. parameters of the loading processes and the smoothing parameters.

In initial investigations of the realized factor loadings in scenarios with different risk factors, we found the estimates for k_i^b to be similar throughout the asset dimension p . Moreover, we found that minor changes in the d.o.f. parameters do not have noteworthy effect on the predictive ability regarding mean forecasts for b_{it+1} . Hence, we preset k_i^b

¹⁰ Note a typo in Moura and Noriller (2019), where they define the scale of the multivariate t as $\frac{k_i^b}{k_i^b - q + 1} S_{it}^b$ by mistake. The error has been fixed here.

¹¹ We set $N_{i1} = 1$ and $S_{i1}^b = I_q$.

in the ongoing of this paper. Similar restrictions are indeed imposed by Uhlig (1997) and Kim (2014) in the context of constant coefficient VARs (see also Moura and Noriller 2019). Kim (2014) suggests to choose values $k_i^b \in [10, 25]$ which mirrors our findings for \hat{k}_i^b . Consequently, we adopt this in the empirical analysis.¹² For given k_i^b the process Ω_{it}^b is asymptotically degenerated if λ_i^b is too large, and explosive if λ_i^b is too small. Hence, it seems reasonable to further tie the smoothing parameter to the d.o.f. via $\lambda_i^b = k_i^b / (k_i^b + 1)$, implying $0.90 < \lambda_i^b < 0.97$, for $k_i^b \in [10, 25]$.

In addition we consider three different restrictions on the smoothing parameters $(\lambda^f, \{\lambda_i^e\})$. They are all considered as distinct model variants and their predictive performance is analyzed in the empirical application below. Each of them has been found to be valuable on its own. Note that all restrictions imply $\lambda^f, \{\lambda_i^e\} \in (0, 1)$ and that they can straightforwardly be imposed during the estimation of the degree of freedom parameters. Here they are exemplified for the factor part. Similar restrictions are later imposed for the residual blocks as well.

The first restriction has been originally proposed by Windle and Carvalho (2014) and is given by

$$\lambda^f = \left(1 + \frac{n^f}{k^f - q - 1}\right)^{-1}, \quad (\text{R1})$$

It implies that the one-step-ahead forecast for the realized factor (residual) covariance based on (12) is obtained as the exponentially weighted moving average (EWMA)

$$E[C_{t+1}^f | C_{1:t}^f] = (1 - \lambda^f)C_t^f + \lambda^f E[C_t^f | C_{1:t-1}^f],$$

which is known to deliver decent (short-term) predictive performance.¹³ Notably, this restriction also results in a martingale for the evolution of latent integrated factor (residual) covariance matrices, i.e., $E[\Sigma_t^f | \Sigma_{t-1}^f] = \Sigma_{t-1}^f$ ($E[\Sigma_{it}^e | \Sigma_{it-1}^e] = \Sigma_{it-1}^e$). A derivation for this result is given in ‘‘Appendix A.2’’.

The second restriction is chosen to induce a random walk behavior for the latent integrated precision matrices in Eq. (6), say $E[\Omega_t^f | \Omega_{t-1}^f] = \Omega_{t-1}^f$, i.e.,

$$\lambda^f = \frac{k^f}{k^f + n^f}. \quad (\text{R2})$$

A similar restriction has been imposed by Uhlig (1997) in the context of singular Matrix-Beta transitions for the latent precision. In the context of stochastic covariance modeling for daily return series, Moura et al. (2020) found (R2) to be performing well in out-of-sample portfolio allocations.

¹² The exact restrictions for k_i^b are postponed to Sect. 5.

¹³ Note that under restriction (R1) the prediction for C_{t+1}^f is similar to the one implied by the RiskMetrics methodology (Morgan 1996). However, in the latter the smoothing parameter is set arbitrarily to 0.96, whilst in the former λ is determined by the distributional shape mirrored in the degree of freedom parameters n^f and k^f .

The third restriction traces back to the work of Shephard (1994), where in an univariate modeling framework akin to Uhlig (1997) the smoothing parameter is chosen as the geometric mean of the latent beta shocks in the precisions’ transition equations. Shephard (1994) argues that this restriction rules out the case of the precision collapsing to zero if $t \rightarrow \infty$. Here, we propose a multivariate extension of this restriction which takes the form

$$\lambda^f = \exp\{E[\log |\Psi_t^f|]/q\}. \tag{R3}$$

In “Appendix A.3” we show that $E[\log |\Psi_t^f|] = \Psi_q^*(k^f/2) - \Psi_q^*((n^f + k^f)/2)$, where $\Psi_q^*(a)$ denotes the d -variate digamma function (see, e.g., Abramowitz and Stegun 1972; Gupta and Nagar 2000) and that (R3) results in a random walk for the log determinant process of the integrated precision matrices.

Figure 1 visualizes the restrictions on the smoothing parameter for an exemplary scenario of 25 assets, which corresponds to the average sector-size in the empirical application. Panels (a)–(c) show contour plots of (R1), (R2) and (R3) on the grid $[26, 300] \times [26, 300]$ for the d.o.f. parameters. The plots indicate that $\lambda \rightarrow 1$ ($\lambda \rightarrow 0$) for n fix at its lower bound and $k \rightarrow \infty$ (k fix at its lower bound and $n \rightarrow \infty$). Hence, the less fat tailed and noisier the observations are, the more the model smooths, and vice versa.

Panels (d)–(f) show contour plots of the differences between the ML estimate for λ (when fixing n and k) and the respective restrictions (R1), (R2) and (R3). The black circles mark the corresponding unrestricted ML estimates of the d.o.f. parameters, i.e., $\hat{n} \approx 89, \hat{k} \approx 175$. The unrestricted MLE for the smoothing parameter is $\hat{\lambda} = 0.63$ with a standard deviation of 0.017. We can see that (R1) and (R3) tend to be smaller than the MLE, whereas (R2) exceeds the MLE slightly, though all differences lie within one standard deviation. The performance of (R1) and (R3) deteriorates if both d.o.f. tend to their lower bound. However, it is common in practice to impose these restrictions when estimating the d.o.f. in order to reduce estimation uncertainty.

Imposing the restrictions stated above reduces the number of parameters to be estimated to $2 \times (1 + S) + p$. This further increases the parsimony but preserves the flexibility of the model at hand and allows for very fast estimation of the unknown model parameters based on numerical optimization routines.¹⁴

3.3 Composite forecasting

Exploiting the decomposition of the realized asset covariance in Eq. (2), we write the forecasting equation for the realized asset covariance C_{t+1}^f by combining the distinct forecasts for $C_{t+1}^f, \{C_{it+1}^e\}$ and $\{b_{it+1}\}$, i.e.,

$$\widehat{C}_{t+1}^r = \widehat{B}_{t+1} \widehat{C}_{t+1}^f \widehat{B}_{t+1}' + \widehat{C}_{t+1}^e, \tag{20}$$

¹⁴ In the empirical application in Sect. 5 we, e.g., consider a model for 225 assets with block-diagonal residual assumption based on nine industry sectors resulting in 245 (705) parameters for the restricted (unrestricted) model. Parallel estimation of all parameters takes < 5 minutes on a standard PC with a 2.3 GHz Intel Core i7 processor in MATLAB.

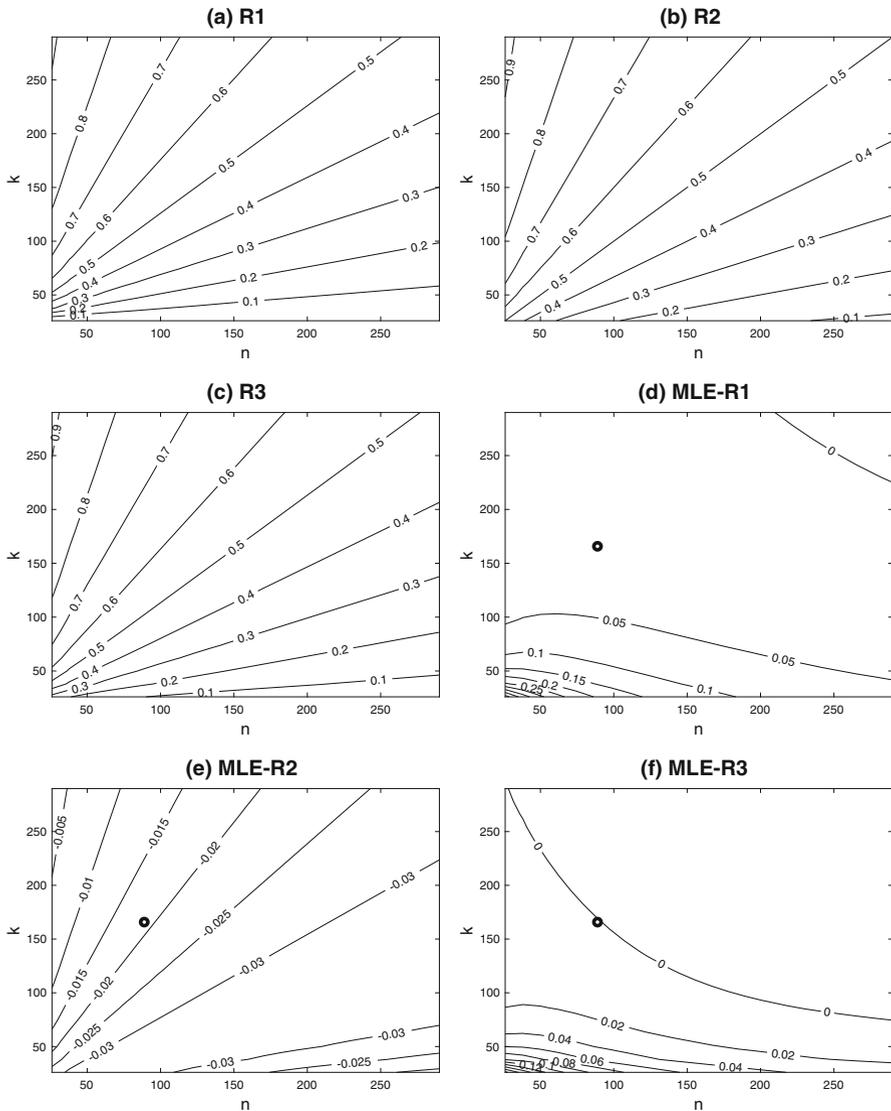


Fig. 1 Smoothing parameter restrictions. **a–c** Smoothing parameter of the UE model under restrictions (R1), (R2) and (R3). **d–f** Differences of the ML estimate for λ (when fixing n and k) and the values obtained under the respective restrictions. The black circles ‘o’ mark the respective unrestricted ML estimates of the d.o.f. parameters

where we use $\widehat{C}_{t+1}^f = E[C_{t+1}^f | C_{1:t}^f]$, $\widehat{C}_{it+1}^e = E[C_{it+1}^e | C_{i1:t}^e]$ and $\widehat{b}_{it+1} = E[b_{it+1} | b_{i1:t}]$ as obtained from the respective predictive moments in (12) and (19). Since the predictions of the factor and residual covariance are p.d., the forecast for the assets covariance will be p.d. by construction as well.

The composite forecasting procedure in Eq. (20) relies on a simple plug-in approach and hence ignores nonlinearities in the term $B_{t+1}C_{t+1}^f B'_{t+1}$ when predicting C_{t+1}^r . As pointed out by an anonymous referee a more general predictor could be obtained by combining the separate forecasts through the optimization of some, possibly strictly consistent, loss function for the assessment of forecasting performance. For example, the two components on the RHS of Eq. (20) could be rescaled by means of scalar coefficients estimated by minimizing the Frobenius norm of the in-sample forecast errors. However, we do not consider such a *calibrated* predictor in the present paper.¹⁵

4 Data

The given data set includes and expands the 60 dimensional data of Gribisch et al. (2020). It consists of 1510 daily observations for 225 stocks traded at the New York Stock Exchange and 12 risk factors, observed between January 3, 2007, and December 31, 2012, comprising a total of 28,203 time series of realized variances and covariances. The stocks are selected by liquidity from the S&P 500 index and are sorted by their sector and industry classification according to the Global Industrial Classification Standard (GICS). For the observed factors we use the market, the high-minus-low price-earnings ratio (HML) and small-minus-big market capitalization (SMB) factors as in the Fama and French (1993) three-factor model. Additionally we consider the sector-specific Spyder Exchange-Traded Funds (SPDR ETFs) for the nine sectors covered by the 225 stocks (Fan et al. 2016; Ait-Sahalia and Xiu 2017; Gribisch et al. 2020). A brief summary of the sector sizes along with the sector-specific ETF tickers and some descriptive statistics are given in Table 1.

The daily realized covariance matrices C_t are computed by using the composite realized kernel method of Lunde et al. (2016) based on 5-minute returns for the Fama–French factors and 1-minute returns for the assets as well as the SPDR ETFs (see Barndorff-Nielsen et al. 2011; Lunde et al. 2016, for further details). Given the joint realized covariance matrices for the assets and factors, we compute according to Eq. (3) the realized residual covariance matrices C_t^e and the realized factor loadings B_t which represent estimates for integrated residual covariance matrices and loadings, for various sets of factors. Figure 2 shows time-series plots of the realized variances for the three Fama–French factors, as well as sector-wise averages of realized asset variances, realized factor loadings and realized residual variances for the industrial, health care and financial sector.

In Fig. 3 we analyze the sparsity pattern for the realized residual covariance matrices. Following Ait-Sahalia and Xiu (2017) and Brito et al. (2018) we determine the economically significant entries of the residual correlation, i.e., we flag each entry with a minimum absolute correlation of 0.15 for at least 1/3 of the sample period. Panel (a) displays the results for the realized asset covariance matrices. Panels (b)–(d) display the results for the realized residual components after removing the common

¹⁵ Initial investigations showed no significant improvements over the composite forecasting procedure. Furthermore, a simulation study has shown that the nonlinearities have only minor effects on the diagonals (which play a minor role compared to the off-diagonal covariance elements) and are therefore negligible. (The results are provided in Supplementary Appendix.)

Table 1 Descriptive statistics for the data set

Sector	ETF	Members	10%	Mean	Median	90%
Energy (E)	XLE	16 (-)	1.27	6.03	3.20	11.64
Materials (M)	XLB	9 (-)	0.90	4.80	2.51	9.97
Industrials (I)	XLI	26 (8)	0.75	3.95	2.10	8.16
Consumer Discretionary (D)	XLY	35 (9)	0.98	5.87	2.88	13.33
Consumer Staples (S)	XLP	20 (7)	0.39	2.07	1.04	4.06
Health Care (H)	XLV	25 (7)	0.55	2.98	1.50	6.18
Financials (F)	XLF	34 (7)	0.82	9.45	2.95	19.85
Information Technologies (T)	XLK	44 (22)	1.10	4.74	2.85	9.71
Utilities (U)	XLU	16 (-)	0.45	2.71	1.27	5.05

The table lists the sectors covered by the data set. In column ETF the symbols of the respective sector ETF are stated. Members denote the number of stocks per sector as classified through their GICS code. The number of assets in the 60D data set of Gribisch et al. (2020) is given in parentheses. The descriptive statistics are average mean and median asset volatility, as well as average lower and upper 10% asset volatility

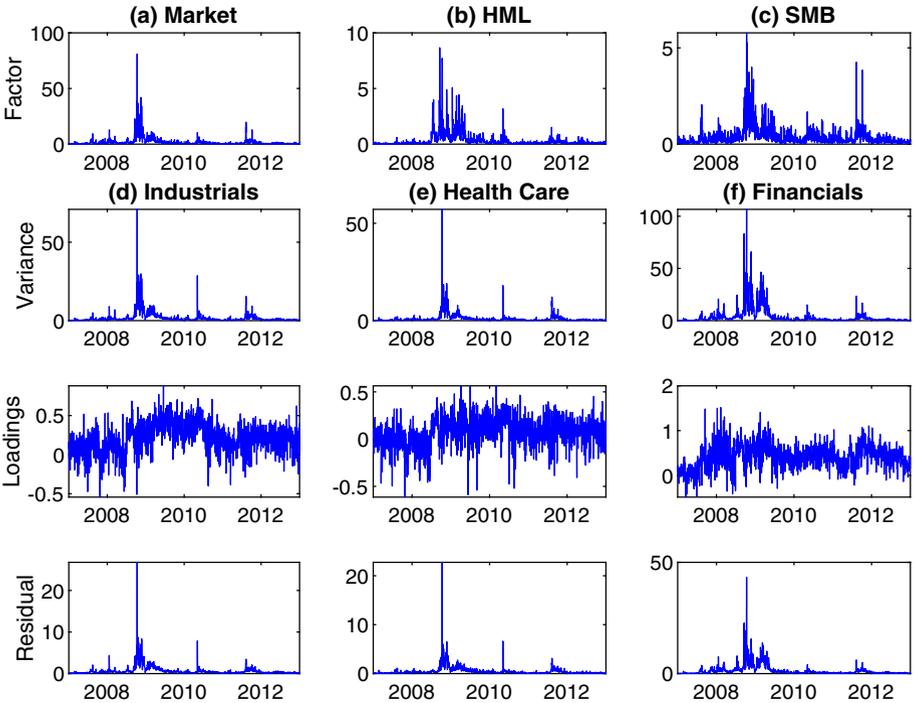


Fig. 2 Time-series plots of selected realized factor and aggregated (residual) asset variances. **a–c** Time series plots of the realized Fama–French factor variances; Columns **d–f** Time series plots of average realized asset variances (top), factor loadings on the market (middle) and residual variances (bottom) for the industrials, health care and financial sector

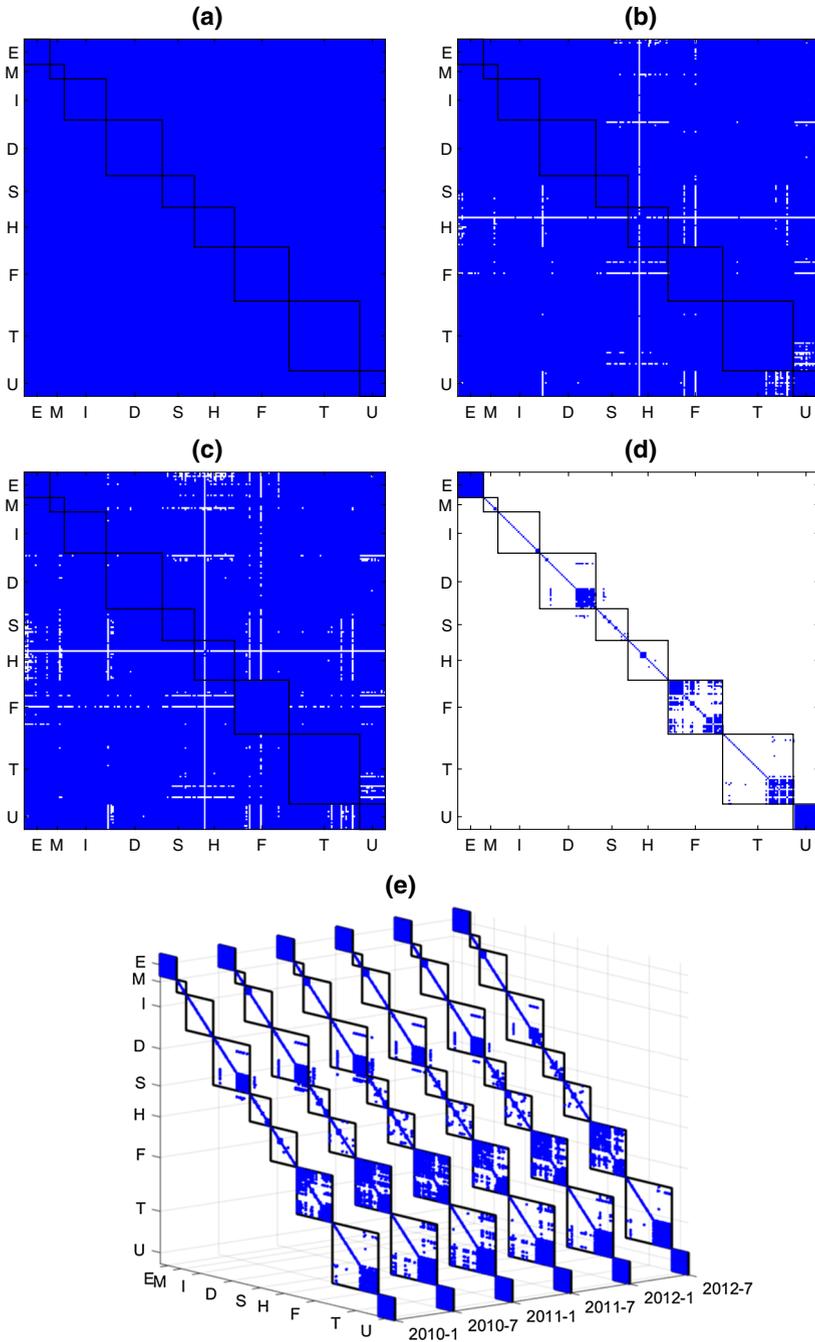


Fig. 3 Sparsity pattern of residual correlations. The figure displays the significant entries of the residual covariance matrices, relative to (a) zero, (b) one (Market), (c) three (Market + FF), and (d) 12 (Market + FF + ETFs) observed factors. e displays the the residual sparsity for 12 factors at different points in time. The sector labels are listed in Table 1

covariation driven by the market factor only, the three Fama–French factors and the Fama–French factors plus nine sector-specific factors based on the ETFs, respectively.

The analysis shows that using only one or three factors still yields very dense residual correlation patterns, indicating not only a strong rejection of the strict diagonality assumption, but also of the block-diagonality based on industry sectors (highlighted as black rectangles). However, by inclusion of the sector ETFs almost all of the significant inter-sectoral and most of the intra-sectoral correlations vanish, leaving behind a somewhat mixed pattern. This could indicate that a block-diagonal specification might model irrelevant correlations whereas a strict diagonal specification still ignores important information.

Both the diagonal and block-diagonal sparsity restrictions are imposed *ex ante* and are not changed during the analysis. In order to investigate whether the sparsity pattern is time-varying we plot the residual pattern for the 12 factor case at different points in time in panel (e). The results indicate that the pattern is rather constant over time.

5 Out-of-sample forecasting analysis

This section reports the out-of-sample forecasting results. Selected in-sample parameter estimation results for the proposed CFSS model and robustness checks are presented in Supplementary Appendix.

5.1 Implementation

We analyze the out-of-sample performance of three different factor structures: a 1-factor model with the market factor only (1F), a 3-factor model based on the Fama–French factors (3F), and a 12-factor model including the Fama–French factors plus the nine sector-specific ETFs (12F). For both the realized factor and residual covariance matrices we fit the unrestricted UE model (R0) and compare its performance to the restricted models (R1), (R2) and (R3) within the CFSS framework. The d.o.f. parameters k_i^b for the stochastic volatility process of the (latent) betas are fixed to their cross-sectional median values based on the in-sample estimates. For completeness, we consider diagonal (D) and block-diagonal residual assumptions based on the GICS classification (S) for each of the factor structures, although this is only (approximately) supported under the 12-factor specification. The diagonal specification sets $S = p$ and $p_i = 1 \forall i$, the block-diagonal specification sets $S = 9$ with sector sizes given in Table 1. Motivated by the findings of Fig. 3, we also consider post-prediction shrinkage of the sector-blocks toward their diagonal. We use a linear shrinkage (LS) approach akin to Ledoit and Wolf (2003), Ledoit and Wolf (2004). The LS prediction reads as

$$\widetilde{\widehat{C}}_{it+1}^e = \alpha_i \{ \widehat{C}_{it+1}^e \odot I_{p_i} \} + (1 - \alpha_i) \widehat{C}_{it+1}^e, \quad (21)$$

where \odot denotes the Hadamard product (see Lütkepohl 1996, p. 3). The optimal shrinkage intensity α_i^* is found by minimizing the expected Frobenius distance between the in-sample LS prediction and the realized residual blocks.¹⁶

The first three years of the data set are used as in-sample and the remainder as out-of-sample periods. We employ a rolling-window size of three years to obtain one-day ($h = 1$), 1-week ($h = 5$), 2-weeks ($h = 10$) and one-month ($h = 22$) ahead predictions. The daily forecasts are obtained by re-estimating each model on a daily basis and forming predictions via Eq. (20). For the multi-step ahead forecasts we rely on a direct forecasting approach, i.e., we use a horizon-specific estimated model where the dependent variable is the multi-period ahead value being predicted (Marcellino et al. 2006).¹⁷ Consequently, we predict the cumulative h -day ahead covariance matrix $C_{t+h}^r = \sum_{i=1}^h C_{t+i}^r$ by applying the CFSS models to the aggregated series $C_{t+h}^f = \sum_{i=1}^h C_{t+i}^f$, $C_{t+h}^e = \sum_{i=1}^h C_{t+i}^e$ and $B_{t+h} = \frac{1}{h} \sum_{i=1}^h B_{t+i}$ and again using Eq. (20). Following Chiriac and Voev (2011) we aggregate the realized components of C_t^r using non-overlapping h -day windows.

The sample period includes the financial crisis of 2008 as well as flash crashes in 2010 and 2011. These events lead to the presence of outliers in the (co)variance time-series which could distort parameter estimation results and affect out-of-sample predictions. In order to mitigate the effect of this comparably extreme events it is nowadays a conventional approach to perform a slight ex-post cleaning on the realized covariance matrices (Callot et al. 2017; Brito et al. 2018). As a means for cleaning the estimation sample we rely on the method proposed in Callot et al. (2017). Each day for which at least 25% of the unique elements of the realized covariance matrix exceed their sample mean up to then by more than four standard errors are flagged for censoring. The flagged matrices are then replaced by the sample average of their ten nearest non-flagged neighbors. In total 22 covariance matrices are flagged, which corresponds to only 1.46% of the whole sample.

It is worth to note that the fraction of flagged matrices can vary depending on the particular realized covariance estimator used. More sophisticated and noise robust realized covariance estimators could mitigate the need for ex-post data cleaning (see, e.g., Fan and Kim 2018). However, a thorough investigation of this problem is going beyond the scope of the paper.

5.2 Competing models

In order to compare the out-of-sample forecasting performance of the proposed composite factor state-space model we consider five alternative state-of-the-art forecasting approaches as benchmarks:

- (1) The Factor LASSO approach of Brito et al. (2018),

¹⁶ That is $\alpha_i^* = \arg \min_{\alpha_i} E[\|\alpha_i \{\widehat{C}_{it}^e \odot I_{p_i}\} + (1 - \alpha_i) \widehat{C}_{it}^e - C_{it}^e\|_F^2]$.

¹⁷ The direct forecasting approach is known to be more robust than the iterated approach as it is less prone to error propagation, which is especially relevant if there exists a model misspecification problem (Andersen et al. 2003; Chiriac and Voev 2011; Bollerslev et al. 2018; Luo and Chen 2020).

- (2) the Factor HEAVY (FHEAVY) model of Sheppard and Xu (2019),¹⁸
- (3) the Realized consistent DCC (Re-cDCC) model of Bauwens et al. (2016),
- (4) the Exponentially Weighted Moving Average (EWMA; Morgan 1996),
- (5) and the random walk (RW) model.

While the Factor LASSO and HEAVY models can be seen as natural competitors, the (non-factor) Re-cDCC is known to be a hard-to-beat benchmark when predicting realized asset covariance matrices. The EWMA and RW models can be seen as standard industry practice. Models (1) and (2) use the decomposition in Eq. (20) to predict C_{t+1}^r , models (3) to (5) are directly applied to the series of realized asset covariance matrices.

The LASSO (and also the EWMA and RW) approach can be implemented as described in Brito et al. (2018). To implement the HEAVY and Re-cDCC model, however, minor restrictions are required for the high-dimensional asset (and factor) setting as considered here. For the factor part in the multi-factor HEAVY model we implement a scalar version of the CAW model (see Golosnoy et al. 2012) and apply a covariance targeting approach (see, e.g., Noureldin et al. 2012) to get rid of the constant part. For the Re-cDCC model we consider the scalar specification for the underlying correlation dynamics. Potential unreliability in the parameter estimation arising from numerically calculating determinants and inverting huge dimensional matrices is circumvented by further restricting the correlation part using the dynamic equicorrelation (DECO) approach. In the DECO model both the determinant and inverse of the correlation matrix are given by closed-form expressions (see also Engle and Kelly 2012). The resulting model (sRe-cDECO hereafter) is advocated by Bauwens et al. (2016) as valuable competitor when handling vast dimensional systems.

Following, e.g., Callot et al. (2017), Opschoor et al. (2017) and Gribisch et al. (2020), we set the EWMA smoothing parameter to 0.96.

5.3 Statistical forecast evaluation

We assess the accuracy of the direct h -step-ahead point forecast with two types of statistical loss functions. First, the root mean squared error (RMSE) based on the Frobenius norm of the forecasting error that is calculated by comparing \widehat{C}_{t+h}^r and the ex-post observed value C_{t+h}^r , is considered (cf. Ledoit et al. 2003). This RMSE is given by

$$\text{RMSE} = \frac{1}{T^*} \sum_t \left[\sum_i (c_{it+h}^r - \widehat{c}_{it+h}^r)^2 + 2 \sum_{i < j} (c_{ijt+h}^r - \widehat{c}_{ijt+h}^r)^2 \right]^{\frac{1}{2}}, \quad (22)$$

where c_{it+h} and c_{ijt+h} denote the realized variance of asset i and the realized covariance between asset i and j , respectively. The corresponding forecasts are denoted by \widehat{c}_{it+h} and \widehat{c}_{ijt+h} , and T^* denotes the number of forecasting periods. In order to see whether a model performs different w.r.t. the different elements in the covariance

¹⁸ Since our focus lies on forecasting the realized measures, we ignore the HEAVY-P equations which link the realized covariance forecast to the conditional (latent) return covariance (see Sheppard and Xu 2019).

matrix we follow Gribisch et al. (2020) and disentangle the RMSE from above into the following two loss functions

$$\text{RMSE}^v = \frac{1}{T^*} \sum_t \left[\sum_i (c_{it+h}^r - \hat{c}_{it+h}^r)^2 \right]^{\frac{1}{2}}, \quad \text{RMSE}^c = \frac{1}{T^*} \sum_t \left[\sum_{i < j} (c_{ijt+h}^r - \hat{c}_{ijt+h}^r)^2 \right]^{\frac{1}{2}}, \quad (23)$$

where RMSE^v uses the variances only and RMSE^c considers the covariances. In addition to the RMSE loss we consider the commonly used QLIKE loss function, i.e.,

$$\text{QLIKE} = \frac{1}{T^*} \sum_t \log |\hat{C}_{t+h}^r| + \text{tr} \left[(\hat{C}_{t+h}^r)^{-1} C_{t+h}^r \right]. \quad (24)$$

While the RMSE losses are symmetric, the QLIKE loss function is an asymmetric loss, penalizing under-predictions more heavily (cf. Luo and Chen 2020). Both loss functions are known to be robust to noisy (co)variance proxies (Patton 2011; Laurent et al. 2013; Sheppard and Xu 2019).

In order to evaluate the statistical significance of differences in the RMSE, RMSE^v , RMSE^c and QLIKE losses across models, we employ the Model Confidence Set (MCS) approach of Hansen et al. (2011). The MCS is computed for the 75% confidence level using a block bootstrap with block length $\lfloor (T^*)^{1/3} \rfloor$ and 10,000 bootstrap replications.

Table 2 reports the RMSE and QLIKE results for the short-, mid- and long-term out-of-sample forecasts. We first compare the performances of the different CFSS model specifications. For one-step and five-step ahead forecasts the 12-factor CFSS model clearly outperforms the single and three-factor settings across all loss functions, regardless of the residual specification considered (though the losses suggest that block-diagonal specifications *S* and *LS* do increase the forecast precision). The biweekly and monthly results reveal that the three-factor models with diagonal and linear-shrinkage-based residuals perform best in predicting long-term covariance matrices, especially for the covariance elements as indicated by RMSE^c . In terms of QLIKE loss, however, including sector ETF-based factors still yields gains in forecast precision. Furthermore, we see that imposing restrictions on the smoothing parameter helps to increase predictive accuracy—while restriction (R2) solely convinces for the QLIKE measure, restriction (R1) performs overall well. Now, we turn to a comparison with the nine competing models. Most striking is that for the ‘full’ RMSE loss all competitors are significantly outperformed by the CFSS model specifications. In particular, for daily and weekly horizons solely the UE-R1-LS specification belongs to the 75% MCS.¹⁹ Similar results can be found for the QLIKE loss, where the UE-R2-LS specification significantly outperforms all competitors for daily forecasts. For multi-step ahead forecasts the sRe-cDECO model and further CFSS specifications, however, perform equally well. The EWMA and RW models are significantly outperformed in any case.

¹⁹ The 75% and 90% MCS are congruent in this case; however, the 90% MCS results are not reported here.

Table 2 continued

Model	$h = 1$			$h = 5$			$h = 10$			$h = 22$						
	RMSE	RMSE ^v	RMSE ^c	QLIKE	RMSE	RMSE ^v	RMSE ^c	QLIKE	RMSE	RMSE ^v	RMSE ^c	QLIKE	RMSE	RMSE ^v	RMSE ^c	QLIKE
UE-R2-D	136.83	32.08	92.69	289.44	115.82	24.79	78.99	132.99	119.44	24.32	81.68	82.74	138.03	27.13	94.72	45.86
UE-R2-S	137.38	32.02	93.26	281.76	118.19	24.46	81.04	131.13	124.76	23.68	86.06	81.57	146.61	25.79	101.59	45.20
UE-R2-LS	135.86	32.02	92.09	280.58	115.82	24.46	79.22	130.98	121.15	23.68	83.32	81.56	140.98	25.79	97.37	45.29
UE-R3-D	138.21	31.56	93.82	288.54	117.30	23.94	80.28	132.83	118.90	22.87	81.59	82.64	132.66	24.18	91.38	45.74
UE-R3-S	137.54	31.47	93.48	282.58	117.86	23.65	80.95	131.30	121.90	22.38	84.19	81.66	138.87	23.32	96.38	45.19
UE-R3-LS	136.57	31.47	92.73	280.58	116.36	23.65	79.80	131.00	119.51	22.38	82.38	81.55	134.79	23.32	93.30	45.19
<i>CFSS 12F</i>																
UE-R0-D	132.52	30.98	89.99	280.83	115.17	23.29	79.10	131.36	124.98	22.04	86.56	81.81	150.81	23.27	105.15	45.28
UE-R0-S	132.53	30.82	90.06	279.58	115.31	23.11	79.24	130.89	125.42	21.66	86.95	81.47	151.83	22.37	106.00	45.08
UE-R0-LS	132.34	30.82	89.92	278.41	115.10	23.11	79.09	130.75	125.13	21.66	86.74	81.42	151.34	22.37	105.65	45.07
UE-R1-D	131.81	30.59	89.56	281.76	112.90	22.80	77.56	131.57	120.45	21.23	83.44	81.93	142.63	21.41	99.54	45.31
UE-R1-S	131.60	30.53	89.44	281.10	112.73	22.73	77.45	131.19	120.55	21.08	83.53	81.63	143.32	21.34	100.04	45.15
UE-R1-LS	131.51	30.53	89.37	279.89	112.63	22.73	77.37	131.04	120.40	21.08	83.42	81.56	143.03	21.34	99.84	45.12
UE-R2-D	134.51	31.40	91.37	280.13	119.30	23.97	81.97	131.13	132.68	23.23	91.94	81.69	164.29	25.74	114.43	45.28
UE-R2-S	134.88	31.22	91.71	278.06	120.02	23.73	82.57	130.52	133.78	22.73	92.84	81.26	165.91	24.52	115.80	45.03
UE-R2-LS	134.56	31.22	91.48	277.23	119.61	23.73	82.27	130.43	133.23	22.73	92.44	81.24	165.06	24.52	115.19	45.04
UE-R3-D	132.72	31.02	90.13	280.72	115.72	23.37	79.49	131.31	125.97	22.16	87.25	81.80	152.68	23.26	106.48	45.27
UE-R3-S	132.78	30.87	90.24	279.40	115.97	23.20	79.71	130.82	126.54	21.80	87.74	81.42	153.86	22.55	107.43	45.06
UE-R3-LS	132.58	30.87	90.09	278.25	115.74	23.20	79.54	130.68	126.21	21.80	87.50	81.37	153.32	22.55	107.04	45.05
Factor Lasso 1F	142.07	30.45	97.03	284.22	117.07	22.98	80.56	131.12	124.51	26.80	84.96	82.24	221.70	86.69	135.58	48.32
Factor Lasso 3F	150.61	30.81	103.09	284.12	123.06	22.80	84.87	131.05	134.02	25.68	92.16	82.13	206.17	57.43	136.20	48.15
Factor Lasso 12F	136.67	30.74	93.09	282.83	113.67	22.80	78.16	130.66	229.97	31.39	160.57	81.95	280.82	92.48	181.21	48.03

Table 2 continued

Model	h = 1			h = 5			h = 10			h = 22						
	RMSE	RMSE ^v	RMSE ^c	QLIKE	RMSE	RMSE ^v	RMSE ^c	QLIKE	RMSE	RMSE ^v	RMSE ^c	QLIKE	RMSE	RMSE ^v	RMSE ^c	QLIKE
Factor HEAVY 1F	138.86	31.31	94.31	295.49	129.61	41.25	83.06	140.99	130.55	45.46	81.36	86.84	140.09	47.86	88.25	47.75
Factor HEAVY 3F	142.63	30.87	97.14	290.54	139.20	35.56	92.01	139.22	136.39	38.47	88.51	85.90	138.84	40.31	88.91	47.32
Factor HEAVY 12F	131.89	30.57	89.67	279.93	118.17	34.69	77.73	135.50	128.76	36.52	84.88	83.79	156.11	38.58	105.06	46.30
sRe-cDECO	174.45	33.39	120.49	279.06	157.91	26.57	109.82	130.48	166.63	26.05	116.24	81.23	189.63	28.83	132.45	45.14
EWMA	155.66	33.14	106.67	283.97	180.53	33.04	125.13	131.57	214.65	40.80	148.65	82.65	264.52	56.00	182.43	46.72
RW	154.12	37.66	104.48	328.31	124.78	26.86	85.41	135.56	131.79	24.09	91.05	83.14	158.97	23.68	110.91	45.58

The table reports the RMSE, RMSE^v, RMSE^c and QLIKE losses as given in Eqs. (22) to (24) for the forecast horizon $h \in \{1, 5, 10, 22\}$. To make the results more comparable the losses have been divided by h . Bold values indicate that the respective model belongs to the 75% model confidence set

Table 3 Daily GMVP forecasting results

Model	Unrestricted GMVP					Restricted GMVP: 150/50					Restricted GMVP: 130/30					Restricted GMVP: Long-only					
	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	
<i>CFSS 1F</i>																					
UE-R0-D	7.73	1.91	0.04	0.48	0.30	7.62	1.90	0.04	0.44	0.30	7.35	1.89	0.04	0.30	0.28	7.75	1.27	0.05	0.00	0.20	
UE-R0-S	6.62	2.54	0.06	0.68	0.60	6.66	2.47	0.06	0.48	0.53	6.74	2.33	0.05	0.30	0.45	7.78	1.54	0.05	0.00	0.30	
UE-R0-LS	6.66	2.50	0.06	0.62	0.54	6.69	2.43	0.05	0.47	0.48	6.74	2.30	0.05	0.30	0.42	7.74	1.52	0.05	0.00	0.28	
UE-R1-D	7.72	1.92	0.04	0.48	0.29	7.61	1.91	0.04	0.44	0.29	7.34	1.89	0.04	0.30	0.27	7.75	1.26	0.05	0.00	0.19	
UE-R1-S	6.62	2.55	0.06	0.68	0.57	6.66	2.47	0.06	0.48	0.50	6.74	2.34	0.05	0.30	0.42	7.78	1.55	0.05	0.00	0.28	
UE-R1-LS	6.65	2.51	0.06	0.64	0.52	6.68	2.45	0.05	0.48	0.47	6.74	2.31	0.05	0.30	0.40	7.75	1.53	0.05	0.00	0.27	
UE-R2-D	7.75	1.91	0.04	0.48	0.29	7.63	1.90	0.04	0.44	0.28	7.36	1.88	0.04	0.30	0.27	7.75	1.26	0.05	0.00	0.19	
UE-R2-S	6.61	2.55	0.06	0.67	0.57	6.66	2.47	0.06	0.48	0.50	6.74	2.33	0.05	0.30	0.42	7.78	1.54	0.05	0.00	0.28	
UE-R2-LS	6.67	2.48	0.05	0.61	0.49	6.70	2.43	0.05	0.47	0.45	6.75	2.29	0.05	0.30	0.39	7.73	1.52	0.05	0.00	0.26	
UE-R3-D	7.73	1.91	0.04	0.48	0.30	7.62	1.91	0.04	0.44	0.30	7.35	1.89	0.04	0.30	0.28	7.75	1.27	0.05	0.00	0.20	
UE-R3-S	6.62	2.55	0.06	0.68	0.60	6.66	2.47	0.06	0.48	0.53	6.74	2.33	0.05	0.30	0.45	7.78	1.54	0.05	0.00	0.30	
UE-R3-LS	6.66	2.49	0.06	0.62	0.53	6.69	2.43	0.05	0.47	0.48	6.74	2.30	0.05	0.30	0.42	7.74	1.52	0.05	0.00	0.28	
<i>CFSS 3F</i>																					
UE-R0-D	7.31	2.24	0.04	0.48	0.31	7.26	2.18	0.04	0.44	0.30	7.10	2.13	0.04	0.30	0.26	7.71	1.30	0.05	0.00	0.17	
UE-R0-S	6.61	2.53	0.06	0.68	0.57	6.66	2.49	0.06	0.48	0.49	6.74	2.36	0.05	0.30	0.40	7.76	1.56	0.05	0.00	0.26	
UE-R0-LS	6.64	2.49	0.06	0.63	0.51	6.68	2.45	0.05	0.48	0.45	6.74	2.34	0.05	0.30	0.38	7.72	1.54	0.05	0.00	0.25	
UE-R1-D	7.30	2.24	0.04	0.48	0.30	7.25	2.18	0.04	0.43	0.29	7.09	2.13	0.04	0.30	0.25	7.72	1.30	0.05	0.00	0.16	
UE-R1-S	6.61	2.53	0.06	0.68	0.54	6.66	2.49	0.06	0.48	0.46	6.74	2.37	0.05	0.30	0.38	7.76	1.56	0.05	0.00	0.25	
UE-R1-LS	6.63	2.50	0.06	0.64	0.49	6.68	2.47	0.05	0.48	0.44	6.74	2.35	0.05	0.30	0.37	7.73	1.54	0.05	0.00	0.24	
UE-R2-D	7.32	2.24	0.04	0.48	0.30	7.26	2.18	0.04	0.44	0.29	7.10	2.13	0.05	0.30	0.25	7.72	1.30	0.05	0.00	0.16	
UE-R2-S	6.61	2.53	0.06	0.68	0.53	6.66	2.49	0.06	0.48	0.46	6.74	2.37	0.05	0.30	0.38	7.76	1.56	0.05	0.00	0.24	

Table 3 continued

Model	Unrestricted GMVP				Restricted GMVP: 150/50				Restricted GMVP: 130/30				Restricted GMVP: Long-only							
	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO					
UE-R2-LS	6.65	2.48	0.06	0.62	0.46	6.69	2.45	0.05	0.47	0.42	6.74	2.33	0.05	0.30	0.35	7.71	1.53	0.05	0.00	0.23
UE-R3-D	7.31	2.24	0.04	0.48	0.31	7.26	2.18	0.04	0.44	0.30	7.10	2.13	0.04	0.30	0.26	7.71	1.30	0.05	0.00	0.17
UE-R3-S	6.61	2.53	0.06	0.68	0.57	6.66	2.49	0.06	0.48	0.49	6.74	2.36	0.05	0.30	0.40	7.76	1.56	0.05	0.00	0.26
UE-R3-LS	6.64	2.49	0.06	0.63	0.50	6.68	2.45	0.05	0.48	0.45	6.74	2.34	0.05	0.30	0.38	7.72	1.54	0.05	0.00	0.25
<i>CFSS 12F</i>																				
UE-R0-D	6.85	2.47	0.08	0.76	0.40	6.76	2.40	0.07	0.50	0.33	6.74	2.26	0.07	0.30	0.26	7.66	1.42	0.07	0.00	0.15
UE-R0-S	6.66	2.47	0.08	0.82	0.52	6.65	2.36	0.07	0.50	0.41	6.63	2.24	0.07	0.30	0.33	7.66	1.50	0.07	0.00	0.19
UE-R0-LS	6.68	2.47	0.08	0.79	0.49	6.66	2.37	0.07	0.50	0.39	6.64	2.24	0.07	0.30	0.31	7.65	1.48	0.07	0.00	0.18
UE-R1-D	6.85	2.47	0.07	0.76	0.38	6.76	2.41	0.07	0.50	0.32	6.74	2.26	0.07	0.30	0.25	7.67	1.41	0.07	0.00	0.14
UE-R1-S	6.67	2.47	0.08	0.82	0.50	6.65	2.36	0.07	0.50	0.39	6.63	2.24	0.07	0.30	0.31	7.66	1.50	0.07	0.00	0.18
UE-R1-LS	6.68	2.47	0.08	0.80	0.47	6.66	2.37	0.07	0.50	0.37	6.64	2.24	0.07	0.30	0.30	7.65	1.49	0.07	0.00	0.18
UE-R2-D	6.85	2.47	0.08	0.76	0.38	6.77	2.40	0.07	0.50	0.31	6.74	2.25	0.07	0.30	0.25	7.67	1.41	0.07	0.00	0.14
UE-R2-S	6.66	2.47	0.08	0.82	0.49	6.65	2.36	0.07	0.50	0.39	6.63	2.24	0.07	0.30	0.31	7.66	1.50	0.06	0.00	0.18
UE-R2-LS	6.68	2.47	0.08	0.79	0.45	6.66	2.37	0.07	0.50	0.36	6.64	2.23	0.07	0.30	0.29	7.65	1.48	0.06	0.00	0.17
UE-R3-D	6.85	2.47	0.08	0.76	0.40	6.76	2.40	0.07	0.50	0.33	6.74	2.26	0.07	0.30	0.26	7.67	1.42	0.07	0.00	0.15
UE-R3-S	6.66	2.47	0.08	0.82	0.53	6.65	2.36	0.07	0.50	0.41	6.63	2.24	0.07	0.30	0.33	7.66	1.50	0.07	0.00	0.19
UE-R3-LS	6.68	2.47	0.08	0.79	0.49	6.66	2.37	0.07	0.50	0.39	6.64	2.24	0.07	0.30	0.31	7.65	1.48	0.07	0.00	0.18
Factor Lasso 1F	7.12	2.58	0.06	0.69	0.70	7.02	2.53	0.05	0.50	0.64	6.88	2.39	0.05	0.30	0.57	7.61	1.40	0.05	0.00	0.41
Factor Lasso 3F	6.93	2.58	0.06	0.66	0.69	6.88	2.51	0.05	0.50	0.62	6.82	2.38	0.05	0.30	0.54	7.65	1.49	0.05	0.00	0.38
Factor Lasso 12F	7.17	2.42	0.08	0.84	0.75	6.97	2.38	0.07	0.50	0.59	6.84	2.32	0.07	0.30	0.48	7.68	1.46	0.06	0.00	0.29

Table 3 continued

Model	Unrestricted GMVP				Restricted GMVP: 150/50				Restricted GMVP: 130/30				Restricted GMVP: Long-only							
	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO					
Factor HEAVY 1F	7.89	1.92	0.04	0.51	0.30	7.74	1.91	0.04	0.46	0.30	7.37	1.86	0.05	0.30	0.28	7.71	1.26	0.05	0.00	0.21
Factor HEAVY 3F	7.43	2.17	0.04	0.49	0.32	7.34	2.13	0.04	0.44	0.31	7.11	2.10	0.04	0.30	0.28	7.70	1.29	0.05	0.00	0.20
Factor HEAVY 12F	6.94	2.45	0.07	0.73	0.48	6.83	2.44	0.07	0.50	0.41	6.81	2.29	0.06	0.30	0.33	7.64	1.43	0.06	0.00	0.19
sRe-cDECO	7.20	2.46	0.11	1.02	0.47	7.14	2.18	0.09	0.50	0.31	7.20	2.04	0.08	0.30	0.24	8.12	1.41	0.07	0.00	0.13
EWMA	6.52	2.27	0.09	0.92	0.23	6.54	2.05	0.07	0.50	0.17	6.65	1.95	0.07	0.30	0.13	7.74	1.28	0.06	0.00	0.06
RW	7.15	2.39	0.08	0.68	2.59	7.18	2.23	0.07	0.48	2.33	7.40	1.99	0.07	0.30	2.05	8.56	1.11	0.07	0.00	1.34

The table reports GMVP summary statistics for the four portfolios under consideration based on 754 one-step-ahead conditional covariance matrix predictions; σ_p stands for the annualized standard deviation of the portfolio returns. The average portfolio concentration CO, short position SP and turnover TO are given by Eqs. (28) to (30). Bold values indicate that the respective model belongs to the 90% model confidence set

Table 4 Weekly GMVP forecasting results

Model	Unrestricted GMVP					Restricted GMVP: 150/50					Restricted GMVP: 130/30					Restricted GMVP: Long-only					
	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	
<i>CFSS IF</i>																					
UE-R0-D	8.81	1.38	0.04	0.48	0.41	8.70	1.41	0.04	0.44	0.41	8.32	1.44	0.04	0.30	0.38	7.68	1.12	0.05	0.00	0.28	
UE-R0-S	7.34	2.02	0.06	0.66	0.68	7.32	2.09	0.05	0.48	0.60	7.25	2.01	0.05	0.30	0.52	7.69	1.33	0.05	0.00	0.35	
UE-R0-LS	7.40	1.97	0.05	0.62	0.61	7.39	2.04	0.05	0.48	0.56	7.30	1.98	0.05	0.30	0.49	7.67	1.33	0.05	0.00	0.33	
UE-R1-D	8.80	1.39	0.04	0.48	0.40	8.69	1.42	0.04	0.44	0.40	8.31	1.45	0.04	0.30	0.38	7.68	1.12	0.05	0.00	0.27	
UE-R1-S	7.34	2.02	0.06	0.66	0.66	7.32	2.09	0.05	0.48	0.59	7.25	2.01	0.05	0.30	0.51	7.70	1.34	0.05	0.00	0.34	
UE-R1-LS	7.39	1.98	0.05	0.62	0.61	7.38	2.05	0.05	0.48	0.56	7.29	1.99	0.05	0.30	0.48	7.69	1.33	0.05	0.00	0.33	
UE-R2-D	8.83	1.38	0.04	0.49	0.40	8.72	1.41	0.04	0.44	0.40	8.33	1.44	0.04	0.30	0.38	7.69	1.12	0.05	0.00	0.27	
UE-R2-S	7.35	2.01	0.06	0.67	0.66	7.34	2.08	0.05	0.49	0.59	7.26	2.00	0.05	0.30	0.50	7.69	1.33	0.05	0.00	0.34	
UE-R2-LS	7.45	1.95	0.05	0.61	0.58	7.43	2.01	0.05	0.47	0.53	7.33	1.96	0.05	0.30	0.47	7.66	1.32	0.05	0.00	0.32	
UE-R3-D	8.81	1.38	0.04	0.48	0.41	8.70	1.42	0.04	0.44	0.41	8.32	1.45	0.04	0.30	0.38	7.69	1.12	0.05	0.00	0.28	
UE-R3-S	7.34	2.02	0.06	0.66	0.69	7.33	2.09	0.05	0.48	0.61	7.25	2.01	0.05	0.30	0.52	7.69	1.34	0.05	0.00	0.35	
UE-R3-LS	7.41	1.97	0.05	0.62	0.61	7.40	2.03	0.05	0.48	0.56	7.30	1.98	0.05	0.30	0.49	7.67	1.33	0.05	0.00	0.34	
<i>CFSS 3F</i>																					
UE-R0-D	8.37	1.59	0.04	0.48	0.42	8.29	1.60	0.04	0.44	0.41	8.04	1.60	0.04	0.30	0.36	7.64	1.09	0.05	0.00	0.23	
UE-R0-S	7.32	2.00	0.06	0.66	0.63	7.29	2.07	0.05	0.49	0.55	7.23	2.01	0.05	0.30	0.46	7.69	1.32	0.05	0.00	0.29	
UE-R0-LS	7.37	1.96	0.05	0.62	0.57	7.35	2.03	0.05	0.48	0.51	7.27	1.98	0.05	0.30	0.43	7.66	1.31	0.05	0.00	0.28	
UE-R1-D	8.37	1.59	0.04	0.47	0.41	8.29	1.60	0.04	0.44	0.40	8.04	1.60	0.04	0.30	0.35	7.63	1.09	0.05	0.00	0.22	
UE-R1-S	7.32	2.00	0.06	0.66	0.61	7.29	2.07	0.05	0.49	0.53	7.23	2.01	0.05	0.30	0.44	7.70	1.32	0.05	0.00	0.29	
UE-R1-LS	7.36	1.97	0.05	0.62	0.57	7.34	2.04	0.05	0.48	0.51	7.26	1.99	0.05	0.30	0.43	7.69	1.31	0.05	0.00	0.28	
UE-R2-D	8.38	1.59	0.04	0.48	0.41	8.30	1.60	0.04	0.44	0.40	8.05	1.60	0.04	0.30	0.35	7.65	1.09	0.05	0.00	0.22	
UE-R2-S	7.33	1.99	0.06	0.66	0.61	7.30	2.06	0.05	0.49	0.53	7.23	2.00	0.05	0.30	0.44	7.69	1.31	0.05	0.00	0.28	

Table 4 continued

Model	Unrestricted GMVP				Restricted GMVP: 150/50				Restricted GMVP: 130/30				Restricted GMVP: Long-only								
	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	
UE-R2-LS	7.40	1.95	0.05	0.61	0.54	7.39	2.01	0.05	0.48	0.49	7.29	1.97	0.05	0.30	0.41	7.66	1.30	0.05	0.00	0.27	
UE-R3-D	8.37	1.59	0.04	0.48	0.42	8.29	1.60	0.04	0.44	0.41	8.04	1.60	0.04	0.30	0.36	7.64	1.09	0.05	0.00	0.23	
UE-R3-S	7.32	2.00	0.06	0.66	0.64	7.29	2.07	0.05	0.49	0.55	7.23	2.01	0.05	0.30	0.46	7.69	1.32	0.05	0.00	0.29	
UE-R3-LS	7.37	1.96	0.05	0.62	0.57	7.36	2.03	0.05	0.48	0.51	7.27	1.98	0.05	0.30	0.43	7.67	1.31	0.05	0.00	0.28	
<i>CFSS 12F</i>																					
UE-R0-D	7.80	1.83	0.07	0.75	0.49	7.72	1.87	0.07	0.50	0.41	7.51	1.85	0.07	0.30	0.34	7.66	1.26	0.07	0.00	0.19	
UE-R0-S	7.58	1.97	0.07	0.81	0.56	7.50	1.97	0.07	0.50	0.45	7.36	1.96	0.07	0.30	0.36	7.64	1.32	0.06	0.00	0.21	
UE-R0-LS	7.61	1.96	0.07	0.79	0.53	7.53	1.97	0.07	0.50	0.43	7.38	1.95	0.07	0.30	0.35	7.63	1.31	0.06	0.00	0.21	
UE-R1-D	7.81	1.83	0.07	0.75	0.48	7.73	1.88	0.07	0.50	0.40	7.51	1.86	0.07	0.30	0.33	7.66	1.27	0.07	0.00	0.19	
UE-R1-S	7.59	1.98	0.07	0.80	0.54	7.51	1.98	0.07	0.50	0.43	7.36	1.96	0.07	0.30	0.35	7.64	1.32	0.06	0.00	0.21	
UE-R1-LS	7.62	1.97	0.07	0.79	0.52	7.54	1.98	0.07	0.50	0.42	7.38	1.96	0.07	0.30	0.34	7.64	1.32	0.06	0.00	0.20	
UE-R2-D	7.80	1.84	0.07	0.76	0.48	7.72	1.87	0.07	0.50	0.40	7.51	1.85	0.07	0.30	0.32	7.67	1.26	0.07	0.00	0.19	
UE-R2-S	7.58	1.97	0.07	0.81	0.54	7.50	1.97	0.07	0.50	0.43	7.36	1.95	0.07	0.30	0.34	7.64	1.32	0.06	0.00	0.20	
UE-R2-LS	7.62	1.96	0.07	0.79	0.50	7.54	1.96	0.07	0.50	0.41	7.39	1.94	0.07	0.30	0.33	7.64	1.31	0.06	0.00	0.20	
UE-R3-D	7.80	1.83	0.07	0.75	0.49	7.72	1.87	0.07	0.50	0.42	7.51	1.85	0.07	0.30	0.34	7.66	1.26	0.07	0.00	0.19	
UE-R3-S	7.58	1.97	0.07	0.81	0.56	7.51	1.97	0.07	0.50	0.45	7.36	1.96	0.07	0.30	0.36	7.64	1.32	0.06	0.00	0.21	
UE-R3-LS	7.62	1.96	0.07	0.79	0.53	7.54	1.97	0.07	0.50	0.43	7.38	1.95	0.07	0.30	0.35	7.64	1.31	0.06	0.00	0.21	
Factor Lasso 1F	7.51	2.11	0.06	0.68	0.68	7.33	2.12	0.05	0.50	0.63	7.17	2.06	0.05	0.30	0.56	7.50	1.43	0.05	0.00	0.39	
Factor Lasso 3F	7.41	2.10	0.05	0.66	0.69	7.25	2.12	0.05	0.50	0.62	7.08	2.08	0.05	0.30	0.53	7.47	1.44	0.05	0.00	0.35	
Factor Lasso 12F	7.61	2.11	0.08	0.84	0.73	7.40	2.05	0.07	0.50	0.57	7.11	2.06	0.07	0.30	0.46	7.41	1.36	0.06	0.00	0.29	

Table 4 continued

Model	Unrestricted GMVP					Restricted GMVP: 150/50					Restricted GMVP: 130/30					Restricted GMVP: Long-only				
	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO
Factor HEAVY 1F	8.60	1.47	0.04	0.49	0.26	8.49	1.48	0.04	0.47	0.27	8.00	1.54	0.04	0.30	0.26	7.55	1.25	0.04	0.00	0.16
Factor HEAVY 3F	8.20	1.51	0.04	0.44	0.27	8.11	1.52	0.04	0.42	0.27	7.81	1.55	0.04	0.30	0.25	7.57	1.25	0.04	0.00	0.16
Factor HEAVY 12F	7.72	1.85	0.06	0.64	0.74	7.64	1.88	0.06	0.49	0.67	7.47	1.87	0.06	0.30	0.55	7.55	1.35	0.06	0.00	0.32
sRe-cDECO	7.10	2.48	0.11	1.00	0.69	7.04	2.28	0.09	0.50	0.46	7.11	2.08	0.08	0.30	0.35	7.59	1.42	0.07	0.00	0.19
EWMA	7.21	2.09	0.09	0.93	0.14	7.14	2.02	0.08	0.50	0.10	7.12	1.97	0.07	0.30	0.07	7.52	1.39	0.07	0.00	0.03
RW	7.43	1.65	0.09	0.93	2.67	7.37	1.73	0.07	0.50	2.04	7.39	1.73	0.07	0.30	1.67	7.77	1.34	0.07	0.00	0.97

The table reports GMVP summary statistics for the four portfolios under consideration based on 150 five-step-ahead conditional covariance matrix predictions; σ_p stands for the annualized standard deviation of the portfolio returns. The average portfolio concentration CO, short position SP and turnover TO are given by Eqs. (28) to (30). Bold values indicate that the respective model belongs to the 90% model confidence set

Further results, including additional loss functions and investigations of the model behavior under different market conditions (i.e., calm and volatile periods), are provided in Supplementary Appendix. These robustness checks confirm the above findings.

5.4 Portfolio construction

In this section we turn to an economic application in out-of-sample portfolio construction. We analyze the forecasting performance for the CFSS model and its competitors based on several variations of the global minimum variance (GMV) and mean-variance-optimal (MV) portfolios in the vein of Lunde et al. (2016) and Moura et al. (2020), imposing short-selling and boundary constraints on the weights to ensure stable and well-diversified portfolios.

For a given asset covariance forecast \widehat{C}_{t+h}^r computed in period t , the GMV portfolio is the solution to the minimization problem

$$\widehat{w}_{t+h} = \arg \min_{w_{t+h}} w'_{t+h} \widehat{C}_{t+h}^r w_{t+h}, \quad \text{s.t.} \quad w'_{t+h} \mathbf{1} = 1, \quad \|w_{t+h}\|_1 \leq 1 + 2s, \quad \|w_{t+h}\|_\infty \leq u, \tag{25}$$

where w_{t+h} is the $p \times 1$ vector of portfolio weights and $\mathbf{1}$ is a $p \times 1$ vector of ones. $\|w_{t+h}\|_1 = \sum_{i=1}^p |w_{it+h}|$, and $\|w_{t+h}\|_\infty = \max_{1 \leq i \leq p} |w_{it+h}|$ denote the L_1 and L_∞ norm of the weight vector, respectively. Here, $s \in [0, 1]$ denotes the percentage that is allowed to be held short, and $u > 0$ denotes the boundary for every distinct position (see Fan et al. 2012; Lunde et al. 2016, for more detailed discussions). Since the constraint minimization problem in Eq. (25) has no closed-form solution we rely on the CVX package in MATLAB of Grant and Boyd (2014) to solve for \widehat{w}_{t+h} .²⁰

To find the MV portfolio Eq. (25) is augmented by the additional constraint $w'_{t+h} m = \mu_{target}$, where m is a signal variable and μ_{target} is the target return. The signal is constructed using the momentum factor of Jegadeesh and Titman (1993), i.e., for each of the p stocks the individual momentum m_i is computed as the geometric average of the previous 252 returns, but excluding the 22 most recent returns. Collecting all the momentum in a vector yields the signal m . The target return is computed as the arithmetic average of the momentum of those stocks that belong to the top-quintile stocks ranked according to momentum (see also Engle et al. 2019; Moura et al. 2020).

For assessing the relative capabilities of the competing models for optimal portfolio allocation, we calculate their out-of-sample portfolio returns $r_t^p = \widehat{w}'_t r_t$ and report five measures, i.e., the portfolio standard deviation

²⁰ However, by dropping the norm constraints the minimization problem has the unique solution $\widehat{w}_{t+h} = (\widehat{C}_{t+h}^r)^{-1} \mathbf{1} / (\mathbf{1}' (\widehat{C}_{t+h}^r)^{-1} \mathbf{1})$ given \widehat{C}_{t+h}^r is positive definite. In large-scale applications as considered here, it might occur that the covariance prediction is not well conditioned, possibly resulting in unrealistically volatile weights due to the inversion of \widehat{C}_{t+h}^r . For the factor model predictions based on Eq. (20) this problem is mitigated. Due to the (block)diagonality of the residual covariance prediction, the p dimensional inverse problem collapses to $\max\{q, p_1, \dots, p_s\}$ dimensions by making use of the Woodbury matrix inversion lemma, i.e. $(C_t^r)^{-1} = (C_t^e)^{-1} - (C_t^e)^{-1} B_t [(C_t^f)^{-1} + B_t' (C_t^e)^{-1} B_t]^{-1} B_t' (C_t^e)^{-1}$ (see Lütkepohl 1996, pp. 29–30).

Table 5 Daily MVP with momentum signal forecasting results

Model	Unrestricted MVP				Restricted MVP: 150/50				Restricted MVP: 130/30				Restricted MVP: Long-only							
	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO					
<i>CFSS IF</i>																				
UE-R0-D	8.34	2.92	0.05	0.67	0.36	8.14	2.80	0.05	0.49	0.34	8.10	2.64	0.06	0.30	0.31	12.48	1.53	0.07	0.00	0.19
UE-R0-S	7.29	3.43	0.07	0.84	0.66	7.53	3.28	0.07	0.50	0.51	7.99	3.05	0.07	0.30	0.42	12.67	1.48	0.08	0.00	0.21
UE-R0-LS	7.35	3.39	0.07	0.79	0.59	7.55	3.26	0.06	0.50	0.48	7.97	3.04	0.06	0.30	0.40	12.63	1.49	0.08	0.00	0.20
UE-R1-D	8.34	2.93	0.05	0.66	0.35	8.14	2.81	0.05	0.49	0.33	8.09	2.65	0.06	0.30	0.30	12.48	1.52	0.07	0.00	0.19
UE-R1-S	7.29	3.43	0.07	0.84	0.63	7.53	3.29	0.07	0.50	0.49	7.98	3.06	0.07	0.30	0.40	12.67	1.48	0.08	0.00	0.20
UE-R1-LS	7.33	3.40	0.07	0.80	0.58	7.54	3.27	0.06	0.50	0.47	7.97	3.06	0.06	0.30	0.38	12.64	1.49	0.08	0.00	0.20
UE-R2-D	8.34	2.92	0.05	0.67	0.35	8.14	2.80	0.05	0.49	0.33	8.10	2.64	0.06	0.30	0.30	12.48	1.52	0.07	0.00	0.19
UE-R2-S	7.28	3.43	0.07	0.84	0.62	7.53	3.28	0.07	0.50	0.48	7.98	3.06	0.07	0.30	0.40	12.67	1.48	0.08	0.00	0.20
UE-R2-LS	7.35	3.38	0.07	0.78	0.54	7.55	3.26	0.06	0.50	0.45	7.95	3.05	0.06	0.30	0.37	12.62	1.49	0.08	0.00	0.20
UE-R3-D	8.34	2.92	0.05	0.67	0.36	8.14	2.80	0.05	0.49	0.34	8.10	2.64	0.06	0.30	0.31	12.48	1.53	0.07	0.00	0.19
UE-R3-S	7.29	3.43	0.07	0.84	0.66	7.54	3.28	0.07	0.50	0.52	7.99	3.05	0.07	0.30	0.42	12.67	1.48	0.08	0.00	0.21
UE-R3-LS	7.35	3.39	0.07	0.79	0.59	7.55	3.26	0.06	0.50	0.48	7.97	3.04	0.06	0.30	0.40	12.63	1.49	0.08	0.00	0.20
<i>CFSS 3F</i>																				
UE-R0-D	8.01	3.12	0.05	0.67	0.38	7.94	2.95	0.05	0.49	0.34	8.04	2.70	0.06	0.30	0.29	12.51	1.55	0.08	0.00	0.18
UE-R0-S	7.30	3.39	0.07	0.85	0.62	7.53	3.24	0.07	0.50	0.48	7.96	3.02	0.07	0.30	0.38	12.65	1.52	0.08	0.00	0.20
UE-R0-LS	7.34	3.36	0.07	0.80	0.56	7.53	3.24	0.06	0.50	0.45	7.94	3.01	0.06	0.30	0.37	12.62	1.53	0.08	0.00	0.19
UE-R1-D	8.01	3.12	0.05	0.66	0.36	7.94	2.95	0.05	0.49	0.32	8.04	2.70	0.06	0.30	0.28	12.52	1.54	0.08	0.00	0.18
UE-R1-S	7.29	3.39	0.07	0.84	0.59	7.52	3.25	0.07	0.50	0.46	7.95	3.03	0.07	0.30	0.37	12.65	1.51	0.08	0.00	0.19
UE-R1-LS	7.32	3.37	0.07	0.81	0.55	7.53	3.24	0.06	0.50	0.44	7.94	3.02	0.06	0.30	0.35	12.63	1.52	0.08	0.00	0.19
UE-R2-D	8.01	3.11	0.05	0.67	0.36	7.93	2.94	0.05	0.49	0.32	8.04	2.71	0.06	0.30	0.28	12.50	1.54	0.08	0.00	0.18

Table 5 continued

Model	Unrestricted MVP					Restricted MVP: 150/50					Restricted MVP: 130/30					Restricted MVP: Long-only					
	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	
UE-R2-S	7.29	3.40	0.07	0.85	0.59	7.52	3.25	0.07	0.50	0.45	7.95	3.03	0.07	0.30	0.36	12.65	1.51	0.08	0.00	0.19	
UE-R2-LS	7.34	3.35	0.07	0.79	0.52	7.54	3.24	0.06	0.50	0.42	7.93	3.02	0.06	0.30	0.34	12.61	1.53	0.08	0.00	0.19	
UE-R3-D	8.01	3.12	0.05	0.67	0.38	7.94	2.95	0.05	0.49	0.34	8.04	2.70	0.06	0.30	0.29	12.51	1.55	0.08	0.00	0.18	
UE-R3-S	7.30	3.39	0.07	0.85	0.62	7.53	3.25	0.07	0.50	0.48	7.96	3.02	0.07	0.30	0.39	12.65	1.52	0.08	0.00	0.20	
UE-R3-LS	7.34	3.36	0.07	0.80	0.56	7.53	3.24	0.06	0.50	0.45	7.94	3.01	0.06	0.30	0.37	12.62	1.53	0.08	0.00	0.19	
<i>CFSS 12F</i>																					
UE-R0-D	7.48	3.41	0.09	0.92	0.46	7.58	3.21	0.08	0.50	0.35	7.87	2.86	0.08	0.30	0.28	12.48	1.54	0.08	0.00	0.17	
UE-R0-S	7.33	3.39	0.09	0.98	0.59	7.48	3.21	0.08	0.50	0.41	7.83	2.93	0.08	0.30	0.33	12.51	1.53	0.08	0.00	0.18	
UE-R0-LS	7.34	3.40	0.09	0.95	0.55	7.49	3.23	0.08	0.50	0.40	7.82	2.92	0.08	0.30	0.32	12.50	1.53	0.08	0.00	0.18	
UE-R1-D	7.49	3.42	0.09	0.92	0.44	7.58	3.22	0.08	0.50	0.33	7.87	2.86	0.08	0.30	0.27	12.49	1.54	0.08	0.00	0.17	
UE-R1-S	7.33	3.39	0.09	0.98	0.56	7.48	3.22	0.08	0.50	0.40	7.83	2.93	0.08	0.30	0.32	12.51	1.53	0.08	0.00	0.17	
UE-R1-LS	7.34	3.39	0.09	0.96	0.53	7.49	3.23	0.08	0.50	0.38	7.82	2.93	0.08	0.30	0.31	12.50	1.53	0.08	0.00	0.17	
UE-R2-D	7.48	3.41	0.09	0.92	0.44	7.57	3.21	0.08	0.50	0.33	7.86	2.86	0.08	0.30	0.27	12.48	1.54	0.08	0.00	0.17	
UE-R2-S	7.32	3.39	0.09	0.98	0.55	7.47	3.22	0.08	0.50	0.39	7.82	2.93	0.08	0.30	0.31	12.51	1.53	0.08	0.00	0.17	
UE-R2-LS	7.34	3.40	0.09	0.95	0.51	7.48	3.23	0.08	0.50	0.37	7.82	2.92	0.08	0.30	0.30	12.50	1.53	0.08	0.00	0.17	
UE-R3-D	7.48	3.41	0.09	0.92	0.46	7.58	3.21	0.08	0.50	0.35	7.87	2.86	0.08	0.30	0.28	12.48	1.55	0.08	0.00	0.17	
UE-R3-S	7.33	3.39	0.09	0.98	0.59	7.48	3.21	0.08	0.50	0.41	7.83	2.93	0.08	0.30	0.33	12.51	1.53	0.08	0.00	0.18	
UE-R3-LS	7.34	3.40	0.09	0.95	0.55	7.49	3.23	0.08	0.50	0.40	7.82	2.92	0.08	0.30	0.32	12.50	1.53	0.08	0.00	0.18	
Factor Lasso 1F	7.73	3.38	0.07	0.82	0.74	7.77	3.27	0.06	0.50	0.62	7.97	3.01	0.06	0.30	0.53	12.57	1.48	0.08	0.00	0.25	

Table 5 continued

Model	Unrestricted MVP						Restricted MVP: 150/50						Restricted MVP: 130/30						Restricted MVP: Long-only						
	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO
Factor Lasso 3F	7.58	3.38	0.07	0.81	0.73	7.68	3.27	0.06	0.50	0.60	7.95	2.98	0.06	0.30	0.51	12.62	1.49	0.08	0.00	0.24	12.62	1.49	0.08	0.00	0.24
Factor Lasso 12F	7.77	3.21	0.09	0.97	0.79	7.74	3.08	0.08	0.50	0.57	8.00	2.79	0.08	0.30	0.46	12.64	1.49	0.08	0.00	0.21	12.64	1.49	0.08	0.00	0.21
Factor HEAVY 1F	8.48	2.81	0.05	0.68	0.35	8.22	2.73	0.05	0.49	0.33	8.10	2.58	0.06	0.30	0.30	12.43	1.49	0.07	0.00	0.19	12.43	1.49	0.07	0.00	0.19
Factor HEAVY 3F	8.13	2.96	0.05	0.66	0.36	8.02	2.86	0.05	0.49	0.33	8.09	2.62	0.06	0.30	0.29	12.48	1.47	0.07	0.00	0.18	12.48	1.47	0.07	0.00	0.18
Factor HEAVY 12F	7.56	3.32	0.08	0.89	0.52	7.59	3.16	0.07	0.50	0.40	7.82	2.90	0.07	0.30	0.32	12.42	1.54	0.08	0.00	0.18	12.42	1.54	0.08	0.00	0.18
sRe-cDECO	7.86	3.02	0.12	1.16	0.51	8.00	2.83	0.09	0.50	0.33	8.29	2.72	0.08	0.30	0.27	12.78	1.47	0.08	0.00	0.14	12.78	1.47	0.08	0.00	0.14
EWMA	7.12	3.37	0.10	1.10	0.29	7.43	3.04	0.08	0.50	0.20	7.83	2.88	0.08	0.30	0.17	12.41	1.50	0.08	0.00	0.14	12.41	1.50	0.08	0.00	0.14
RW	7.85	3.43	0.09	0.87	2.73	7.94	3.37	0.08	0.50	2.27	8.34	3.08	0.08	0.30	1.89	12.90	1.54	0.08	0.00	0.77	12.90	1.54	0.08	0.00	0.77

The table reports MVP with momentum signal summary statistics for the four portfolios under consideration based on 754 one-step-ahead conditional covariance matrix predictions; σ_p stands for the annualized standard deviation of the portfolio returns. The average portfolio concentration CO, short position SP and turnover TO are given by Eqs. (28) to (30). Bold values indicate that the respective model belongs to the 90% model confidence set

Table 6 Weekly MVP with momentum signal forecasting results

Model	Unrestricted MVP				Restricted MVP: 150/50				Restricted MVP: 130/30				Restricted MVP: Long-only							
	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO					
<i>CFSS IF</i>																				
UE-R0-D	9.45	2.31	0.05	0.67	0.52	8.98	2.31	0.05	0.49	0.50	8.67	2.27	0.06	0.30	0.47	12.03	1.44	0.07	0.00	0.34
UE-R0-S	7.91	2.95	0.07	0.83	0.78	8.03	2.85	0.06	0.50	0.63	8.10	2.76	0.06	0.30	0.53	12.06	1.40	0.08	0.00	0.35
UE-R0-LS	8.00	2.89	0.06	0.78	0.71	8.09	2.80	0.06	0.50	0.60	8.14	2.72	0.06	0.30	0.51	12.05	1.41	0.08	0.00	0.35
UE-R1-D	9.44	2.31	0.05	0.66	0.51	8.97	2.32	0.05	0.49	0.49	8.65	2.28	0.06	0.30	0.46	12.03	1.44	0.07	0.00	0.34
UE-R1-S	7.92	2.95	0.07	0.82	0.76	8.03	2.85	0.06	0.50	0.62	8.10	2.76	0.06	0.30	0.52	12.08	1.40	0.08	0.00	0.35
UE-R1-LS	7.98	2.91	0.06	0.79	0.70	8.08	2.82	0.06	0.50	0.59	8.14	2.73	0.06	0.30	0.51	12.07	1.41	0.08	0.00	0.34
UE-R2-D	9.46	2.29	0.05	0.67	0.51	8.99	2.30	0.05	0.49	0.50	8.68	2.26	0.06	0.30	0.46	12.03	1.44	0.07	0.00	0.34
UE-R2-S	7.93	2.94	0.07	0.83	0.76	8.04	2.84	0.06	0.50	0.61	8.12	2.75	0.06	0.30	0.52	12.05	1.40	0.08	0.00	0.35
UE-R2-LS	8.05	2.86	0.06	0.77	0.67	8.13	2.78	0.06	0.50	0.57	8.17	2.70	0.06	0.30	0.50	12.04	1.41	0.08	0.00	0.34
UE-R3-D	9.45	2.31	0.05	0.67	0.53	8.98	2.31	0.05	0.49	0.51	8.66	2.27	0.06	0.30	0.47	12.03	1.44	0.07	0.00	0.34
UE-R3-S	7.92	2.95	0.07	0.83	0.78	8.03	2.85	0.06	0.50	0.63	8.10	2.76	0.06	0.30	0.53	12.06	1.40	0.08	0.00	0.35
UE-R3-LS	8.01	2.89	0.06	0.78	0.71	8.10	2.80	0.06	0.50	0.60	8.15	2.72	0.06	0.30	0.51	12.05	1.41	0.08	0.00	0.35
<i>CFSS 3F</i>																				
UE-R0-D	9.15	2.39	0.05	0.66	0.54	8.80	2.38	0.05	0.49	0.49	8.64	2.29	0.06	0.30	0.44	12.05	1.44	0.07	0.00	0.33
UE-R0-S	7.94	2.90	0.07	0.83	0.73	8.06	2.81	0.06	0.50	0.59	8.14	2.70	0.06	0.30	0.49	12.04	1.41	0.08	0.00	0.34
UE-R0-LS	8.01	2.85	0.06	0.79	0.67	8.10	2.78	0.06	0.50	0.56	8.18	2.67	0.06	0.30	0.47	12.03	1.42	0.08	0.00	0.34
UE-R1-D	9.14	2.39	0.05	0.66	0.52	8.79	2.38	0.05	0.49	0.48	8.62	2.29	0.06	0.30	0.43	12.06	1.45	0.07	0.00	0.33
UE-R1-S	7.94	2.90	0.07	0.82	0.71	8.05	2.81	0.06	0.50	0.57	8.14	2.70	0.06	0.30	0.48	12.06	1.41	0.08	0.00	0.34
UE-R1-LS	7.99	2.87	0.07	0.79	0.67	8.09	2.79	0.06	0.50	0.55	8.18	2.68	0.06	0.30	0.46	12.05	1.42	0.08	0.00	0.34
UE-R2-D	9.16	2.38	0.05	0.66	0.53	8.81	2.37	0.05	0.49	0.49	8.66	2.28	0.06	0.30	0.43	12.05	1.44	0.08	0.00	0.33
UE-R2-S	7.95	2.90	0.07	0.83	0.71	8.07	2.80	0.06	0.50	0.57	8.15	2.69	0.06	0.30	0.48	12.04	1.41	0.08	0.00	0.34

Table 6 continued

Model	Unrestricted MVP					Restricted MVP: 150/50					Restricted MVP: 130/30					Restricted MVP: Long-only					
	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	
UE-R2-LS	8.04	2.83	0.06	0.78	0.64	8.13	2.77	0.06	0.50	0.54	8.21	2.65	0.06	0.30	0.46	12.02	1.42	0.08	0.00	0.33	
UE-R3-D	9.15	2.39	0.05	0.66	0.54	8.80	2.38	0.05	0.49	0.50	8.63	2.29	0.06	0.30	0.44	12.05	1.45	0.07	0.00	0.33	
UE-R3-S	7.94	2.90	0.07	0.83	0.74	8.06	2.81	0.06	0.50	0.59	8.14	2.70	0.06	0.30	0.49	12.05	1.41	0.08	0.00	0.34	
UE-R3-LS	8.01	2.85	0.06	0.78	0.67	8.11	2.78	0.06	0.50	0.56	8.19	2.67	0.06	0.30	0.47	12.03	1.42	0.08	0.00	0.34	
<i>CFSS 12F</i>																					
UE-R0-D	8.43	2.70	0.08	0.91	0.61	8.47	2.69	0.08	0.50	0.49	8.58	2.53	0.08	0.30	0.42	12.03	1.40	0.08	0.00	0.32	
UE-R0-S	8.14	2.89	0.09	0.96	0.68	8.29	2.79	0.08	0.50	0.52	8.36	2.68	0.08	0.30	0.43	12.01	1.40	0.08	0.00	0.32	
UE-R0-LS	8.18	2.87	0.08	0.94	0.65	8.31	2.78	0.07	0.50	0.50	8.39	2.65	0.08	0.30	0.43	12.01	1.40	0.08	0.00	0.32	
UE-R1-D	8.42	2.70	0.08	0.90	0.60	8.47	2.69	0.08	0.50	0.48	8.57	2.54	0.08	0.30	0.41	12.03	1.41	0.08	0.00	0.31	
UE-R1-S	8.15	2.89	0.09	0.96	0.66	8.30	2.79	0.07	0.50	0.51	8.36	2.68	0.08	0.30	0.42	12.01	1.40	0.08	0.00	0.32	
UE-R1-LS	8.19	2.88	0.08	0.94	0.64	8.31	2.79	0.07	0.50	0.50	8.39	2.66	0.08	0.30	0.42	12.01	1.40	0.08	0.00	0.31	
UE-R2-D	8.42	2.70	0.09	0.92	0.60	8.47	2.69	0.08	0.50	0.48	8.59	2.53	0.08	0.30	0.41	12.03	1.40	0.08	0.00	0.31	
UE-R2-S	8.14	2.89	0.09	0.96	0.66	8.29	2.79	0.08	0.50	0.50	8.36	2.67	0.08	0.30	0.42	12.01	1.40	0.08	0.00	0.31	
UE-R2-LS	8.18	2.86	0.08	0.94	0.62	8.31	2.78	0.07	0.50	0.49	8.40	2.64	0.08	0.30	0.41	12.01	1.40	0.08	0.00	0.31	
UE-R3-D	8.42	2.70	0.08	0.91	0.61	8.47	2.69	0.08	0.50	0.49	8.58	2.54	0.08	0.30	0.42	12.03	1.40	0.08	0.00	0.32	
UE-R3-S	8.15	2.89	0.09	0.96	0.68	8.29	2.79	0.08	0.50	0.52	8.36	2.68	0.08	0.30	0.43	12.01	1.40	0.08	0.00	0.32	
UE-R3-LS	8.18	2.87	0.08	0.94	0.65	8.31	2.78	0.07	0.50	0.51	8.39	2.65	0.08	0.30	0.43	12.01	1.40	0.08	0.00	0.32	
Factor Lasso 1F	8.09	2.96	0.07	0.82	0.74	8.03	2.88	0.06	0.50	0.64	8.03	2.77	0.06	0.30	0.55	12.08	1.35	0.08	0.00	0.35	
Factor Lasso 3F	8.01	2.95	0.07	0.81	0.75	7.97	2.90	0.06	0.50	0.63	8.01	2.79	0.06	0.30	0.53	12.06	1.39	0.08	0.00	0.34	
Factor Lasso 12F	8.16	3.00	0.09	0.97	0.80	8.03	2.93	0.08	0.50	0.59	8.06	2.71	0.08	0.30	0.49	11.97	1.41	0.08	0.00	0.31	

Table 6 continued

Model	Unrestricted MVP						Restricted MVP: 150/50						Restricted MVP: 130/30						Restricted MVP: Long-only					
	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP	TO	σ_p	SR	CO	SP
Factor HEAVY 1F	9.02	2.28	0.05	0.66	0.38	8.81	2.22	0.05	0.50	0.38	8.62	2.01	0.05	0.30	0.36	12.34	1.32	0.07	0.00	0.29				
Factor HEAVY 3F	8.78	2.25	0.05	0.62	0.37	8.65	2.20	0.05	0.49	0.37	8.61	1.99	0.05	0.30	0.35	12.31	1.28	0.07	0.00	0.28				
Factor HEAVY 12F	8.17	2.66	0.07	0.80	0.78	8.26	2.57	0.07	0.50	0.65	8.43	2.31	0.07	0.30	0.54	12.02	1.43	0.08	0.00	0.33				
sRe-cDECO	7.66	3.14	0.12	1.15	0.79	7.84	2.87	0.09	0.50	0.52	8.20	2.58	0.08	0.30	0.43	12.14	1.37	0.08	0.00	0.29				
EWMA	7.65	3.09	0.10	1.09	0.32	7.90	2.89	0.08	0.50	0.28	8.13	2.68	0.08	0.30	0.26	12.09	1.43	0.08	0.00	0.28				
RW	7.98	2.80	0.11	1.11	2.83	8.15	2.78	0.08	0.50	1.93	8.13	2.75	0.08	0.30	1.51	12.10	1.46	0.08	0.00	0.59				

The table reports MVP with momentum signal summary statistics for the four portfolios under consideration based on 150 five-step-ahead conditional covariance matrix predictions; σ_p stands for the annualized standard deviation of the portfolio returns. The average portfolio concentration CO, short position SP and turnover TO are given by Eqs. (28) to (30). Bold values indicate that the respective model belongs to the 90% model confidence set

$$\sigma_p = \sqrt{\frac{1}{T^*} \sum_t (r_t^p - \mu_p)^2}, \quad \text{with } \mu_p = \frac{1}{T^*} \sum_t r_t^p, \quad (26)$$

the Sharpe ratio

$$\text{SR} = \mu_p / \sigma_p, \quad (27)$$

the average portfolio concentration

$$\text{CO} = \frac{1}{T^*} \sum_t \left(\sum_{i=1}^p \hat{w}_{it}^2 \right)^{1/2}, \quad (28)$$

the average total short positions

$$\text{SP} = \frac{1}{T^*} \sum_t \sum_{i=1}^p |w_{it}| \mathbb{I}_{w_{it} < 0}, \quad (29)$$

and the average turnover rates

$$\text{TO} = \frac{1}{T^*} \sum_t \left| \hat{w}_{it} - \hat{w}_{it-1} \frac{1 + r_{it-1}}{1 + r_{t-1}^p} \right|. \quad (30)$$

The reported portfolio standard deviations and Sharpe ratios are annualized by multiplication with $\sqrt{[252/h]}$. Statistical significance of the differences in portfolio standard deviations is assessed by the MCS approach using the squared demeaned portfolio returns as loss series. For a more detailed discussion of the measures in Eqs. (28)–(30) the reader is referred to Hautsch et al. (2011), Bauwens et al. (2016), Bollerslev et al. (2018).

In the present paper we follow Brito et al. (2018) and Lunde et al. (2016) and consider three restricted portfolio construction problems, i.e., a *150/50 portfolio* ($s = 0.50$), a *130/30 portfolio* ($s = 0.30$) and a *long-only portfolio* ($s = 0.00$). The latter portfolios are constructed using a boundary value of $u = 0.10$ for each weight, such that at least ten assets are included. In addition, unrestricted (G)MV portfolios are considered. We focus on daily allocations and consider weekly horizons as robustness check.

The daily GMV portfolio allocation results are collected in Table 3. For the unrestricted portfolio the EWMA has the lowest standard deviation. However, the 1F-S and 3F-S CFSS models cannot be outperformed significantly while yielding higher Sharpe ratios and more diversified portfolios with less extreme short positions. By imposing 50% short-selling constraints the performance of the 12F- and LS-based residual models improves with comparably lower turnover for the CFSS models. By imposing stronger (30% and 0%) short-selling constraints the portfolio standard deviations get hardly distinguishable. Though comparing the CFSS model specifications to the LASSO and HEAVY competitors, we find better performance regarding the

SR and *TO* results for the CFSS. Overall, the lowest turnover is obtained with the EWMA. The highest turnover is obtained for the random walk, followed by the Factor LASSO specifications. The weekly portfolio results are summarized in Table 4. They indicate similar allocations. The lowest standard deviations are now obtained by the sRe-cDECO and EWMA. However, the CFSS with block residual belongs to the MCS in every scenario, whereas the factor HEAVY is only included for long-only portfolios.

The daily MV portfolio allocation results are collected in Table 5. The results are qualitatively similar to those obtained in the GMV exercise. For the unrestricted allocations the EWMA shows the lowest standard deviation. The second best models included in the MCS are the CFSS 1F-R1-S, 1F-R2-S and 3F-R2-S specifications, respectively. Compared to the EWMA all three have higher *SR* as well as lower *CO* and *SP*. Imposing short-selling constraints ameliorates the performance of the 12F CFSS model specifications in terms of lower standard deviations and lower turnover values than the 1F and 3F specifications. As robustness check, Table 6 shows weekly allocation results. The results do not differ substantially in comparison with the daily allocations.

6 Conclusion

In this paper we propose a composite factor state-space (CFSS) model for the prediction of vast-dimensional realized covariance matrices of asset returns. The model exploits an observed factor structure of the realized covariances of asset returns. Its components are found by a block LDL decomposition of the joint realized covariance matrix of the observed risk factors and assets. This yields individual time-series for the realized factor covariances, the realized residual covariances and the matrices of realized factor loadings. An assumption of independence is made, which admits a straightforward factorization of the likelihood function, reducing model complexity and allowing for fast parameter estimation and prediction of asset covariance matrices in scalable dimensions.

We apply the factor model to a data set of daily realized covariance matrices for 225 NYSE-traded stocks. The observed risk factors we consider are the CAPM market factor, the Fama and French (1993) HML and SMB factors as well as sector-specific ETFs. It turns out that for justifying an approximate factor structure based on sector-blocks for the residual components it is critical to include all those factors.

In an extensive out-of-sample application including point forecasts and predictions of the global minimum variance and mean variance optimal portfolios the CFSS model shows superior forecasting performance and outperforms several competing models in almost all forecast horizons. We conclude that the CFSS approach is a valuable tool for modeling and forecasting time-series of vast-dimensional realized covariance matrices.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00181-022-02245-1>.

Acknowledgements I want to thank participants of the 40th International Symposium on Forecasting (Virtual ISF 2020), Laura Reh, Simon Umbach and Hera Reh for valuable comments and steady support. I also thank the editor Prof. Kunst and two anonymous referees for useful comments, which helped to improve the paper.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The author declares that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

A Further derivations for the UE model

In the following we derive several results for the model at hand, which were not considered in the original work of Windle and Carvalho (2014). First, we derive a closed-form formula for the second-order predictive moment, which, e.g., enables the calculation of confidence intervals for the point prediction of the covariance elements. Second, to improve the interpretability of the UE we derive the underlying dynamic process for the latent integrated covariance matrix, as implied by the Matrix-Beta transition for the integrated precision. Third, we derive the autoregressive dynamics of the logarithmic determinant of the latent precision process. For example, Philipov and Glickman (2006) use these dynamics to deduce stationarity conditions for their matrix process.

For notational convenience reconsider the UE model without any superscript indicating correspondence to the factor or residual part. For any realized covariance C_t of dimension $d \times d$ the UE model is written as

$$C_t \mid \Sigma_t \sim \mathcal{W}_d(n, \Sigma_t/n), \tag{31}$$

$$\Omega_t = \mathcal{U}(\Omega_{t-1})' \Psi_t \mathcal{U}(\Omega_{t-1}) / \lambda, \quad \Psi_t \sim \mathcal{B}_d^I\left(\frac{k}{2}, \frac{n}{2}\right), \tag{32}$$

where $\Omega_t = \Sigma_t^{-1}$, with initial condition $\Omega_1 \sim \mathcal{W}_d(k, (nS_0)^{-1}/\lambda)$. In general the updating and filtering distributions for Ω_t are given by $\Omega_t \mid C_{1:t} \sim \mathcal{W}_d(n+k, (nS_t)^{-1})$ and $\Omega_{t+1} \mid C_{1:t} \sim \mathcal{W}_d(k, (nS_t)^{-1}/\lambda)$, respectively, where $S_t = \lambda S_{t-1} + C_t$ (see Windle and Carvalho 2014, Prop. 1).

A.1 Second-order predictive moments

The first-order predictive moment of C_{t+1} given the information set $C_{1:t}$ is readily obtained by the law of iterated expectations (Konno 1991; Windle and Carvalho 2014)

$$E[C_{t+1} | C_{1:t}] = E[E[C_{t+1} | \Sigma_{t+1}] | C_{1:t}] = E[\Omega_{t+1}^{-1} | C_{1:t}] = \frac{\lambda n}{k - d - 1} S_t.$$

It exists for $k > d + 1$. The second-order predictive moment, i.e., the covariance of $\text{vec}(C_{t+1})$ given $C_{1:t}$, can be found using the law of total variance and the results in Muirhead (2005) and von Rosen (1988) for the moments of (inverted) Wishart matrices

$$\begin{aligned} & \text{Cov}[\text{vec}(C_{t+1}) | C_{1:t}] \\ &= E[\text{Cov}[\text{vec}(C_{t+1}) | \Sigma_{t+1}] | C_{1:t}] + \text{Cov}[E[\text{vec}(C_{t+1}) | \Sigma_{t+1}] | C_{1:t}] \\ &= E\left[n^{-1}(I_{d^2} + K_{dd})(\Omega_{t+1} \otimes \Omega_{t+1}) | C_{1:t}\right] + \text{Cov}[\text{vec}(\Omega_{t+1}) | C_{1:t}] \\ &= \frac{1}{n}(I_{d^2} + K_{dd})E[\Omega_{t+1} \otimes \Omega_{t+1} | C_{1:t}] + E[\text{vec}(\Omega_{t+1})\text{vec}(\Omega_{t+1})' | C_{1:t}] \\ &\quad - E[\text{vec}(\Omega_{t+1}) | C_{1:t}]E[\text{vec}(\Omega_{t+1})' | C_{1:t}]' \\ &= \frac{(k + n - d - 1)}{n(k - d)(k - d - 1)(k - d - 3)} \left[\lambda^2 n^2 (I_{d^2} + K_{dd})(S_t \otimes S_t) \right. \\ &\quad \left. + \frac{2\lambda^2 n^2}{(k - d - 1)} \text{vec}(S_t)\text{vec}(S_t)' \right] \\ &= \frac{\lambda^2 n(k + n - d - 1)}{(k - d)(k - d - 1)(k - d - 3)} \left[(I_{d^2} + K_{dd})(S_t \otimes S_t) + \frac{2}{(k - d - 1)} \text{vec}(S_t)\text{vec}(S_t)' \right] \end{aligned}$$

where K_{dd} denotes the $d^2 \times d^2$ commutation matrix with $K_{dd}\text{vec}(A) = \text{vec}(A')$ (see Lütkepohl 1996, Chapter 9). The second-order moment exists if $k > d + 3$.

A.2 Underlying latent dynamic process for the integrated covariance

The Matrix-Beta type-I transition equation in Eq. (32) implies a shifted Matrix-F transition for the underlying latent integrated covariance matrix Σ_t itself. Defining $\Omega_t^{-1} = \Sigma_t$, we have

$$\begin{aligned} \Omega_t^{-1} &= \lambda(\mathcal{U}(\Omega_{t-1}))^{-1}\Psi_t^{-1}(\mathcal{U}(\Omega_{t-1}))^{-1} \\ &= \lambda(\mathcal{U}(\Omega_{t-1}))^{-1}(\Psi_t^{-1} - I_d + I_d)(\mathcal{U}(\Omega_{t-1}))^{-1} \\ &= \lambda(\mathcal{U}(\Omega_{t-1}))^{-1}(\Psi_t^{-1} - I_d)(\mathcal{U}(\Omega_{t-1}))^{-1} + \lambda(\mathcal{U}(\Omega_{t-1})'\mathcal{U}(\Omega_{t-1}))^{-1} \\ &= \lambda(\mathcal{U}(\Omega_{t-1}))^{-1}\Upsilon_t(\mathcal{U}(\Omega_{t-1}))^{-1} + \lambda\Omega_{t-1}^{-1}, \end{aligned}$$

where $\Upsilon_t = (\Psi_t^{-1} - I_d)$ follows a d -dimensional Matrix-Beta type-II distribution (also known as Matrix-F distribution), say $\Upsilon_t \sim \mathcal{B}_d^{II}(\frac{n}{2}, \frac{k}{2})$, with expectation $E[\Upsilon_t] = \frac{n}{k-d-1}I_d$ (Gupta and Nagar 2000, Theorems 5.3.5 and 5.3.20), and I_d denotes the d -dimensional identity matrix. It follows that the conditional distribution of $\Omega_t^{-1} = \Sigma_t$

given Ω_{t-1} is of Matrix- F type shifted by the factor $\lambda\Omega_{t-1}^{-1}$, and the conditional expectation is given by

$$E[\Omega_t^{-1} \mid \Omega_{t-1}] = \lambda \left(1 + \frac{n}{k-d-1} \right) \Omega_{t-1}^{-1}.$$

By restricting the parameter λ through (R1) the Ω_t^{-1} process becomes a martingale with conditional expectation $E[\Omega_t^{-1} \mid \Omega_{t-1}] = \Omega_{t-1}^{-1}$.

A.3 Log determinant process for the latent precision

From Eq. (32) the evolution of the logarithmic determinant of Ω_t follows as

$$\begin{aligned} \log |\Omega_t| &= \log |\mathcal{U}(\Omega_{t-1})' \Psi_t \mathcal{U}(\Omega_{t-1}) / \lambda| \\ &= \log \left\{ |\mathcal{U}(\Omega_{t-1})' \mathcal{U}(\Omega_{t-1})| \cdot |\Psi_t| \cdot \lambda^{-d} \right\} \\ &= \log |\Omega_{t-1}| + \log |\Psi_t| - d \log \lambda. \end{aligned}$$

Setting $\lambda = \exp\{E[\log |\Psi_t|] / d\}$ according to restriction (R3) results in a random walk process for the log determinant as

$$\log |\Omega_t| = \log |\Omega_{t-1}| + (\log |\Psi_t| - E[\log |\Psi_t|]) = \log |\Omega_{t-1}| + \tilde{\psi}_t,$$

with $\tilde{\psi}_t = \log |\Psi_t| - E[\log |\Psi_t|]$ being a white noise increment. Hence, the $\log |\Omega_t|$ process is a martingale with conditional expectation $E[\log |\Omega_t| \mid \Omega_{t-1}] = \log |\Omega_{t-1}|$. The unconditional expectation for the logarithmic determinant of Ψ_t follows from $\Psi_t \sim \mathcal{B}_d^I(\frac{k}{2}, \frac{n}{2})$ as

$$\begin{aligned} E[\log |\Psi_t|] &= \left[\frac{\Gamma_d(\frac{n}{2}) \Gamma_d(\frac{k}{2})}{\Gamma_d(\frac{n+k}{2})} \right]^{-1} \int_{0 < \Psi_t < I_d} (\log |\Psi_t|) |\Psi_t|^{(k-d-1)/2} |I_d - \Psi_t|^{(n-d-1)/2} d\Psi_t \\ &= 2 \left[\frac{\Gamma_d(\frac{n}{2}) \Gamma_d(\frac{k}{2})}{\Gamma_d(\frac{n+k}{2})} \right]^{-1} \int_{0 < \Psi_t < I_d} \frac{\partial}{\partial k} |\Psi_t|^{(k-d-1)/2} |I_d - \Psi_t|^{(n-d-1)/2} d\Psi_t \\ &= 2 \left[\frac{\Gamma_d(\frac{n}{2}) \Gamma_d(\frac{k}{2})}{\Gamma_d(\frac{n+k}{2})} \right]^{-1} \frac{\partial}{\partial k} \int_{0 < \Psi_t < I_d} |\Psi_t|^{(k-d-1)/2} |I_d - \Psi_t|^{(n-d-1)/2} d\Psi_t \\ &= 2 \left[\frac{\Gamma_d(\frac{n}{2}) \Gamma_d(\frac{k}{2})}{\Gamma_d(\frac{n+k}{2})} \right]^{-1} \frac{\partial}{\partial k} \left[\frac{\Gamma_d(\frac{n}{2}) \Gamma_d(\frac{k}{2})}{\Gamma_d(\frac{n+k}{2})} \right] \\ &= 2 \frac{\partial}{\partial k} \log \left[\frac{\Gamma_d(\frac{n}{2}) \Gamma_d(\frac{k}{2})}{\Gamma_d(\frac{n+k}{2})} \right] \\ &= \Psi_d^* \left(\frac{k}{2} \right) - \Psi_d^* \left(\frac{k+n}{2} \right), \end{aligned}$$

with $\Psi_d^*(a) = \partial \log \Gamma_d(a) / \partial a$ being the d -variate digamma function (see, e.g., Abramowitz and Stegun 1972; Gupta and Nagar 2000).

References

- Abramowitz M, Stegun I (1972) Handbook of mathematical functions. Number 55, 9 edn. Dover Publications Inc., New York
- Aït-Sahalia Y, Xiu D (2017) Using principal component analysis to estimate a high-dimensional factor model with high-frequency data. *J Econom* 201(2):384–399
- Andersen TG, Bollerslev T, Diebold FX, Labys P (2003) Modeling and forecasting realised volatility. *Econometrica* 71:579–625
- Asai M, McAleer M (2015) Forecasting co-volatilities via factor models with asymmetry and long memory in realized covariance. *J Econom* 189(2):251–262
- Barndorff-Nielsen OE, Shephard N (2004) Econometric analysis of realised covariation: high frequency covariance, regression and correlation in financial economics. *Econometrica* 72(3):885–925
- Barndorff-Nielsen OE, Hansen PR, 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. *J Econom* 162(2):149–169
- Bauer GH, Vorkink K (2011) Forecasting multivariate realized stock market volatility. *J Econom* 160(1):93–101
- Bauwens L, Braione M, Storti G (2016) Multiplicative conditional correlation models for realized covariance matrices. CORE working paper (2016/41)
- Bauwens L, Braione M, Storti G (2017) A dynamic component model for forecasting high-dimensional realized covariance matrices. *Econom Stat* 1:40–61
- Blasques F, Lucas A, Opschoor A, Rossini L (2021) Tail heterogeneity for dynamic covariance-matrix valued random variables: the F-Riesz distribution. Tinbergen Institute Discussion Paper
- Bollerslev T, Patton AJ, Quaedvlieg R (2018) Modeling and forecasting (un)reliable realized covariances for more reliable decisions. *J Econom* 207(1):71–91
- Brito D, Medeiros MC, Ribeiro R (2018) Forecasting large realized covariance matrices: The benefits of factor models and shrinkage. SSRN Working Paper
- Callot LAF, Kock AB, Medeiros MC (2017) Modeling and forecasting large realized covariance matrices and portfolio choice. *J Appl Econom* 32(1):140–158
- Chiriac R, Voev V (2011) Modelling and forecasting multivariate realized volatility. *J Appl Econom* 26(6):922–947
- Corsi F (2009) A simple approximative long-memory model of realized volatility. *J Financ Econom* 7(2):174–196
- Engle R, Kelly B (2012) Dynamic equicorrelation. *J Bus Econ Stat* 30(2):212–228
- Engle R, Ledoit O, Wolf M (2019) Large dynamic covariance matrices. *J Bus Econ Stat* 38(2):363–375
- Fama EF, French KR (1993) Common risk factors in the returns on stocks and bonds. *J Financ Econ* 33(1):3–56
- Fan J, Kim D (2018) Robust high-dimensional volatility matrix estimation for high-frequency factor model. *J Am Stat Assoc* 113(523):1268–1283
- Fan J, Zhang J, Yu K (2012) Vast portfolio selection with gross-exposure constraints. *J Am Stat Assoc* 107(498):592–606
- Fan J, Furger A, Xiu D (2016) Incorporating global industrial classification standard into portfolio allocation: a simple factor-based large covariance matrix estimator with high-frequency data. *J Bus Econ Stat* 34(4):489–503
- Golosnoy V, Gribisch B, Liesenfeld R (2012) The conditional autoregressive Wishart model for multivariate stock market volatility. *J Econom* 167(1):211–223
- Gourieroux C, Jasiak J, Sufana R (2009) The Wishart autoregressive process of multivariate stochastic volatility. *J Econom* 150(2):167–181
- Grant M, Boyd S (2014) CVX: Matlab software for disciplined convex programming, version 2.1. <http://cvxr.com/cvx>

- Gribisch B, Hartkopf JP (2022) Modeling realized covariance measures with heterogeneous liquidity: a generalized matrix-variate Wishart state-space model. *J Econom* (forthcoming). <https://doi.org/10.1016/j.jeconom.2022.01.007>
- Gribisch B, Hartkopf JP, Liesenfeld R (2020) Factor state-space models for high-dimensional realized covariance matrices of asset returns. *J Empir Financ* 55(1):1–20
- Gupta AK, Nagar DK (2000) Matrix variate distributions. Monographs and Surveys in Pure and Applied Mathematics 104. Chapman & Hall, Boca Raton
- Hansen PR, Lunde A, Nason JM (2011) The model confidence set. *Econometrica* 79:453–497
- Hautsch N, Kyj LM, Malec P (2011) The merit of high-frequency data in portfolio allocation. SSRN Working paper. <https://doi.org/10.2139/ssrn.1926098>
- Jegadeesh N, Titman S (1993) Returns to buying winners and selling losers: implications for stock market efficiency. *J Finance* 48(1):65–91
- Jin X, Maheu JM (2013) Modeling realized covariances and returns. *J Financ Econom* 11(2):335–369
- Jin X, Maheu JM (2016) Bayesian semiparametric modeling of realized covariance matrices. *J Econom* 192(1):19–39
- Jin X, Maheu JM, Yang Q (2019) Bayesian parametric and semiparametric factor models for large realized covariance matrices. *J Appl Econom* 34:641–660
- Kim D (2014) Maximum likelihood estimation for vector autoregressions with multivariate stochastic volatility. *Econ Lett* 123:282–286
- Konno Y (1991) A note on estimating eigenvalues of scale matrix of the multivariate F-distribution. *Ann Inst Stat Math* 43(1):157–165
- Laurent S, Rombouts JVK, Violante F (2013) On loss functions and ranking forecasting performances of multivariate volatility models. *J Econom* 173(1):1–10
- Ledoit O, Wolf M (2003) Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J Empir Financ* 10:603–621
- Ledoit O, Wolf M (2004) A well-conditioned estimator for large-dimensional covariance matrices. *J Multivar Anal* 88:365–411
- Ledoit O, Santa-Clara P, Wolf M (2003) Flexible multivariate GARCH modeling with an application to international stock markets. *Rev Econ Stat* 85(3):735–747
- Lunde A, Shephard N, Sheppard K (2016) Econometric analysis of vast covariance matrices using composite realized kernels and their application to portfolio choice. *J Bus Econ Stat* 34(4):504–518
- Luo J, Chen L (2020) Realized volatility forecast with Bayesian random compressed multivariate HAR model. *Int J Forecast* 36:781–799
- Lütkepohl H (1996) *Handbook of matrices*. Wiley, Chichester
- Marcellino M, Stock JH, Watson MW (2006) A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *J Econom* 135:499–526
- Morgan JP (1996) *Riskmetrics*, 4th edn. techreport, J. P. Morgan, New York
- Moura GV, Noriller MR (2019) Maximum likelihood estimation of a TVP-VAR. *Econ Lett* 174:78–83
- Moura GV, Santos A, Ruiz E (2020) Comparing high-dimensional conditional covariance matrices: implications for portfolio selection. *J Bank Finance* 118:1–13
- Muirhead RJ (2005) *Aspects of multivariate statistical theory*. Wiley, Hoboken
- Noureddin D, Shephard N, Sheppard K (2012) Multivariate high-frequency-based volatility (HEAVY) models. *J Appl Econom* 27(6):907–933
- Opschoor A, Lucas A (2019) Fractional integration and fat tails for realized covariance kernels. *J Financ Econom* 17(1):66–90
- Opschoor A, Janus P, Lucas A, van Dijk D (2017) New HEAVY models for fat-tailed realized covariances and returns. *J Bus Econ Stat* 36(4):643–657
- Park S, Linton O (2012) *Handbook of volatility models and their applications*, chapter realized volatility: theory and applications. Wiley, Hoboken, pp 317–345
- Patton AJ (2011) Volatility forecast comparison using imperfect volatility proxies. *J Econom* 160(1):246–256
- Philipov A, Glickman ME (2006) Factor stochastic volatility via Wishart processes. *Econom Rev* 25(2–3):311–334
- Shephard N (1994) Local scale models: state space alternative to integrated GARCH processes. *J Econom* 60:181–202
- Sheppard K, Xu W (2019) Factor high-frequency based volatility (HEAVY) models. *J Financ Econom* 17(1):33–65

- Tao M, Wang Y, Yao Q, Zou J (2011) Large volatility matrix inference via combining low-frequency and high-frequency approaches. *J Am Stat Assoc* 106:1025–1040
- Uhlig H (1994) On singular Wishart and singular multivariate Beta distributions. *Ann Stat* 22:395–405
- Uhlig H (1997) Bayesian vector autoregressions with stochastic volatility. *Econometrica* 65:59–73
- Vasallo D, Buccheri G, Corsi F (2021) A DCC-type approach for realized covariance modeling with score-driven dynamics. *Int J Forecast* 37(2):569–586
- von Rosen D (1988) Moments for the inverted Wishart distribution. *Scand J Stat* 15(2):97–109
- Windle J, Carvalho M (2014) A tractable state-space model for symmetric positive-definite matrices. *Bayesian Anal* 9(4):759–792
- Zhou J, Jiang F, Zhu K, Li WL (2019) Time series models for realized covariance matrices based on the Matrix-F distribution. arXiv preprint, [arXiv:1903.12077](https://arxiv.org/abs/1903.12077)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.