Factor State-Space Models for High-Dimensional Realized Covariance Matrices of Asset Returns

Bastian Gribisch

_Institute of Econometrics and Statistics, University of Cologne, Germany_

Jan Patrick Hartkopf

_Cologne Graduate School, University of Cologne, Germany_

Roman Liesenfeld*

_Institute of Econometrics and Statistics, University of Cologne, Germany_

(August 17, 2018)

Abstract

We propose a dynamic factor state-space model for high-dimensional covariance matrices of asset returns. It makes use of observed risk factors and assumes that the latent integrated joint covariance matrix of the assets and the factors is observed through their realized covariance matrix with a Wishart measurement density. For the latent integrated covariance matrix of the assets we impose a strict factor structure allowing for dynamic variation in the covariance matrices of the factors and the residuals components as well in the factor loadings. This factor structure translates into a factorization of the Wishart measurement density which facilitates statistical inference based on simple Bayesian MCMC procedures making the approach scalable w.r.t. the number of assets. An empirical application to realized covariance matrices for 60 NYSE traded stocks using the Fama-French factors and sector-specific factors represented by exchange traded funds (ETF’s) shows that the model performs very well in- and out of sample.

*JEL classification:* C32, C38, C51, C58, G17

_Keywords:_ Factor model, Realized covariance, State-space model, Bayesian inference, Wishart distribution.

*Corresponding address: Institute of Econometrics and Statistics, University of Cologne, Universitätsstr. 22a, D-50937 Cologne, Germany. Tel.: +49(0)221-470-2813; fax: +49(0)221-470-5074. E-mail address: liesenfeld@statistik.uni-koeln.de (R. Liesenfeld)
1 Introduction

Modeling and forecasting covariance matrices of asset returns is important in risk management and portfolio allocation. Recent contributions to this field increasingly make use of realized covariance matrices which provide non-parametric ex-post estimates for the latent integrated covariance matrices of asset returns, and develop dynamic time-series models for those estimates\(^1\). Pioneering approaches to modeling and predicting realized covariance matrices are found in Gourieroux et al. (2009), Chiriac and Voev (2011), Bauer and Vorkink (2011), Noureldin et al. (2012), Golosnoy et al. (2012) and Jin and Mahcu (2013). However, the models developed in those studies typically suffer from a proliferation of parameters in high dimensional applications and the estimation of their parameters becomes rapidly difficult as the number of assets increases so that they often have limited practical relevance in realistic financial applications. Strategies which have been proposed to overcome those difficulties include the design of models such that their parameters can be iteratively estimated by multistep procedures (Bauwens et al., 2012, 2016), the use of LASSO (least absolute shrinkage and selection operator) type estimation techniques (Callot et al., 2017), or the application of sparse factor structures for the assets' covariance matrix (Tao et al., 2011, Sheppard and Xu, 2014, Asai and McAleer, 2015, Jin et al., 2017). The particular appeal of using a factor approach relative to the other alternatives is that factor models can be economically motivated and have a long established history in explaining the variation of financial returns. Prominent examples thereof are the Capital Asset Pricing Model (CAPM) of Sharpe (1964) andLintner (1965), the Arbitrage Pricing Theory (APT) of Ross (1976) and the three-factor model of Fama and French (1993).

In this paper we also adopt a factor approach and propose a dynamic factor state-space model for high-dimensional covariance matrices of asset returns which can be easily statistically analyzed by a combination of fairly standard Bayesian Monte-Carlo Markov-Chain (MCMC) procedures. It makes use of observed risk factors (such as those in the Fama-French model) and takes the joint integrated covariance matrix of the assets and factors as a latent state variable which is observed through their noisy realized covariance matrix with a Wishart measurement density. For the marginal integrated covariance matrix of the assets we impose a strict factor structure decomposing it into a low-rank component driven by the factors’ integrated covariance and a sparse diagonal matrix for

\(^1\)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. (2015).
the residual components. In this matrix factor decomposition we allow for dynamic variation in the (co)variances of the factors and residual components as well as in the factor loadings. The key to a simple statistical analysis lies in the property of the Wishart measurement density that it factorizes under an observed strict factor structure in such a way that the proposed multivariate state-space model for the covariance matrix can be devoted into conditionally independent low-dimensional state space models. Thus, we can rely on a simple MCMC approach and exploit for its implementation computational parallelization techniques making the approach fully scalable w.r.t. the number of assets.

Our Wishart factor state-space (WFSS) model builds upon and generalizes the model of Sheppard and Xu (2014). Their factor HEAVY (high-frequency based volatility) model also combines a Wishart density for the realized covariance matrix with an observed strict factor structure. However, in its application the authors use a single-factor structure based on the market factor in the CAPM model. In our approach we consider additional risk factors and utilize the three Fama-French factors as well as sector-specific Exchange Traded Funds (ETF’s). Such an extension can be expected to be critical since a single factor may not suffice to justify the approximation of the covariance matrix for the residual component by a diagonal matrix. Empirical evidence that the CAPM market factor is by far not sufficient to eliminate the correlation in the residual component and that the combination of the three Fama-French factors with sector-specific ETFs substantially improves the sparsity of the residual correlation is provided by Fan et al. (2016) and is also confirmed by our empirical results in Section 4. Our Bayesian factor state space approach also differs from the factor HEAVY model in that the latter is an observation-driven GARCH-type framework statistically analyzed by a two-step quasi ML procedure. This two-step approach is required as the joint estimation of all the parameters (including the degrees of freedom of the Wishart distribution) becomes computationally difficult in high-dimensional applications. Last but not least, the factor HEAVY model as implemented by Sheppard and Xu (2014) uses low-order GARCH(1,1) recursions so that it ignores potential long-memory type dynamics in the realized (co)variances of asset returns which is often found in empirical applications (see, e.g., Corsi et al., 2012, Bekierman and Manner, 2017 and the literature cited therein). In order to accommodate potential long-memory we endow the state variables directing the variances in the WFSS model with heterogenous autoregressive (HAR) processes which are known to provide effective approximations of long-memory dynamics (Corsi, 2009).
An alternative to the observed factor approach with explicit exogenously supplied factors as adopted in this paper is to rely on implicit factors extracted from realized covariance data as proposed by Tao et al. (2011) and Asai and McAleer (2015). They first construct from the realized covariance matrices of stock returns via an eigenanalysis the covariance matrices of common factors for which they build in a second step dynamic time-series models. In contrast to our approach, this two-step procedure is based on the assumption that both the factor loadings as well as the residual covariance matrix are time-invariant which restricts the flexibility to account for nontrivial contemporaneous and dynamic interactions in the (co)variances of asset returns. The same applies to the factor approach of Jin et al. (2017) which imposes an eigenvalue decomposition directly in the dynamic model for the realized covariance matrix of stock returns.

As a preview of our main empirical results, we apply the WFSS model to daily covariance matrices for the returns of 60 NYSE traded stocks and find by using Bayesian model comparisons that the single factor approach based on the CAPM market factor is empirically rejected in favor of a factor structure including the three Fama-French factors and the sector-specific ETF factors. Most importantly, this extended factor structure combined with a WFSS allowing for dynamically varying factor loadings and long-memory type dependence is extremely useful in explaining the observed dynamic variation in the covariance matrix of the stock returns. We also run out-of-sample forecasts and illustrate the predictive performance of our approach relative to competing models along several dimensions: Accuracy of the (co)variance predictions, reliability of density forecast and the ability to produce predictions for the Value-at-Risk and the global-minimum variance portfolio. Our results show that our WFSS model performs favorable in nearly all dimensions relative to its competitors.

The rest of the paper is organized as follows: Section 2 introduces the baseline model. The proposed MCMC procedure for the Bayesian posterior analysis and model comparisons as well as the construction of forecasts are discussed in Section 3. Section 4 presents the empirical application to NYSE data and Section 5 concludes.
2 The Model

2.1 Factor structure

Consider the $m \times m$ realized covariance matrix $C_t$ used to approximate the joint period-$t$ integrated covariance matrix $\Sigma_t$ for a vector of log-prices for $p$ individual assets together with $q$ observable risk factors with $m = p + q$. Here, $p$ is assumed to be substantially larger than $q$. Let $\Sigma_t$ and $C_t$ be partitioned as

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

(1)

where $\Sigma_t^r$ denotes the $p \times p$ integrated covariance matrix for the assets, $\Sigma_t^f$ the $q \times q$ integrated covariance matrix for the factors and $\Sigma_t^{rf}$ the $q \times p$ matrix of the integrated covariances between the factors and assets. $C_t$ in Eq. (1) is partitioned conformably with $\Sigma_t$ so that $C_t^r$ is the realized covariance matrix of the assets, $C_t^f$ that of the factors and $C_t^{rf}$ the matrix of the realized covariances between the factors and the assets.

Our aim is to predict the potentially large-dimensional realized covariance matrix $C_t^r$. For this purpose, we propose a joint dynamic model for the realized (co)variances of the factors and assets in $C_t$. This model is based on a decomposition of the assets’ integrated covariance matrix $\Sigma_t^r$ into a low-rank component driven by the integrated covariance matrix of the observed risk factors $\Sigma_t^f$ and a residual component. Such a decomposition obtains by imposing a continuous-time factor model for the asset prices relating them to the factor prices and assuming that the factor prices and the residual components of the asset prices are uncorrelated (Fan et al., 2016 and Aït-Sahalia and Xiu, 2015). The resulting integrated covariance matrix for the assets obtains as

$$
\Sigma_t^r = B_t \Sigma_t^f B_t' + \Sigma_t^e,
$$

(2)

where $B_t$ denotes the factor loading matrix of size $p \times q$ and $\Sigma_t^e$ is the integrated covariance matrix of the residual components of the factor model. Here, we assume a strict factor model such that $\Sigma_t^e$ is a diagonal matrix. This ensures that the number of parameters increases only linearly in the number of assets which is obviously desirable in (very) high-dimensional applications. Clearly, assuming $\Sigma_t^e$ to be diagonal appears to be fairly restrictive, especially, for a small number of observed factors like in the one-factor CAPM and the three-factor model of Fama and French (1993), where we cannot expect...
the residual correlations to be negligible. However, when augmenting those small factor models by
including sector specific Exchange Traded Funds (ETFs) as observed additional factors the sparsity
of the residual correlation significantly improves (see Fan et al., 2016, and the empirical results in
Section 4.1)

In our approach, we allow the factor risk premia in \( \Sigma^f \) and the idiosyncratic risks in \( \Sigma^e \) as well
as the betas given by the factor loadings in \( B \) to be time-varying. While time-variation in the factor
risk premia and the idiosyncratic risk, now is well accepted, it is less so for the betas. However, the
recent empirical literature reports increasing evidence of dynamically varying betas in a one-factor
CAPM model (Andersen et al., 2005, Ang and Chen, 2007, Ghysels and Jacquier, 2006, Sheppard
and Xu, 2014, Kalnina, 2015) as well in a Fama-French three-factor model (Bollerslev and Zhang,
2003, Engle, 2017). A further justification for time-varying betas, is that they obtain as population
regression coefficients \( B_t = \Sigma_t^{fr} (\Sigma_t^f)^{-1} \) so that their time-invariance would require to impose severe
restrictions on the dynamics of the joint integrated covariance matrix \( \Sigma_t \) (Engle, 2017).

2.2 Wishart Factor State-Space Model

Taking the integrated covariance matrix \( \Sigma_t \) with its components as a latent state variable observed
through the noisy realized covariance matrix \( C_t \) it is reasonable to model \( C_t \) by a state-space approach
with a measurement density \( f(C_t|\Sigma_t) \) relating the measurements \( C_t \) to the states \( \Sigma_t \) and a transition
density \( f(\Sigma_t|\Sigma_{t-1}) \) for the time-varying \( \Sigma_t \) to be designed to approximate the observed dynamics
of \( C_t \). The notation \( A_{s,t} \) is used to denote the collection \( \{A_s, \ldots, A_t\} \).

A natural selection for the measurement density \( f(C_t|\Sigma_t) \) is that of a central \( m \)-dimensional
Wishart distribution, \( C_t|\Sigma_t \sim W_m(n, \Sigma_t/n) \), where \( n \geq m \) is the scalar degree of freedom and \( \Sigma_t/n \) is the scale matrix (see Philipov and Glickman, 2006, Golosnoy et al., 2012, and Noureldin et al.,
2012, for applications of the Wishart distribution to realized covariance matrices). The scale matrix
is normalized by \( n \) so that the conditional expectation of \( C_t \) is given by \( E(C_t|\Sigma_t) = \Sigma_t \). The density
function of this Wishart distribution is

\[
f_W(C_t|n, \Sigma_t/n) = \frac{|C_t|^{(n-m-1)/2} |\Sigma_t/n|^{-n/2}}{2^{nm/2} \pi^{m(m-1)/4} \prod_{i=1}^{m} \Gamma((n + 1 - i)/2)} \exp \left\{ -\frac{n}{2} \text{tr}(\Sigma_t^{-1} C_t) \right\}, \tag{3}
\]

where \( \Gamma(\cdot) \) denotes the Gamma function.
An important advantage of using a Wishart measurement density for \( C_t \) is that it admits a parametrization reflecting the factor decomposition of the integrated covariance of the assets \( \Sigma_t^r \) in Eq. (2), which when combined with a diagonal form of the integrated residual covariance \( \Sigma_t^e \) translates into a convenient factorization of the measurement density. This greatly simplifies the statistical inference in high-dimensional applications. In particular, using the partitioning of \( C_t \) and \( \Sigma_t \) in Eq. (1) and taking

\[
C_t^e = C_t^r - C_t^{fr}(C_t^f)^{-1}C_t^{rf}, \quad \Sigma_t^e = \Sigma_t^r - \Sigma_t^{fr}(\Sigma_t^f)^{-1}\Sigma_t^{rf}, \quad B_t = \Sigma_t^{fr}(\Sigma_t^f)^{-1},
\]

the Wishart density function for \( C_t \) in Eq. (3) factorizes into the product of a Wishart density for \( C_t^f \), a conditional Gaussian density for \( C_t^{fr} \) given \( C_t^f \) and an independent Wishart density for \( C_t^e \) (Muirhead, 2005, Theorem 3.2.10),

\[
f_W(C_t|n, \Sigma_t/n) = f_W(C_t^f|n, \Sigma_t^f/n) \times f_{MN}(C_t^{fr}|B_tC_t^f, (\Sigma_t^e \otimes C_t^f)/n) f_W(C_t^e|n - q, \Sigma_t^e/n).
\]

Here the function \( f_{MN} \) is the density of a matrix-variate normal distribution for \( C_t^{fr} \) with mean and covariance for \( \text{vec}(C_t^f) \) given by \( \text{vec}(C_t^f B_t^\prime) \) and \( (\Sigma_t^e \otimes C_t^f)/n \), respectively.

Under the assumption that the integrated residual covariance is diagonal with \( \Sigma_t^e = \text{diag}(\sigma_{et}^1, \ldots, \sigma_{et}^p) \) it follows from the decomposition in Eq. (5) that the Wishart density for \( C_t \) as a function of the latent states \( \Sigma_t = (\Sigma_t^f, B_t, \Sigma_t^e) \) obtains as

\[
f_W(C_t|n, \Sigma_t/n) \propto \left| \Sigma_t^f \right|^{-n/2} \exp \left\{ - \frac{n}{2} \text{tr} \left[ (\Sigma_t^f)^{-1}C_t^f \right] \right\} \prod_{i=1}^p (\sigma_{it}^e)^{-n/2} \exp \left\{ - \frac{n}{2\sigma_{it}^e} (\beta_{it}^\prime C_t^f \beta_{it} - 2\beta_{it}^\prime c_{it}^e + c_{it}^e) \right\},
\]

where \( \beta_{it} = (\beta_{i1t}, \ldots, \beta_{iqt})^\prime \) denotes the vector of loadings of asset \( i \) on the \( q \) factors such that \( B_t = (\beta_{1t}, \ldots, \beta_{pt})^\prime \) and \( c_{it}^{rf} = (c_{i1t}^{rf}, \ldots, c_{iqt}^{rf})^\prime \) is the vector of realized covariances between the \( q \) factors and the \( i \)'th asset with \( C_t^{rf} = (c_{i1t}^{rf}, \ldots, c_{iqt}^{rf}) \). The scalar \( c_{it}^e \) is the \( i \)th diagonal element of \( C_t^e \) representing the realized variance of asset \( i \). The component in the first bracket in Eq. (6) is the kernel of the Wishart density for the realized factor covariance \( C_t^f \) as given in Eq. (5), while the component in the second bracket is obtained from the multiplication of the Wishart density kernel for
the realized residual covariance $C_e^t$ by the conditional Gaussian density kernel for $C_{fr}^t$, given $C_f^t$. As a result of the diagonal form of $\Sigma^t$, this second component factorizes into $p$ (functionally) independent factors, one for each asset.

In order to complete the factor state-space model, we specify the transition densities for the latent time-varying integrated covariance matrices ($\Sigma_f^t$, $\Sigma_e^t$) and the factor loadings in $B_t$. To accommodate the observed dynamics in the realized covariance matrix with its typically strong persistence, especially in the realized variances, we combine simple AR(1) processes for the factor loadings with heterogeneous autoregressive (HAR) processes for the integrated variances. Those HAR processes as introduced by Corsi (2009) provide simple yet effective approximations of long-memory type persistence and are well-suited for modelling realized variances (see, e.g., Bekierman and Manner, 2017 and the literature cited therein).

For the logs of the $p$ elements in the diagonal idiosyncratic covariance matrix $\Sigma^t$ denoted by $x_{it}^e = \ln \sigma_{it}^e$ we assume mutually independent Gaussian HAR processes of the form

$$x_{it}^e - \gamma_i = \phi_{i1}^e \bar{x}_{i[t-1:t-1]}^e + \phi_{i2}^e \bar{x}_{i[t-1:t-5]}^e + \phi_{i3}^e \bar{x}_{i[t-1:t-22]}^e + \nu_i^e \eta_{it},$$

$$\eta_{it} \sim N(0, 1), \quad i = 1, \ldots, p,$$

where $\bar{x}_{i[t-1:t-h]}^e = \sum_{\tau=1}^h (x_{it-\tau}^e - \gamma_i^e)/h$ for $h = 1, 5, 22$ represents daily, weekly and monthly lags, respectively. The parameters are $\theta_i^e = (\gamma_i^e, \phi_{i1}^e, \phi_{i2}^e, \phi_{i3}^e, \nu_i^e)$, and for the restriction $\phi_{i2}^e = \phi_{i3}^e = 0$ the HAR process in Eq. (7) reduces to a standard AR(1). For the $pq$ factor loadings in $B_t$ we assume the following independent standard Gaussian AR(1) processes:

$$\beta_{ikt} - \gamma_{ik} = \phi_{ik}^\beta (\beta_{ikt} - \gamma_{ik}) + \nu_{ik}^\beta \eta_{ikt}, \quad \eta_{ikt} \sim N(0, 1), \quad i = 1, \ldots, p, \quad k = 1, \ldots, q,$$

parameterized by $\theta_{ik}^\beta = (\gamma_{ik}^\beta, \phi_{ik}^\beta, \nu_{ik}^\beta)$.

The factor covariance matrix $\Sigma^f_t$ is Cholesky-decomposed into

$$\Sigma^f_t = L_t^{-1} D_t L_t^{-1'},$$

8
where $D_t$ is a diagonal matrix and $L_t$ is a lower-triangular matrix with unit diagonal elements, say,

$$
D_t = \text{diag}(\sigma_{1t}^f, \ldots, \sigma_{qt}^f), \quad L_t = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \ell_{21t} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{q1t} & \ell_{q2t} & \cdots & 1 \end{pmatrix} = \begin{pmatrix} \ell_{11t}^f \\ \ell_{21t}^f \\ \vdots \\ \ell_{qt}^f \end{pmatrix}.
$$

In order to allow for sufficient flexibility in accounting for the observed dynamics in the factor covariance we assume for the logs of the $q$-elements in $D_t$, denoted by $x_{kt}^f = \ln(\sigma_{kt}^f)$ independent Gaussian HAR processes, that is

$$
x_{kt}^f - \gamma_k^f = \phi_{k1}^f x_{[t-1:t-1]}^f + \phi_{k2}^f x_{[t-1:t-5]}^f + \phi_{k3}^f x_{[t-1:t-22]}^f + \nu_k^f \eta_{kt}^f, \quad \eta_{kt}^f \sim \mathcal{N}(0, 1), \quad k = 1, \ldots, q,
$$

with parameters $\theta_k^f = (\gamma_k^f, \phi_{k1}^f, \phi_{k2}^f, \phi_{k3}^f, \nu_k^f)$, and for the $(q - 1)/2$ free elements in the matrix of pseudo-loadings $L_t$ independent Gaussian AR(1) processes,

$$
\ell_{kj}^f - \gamma_{kj}^f = \phi_{kj}^f (\ell_{kj-1}^f - \gamma_{kj}^f) + \nu_{kj}^f \eta_{kj}^f, \quad \eta_{kj}^f \sim \mathcal{N}(0, 1), \quad k > j = 1, \ldots, q - 1,
$$

where $\theta_{kj}^f = (\gamma_{kj}^f, \phi_{kj}^f, \nu_{kj}^f)$.

For later reference we note that a HAR process of the form as given in Eqs. (7) and (11) can be written as the following restricted AR(22):

$$
x_t - \gamma = \left[ \phi_1 + \phi_2 + \phi_3 \right] (x_{t-1} - \gamma) + \left[ \phi_2 + \phi_3 \right] (x_{t-2} - \gamma) + \cdots + \left[ \phi_2 + \phi_3 \right] (x_{t-5} - \gamma)
$$

$$
+ \phi_3 (x_{t-6} - \gamma) + \cdots + \phi_3 (x_{t-22} - \gamma) + \nu \eta_t,
$$

where we have omitted the indices for the assets and factors.

The vector of parameters in this Wishart factor state space (WFSS) model as defined by Eqs. (6)-(12) consists of $5[p + q] + 3[pq + q(q - 1)/2] + 1$ parameters, which are the set of HAR-parameters $(\gamma, \phi_1, \phi_2, \phi_3, \nu)$ for the $[p + q]$ state processes $\{x_{kt}^e\}$ and $\{x_{kt}^f\}$, the set of AR(1) parameters $(\gamma, \phi, \nu)$ for the $[pq + q(q - 1)/2]$ state processes $\{\beta_{ikt}\}$ and $\{\ell_{kj}^f\}$, and the degree of freedom $n$ of the Wishart measurement density. Due to the factor structure, both the number of state processes as well as the
number of parameters are (for a given set of factors) linear in the number of assets \( p \). Still, for a large number of assets, the actual amount of parameters appears to be fairly large. For instance, with \( p = 60 \) assets, \( q = 9 \) observable factors as in our application below we have for the unrestricted WFSS model 645 state processes and 2074 parameters to be estimated. However, as we use a data set covering \( T = 1510 \) trading days with \((p + q)(p + q + 1)T/2 = 3,646,650\) (co)variance observations we have 1758 observations per parameter which can be expected to provide enough information for a reliable statistical inference.

In order to efficiently handle the large number of parameters in the Bayesian MCMC posterior analysis we take full advantage of the factorization of the measurement density in Eq. (6). The fact that its factor component as a function in the factor-specific states \( \Sigma_f^t \) and its \( p \) asset components as functions in the asset-specific states \( \{\sigma_{it}, \beta_{it}\} \) are mutually functionally independent together with the independent priors for the state processes given in Eqs. (7)-(12) allows us to use a simple MCMC approach and to exploit for its implementation computational parallelization techniques. This makes our approach fully scalable w.r.t. the number assets \( p \).

Using the WFSS model for the joint covariance matrix of the assets and the observed risk factors we can perform forecasting of the assets’ integrated covariance matrix \( \Sigma_t^r \) as further detailed in Section 3.3 below. Our WFSS framework also allows us to test the hypothesis that the factor loadings in \( B_t \) are time-invariant, which is the case if in the AR specification for the betas \( \beta_{ikt} \) in Eq. (8) \( \phi_{ik}^\beta = 0 \) and \( \nu_{ik}^\beta \to 0 \) so that \( \beta_{ikt} = \gamma_{ik}^\beta \forall t \). Likewise, the framework enables us to analyze by means of corresponding model comparisons the relative importance of the different factors under consideration in explaining and predicting the observed variation of the asset returns. Instrumental for such an analysis is the conditional density of \( C_t^r \) given \((C_t^{fr}, C_t^f, \Sigma_t)\) for the respective set of included factors.

As discussed further below in Section 3.2 this density is required to obtain the conditional likelihood for the observed realized covariance matrices of the assets. According to Eqs. (4) and (5) the realized covariance for the assets can be represented as \( C_t^r = C_t^{fr}(C_t^f)^{-1}C_t^{r,f} + C_t^e \) with \( C_t^e \sim W_p(n - q, \Sigma_t^e/n) \) and Jacobian \( dC_t^r = dC_t^e \), so that the conditional density for \( C_t^r \) is given by

\[
f(C_t^r \mid C_t^{fr}, C_t^f, \Sigma_t) = f(C_t^r \mid \Sigma_t) = f_W(C_t^e \mid n - q, \Sigma_t^e/n).
\]
2.3 A note on the Wishart assumption for the measurement density

Asset returns are typically subject to the presence of outliers generating fat tails in the corresponding covariance measures (Opschoor et al., 2017). Under the assumed conditional Wishart distribution, with its fairly thin tails, significant fat-tail behavior can only originate from the unconditional variation of the latent states \((\Sigma_f^t, \Sigma_e^t, B_t)\), but this may not suffice to fully capture the tail behavior of realized covariance data. In such cases, the Wishart can be usefully replaced by a distribution allowing for conditional heavy tails such as the matrix-\(F\) distribution which obtains from a Wishart-inverted-Wishart mixture (Konno, 1991). A successful application of the matrix-\(F\) distribution to realized covariances is found in Opschoor et al. (2017) who combine it with an observation-driven generalized autoregressive score (GAS) approach. The matrix-\(F\) nests the Wishart distribution as a special case and admits a parametrization based on the strict factor decomposition in Eq. (2) leading (when conditioned on the mixing variable) to a factorization of the density for the realized covariance matrix \(C_t\) which is of the same form as under the Wishart (see Eqs. 5 and 6). This makes the matrix-\(F\) distribution easily applicable for a straightforward fat-tailed generalization of the WFSS model. However, the results we obtained for an initial posterior analysis of this generalization shows that there is no evidence against the Wishart in favor of the fat-tailed matrix-\(F\) distribution\(^2\). Actually, the posterior estimates for all the parameters of the matrix-\(F\) generalization are virtually equal to their values obtained for the fitted WFSS model. This indicates that the marginalization of the conditional Wishart distribution for \(C_t\) w.r.t. the latent state variables suffices to capture the tail behavior of the realized covariance data. This is illustrated in Figure 1 where we plot the histogram of the observed realized variances of the Citigroup and Caterpillar stock together with their unconditional distribution predicted under the fitted WFSS model defined by Eqs. (6)-(12) for 60 assets using 9 factors (see Section 4 below).

\(^2\)Results are not presented here but are available upon request.
3 Bayesian Posterior Analysis and Forecasting

We utilize MCMC methods for a Bayesian posterior analysis of the WFSS model and use the Gibbs approach to simulate from the joint posterior of the parameter and states

$$\pi(\{\theta^e_i\}, \{\theta^B_i\}, \{\theta^f_k\}, \{\theta^\ell_{kj}\}, n, \{x^e_{i,1:T}\}, \{x^f_{k,1:T}\}, \{\ell_{kj,1:T}\}|C_{1:T}).$$  \hspace{1cm} (15)

The factorization of the measurement density (6) combined with the independent priors for the state processes as specified by the state-transitions (7)-(12) imply that the WFSS model can be devoted in \(p+q\) conditionally independent state-space models, one for each of the \(q\) factors and \(p\) assets. This allows us to update within the Gibbs approach their respective state processes and their state specific parameters factor-by-factor and asset-by-asset.

3.1 MCMC algorithm

Our proposed MCMC implementation for the WFSS model consists of the following steps:

1.) Updating \(\{\theta^f_k\}, \{x^f_{k,1:T}\}\): For \(\theta^f_k\), the parameters of the Gaussian HAR-models for the factor state processes \(\{x^f_{k,1:T}\}\) as given in Eq. (11), we select independent natural conjugate Normal-inverted-Gamma priors. Thus we can directly simulate from their full conditional posteriors \(\pi(\theta^f_k|x^f_{k,1:T}), k = 1, \ldots, q\).

The measurement density in Eq. (6) together with the state transitions in Eqs. (9)-(11) for \(\Sigma^f_t\) define conditionally independent nonlinear non-Gaussian state-space models for the \(q\) factor state processes \(\{x^f_{k,1:T}\}\) given the pseudo loadings in \(L_t\). The resulting \(q\) full conditional posteriors for \(\{x^f_{k,1:T}\}\) are

$$\pi(x^f_{k,1:T}|\ell_{k,1:T}, \theta^f_k, n, C_{1:T}) \propto \prod_{t=1}^{T} \exp \left\{ -\frac{n}{2} \left[ x^f_{kt} + (\ell^t_{kt} C^f_t \ell_{kt}) \exp(-x^f_{kt}) \right] \right\}$$

$$\times f_N \left(x^f_{kt}|\gamma^f_k + \phi^f_{k1} x^f_{k[t-1:t-1]} + \phi^f_{k2} x^f_{k[t-1:t-5]} + \phi^f_{k3} x^f_{k[t-1:t-22]}, \nu^f_k, \sigma^2 \right), \quad k = 1, \ldots, q,$$

where \(f_N(\cdot|\mu, \sigma^2)\) denotes a Gaussian density with mean \(\mu\) and variance \(\sigma^2\). To sample a full trajectory \(x^f_{k,1:T}\) from its posterior in one block we use the Particle Gibbs (PG) procedure based on the Bootstrap Particle Filter (BPF) combined with Ancestor Sampling (AS). In a nutshell, the PG as
proposed by Andrieu et al. (2010) is a standard Gibbs sampler where we can use the BPF (Gordon et al., 1993) inside the Gibbs procedure in order to propose approximate samples from the posterior in such a way that the ‘ideal’ but infeasible Gibbs sampler is approximated. In order to improve the mixing of the resulting PG algorithm we combine the BPF with AS as recently proposed by Lindsten et al. (2014). A detailed description of this PG-AS procedure is also found in Grothe et al. (2017). Since the $q$ state processes with their specific parameters $\{(x_{k,1:T}^f, \theta_k^f)\}$ are conditionally independent they can be updated by first drawing in parallel the $q$ parameter vectors and then running in parallel the PG-AS for the $q$ state processes.

2.) Updating $\{\theta_{k,j}^f\}, \{\ell_{k,j;1:T}\}$: For the parameters $\{\theta_{k,j}^f\}$ of the $q(q-1)/2$ Gaussian AR-processes of the pseudo loadings $\ell_{k,j}$ in Eq. (12) we select Normal-inverted-Gamma priors so that they can be directly sampled from their full conditional posteriors $\pi(\theta_{k,j}^f|\ell_{k,j;1:T})$, $k > j = 1, \ldots, q - 1$.

Let $\tilde{\ell}_{kt} = (\ell_{k1}, \ldots, \ell_{kk-1})'$ denote the $(k - 1)$-dimensional vector consisting of the unrestricted elements in the vector of pseudo loadings $\ell_{kt}$ in Eq. (10) such that $\ell_{kt} = (\tilde{\ell}_{kt}', 1, 0, \ldots, 0)'$, $k = 2, \ldots, q$. Then the measurement density (6) together with the state transitions for the $\ell_{kjt}$’s in Eq. (12) define $q - 1$ conditionally independent linear Gaussian state-space models for the $(k - 1)$-dimensional processes $\{\tilde{\ell}_{k;1:T}\}$ given $\{x_{k;1:T}^f\}$. The corresponding full conditional posterior of $\tilde{\ell}_{k,1:T}$ has the following particular linear Gaussian form in $\tilde{\ell}_{kt}$ given $\tilde{\ell}_{kt-1}$:

$$
\pi(\tilde{\ell}_{k,1:T}|x_{k,1:T}^f, \{\theta_{k,j}^f\}, n, C_{1:T}) \propto \prod_{t=1}^T \exp \left\{-\frac{n}{2} \left[ \tilde{\rho}_{kt} \tilde{C}_{kt} \tilde{\ell}_{kt} - 2(\tilde{c}_{kt}') \tilde{\ell}_{kt} \right] \exp(-x_{kt}^f) \right\}
\times f_N(\tilde{\ell}_{kt} | \gamma_k^f + \Phi_k^f (\tilde{\ell}_{kt} - \gamma_k^f), \Sigma_k^f), \quad k = 2, \ldots, q,
$$

with $\gamma_k^f = (\gamma_{k1}^f, \ldots, \gamma_{kk-1}^f)'$, $\Phi_k^f = \text{diag}(\phi_{k1}^f, \ldots, \phi_{kk-1}^f)$ and $\Sigma_k^f = \text{diag}(\nu_{k1}^f, \ldots, \nu_{kk-1}^f)$. The matrix $\tilde{C}_{kt}$ denotes the upper left block of $C_k^f$ consisting of its first $k - 1$ rows and columns, and $\tilde{c}_{kt}$ is the column vector consisting of the first $k - 1$ elements of $k$’th row of $C_k^f$. To simulate the $k - 1$ trajectories in $\tilde{\ell}_{k,1:T}$ in one block we can straightforwardly apply the forward-filtering backward-sampling (FFBS) procedure of de Jong and Shephard (1995) for linear Gaussian state-space models. The conditional independence of the state processes $\{\tilde{\ell}_{k,1:T}\}$ allows us to update them together with their specific parameters $\{(\theta_{k1}^f, \ldots, \theta_{kk-1}^f)\}$ in parallel.

---

3 An alternative to the PG-AS procedure for sampling $x_{k,1:T}^f$ is the auxiliary mixture sampler based upon the Kalman filter as proposed by Kim et al. (1998). We also experimented with this alternative and the MCMC results we obtained are virtually the same as those based upon the PG-AS.
3.) Updating $\{\theta_i^e\}, \{x_{i,1:T}^e\}$: For $\{\theta_i^e\}$, the parameters of the $p$ HAR-processes for the log of the idiosyncratic variances in Eq. (7), we use normal-inverted-Gamma priors so that we can directly simulate from their full conditional posteriors $\pi(\theta_i^e|x_{i,1:T}^e), \ i = 1, \ldots, p$.

The measurement density in Eq. (6) together with the state transitions in Eq. (7) define conditionally independent nonlinear non-Gaussian state-space models for the $p$ state processes for the logs of the idiosyncratic variances $\{x_{i,1:T}^e\}$ given the loadings $\{\beta_{i,1:T}\}$. The corresponding full conditional posteriors for $\{x_{i,1:T}^e\}$ are

$$
\pi(x_{i,1:T}^e|\beta_{i,1:T}, \theta_i^e, n, C_{1:T}) \propto \prod_{t=1}^{T} \exp \left\{ -\frac{n}{2} \left[ x_{it}^e + (\beta'_{it} C_{it}^f \beta_{it} - 2 \beta'_{it} c_{it}^f + c_{it}^T) \exp(-x_{it}^e) \right] \right\} \times f_N\left(x_{it}^e | \gamma_{it}^e + \phi_{it}^e \bar{x}_{i[t-1:t-1]} + \phi_{it}^e \bar{x}_{i[t-1:t-5]} + \phi_{it}^e \bar{x}_{i[t-1:t-22]}, [\nu_{it}^e]^2\right), \ i = 1, \ldots, p.
$$

To sample the full trajectory $x_{i,1:T}^e$ from its posterior in one block we use the PG-AS procedure. Here again we can exploit the inherent parallel structure in the updating step for $\{(x_{i,1:T}^e, \theta_i^e)\}$.

4.) Updating $\{\theta_{ik}^\beta\}, \{\beta_{i,1:T}\}$: As for all the other AR parameters we select for those of the $pq$ time-varying factor loadings in Eq. (8) Normal-inverted-Gamma priors so that we can simulate directly from their full conditional posteriors $\pi(\theta_{ik}^\beta|\beta_{ik,1:T}), \ i = 1, \ldots, p, \ k = 1, \ldots, q$.

Since the measurement density in Eq. (6) together with the state transitions of the factor loadings $\beta_{ikt}$ define $p$ conditionally independent linear Gaussian state space models for $\{\beta_{i,1:T}\}$ given $\{x_{i,1:T}^e\}$, we can easily simulate by the FFBS procedure each $\beta_{i,1:T}$ from its full conditional posterior

$$
\pi(\beta_{i,1:T}|x_{i,1:T}^e, \{\theta_{ik}^\beta\}, n, C_{1:T}) \propto \prod_{t=1}^{T} \exp \left\{ -\frac{n}{2} \left( \beta'_{it} C_{it}^f \beta_{it} - 2 \beta'_{it} c_{it}^f \right) \exp(-x_{it}^e) \right\} \times f_N(\beta_{it} | \gamma_{it}^\beta + \Phi_{it}^\beta (\beta_{it} - \gamma_{it}^\beta), \Sigma_{it}^\beta), \ i = 1, \ldots, p,
$$

where $\gamma_{it}^\beta = (\gamma_{i1}^\beta, \ldots, \gamma_{iq}^\beta)'$, $\Phi_{it}^\beta = \text{diag}(\phi_{i1}^\beta, \ldots, \phi_{iq}^\beta)$ and $\Sigma_{it}^\beta = \text{diag}([\nu_{i1}^{\beta}]^2, \ldots, [\nu_{iq}^{\beta}]^2)$. As in the previous steps the update of $\{(\beta_{i,1:T}, \theta_{i1}^\beta, \ldots, \theta_{iq}^\beta)\}$ can be parallelized.

5.) Updating $n$: In the last step we simulate the Wishart degrees of freedom parameter. We select for $n > m$ an uniform prior, $\text{pr}(n)$, defined on a discrete grid. Thus we can directly simulate from
its full conditional posterior which is given by a multinomial distribution, i.e.

\[
\pi(n|\{x_e^{i,1:T}\}, \{\beta_{i,1:T}\}, \{x_k^{f,1:T}\}, \{\ell_k^{l,1:T}\}, C_{1:T}) \propto \prod_{t=1}^{T} f_{W}(C_t|n, \Sigma_t/n) \text{pr}(n).
\]

(20)

Our MCMC algorithm repeatedly cycles through Step 1) to 5). After dropping the draws from the first cycles as burn-in we use the draws from the next \(S\) cycles for the purpose of approximating the joint posterior in Eq. (15). Bayesian point estimates (posterior means) of the model parameters and latent state variables are then obtained as sample averages over the corresponding Gibbs draws.

3.2 Model comparison

For the purpose of comparing alternative WFSS model specifications obtained under different sets of risk factors included in \(C_f^T\) and/or different restrictions imposed on the state transitions densities in Eqs. (7)-(12) we rely upon the Deviance Information Criterion (DIC) based on the likelihood function (Spiegelhalter et al., 2002). For models with alternative sets of factors the joint realized covariances for the assets and the factors \(C_t\) consist of different set of variables which prevents the use of the likelihood function for \(C_{1:T}\) for model comparisons. Hence, we rely on the DIC based on the conditional likelihood function of the assets’ realized covariances \(C_r^T\) given \(C_{fr}^T\) and \(C_f^T\).

Let \(\theta = (\{\theta^e_i\}, \{\theta^\beta_{ik}\}, \{\theta^f_k\}, \{\theta^\ell_{kj}\}, n)\) denote the list of all parameters. Then the conditional DIC is given by

\[
\text{DIC} = -2 \log p(C_{1:T}^r|C_{fr}^T, C_f^T; \hat{\theta}) + 2p_D, \tag{21}
\]

with small values of the criterion preferred. The term \(p(C_{1:T}^r|C_{fr}^T, C_f^T; \hat{\theta})\) represents the conditional likelihood function evaluated at the posterior estimates for \(\hat{\theta}\) rewarding good fits, and \(p_D\) is the effective sample size, penalizing good fits achieved by means of excessively rich parameterizations. The effective sample size is defined as

\[
p_D = -2 \left[ E_{\text{post}}[\log p(C_{1:T}^r|C_{fr}^T, C_f^T; \theta]) - \log p(C_{1:T}^r|C_{fr}^T, C_f^T; \hat{\theta}) \right], \tag{22}
\]

where \(E_{\text{post}}[\log p(C_{1:T}^r|C_{fr}^T, C_f^T; \theta)]\) is the mean of the conditional log-likelihood function taken w.r.t. the posterior distribution of \(\theta\).
Based on the result for the conditional density of $C^r_t$ given $(C^{fr}_t, C^f_t)$ provided in Eq. (14) the conditional likelihood function in Eqs. (21) and (22) obtains as

$$p(C^r_{1:T} | C^{fr}_{1:T}, C^f_{1:T}; \theta) = \int \prod_{t=1}^{T} f(C^e_t | \Sigma^e_t; \theta) f(\Sigma^e_t | \Sigma^e_{1:t-1}; \theta) \, d\Sigma^e_{1:T}. \quad (23)$$

Under the Wishart density for $f(C^e_t | \Sigma^e_t; \theta)$ together with the diagonal assumption $\Sigma^e_t = \text{diag}(\exp\{x^e_{1t}\}, \ldots, \exp\{x^e_{pt}\})$ and the independent priors for the state processes $\{x^e_{i,1:T}\}$ the conditional likelihood in Eq. (23) as a function in $\{x^e_{i,1:T}\}$ factorizes into $p$ asset-specific components which are functionally independent, so that

$$p(C^r_{1:T} | C^{fr}_{1:T}, C^f_{1:T}; \theta) \propto \prod_{i=1}^{p} \int \prod_{t=1}^{T} \exp \left\{ -\frac{1}{2} \left[ (n-q)x^e_{it} + nc^e_{it} \exp(-x^e_{it}) \right] \right\} \times f_N \left( x^e_{it} \left| \gamma^e_{it} + \phi^e_{1t}x^e_{i[t-1:t-1]} + \phi^e_{2t}x^e_{i[t-1:t-5]} + \phi^e_{3t}x^e_{i[t-1:t-22]}, [\nu^e_{i}]^2 \right| \right) \, dx^e_{i,1:T}, \quad (24)$$

where $c^e_{it}$ denotes the $i$'th diagonal element of the realized idiosyncratic covariance matrix $C^e_t$. For a given value of $\theta$ the $p$ integrals w.r.t. the $x^e_{i,1:T}$'s can be taken as likelihood functions of independent univariate nonlinear non-Gaussian state space models so that they can be easily evaluated in parallel using the standard BPF. Using this BPF for the likelihood evaluation we can estimate the posterior mean of the conditional log-likelihood function in Eq. (22) by the arithmetic mean over the Gibbs draws of the parameters $\{\theta^{(i)}\}_{i=1}^{S}$, that is

$$\hat{E}_{\text{post}} \left[ \log p(C^r_{1:T} | C^{fr}_{1:T}, C^f_{1:T}; \theta) \right] = \frac{1}{S} \sum_{i=1}^{S} \log p(C^r_{1:T} | C^{fr}_{1:T}, C^f_{1:T}; \theta^{(i)}). \quad (25)$$

### 3.3 Forecasting

Using the Gibbs sampler outlined in Section 3.1 for a fixed value of parameters $\theta$, we can perform out-of-sample point- and density forecasting for the realized covariance matrix of asset returns $C^r_{t+1}$. A density forecast for $C^r_{t+1}$ obtains as

$$p(C^r_{t+1} | C_{1:t}; \theta) = \int f(C^r_{t+1} | \Sigma_{t+1}; \theta) \, f(\Sigma_{t+1} | \Sigma_{1:t}; \theta) \, \pi(\Sigma_{1:t} | C_{1:t}; \theta) \, d\Sigma_{1:t+1}. \quad (26)$$
where \( \pi(\Sigma_{1:t}|C_{1:t}; \theta) \) denotes the posterior density of the state variables in \( \Sigma_{1:t} \) for the observed data up to period \( t \), and \( f(C_{t+1}^r|\Sigma_{t+1}; \theta) \) is the marginal density for the realized covariance of the assets. Under the Wishart assumption for \( C_t \), this marginal density for \( C_{t+1}^r \) is itself a Wishart density given by \( f_W(C_{t+1}^r|n, \Sigma_{t+1}^r/n) \) (Muirhead, 2005, Corollary 3.2.6). The forecasting density (26) evaluated at the ex-post observed value for \( C_{t+1}^r \) defines the period \( t+1 \) predictive likelihood. For its computation we set the parameters \( \theta \) equal to their posterior mean estimates based on the data observed until period \( t \).

Note that the density forecast in Eq. (26) and its moments do not account for parameter estimation uncertainty. An alternative which would account for this uncertainty is to use the standard Bayesian predictive density (Geweke, 2005). While it can be straightforwardly implemented for the WFSS model using Gibbs draws from the joint posterior distribution of the state variables and the parameters, we decided to perform forecasting based on the density forecast in Eq. (26), because we shall compare the predictive performance of the WFSS model with that of alternative forecasting approaches for which it is not clear how to account for estimation uncertainty.

The predictive likelihood according to Eq. (26) can be approximated via MC integration, i.e.,

\[
p(C_{t+1}^r|C_{1:t}; \theta) \simeq \frac{1}{S} \sum_{i=1}^{S} f(C_{t+1}^r|\Sigma_{t+1}^{(i)}; \theta),
\]

(27)

where \( \{\Sigma_{t+1}^{(i)}\} \) are simulated draws from the convolution \( f(\Sigma_{t+1}|\Sigma_{1:t}; \pi(\Sigma_{1:t}|C_{1:t}; \theta)) \) based on Gibbs simulations from \( \pi(\Sigma_{1:t}|C_{1:t}; \theta) \). Using the simulated draws \( \{\Sigma_{t+1}^{(i)}\} \) the point forecast of \( C_{t+1}^r \) given by \( \mathbb{E}(C_{t+1}^r|C_{1:t}; \theta) = \mathbb{E}(\Sigma_{t+1}^r|C_{1:t}; \theta) \) can be approximated by

\[
\mathbb{E}(C_{t+1}^r|C_{1:t}; \theta) \simeq \frac{1}{S} \sum_{i=1}^{S} \Sigma_{t+1}^{r(i)}.
\]

(28)

While the simulated draws \( \{\Sigma_{t+1}^{(i)}\} \) required for the computation of the predictive likelihood in Eq. (27) can be straightforwardly obtained, the evaluation of the \((p \times p)\)-dimensional measurement density \( f(C_{t+1}^r|\Sigma_{t+1}^{(i)}; \theta) = f_W(C_{t+1}^r|n, \Sigma_{t+1}^{r(i)}/n) \) in high-dimensional applications turns out to be numerically very unstable. In fact, when the Wishart measurement density is high-dimensional as in our application for \( p = 60 \) assets its evaluation suffers from frequent floating-point underflows (see Kastner, 2017 for a discussion of similar computational problems). Thus, instead of using the predic-
tive likelihood for the covariance matrix $C_{t+1}^r$ itself, we exploit the property of Wishart distributions of being closed under linear transformations and rely on the predictive likelihood for $\iota'C_{t+1}^r\iota$ where $\iota$ denotes a vector full of ones. This transformation can be interpreted as the return variance of an equally weighted portfolio. Its predictive likelihood obtains by replacing in Eqs. (26) and (27) the multivariate Wishart $f(C_{t+1}^r|\Sigma_{t+1}; \theta) = f_W(C_{t+1}^r|n, \Sigma_{t+1}/n)$ by the corresponding univariate density $f(\iota'C_{t+1}^r\Sigma_{t+1}; \theta)$ which is a one-dimensional Wishart given by $f_W(\iota'C_{t+1}^r\iota|n, \iota'\Sigma_{t+1}\iota/n)$ (Muirhead, 2005, Theorem 3.2.5).

4 Empirical application

4.1 Data

We use the WFSS model to analyze the dynamics of the daily realized covariance matrix for 60 stocks traded at the New York Stock Exchange. The stocks are selected by liquidity from the S&P 500 index and sorted by their sector and industry classification according to the Global Industrial Classification Standard (GICS). The list of stocks covering six industry sectors is provided in Table 1. For observed risk factors we use the market, high-minus-low price-earnings ratio (HML) and small-minus-big market capitalization (SMB) factor in the Fama-French 3-factor model. In addition, we consider the sector-specific Spyder Exchange-Traded Funds (SPDR ETFs) for the six sectors covered by the 60 stocks: XLI (Industrials), XLY (Consumer Discretionary), XLP (Consumer Staples), XLV (Health Care), XLF (Financials), and XLK (Information Technologies).

The daily realized covariance matrices $C_t$ are computed 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 sector-specific ETF’s. The data comprises 2415 time series of realized variances and covariances for the sample period from January 3, 2007 to December 31, 2012, covering 1510 trading days. See Figure 2 for time-series plots of the realized variance for two randomly selected stocks (Citygroup and Caterpillar), the ETFs for the sectors of those two stocks (XLF and XLI), and the three Fama-French factors. Figure 2 also provides the sample autocorrelation functions (ACF) of the realized variance for the two selected stocks which indicate a very strong serial correlation.

Such a strong serial correlation we find for all assets and risk factors.

---

4 We are grateful to Dacheng Xiu for kindly providing us with the intraday Fama-French factor return data. The intraday return data for the 60 stocks and 6 ETFs has been obtained from QuantQuote.com.
Given the joint realized covariance matrices for the assets and the factors $C_t$ we can compute according to Eqs. (1) and (4) the realized residual covariance matrices $C^e_t$ which represent estimates for integrated residual covariance matrices $\Sigma^e_t$. In order to assess the realized residual sparsity we compute from $C^e_t$ for various sets of factors the time average of the corresponding residual correlation matrices. Figure 3 shows heat plots of this average residual correlation matrix obtained from using no factors, the CAPM factor (market), the three Fama-French factors (market, HML, SMB), and the three Fama-French factor plus the six EFTs. These plots illustrate dense asset correlations ranging from 0.18 to 0.52 (no factors). We also observe that these correlations are significantly reduced when including risk factors. Under the 9-factor case (Fama-French plus ETFs) the resulting residual correlation matrix is fairly close to diagonal with a maximum absolute average correlation of 0.22. These results are consistent with those reported by Fan et al. (2016) and justify the use of a strict factor approximation to the integrated covariance matrix of the assets $\Sigma^r_t$, especially, when combined with a Fama-French plus ETF factor structure.

4.2 Estimation results

Using the complete sample covering 1510 trading days we estimate the WFSS model for the following three factor structures: A 1-factor model with the market factor in the CAPM (1F), a 3-factor model based on the Fama-French factors (3F), and a 9-factor model including the Fama-French factors plus the six sector-specific ETFs (9F). For each of the factor structures we take the unrestricted WFSS specification defined in Eqs. (3)-(12) and compare it to model specifications obtained under the restrictions that the factor loadings are time-invariant ($\beta_{ikt} = \beta_{ik}$) and/or the log (pseudo) variances $x^e_{it}$ and $x^f_{kt}$ have a simple short-memory AR(1) dynamics ($\phi^e_{12} = \phi^e_{13} = 0$ and $\phi^f_{k2} = \phi^f_{k3} = 0$). In total we compare for each of the factor structure 1F, 3F and 9F four WFSS specifications: The unrestricted one with time-varying loadings and HAR variance dynamics (HAR-$\nu\beta$), the WFSS with constant loadings and HAR variance dynamics (HAR-c$\beta$), time-varying loadings and AR(1) variance dynamics (AR1-$\nu\beta$), and constant loadings and AR(1) variance dynamics (AR1-c$\beta$).

The prior assumptions we use for the parameters are fairly uninformative. For the $(\gamma, \phi_1, \phi_2, \phi_3, \nu^2)$ parameters in each of the Gaussian HAR processes $\{x^e_{it}\}$ and $\{x^f_{kt}\}$ we assume independent conjugate Normal-inverted Gamma priors with hyper-parameters selected such that $E(\nu^2) = 0.2$ and $\text{Var}(\nu^2) = 0.0156$, $E(\phi_1, \phi_2, \phi_3) = (0.3, 0.3, 0.3)$ and $\text{Cov}(\phi_1, \phi_2, \phi_3) = \text{diag}(0.1, 0.1, 0.1)$, $E(\gamma^*) = 0$
and \( \text{Var}(\gamma^*) = 20 \), where \( \gamma^* = \gamma(1 - \phi_1 - \phi_2 - \phi_3) \). Likewise we use for the \((\gamma, \phi, \nu^2)\) parameters in each of the Gaussian AR(1) processes \( \{\beta_{kk'}\} \) and \( \{\ell_{k,t}\} \) independent conjugate Normal-inverted Gamma priors with \( \text{E}(\nu^2) = 0.2 \) and \( \text{Var}(\nu^2) = 0.0156 \), \( \text{E}(\phi) = 0.86 \) and \( \text{Var}(\phi) = 0.1 \), \( \text{E}(\gamma^*) = 0 \) and \( \text{Var}(\gamma^*) = 20 \), where \( \gamma^* = \gamma(1 - \phi) \). This prior we also assume for the parameters of the Gaussian HAR processes \( \{x_{e,t}\} \) and \( \{x_{f,t}\} \) under the AR(1) restriction. For the Wishart degrees of freedom \( n \) we select a discrete uniform prior on the interval \((p + q, 350]\) with 1000 equally distant grid points.

For parameter estimation we run the MCMC algorithm proposed in Section 3.1 for 15,000 iterations, where the first 5,000 are discarded. The PG-AS procedure in the MCMC update step for the nonlinear state-trajectories \( \{x_{f,k,T}\} \) and \( \{x_{e,i,T}\} \) is implemented using 50 particles (see steps 1. and 3.) of the MCMC algorithm. For the computation of the DIC criterion in Section 3.2 we run the BPF using 25,000 particles. The MCMC algorithm is implemented in MATLAB. The average CPU computing time per MCMC iteration ranges from 2.4 (1F-AR1-c-\( \beta \)) to 5.9 sec (9F-HAR-\( \nu \beta \)) on an intel i7 3.4GHz processor with 4 cores. In order to evaluate the sampling efficiency of the proposed MCMC procedure for estimating the parameters of the WFSS model we compute the realized inefficiency factors for the posterior samples of the parameters\(^5\). Its benchmark value expected under prefect mixing of the MCMC draws for the parameters is equal to one. For the unrestricted full WFSS model (9F-HAR-\( \nu \beta \)) the values of the inefficiency factor across all the parameters range from 1.5 to 17.36 with an average value of 3.44 indicating a high sampling efficiency with a very fast mixing rate.

Table 3 contains the DIC values for the WFSS specifications. They show that of all factor structures the 9-factor one including the Fama-French factors and the ETFs yields the best trade-off between goodness of fit for the realized covariance matrix of the assets and parametric simplicity. This applies to all model specifications and is consistent with the results of our initial analysis showing that under the 9-factor structure the time average of the realized residual correlation matrices is closest to diagonal (see Section 4.1). Hence, these results who strong evidence against the 1-factor CAPM structure as used by Sheppard and Xu (2014) in their factor HEAVY model. We also find that the gain in terms of DIC values obtained from adding to a 3-factor Fama-French model the six sector-specific ETFs is substantially larger than that we obtain when moving from a 1-factor CAPM model

\(^5\)The inefficiency factor for the posterior sample of a parameter is defined as \( \text{IF} = 1 + 2B/(B-1) \sum_{j=1}^B K(j/B) \hat{\rho}(j) \), where \( \hat{\rho}(j) \) denotes the lag \( j \) sample autocorrelation of the MCMC draws of the parameter, \( K(\cdot) \) is the Parzen kernel function and \( B \) is the bandwidth which we set equal to \( B = 100 \) (for details, see Kim et al., 1998).
to a 3-factor structure by adding the two Fama-French factors HML and SMB. Hence, the sector-specific risk factors when considered together with the Fama-French factors appear to be extremely useful in explaining the daily variation of the assets’ covariance matrix. As for the dynamic structure, we observe that WFSSs with dynamically varying factor loadings uniformly outperform those with constant loadings, and that WFSSs with long-memory type HAR dynamics for the variances strictly dominate their short-memory AR(1) counterparts. This evidence in favor of time-varying loadings is fully consistent with the results reported in Bollerslev and Zhang (2003), Kalina (2015), Sheppard and Xu (2014), and Engle (2017). Our DIC results for the loadings are also particularly remarkable in that they reveal that the improvement in model fit achieved by allowing the loadings to be time-varying does not appear to be compromised by the involved significant inflation of parameters (e.g. 994 for the 9F-HAR-c$\beta$ versus 2,074 in the 9F-HAR-v$\beta$). Of all 12 considered WFSS models, the DIC-preferred specification for the assets’ covariance matrix is the unrestricted full 9F-HAR-v$\beta$ WFSS.

Figure 4 summarizes the parameter estimates for the DIC-preferred 9F-HAR-v$\beta$ model. It shows box-plots of the posterior mean values for all the autoregressive coefficients of the AR(1)-processes ($\phi$) and HAR-processes ($\phi_1, \phi_2, \phi_3$) directing the integrated (co)variances in ($\Sigma_f^t$, $\Sigma_e^t$) and factor loadings in $B_t$ (see Eqs. 7-12). It also provides the posterior mean values for the maximal root of the restricted AR(22) representation of the HAR processes ($z$), as well as the posterior mean values of the stationary variance ($V$) of the AR(1) and HAR processes\(^6\). As for the maximal estimated roots of the HAR processes for the assets’ residual variances ($z^e$) and the diagonals of the factor covariance matrix ($z^f$), we observe that they are close to unity implying stationarity though a very strong persistence. The estimates of the stationary variances $V^\beta$ and AR(1) roots $\phi^\beta$ reveal that the assets’ loadings on the risk factors typically exhibit a relatively small, yet significant variation and a moderate persistence. This finding is in line with the DIC values in Table 3 preferring dynamically varying factor loadings over constant ones. In Figure 2 we plot the ACFs for the realized variance of Citigroup and Caterpillar predicted under the fitted 9F-HAR-v$\beta$ WFSS together with their observed sample counterparts. They illustrate that this model combining long-memory type HAR dynamics for the variances and short-memory AR(1) dynamics for the loadings is well able to account for the observed strongly persistent movements of the assets’ realized variances. Similar results we

\(^6\)For the AR(1) processes the stationary variance obtains as $V = \nu^2/(1-\phi^2)$ and for the HAR processes analogously from their restricted AR(22) representation in Eq. (13) (see Hamilton, 1994, p. 59)
obtain for all assets. Figure 5 displays the smoothed estimates of the total integrated variance \( (\sigma_r^2) \), the residual variance \( (\sigma_e^2) \) and the loadings on the Fama-French and sector-specific risk factor \( (\beta_t) \) for the Citigroup stock under the 9F-HAR-v\( \beta \) WFSS together with the respective observed realized variances and loadings. The smoothed estimates are obtained as posterior mean values of the integrated variances and loadings computed from the MCMC posterior draws of the state variables. Unsurprisingly, we see that the smoothed estimates follow closely the variation in their realized counterparts.

In order to quantify the relative importance of the risk factors we provide in Figure 5 their estimated explained contributions to the assets’ total variances under the HAR-v\( \beta \) WFSS with 1 (market), 3 (market+HML+SMB) and 9 factors (market+HML+SMB+ETFs), respectively. For asset \( i \) the estimated factor contribution under a given factor structure is computed as

\[
1 - \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}(\sigma_{it}^2 | C_{1:T}) / c_{it}^2,
\]

where \( \mathbb{E}(\sigma_{it}^2 | C_{1:T}) \) is the posterior mean of the integrated residual variance and \( c_{it}^2 \) the total realized variance. The explained contribution of the market risk under the 1-factor CAPM is quite substantial and varies between 18% and 35%. Combining the market risk with the HML and SMB factor only moderately increases the explained contribution to a range between 23% and 37%, while a significant increase in that contribution up to a level of 51% is obtained by adding the ETFs to the market, HML and SMB. This corroborates our previous results on the relative importance of the HML, SMB and sector-specific risk factors.

### 4.3 Out-of-sample forecasting results

We now analyze the out-of-sample forecasting performance of the WFSS for the realized covariance matrix of asset returns \( C_r^T \) and compare it to alternative forecasting approaches. As alternative approaches we consider the Exponentially Weighted Moving Average (EWMA) model (Morgan, 1996) and the scalar Realized consistent Dynamic Conditional Correlation (sRe-cDCC) model of Bauwens et al. (2016). The EWMA is popular in industry practice while the sRe-cDCC model is found by Bauwens et al. to be highly effective in predicting large covariance matrices of asset returns outperforming alternative popular high-dimensional forecasting models.

The EWMA is given by

\[
\mathbb{E}(C_{t+1}^r | C_{1:t}^r) = (1 - \lambda)C_t^r + \lambda \mathbb{E}(C_t^r | C_{1:t-1}^r),
\]
where we set the smoothing parameter $\lambda$ to its typically selected value of 0.96 (Callot et al., 2017). The sRe-DCC model assumes for $C_t^r$ given $C_{1:t-1}^r$ a Wishart distribution, $C_t^r \mid C_{1:t-1}^r \sim W_p(n, S_t/n)$, where the scale matrix is factorized as

$$S_t = AV_t R_t V_tA',$$

with $A$ being the lower triangular Cholesky factor of $\text{E}(C_t^r) = \text{E}(S_t)$. The matrix $V_t$ is defined as $V_t = \text{diag}(\sqrt{v_{1t}}, \ldots, \sqrt{v_{pt}})$, where $v_{it}$ is the conditional variance of asset $i$, and $R_t$ is the conditional correlation matrix obtained as $R_t = (Q_t \odot I_p)^{-1/2}Q_t(Q_t \odot I_p)^{-1/2}$, where $\odot$ denotes the Hadamard element-by-element product. Both the $v_{it}$’s and $Q_t$ are endowed with GARCH-type recursions of the form

$$v_{it} = (1 - \kappa_i - \delta_i) + \kappa_i c_{it-1}^* + \delta_i v_{it-1}, \quad i = 1, \ldots, p,$$

$$Q_t = (1 - \alpha - \varphi)I_p + \alpha C_{t-1}^Q + \varphi Q_{t-1},$$

where $c_{it}^*$ is the $i$th diagonal element of $C_t^* = A^{-1}C_t^r A^{-1}'$ and $C_t^Q = (Q_t \odot I_p)^{1/2}V_t^{-1}C_t^r V_t^{-1}(Q_t \odot I_p)^{1/2}$, with $I_p$ denoting the $p$-dimensional identity matrix. The model parameters are given by the scalars $\{\kappa_i\}, \{\delta_i\}, \alpha, \varphi$ and are estimated by the three-step Quasi-ML (QML) procedure proposed by Bauwens et al. (2016).

In our forecast experiments we focus on 1-day-ahead predictions. They are obtained by re-estimating the model parameters every 10 trading days on a rolling 4-year window with 1008 daily observations and then producing a sequence of new 1-day-ahead forecasts based on the updated parameter estimates. We consider two out-of-sample forecasting periods each covering one year with 251 trading days (see the shaded areas in the time series plots in Figure 2). The first period is the year 2011 where the volatility ranges from a small to a relatively high level, and the second period covers the year 2012 with a constantly fairly low volatility.

### 4.3.1 Statistical forecast evaluation

For statistical forecast evaluation we rely on the predictive likelihood for the assets’ realized variances and the accuracy of the point forecasts for the covariance matrix (see Section 3.3).

To assess the point forecast accuracy we follow Ledoit et al. (2003) and use the root-mean-squared-
error (RMSE) based on the Frobenius norm comparing the covariance matrix forecast \( \hat{C}_t \) and the ex-post observed value for \( C_t \). This RMSE is given by

\[
\text{RMSE} = \frac{1}{T^*} \sum_t \| C_t - \hat{C}_t \| = \frac{1}{T^*} \sum_t \left[ \sum_i (c_{it} - \hat{c}_{it})^2 + 2 \sum_{i<j} (c_{ijt} - \hat{c}_{ijt})^2 \right]^{1/2}, \tag{33}
\]

where \( c_{it} \) and \( c_{ijt} \) denote the realized variance of asset \( i \) and the realized covariance between asset \( i \) and \( j \), respectively, and \( \hat{c}_{it} \) and \( \hat{c}_{ijt} \) their forecasts\(^7\). \( T^* \) is the number of forecast periods. In order to disentangle the forecast performance w.r.t. the different elements in the covariance matrix we also compute the RMSE separately for the variances and covariances, i.e.,

\[
\text{RMSE}^V = \frac{1}{T^*} \sum_t \left[ \sum_i (c_{it} - \hat{c}_{it})^2 \right]^{1/2}, \quad \text{RMSE}^C = \frac{1}{T^*} \sum_t \left[ \sum_{i<j} (c_{ijt} - \hat{c}_{ijt})^2 \right]^{1/2}. \tag{34}
\]

For assessing the significance of differences in the RMSE, RMSE\(^V\) and RMSE\(^C\) losses across models we rely on the model confidence set (MCS) approach of Hansen et al. (2011). The MCS identifies the model or set of models having the best forecasting performance at a given confidence level. The MSC is computed for confidence levels of 75\% and 90\% using a block bootstrap with block length \( \lfloor (T^*)^{1/3} \rfloor \) and 10,000 bootstrap replications.

The RMSE results for the out-of-sample periods 2011 and 2012 as well as for the two periods aggregated together are summarized in Table 3. They reveal that for forecasting the assets’ covariance matrix based on the WFSS approach it is important to make use of all the 9 available risk factors and to account for time-variation in the factor loadings. In fact, for the fairly volatile period 2011 it is the 9-factor WFSS with time-varying loadings and AR(1) variance dynamics (9F-AR1-v\( \beta \)) which exhibits among all WFSS models the smallest RMSE loss and for the low-volatility period 2012, the 9-factor WFSS with time-varying loadings and HAR variance dynamics (9F-HAR-v\( \beta \)). The later is also the best performing WFSS for the aggregated period 2011-2012. Comparing the RMSE losses with RMSE\(^V\) and RMSE\(^C\) values indicates that the comparatively good predictive performance of those two 9-factor WFSS models with time-varying loadings is due to their ability to produce relatively precise forecasts, especially for the covariances. However, if it is the forecast performance only for the variances, the best WFSS is the 9-factor model with constant loadings and HAR dynamics (9F-

\(^7\)Alternative frequently applied loss functions are the Stein loss, the QLIKE and the von Neumann divergence (Bauwens et al., 2014). However, for parameter-driven state-space models as the WFSS these loss functions are not guaranteed to provide consistent rankings of the forecasting performance (Laurent et al., 2013).
HAR-cβ). As for the comparison of the WFSS approach with the competing alternatives, we find that for the periods 2011 and 2011-2012 the RMSE loss of the best-performing WFSS is at the 75% MCS-confidence level significantly smaller than that of the EWMA but is on par with the loss of the sRe-cDCC model. For the 2012 period both, the EWMA and sRe-cDCC are at the 75% level significantly outperformed by the best WFSS. Hence, overall our WFSS approach with 9 factors and dynamically varying loadings performs favorably in relation to the competing models.

Figure 7 shows the time-series plots of the predictive likelihood. ....

4.3.2 Value-at-risk forecasts

For an economic evaluation of predictive performance we consider Value-at-Risk (VaR) forecasts for portfolios constructed from the \( p \) assets. For simplicity, we focus on an equally-weighted portfolio with return \( \omega' r_t \), where \( r_t \) denotes the \( p \)-dimensional vector of the period-\( t \) asset returns and \( \omega \) is the vector of equal portfolio weights. Assuming for the 1-period-ahead predictive distribution of the portfolio returns a normal \( \mathcal{N}(0, \omega' \hat{C}_t \omega) \), the predicted period-\( t \) portfolio VaR at level \( \alpha \) obtains as

\[
\text{VaR}_t(\alpha) = z_\alpha \sqrt{\omega' \hat{C}_t \omega},
\]

where \( z_\alpha \) is the \( \alpha\% \) quantile of a standard normal distribution. The normal distribution for the portfolio returns can also be replaced by a more flexible student-\( t \). However, the results we obtained under the student-\( t \) (not reported here) do not qualitatively differ from those for the normal distribution.

For assessing the accuracy of the predicted VaR we follow Chib et al. (2006) and test for unconditional and conditional coverage based on the ‘hit-indicator’ variable \( I_t = 1[\omega' r_t \leq \text{VaR}_t(\alpha)] \) for \( t = 1, \ldots, T^* \), signaling that the realized portfolio return is lower or equal than the predicted VaR. The hypothesis of correct unconditional coverage can be tested with the likelihood-ratio statistic (Kupiec, 1995)

\[
\text{LR}_{UC} = 2 \left\{ \ln \left[ \hat{\alpha}^{\hat{\alpha} T^*} (1 - \hat{\alpha})^{(1 - \hat{\alpha}) T^*} \right] - \ln \left[ \alpha^{\alpha T^*} (1 - \alpha)^{(1 - \alpha) T^*} \right] \right\},
\]

where \( \hat{\alpha} \) is the hit-rate defined as \( \hat{\alpha} = \sum_{t=1}^{T^*} I_t / T^* \). Under the hypothesis that the observed hit-rate \( \hat{\alpha} \) is equal to the nominal level \( \alpha \) (correct unconditional coverage) the statistic \( \text{LR}_{UC} \) is distributed asymptotically as a \( \chi^2_{(1)} \). Conditional coverage can be tested by jointly testing for unconditional
coverage and serial independence of the hit-indicator sequence \( \{I_t\} \) with the statistic (Christoersen, 1998)

\[
LR_{CC} = 2 \left\{ \ln \left[ (1 - \pi_{01})^{T_{01}} \pi_{01}^{T_{01}} (1 - \pi_{11})^{T_{11}} \pi_{11}^{T_{11}} \right] - \ln [ \alpha^{\hat{\alpha} T^*} (1 - \alpha)^{(1 - \hat{\alpha}) T^*} ] \right\},
\]

(37)

where \( T_{ij} \) denotes the number of cases for which we observe \( I_t = j \) and \( I_{t-1} = i \) for \( i, j \in \{0, 1\} \) while \( \pi_{01} = T_{01}/(T_{01} + T_{00}) \) and \( \pi_{11} = T_{11}/(T_{10} + T_{11}) \). Under the joint hypothesis of correct unconditional coverage and serial independence \( LR_{CC} \) is distributed asymptotically as a \( \chi^2 \).

The \( p \)-values of the coverage tests for the predicted 1% and 5% VaR are shown in Table 4. They reinforce our earlier results on the forecast accuracy for the covariance matrix of the assets. The best performing WFSS models for both the 1% and 5% VaR are those with all 9 risk factors and time-varying loadings (9F-AR1-v\( \beta \) and 9F-HAR-v\( \beta \)). For the year 2011 as well as the year 2012 they pass the tests for unconditional and conditional coverage at the 1% significance level. However, for the aggregated period 2011-2012 both models fail to pass the unconditional coverage test for the 1% VaR as they significantly underestimate the VaR, but note that this applies to all models. Overall, the performance of the 9-factor WFSS with time varying loadings is better than that of the EWMA and not worse than that of sRe-cDCC model.

4.3.3 Global-minimum-variance-portfolio forecasts

As a further economic experiment designed to evaluate the forecasting performance we use the predicted covariance matrices of the assets to construct optimal investment portfolios (Bauwens et al., 2016 and Callot et al. 2016). For portfolio allocation we consider a strategy based on the global-minimum-variance portfolio (GMVP) which has the advantage relative to minimum-variance portfolios that its ex-ante portfolio weights only depend on the covariance matrix. When constructing the GMVP we exclude short-selling by imposing the portfolio weights to be non-negative. This is rather typical in high-dimensional portfolio allocation problems and acts like a regularization device for reducing the impact of errors in covariance matrix forecasts on the allocation (Frost and Savarino, 1986, Jagannathan and Ma, 2003 and Bauwens et al., 2016).

For a given covariance matrix forecast \( \hat{C}_t \) computed in period \( t-1 \), the GMVP weights \( \hat{w}_t \) obtain
by solving the minimization problem

\[ \hat{w}_t = \arg \min_{w_t} \mathbf{w}_t' \hat{\mathbf{C}}_t \mathbf{w}_t, \quad \text{subject to} \quad \sum_{i=1}^{p} w_{it}, \quad w_{it} > 0 \ \forall i, \]  

(38)

where \( w_t = (w_{1t}, \ldots, w_{pt})' \) is the vector of period-\( t \) portfolio weights to be selected in period \( t - 1 \).

For assessing the relative capabilities of the competing models for optimal portfolio allocation we use the following four measures:

(i) The average out-of-sample portfolio return defined as \( \hat{\mu}_p = \frac{\sum_{t=1}^{T^*} r_{pt}}{T^*} \) with \( r_{pt} = \hat{w}_t' \mathbf{r}_t \);

(ii) The accumulated portfolio return over the out-of-sample window \( \hat{\mu}_p^{\text{cum}} = \left[ \prod_{t=1}^{T^*} (1 + r_{pt}) \right] - 1 \);

(iii) The portfolio return standard deviation over the out-of-sample window \( \hat{\sigma}_p = \sqrt{\frac{\sum_{t=1}^{T^*} (r_{pt} - \hat{\mu}_p)^2}{T^*}} \);

(iv) The Sharpe ratio \( \text{SR}_p = \frac{\hat{\mu}_p}{\hat{\sigma}_p} \).

The results are summarized in Table 5. As expected, for the low-volatility period 2012 the portfolio return standard deviations are uniformly smaller across all forecasting models than for the more turbulent year 2011. On the other hand, all models generate for 2012 also a lower average and accumulated return than for 2011. As for the comparison of the competing models, we find that the EWMA and sRe-cDCC are outperformed in terms of all measures by the respective best WFSS. In terms of minimal portfolio risk, the 9-factor WFSS HAR model with constant factor loadings (9F-HAR-c\( \beta \)) shows across all periods the best performance, though the differences to the other models are typically not very large. The highest average and accumulated return as well as the largest Sharpe ratio are obtained by portfolio allocations based on WFSSs with time-varying loadings: For period 2011 it is the 3-factor WFSS HAR and for 2012 the 1-factor WFSS AR1 model.

5 Conclusion
References


<table>
<thead>
<tr>
<th>#</th>
<th>Symbol</th>
<th>Company</th>
<th>Sector</th>
<th>#</th>
<th>Symbol</th>
<th>Company</th>
<th>Sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>BA</td>
<td>Boeing Company</td>
<td>I</td>
<td>31</td>
<td>MDT</td>
<td>Medtronic Inc.</td>
<td>H</td>
</tr>
<tr>
<td>2</td>
<td>CAT</td>
<td>Caterpillar, Inc.</td>
<td>I</td>
<td>32</td>
<td>BAC</td>
<td>Bank of America Corporation</td>
<td>F</td>
</tr>
<tr>
<td>3</td>
<td>DE</td>
<td>Deere &amp; Company</td>
<td>I</td>
<td>33</td>
<td>C</td>
<td>Citigroup Inc.</td>
<td>F</td>
</tr>
<tr>
<td>4</td>
<td>EMR</td>
<td>Emerson Electric Company</td>
<td>I</td>
<td>34</td>
<td>FITB</td>
<td>Fifth Third Bankcorp</td>
<td>F</td>
</tr>
<tr>
<td>5</td>
<td>GE</td>
<td>General Electric Company</td>
<td>I</td>
<td>35</td>
<td>JPM</td>
<td>J P Morgan Chase &amp; Co.</td>
<td>F</td>
</tr>
<tr>
<td>6</td>
<td>GLW</td>
<td>Coming Incorporated</td>
<td>I</td>
<td>36</td>
<td>AXP</td>
<td>American Express Company</td>
<td>F</td>
</tr>
<tr>
<td>7</td>
<td>HON</td>
<td>Honeywell International Inc.</td>
<td>I</td>
<td>37</td>
<td>GS</td>
<td>Goldman Sachs Group, Inc.</td>
<td>F</td>
</tr>
<tr>
<td>8</td>
<td>LB</td>
<td>La Barge Inc.</td>
<td>I</td>
<td>38</td>
<td>MS</td>
<td>Morgan Stanley</td>
<td>F</td>
</tr>
<tr>
<td>9</td>
<td>LMT</td>
<td>Lockheed Martin Corporation</td>
<td>I</td>
<td>39</td>
<td>ADBE</td>
<td>Adobe Systems Incorporated</td>
<td>T</td>
</tr>
<tr>
<td>10</td>
<td>AMZN</td>
<td>Amazon.com, Inc.</td>
<td>D</td>
<td>40</td>
<td>ADP</td>
<td>Automatic Data Processing</td>
<td>T</td>
</tr>
<tr>
<td>11</td>
<td>BBBY</td>
<td>Bed Bath &amp; Beyond Inc.</td>
<td>D</td>
<td>41</td>
<td>CTXS</td>
<td>Citrix Systems, Inc.</td>
<td>T</td>
</tr>
<tr>
<td>12</td>
<td>BBY</td>
<td>Best Buy Co., Inc.</td>
<td>D</td>
<td>42</td>
<td>ES</td>
<td>Electronic Arts, Inc.</td>
<td>T</td>
</tr>
<tr>
<td>13</td>
<td>GPS</td>
<td>Gap, Inc.</td>
<td>D</td>
<td>43</td>
<td>EBAY</td>
<td>Ebay, Inc.</td>
<td>T</td>
</tr>
<tr>
<td>14</td>
<td>HD</td>
<td>Home Depot, Inc.</td>
<td>D</td>
<td>44</td>
<td>EMC</td>
<td>EMC Corporation MA</td>
<td>T</td>
</tr>
<tr>
<td>15</td>
<td>JCP</td>
<td>J.C. Penney Company</td>
<td>D</td>
<td>45</td>
<td>IBM</td>
<td>International Business Machines</td>
<td>T</td>
</tr>
<tr>
<td>16</td>
<td>KSS</td>
<td>Kohl’s Corporation</td>
<td>D</td>
<td>46</td>
<td>INTU</td>
<td>Intuit, Inc.</td>
<td>T</td>
</tr>
<tr>
<td>17</td>
<td>LOW</td>
<td>Lowe’s Companies, Inc.</td>
<td>D</td>
<td>47</td>
<td>MSFT</td>
<td>Microsoft Corporation</td>
<td>T</td>
</tr>
<tr>
<td>18</td>
<td>COST</td>
<td>Costco Wholesale Corporation</td>
<td>S</td>
<td>48</td>
<td>NTAP</td>
<td>NetApp, Inc.</td>
<td>T</td>
</tr>
<tr>
<td>19</td>
<td>CVS</td>
<td>CVS Caremark Group</td>
<td>S</td>
<td>49</td>
<td>ORCL</td>
<td>Oracle Corporation</td>
<td>T</td>
</tr>
<tr>
<td>20</td>
<td>KR</td>
<td>Kroger Company</td>
<td>S</td>
<td>50</td>
<td>ADI</td>
<td>Analog Devices, Inc.</td>
<td>T</td>
</tr>
<tr>
<td>21</td>
<td>CAG</td>
<td>ConAgra, Inc.</td>
<td>S</td>
<td>51</td>
<td>ALTR</td>
<td>Altera Corporation</td>
<td>T</td>
</tr>
<tr>
<td>22</td>
<td>KO</td>
<td>Coca-Cola Company</td>
<td>S</td>
<td>52</td>
<td>AMAT</td>
<td>Applied Materials, Inc.</td>
<td>T</td>
</tr>
<tr>
<td>23</td>
<td>MO</td>
<td>Altria Group</td>
<td>S</td>
<td>53</td>
<td>AMD</td>
<td>Advanced Micro Devices, Inc.</td>
<td>T</td>
</tr>
<tr>
<td>24</td>
<td>PEP</td>
<td>PepsiCo, Inc.</td>
<td>S</td>
<td>54</td>
<td>BRCM</td>
<td>Broadcom Corporation</td>
<td>T</td>
</tr>
<tr>
<td>25</td>
<td>AG</td>
<td>Agilent Technologies Inc.</td>
<td>H</td>
<td>55</td>
<td>INTC</td>
<td>Intel Corporation</td>
<td>T</td>
</tr>
<tr>
<td>26</td>
<td>ABT</td>
<td>Abbott Laboratories</td>
<td>H</td>
<td>56</td>
<td>KLAC</td>
<td>KLA-Tencor Corporation</td>
<td>T</td>
</tr>
<tr>
<td>27</td>
<td>BAX</td>
<td>Baxter International Inc.</td>
<td>H</td>
<td>57</td>
<td>LLTC</td>
<td>Linear Technology Corporation</td>
<td>T</td>
</tr>
<tr>
<td>28</td>
<td>BMY</td>
<td>Bristol-Myers Squibb Company</td>
<td>H</td>
<td>58</td>
<td>MU</td>
<td>Micron Technology, Inc.</td>
<td>T</td>
</tr>
<tr>
<td>29</td>
<td>CAH</td>
<td>Cardinal Health Inc.</td>
<td>H</td>
<td>59</td>
<td>NVDA</td>
<td>NVIDIA Corporation</td>
<td>T</td>
</tr>
<tr>
<td>30</td>
<td>JNJ</td>
<td>Johnson &amp; Johnson</td>
<td>H</td>
<td>60</td>
<td>QCOM</td>
<td>QUALCOMM Incorporated</td>
<td>T</td>
</tr>
</tbody>
</table>

Table 1: List of the stocks included in the data set. Stocks are selected by liquidity from the S&P 500 index and sorted by their sector and industry classification according to the Global Industrial Classification Standard (GICS). The sector labels are: (I) Industrials; (D) Consumer Discretionary; (S) Consumer Staples; (H) Health Care; (F) Financials; (T) Information Technologies.
<table>
<thead>
<tr>
<th>WFSS model</th>
<th># params</th>
<th># obs./param</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1F-AR1-cβ</td>
<td>244</td>
<td>11,703</td>
<td>-2,860,015</td>
</tr>
<tr>
<td>3F-AR1-cβ</td>
<td>379</td>
<td>8,032</td>
<td>-3,414,454</td>
</tr>
<tr>
<td>9F-AR1-cβ</td>
<td>856</td>
<td>4,260</td>
<td>-6,030,216</td>
</tr>
<tr>
<td>1F-AR1-vβ</td>
<td>364</td>
<td>7,845</td>
<td>-2,868,559</td>
</tr>
<tr>
<td>3F-AR1-vβ</td>
<td>739</td>
<td>4,119</td>
<td>-3,469,892</td>
</tr>
<tr>
<td>9F-AR1-vβ</td>
<td>1,936</td>
<td>1,884</td>
<td>-6,266,725</td>
</tr>
<tr>
<td>1F-HAR-cβ</td>
<td>366</td>
<td>7,802</td>
<td>-2,874,627</td>
</tr>
<tr>
<td>3F-HAR-cβ</td>
<td>505</td>
<td>6,028</td>
<td>-3,428,088</td>
</tr>
<tr>
<td>9F-HAR-cβ</td>
<td>994</td>
<td>3,669</td>
<td>-6,041,832</td>
</tr>
<tr>
<td>1F-HAR-vβ</td>
<td>486</td>
<td>5,875</td>
<td>-2,882,791</td>
</tr>
<tr>
<td>3F-HAR-vβ</td>
<td>865</td>
<td>3,519</td>
<td>-3,483,658</td>
</tr>
<tr>
<td>9F-HAR-vβ</td>
<td>2,074</td>
<td>1,758</td>
<td>-6,279,064</td>
</tr>
</tbody>
</table>

Table 2: Deviance information criteria (DIC) for the WFSS model specifications. The likelihood for the DICs given in Eq. (23) is computed using a bootstrap particle filter with 25,000 particles. # params is the number of parameters and # obs/param the number of observations per parameter.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>RMSE(^V)</td>
<td>RMSE(^C)</td>
<td>RMSE</td>
<td>RMSE(^V)</td>
<td>RMSE(^C)</td>
<td>RMSE</td>
<td>RMSE(^V)</td>
<td>RMSE(^C)</td>
</tr>
<tr>
<td>WFSS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1F-AR1-c(\beta)</td>
<td>56.98</td>
<td>18.28</td>
<td>37.02</td>
<td>25.89</td>
<td>12.62</td>
<td>15.41</td>
<td>41.50</td>
<td>15.46</td>
<td>26.26</td>
</tr>
<tr>
<td>3F-AR1-c(\beta)</td>
<td>58.57</td>
<td>18.23</td>
<td>38.15</td>
<td>26.06</td>
<td>12.54</td>
<td>15.55</td>
<td>42.38</td>
<td>15.40</td>
<td>26.90</td>
</tr>
<tr>
<td>9F-AR1-c(\beta)</td>
<td>52.33</td>
<td>17.82</td>
<td>33.93</td>
<td>25.08</td>
<td>12.11</td>
<td>15.06</td>
<td>38.76</td>
<td>14.97</td>
<td>24.53</td>
</tr>
<tr>
<td>1F-AR1-v(\beta)</td>
<td>52.41</td>
<td>19.10</td>
<td>33.42</td>
<td>25.17</td>
<td>13.03</td>
<td>14.65</td>
<td>38.84</td>
<td>16.08</td>
<td>24.08</td>
</tr>
<tr>
<td>3F-AR1-v(\beta)</td>
<td>53.11</td>
<td>18.67</td>
<td>34.01</td>
<td>25.23</td>
<td>12.83</td>
<td>14.76</td>
<td>39.22</td>
<td>15.76</td>
<td>24.42</td>
</tr>
<tr>
<td>9F-AR1-v(\beta)</td>
<td><strong>50.58</strong></td>
<td>18.39</td>
<td><strong>32.51</strong></td>
<td>25.14</td>
<td>12.55</td>
<td>14.95</td>
<td>37.91</td>
<td>15.48</td>
<td>23.77</td>
</tr>
<tr>
<td>1F-HAR-c(\beta)</td>
<td>58.35</td>
<td>16.77</td>
<td>38.69</td>
<td>25.37</td>
<td>10.67</td>
<td>15.80</td>
<td>41.14</td>
<td>13.59</td>
<td>26.74</td>
</tr>
<tr>
<td>3F-HAR-c(\beta)</td>
<td>59.18</td>
<td>16.73</td>
<td>39.30</td>
<td>25.52</td>
<td>10.66</td>
<td>15.91</td>
<td>41.62</td>
<td>13.56</td>
<td>27.09</td>
</tr>
<tr>
<td>9F-HAR-c(\beta)</td>
<td><strong>53.46</strong></td>
<td><strong>16.68</strong></td>
<td>35.22</td>
<td>23.84</td>
<td><strong>10.54</strong></td>
<td>14.71</td>
<td>38.00</td>
<td><strong>13.48</strong></td>
<td>24.52</td>
</tr>
<tr>
<td>1F-HAR-v(\beta)</td>
<td>53.62</td>
<td>17.14</td>
<td>35.09</td>
<td>24.09</td>
<td>10.82</td>
<td>14.75</td>
<td>38.21</td>
<td>13.84</td>
<td>24.47</td>
</tr>
<tr>
<td>3F-HAR-v(\beta)</td>
<td>54.36</td>
<td>16.97</td>
<td>35.65</td>
<td>24.61</td>
<td>10.75</td>
<td>15.17</td>
<td>38.83</td>
<td>13.72</td>
<td>24.96</td>
</tr>
<tr>
<td>9F-HAR-v(\beta)</td>
<td><strong>51.24</strong></td>
<td>16.87</td>
<td><strong>33.54</strong></td>
<td><strong>23.02</strong></td>
<td>10.63</td>
<td><strong>14.03</strong></td>
<td><strong>36.51</strong></td>
<td><strong>13.61</strong></td>
<td><strong>23.36</strong></td>
</tr>
<tr>
<td>EWMA</td>
<td>57.78</td>
<td>18.64</td>
<td>38.02</td>
<td>23.15</td>
<td>10.81</td>
<td>14.14</td>
<td>41.62</td>
<td>14.88</td>
<td>26.93</td>
</tr>
<tr>
<td>sRe-cDCC</td>
<td>53.13</td>
<td>17.92</td>
<td>34.60</td>
<td>24.94</td>
<td>11.27</td>
<td>15.43</td>
<td>39.12</td>
<td>14.60</td>
<td>25.07</td>
</tr>
</tbody>
</table>

Table 3: Evaluation of forecasting accuracy. The table reports the RMSE, RMSE\(^V\), and RMSE\(^C\) losses as given in Eqs. (33) and (34). Bold figures indicate the smallest loss; Light grey-shaded cells indicate that the respective model is in the 90% model confidence set, dark grey-shaded cells indicate that the respective model belongs to the 75% model confidence set.
<table>
<thead>
<tr>
<th>Model</th>
<th>Period 2011 LR&lt;sub&gt;UC&lt;/sub&gt;</th>
<th>LR&lt;sub&gt;CC&lt;/sub&gt;</th>
<th>Period 2012 LR&lt;sub&gt;UC&lt;/sub&gt;</th>
<th>LR&lt;sub&gt;CC&lt;/sub&gt;</th>
<th>Period 2011 - 2012 LR&lt;sub&gt;UC&lt;/sub&gt;</th>
<th>LR&lt;sub&gt;CC&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>WFSS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1F-AR1-c&lt;sub&gt;β&lt;/sub&gt;</td>
<td>0.00 0.05</td>
<td>0.00 0.02</td>
<td>0.00 0.01</td>
<td>0.00 0.01</td>
<td>0.00 0.00</td>
<td>0.00 0.00</td>
</tr>
<tr>
<td>3F-AR1-c&lt;sub&gt;β&lt;/sub&gt;</td>
<td>0.00 0.00</td>
<td>0.00 0.01</td>
<td>0.00 0.00</td>
<td>0.00 0.01</td>
<td>0.00 0.00</td>
<td>0.00 0.00</td>
</tr>
<tr>
<td>9F-AR1-c&lt;sub&gt;β&lt;/sub&gt;</td>
<td>0.01 0.14</td>
<td>0.02 0.08</td>
<td>0.02 0.21</td>
<td>0.05 0.13</td>
<td>0.00 0.03</td>
<td>0.00 0.05</td>
</tr>
<tr>
<td>1F-AR1-v&lt;sub&gt;β&lt;/sub&gt;</td>
<td>0.00 0.14</td>
<td>0.00 0.08</td>
<td>0.00 0.01</td>
<td>0.00 0.03</td>
<td>0.00 0.00</td>
<td>0.00 0.02</td>
</tr>
<tr>
<td>3F-AR1-v&lt;sub&gt;β&lt;/sub&gt;</td>
<td>0.00 0.05</td>
<td>0.00 0.02</td>
<td>0.00 0.00</td>
<td>0.00 0.01</td>
<td>0.00 0.00</td>
<td>0.00 0.00</td>
</tr>
<tr>
<td>9F-AR1-v&lt;sub&gt;β&lt;/sub&gt;</td>
<td>0.02 0.49</td>
<td>0.05 0.30</td>
<td>0.06 0.32</td>
<td>0.14 0.61</td>
<td>0.00 0.24</td>
<td>0.01 0.37</td>
</tr>
<tr>
<td>1F-HAR-c&lt;sub&gt;β&lt;/sub&gt;</td>
<td>0.00 0.01</td>
<td>0.00 0.02</td>
<td>0.00 0.01</td>
<td>0.00 0.01</td>
<td>0.00 0.00</td>
<td>0.00 0.00</td>
</tr>
<tr>
<td>3F-HAR-c&lt;sub&gt;β&lt;/sub&gt;</td>
<td>0.00 0.01</td>
<td>0.00 0.01</td>
<td>0.00 0.00</td>
<td>0.00 0.01</td>
<td>0.00 0.00</td>
<td>0.00 0.00</td>
</tr>
<tr>
<td>9F-HAR-c&lt;sub&gt;β&lt;/sub&gt;</td>
<td>0.01 0.02</td>
<td>0.02 0.08</td>
<td>0.01 0.02</td>
<td>0.02 0.02</td>
<td>0.00 0.00</td>
<td>0.00 0.00</td>
</tr>
<tr>
<td>1F-HAR-v&lt;sub&gt;β&lt;/sub&gt;</td>
<td>0.00 0.02</td>
<td>0.00 0.08</td>
<td>0.01 0.02</td>
<td>0.02 0.02</td>
<td>0.00 0.00</td>
<td>0.00 0.00</td>
</tr>
<tr>
<td>3F-HAR-v&lt;sub&gt;β&lt;/sub&gt;</td>
<td>0.00 0.01</td>
<td>0.00 0.02</td>
<td>0.00 0.01</td>
<td>0.00 0.01</td>
<td>0.00 0.00</td>
<td>0.00 0.00</td>
</tr>
<tr>
<td>9F-HAR-v&lt;sub&gt;β&lt;/sub&gt;</td>
<td>0.01 0.14</td>
<td>0.02 0.27</td>
<td>0.02 0.08</td>
<td>0.05 0.03</td>
<td>0.00 0.01</td>
<td>0.00 0.01</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.00 0.22</td>
<td>0.00 0.03</td>
<td>0.02 0.32</td>
<td>0.05 0.13</td>
<td>0.00 0.12</td>
<td>0.00 0.00</td>
</tr>
<tr>
<td>sRes-cDCC</td>
<td>0.06 0.49</td>
<td>0.15 0.78</td>
<td>0.02 0.67</td>
<td>0.05 0.55</td>
<td>0.00 0.55</td>
<td>0.00 0.46</td>
</tr>
</tbody>
</table>

Table 4: VaR forecasting results. The table reports the p-values for the unconditional coverage test LR<sub>UC</sub> and conditional coverage test LR<sub>CC</sub> for VaR predictions at the 1% and 5% level. Bold figures indicate that the null hypothesis cannot be rejected at the 1% significance level.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\sigma}_p \times 10^3$</td>
<td>$\hat{\mu}_p \times 10^2$</td>
<td>$\hat{\mu}_p^{cum}$</td>
</tr>
<tr>
<td></td>
<td>$\hat{\sigma}_p \times 10^3$</td>
<td>$\hat{\mu}_p \times 10^2$</td>
<td>$\hat{\mu}_p^{cum}$</td>
</tr>
<tr>
<td>WFSS-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1F-AR1-c$\beta$</td>
<td>6.78</td>
<td>4.17</td>
<td>6.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9.91</td>
</tr>
<tr>
<td>3F-AR1-c$\beta$</td>
<td>6.75</td>
<td>4.45</td>
<td>6.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10.7</td>
</tr>
<tr>
<td>9F-AR1-c$\beta$</td>
<td>6.56</td>
<td>4.58</td>
<td>6.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11.1</td>
</tr>
<tr>
<td>1F-AR1-v$\beta$</td>
<td>6.66</td>
<td>4.48</td>
<td>6.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10.8</td>
</tr>
<tr>
<td>3F-AR1-v$\beta$</td>
<td>6.67</td>
<td>4.44</td>
<td>6.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10.6</td>
</tr>
<tr>
<td>9F-AR1-v$\beta$</td>
<td>6.79</td>
<td>4.51</td>
<td>6.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10.8</td>
</tr>
<tr>
<td>1F-HAR-c$\beta$</td>
<td>6.77</td>
<td>3.89</td>
<td>5.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9.19</td>
</tr>
<tr>
<td>3F-HAR-c$\beta$</td>
<td>6.73</td>
<td>4.16</td>
<td>6.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9.91</td>
</tr>
<tr>
<td>9F-HAR-c$\beta$</td>
<td>6.55</td>
<td>4.15</td>
<td>6.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9.90</td>
</tr>
<tr>
<td>1F-HAR-v$\beta$</td>
<td>6.56</td>
<td>4.33</td>
<td>6.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10.4</td>
</tr>
<tr>
<td>3F-HAR-v$\beta$</td>
<td>6.61</td>
<td>4.65</td>
<td>7.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11.2</td>
</tr>
<tr>
<td>9F-HAR-v$\beta$</td>
<td>6.63</td>
<td>4.56</td>
<td>6.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11.0</td>
</tr>
<tr>
<td>EWMA</td>
<td>6.72</td>
<td>3.32</td>
<td>4.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7.72</td>
</tr>
<tr>
<td>sRe-cDCC</td>
<td>6.70</td>
<td>4.33</td>
<td>6.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10.4</td>
</tr>
</tbody>
</table>

Table 5: GMVP forecasting results. Bold figures indicate the best models.
Figure 1: Panels (a) and (c): Histograms of the realized variances for the Citigroup and Caterpillar stocks and their unconditional distribution predicted under the fitted WFSS model (9F-HAR-v_β) defined by Eqs. (6)-(12) for 60 assets using 9 factors (red solid line). The predicted distributions are obtained from kernel density estimates based on simulated data from the fitted WFSS model for T = 50,000 time periods. Panels (b) and (d): Section of the histograms in panels (a) and (c) for the ranges [10, 30] and [20, 50], respectively.
Figure 2: Panels (a)-(h): Time series plots of the realized variances; The gray shaded areas mark the two out-of-sample windows used in the forecasting experiments in Section 4.3. Panels (h) and (i): Sample ACF of the realized variance (blue line) and predicted ACF under the fitted 9F-HAR-$\gamma \beta$ WFSS model (red line). Predicted ACF is obtained from simulated data from the fitted WFSS model for $T = 50,000$ time periods.
Figure 3: Heat plots of the time-average of daily realized (residual) correlation matrices for 60 assets in the S&P500. Panel (a): unconditional without factors; Panel (c): conditional on the market factor; Panel (b): conditional on the 3 Fama-French factors; Panel (d): conditional on the 3 Fama-French factors and 6 sector-specific ETF’s. The black squares along the diagonal indicate sector blocks. The sector labels are: (I) Industrials; (D) Consumer Discretionary; (S) Consumer Staples; (H) Health Care; (F) Financials; (T) Information Technologies.
Figure 4: Box-plots of the posterior mean values for the autoregressive coefficients (panel a), maximal HAR roots in modulus (panel b) and stationary variances (panel c) for the AR(1) and HAR processes of the 9F-HAR-\(\beta\)WFSS model directing the integrated (co)variances in \((\Sigma_t^f, \Sigma_t^e)\) and factor loadings in \(B_t\) (see Eqs. 7-12). The sector labels are: (I) Industrials; (D) Consumer Discretionary; (S) Consumer Staples; (H) Health Care; (F) Financials; (T) Information Technologies.
Figure 5: Time series plots of the smoothed estimates for the total integrated variance $\sigma_t^r$, integrated residual variance $\sigma_t^e$, and integrated loadings $\beta_t$ on the market-, HML-, SMB- and financial sector specific risk factor for the Citigroup stock under the 9-factor WFSS-tv$\beta$-HAR model (red line), together with their respective observed realized counterparts (blue line).
Figure 6: Bar plots of the fractions of realized stock return variation of the 60 assets explained by the risk factors, computed as $1 - \frac{1}{T} \sum_{t=1}^{T} E(\sigma_{it}^2 | C_{1:T})/\sigma_{it}^2$, where $E(\sigma_{it}^2 | C_{1:T})$ is the posterior mean of the integrated residual obtained under the fitted HAR-v$\beta$ WFSS model with the 1 (market), 3 (market+HML+SMB) and 9 factors (market+HML+SMB+ETFs), respectively. White bar: Explained variation by the market; White and red bar: Explained variation by the market+HML+SMB; white, red and blue bar: explained variation by the market+HML+SMB+ETFs. The sector labels are: (I) Industrials; (D) Consumer Discretionary; (S) Consumer Staples; (H) Health Care; (F) Financials; (T) Information Technologies.