Journal of Banking and Finance 138 (2022) 106404

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

Modeling and forecasting realized portfolio weights

Vasyl Golosnoy[a,*], Bastian Gribisch[b]

[a] Faculty of Management and Economics, Ruhr-Universität Bochum, Universitätstraße 150, Bochum 44780, Germany
[b] Institute of Econometrics and Statistics, Universität zu Köln, Germany

ARTICLE INFO

Article history:
Received 25 August 2020
Accepted 3 January 2022
Available online 6 January 2022

JEL classification:
C32
C58
G11
G17

Keywords:
M-estimation
Lasso
Realized covariances
Realized GMVP
VARFIMA

ABSTRACT

We propose direct multiple time series models for predicting high dimensional vectors of observable realized global minimum variance portfolio (GMVP) weights computed based on high-frequency intraday returns. We apply Lasso regression techniques, develop a class of multiple AR(FI)MA models for realized GMVP weights, suggest suitable model restrictions, propose M-type estimators and derive the statistical properties of these estimators. In the empirical analysis for portfolios of 225 stocks from the S&P 500 we find that our direct models effectively minimize either statistical or economic forecasting losses both in- and out-of-sample as compared to relevant alternative approaches.

1. Introduction

Portfolio selection is of importance in both empirical finance and financial econometrics. In this paper we consider a risk averse mean-variance investor with a desire to hold a portfolio of risky assets where the number of assets k is finite but sufficiently large in order to achieve broad diversification. For computing the optimal portfolio composition our investor needs to estimate and to predict both the mean vector and the covariance matrix of the asset returns. However, since there are well-documented difficulties to obtain reliable estimators and predictors of expected returns (see Merton, 1980), using mean estimates for calculating optimal portfolio weights could lead to unreasonable portfolio compositions (cf. Best and Grauer, 1991; Jagannathan and Ma, 2003).

The weights of the global minimum variance portfolio (GMVP) depend solely on the covariance matrix, which can be precisely estimated and is fairly predictable on at least short and medium investment horizons (see e.g. Bauwens et al., 2006). The GMVP is of considerable importance in capital market theory as it is the starting point of the Markowitz mean-variance efficient frontier. Moreover, all efficient portfolios can be written as a linear combination of the GMVP weights and weights for any other efficient portfolio.

In this paper we exploit intraday high-frequency returns to compute precise and consistent ex-post estimates of the daily GMVP compositions which we label *realized GMVP weights*. Realized GMVP weights are non-linear functions of realized covariance matrices. While Golosnoy et al. (2020) consider sequential on-line monitoring of realized GMVP weights and Golosnoy et al. (2019) propose naive (ad-hoc) exponential smoothing for forecasting the realized GVMP compositions, in this paper we develop a rather general class of linear vector autoregressive (VAR) specifications, which is applicable for a direct modeling and forecasting of the realized GMVP composition for several hundred risky assets. In order to cope with the curse of dimensionality, we exploit the Least Absolute Shrinkage and Selection Operator (Lasso)-approach of Tibshirani (1996) for analyzing the dynamic structure of the realized GMVP weight vectors and selecting sparse high dimensional VAR-specifications. Motivated by the obtained Lasso estimates we follow the objective to achieve further dimensionality reduction and consider both scalar and diagonal versions of the vector autoregressive fractionally integrated moving average (VARFIMA) model, the VARMA(1,1) approach and the multiple heterogenous autoregressive (MHAR) model (cf. Bubak et al., 2011). For estimating these direct models we elaborate M-type estimators specifically designed to minimize the economic loss criterium related to the realized portfolio variance. We derive the stochastic properties of these estimators and provide the corresponding asymptotic distributions.

* Corresponding author.
E-mail addresses: vasyl.golosnoy@rub.de (V. Golosnoy), bastian.gribisch@statistik.uni-koeln.de (B. Gribisch).

https://doi.org/10.1016/j.jbankfin.2022.106404
0378-4266/

Our direct modeling approach is related to Brandt et al. (2009) and Reh et al. (2019), who model portfolio weights as functions of exogenous regressors such as observed firm characteristics or past daily asset returns. However, in contrast to the former studies our realized GMVP framework incorporates high-frequency return information and therefore allows for a more precise measurement of the highly predictable GMVP dynamics (cf. Chiriac and Voev, 2011). Moreover, as opposed to Reh et al. (2019) our time series models are linear. This allows for a straightforward application of flexible modeling devices like Lasso- or ARMA-type dynamics, which prove to be suitable for modeling realized GMVP weights.

Our direct GMVP approach is an alternative to the common (indirect) way of predicting GMVP weights based on covariance matrix forecasts obtained from multivariate volatility models. The latter require dynamic multiple (typically non-linear) time series modeling for daily realized covariance matrices. The GMVP forecast is then obtained by plugging the resulting covariance prediction into the GMVP formula, see e.g. Callot et al. (2017); Chiriac and Voev (2011); DeNard et al. (2021); Golosnoy et al. (2012); Gribisch et al. (2020); Jin and Maheu (2013); Noureldin et al. (2012); Hautsch et al. (2015); Hafner et al. (2020). The class of *indirect* approaches for predicting the GMVP compositions, however, has the following disadvantages: first, it is subject to the curse of dimensionality as the total number of (co)variance series grows quadratically in the number of assets: a realized covariance matrix for k assets comprises $k(k+1)/2$ distinct variances and covariances to be modeled. Accordingly, time series models for realized covariance matrices are either heavily parameterized or require strong parametric restrictions. In contrast, the vector of realized GMVP weights contains only k elements to model and is, therefore, much less prone do the dimensionality issue. Further, it turns out to be quite challenging to model highly persistent covariance matrix processes with complex autoregressive dynamics, and, at the same time, to guarantee that realized covariance matrix forecasts are positive definite. For this reason indirect approaches are particularly vulnerable to model misspecifications when not all of the relevant properties of the data generating process are taken into account: a multivariate volatility model might e.g. capture relevant parts of the (co)variance dynamics but at the same time might miss some important aspects of the dynamics in the GMVP weights. Moreover, indirect GMVP forecasts are not necessarily optimal as they are non-linear functions of covariance matrix forecasts.

Our direct realized GMVP approach contributes to the existing literature as follows: we propose a class of linear multiple time series approaches for modeling realized portfolio weights which employs high-frequency asset returns in order to efficiently exploit intraday information to estimate the GMVP compositions. In order to alleviate the curse of dimensionality we explicitly model the observed time series dynamics of the realized weight vectors by sparse high dimensional time series specifications estimated by minimizing either statistical or economic GMVP loss criteria. Our models are linear by nature, which allows for a straightforward implementation of well approved multivariate time series specifications like ARFIMA/HAR models or Lasso model selection procedures. We do not have to care about non-linear restrictions in order to guarantee positive definiteness of covariance matrix forecasts, our model specifications are tailored to fit the dynamics of the GMVP proportions, and the direct modeling of conditional expectations of the GMVP weights immediately provides mean squared optimal forecasts for the GMVP composition. Our direct approaches are finally easy to estimate. Although we focus on the GMVP, our methodology could also be applied to other portfolio rules with observable weights, e.g. those with shrinkage (cf. DeMiguel et al., 2009; Bollerslev et al., 2019) or those resulting from a combination of various portfolio recommendations (cf. Tu and Zhou, 2011). We apply our dynamic realized GMVP approach to time series of daily realized GMVP compositions for 225 NYSE traded stocks and find that it performs very well compared to a wide range of benchmarks from the literature. We furthermore show that combining our direct approaches with indirect ones as in Golosnoy et al. (2019) can lead to further improvements in the GMVP prediction.

The rest of the paper is organized as follows. In Section 2 we discuss the consistent realized estimation of daily covariance matrices and the corresponding GMVP weights based on intraday asset returns. In Section 3 we present our direct multiple time series approach for realized GMVP weights, and discuss both restricted and unrestricted model specifications. In Section 4 we introduce M-type estimators for our direct models and derive the statistical properties of the obtained estimators. The empirical study is provided in Section 5, while Section 6 concludes the paper. The proofs are placed in the Appendix.

2. GMVP weights and their realized estimators

Let $p(\tau)$ denote a k-dimensional vector of log-prices following a Brownian stochastic semimartingale with $k \times k$-dimensional spot covariance matrix $\Theta(\tau)$, where $\tau \in \mathbb{R}_+$ represents continuous time. We treat the number of assets k as fixed but sufficiently large for achieving diversification benefits.

We set the trading day's length to the unit interval and define the integrated covariance matrix for day t as $\Sigma_t = \int_{t-1}^{t} \Theta(\tau)\, d\tau$, which is assumed to be positive definite for all t. The daily vector of asset returns is given by $p(t) - p(t-1) = x_t$ such that $x_t \sim (\mu_t, \Sigma_t)$.

A risk averse Markowitz-type investor should choose a k-dimensional vector of portfolio weights w^* with portfolio return $x_t' w^*$ by maximizing for given (μ_t, Σ_t) the mean-variance objective function

$$\max_{w^*} \left[E[x_t' w^*] - \frac{\alpha}{2} Var[x_t' w^*] \;=\; \mu_t' w^* - \frac{\alpha}{2} w^{*\prime} \Sigma_t w^* \right] \qquad \text{w.r.t.} \quad \iota_k' w^* = 1, \tag{1}$$

where $\alpha > 0$ is the degree of the investor's risk attitude and ι_k is a k-dimensional vector of ones. The solution of the maximization task in (1) is given by (cf. Okhrin and Schmid, 2006)

$$\mathfrak{w}_t^* = \frac{\Sigma_t^{-1} \iota_k}{\iota_k' \Sigma_t^{-1} \iota_k} + \frac{1}{\alpha} \left(\Sigma_t^{-1} - \frac{\Sigma_t^{-1} \iota_k \iota_k' \Sigma_t^{-1}}{\iota_k' \Sigma_t^{-1} \iota_k} \right) \mu_t. \tag{2}$$

Since from the forecaster's point of view the optimal portfolio composition in (2) has to be selected given the information set available up to period $t-1$, the investor requires proper predictors of μ_t and Σ_t. However, the vector of expected returns μ_t is rather difficult to estimate and to predict at least in the short run (Merton, 1980). For these reasons using imprecise estimators of μ_t in (2) often leads to unreasonable recommendations for portfolio weights (cf. Best and Grauer, 1991; Welch and Goyal, 2008). Hence, the expression for the optimal weights in (2) is not very useful in empirical portfolio selection due to problems related to the measurement of the mean vector μ_t.

In contrast to the mean vector, the covariance matrix Σ_t can be precisely estimated and is fairly predictable (cf. Bauwens et al., 2006). The GMVP composition then obtains by setting the vector μ_t in Eq. (2) to zero, resulting in the GMVP weight vector

$$\omega_t^* = \Sigma_t^{-1} \iota_k \;/\; \iota_k' \Sigma_t^{-1} \iota_k, \tag{3}$$

which can be seen as a reasonable alternative to $\mathfrak{w}_t^*$ in practical portfolio decisions of investors with different risk attitudes. Since the GMVP is the starting point of the mean-variance efficient frontier, its composition determines the frontier's location and, hence, the GMVP is of importance for all mean-variance investors. We note that some of the weights in the vector ω_t^* can become negative as short positions are allowed for the realized GMVP. In the context of realized GMVP forecasting Golosnoy et al. (2019) find that no-short-selling constraints worsen the prediction of GMVP weights compared to the unrestricted settings.

The GMVP vector ω_t^* depends on the integrated covariance matrix Σ_t which is not directly observable and has to be estimated from the data. In this paper we consider realized estimators of the daily GMVP weights based on intraday asset return information. In order to construct a realized estimator, assume that we observe m equally spaced intraday vectors of log-prices $p(\cdot)$ per trading day. Then the jth intraday return vector at day t is given by

$$x_{t,j} = p\big(t-1+j/m\big) - p\big(t-1+(j-1)/m\big), \qquad j = 1,\dots,m.$$

These high-frequency (intraday) returns serve for the construction of the realized covariance matrix R_t, which is a precise non-parametric ex-post estimate of Σ_t. The simplest form of the realized covariance matrix is given by

$$R_t = \sum_{j=1}^{m} x_{t,j}x'_{t,j}. \tag{4}$$

Under rather general assumptions on the stochastic process of the log-prices $p(\tau)$, Barndorff-Nielsen and Shephard (2004, p. 892) show that (i) the matrix $R_t = \sum_{j=1}^{m} x_{t,j}x'_{t,j}$ is a consistent estimator of Σ_t for the number of intraday returns $m \to \infty$; (ii) conditional on the day t path of the spot covariance $\Theta(\cdot)$ it holds for the elements of this realized covariance matrix:

$$m^{1/2}\Big(vech(R_t) - vech(\Sigma_t)\Big) \xrightarrow{L} \mathcal{N}(0,\Pi_t) \qquad \text{for} \qquad m \to \infty, \tag{5}$$

where Π_t is a stochastic matrix depending on $\Theta(\cdot)$. Integrating over $\Theta(\cdot)$ then leads to an asymptotic mixed normal law for $vech(R_t)$ conditional on $vech(\Sigma_t)$. For recent theoretical results on light-tailed mixture distributions see e.g. Klüppelberg and Seifert (2019, 2020).

Further refinements of realized covariance measures as e.g. the multivariate realized kernel estimator of Barndorff-Nielsen et al. (2011), the composite realized kernel estimator of Lunde et al. (2016) or the CholCov estimator of Boudt et al. (2017) are proposed for a robust measurement of Σ_t in the presence of asynchronous intraday returns and market microstructure noise.

For a given realized covariance matrix R_t the realized GMVP weight vector is computed as

$$w_t^* = R_t^{-1}\iota_k \,/\, \iota_k' R_t^{-1}\iota_k. \tag{6}$$

This realized vector w_t^* is observable at the end of the investment period, in our case at the end of the trading day t. As the elements of w_t^* sum up to unity due to $\iota_k' w_t^* = 1$, there exists a non-degenerate distribution for any $\ell = (k-1)$-dimensional subvector of w_t^*. Hence, in the following we model ℓ-dimensional GMVP vectors w_t which are obtained by skipping the last entry of w_t^*.

Based on the limit result in (5) we provide the asymptotic distribution of the realized GMVP weights in the next proposition which generalizes the result of Theorem 1 in Golosnoy et al. (2020).

Proposition 1. *Let R_t denote the consistent realized estimator of the $k\times k$-dim. covariance matrix Σ_t defined in Eq. (4). Then the k-dim. realized GMVP vector $w_t^* = R_t^{-1}\iota_k/\iota_k' R_t^{-1}\iota_k$ is a consistent estimator of the GMVP vector ω_t^*. Next, given that $m^{1/2}(vech(R_t) - vech(\Sigma_t)) \xrightarrow{L} \mathcal{N}(0,\Pi_t)$ for the number of intraday returns $m\to\infty$ the asymptotic distribution of the ℓ-dimensional realized GMVP vector w_t conditional on the day t path of $\Theta(\cdot)$ is given as*

$$m^{1/2}(w_t - \omega_t) \xrightarrow{L} \mathcal{N}(0,\Omega_t), \tag{7}$$

where Ω_t is the first $\ell\times\ell$ submatrix of the $k\times k$-dimensional positive semi-definite matrix Ω_t^*:

$$\Omega_t^* = G_t' \,\Pi_t\, G_t, \qquad G_t = \partial\omega_t^{*'}/\partial vech(\Sigma_t),$$
$$G_t = \frac{1}{2}\Big[\Sigma_t^{-1}\iota_k\ (\iota_k'\otimes\iota_k') - (\iota_k'\Sigma_t^{-1}\iota_k)\ (\iota_k'\otimes I_k)\Big]\ [I_{k^2} + K_k]$$
$$(\Sigma_t^{-1}\otimes\Sigma_t^{-1})\ D_k/(\iota_k'\Sigma_t^{-1}\iota_k)^2.$$

I_{k^2} denotes the identity matrix of dimension $k^2\times k^2$, D_k the duplication and K_k the commutation matrix (Lütkepohl, 1996, p. 9). Since the matrix Ω_t is unconditionally stochastic as it depends on Π_t, Eq. (7) gives a mixed normal asymptotic distribution for w_t conditional on ω_t.

While the result in Proposition 1 refers to the basic realized covariance measure R_t in (4), there exist similar mixed normal asymptotics for refined realized covariance measures such as the multivariate realized kernel of Barndorff-Nielsen et al. (2011) or the CholCov-estimator of Boudt et al. (2017), though typically with slower speed of convergence due to the underlying robustness properties of these estimators. The result in Proposition 1 should then be adjusted by replacing the speed of convergence and the conditional covariance Π_t according to the respective estimator.

For vast-dimensional applications with thousands of assets regularized realized covariance matrices have recently been proposed by Fan et al. (2016); Lam and Feng (2018), and Bollerslev et al. (2019). Under specific regularity conditions these estimators are consistent for $m\to\infty$ and $k\to\infty$ jointly, but distributional limit results are typically not available. In this paper we do not focus explicitly on such vast-dimensional scenarios, although our time series methodology stays applicable for $k\to\infty$ as well, as long as the realized GMVP weight vector is computed using properly regularized realized covariance measures.

To summarize, equipped with intraday asset return information, the investor can interpret the realized covariance matrix R_t and the corresponding realized GMVP weight vector w_t as precise observable daily measures of Σ_t and ω_t, respectively.

3. Direct modeling of realized GMVP weights

Our aim is to investigate how historically observed realized GMVP weights could help for a proper prediction of the GMVP weights for the next trading day. For this purpose we analyze *direct* time series specifications for the realized GMVP vector w_t given the information set $\mathcal{F}_{t-1} = \{w_{t-1}, w_{t-2},\dots\}$ as an alternative to the commonly applied *indirect* forecasting of GMVP weights by plugging in realized covariance predictors from multivariate covariance models into the GMVP formula, as e.g. in Bauwens et al. (2017); Bollerslev et al. (2018); Callot et al. (2017); Chiriac and Voev (2011). The direct modeling of realized GMVP weights alleviates the curse of dimensionality in covariance modeling and forecasting and allows to apply reduced-form multiple time series models specifically designed to reflect their autoregressive dynamics.

A further advantage of the direct modeling of daily realized GMVP weights compared to realized covariances is the empirical finding that the former exhibit much less heteroskedasticity and autocorrelation, see Fig. 1, which makes it easier to find an appropriate dynamic specification. This feature can be explained within the class of two-component volatility models with a persistent scalar long-term component common to multiple (co)variance series (cf. Engle and Rangel, 2008; Belke and Beckmann, 2015). Then, following e.g. the approach of Hafner and Linton (2010), one could write the daily covariance matrix as a product of a scalar – possibly non-stationary – secular component ζ_t and the remaining short-term covariance matrix $\dot\Sigma_t$ such that $\Sigma_t = \zeta_t\cdot\dot\Sigma_t$, where $\dot\Sigma_t$ follows an ergodic and weakly stationary process with substantially reduced persistence as compared to Σ_t. The secular component ζ_t

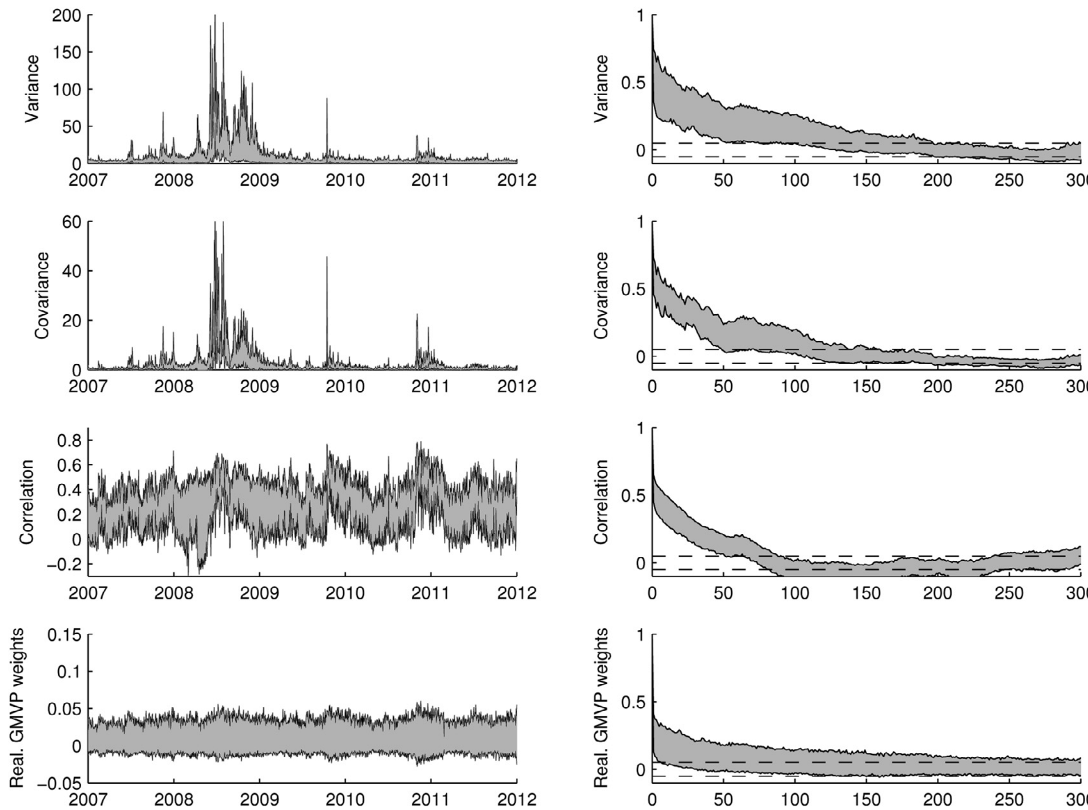

Fig. 1. Envelop plots of realized variances, covariances, correlations and GMVP weights and their sample ACFs: 5%- and 95%-sample quantiles of 225 variance, 25,200 covariance/correlation and 225 GMVP weight time series (left) and of the sample ACFs for 300 lags (right). Dashed lines denote 95% Bartlett confidence bounds for no autocorrelation.

can be interpreted as the variance of a common market factor motivated theoretically by the CAPM.

Empirically, the common long-run pattern in the realized (co)variance series is clearly observed in the data, see Fig. 1. Then by considering the GMVP vector ω^*, the long term scalar component ζ_t shortens:

$$\omega_t^* = \Sigma_t^{-1}\iota_k \,/\, \iota_k'\Sigma_t^{-1}\iota_k = \dot{\Sigma}_t^{-1}\iota_k \,/\, \iota_k'\dot{\Sigma}_t^{-1}\iota_k, \tag{8}$$

such that the impact of the persistent common component vanishes, leaving only the dynamics of $\dot{\Sigma}_t$. The same argument also holds for the realized GMVP weights w_t making them a rather convenient object for multiple time series analysis.

3.1. Multiple time series models for realized GMVP weights

We now introduce direct multiple time series specifications for the realized GMVP weights w_t. Note that although we focus on the realized GMVP, our analysis is applicable to observable weights from other portfolio rules or recommendations as well, e.g. for portfolios based on robust (shrinkage) estimation procedures (cf. Frahm and Memmel, 2010), or on methods for combining various portfolio rules (Tu and Zhou, 2011).

We start from a rather general ℓ-dimensional vector autoregressive (VAR) process of order p:

$$w_t = \upsilon_t + u_t, \qquad \text{with} \qquad u_t \overset{iid}{\sim} (0, \Sigma_u), \tag{9}$$

$$\upsilon_t = E[w_t|\mathcal{F}_{t-1}] = \gamma + \sum_{i=1}^{p} B_p w_{t-p}, \tag{10}$$

where the covariance matrix Σ_u is assumed to be positive definite. The structural model parameters in (10) are the ℓ-dimensional vector $\gamma = (\gamma_1, \ldots, \gamma_\ell)'$ and the $\ell \times \ell$-dimensional matrices B_i, $i = 1, \ldots, p$. This VAR process is stable and weakly stationary if and only if $\det\big(I_\ell - B_1 z - \ldots - B_p z^p\big) \neq 0$ for all $|z| \leq 1$ (Lütkepohl, 2005, p. 16).

The VAR(p) process of Eqs. (9) and (10) allows for rich dynamics of the GMVP compositions as for large k every entry of the weight vector can potentially depend on thousands of variables, including own lags and lagged weights of other assets. This may lead to equations with much more parameters than observations, rendering standard estimation techniques like OLS as imprecise or even infeasible. Furthermore, the autoregressive VAR parameter matrices may contain many insignificant coefficients, making a standard OLS-based VAR approach prone to overfitting, which may worsen the out-of-sample forecasting performance. For these reasons we need sensible model restrictions on the general VAR(p) dynamics in (9).

The Lasso technique of Tibshirani (1996) is a popular data-driven approach suitable to unveil potentially complex dependence structures and to introduce the desired degree of sparsity to a VAR system. For recent applications of the Lasso to high dimensional covariance modeling and indirect GMVP forecasting see also Kock and Callot (2015) and Callot et al. (2017).

A further advantage of the Lasso is that it is feasible even when the number of parameters to be estimated is larger than the sample size T. This would easily be the case if the number of assets k is of similar magnitude (or even higher) as the sample size T, i.e. in vast-dimensional portfolios with thousands of assets (cf. Hautsch and Voigt, 2019).

3.2. The Lasso procedure

We estimate the VAR process by applying the Lasso procedure of Tibshirani (1996) separately to each equation of the VAR system from Eqs. (9) and (10). Let $Z_t = (1, w'_{t-1}, \dots, w'_{t-p})'$ denote the $(\ell p + 1)$ dimensional vector of explanatory variables and $Z = (Z_T, \dots, Z_1)'$ the $T \times (\ell p + 1)$ matrix of covariates. Next, denote the time series of realized GMVP weights of the ith asset by $w_i = (w_{T,i}, \dots, w_{1,i})'$ and write the corresponding Lasso objective for the ith VAR equation as the penalized OLS

$$\mathcal{L}_P(\theta_i) = (1/T) \cdot ||w_i - Z\theta_i||^2 + 2\lambda_i \sum_{j=1}^{\ell p} |\beta_{i,j}|, \tag{11}$$

where $\theta_i = (\gamma_i, \beta_i')'$ denotes the $(\ell p + 1)$-dimensional vector summarizing the coefficients of the ith VAR equation in (9), the vector $\beta_i = (\beta_{i,1}, \dots, \beta_{i,\ell p})'$ comprises the according slope coefficients, and $|| \cdot ||^2$ denotes the Euclidean norm.

The estimates $\hat{\theta}_i$ are obtained via minimizing the objective (11) over θ_i for a sequence of penalty parameters λ_i. The final value of λ_i is then chosen according to the Bayesian Information Criterion (BIC), see e.g. Kock and Callot (2015), and Callot et al. (2017). Note that the Lasso allows for fast estimation and variable selection in a single step. For further details on the Lasso estimation procedure we refer to the studies of Zhao and Yu (2006); Bickel et al. (2009); Kock and Callot (2015). Although the Lasso procedure selects the correct model only asymptotically under quite restrictive conditions on the covariates, it serves as an effective screening device in order to uncover the most relevant dependence structures in the data.

As e.g. demonstrated by Belloni and Chernozhukov (2013) the performance of the Lasso may be further improved via Post-Lasso OLS estimation: after identifying by Lasso the nonzero VAR coefficients, the VAR model with only the nonzero coefficients included is estimated via OLS in order to obtain the final (restricted) model specification. The Post-Lasso approach is known to offer advantages compared to the conventional Lasso due to improved rates of convergence and smaller bias (cf. Belloni and Chernozhukov, 2013). In our empirical application in Section 5 we consider Post-Lasso OLS estimation as a competitor to the standard Lasso approach in Eq. (11).

3.3. VARFIMA and MHAR specifications

Our empirical results on the Lasso estimates reported in Section 5 indicate persistent dynamics for the series of realized GMVP weights, which are mostly captured by the first twenty lags of the realized GMVP series. Moreover, many of the weight series show signs of long-memory indicated by a hyperbolic-like decay of the corresponding sample autocorrelation functions (acf), see Fig. 1. Motivated by these empirical results and since the estimation of unrestricted VAR(FI)MA models is infeasible in high-dimensional settings we now propose diagonal and scalar VARFIMA models as sparse alternatives to the Lasso. The VARFIMA(p, d, q) process for the realized weight vector is defined as (cf. Chiriac and Voev, 2011)

$$\Phi(L)D(L)(w_t - \gamma) = \Xi(L)u_t, \qquad u_t \overset{iid}{\sim} (0, \Sigma_u), \tag{12}$$

where γ is a vector of constants, $\Phi(L) = I_\ell - \Phi_1 L - \ldots - \Phi_p L^p$, and $\Xi(L) = I_\ell - \Xi_1 L - \ldots - \Xi_q L^q$ are matrix lag polynomials and $D(L) = diag\{(1-L)^{d_1}, \dots, (1-L)^{d_\ell}\}$, where $d_1, \dots, d_\ell$ are the degrees of fractional integration of each of the ℓ elements of the vector w_t. If $d_i < 1/2$ for all i-s and the roots of $\Phi(L)$ and $\Theta(L)$ are outside the unit circle, the ARFIMA process in (12) is weakly stationary with an infinite order AR representation. The AR and MA matrices Φ_i and Ξ_i are assumed to be diagonal in order to preserve a sparse parametrization. The workhorse in our empirical study is the VARFIMA($1, d, 1$) model. The scalar model version restricts the diagonal AR and MA coefficients as well as the degrees of fractional integration $\{d_i\}_{i=1}^{\ell}$ to be identical over all ℓ elements of w_t, a specification which turns out to be supported by our data in Section 5. As nested special cases of the VARFIMA($1, d, 1$) approach we also consider diagonal and scalar VAR(1) and VARMA(1,1) models. The heterogeneous autoregressive (HAR) process introduced by Corsi (2009) is well known to provide a flexible approximation to ARFIMA dynamics via cascade-like regression structures involving up to 20 lags of the observable data. HAR processes are sparse but effective modeling devices which are easy to estimate and typically have good out-of-sample forecasting properties for long-memory processes. We therefore enlarge our set of time series models for realized GMVP weights by diagonal and scalar multiple HAR (MHAR) specifications:

$$w_t = \gamma + B_1 w_{t-1} + B_2 w_{t-1}^{(w)} + B_3 w_{t-1}^{(m)} + u_t, \tag{13}$$

where $w_{t-1}^{(w)} = (1/5) \cdot \sum_{\tau=t-5}^{t-1} w_\tau$ and $w_{t-1}^{(m)} = (1/20) \cdot \sum_{\tau=t-20}^{t-1} w_\tau$ are weekly and monthly averages, respectively. Diagonal and scalar versions of the MHAR model are then obtained by restricting the autoregressive coefficient matrices in (13) to be either diagonal or scalars.

VARFIMA and HAR models are typically estimated by ML or (non)linear LS techniques. Regarding the ML methodology we note that exact finite sample results for the conditional distribution of w_t given $\mathcal{F}_{t-1}$ strongly depend on the distributional assumption for the realized covariance measure R_t. For example, Golosnoy et al. (2012) assume a Wishart distribution for R_t conditional on Σ_t, i.e. $R_t \sim \mathcal{W}_k(\nu, \Sigma_t/\nu)$, where ν is the degree of freedom parameter and Σ_t is assumed to be measurable given $\mathcal{F}_{t-1}$. Golosnoy et al. (2020) then show that under this Wishart assumption and for $\nu \geq k-1 \geq 1$ the vector of GMVP weights follows a $k-1$-variate elliptical t-distribution, which could be used in order to compute the likelihood function (cf. Muirhead, 1982, p. 48).

Although the Wishart assumption is convenient for low-dimensional realized covariance matrices (cf. Gorgi et al., 2019), it appears quite restrictive particularly for high dimensional specifications (cf. Jin and Maheu, 2013; Gribisch and Hartkopf, 2021). An alternative way to obtain an ML-estimator is to rely on asymptotic results. In Proposition 1 we show that for the basic realized covariance measure in Eq. (4) the vector w_t follows asymptotically an ℓ-dimensional mixed normal distribution. Unfortunately, expressions for the asymptotic covariance matrix of this distribution are not straightforward and depend on unknown quantities, see Barndorff-Nielsen et al. (2011). Furthermore, the corresponding limit theorems for the realized covariance matrix are not readily available for some recently proposed refinements of realized covariance estimators, such as e.g. the robust composite realized kernel estimator of Lunde et al. (2016) or the regularized realized covariance measure of Fan et al. (2016) with asymptotics involving both $m \to \infty$ and $k \to \infty$.

As a consequence, we refrain from applying the ML methodology and concentrate on non-parametric (non)linear least squares estimators for the VARFIMA and HAR models, which directly minimize the MSE of the in-sample GMVP predictions

$$L_S(\theta) = (1/T) \cdot \sum_{t=1}^{T} (w_t - \nu_t(\theta))'(w_t - \nu_t(\theta)), \tag{14}$$

where the vector θ comprises the model parameters of the AR(FI)MA or MHAR processes.

LS estimators for the ARFIMA parameters are obtained by truncating the corresponding infinite order AR representation at 1000 lags and initializing the lag structure by the unconditional mean of the GMVP weights at the parameter values, see

Janus et al. (2014) and Opschoor and Lucas (2019) for similar implementations. The LS objective function is minimized by numerical optimization techniques whereas the asymptotic standard errors are computed using the robust QML approach discussed at the end of Section 4.

As an alternative to the LS estimation of diagonal and scalar HAR model specifications with fixed cascade-like regression structure as in (13) we finally consider the data-driven Lasso-HAR approach of Audrino et al. (2019) and Wilms et al. (2021) which starts with the general cascade regression

$$w_t = \gamma + \sum_{i=1}^{p} \beta_i \sum_{j=1}^{i} w_{t-j} + u_t. \tag{15}$$

We then estimate the model parameters in the $\ell \times \ell$ coefficient matrices β_i equation-wise by the Lasso approach with $p = 100$.

4. M-type estimators

The LS estimators discussed above are obtained by minimizing the statistical loss criterion in Eq. (14) over the parameter vector θ indexing the underlying time series model. From the perspective of the GMVP investor with an objective related to the realized portfolio variance $(w_t^*)' R_t w_t^*$ it would also be of interest to obtain the parameter vector θ from the task

$$\min_{\theta} \left[\mathcal{M}(\theta) = \sum_{t=1}^{T} (\upsilon_t^*)' R_t \upsilon_t^* = \upsilon^*(\theta)' R \upsilon^*(\theta) \right], \tag{16}$$

with the $kT \times kT$-dimensional matrix $R = \text{diag}(R_1, \ldots, R_T)$. The vector υ_t^* is the model-based forecast of w_t^*, obtained as $\upsilon_t^* = e_k + \tilde{I}_{k\times\ell} \upsilon_t$ with $\tilde{I}_{k\times\ell} = [I_\ell, -\iota_\ell]'$ and the k-dimensional vector e_k with 1 on the kth position and 0 otherwise. The criterion in (16) corresponds to an economic loss function as opposed to the statistical loss function in (14).

The optimization problem in (16) is equivalent to

$$\min_{\theta} \left[\mathcal{M}^*(\theta) = \sum_{t=1}^{T} (w_t^* - \upsilon_t^*)' R_t (w_t^* - \upsilon_t^*) = (w^* - \upsilon^*)' R (w^* - \upsilon^*) = (u^*)' R u^* \right], \tag{17}$$

which resembles a GLS-type objective with an endogenous weight matrix R. We denote the M-estimator obtained by minimizing the economic loss (17) by $\hat{\theta}_{\mathcal{M}}$.

For a linear VAR(p) process as in Eqs. (9) and (10) we obtain $\hat{\theta}_{\mathcal{M}}$ in closed form. For notational convenience we summarize the VAR parameters to the matrix $B = (\gamma, B_1, \ldots, B_p)$ of dimension $\ell \times (\ell p + 1)$ and write the $(\ell^2 p + \ell)$-dimensional vector $\beta = \text{vec}(B)$. Next, we define the matrix $W = (w_1, \ldots, w_T)$ of dimension $(\ell \times T)$, and the vector $Z_t = (1, w_t', \ldots, w_{t-p+1}')'$ of dimension $(\ell p + 1)$ such that the matrix $Z = (Z_0, \ldots, Z_{T-1})$ is of dimension $(\ell p + 1) \times T$, and $U = (u_1, \ldots, u_T)$ of dimension $\ell \times T$. Based on the representation $W = BZ + U$ (Lütkepohl, 2005, p. 70) we obtain

$$\begin{aligned} w := vec(W) &= vec(BZ) + vec(U) = (Z' \otimes I_\ell) vec(B) + vec(U) \\ &= (Z' \otimes I_\ell)\beta + u. \end{aligned} \tag{18}$$

We provide the M-estimator $\hat{\beta}_{\mathcal{M}}$ for the VAR(p) process in (18) in the following proposition.

Proposition 2. *For the VAR(p) in (18) the M-estimator of β minimizing the loss (17) is*

$$\hat{\beta}_{\mathcal{M}} = [\tilde{Z}\tilde{Z}']^{-1} \tilde{Z} R^{1/2} (w^* - \iota_T \otimes e_k) = \beta + [\tilde{Z}\tilde{Z}']^{-1} \tilde{Z} R^{1/2} (I_T \otimes \tilde{I}_{k\times\ell}) u,$$

$$\tilde{Z} = (Z \otimes \tilde{I}_{k\times\ell}') R^{1/2},$$

where $\tilde{I}_{k\times\ell} = [I_\ell, -\iota_\ell]'$, $e_k = (0, \ldots, 0, 1)'$, and $R^{1/2}$ is the square root matrix with $R = R^{1/2} R^{1/2}$.

As we are interested in sparse estimation of VAR(p) models, we also consider the possibility to apply the Lasso principle for the M-estimation procedure.

Remark 1. Rewriting the task in (17) using the VAR(p) structure in (18) gives

$$\begin{aligned} \mathcal{M}^*(\theta) &= (\tilde{w} - \tilde{Z}'\beta)'(\tilde{w} - \tilde{Z}'\beta) = \|\tilde{w} - \tilde{Z}'\beta\|^2, \\ &\text{with} \quad \tilde{w} = (R^{1/2})'(I_T \otimes \tilde{I}_{k\times\ell}) w. \end{aligned} \tag{19}$$

The objective in (19) combined with the Lasso penalty for the slope parameters gives the objective function for Lasso M-estimation of the VAR(p) process.

The result in Remark 1 allows to implement the Lasso for one-step estimation of the complete VAR(p) model including the whole set of $\ell^2 p$ autoregressive coefficients, as opposed to the equation-wise Lasso in (11). However, while one-step Lasso is certainly possible for low-dimensional applications, the computation of the Lasso objective in (19) turns out to be numerically prohibitive in high dimensional environments due to the vast number of model parameters as compared to equation-wise Lasso estimation. We therefore do not apply Lasso M-estimation of the VAR(p) model in our empirical application with $k = 225$ assets.

The estimator $\hat{\beta}_{\mathcal{M}}$ for the VAR(p) process is motivated from the economic perspective, but appears to be statistically inconsistent, because non-zero correlations between R_t and w_t^* cause an endogeneity problem. Consistent M-estimators can be obtained by replacing the matrix R in Eq. (16) by the matrix $\tilde{R} = \text{diag}(\tilde{R}_1, \ldots, \tilde{R}_T)$ with $\tilde{R}_t = E[R_t | \mathcal{F}_{t-1}]$, or by any other measurable function of $\mathcal{F}_{t-1}$. For the upcoming empirical application we employ $\tilde{R}_t = R_{t-1}$, i.e. the popular random walk forecast, which is motivated by the strong autoregressive persistence of realized covariance matrices.[1] We denote the estimator $\tilde{\theta}_{\mathcal{M}}$ obtained by minimizing (16) with R_t replaced by $\tilde{R}_t$ as the $\tilde{M}$-estimator in order to distinguish it from the $\hat{M}$-estimator $\hat{\theta}_{\mathcal{M}}$.

For the VAR(p) process we obtain the consistent estimator $\tilde{\beta}_{\mathcal{M}}$ in the following proposition.

Proposition 3. *For the VAR(p) model in (18) the M-estimator of β which minimizes the objective function in Eq. (17) with R_t replaced by $\tilde{R}_t = R_{t-1}$ obtains as*

$$\tilde{\beta}_{\mathcal{M}} = [\tilde{\tilde{Z}}\tilde{\tilde{Z}}']^{-1} \tilde{\tilde{Z}} \tilde{R}^{1/2} (w^* - \iota_T \otimes e_k) = \beta + [\tilde{\tilde{Z}}\tilde{\tilde{Z}}']^{-1} \tilde{\tilde{Z}} \tilde{R}^{1/2} (I_T \otimes \tilde{I}_{k\times\ell}) u,$$

where $\tilde{R} = \text{diag}(R_0, \ldots, R_{T-1})$, $\tilde{\tilde{Z}} = (Z \otimes \tilde{I}_{k\times\ell}') \tilde{R}^{1/2}$, $\tilde{I}_{k\times\ell} = [I_\ell, -\iota_\ell]'$ and $e_k = (0, \ldots, 0, 1)'$.

Additionally, assume that the matrix $\tilde{\Gamma} = \text{plim}(\tilde{\tilde{Z}}\tilde{\tilde{Z}}'/T)$ exists and is non-singular. Then for $T \to \infty$ the estimator $\tilde{\beta}_{\mathcal{M}}$ is consistent with the asymptotic distribution

$$\begin{aligned} &\sqrt{T}(\tilde{\beta}_{\mathcal{M}} - \beta) \xrightarrow{L} \mathcal{N}(0, \tilde{\Gamma}^{-1} \tilde{\Psi} \tilde{\Gamma}^{-1}), \\ &\quad \text{with} \quad \tilde{\Psi} = \text{plim}\left(\tilde{\tilde{Z}} \tilde{R}^{1/2} (I_T \otimes (\tilde{I}_{k\times\ell} \Sigma_u \tilde{I}_{k\times\ell}')) \tilde{R}^{1/2} \tilde{\tilde{Z}}'/T\right). \end{aligned}$$

Although the closed-form results in Propositions 2 and 3 are appealing for low-dimensional settings, calculation of the VAR(p)

[1] As an alternative to the random walk forecast for $\tilde{R}_t$ we also investigated forecasts generated by a fitted scalar CAW(1,1) model (cf. Golosnoy et al., 2012), which however led to slightly worse forecasting results under the M-estimation scheme. This finding may be explained by the additional estimation uncertainty introduced via the autoregressive scalar CAW parameters. We therefore concentrate on the random walk forecast primarily due to its intuitiveness and implementation simplicity.

M-estimator and its asymptotic covariance matrix involve matrix-manipulations which are computationally infeasible in high dimensional settings. For our high dimensional scenario in the empirical application we therefore resort to numerical optimization of the VAR(p) M-estimation objective and compute asymptotic standard errors via the QML covariance. In particular, we obtain parameter estimates for the AR(FI)MA and HAR models by numerical optimization of the objective functions in Eqs. (16) and (17). Since these estimators can be interpreted as Gaussian QML estimators (cf. Wooldridge, 1994), inference can be based on the robust asymptotic QML covariance matrix, which is obtained as

$$\widehat{ACov}(\hat{\theta}) = (1/T)\cdot\left[\hat{D}_T\hat{S}_T^{-1}\hat{D}_T'\right]^{-1}$$
$$\text{with}\quad \hat{S}_T = (1/T)\cdot\sum_{t=1}^{T} h_t h_t', \quad \hat{D}_T = (1/T)\cdot H, \tag{20}$$

where h_t is the vector of the quasi-likelihood scores and H is the corresponding positive definite Hessian matrix evaluated at the parameter estimates (see e.g. Lee and Hansen, 1994). The asymptotic covariance in (20) coincides with the GMM covariance under zero mean Gaussian ML scores and is valid as far as the conditional mean of the GMVP weight vector process is correctly specified.

Finally note that LS and M-estimators of scalar/diagonal MHAR and ARFIMA models are applicable to vast-dimensional settings with $k > T$, as long as the total number of observations $k \cdot T$ exceeds the number of model parameters. Consistent regularized realized estimators for vast-dimensional covariance matrices are discussed in Fan et al. (2016) and Bollerslev et al. (2019).

5. Empirical study

We now apply the direct time series models for realized GMVP weights to empirical data and compare the performance of the direct models with several prominent indirect competitors both in sample and out of sample. We consider a GMVP based on $k = 225$ stocks selected by liquidity from the S&P 500 which are assigned to one of nine industry sectors according to the GICS (cf. Fan et al., 2016). The information on the stocks and industry classification is given in the supplementary online Appendix. The sample starts at January 3, 2007, and ends on December 31, 2012, with $T = 1510$ trading days; 1-minute closing prices between 9:30 a.m. and 4:00 p.m local time are obtained from QuantQuote.com. We compute daily realized covariance matrices R_t by the data efficient and microstructure noise robust composite realized kernel method of Lunde et al. (2016).[2] As this estimator involves an eigenvalue cleaning procedure, it is also applicable in situations when the number of stocks exceeds the number of intraday price observations (see e.g. Bollerslev et al., 2019, p. 118). Our data set then comprises $(k^2 + k)/2$ =25,425 time series of distinct realized variances and covariances.[3]

In order to visualize the data we provide in Fig. 1 envelop plots of 5% and 95% sample quantiles of the (co)variance, correlation and realized GMVP weight series at each point in time (left panel) together with the associated sample ACFs (right panel). The (co)variance data shows a slowly decaying serial correlation pattern and indicates the presence of long memory type persistence which may be well approximated by cascade models such as the HAR model of Corsi (2009). This long memory type of dependence also applies to the realized correlations and realized GMVP weights whereby the latter show a slow decay of the ACF, albeit at a substantially reduced level. The realized GMVP weights also appear more stable over time and exhibit much less heteroskedasticity compared to the (co)variance time series. Note that the reduced variability in the realized weights as compared to the realized covariance series is expected to result in a higher predictive power of our direct models for portfolio weights. This leads to smaller forecasting errors, and, hence, induces a smaller turnover, which is of importance under proportional transaction costs.

5.1. Competing models

We consider two major classes of approaches for predicting the GMVP weight vectors: first, direct approaches based on realized GMVP weights and, second, indirect approaches based on daily realized covariance matrix predictions which are used to obtain the GMVP forecasts.

As for the direct GMVP approaches, we consider both fully parameterized VAR and MHAR-type specifications estimated equation-wise by Lasso techniques, as well as sparsely parameterized diagonal and scalar VAR(FI)MA and MHAR models. The latter approaches are often applied as flexible approximations for strongly persistent autoregressive dynamics as observed for the realized GMVP weight series in Fig. 1. We obtain overall eleven direct GMVP competitor models, namely

- (i)–(iii) the Lasso-based approaches: Lasso-w (equation-wise Lasso estimation of the full VAR(100) model according to Eq. (11) of Section 3.2), Post-Lasso-w which corresponds to the former with OLS estimation applied subsequently to the Lasso-selected model structure (Post-Lasso OLS), and the fully data driven Lasso-w-HAR as given in Eq. (15) with $p = 100$;
- (iv)–(vii) diagonal VARFIMA(1,d,1), VAR(1), VARMA(1,1), and MHAR models estimated by LS;
- (viii)-(xi) scalar VARFIMA(1,d,1), VAR(1), VARMA(1,1), and MHAR models estimated by LS, $\widehat{\mathcal{M}}$ and $\widetilde{\mathcal{M}}$.

All estimations are based on data for the periods $t = \tau + 1, \ldots, T$, where the first $\tau = 20$ observations are reserved in order to initialize the VAR and MHAR model recursions and to approximate the initial forecasts by the arithmetic mean from the pre-sample, $(1/\tau)\sum_{t=1}^{\tau} w_t$. We then model the demeaned weight vectors $\tilde{w}_t^* = w_t^* - (1/T)\sum_{t=1}^{T} w_t^*$ in order to eliminate the vector of intercepts in the autoregressive specifications, which corresponds to the covariance targeting approach for indirect GMVP models (cf. Bauwens et al., 2017). For fitting the VAR, VAR(FI)MA and MHAR models we apply both LS- and M-estimators. The corresponding objective functions are minimized by Quasi-Newton numerical routines with BFGS updating of the Hessian matrix. For LS estimation of the diagonal model specifications we run parallel regressions for all ℓ realized weight series. M-estimation of diagonal VAR(FI)MA and MHAR models is not manageable in an economically sensible amount of time because parallelization is not possible for the M-estimation approach. For this reason we refrain from applying M-estimation to the diagonal VAR(FI)MA and MHAR models and provide M-estimation results only for their scalar counterparts.

As competing indirect GMVP forecasting models we consider four recently proposed specifications which are suitable for high-dimensional realized covariance data.

- (xii) The Factor-HEAVY model of Sheppard and Xu (2019), denoted by 3F-HEAVY. The model assumes a conditional Wishart

[2] Note that the composite realized kernel is not robust with respect to discrete jumps in the intraday price process, just as the recent data-efficient realized covariance matrix estimators proposed by Boudt et al. (2017); Brownlees et al. (2018); Bollerslev et al. (2018, 2019). In fact, Christensen et al. (2014) discover that the effect of jumps on realized measures is minor for data sampled at higher frequencies than every 5 min. Similar evidence is reported in Dette et al. (2021) for 1-minute data. Moreover, the effect of common jumps affecting all the realized (co)variances cancels out when computing the realized GMVP weights, see Eq. (8).

[3] For our data set with minute closing prices, i.e. typically 390 price observations per trading day, the median (computed over the 1510 trading days) of the minimum number of refresh-time-synchronized returns within the composite realized kernel amounts to 351, implying overall precise covariance estimates.

Table 1
Lasso results on significant (i.e. non-zero) cross-effects between industry sectors. The ith row of the table shows the fraction of total significant sector spillovers coming from the jth column sector. Industry sectors labels: (E) energy; (M) materials; (I) industrials; (D) consumer discretionary; (S) consumer staples; (H) health care; (F) financials; (T) information technologies; (U) utilities.

From :	E	M	I	D	S	H	F	T	U
to E	0.1646	0.0506	0.0759	0.1646	0.1139	0.3291	0.0506	0.0506	0
to M	0.0167	0.4167	0.0167	0.1833	0	0.0833	0.1833	0.1000	0
to I	0.0435	0.0124	0.0994	0.2981	0.0186	0.1801	0.2050	0.0870	0.0559
to D	0.0169	0.0135	0.0608	0.2601	0.0473	0.1892	0.3041	0.0608	0.0473
to S	0.0426	0.0355	0.0638	0.1418	0.1206	0.1631	0.1915	0.1915	0.0496
to H	0.0584	0.0117	0.0545	0.1984	0.0623	0.3307	0.1751	0.0973	0.0117
to F	0.0229	0.0229	0.0659	0.1433	0.0487	0.0716	0.5014	0.0888	0.0344
to T	0.0377	0.0293	0.0879	0.2008	0.0753	0.1130	0.1883	0.2092	0.0586
to U	0.0147	0.0147	0.1029	0.2647	0.0735	0.0294	0.2206	0.1618	0.1176

Table 2
In-sample LS-, $\widehat{\mathcal{M}}$- and $\widetilde{\mathcal{M}}$-estimation results for the scalar VAR(1), VARMA(1,1), MHAR and ARFIMA models.

Model	Scal VAR(1)		Scal VARMA(1,1)		Scal MHAR		Scal ARFIMA(1,d,1)	
LS estimation results								
param.	est	std	est	std	est	std	est	std
ϕ_1 or b_1	0.2843	(0.0062)	0.9568	(0.0038)	0.1123	(0.0061)	0.3750	(0.0148)
θ_1 or b_2			0.8358	(0.0070)	0.2065	(0.0130)	0.5880	(0.0218)
b_3 or d					0.4342	(0.0208)	0.3855	(0.0181)
$\widehat{\mathcal{M}}$-estimation results								
param.	est	std	est	std	est	std	est	std
ϕ_1 or b_1	0.9016	–	0.3379	–	0.1963	–	-0.0272	–
θ_1 or b_2			0.7654	–	0.7313	–	0.3841	–
b_3 or d					0.5531	–	0.7999	–
$\widetilde{\mathcal{M}}$-estimation results								
param.	est	std	est	std	est	std	est	std
ϕ_1 or b_1	0.4261	(0.0140)	0.9779	(0.0146)	0.1339	(0.0152)	0.3053	(0.0851)
θ_1 or b_2			0.8127	(0.0322)	0.2549	(0.0409)	0.6940	(0.0516)
b_3 or d					0.5007	(0.0548)	0.5970	(0.0894)

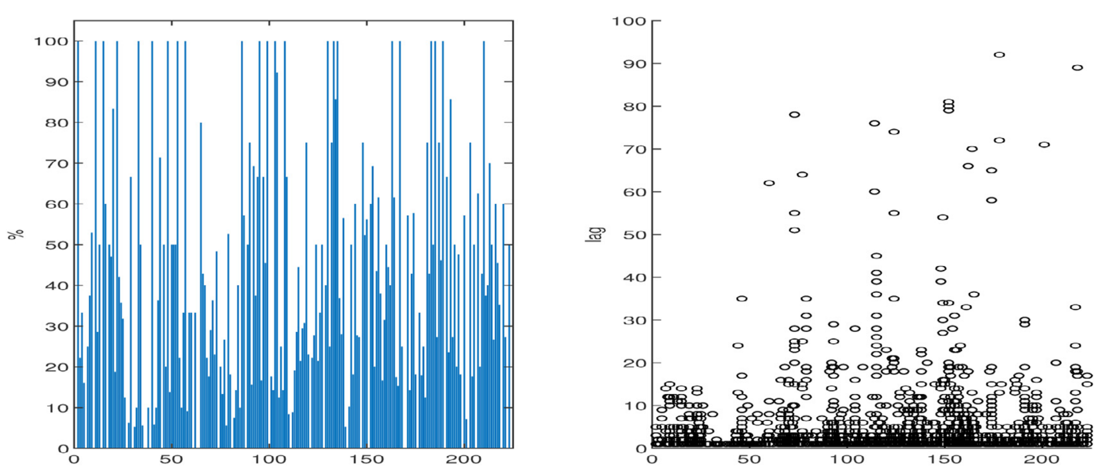

Fig. 2. Lasso estimation results. Left panel: Fraction of significant (non-zero) 'own-lag' coefficients of the ith VAR equation, $i = 1, \ldots, 224$, relative to all significant coefficients of the respective equation. Right panel: significant 'own-lags' of each GMVP weight series $i = 1, \ldots, 224$.

distribution for the joint realized covariance matrix of the underlying asset returns and observed risk factors, and employs a factor decomposition for the asset covariance matrix conditional on the factor (co)variances. The Factor-HEAVY then imposes a diagonal idiosyncratic covariance matrix and features dynamic factor (co)variances, dynamic loadings and dynamic idiosyncratic variances (see the HEAVY-M recursions in Eq. 15 of Sheppard and Xu, 2019, for details). For the factor covariances we use realized (co)variance data for the three Fama-French factors computed as composite realized kernels using intraday return data for the S&P 500 and the HML and SMB factors.[4] The model is estimated via QML as in Sheppard and Xu (2019).

- (xiii) The Factor-Lasso approach of Brito et al. (2018), denoted by 3F-Lasso, which employs a factor decomposition for the realized asset covariance matrix conditional on realized factor (co)variances with block diagonal idiosyncratic covariance matrix. We use Lasso VAR regressions for the matrix-log of the idiosyncratic covariance blocks and the factor covariance ma-

[4] We are grateful to Dacheng Xiu for kindly providing us with the intraday Fama-French factor return data which have been originally used in Fan et al. (2016) and Gribisch et al. (2020).

trix with the 3 Fama-French factors as in (xii). The loadings are modeled by independent HAR recursions estimated via OLS, see Eqs. (7), (9), (15), and (17) in Brito et al. (2018).

- (xiv) The DCC-NL model of Engle et al. (2019) which employs the nonlinear shrinkage approach of Ledoit and Wolf (2012) for the correlation targeting matrix. The model is estimated using QML based on daily asset returns and does not exploit high-frequency return information. The DCC-NL approach is suitable for vast dimensional portfolio problems and is the current benchmark for predicting covariance matrices and GMVP weights based on daily asset returns (cf. Moura et al., 2020).
- (xv) The scalar CAW(1,1) model of Golosnoy et al. (2012) which constitutes a particularly parsimonious way of modeling realized covariance matrices. The sparse parameterization of the scalar CAW allows us to consider both ML and $\widehat{\mathcal{M}}$ estimation.

Since our empirical application involves 225 assets we are forced to consider a set of competing models, which can be reasonably fitted to high-dimensional covariance data. The set of available benchmark models is therefore restricted and potential candidates effectively boil down to sparse factor approaches and covariance models with heavy parameter restrictions, such as the DCC and the scalar CAW illustrated above. The Factor-HEAVY model and the Factor-Lasso approach thereby rely on an information set which is different to the direct GMVP models, namely realized covariance data on the Fama-French factors. While common factors are clearly present in the realized covariance data, these factors tend to 'cancel out' when computing the GMVP weights (compare our discussion in Section 3). In fact an initial PCA analysis for the realized weights did not indicate a dominant factor structure and accordingly we do not consider direct factor-based approaches for the realized GMVP weights.

Finally, recent results of Golosnoy et al. (2019) indicate that despite the apparent advantages of the direct GMVP approach it might be worthwhile to combine sparsely parameterized direct and indirect GMVP predictors in a linear way in order to further improve the GMVP forecasting performance via taking "the best of both worlds". As the final GMVP modeling approach (xvi) we therefore consider a linear combination of direct scalar VARMA(1,1) forecasts and indirect scalar CAW(1,1) forecasts for the GMVP weights (labeled *VARMA-scalCAW combi*). We then write the GMVP forecast as

$$v_t = (1-\psi)v_t^{(\mathrm{caw})} + \psi v_t^{(\mathrm{varma})}, \qquad (21)$$

where $v_t^{(\mathrm{caw})}$ and $v_t^{(\mathrm{varma})}$ are GMVP weight forecasts obtained from the indirect scalar CAW(1,1) and the direct scalar VARMA(1,1) models, respectively. The fraction $\psi \in [0, 1]$ corresponds to the weight of the direct approach. The intercept vector for the VARMA(1,1) and the target matrix for the scalar CAW(1,1) recursions are obtained (as usual) by targeting procedures as in Bauwens et al. (2017). The remaining five parameters – two autoregressive parameters for the scalar VARMA(1,1), two autoregressive parameters for the scalar CAW(1,1), and the combination factor ψ – are selected in a one-step procedure by minimizing the LS- or M-estimation objective functions in Eqs. (14) and (16), respectively, via numerical quasi-Newton techniques.

While we combine VARMA(1,1) and scalar CAW forecasts it is of course possible to consider other model combinations as well. However, as illustrated in the upcoming sections, the scalar VARMA and CAW models show the overall best economic in- and out-of-sample GMVP performance. Moreover, they both imply particularly sparse VARMA-type specifications for the realized weights and the realized covariances, respectively, allowing for joint one-step optimization of the objective functions with respect to all involved model parameters of the combination approach.

Table 3
Evaluation of the in-sample forecasting accuracy: average differences of the forecasted and realized GMVP variances ($\mathcal{L}_E$, see Eq. (22)) and MSEs ($\mathcal{L}_S$, see Eq. (23)) for realized GMVP weights along with the corresponding MCS *p*-values in parenthesis. Bold numbers: minimum values for the respective loss functions.

Model	Post-Lasso-*w*	Lasso-*w*-HAR	Lasso-*w*	Diag ARFIMA $(1,d,1)$	Diag VAR(1)	Diag VARMA (1,1)	Diag MHAR	Scal ARFIMA $(1,d,1)$	Scal VAR(1)	Scal VARMA (1,1)	Scal MHAR	3F-HEAVY	3F-Lasso	DCC-NL	scal CAW	VARMA-scalCAW combi
	Economic Loss $\mathcal{L}_E$															
LS/ML/Lasso	0.3454	0.3794	0.4298	0.3463	0.4358	0.3416	0.3432	0.3411	0.4347	0.3432	0.3417	0.3715	0.2781	0.4996	0.2253	
0.3414	(0.0006)	(0.0006)	(0.0005)	(0.0015)	(0)	(0.0007)	(0.0009)	(0.0057)	(0.0006)	(0.0051)	(0.0051)	(0.0006)	(0.0094)	(0)	(0.0099)	(0.0015)
$\widehat{\mathcal{M}}$-estim.	(–)	(–)	(–)	(–)	(–)	(–)	(–)	0.2910	0.3348	0.2674	0.2674	(–)	(–)	(–)	0.2238	**0.2183**
	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(0.0062)	(0.0051)	(0.0094)	(0.0094)	(–)	(–)	(–)	(0.0144)	(1)
$\widetilde{\mathcal{M}}$-estim.	(–)	(–)	(–)	(–)	(–)	(–)	(–)	0.3176	0.3941	0.3149	0.3181	(–)	(–)	(–)	(–)	(–)
	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(0.0062)	(0.0003)	(0.0062)	(0.0062)	(–)	(–)	(–)	(–)	(–)
	Statistical Loss $\mathcal{L}_S$															
LS/ML/Lasso	0.0561	0.0485	0.0514	0.0475	0.0516	0.0479	0.0476	**0.0474**	0.0515	0.0478	0.0475	0.0544	0.0588	0.1674	0.0794	0.0476
	(0)	(0.0067)	(0)	(0.0983)	(0)	(0.0191)	(0.0853)	(1)	(0)	(0.0853)	(0.0853)	(0)	(0)	(0)	(0)	(0.2914)
$\widehat{\mathcal{M}}$-estim.	(–)	(–)	(–)	(–)	(–)	(–)	(–)	0.0527	0.0728	0.0584	0.0579	(–)	(–)	(–)	0.0781	0.0683
	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(0)	(0)	(0)	(0)	(–)	(–)	(–)	(0)	(0)
$\widetilde{\mathcal{M}}$-estim.	(–)	(–)	(–)	(–)	(–)	(–)	(–)	0.0477	0.0526	0.04826	0.0478	(–)	(–)	(–)	(–)	(–)
	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(0.0853)	(0)	(0.0004)	(0.0191)	(–)	(–)	(–)	(–)	(–)

Table 4
Descriptive statistics for the in-sample weight forecasts.

model	$\lvert v_i^{(\bullet)}\rvert \times 100$	$\max v_i^{(\bullet)} \times 100$	$\min v_i^{(\bullet)} \times 100$	$\sum v_i^{(\bullet)}\ \mathbf{1}(v_i^{(\bullet)} < 0)$	$\sum \mathbf{1}(v_i^{(\bullet)} < 0)/225$	$\sum \lvert v_{i,t}^{(\bullet)} - v_{i,t-1}^{(\bullet)}\rvert$	$S_p^{(\bullet)}$
Post-Lasso-w	0.7459	8.8777	−5.1704	−0.3391	0.4280	0.7269	10.1642
Lasso-w-HAR	0.6504	5.7735	−1.9015	−0.2317	0.3982	0.1689	10.6758
Lasso-w	0.6383	5.8448	−2.8558	−0.2181	0.3909	0.3035	11.0602
Diag ARFIMA $(1, d, 1)$	0.6893	6.4761	−1.5455	−0.2755	0.4108	0.3544	10.4182
Diag VAR(1) LS	0.6434	6.0466	−1.6537	−0.2238	0.3889	0.6766	11.1536
Diag VARMA(1,1) LS	0.6892	6.5149	−1.6078	−0.2753	0.4107	0.2383	10.2954
Diag MHAR LS	0.6901	6.5294	−1.5286	−0.2763	0.4113	0.3733	10.3036
Scal ARFIMA $(1, d, 1)$	0.6951	6.3303	−0.9476	−0.2820	0.4151	0.3677	10.3566
Scal VAR(1) LS	0.6433	5.7388	−0.7909	−0.2238	0.3865	0.7302	11.1350
Scal VARMA(1,1) LS	0.6887	6.3354	−0.9279	−0.2748	0.4117	0.2415	10.3661
Scal MHAR LS	0.6906	6.3470	−0.9347	−0.2769	0.4119	0.3941	10.3451
Scal ARFIMA $(1, d, 1)$ $\widehat{\mathcal{M}}$	0.8196	8.6289	−1.2918	−0.4220	0.4553	0.8200	9.7501
Scal VAR(1) $\widehat{\mathcal{M}}$	0.9649	11.8658	−1.7267	−0.5855	0.4836	2.3159	10.0103
Scal VARMA(1,1) $\widehat{\mathcal{M}}$	0.9540	9.9954	−1.6836	−0.5733	0.4797	0.6825	9.8382
Scal MHAR $\widehat{\mathcal{M}}$	0.9536	9.9294	−1.6858	−0.5728	0.4792	0.8295	9.8380
Scal ARFIMA $(1, d, 1)$ $\widetilde{\mathcal{M}}$	0.7301	6.9003	−1.0392	−0.3214	0.4288	0.4498	10.1067
Scal VAR(1) $\widetilde{\mathcal{M}}$	0.6958	7.0089	−0.9557	−0.2827	0.4098	1.0945	10.6542
Scal VARMA(1,1) $\widetilde{\mathcal{M}}$	0.7327	7.0291	−1.0456	−0.3242	0.4291	0.3308	10.0776
Scal MHAR $\widetilde{\mathcal{M}}$	0.7270	6.9143	−1.0323	−0.3179	0.4271	0.4718	10.1161
3F-HEAVY	0.9026	7.4487	−1.2425	−0.5154	0.5031	0.3870	10.7135
3F-Lasso	1.0259	9.1553	−2.0498	−0.6542	0.4890	0.5642	10.1041
DCC-NL	1.7077	16.8737	−5.1513	−1.4212	0.5003	0.8693	9.8048
scal CAW(1,1)	1.2830	11.9266	−3.2088	−0.9434	0.4971	0.7547	9.5446
scal CAW(1,1) $\widehat{\mathcal{M}}$	1.2780	11.8398	−3.1816	−0.9378	0.4941	0.6117	9.6101
VARMA-scalCAW combi, LS	1.1739	11.0340	−2.7733	−0.8206	0.4932	0.6556	9.3982
VARMA-scalCAW combi, $\widehat{\mathcal{M}}$	0.6986	6.4122	−0.9966	−0.2860	0.4127	0.2636	10.3181

5.2. Estimation results and in sample comparison

We start with the discussion of the model selection results obtained by the Lasso-w VAR(p) approach with maximum lag order $p = 100$. The Lasso estimates are computed as detailed in Section 3.2 using the Matlab routine *lasso*. In the left panel of Fig. 2 we depict the fraction of significant (i.e., remaining non-zero) 'own-lag' coefficients of the ith VAR equation, $i = 1, \ldots, 224$, i.e. coefficients for lags of the ith GMVP weight in the ith VAR equation, relative to all significant coefficients of the respective equation. The results imply that cross-effects between the GMVP weights for different assets appear overall highly relevant for many assets under consideration.

In Table 1 we provide a descriptive analysis of the relative importance of cross-sectional effects. In particular, we report for all 9 GICS sectors the fraction of non-zero spillover coefficients capturing cross-effects from GMVP weights for assets from the column sectors to the row sectors, hence, the numbers in each row sum up to one. The results do not indicate a clear sectoral pattern but confirm the importance of cross-effects as reflected by the Lasso VAR estimates. Particularly pronounced are the spillover effects from the consumer discretionary and financial sectors.

The significant own-lags of each GMVP weight are depicted in the right panel of Fig. 2. The first 20 lags appear to be the most important drivers of the autoregressive dynamics. For some weight series we obtain significant autoregressive coefficients for more than 40 lags and even beyond 90 lags, indicating a long-memory type of persistence pattern in the data, which motivates the analysis of ARFIMA and HAR dynamics for the realized GMVP weights.

In Table 2 we provide the full sample parameter estimates and standard errors for the scalar VAR(1), VARMA(1,1), MHAR, and ARFIMA(1,d,1) models under the LS-, $\widehat{\mathcal{M}}$-, and $\widetilde{\mathcal{M}}$ estimation schemes. While the LS- and $\widetilde{\mathcal{M}}$-estimates imply model stability for almost all model specifications, this is not the case for the scalar VARMA, MHAR and ARFIMA models under the $\widehat{\mathcal{M}}$ estimation approach and for the ARFIMA model under the $\widetilde{\mathcal{M}}$ setting. Moreover, while the consistent $\widetilde{\mathcal{M}}$-estimates appear overall close to those obtained by LS, the $\widehat{\mathcal{M}}$-estimation scheme results in parameter estimates which are substantially different from their LS counterparts. This finding could be explained by the inconsistency of $\widehat{\mathcal{M}}$-estimation induced by the direct minimization of the GMVP variance in (17). The estimates of the ARFIMA d parameters imply pronounced long memory in the series of GMVP weights. For the diagonal ARFIMA specification the d-estimates range from 0.13 to 0.79 with a median value of 0.34. Note that 75% of the d-estimates are statistically significant at the 1% level.

We now turn to the analysis of the relative fit of the introduced direct GMVP specifications and their indirect competitor models via an in-sample forecasting application. We denote the one period ahead GMVP forecast from each of the considered models by $v_t^* = E[w_t^* | \mathcal{F}_{t-1}]$ and compare this forecasted weight vector to the ex-post realization w_t^*. For $t = \tau + 1, \ldots, T$ we use both economic and statistical loss functions to evaluate the forecasting performance.

From the *economic* perspective, the best GMVP forecast should provide the smallest realized GMVP variance with the associated loss

$$\mathcal{L}_E = \frac{1}{T-\tau} \sum_{t=\tau+1}^{T} \left((v_t^*)' R_t v_t^* - (w_t^*)' R_t w_t^* \right) = \frac{1}{T-\tau} \sum_{t=\tau+1}^{T} \left((v_t^*)' R_t v_t^* - 1/(\iota_k' R_t^{-1} \iota_k) \right), \tag{22}$$

which is a natural criterion for the evaluation of the forecasting performance from the minimum-variance investor's perspective (see e.g. Chiriac and Voev, 2011). As an additional *statistical* criterion we consider the mean squared error (MSE) computed for all k weight forecasts:

$$\mathcal{L}_S = \frac{1}{T-\tau} \sum_{t=\tau+1}^{T} (w_t^* - v_t^*)'(w_t^* - v_t^*). \tag{23}$$

In order to assess the statistical significance of ex-post realized losses we implement the model confidence set (MCS) approach of Hansen et al. (2011) which identifies the set of models with the best forecasting performance at a given confidence level. The MCSs are computed by sequential application of Diebold–Mariano-type tests for equal predictive ability, where the null distribution of the

Table 5
Evaluation of the out-of-sample forecasting accuracy: average out-of-sample differences of the forecasted and realized GMVP variances ($\mathcal{L}_E$, see Eq. (22)) and MSEs ($\mathcal{L}_S$, see Eq. (23)) for realized GMVP weights along with the corresponding MCS p-values in parenthesis. Bold numbers: minimum values for the respective loss functions.

Model	Post-Lasso-w	Lasso-w-HAR	Lasso-w	Diag ARFIMA $(1,d,1)$	Diag VAR(1)	Diag VARMA (1,1)	Diag MHAR	Scal ARFIMA $(1,d,1)$	Scal VAR(1)	Scal VARMA (1,1)	Scal MHAR	3F-HEAVY	3F-Lasso	DCC-NL	scal CAW	VARMA-scalCAW combi
	Complete															
	Economic Loss $\mathcal{L}_E$															
LS/ML/Lasso	0.2017	0.1906	0.2242	0.1587	0.2094	0.1601	0.1588	0.1577	0.2121	0.1622	0.1603	0.1914	0.1357	0.2544	0.1219	0.1615
	(0.0011)	(0.0015)	(0.0001)	(0.0054)	(0.0007)	(0.0054)	(0.0040)	(0.0054)	(0.0012)	(0.0054)	(0.0054)	(0.0007)	(0.0091)	(0)	(0.0091)	(0.0064)
$\widehat{\mathcal{M}}$-estim.	(-)	(-)	(-)	(-)	(-)	(-)	(-)	0.1372	0.1562	0.1266	0.1258	(-)	(-)	(-)	0.1183	**0.1139**
	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(0.0059)	(0.0054)	(0.0091)	(0.0091)	(-)	(-)	(-)	(0.0067)	(1)
$\widetilde{\mathcal{M}}$-estim.	(-)	(-)	(-)	(-)	(-)	(-)	(-)	0.1469	0.1903	0.1479	0.1482	(-)	(-)	(-)	(-)	(-)
	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(0.0088)	(0.0007)	(0.0054)	(0.0059)	(-)	(-)	(-)	(-)	(-)
	Statistical Loss $\mathcal{L}_S$															
LS/ML/Lasso	0.0660	0.0504	0.0552	0.0485	0.0522	0.0489	0.0486	**0.0478**	0.0517	0.0482	0.0480	0.0536	0.0588	0.1640	0.0805	0.0481
	(0)	(0)	(0)	(0.0002)	(0)	(0)	(0)	(1)	(0)	(0.0002)	(0.0002)	(0)	(0)	(0)	(0)	(0.0004)
$\widehat{\mathcal{M}}$-estim.	(-)	(-)	(-)	(-)	(-)	(-)	(-)	0.0537	0.0745	0.0588	0.0583	(-)	(-)	(-)	0.0783	0.0687
	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(0)	(0)	(0)	(0)	(-)	(-)	(-)	(0)	(0)
$\widetilde{\mathcal{M}}$-estim.	(-)	(-)	(-)	(-)	(-)	(-)	(-)	0.0482	0.0532	0.0487	0.0484	(-)	(-)	(-)	(-)	(-)
	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(0.0002)	(0)	(0)	(0.0002)	(-)	(-)	(-)	(-)	(-)

test statistic is obtained by a block bootstrap with 10^4 bootstrap samples of loss differences and a block length of $\lfloor (T-\tau)^{1/3} \rfloor = 14$. For further details we refer to Hansen et al. (2011).

In Table 3 we evaluate the full sample forecasting accuracy under the economic and statistical loss functions in Eqs. (22) and (23), respectively. We start with comparing the direct GMVP approaches: The lowest economic losses $\mathcal{L}_E$ are obtained for the scalar VARMA(1,1) and the scalar MHAR model estimated by $\widehat{\mathcal{M}}$. Interestingly, all three direct Lasso approaches appear not competitive at all, although these models account for cross-effects, which are found to be present in the data, see Fig. 2 and Table 1. Even diagonal model specifications perform worse than their scalar counterparts; the data apparently favors models which are able to accommodate the persistent autocorrelation pattern of the realized weight series as indicated by the respective ACF envelop plot in the lower right part of Fig. 1. Accordingly, the persistent VARMA(1,1) and MHAR models clearly outperform the short-memory VAR(1) approach. Models estimated by the consistent $\widetilde{\mathcal{M}}$-approach are dominated by those estimated by $\widehat{\mathcal{M}}$, which is not unexpected, since the latter comes close to a direct minimization of the economic loss function $\mathcal{L}_E$. Overall the LS-based direct approaches are clearly outperformed by the $\widehat{\mathcal{M}}$-based estimators.

Turning to the economic losses for the indirect GMVP forecasting models, we find that the 3F-HEAVY and DCC-NL models are overall not competitive. Notably the DCC-NL performs even worse than the short-memory VAR(1) approach, a finding which could be explained by the fact that the DCC-NL exploits only daily asset return information for the forecasting of the GMVP weight series. Interestingly, the sparsely parameterized scalar CAW(1,1) model shows the overall best forecasting performance for the economic GMVP loss - even compared to the direct GMVP approaches. This finding might be an artefact of potential nonlinear GMVP dynamics which are not captured by the direct GMVP approaches. Further, we consider a model combination of the indirect scalar CAW(1,1) and the direct VARMA(1,1) as discussed in Eq. (21) of Section 5.1. The combination approach turns out to perform overall best under $\widehat{\mathcal{M}}$-estimation with an MCS p-value of one as compared to a p-value of 0.0144 for the $\widehat{\mathcal{M}}$-fitted scalar CAW.[5] The estimate of the combination weight ψ amounts to 0.2476, implying a relative weight of almost 25% for the direct VARMA(1,1) in the linear combination setting.

The lowest statistical loss $\mathcal{L}_S$ is obtained for the scalar ARFIMA model estimated by LS. The 95% MCS further includes the diagonal AFRIMA, diagonal MHAR, scalar VARMA(1,1) and scalar MHAR models, all estimated by LS, and the scalar ARFIMA model estimated by $\widetilde{\mathcal{M}}$. We observe that the LS-based scalar model specifications perform somewhat better than the corresponding diagonal ones under the statistical LS loss. This finding appears counterintuitive but is explained by the fact that we estimate the autoregressive specifications for the first $k-1$ weights only, but evaluate the loss $\mathcal{L}_S$ for all k weights jointly, compare Eqs. (14) and (23). The Lasso-based models are again not competitive, and models fitted by M-estimation perform significantly worse than those fitted by LS. Moreover the consistent $\widetilde{\mathcal{M}}$-estimators result in lower losses as compared to their $\widehat{\mathcal{M}}$ counterparts, which may be explained by the in-consistency of the $\widehat{\mathcal{M}}$ estimation scheme. Remarkably the 95% MCS contains not a single indirect GMVP approach. The LS-estimated scalar CAW(1,1)-scalar VARMA(1,1) model combination refers to the 10% MCS.

We note that our empirical results inevitably depend to some extent on the precision of the underlying realized weights and covariances matrices, which is largely determined by the number of

[5] Note that it is not possible to fit the 3F-HEAVY, 3F-Lasso and DCC-NL models via the $\widehat{\mathcal{M}}$-estimation approach due to the vast number of model parameters involved.

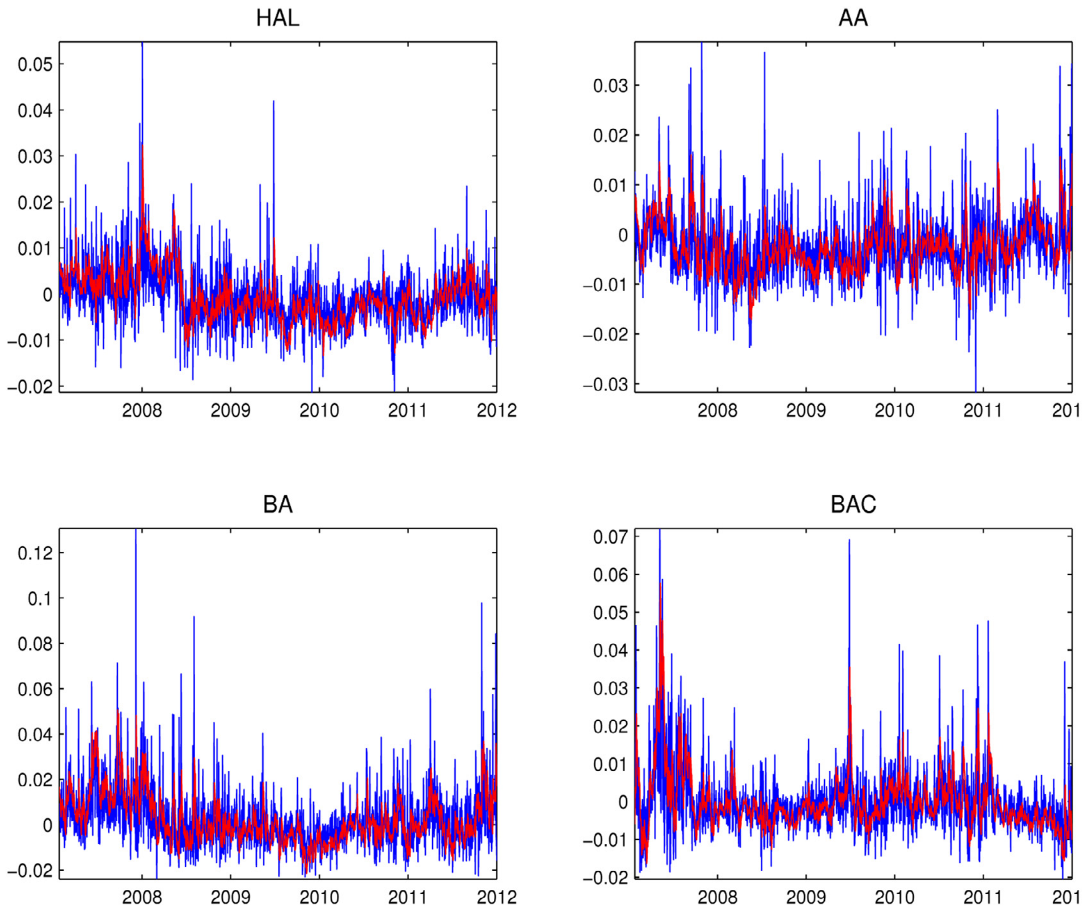

Fig. 3. In-sample fit: $\widehat{\mathcal{M}}$-estimated scalar VARMA(1,1), red; realized GMVP weights, blue: Halliburton (HAL, energy), Alcoa (AA, materials), Boeing (BA, industrials), Bank of America (BAC, financials). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

intraday returns, the microstructure noise level and the dimension of the underlying asset portfolio. In order to analyze the robustness of our empirical results we therefore conducted a small simulation exercise involving the simulation of up to 1000-dimensional realized weight vectors for various liquidity and microstructure noise levels for the underlying intraday returns. As a result we found that our major findings in Table 3 are overall robust for various return vector dimensions, liquidity and noise levels. The simulation results are available upon request.

In Fig. 3 we depict the series of realized GMVP weights for Halliburton (HAL), Alcoa (AA), Boeing (BA) and Bank of America (BAC) together with the corresponding in-sample forecasts from the scalar VARMA(1,1) model under $\widehat{\mathcal{M}}$-estimation. Although being inconsistently estimated, the fitted values in red capture successfully the observed dynamics of the realized weights in blue.

In Table 4 we report descriptive statistics for the forecasted GMVP weights in analogy to the analysis of Brandt et al. (2009). In particular, we compute for each approach $(\bullet)$ the average absolute portfolio weight, the average minimum and maximum weights, the average sum of negative weights in the portfolio, the average fraction of negative weights in the portfolio, and the portfolio's turnover. Additionally, we measure the in-sample performance in terms of the empirical standard deviation $S_P^{(\bullet)}$ computed from the "optimized" minimum variance portfolio returns $x_t^{(\bullet)} = x_t' \upsilon_t^{*(\bullet)}$. The later criterion is widely used in the financial literature for performance evaluation, see e.g. Brandt et al. (2009) or Engle and Kelly (2012). We obtain

$$(S_P^{(\bullet)})^2 = (T-\tau)^{-1} \sum_{t=\tau+1}^{T} \left(x_t^{(\bullet)} - \bar{x}^{(\bullet)}\right)^2,$$

where $\bar{x}^{(\bullet)} = (T-\tau)^{-1} \sum_{t=\tau+1}^{T} x_t^{(\bullet)}$ denotes the average portfolio return for the approach $(\bullet)$.

Based on the annualized values of $S_P^{(\bullet)}$ (reported in %) as a closely related alternative to the economic loss $\mathcal{L}_E$ we again obtain the model combination of the scalar CAW(1,1) and the scalar VARMA(1,1) as the best GMVP forecasting approach. Remarkably, the direct GMVP models estimated by LS- and $\widetilde{\mathcal{M}}$ lead to much less portfolio re-balancing turnover as compared to the $\widehat{\mathcal{M}}$-estimated models. They however show higher portfolio standard deviations $S_P^{(\bullet)}$ which indicates a trade-off between active variance minimization and paying higher proportional transaction costs.

5.3. Out-of-sample results

In order to evaluate the out-of-sample performance we compute one-period ahead forecasts υ_t of the GMVP weight vectors w_t for the 502 trading days from January 3, 2011 to December 31, 2012, which cover both (relatively) high as well as fairly low volatility levels and resemble stock market fluctuations which are typical for the post-subprime mortgage crises period. We evaluate the forecasting performance using the economic and statistical loss functions from Eqs. (22) and (23) computed for the out-of-sample forecasting periods, and report the out-of-sample standard deviations $S_P^{(\bullet)}$ and further descriptive measures on the GMVP weight forecasts. All models are re-estimated every new trading day using

Table 6
Descriptive statistics for the out-of-sample weight forecasts.

model	$\|v_i^{(\bullet)}\| \times 100$	max $v_i^{(\bullet)} \times 100$	min $v_i^{(\bullet)} \times 100$	$\sum v_i^{(\bullet)}\, \mathbf{1}(v_i^{(\bullet)} < 0)$	$\sum \mathbf{1}(v_i^{(\bullet)} < 0)/225$	$\sum \|v_{i,t}^{(\bullet)} - v_{i,t-1}^{(\bullet)}\|$	$S_p^{(\bullet)}$
Post-Lasso-w	0.7369	10.1596	−2.0105	−0.3290	0.4343	0.8026	7.3438
Lasso-w-HAR	0.6436	4.9666	−0.9698	−0.2241	0.3977	0.1857	7.2540
Lasso-w	0.6354	5.8379	−0.9868	−0.2149	0.3984	0.3170	7.7289
Diag ARFIMA $(1, d, 1)$	0.6918	5.6534	−1.4585	−0.2783	0.4089	0.3725	6.8031
Diag VAR(1) LS	0.6452	5.6846	−1.1656	−0.2259	0.3922	0.6960	7.6435
Diag VARMA(1,1) LS	0.6917	5.6670	−1.7238	−0.2782	0.4094	0.2596	6.7964
Diag MHAR LS	0.6915	5.6941	−1.4263	−0.2780	0.4101	0.3878	6.7695
Scal ARFIMA $(1, d, 1)$	0.6977	5.5985	−0.9837	−0.2849	0.4125	0.3801	6.8024
Scal VAR(1) LS	0.6474	5.5845	−0.8156	−0.2283	0.3866	0.7634	7.7533
Scal VARMA(1,1) LS	0.6879	5.5915	−0.9539	−0.2739	0.4090	0.2551	6.8396
Scal MHAR LS	0.6906	5.6371	−0.9651	−0.2770	0.4095	0.4081	6.8272
Scal ARFIMA $(1, d, 1)$ $\widehat{\mathcal{M}}$	0.8206	7.8540	−1.3194	−0.4231	0.4524	0.8333	6.5041
Scal VAR(1) $\widehat{\mathcal{M}}$	0.9781	11.2781	−1.8058	−0.6004	0.4822	2.3763	6.9575
Scal VARMA(1,1) $\widehat{\mathcal{M}}$	0.9513	8.8071	−1.7125	−0.5702	0.4738	0.7058	6.5570
Scal MHAR $\widehat{\mathcal{M}}$	0.9500	8.6793	−1.7224	−0.5688	0.4725	0.8395	6.5324
Scal ARFIMA $(1, d, 1)$ $\widetilde{\mathcal{M}}$	0.7300	6.0534	−1.0731	−0.3213	0.4252	0.4658	6.6398
Scal VAR(1) $\widetilde{\mathcal{M}}$	0.7008	6.7393	−0.9883	−0.2884	0.4087	1.1294	7.4192
Scal VARMA(1,1) $\widetilde{\mathcal{M}}$	0.7307	6.1638	−1.0713	−0.3221	0.4249	0.3374	6.6236
Scal MHAR $\widetilde{\mathcal{M}}$	0.7257	6.0700	−1.0637	−0.3164	0.4225	0.4855	6.6438
3F-HEAVY	0.8815	6.4833	−1.1914	−0.4917	0.4938	0.3672	7.6264
3F-Lasso	1.0137	7.5656	−1.9815	−0.6404	0.4885	0.5264	6.9030
DCC-NL	1.7364	14.6582	−5.4586	−1.4535	0.4995	0.9059	7.7905
scal CAW(1,1)	1.2886	11.2420	−3.3181	−0.9497	0.4924	0.7702	6.3516
scal CAW(1,1) $\widehat{\mathcal{M}}$	1.2768	11.0393	−3.2507	−0.9364	0.4874	0.6357	6.3175
VARMA-scalCAW combi, LS	0.6962	5.6623	−1.0015	−0.2832	0.4086	0.3249	6.8220
VARMA-scalCAW combi, $\widehat{\mathcal{M}}$	1.1689	10.0462	−2.7872	−0.8151	0.4866	0.6632	6.2854

expanding estimation windows. New forecasts are then generated based on the updated parameter estimates.

Table 5 reports the out-of-sample forecasting results obtained under the $\mathcal{L}_E$ and $\mathcal{L}_S$ loss functions. For the economic loss the forecasting results are overall very similar to our in-sample findings: Among the direct GMVP approaches the scalar MHAR and the scalar VARMA(1,1) estimated by $\widehat{\mathcal{M}}$ show the best performance. The diagonal models are again outperformed by their scalar counterparts - a finding which may be attributed to the additional estimation uncertainty affecting the point forecasts and inflating the variance of the forecast errors. The best results within the class of indirect GMVP models are obtained for the scalar CAW(1,1) estimated by $\widehat{\mathcal{M}}$ and the 3F-Lasso. The $\widehat{\mathcal{M}}$-fitted combination of the scalar CAW(1,1) and the scalar VARMA(1,1) again provides the overall smallest economic loss.

Turning to the statistical loss, we find that all indirect GMVP approaches are – not surprisingly – significantly outperformed by their direct GMVP competitors. The overall smallest statistical loss is obtained for the scalar ARFIMA model (the only constituent of the 99% MCS), closely followed by the scalar MHAR, both fitted by LS. The smallest out-of-sample standard deviation $S_p^{(\bullet)}$ is again obtained for the model combination approach under the $\widehat{\mathcal{M}}$ estimation scheme, see Table 6.

Based on both in- and out-of-sample results we conclude that the $\widehat{\mathcal{M}}$-fitted combination of the scalar VARMA(1,1) and the scalar CAW(1,1) performs best under the economic loss, whereas the LS-estimated scalar ARFIMA turns out to perform best under the statistical loss function. These models for realized GMVP weights show overall sound forecasting results while being both parsimonious in parametrization and easy to estimate.

6. Summary

We develop direct multiple linear time series models in order to forecast high dimensional vectors of realized global minimum variance portfolio (GMVP) weights computed based on intraday high-frequency returns. The direct forecasting of realized GMVP weights constitutes an alternative to the conventional (indirect) two step approach, which consists of (first) developing multivariate volatility models for the prediction of high dimensional realized covariance matrices, and (second) plugging the obtained covariance forecasts into the GMVP formula. The main advantages of our direct approaches are their applicability to high dimensional portfolio selection problems, the resulting MSE-optimal forecasts, as well as a convenient linear modeling where it is easy to impose interpretable parameter restrictions.

For the prediction of realized GMVP weight vectors we consider the general class of vector AR(FI)MA models and apply Lasso regression techniques in order to impose reasonable parameter restrictions and to mitigate the curse of dimensionality. We further derive M-type estimators specifically designed to minimize the economic loss criterium related to the realized portfolio variance and deduce the statistical properties of the obtained estimators."?>

We provide an empirical application to a high dimensional realized GMVP constructed from 225 stocks selected from the S&P 500. As benchmark models we consider a set of popular indirect approaches. Our main empirical findings are as follows: Under the economic GMVP variance loss function the $\widehat{\mathcal{M}}$-fitted combination of the direct scalar VARMA(1,1) and the indirect scalar CAW(1,1) turns out to perform best in-sample as well as out-of-sample. Hence, in terms of the economic GMVP loss we do not find a clear unique winner within the classes of direct and indirect GMVP approaches. This finding might be due to potential nonlinear GMVP dynamics which are not captured by the direct GMVP approaches but are well approximated by the indirect GMVP models. The result is also in line with the findings of Golosnoy et al. (2019) who show that even a naive combination of direct and indirect approaches for realized GMVP weights based on simple scalar exponential smoothing improves the goodness of forecasts for a broad set of common performance measures. Under the statistical MSE loss the best forecasting results are obtained for the LS-estimated direct scalar VARFIMA model, clearly outperforming the indirect GMVP forecasting models. The direct GMVP approach allows for an improved prediction of the location of the Markowitz mean-variance efficient frontier which is of substantial importance for portfolio investors.

Credit author statement

Both authors (Vasyl Golosnoy and Bastian Gribisch) have contributed all aspects of the paper in equal proportions.

Declaration of Competing Interest

None.

Acknowledgments

this work has been partly supported by the research project number 415503985 of the German Research Foundation (DFG).

Appendix A.

Proof of Proposition 1.. For $\omega_t^* = \Sigma_t^{-1}\iota_k/\iota_k'\Sigma_t^{-1}\iota_k$ we obtain by the multiplication rule

$$\frac{\partial\omega_t^*}{\partial \mathrm{vech}(\Sigma_t)'} = \frac{\partial(\Sigma_t^{-1}\iota_k)}{\partial \mathrm{vech}(\Sigma_t)'}\ (\iota_k'\Sigma_t^{-1}\iota_k)^{-1} + (\Sigma_t^{-1}\iota_k)\ \frac{\partial(\iota_k'\Sigma_t^{-1}\iota_k)^{-1}}{\partial \mathrm{vech}(\Sigma_t)'}.$$

With Lütkepohl (1996), p. 183 (3) and p. 198, 10.6 (2), and $D_kD_k^+ = \frac{1}{2}(I_{k^2}+K_k)$ we obtain

$$\begin{aligned}\frac{\partial(\Sigma_t^{-1}\iota_k)}{\partial \mathrm{vech}(\Sigma_t)'} &= \frac{\partial(\Sigma_t^{-1}\iota_k)}{\partial \mathrm{vech}(\Sigma_t^{-1})'}\frac{\partial \mathrm{vech}(\Sigma_t^{-1})}{\partial \mathrm{vech}(\Sigma_t)'}\\ &= \frac{\partial \mathrm{vec}(\Sigma_t^{-1}\iota_k)}{\partial \mathrm{vec}(\Sigma_t^{-1})'}\ \frac{\partial \mathrm{vec}(\Sigma_t^{-1})}{\partial \mathrm{vech}(\Sigma_t^{-1})'}\ \frac{\partial \mathrm{vech}(\Sigma_t^{-1})}{\partial \mathrm{vech}(\Sigma_t)'}\\ &= (\iota_k'\otimes I_k)\ D_k\ (-D_k^+(\Sigma_t^{-1}\otimes\Sigma_t^{-1})D_k)\\ &= -\frac{1}{2}(\iota_k'\otimes I_k)\ [I_{k^2}+K_k]\ (\Sigma_t^{-1}\otimes\Sigma_t^{-1})\ D_k,\end{aligned}$$

where D_k^+ denotes the Moore-Penrose inverse of D_k. Accordingly,

$$\begin{aligned}\frac{\partial(\iota_k'\Sigma_t^{-1}\iota_k)^{-1}}{\partial \mathrm{vech}(\Sigma_t)'} &= -(\iota_k'\Sigma_t^{-1}\iota_k)^{-2}\ \frac{\partial(\iota_k'\Sigma_t^{-1}\iota_k)}{\partial \mathrm{vech}(\Sigma_t)'}\\ &= -(\iota_k'\Sigma_t^{-1}\iota_k)^{-2}\ \frac{\partial(\iota_k'\Sigma_t^{-1}\iota_k)}{\partial \mathrm{vec}(\Sigma_t^{-1})'}\ \frac{\partial \mathrm{vec}(\Sigma_t^{-1})}{\partial \mathrm{vech}(\Sigma_t^{-1})}\\ &\quad \frac{\partial \mathrm{vech}(\Sigma_t^{-1})}{\partial \mathrm{vech}(\Sigma_t)'}\\ &= -(\iota_k'\Sigma_t^{-1}\iota_k)^{-2}\ (\iota_k'\otimes\iota_k')\ D_k\ (-D_k^+(\Sigma_t^{-1}\otimes\Sigma_t^{-1})D_k)\\ &= (1/2)\cdot(\iota_k'\Sigma_t^{-1}\iota_k)^{-2}\ (\iota_k'\otimes\iota_k')\ [I_{k^2}+K_k]\ (\Sigma_t^{-1}\otimes\Sigma_t^{-1})\ D_k.\end{aligned}$$

We then obtain

$$\begin{aligned}\frac{\partial\omega_t^*}{\partial \mathrm{vech}(\Sigma_t)'} &= \frac{1}{2(\iota_k'\Sigma_t^{-1}\iota_k)^2}\\ &\left\{\Sigma_t^{-1}\iota_k\ (\iota_k'\otimes\iota_k') - (\iota_k'\Sigma_t^{-1}\iota_k)\ (\iota_k'\otimes I_k)\right\}\\ &[I_{k^2}+K_k]\ (\Sigma_t^{-1}\otimes\Sigma_t^{-1})\ D_k,\end{aligned}$$

which completes the proof. □

Proof of Proposition 2.. We consider at each t the realized portfolio variance $(v_t^*)'R_tv_t^*$ of the GMVP given the information set $\mathcal{F}_{t-1}$. The sum of the realized portfolio variances is to be minimized with respect to β, i.e.

$$\min_\beta\left[\mathcal{M}(\beta) = \sum_{t=1}^T v_t^*(\beta)'R_tv_t^*(\beta) = v^*(\beta)'Rv^*(\beta)\right], \tag{24}$$

where the matrix R of dimension $kT\times kT$ is given as $R = \mathrm{diag}(R_1,\dots,R_T)$. The resulting $\mathcal{M}$-type (cf. Hayashi, 2000, p. 447) estimator $\widehat{\beta}_M$ is derived as follows. First note that $v_t^* = e_k + \tilde{I}_{k\times\ell}v_t$ with the matrix $\tilde{I}_{k\times\ell} = [I_\ell, -\iota_\ell]'$ and k-dimensional vector e_k with 1 on the kth position and a 0 otherwise. Then, using the vector representation $v = (Z'\otimes I_\ell)\beta$, we write

$$\begin{aligned}v^* &= \iota_T\otimes e_k + (I_T\otimes\tilde{I}_{k\times\ell})v = \iota_T\otimes e_k + (I_T\otimes\tilde{I}_{k\times\ell})(Z'\otimes I_\ell)\beta\\ &= \iota_T\otimes e_k + (Z'\otimes\tilde{I}_{k\times\ell})\beta,\end{aligned}$$

such that $u^* = w^* - v^* = (I_T\otimes\tilde{I}_{k\times\ell})u$. Next, observe that the optimization problem in (16) is equivalent to

$$\min_\beta\left[\mathcal{M}^*(\beta) = \sum_{t=1}^T(w_t^*-v_t^*)'R_t(w_t^*-v_t^*) = (w^*-v^*)'R(w^*-v^*) = (u^*)'Ru^*\right] \tag{25}$$

as it holds that

$$(w_t^*-v_t^*)'R_t(w_t^*-v_t^*) = (w_t^*)'R_tw_t^* - 2(v_t^*)'R_tw_t^* + (v_t^*)'R_tv_t^*,$$

where only the last component depends on β due to $R_tw_t^* = \iota_k/(\iota_k'R_t^{-1}\iota_k)$ and $(v_t^*)'\iota_k = 1$.

With $u^* = w^* - \iota_T\otimes e_k - (Z'\otimes\tilde{I}_{k\times\ell})\beta$, we get

$$\min_\beta\left[\mathcal{M}(\beta) = (w^*-\iota_T\otimes e_k-(Z'\otimes\tilde{I}_{k\times\ell})\beta)'R(w^*-\iota_T\otimes e_k-(Z'\otimes\tilde{I}_{k\times\ell})\beta)\right].$$

We denote $\tilde{Z} = (Z\otimes\tilde{I}_{k\times\ell}')R^{1/2}$ and obtain the solution

$$\begin{aligned}\hat{\beta}_{\mathcal{M}} &= [(Z\otimes\tilde{I}_{k\times\ell}')R(Z'\otimes\tilde{I}_{k\times\ell})]^{-1}(Z\otimes\tilde{I}_{k\times\ell}')R(w^*-\iota_T\otimes e_k)\\ &= \beta + [\tilde{Z}\tilde{Z}']^{-1}\tilde{Z}R^{1/2}(I_T\otimes\tilde{I}_{k\times\ell})u = \beta + [\tilde{Z}\tilde{Z}']^{-1}\tilde{Z}R^{1/2}u^*.\end{aligned}$$

This completes the proof. □

Proof of Proposition 3. In order to derive the stochastic properties of the M-estimator, assume that the ℓ-dimensional VAR process in (18) is stable. Additionally, assume that $E(u_t) = 0$, $E[u_tu_t'] = \Sigma_u$ positive definite and u_t, u_s are uncorrelated as well as that all fourth moments exist and are bounded. Note that the first and third condition are fulfilled by construction since $u_t = w_t - E[w_t|\mathcal{F}_{t-1}]$ is a martingale difference sequence.

Next, assume that $\mathrm{plim}(\tilde{Z}\tilde{Z}'/T) = \tilde{\Gamma}$ of dimension $(\ell p+1)\times(\ell p+1)$ exists and is non-singular, and consider the expression

$$\begin{aligned}\tilde{Z}\widetilde{R}^{1/2}u^* &= (Z\otimes\tilde{I}_{k\times\ell}')\widetilde{R}u^* = (Z\otimes\tilde{I}_{k\times\ell}')vec(\tilde{U}^*)\\ &= \mathrm{vec}(\tilde{I}_{k\times\ell}'\tilde{U}^*Z') = (I_{\ell p+1}\otimes\tilde{I}_{k\times\ell}')\mathrm{vec}(\tilde{U}^*Z')\\ &= (I_{\ell p+1}\otimes\tilde{I}_{k\times\ell}')\sum_{t=1}^T\mathrm{vec}(\tilde{u}_t^*Z_{t-1}')\\ &= (I_{\ell p+1}\otimes\tilde{I}_{k\times\ell}')\sum_{t=1}^T\mathrm{vec}(\widetilde{R}_tu_t^*Z_{t-1}'),\end{aligned}$$

where $\mathrm{vec}(\tilde{U}^*) = \widetilde{R}(I_T\otimes\tilde{I}_{k\times\ell})u$. Note that $\mathrm{vec}(\widetilde{R}_tu_t^*Z_{t-1}')$ is a martingale difference sequence which could be shown by applying the law of iterated expectations:

$$E[\mathrm{vec}(\widetilde{R}_tu_t^*Z_{t-1}')] = \mathrm{vec}(E[\widetilde{R}_tE[u_t^*|\mathcal{F}_{t-1}]Z_{t-1}']) = 0.$$

Due to the central limit theorem for ergodic martingale difference sequences (cf. Hayashi, 2000, p. 106) it holds that

$$\begin{aligned}&(1/\sqrt{T})(Z\otimes\tilde{I}_{k\times\ell}')\widetilde{R}(I_T\otimes\tilde{I}_{k\times\ell})u\\ &= (1/\sqrt{T})\tilde{Z}\widetilde{R}^{1/2}(I_T\otimes\tilde{I}_{k\times\ell})u\ \xrightarrow[T\to\infty]{L}\ \mathcal{N}(0,\widetilde{\Psi}),\end{aligned} \tag{26}$$

$$\text{with}\quad \widetilde{\Psi} = \mathrm{plim}\big(\tilde{Z}\widetilde{R}^{1/2}(I_T\otimes(\tilde{I}_{k\times\ell}\Sigma_u\tilde{I}_{k\times\ell}'))\widetilde{R}^{1/2}\tilde{Z}'/T\big). \tag{27}$$

Consistency follows by

$$\begin{aligned}&\mathrm{plim}(\hat{\beta}_{\mathcal{M}}-\beta) = \mathrm{plim}([\tilde{Z}\tilde{Z}'/T]^{-1})\ \mathrm{plim}(\tilde{Z}\widetilde{R}^{1/2}(I_T\otimes\tilde{I}_{k\times\ell})u/T)\\ &= \tilde{\Gamma}^{-1}\cdot 0 = 0,\end{aligned}$$

according to the Slutsky theorem. The asymptotic distribution is obtained for $T \to \infty$ as

$$\sqrt{T}(\hat{\beta}_{\mathcal{M}} - \beta) \xrightarrow{L} \mathcal{N}(0, \tilde{\Gamma}^{-1}\widetilde{\Psi}\tilde{\Gamma}^{-1}).$$

This completes the proof. □

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.jbankfin.2022.106404

References

Audrino, F., Huang, C., Okhrin, O., 2019. Flexible HAR model for realized volatility. Stud. Nonlinear Dyn. Econom. 20170080, 1–22.
Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2011. Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. J. Econom. 162, 149–169.
Barndorff-Nielsen, O.E., Shephard, N., 2004. Econometric analysis of realised covariation: high frequency covariance, regression and correlation in financial economics. Econometrica 72, 885–925.
Bauwens, L., Braione, M., Storti, G., 2017. A dynamic component model for forecasting high-dimensional realized covariance matrices. Econom. Stat. 1, 40–61.
Bauwens, L., Laurent, S., Rombouts, J.V.K., 2006. Multivariate GARCH models: a survey. J. Appl. Econom. 21, 79–109.
Belke, A., Beckmann, J., 2015. Monetary policy and stock prices – cross-country evidence from cointegrated VAR models. J. Bank. Finance 54, 254–265.
Belloni, A., Chernozhukov, V., 2013. Least squares after model selection in high-dimensional sparse models. Bernoulli 19, 521–547.
Best, M., Grauer, R., 1991. On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. Rev. Financ. Stud. 4, 315–342.
Bickel, P.J., Ritov, Y., Tsybakov, A.B., 2009. Simultaneous analysis of Lasso and Dantzig selector. Ann. Stat. 37 (4), 1705–1732.
Bollerslev, T., Meddahi, N., Nyawa, S., 2019. High-dimensional multivariate realized volatility estimation. J. Econom. 212, 116–136.
Bollerslev, T., Patton, A.J., Quaedvlieg, R., 2018. Modeling and forecasting (un)reliable realized covariances for more reliable financial decisions. J. Econom. 207, 71–91.
Boudt, K., Laurent, S., Lunde, A., Quaedvlieg, R., Sauri, O., 2017. Positive semidefinite integrated covariance estimation, factorisations and asynchronicity. J. Econom. 196, 347–367.
Brandt, M.W., Santa-Clara, P., Valkanov, R., 2009. Parametric portfolio policies: exploiting characteristics in the cross-section of equity returns. Rev. Financ. Stud. 22, 3411–3447.
Brito, D. S., Medeiros, M. C., Ribeiro, R. M., 2018. Forecasting large realized covariance matrices: the benefits of factor models and shrinkage. Working paper.
Brownlees, C., Nualart, E., Sun, Y., 2018. Realized networks. J. Appl. Econom. 33 (7), 986–1006.
Bubak, V., Kocenda, E., Zikes, F., 2011. Volatility transmission in emerging European foreign exchange markets. J. Bank. Finance 35 (11), 2829–2841.
Callot, L.A., Kock, A.B., Medeiros, M.C., 2017. Modeling and forecasting large realized covariance matrices and portfolio choice. J. Appl. Econom. 32, 140–158.
Chiriac, R., Voev, V., 2011. Modelling and forecasting multivariate realized volatility. J. Appl. Econom. 26, 922–947.
Christensen, K., Oomen, R., Podolskij, M., 2014. Fact or friction: jumps at ultra high frequency. J. Financ. Econ. 114 (3), 576–599.
Corsi, F., 2009. A simple approximative long-memory model of realized volatility. J. Financ. Econom. 7, 174–196.
DeMiguel, V., Garlappi, L., Nogales, F.J., Uppal, R., 2009. A generalized approach to portfolio optimization: improving performance by constraining portfolio norms. Manag. Sci. 55, 798–812.
DeNard, G., Ledoit, O., Wolf, W., 2021. Factor models for portfolio selection in large dimensions: the good, the better and the ugly. Journal of Financial Econometrics 19(2): 236-257.
Dette, H., Golosnoy, V., Kellermann, J., 2021. Correcting intraday periodicity bias in realized volatility measures. Econom. Stat.. in press. doi:10.1016/j.ecosta.2018.08.003.
Engle, R.F., Kelly, B., 2012. Dynamic equicorrelation. J. Bus. Econ. Stat. 30, 212–228.
Engle, R.F., Lediot, O., Wolf, M., 2019. Large dynamic covariance matrices. J. Bus. Econ. Stat. 37, 363–375.
Engle, R.F., Rangel, J.G., 2008. The spline GARCH model for unconditional volatility and its global macroeconomic causes. Rev. Financ. Stud. 21, 1187–1222.
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, 489–503.
Frahm, G., Memmel, C., 2010. Dominating estimators for minimum–variance portfolios. J. Econom. 159, 289–302.
Golosnoy, V., Gribisch, B., Liesenfeld, R., 2012. The conditional autoregressive Wishart model for multivariate stock market volatility. J. Econom. 167, 211–223.
Golosnoy, V., Gribisch, B., Seifert, M.I., 2019. Exponential smoothing of realized portfolio weights. J. Empir. Finance 53, 222–237.
Golosnoy, V., Schmid, W., Seifert, M.I., Lazariv, T., 2020. Statistical inferences for realized portfolio weights. Econom. Stat. 14, 49–62.
Gorgi, P., Hansen, P.R., Janus, P., Koopman, S.J., 2019. Realized Wishart-GARCH: a score-driven multi-asset volatility model. J. Financ. Econom. 17, 1–32.
Gribisch, B., Hartkopf, J.P., 2021. Modeling realized covariance measures with heterogeneous liquidity: a generalized matrix-variate wishart state-space Model. In: SSRN Working Paper.
Gribisch, B., Hartkopf, J.P., Liesenfeld, R., 2020. Factor state-space models for high--dimensional realized covariance matrices of asset returns. J. Empir. Finance 55, 1–20.
Hafner, C.M., Linton, O.B., 2010. Efficient estimation of a multivariate multiplicative volatility model. J. Econom. 159, 55–73.
Hafner, C.M., Linton, O.B., Tang, H., 2020. Estimation of a multiplicative correlation structure in the large dimensional case. J. Econom. 217, 431–470.
Hansen, P.R., Lunde, A., Nason, J.M., 2011. The model confidence set. Econometrica 79, 453–497.
Hautsch, N., Kyi, L., Malec, P., 2015. Do high-frequency data improve high-dimensional portfolio allocations? J. Appl. Econom. 30, 263–290.
Hautsch, N., Voigt, S., 2019. Large-scale portfolio allocation under transaction costs and model uncertainty. J. Econom. 212, 221–240.
Hayashi, F., 2000. Econometrics. Princeton University Press, New Jersey.
Jagannathan, R., Ma, T., 2003. Risk reduction in large portfolios: why imposing the wrong constraints helps. J. Finance 58, 1651–1683.
Janus, P., Koopman, S.J., Lucas, A., 2014. Long memory dynamics for multivariate dependence under heavy tails. J. Empir. Finance 29, 187–206.
Jin, X., Maheu, J.M., 2013. Modelling realized covariances and returns. J. Financ. Econom. 11, 335–369.
Klüppelberg, C., Seifert, M.I., 2019. Financial risk measures for a network of individual agents holding portfolios of light-tailed objects. Finance Stoch. 23 (4), 795–826.
Klüppelberg, C., Seifert, M.I., 2020. Explicit results on conditional distributions of generalized exponential mixtures. J. Appl. Probab. 57 (3), 760–774.
Kock, A.B., Callot, L., 2015. Oracle inequalities for high dimensional vector autoregressions. J. Econom. 186, 325–344.
Lam, C., Feng, P., 2018. A nonparametric eigenvalue-regularized integrated covariance matrix estimator for asset return data. J. Econom. 206, 226–257.
Ledoit, O., Wolf, M., 2012. Nonlinear shrinkage estimation of large-dimensional covariance matrices. Ann. Stat. 40, 1024–1060.
Lee, S.-W., Hansen, B.E., 1994. Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. Econom. Theory 10, 29–52.
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, 504–518.
Lütkepohl, H., 1996. Handbook of Matrices. Wiley, Chichester, UK.
Lutkepohl, H., 2005. New Introduction to Multiple Time Series Analysis. Springer, Berlin, Germany.
Merton, R.C., 1980. On estimating the expected return on the market: an exploratory investigation. J. Financ. Econ. 8, 323–361.
Moura, G.V., Santos, A.A.P., Ruiz, E., 2020. Comparing high-dimensional conditional covariance matrices: implications for portfolio selection. J. Bank. Finance 118, 105882.
Muirhead, R.J., 1982. Aspects of Multivariate Statistical Theory. Wiley, New Jersey.
Noureldin, D., Shephard, N., Sheppard, K., 2012. Multivariate high frequency based volatility (HEAVY) models. J. Appl. Econom. 27, 907–933.
Okhrin, Y., Schmid, W., 2006. Distributional properties of optimal portfolio weights. J. Econom. 134, 235–256.
Opschoor, A., Lucas, A., 2019. Fractional integration and fat tails for realized covariance kernels. J. Financ. Econom. 17 (1), 66–90.
Reh, L., Krüger, F., Liesenfeld, R., 2019. Dynamic modeling of the global minimum variance portfolio. Working paper.
Sheppard, K., Xu, W., 2019. Factor high-frequency-based volatility (HEAVY) models. J. Financ. Econom. 17, 33–65.
Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B 58, 267–288.
Tu, J., Zhou, G., 2011. Markowitz meets talmud: a combination of sophisticated and naive diversification strategies. J. Financ. Econ. 99, 204–215.
Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. Rev. Financ. Stud. 21, 1455–1508.
Wilms, I., Rombouts, J., Croux, C., 2021. Multivariate volatility forecasts for stock market indices. Int. J. Forecast. 37, 484–499.
Wooldridge, J., 1994. Estimation and inference for dependent processes. In: Engle, R., McFadden, D. (Eds.), Handbook of Econometrics, vol. IV, pp. 2639–2738.
Zhao, P., Yu, B., 2006. On model selection consistency of lasso. J. Mach. Learn. Res. 7, 2541–2563.