Journal of Banking and Finance 178 (2025) 107505

Contents lists available at ScienceDirect

Journal of Banking and Finance

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

Large dynamic covariance matrices and portfolio selection with a heterogeneous autoregressive model

Igor Honig, Felix Kircher *

Chair of Statistics and Risk Management, Faculty of Business, Economics and Management Information Systems, University of Regensburg, 93040 Regensburg, Germany

ARTICLE INFO

JEL classification:
C22
C51
C58
G11

Keywords:
Time-varying covariance matrix
High dimensions
Heterogeneous autoregressive model
Minimum variance portfolio
Markowitz portfolio optimization

ABSTRACT

We propose a novel framework for modeling large dynamic covariance matrices via heterogeneous autoregressive volatility and correlation components. Our model provides direct forecasts of monthly covariance matrices and is flexible, parsimonious and simple to estimate using standard least squares methods. We address the problem of parameter estimation risks by employing nonlinear shrinkage methods, making our framework applicable in high dimensions. We perform a comprehensive empirical out-of-sample analysis and find significant statistical and economic improvements over common benchmark models. For minimum variance portfolios with over a thousand stocks, the annualized portfolio standard deviation improves to 8.92% compared to 9.75–10.43% for DCC-type models.

1. Introduction

Multivariate volatility modeling and forecasting are central to risk and asset management, where the covariance matrix of returns is a critical input for optimal portfolio selection, dating back to Markowitz (1952). In practice, estimation of the covariance matrix poses two fundamental challenges. First, problems arise in high dimensions, i.e., when the number of assets is large relative to the sample size. Specifically, estimation risks emerge when parameters can only be estimated imprecisely from historical data, leading to unreliable portfolio weights and suboptimal out-of-sample portfolio performance. Second, since covariance matrices are typically estimated using daily returns, a temporal mismatch occurs in the context of asset management, where forecasts over longer horizons (e.g., one month) are more relevant. This requires multi-step forecasting and temporal aggregation techniques (see, e.g., De Nard et al., 2021), which increases forecasting complexity and reduces forecast accuracy if the model is misspecified (see, e.g., Ghysels et al., 2019). Despite progress on both ends, no current model simultaneously addresses the challenges of high dimensionality and the temporal mismatch. However, this is highly relevant for portfolio managers as it has the potential for significant performance gains for individual stock portfolios.

In this paper, we propose a new framework for modeling large dynamic covariance matrices for optimal portfolio selection. Our model produces direct forecasts of the covariance matrix at common portfolio rebalancing frequencies (e.g., monthly), while exploiting the information contained in daily returns. By employing shrinkage estimators of the covariance matrix, our model is applicable in settings with more than a thousand assets. In an empirical out-of-sample analysis, we show that our new framework yields statistically and economically significant improvements over current benchmark models for the construction of high-dimensional minimum variance portfolios.

Our framework is motivated by the popular heterogeneous autoregressive (HAR) model of Corsi (2009) and builds on insights from Bollerslev et al. (2018). We start by constructing realized volatility features based on different aggregation levels that reflect economically motivated short- and long-term horizons. We then specify future monthly volatilities as a linear function of these volatility features and jointly estimate their dynamics via least squares to exploit the commonalities in realized volatilities across financial assets (see Bollerslev et al., 2018). This approach yields a parsimonious component structure that approximates the apparent long memory properties of monthly volatilities. To enable separate dynamics in the correlations, we employ a similar component structure to model future monthly correlations based on new realized correlation features. In doing so, we obtain a versatile modeling framework for monthly covariance matrices based on daily returns that is computationally efficient and simple to implement.

* Corresponding author.
E-mail addresses: igor.honig@ur.de (I. Honig), felix.kircher@ur.de (F. Kircher).

https://doi.org/10.1016/j.jbankfin.2025.107505
Received 5 July 2023; Accepted 18 June 2025
Available online 2 July 2025

Our main contribution is twofold. The first contribution is our new model that we refer to as the "Multivariate Heterogeneous Exponential" (MHEx) model, extending the class of HExp volatility models by Bollerslev et al. (2018) to the multivariate case. Our approach to decomposing the return covariance matrix into (realized) volatilities and correlations draws inspiration from dynamic correlation models (as in Engle, 2002; Oh and Patton, 2016). A distinct feature of our model is its ability to produce smooth volatility and correlation dynamics. We achieve this by applying an exponential weighting scheme when constructing the realized volatility and correlation features. These enhancements offer increased flexibility and precision in estimating the covariance dynamics compared to the standard HAR approach. By directly modeling and forecasting monthly covariance matrices, we circumvent the problems associated with temporal mismatch and forecast aggregation that arise from daily models.

Our second contribution is to incorporate shrinkage estimators into our framework. Most applications of HAR models consider a small number of assets, say ten, where shrinkage is not essential. Shrinkage estimators of the covariance matrix, existing in linear (Ledoit and Wolf, 2003, 2004a,b) and nonlinear (Ledoit and Wolf, 2012, 2017, 2020, 2022) forms, have been devised to mitigate estimation risks. In general, this is achieved by reducing the variance of the sample covariance matrix estimator at the expense of introducing bias. We show how to apply the nonlinear shrinkage function of Ledoit and Wolf (2020) to our realized correlation features, i.e., the components that determine future monthly correlations. This enables our framework to handle high-dimensional settings, which is important for the selection of individual stock portfolios.

We demonstrate the effectiveness of our MHEx model in an empirical out-of-sample analysis. Our set of benchmark models includes: (a) the dynamic conditional correlation model with nonlinear shrinkage (DCC-NL) by Engle et al. (2019); (b) the dynamic approximate one-factor model (AFM-DCC-NL) by De Nard et al. (2021); (c) a variant of the realized DCC model by Bollerslev et al. (2020b) with nonlinear shrinkage (RDCC-NL); (d) a scalar BEKK model inspired by Engle and Kroner (1995) with nonlinear shrinkage (BEKK-NL); and variants of these models. We use five data sets, comprising predefined portfolios and individual stocks, with asset universes ranging from 10 to 1500. Academic studies commonly examine predefined portfolios, whereas portfolio selection involving large numbers of individual stocks is particularly relevant for asset managers, as noted by Ledoit and Wolf (2017). We construct minimum variance portfolios and show that the gains from using our MHEx model over a static covariance matrix model increase with the number of assets. Our MHEx model consistently outperforms all benchmark models, achieving statistical significance at the 1% level across all data sets. For example, for 100 portfolios based on size and book-to-market characteristics, the annualized standard deviation of minimum variance strategies improves to 8.62% using the MHEx model compared to 9.49% (9.27%) using the (R)DCC-NL model. The AFM-DCC-NL and BEKK-NL benchmark models perform similarly to the RDCC-NL model, with portfolio standard deviations of 9.21 and 9.25%, respectively. We also impose restrictions on the DCC-NL model parameters to jointly estimate the variance dynamics for all assets, which leads to improvements in out-of-sample performance. However, we find that the MHEx model consistently outperforms this restricted DCC-NL model.

Although the dynamic covariance matrix models offer superior forecasting performance, they tend to produce unstable and extreme portfolio weights for large stock portfolios. In particular, the unconstrained minimum variance portfolios exhibit high monthly leverage as well as turnover, which often exceed 200% and 100%, respectively. For enhanced practical applicability, we analyze minimum variance portfolios incorporating leverage and weight constraints. Under such portfolio constraints, the relative outperformance of the MHEx model over the (R)DCC-NL model increases consistently with the number of assets and is statistically significant at the 1% level for all but the smallest data set. Specifically, for portfolios comprising the S&P 1500 constituents, the standard deviation of a 130/30 minimum variance portfolio (with short positions limited to 30% of the initial capital) is reduced from 10.43% (9.75%) for the (R)DCC-NL model to 8.92% using the MHEx model. We find similar outperformance of the MHEx model relative to the AFM-DCC-NL, BEKK-NL, and restricted DCC-NL models. In summary, our MHEx model provides significant value in estimating dynamic covariance matrices, thereby facilitating minimum variance (or mean–variance) portfolio construction in high dimensions.

Our paper adds to the existing literature on modeling dynamic covariance matrices in high dimensions. In the context of dynamic conditional correlation models, De Nard et al. (2022) enhance the DCC-NL model by incorporating daily OHLC data. Another common model class are factor copulas (see, e.g., Creal and Tsay, 2015; Opschoor et al., 2021; Oh and Patton, 2023) and factor models (see, e.g., Bollerslev et al., 2019; Gribisch et al., 2020; De Nard et al., 2021; Alves et al., 2024), which build on common observable or latent systematic factors and represent an effective technique for dimension reduction. For instance, Alves et al. (2024) estimate a vectorized HAR model with the LASSO to improve the accuracy of covariance matrix forecasts and minimum variance portfolio estimates. However, few studies, such as De Nard et al. (2021), consider forecasts beyond the daily frequency and explicitly target the monthly forecast horizon, which is the central focus of our analysis. Our paper thus provides pertinent methods and insights for medium-term risk modeling and forecasting in large-scale portfolio settings.

The paper is organized as follows. Section 2 describes our modeling and estimation framework for large dynamic covariance matrices and monthly portfolio selection. Section 3 details our empirical approach and discusses the main results of our out-of-sample portfolio analyses. Section 4 concludes.

2. Modeling large dynamic covariance matrices

Specifying a covariance matrix model is crucial for determining optimal portfolio weights in practice. Since the true covariance matrix is not directly observable, we rely on realized measures constructed from higher frequency daily returns. The realized covariance matrix is a consistent estimator of the true covariance matrix for ever finer sampling intervals (see Barndorff-Nielsen and Shephard, 2004). However, two primary limitations prevent us from using high-frequency intraday data for the construction of lower frequency realized measures, which is common for large stocks or stock indices (see, e.g., Bollerslev et al., 2020a, 2022; Andersen et al., 2023). First, high-frequency intraday data are unlikely to be available for the large number of individual stocks contained in the broad market or segment indices considered in this paper. Second, high-frequency intraday data are not available for the long time periods of more than 40 or 70 years analyzed in our empirical application. For these reasons, we construct monthly realized measures using widely available long histories of daily asset returns. We draw on literature that utilizes similar monthly realized (co)variation measures based on daily returns for applications such as portfolio construction and asset pricing, see Bollerslev et al. (2022), Bekaert et al. (2022), or Moreira and Muir (2017).[1]

Throughout the paper, we denote by $t = 1, \dots, T$ the monthly time index, where T is the sample size of monthly observations of a given data set. The subscript $i = 1, \dots, N_t$ indexes assets in month t, with N_t indicating that the number of investable assets varies over time. Monthly log returns are represented by the vector $r_t = (r_{1,t}, \dots, r_{N_t,t})'$, and daily log returns by $r_{t,k} = (r_{1,t,k}, \dots, r_{N_t,t,k})'$ for day k in month t.

[1] In prior works, such measures of monthly realized (co)variances, dating back to Merton (1980), have been employed in French et al. (1987), Brailsford and Faff (1996), Jagannathan and Ma (2003), and Ang et al. (2009).

2.1. Heterogeneous autoregressive volatilities and correlations

We employ a decomposition of the true but latent $(N_t \times N_t)$ monthly covariance matrix of asset returns $\Sigma_t = D_t R_t D_t$, where D_t is the diagonal matrix of N_t return volatilities and R_t is the correlation matrix of asset returns. To build our model, we apply an analogous decomposition to the ex-post covariance matrix estimator for Σ_t, which we refer to as RCOV_t, i.e.,

$$\mathrm{RCOV}_t = \mathrm{Diag}\left(\mathrm{RV}_t\right) \mathrm{RCOR}_t \, \mathrm{Diag}\left(\mathrm{RV}_t\right), \tag{1}$$

where $\mathrm{Diag}\left(\mathrm{RV}_t\right)$ denotes the diagonal matrix of N_t monthly realized volatilities, and RCOR_t is the monthly realized correlation matrix.

In each month t, we have N_t investable assets with daily log returns $r_{t,k}$ collected over a total of K_t days. We compute the vector of monthly realized volatilities as the sum of daily squared log returns,

$$\mathrm{RV}_t = \sqrt{\sum_{k=1}^{K_t} r_{t,k}^2}, \tag{2}$$

i.e., we do not subtract the square of average daily returns. Instead, as is common in the finance literature, we assume that daily returns $r_{t,k}$ have an expected value of approximately zero and are uncorrelated.

In practice, the RV_t estimator will not be an error-free measure of the true monthly volatilities. Rather, the estimates obtained are subject to estimation errors that result from a natural upper bound on the values of K_t. Given these data limitations, we follow the literature (see, e.g., Barndorff-Nielsen and Shephard, 2002; Bollerslev et al., 2016) and similarly assume that the monthly RV_t^2 estimates can be expressed as the true monthly variance of each asset plus an estimation error. The consistency property of RV_t^2 for the true monthly variances, coupled with the assumption of serially uncorrelated estimation errors, provides a simple way to model the latent monthly variances via reduced form models for the observable realized variances. For example, assuming that each asset's true monthly variance evolves according to a stationary AR(1) process, then (under strong conditions) the dynamics of the RV_t^2 estimates allows for an ARMA(1,1) representation (see, e.g., Bollerslev et al., 2016). In an empirical application, Bekaert et al. (2025) show that an ARMA(1,1) model performs well for monthly idiosyncratic realized volatility forecasting with a large number of stocks. For our framework, we exploit the fact that the dynamics of an ARMA(1,1) model – or a more general ARMA(p, q) model – can be well approximated by a long-lag autoregressive model, for which we propose a modified and refined version as described below.

Specifically, for the realized volatility dynamics, we consider a linear process

$$\mathrm{RV}_t = \sum_{m \in \mathcal{M}_{\mathrm{RV}}} \phi_m \, \mathrm{ExpRV}_{t-1}^m + \varepsilon_t, \tag{3}$$

where ϕ_m, $m \in \mathcal{M}_{\mathrm{RV}}$, are scalar parameters, and ε_t is an error term. $\mathcal{M}_{\mathrm{RV}}$ denotes the set of different centers of mass for the exponential weights in the variables ExpRV_t^m as specified below. The model in (3) is formulated in terms of $|\mathcal{M}_{\mathrm{RV}}|$ exponentially-weighted realized volatility measures, which are constructed based on past squared daily returns,

$$\mathrm{ExpRV}_t^m = \sqrt{\overline{K} \sum_{l=0}^{L-1} \sum_{k=1}^{K_{t-l}} w_{t,j}^m \, r_{t-l,k}^2}, \tag{4}$$

where $j = \sum_{r=0}^{l} K_{t-r} - k$ indexes exponentially decreasing weights (for $j \to \infty$) for return observations in the past L months, including the current month. The weights in (4) are defined as

$$w_{t,j}^m = \frac{(1 + 1/m)^{-j}}{\sum_{k=0}^{\mathcal{K}_t - 1} (1 + 1/m)^{-k}}, \tag{5}$$

for $j = 0, \ldots, \mathcal{K}_t - 1$, where $\mathcal{K}_t = \sum_{l=0}^{L-1} K_{t-l}$ is the total number of days over the previous L months. The parameter m determines how quickly the exponential weights decay to zero. For large $\mathcal{K}_t$, m is equivalent to the center of mass defined as $\sum_{j=0}^{\mathcal{K}_t - 1} w_{t,j}^m \, j$ for weights $w_{t,j}^m$. Thus, the parameter m corresponds to the weighted average lag of days used to calculate ExpRV_t^m, reflecting the component's average time horizon.

For our empirical analysis, we choose common time horizons for the parameter m of the ExpRV_t^m variables, i.e., we specify $\mathcal{M}_{\mathrm{RV}} = \{1, 5, 20, 60, 120, 250, \infty\}$. The special case $m = \infty$ implies equal-weighting of the squared daily returns in (4), hence ExpRV_t^∞ is equivalent to the mean of squared daily returns over the past L months. We choose $L = 60$, i.e., five years, and use ExpRV_t^∞ as the long-run volatility component, similar to Bollerslev et al. (2018). In general, the choice of $\mathcal{M}_{\mathrm{RV}}$ depends on the research question at hand. We opt for a rather large number of components because we did not encounter problems with overfitting in our empirical analysis. To match the monthly horizon of the realized volatility RV_t modeled in (3), we apply a common scaling factor of $\overline{K} = 21$ in the definition of the ExpRV_t^m variables in (4), following the convention that there are on average $\overline{K} = 21$ trading days per month.

Our modeling approach for RV_t is inspired by the HExp model as considered in Bollerslev et al. (2018) for intraday realized volatilities. The model resembles the structure of the univariate HAR model of Corsi (2009), which relies on a small number of predetermined volatility components to conveniently restrict the infinite-dimensional space of coefficients of a general AR(∞) process for RV_t. Although the HAR model has emerged as a widely used framework in many empirical studies, it has been refined in Bollerslev et al. (2018) to better capture important properties of financial time series by modifying the construction of the realized volatility components. In the original HAR model, the components are sample averages of realized volatilities over various time horizons (e.g., a day, a week, a month etc.), which results in a step-like function for the lagged RV_t's. In contrast, using exponential weights as in (4) achieves smoothness in the lag function, with more recent daily observations having a greater impact on future volatility. Because only the construction of the model components is affected, and not the model structure itself, this approach maintains the benefits of the original HAR model. That is, our model in (3) can capture the long memory behavior of return volatilities. Our approach shares characteristics with the MIDAS approach using exponential Almon lags, as in Ghysels et al. (2019). Specifically, both forecast monthly variances (volatilities) directly based on exponentially weighting lagged daily realized variances (squared returns). However, the MIDAS approach requires estimating the parameters of the weighting function, whereas we use a set of predetermined volatility components, rendering our model highly flexible and simpler to estimate. For multivariate modeling of realized volatilities, Bollerslev et al. (2018) find that volatilities of different financial assets behave similarly over time, so that imposing common model parameters yields out-of-sample forecasting improvements. Because this could be confirmed in our empirical application, we specify the model as in (3) with common parameters $\phi_m, m \in \mathcal{M}_{\mathrm{RV}}$, for all assets.

We extend the above framework to the modeling of the dynamic realized correlation matrix RCOR_t. Analogous to the computation of realized volatilities in (2), we obtain the realized covariance matrix for month t from daily returns as

$$\mathrm{RCOV}_t = \sum_{k=1}^{K_t} r_{t,k} \, r'_{t,k}. \tag{6}$$

Based on (1), we compute the monthly realized correlation matrix

$$\mathrm{RCOR}_t = \mathrm{Diag}\left(\mathrm{RV}_t\right)^{-1} \mathrm{RCOV}_t \, \mathrm{Diag}\left(\mathrm{RV}_t\right)^{-1}, \tag{7}$$

using (2) and (6). Due to the symmetry of the correlation matrix, only the dynamics of its lower triangular elements must be specified. We assume a linear process that is driven by lagged exponentially-weighted realized correlation components, i.e.,

$$\psi\left(\mathrm{RCOR}_t\right) = \sum_{m \in \mathcal{M}_{\mathrm{RCOR}}} \gamma_m \, \psi\left(\mathrm{ExpRCOR}_{t-1}^m\right) + \nu_t, \tag{8}$$

where γ_m, $m \in \mathcal{M}_{\text{RCOR}}$, are scalar parameters, and v_t is an error term. $\mathcal{M}_{\text{RCOR}}$ is the set of centers of mass for the realized correlation components, and $\psi(\cdot)$ denotes the operator stacking all elements below the main diagonal of the correlation matrix into an $N_t(N_t - 1)/2$-dimensional vector. The regressors in (8) are defined as

$$\text{ExpRCOR}_t^m = \text{Diag}\left(\text{ExpRV}_t^m\right)^{-1} \text{ExpRCOV}_t^m \ \text{Diag}\left(\text{ExpRV}_t^m\right)^{-1} \quad (9)$$

with ExpRV_t^m as in (4) and

$$\text{ExpRCOV}_t^m = \overline{K} \sum_{l=0}^{L-1} \sum_{k=1}^{K_{t-l}} w_{t,j}^m \, r_{t-l,k} \, r'_{t-l,k}, \quad (10)$$

where $w_{t,j}^m$ is given in (5).

For the correlation model, we specify $\mathcal{M}_{\text{RCOR}} = \{10, 20, 60, 120, 250, \infty\}$, leaving out smaller values for m because they lead to rather noisy behavior of the corresponding ExpRCOV_t^m components with $m < 10$. Similar to the ExpRV_t^m variables in (4), we scale the ExpRCOV_t^m variables in (10) by $\overline{K} = 21$ to match the monthly horizon of RCOV_t defined in (6).

By construction, the model assumes the same dynamics for all entries in the realized correlation matrix. This reduces the number of parameters to be estimated and makes our model (8) less prone to overfitting, similar to the DCC model (discussed in Section 2.4). Moreover, this guarantees positive definiteness of the correlation matrix for suitable restrictions on γ_m. Naturally, the correlation model (8) shares the main features of the volatility model (3). The use of exponentially-weighted realized correlation components provides a parsimonious, flexible, and smooth specification of the dependence structure of realized correlations upon their own past. The regression coefficients γ_m determine the relative influence, or weights, of the short- and long-term realized correlation components used to approximate the apparent long memory behavior of stock return correlations. We refer to the framework presented above as the "Multivariate Heterogeneous Exponential", or MHEx, model. Its name is derived from the HExp volatility model in Bollerslev et al. (2018) and the exponential weighting in the definition of the realized volatility and correlation components.

We identify three key advantages of our proposed framework. First, because the focus is on portfolio optimization, we require forecasts of the covariance matrix for the next month. We obtain these forecasts directly using (3), (8), and (1), thereby avoiding the need for a forecast aggregation scheme based on recursively obtained daily covariance matrix estimates (as applied in De Nard et al., 2021). Second, we achieve high flexibility by modeling the dynamics of realized volatilities based on multiple exponentially-weighted components covering a wide range of short- and long-term realized volatility estimates. The same applies to the dynamic modeling of realized correlations, including the use of shrinkage estimators discussed in the next section to counteract the curse of dimensionality. Third, in particular for the realized correlation model, the use of predetermined components provides the advantage of simple implementation and computationally fast estimation. This is because the corresponding regression coefficients γ_m in (8) can be obtained with a suitable least squares estimator.

Remark 2.1. Our modeling framework is specifically formulated from the perspective of an investor who updates portfolio weights at a monthly frequency based on daily data. In fact, given (3) and (8), our methodology can be characterized as a mixed-frequency approach, since monthly realized volatilities and correlations are determined by variables measured at a higher (daily) frequency. Consequently, while our framework can be easily adapted to model covariance matrices at lower frequencies (e.g., quarterly, semi-annually, etc.), it cannot be used in exactly the same way for daily portfolio choice, which could be another relevant practical application. For such an application, our modeling framework needs to be modified to accommodate high-frequency intraday data. In practice, such data may not always be available for larger asset universes and, in particular, over long time periods as studied in this paper.

2.2. Incorporating shrinkage estimators

It is well-known that in high dimensions – or for comparatively large N/T – the sample estimator of the covariance matrix, or the correlation matrix, is inadequate. This is because the number of free parameters in the matrix grows with N^2, allowing for reliable estimates only when T is an order of magnitude larger than N. Estimation errors become more severe for smaller T, however, they can be mitigated by adding structure to the covariance matrix. The attempt to find an optimal balance between bias and variance of a covariance matrix estimator has led to shrinkage estimators. Linear shrinkage estimators were introduced in Ledoit and Wolf (2004b) and extended to nonlinear shrinkage estimators, as in Ledoit and Wolf (2017) in the context of portfolio selection. Against this background, our goal is to make the MHEx model from Section 2.1 applicable to large portfolios. To do so, we propose a method to perform shrinkage on the exponentially-weighted (realized) correlation matrices.

Shrinkage estimators of the sample covariance matrix are developed under the assumption that observations are equally-weighted. We establish the connection between exponential and equal weighting of the data through the notion of effective sample size. Let $E_1 = \sum_{j=0}^{\mathcal{K}-1} w_j X_j$ be an estimator based on exponential weighting with weights w_j, and $E_2 = \frac{1}{\mathcal{J}} \sum_{j=0}^{\mathcal{J}-1} X_j$ be an estimator based on equal weighting, where X_j is a stationary series of serially uncorrelated random variables, which could, e.g., be the product of standardized returns as for a correlation measure. Shrinkage is designed to reduce estimation risk by shrinking the variance of an estimator in exchange for increased bias. Therefore, analyzing the unconditional variances of E_1 and E_2, we get that $\mathbb{V}\text{ar}\left(E_1\right) = \sigma_X^2 \sum_{j=0}^{\mathcal{K}-1} w_j^2 = \sigma_X^2/\mathcal{J} = \mathbb{V}\text{ar}\left(E_2\right)$ when $\mathcal{J} = 1/\sum_{j=0}^{\mathcal{K}-1} w_j^2$, where $\sigma_X^2 = \mathbb{V}\text{ar}\left(X_j\right)$ and $\mathbb{V}\text{ar}$ denotes the unconditional variance. The term $1/\sum_{j=0}^{\mathcal{K}-1} w_j^2$ represents the effective sample size for weights w_j, which gives the sample size of an equally-weighted statistic with the same variance or estimation risk as the w_j-weighted statistic. For the exponential weights in (5) we have that

$$\frac{1}{\sum_{j=0}^{\mathcal{K}_t-1} \left(w_{t,j}^m\right)^2} \to 2\,m+1 \text{ for } \mathcal{K}_t \to \infty.$$

Therefore, to apply nonlinear shrinkage to ExpRCOR_t^m in (9), we adopt the nonlinear shrinkage function from the sample correlation matrix based on the past $2m+1$ (equally-weighted) observations, which also imply a center of mass of m. The estimation risk and, as further shown below, the structure of the two correlation matrices based on the same data are comparable, indicating that the optimal shrinkage intensity is similar and the nonlinear shrinkage function can be seamlessly transferred to ExpRCOR_t^m.

In our framework, we use the analytical nonlinear shrinkage estimator of Ledoit and Wolf (2020), because it is computationally fast and performs well for both small and large samples. The main inputs for the Ledoit and Wolf (2020) nonlinear shrinkage estimator are the sample size relative to the matrix dimension and the vector of eigenvalues. The nonlinear shrinkage function transforms the sample eigenvalues based on a nonlinear function such that large eigenvalues tend to decrease while small eigenvalues increase, see Ledoit and Wolf (2017, 2020). Therefore, provided that the eigenvalue structure of both covariance matrices is reasonably similar, we can readily apply the eigenvalue transformation implied by the nonlinear shrinkage estimator of the sample correlation matrix based on $2m+1$ observations to ExpRCOR_t^m.

Fig. 1 illustrates this relationship. The left plot shows the 150 largest eigenvalues of the sample correlation matrix based on 121 observations for 492 firms in the S&P 500 index in 12/2022, with available data over the past five years, along with the ExpRCOR_t^m estimator with $m = 60$. The eigenvalue structure is noticeably similar and the largest eigenvalues are almost identical. Only the smaller eigenvalues of ExpRCOR_t^m decay much slower because we use a five-year estimation window, which means that observations older than 121

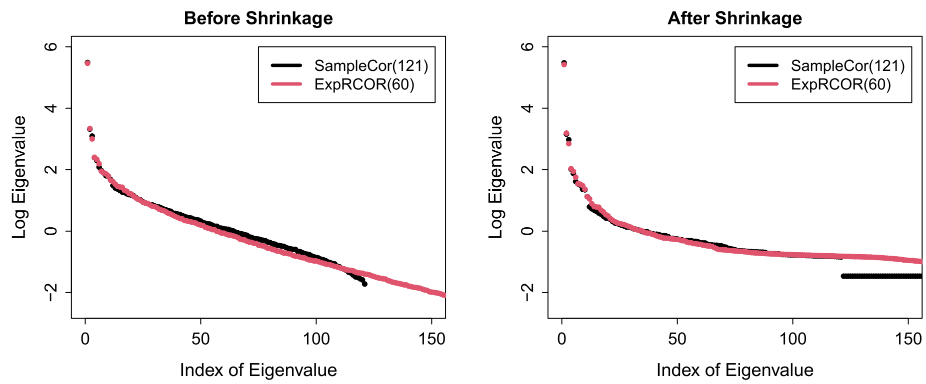

Fig. 1. Similarity of eigenvalues before and after shrinkage. Left: log eigenvalues of the sample correlation matrix based on 121 observations and of ExpRCOR_t^m with $m = 60$. Right: log eigenvalues of the nonlinearly shrunk sample correlation matrix based on 121 observations and of ExpRCOR_t^m with $m = 60$ after applying the nonlinear shrinkage function.

days are not discarded completely, as in the SampleCor(121) estimator, but receive successively less weight. The right plot shows the largest eigenvalues of both correlation matrices after nonlinear shrinkage. We apply the nonlinear estimator to standardized returns to shrink the sample correlation matrix based on $2m + 1 = 121$ observations. We linearly interpolate, as well as extrapolate for the largest eigenvalues, the estimated nonlinear shrinkage function, which maps the original eigenvalues of the sample correlation matrix to the eigenvalues of the nonlinearly shrunk sample correlation matrix. We then use this nonlinear shrinkage function to shrink ExpRCOR_t^m. The result, as shown in Fig. 1, again reveals a similar eigenvalue structure compared to the shrunk sample correlation matrix and a slower decay for the small eigenvalues of the shrunk ExpRCOR_t^m matrix. We note that the eigenvectors of ExpRCOR_t^m remain unchanged, as is the case in the nonlinear shrinkage approach of Ledoit and Wolf (2017, 2020). We employ this shrinkage methodology for all ExpRCOR_t^m, as defined in (9), in our practical applications of the MHEx model.

2.3. Parameter estimation

In the following, we address the parameter estimation of our proposed MHEx model. For each month t, the volatility model (3) is estimated via least squares with the following constraints on the regression coefficients, i.e., the optimization problem for $\boldsymbol{\phi} = (\phi_m)_{m \in \mathcal{M}_{\text{RV}}}$ is

$$\min_{\boldsymbol{\phi}} \sum_{s=t-T_{\text{E}}}^{t-1} \left(\text{RV}_s - \sum_{m \in \mathcal{M}_{\text{RV}}} \phi_m \, \text{ExpRV}_{s-1}^m \right)' \left(\text{RV}_s - \sum_{m \in \mathcal{M}_{\text{RV}}} \phi_m \, \text{ExpRV}_{s-1}^m \right) \tag{11}$$

$$\text{subject to} \sum_{m \in \mathcal{M}_{\text{RV}}} \phi_m = 1, \quad \phi_m \geq 0 \ \forall m \in \mathcal{M}_{\text{RV}},$$

where T_{E} denotes the sample size used to estimate the MHEx model. While the ExpRV_t^m variables are all non-negative, we must also restrict the coefficients ϕ_m to be non-negative to ensure weakly positive estimates of volatilities. The sum-to-one constraint for ϕ_m, $m \in \mathcal{M}_{\text{RV}}$, facilitates an interpretation in terms of forecast averaging, where the ExpRV_t^m can be viewed as different individual forecasts of future volatility combined by weighted averaging.

We then estimate the correlation model (8) via least squares with the following constraints on the regression coefficients, i.e., the optimization problem for $\boldsymbol{\gamma} = (\gamma_m)_{m \in \mathcal{M}_{\text{RCOR}}}$ is

$$\min_{\boldsymbol{\gamma}} \sum_{s=t-T_{\text{E}}}^{t-1} \left(\psi\left(\text{RCOR}_s\right) - \sum_{m \in \mathcal{M}_{\text{RCOR}}} \gamma_m \, \psi\left(\text{ExpRCOR}_{s-1}^m\right) \right)' \left(\psi\left(\text{RCOR}_s\right) - \sum_{m \in \mathcal{M}_{\text{RCOR}}} \gamma_m \, \psi\left(\text{ExpRCOR}_{s-1}^m\right) \right) \tag{12}$$

$$\text{subject to} \sum_{m \in \mathcal{M}_{\text{RCOR}}} \gamma_m = 1, \quad \gamma_m \geq 0 \ \forall m \in \mathcal{M}_{\text{RCOR}}.$$

We estimate the model with a non-negativity and sum-to-one constraint on the parameters γ_m.[2] In (12) and (13), ExpRCOR_t^m refers to the exponentially-weighted realized correlation matrix (9) after application of the nonlinear shrinkage methodology described in Section 2.2. The number of observations in $\psi\left(\text{ExpRCOR}_t^m\right)$ grows rapidly in high dimensions (i.e., for large N_t), resulting in a huge sample size for the least squares estimator (12). For numbers of assets in the hundreds we therefore replace the ψ operator with a $\widetilde{\psi}$ operator, stacking only the elements of the first off-diagonal into an $N_t - 1$ dimensional vector. The idea behind this is similar to the composite likelihood method, as used in Engle et al. (2019) and discussed in Section 2.4. In our application, the differences in the coefficient estimates for γ_m using $\widetilde{\psi}$ instead of ψ in (12) are marginal, but the runtime of the analysis improves considerably.

Given the estimated coefficients $\widehat{\phi}_m, m \in \mathcal{M}_{\text{RV}}$, and $\widehat{\gamma}_m, m \in \mathcal{M}_{\text{RCOR}}$, we predict monthly volatilities $\widehat{D}_t$, correlation matrices $\widehat{R}_t$, and covariance matrices $\widehat{\Sigma}_t$ as

$$\begin{aligned} \widehat{D}_t &\equiv \text{Diag}\left(\widehat{\text{RV}}_t\right) \quad \text{where} \quad \widehat{\text{RV}}_t = \sum_{m \in \mathcal{M}_{\text{RV}}} \widehat{\phi}_m \, \text{ExpRV}_{t-1}^m, \\ \widehat{R}_t &\equiv \widehat{\text{RCOR}}_t = \sum_{m \in \mathcal{M}_{\text{RCOR}}} \widehat{\gamma}_m \, \text{ExpRCOR}_{t-1}^m \quad \text{and} \\ \widehat{\Sigma}_t &\equiv \widehat{D}_t \widehat{R}_t \widehat{D}_t. \end{aligned} \tag{13}$$

The estimated correlation matrix $\widehat{R}_t$ and the covariance matrix $\widehat{\Sigma}_t$ are guaranteed to be valid covariance matrices because, given the constraints on the coefficients $\widehat{\gamma}_m$, $\widehat{\text{RCOR}}_t$ is a convex combination of positive definite matrices.

2.4. Comparison with alternative covariance matrix models

We now discuss two competing covariance matrix models considered in our empirical analysis and highlight their differences with our proposed framework.

2.4.1. The DCC-NL model

The DCC-NL model as proposed by Engle et al. (2019), and used in De Nard et al. (2021, 2022), is estimated in two steps via (quasi)

[2] We have also considered the sparsity-encouraging LASSO, allowing the sum of coefficients γ_m to be less than one. The intuition here is to shrink the estimated correlations towards zero, which is a common approach to address estimation risks in large covariance matrices, similar to the linear shrinkage approach in Ledoit and Wolf (2004b). As this did not yield superior statistical and economic performance compared to the optimization problem formulated in (12), these results are not reported.

maximum likelihood by assuming a multivariate normal distribution for the vector of daily asset returns. In the first step, the asset-specific daily conditional variances are estimated using a GARCH(1,1) model,

$$\sigma^2_{i,t,k} = \omega_i + \alpha_i\, r^2_{i,t,k-1} + \beta_i\, \sigma^2_{i,t,k-1}, \quad i = 1, \ldots, N_t, \tag{14}$$

with GARCH parameters $(\omega_i, \alpha_i, \beta_i)$, where we assume zero means for the daily log returns $r_{t,k}$.[3] In the second step, the dynamic conditional correlation matrix is estimated using correlation targeting as in Engle et al. (2019). That is, the DCC parameters (C, α, β) in

$$Q_{t,k} = (1 - \alpha - \beta)\, C + \alpha\, s_{t,k-1} s'_{t,k-1} + \beta\, Q_{t,k-1},$$
$$R_{t,k} = \text{Diag}\left(Q_{t,k}\right)^{-1/2} Q_{t,k}\ \text{Diag}\left(Q_{t,k}\right)^{-1/2}$$

are estimated by first replacing the unconditional correlation matrix C by a nonlinear shrinkage estimate of the covariance matrix of the estimated standardized returns $\hat{s}_{t,k} = r_{t,k}/\hat{\sigma}_{t,k}$.

In large dimensions, Engle et al. (2019) employ the composite likelihood method (Pakel et al., 2021), which circumvents the time-consuming inversion of the conditional correlation matrix $R_{t,k}$ by assembling the likelihood for the full model from the likelihoods of adjacent pairs of assets only. Note that we take a similar approach when estimating the MHEx correlation model as described in Section 2.3, where only the correlations of adjacent pairs are considered. When estimating the DCC-NL model, we use the nonlinear shrinkage estimator from Ledoit and Wolf (2020) for $\hat{C}$, and the composite likelihood whenever $N_t > 100$.

To obtain monthly covariance matrix forecasts, we follow De Nard et al. (2021, 2022) and make multi-step forward predictions for each day k within the next month, assuming $\overline{K} = 21$ days within a month. The covariance matrix forecast for the whole month t is then derived by aggregating the daily forecasts as in $\hat{\Sigma}_t = \sum_{k=1}^{\overline{K}} \hat{\Sigma}_{t,k}$ with $\hat{\Sigma}_{t,k} = \hat{D}_{t,k}\, \hat{R}_{t,k}\, \hat{D}_{t,k}$. Here, $\hat{D}_{t,k}$ is the diagonal matrix of multi-step variance forecasts, and $\hat{R}_{t,k}$ the multi-step correlation matrix forecast for day k.

To highlight the increased flexibility of our approach compared to the DCC-NL model, consider the first day of month t. The GARCH(1,1) model allows the ARCH(∞) representation

$$\sigma^2_{i,t,1} = \frac{\omega_i}{1 - \beta_i} + \alpha_i \sum_{l=1}^{\infty} \sum_{s=1}^{K_{t-l}} \beta_i^j\, r^2_{i,t-l,s}, \quad i = 1, \ldots, N_t,$$

with $j = \sum_{r=1}^{l} K_{t-r} - s$, for $\omega_i, \alpha_i, \beta_i \geq 0$ and $\alpha_i + \beta_i < 1$. This equation can be interpreted as the long-run variance plus an exponentially-weighted realized variance component whose smoothing parameter or decay rate β_i is asset-specific and estimated along with the other model parameters. This is in contrast to our MHEx volatility model from Section 2.1, which contains multiple exponentially-weighted components with different pre-specified decay rates. Provided that these pre-specified parameters cover the space reasonably well, we gain in flexibility, allowing us to model the higher impact of more recent returns as well as the apparent long memory properties of realized volatilities. The GARCH(1,1) model has to compromise in this regard with only a single smoothing parameter. To some extent, this can be mitigated by considering higher order or multi-component GARCH models. However, model selection and estimation then becomes increasingly cumbersome. The same reasoning applies to the DCC model for dynamic correlations compared to our MHEx correlation model from Section 2.1.

2.4.2. The RDCC-NL model

While the DCC-NL model is specified on daily returns, and monthly forecasts are derived by aggregating daily forecasts over the next month, a DCC-type model estimated directly on monthly returns could be preferable given our focus on monthly portfolio selection. DCC-type models, which include high-frequency data to improve their forecasting ability, are often called Realized DCC (RDCC) models, as in Bollerslev et al. (2020b) and Bauwens and Xu (2023). We define a RDCC-NL model analogously to Bollerslev et al. (2020b), however, we specify the model on monthly returns, use daily data within the model (instead of intraday data), and apply a nonlinear shrinkage estimator for the correlation matrix. The RDCC-NL model yields direct monthly covariance matrix forecasts based on daily data, thus allowing a more direct comparison with our MHEx model from Section 2.1. The RDCC-NL model is defined by the following equations:

$$\sigma^2_t = \left(1 - \alpha_v - \beta_v\right) V + \alpha_v\, \text{RV}^2_{t-1} + \beta_v\, \sigma^2_{t-1}, \tag{15}$$
$$Q_t = (1 - \alpha_q - \beta_q)\, C + \alpha_q\, \text{RCOR}_{t-1} + \beta_q\, Q_{t-1},$$
$$R_t = \text{Diag}\left(Q_t\right)^{-1/2} Q_t\ \text{Diag}\left(Q_t\right)^{-1/2},$$

where RV_t and RCOR_t are defined as in (2) and (7) in Section 2.1, V denotes the vector of long-run variances of the considered assets, C is the long-run correlation matrix, and σ^2_t and R_t denote the vector of monthly conditional variances and the conditional correlation matrix, i.e., the conditional variances and correlation matrix of (simple) asset returns for month t.

In the RDCC-NL model specification, the lagged monthly realized volatilities RV_{t-1} (based on daily returns within month $t-1$) replace the lagged monthly returns that would be used in a DCC model specified on monthly data. Similarly, the lagged monthly realized correlation matrix RCOR_{t-1} replaces the outer product of lagged standardized monthly returns in the definition of the correlation dynamics. As a result, the RDCC-NL model uses the information from within-month returns more efficiently than a DCC model applied to monthly data, while allowing direct predictions of monthly covariance matrices. With the RDCC-NL model specification in (15), the unconditional variances and correlation matrix of monthly returns are essentially given by V and C, assuming that the unconditional levels of RV^2_t and RCOR_t are approximately equal to V and C. Because the specification of the model on monthly returns implies that fewer observations are available for model estimation compared to the DCC-NL model estimated on daily returns, we specify α_q and β_q as common parameters for all assets in the variance equation in (15).

We estimate the RDCC-NL model similarly to the DCC-NL model via a two-step (quasi) maximum likelihood approach with variance and correlation targeting for V and C. That is, we estimate V as the mean of squared daily returns over the estimation period (scaled by $\overline{K} = 21$ to conform to a monthly horizon), and we estimate C as the correlation matrix of daily returns via the nonlinear shrinkage estimator of Ledoit and Wolf (2020), which is applied to daily returns scaled to unit variance. In the first step, with V and C fixed, the common parameters α_v and β_v of the variance equation are estimated by maximizing the mean log likelihood of demeaned monthly (simple) returns of all assets. In the second step, the parameters α_q and β_q of the correlation equation are estimated based on the devolatized (via the estimated volatilities from the previous step) monthly (simple) returns.[4] In large dimensions, such as for single stock data sets with $N_t > 100$, we use the composite likelihood method for parameter estimation of the RDCC-NL model, as described above for the DCC-NL model.

[3] For the first day of each month, when $k = 1$, the right-hand side (in slight abuse of notation) refers to the last day of the previous month.

[4] In our empirical application, we have tested different specifications of the RDCC-NL model, i.e., additionally including the lagged (standardized) monthly returns in the specifications in (15), explicitly subtracting the mean of RCOR_{t-1} in the intercept specification, such as in $\left(1 - \beta_q\right) C - \alpha_q \overline{\text{RCOR}}$, estimating V and C via daily or monthly returns, and applying nonlinear shrinkage to RCOR_{t-1}. The specification described in this section proved to be the most effective.

Table 1
Data sets used in the empirical analysis, with full time period and start of the out-of-sample (OOS) test period.

#	Data set	Abbreviation	Time period	Start OOS period
1	10 US industry portfolios	FF10	01/1950–12/2022	01/1970
2	30 US industry portfolios	FF30	01/1950–12/2022	01/1970
3	100 US Size-BM portfolios	FF100	01/1950–12/2022	01/1970
4	S&P 500 index constituents	SP500	01/1980–12/2022	01/1995
5	S&P 1500 index constituents	SP1500	01/1980–12/2022	01/1995

2.5. Portfolio optimization

In our empirical analysis, we focus on minimum variance portfolio optimization. As minimum variance portfolios require only the covariance matrix as an input, they provide a natural testing ground for covariance matrix forecasts, where the resulting portfolio strategies can then be evaluated based on their realized out-of-sample variances. The weights of the population minimum variance portfolio are given by

$$w_{\mathrm{MINV},t} = \frac{\Sigma_t^{-1} 1_{N_t}}{1'_{N_t} \Sigma_t^{-1} 1_{N_t}}, \tag{16}$$

where 1_{N_t} denotes an $N_t \times 1$ vector of ones. The population covariance matrix Σ_t must be replaced by an estimate $\hat{\Sigma}_t$ to obtain feasible portfolio weights $\hat{w}_{\mathrm{MINV},t}$ in practice.

When implementing (16), $\hat{w}_{\mathrm{MINV},t}$ can take large positive and negative values. This is often undesirable because it leads to a high concentration in certain assets. Therefore, portfolio optimization is usually implemented with constraints on the resulting portfolio weights. The minimum variance portfolio optimization problem with a leverage constraint and bounds on portfolio weights can only be solved numerically and is formulated as

$$\min_w w' \Sigma_t w \tag{17}$$

$$\text{subject to} \quad w' 1_{N_t} = 1, \ \sum_{i=1}^{N_t} |w_i| \le \ell \text{ and } w_i \in \left[w_{\min}, w_{\max}\right], \quad i = 1, \dots, N_t,$$

where ℓ is the maximum gross leverage allowed, and $w_{\min}$ and $w_{\max}$ are the lower and upper bounds on portfolio weights, respectively.

Remark 2.2. The MHEx and DCC-NL approaches provide predictions for covariance matrices of monthly log returns, because they rely on daily log returns summing to monthly log returns. However, for portfolio optimization as described in this section, Σ_t represents the covariance matrix of simple returns. For our practical application, the predicted covariance matrix of log returns provides a suitable approximation for the covariance matrix of simple returns, since the differences are negligible relative to forecasting errors and the typical differences between the covariances of different equities.[5] Thus, we use the predictions of the MHEx and DCC-NL approaches defined in the preceding sections for $\hat{\Sigma}_t$.

3. Empirical analysis

Our empirical analysis is designed as a (pseudo) out-of-sample test of the covariance matrix models described in Section 2 (and its variants described in Section 3.3). Because our practical application focuses on portfolio optimization, the primary objective of this analysis is to compare the out-of-sample performance of estimated minimum variance portfolios. We perform backtesting exercises for five data sets with an increasing number of assets ranging from ten to over a thousand to evaluate the usefulness of the MHEx model for a wide range of potential applications.

3.1. Data sets

Our empirical analysis includes five data sets with corresponding time periods given in Table 1. The first three (FF) portfolio data sets are available on Kenneth R. French's website, and the last two stock market index-based (SP) data sets are from Refinitiv Datastream. For each data set, we obtain data on daily total returns. As given in Table 1, we conduct the out-of-sample analysis (i) for applications to predefined portfolios, which implies a smaller number of assets to be modeled, such as ten or 30 industry portfolios or 100 stock characteristics-based portfolios, and (ii) for applications to single stocks, where the number of individual assets can be very large, such as the constituents of the S&P 500 and S&P 1500 indices.

The FF100 data set (see Table 1) contains a small number of missing values at the beginning of the sample. We therefore replace each missing value with the average value of the neighboring portfolios. End-of-month constituent lists for the SP data sets are available starting in 09/1989 for the S&P 500 and in 12/1994 for the S&P 1500. For the months prior to that, we use the first available constituent list. This only affects the estimation periods, where the first constituent list is always included, while the out-of-sample test periods begin after the first list is available.

We perform portfolio optimization using only stocks that are index constituents, i.e., the portfolio built for month t consists of the index constituents at the end of month $t-1$. This approach avoids a look-ahead bias and excludes micro-cap and illiquid stocks from investment. We require the stocks to have a complete daily return history for the past 60 months in order to make predictions with the MHEx model and to estimate the benchmark models given in Section 3.3. Due to this restriction, the number of stocks in the portfolio is generally less than 500 and 1500 for the SP500 and SP1500 data sets, respectively. Throughout the out-of-sample period, our SP500 portfolio contains an average of 480 stocks, and our SP1500 portfolio contains an average of 1360 stocks. While initially there are fewer stocks available with the required data history, the number of stocks within the last 20 years of the out-of-sample period is close to the total number of index constituents.

3.2. Empirical approach for parameter estimation and portfolio strategies

The (pseudo) out-of-sample analysis is conducted as follows. Data sets are divided into an initial estimation period and an out-of-sample test period. For the FF data sets starting in 1950, the out-of-sample test period starts in 1970 and ends in 2022, for the SP data sets, the out-of-sample test period starts in 1995 and ends in 2022 (see Table 1). Given the 5 years required to calculate the variables for the MHEx model, initial estimation periods are 15 and 10 years, respectively. We extend the estimation window each month until the window reaches its final size of $T_{\mathrm{E}} = 360$ months, i.e., 30 years.[6] We estimate the MHEx model

[5] For example, assuming a multivariate normal distribution for monthly log returns with expected returns of 0.005 and standard deviations of 0.05 (i.e., 6% and 17% per annum) and a correlation of 0.5, the relative difference between the simple return covariance and the log return covariance is 1.4% (based on the properties of the lognormal distribution). As this affects all assets to some extent, it has virtually no effect on the estimated portfolio weights.

[6] The empirical results are robust to the length of the estimation window. A further expanding window yields virtually identical results (or marginally improves the performance of the MHEx model), but leads to slightly longer run times for estimation.

from Section 2.1 before the start of each month of the out-of-sample test period as described in Section 2.3. We predict volatilities and correlations and compute the covariance matrix as given in (13). The predicted covariance matrix is used as an input for portfolio optimization. We determine minimum variance portfolios without constraints on portfolio weights as in (16). In practice, it may be desirable to limit the size of individual portfolio positions or to impose restrictions on short sales. Therefore, we also perform the analysis for constrained minimum variance portfolios as given in (17).

3.3. Benchmark models

In the empirical analysis, we compare the MHEx approach for the estimation of large dynamic covariance matrices with other state-of-the-art methods[7]:

- **NL** is a static covariance matrix estimator using the nonlinear covariance matrix shrinkage method from Ledoit and Wolf (2020) and a fixed estimation window as in Ledoit and Wolf (2017). We scale the static covariance matrix estimated from five years of daily returns (as for the DCC-NL model) by $\overline{K} = 21$ to obtain a prediction for the next month.
- **DCC-NL** is the well-established DCC model combined with a nonlinear shrinkage estimator for the unconditional covariance matrix, as introduced in Engle et al. (2019). Following Engle et al. (2019) and De Nard et al. (2021, 2022), we use the past five years of daily data to estimate the DCC-NL covariance model. We obtain a prediction for the next month from the DCC-NL model by aggregating the DCC-NL covariance matrix forecasts over $\overline{K} = 21$ days, following De Nard et al. (2021, 2022).
- **JDCC-NL** is similar to the DCC-NL model, but more restrictive with respect to the time-varying return variances which are now modeled jointly rather than on an asset-by-asset basis. The intuition is to impose similar constraints on the model for the conditional daily return variances as we do in our MHEx model for monthly realized volatilities. Formally, the JDCC-NL model imposes additional constraints on the two GARCH(1,1) parameters (α_i, β_i) by setting $\alpha_i = \alpha$ and $\beta_i = \beta$ $\forall i = 1, \dots, N_t$, letting only the intercept ω_i to be asset-specific, see (14).
- **AFM-DCC-NL** is the approximate single-factor (or AFM1-DCC-NL) model of De Nard et al. (2021) that imposes cross-sectional constraints on the (co)variances via a factor structure for the vector of asset returns using the Fama–French market factor, and then applies the DCC-NL model for the dynamic covariance matrix of the *residuals* from the factor model. Similar to the DCC-NL model, we use the past five years of daily data to estimate the AFM-DCC-NL covariance matrix model and obtain a prediction for the next month by aggregating the AFM-DCC-NL covariance matrix forecasts over $\overline{K} = 21$ days.
- **AFM-JDCC-NL** is similar to the AFM-DCC-NL model, but with additional constraints on the parameters governing the time-varying residual variances, which are estimated jointly in a similar manner to the JDCC-NL model.
- **RDCC-NL** is a Realized DCC model analogously to Bollerslev et al. (2020b), but applied to monthly returns, using daily data within the model and applying the nonlinear shrinkage estimator of Ledoit and Wolf (2020) to the correlation matrix. Because the RDCC-NL model is estimated based on monthly data, we use an estimation window of 30 years as for the MHEx model, see Section 3.2. The initial estimation periods at the beginning of the out-of-sample periods are 20 and 15 years for the FF and SP data sets, respectively, which are extended over time until the estimation periods reach 30 years.[8] The RDCC-NL model provides direct predictions of conditional variances and correlation matrices for the next month based on (15), which we combine to a forecast of the monthly covariance matrix.
- **BEKK-NL** is the scalar BEKK(1,1) model inspired by Engle and Kroner (1995) where the true but latent return covariance matrix $\Sigma_{t,k}$ for day k in month t is modeled as[9]

$$\Sigma_{t,k} = (1-\alpha-\beta)\,\Sigma + \alpha\, r_{t,k-1} r_{t,k-1}' + \beta\, \Sigma_{t,k-1}.$$

  Here, α and β are positive scalars. As in the DCC-NL model, the parameters are estimated via (quasi) maximum likelihood assuming a multivariate normal distribution for the vector of daily returns, and in large dimensions the composite likelihood method is used, see Section 2.4.1. The unconditional covariance matrix Σ is estimated with the nonlinear shrinkage estimator of Ledoit and Wolf (2020). Among the dynamic benchmark models, the BEKK-NL model imposes the strongest restrictions on the dynamics of individual daily variances and covariances, and it is feasible in high dimensions by the use of shrinkage methods.[10] We use the past five years of daily data to estimate the BEKK-NL covariance model and obtain a prediction for the next month from the BEKK-NL model by aggregating the BEKK-NL covariance matrix forecasts over $\overline{K} = 21$ days.
- **AFM-BEKK-NL** applies a single-factor model to the vector of asset returns, similar to the AFM-DCC-NL model, but uses the BEKK-NL model for the dynamic covariance matrix of the *residuals* from the factor model. The past five years of daily data are used to estimate the AFM-BEKK-NL covariance matrix model and to obtain a prediction for the next month by aggregating its covariance matrix forecasts over $\overline{K} = 21$ days.
- **EW** is an equally-weighted portfolio of assets which does not require any information about the return covariance matrix and is often used as a simple benchmark portfolio in empirical applications, see, e.g., DeMiguel et al. (2009) or Ledoit and Wolf (2017).

3.4. Performance evaluation

We evaluate the goodness-of-fit of the estimated models via mean squared errors (and similarly via mean absolute errors) for predicted volatilities, correlations, and covariances:

$$\mathrm{MSE}_{\mathrm{VOL}} = \frac{1}{T - t_{\mathrm{oos}} + 1} \sum_{t=t_{\mathrm{oos}}}^{T} \left(\widehat{\mathrm{RV}}_t - \mathrm{RV}_t\right)' \left(\widehat{\mathrm{RV}}_t - \mathrm{RV}_t\right) / N_t,$$

$$\mathrm{MSE}_{\mathrm{COR}} = \frac{1}{T - t_{\mathrm{oos}} + 1} \sum_{t=t_{\mathrm{oos}}}^{T} \|\widehat{\mathrm{RCOR}}_t - \mathrm{RCOR}_t\|_F^2 / N_t^2,$$

$$\mathrm{MSE}_{\mathrm{COV}} = \frac{1}{T - t_{\mathrm{oos}} + 1} \sum_{t=t_{\mathrm{oos}}}^{T} \|\widehat{\Sigma}_t - \mathrm{RCOV}_t\|_F^2 / N_t^2,$$

[7] We thank the editor and anonymous referees for their helpful suggestions on additional benchmark models.

[8] For the SP data sets, it is ensured that the single stocks in the portfolio have a complete five-year return history (see Section 3.1). To estimate the model for more than five years, we divide the estimation period into five-year subperiods, in which the RDCC-NL model is applied to the index constituents at the end of each subperiod.

[9] For the first day of each month, when $k = 1$, the right-hand side (in slight abuse of notation) refers to the last day of the previous month.

[10] As an additional benchmark model, we have examined a variant of the BEKK model by Bekaert and Wu (2000) with nonlinear shrinkage. We implemented the model, which includes a market index in the covariance matrix, by suitably restricting the parameter space to ensure its applicability in high dimensions. In our analysis, the model's performance is similar to that of the BEKK-NL model, outperforming the AFM-BEKK-NL model. However, due to space constraints and given its parsimony, we only report the results for the BEKK-NL model.

where t_{oos} indexes the first month of the out-of-sample period, $\widehat{\text{RV}}_t$, $\widehat{\text{RCOR}}_t$, and $\widehat{\Sigma}_t$ are the predictions of volatilities, correlation and covariance matrices, as given in (13) for the MHEx model, and $\|\cdot\|_F$ denotes the Frobenius norm of a matrix.

For each month t in the out-of-sample period, we calculate monthly realized portfolio returns as $\tilde{r}_t^{\text{PF}} = \hat{\omega}_t' \tilde{r}_t$, where $\hat{\omega}_t$ are estimated optimal portfolio weights for month t and $\tilde{r}_t$ denotes the vector of monthly simple returns of the N_t assets in the portfolio. We calculate daily realized portfolio returns as $\tilde{r}_{t,k}^{\text{PF}} = \hat{\omega}_{t,k}' \tilde{r}_{t,k}$ for days $k = 1, \dots, K_t$, where $\hat{\omega}_{t,k} = \frac{\hat{\omega}_t \circ \prod_{j=1}^{k-1}(1+\tilde{r}_{t,j})}{\hat{\omega}_t'\left(\prod_{j=1}^{k-1}(1+\tilde{r}_{t,j})\right)}$ reflects that the portfolio weights $\hat{\omega}_t$ from the beginning of the month change slightly as each day of the month passes, $\tilde{r}_{t,k}$ denote daily simple returns of the N_t assets in the portfolio, and $\circ$ is the Hadamard product, which also applies to the product operator $\prod$.

As portfolio performance statistics, we calculate annualized average portfolio returns (AV), standard deviations (SD), and Sharpe ratios (SR) from daily returns as

$$\text{AV} = \frac{252}{\sum_{t=t_{\text{oos}}}^{T} K_t} \sum_{t=t_{\text{oos}}}^{T} \sum_{k=1}^{K_t} \tilde{r}_{t,k}^{\text{PF}},$$

$$\text{SD} = \sqrt{\frac{252}{\sum_{t=t_{\text{oos}}}^{T} K_t} \sum_{t=t_{\text{oos}}}^{T} \sum_{k=1}^{K_t} \left(\tilde{r}_{t,k}^{\text{PF}} - \underline{\tilde{r}}^{\text{PF}}\right)^2},$$

$$\text{SR} = \frac{\frac{252}{\sum_{t=t_{\text{oos}}}^{T} K_t} \sum_{t=t_{\text{oos}}}^{T} \sum_{k=1}^{K_t} \tilde{r}_{t,k}^{\text{PF}} - \tilde{r}_{t,k}^{\text{Rf}}}{\sqrt{\frac{252}{\sum_{t=t_{\text{oos}}}^{T} K_t} \sum_{t=t_{\text{oos}}}^{T} \sum_{k=1}^{K_t} \left(\tilde{r}_{t,k}^{\text{PF}} - \tilde{r}_{t,k}^{\text{Rf}} - \left(\underline{\tilde{r}}^{\text{PF}} - \underline{\tilde{r}}^{\text{Rf}}\right)\right)^2}},$$

where the underline (e.g. in $\underline{\tilde{r}}^{\text{PF}}$) represents average returns and $\tilde{r}_{t,k}^{\text{Rf}}$ denotes the daily risk-free rate, which we obtain from Kenneth R. French's website.

For the estimated minimum variance portfolios, we test if the differences in out-of-sample variances are statistically significant using the HAC inference method from Ledoit and Wolf (2011). The primary comparison is between our dynamic MHEx model and the dynamic benchmark models given in Section 3.3.

To analyze the properties and implementability of the portfolio strategies, we calculate the average gross portfolio leverage (LEV), the average proportion of negative portfolio weights (NEG), and the average monthly portfolio turnover (TO) as

$$\begin{aligned}
\text{LEV} &= \frac{1}{T - t_{\text{oos}} + 1} \sum_{t=t_{\text{oos}}}^{T} \sum_{i=1}^{N_t} |\hat{\omega}_{i,t}|, \\
\text{NEG} &= \frac{1}{T - t_{\text{oos}} + 1} \sum_{t=t_{\text{oos}}}^{T} \frac{1}{N_t} \sum_{i=1}^{N_t} 1_{\{\hat{\omega}_{i,t}<0\}}, \qquad (18) \\
\text{TO} &= \frac{1}{T - t_{\text{oos}}} \sum_{t=t_{\text{oos}}+1}^{T} \sum_{i=1}^{N_t} |\hat{\omega}_{i,t} - \hat{\omega}_{i,t-1}^{+}|,
\end{aligned}$$

where $\hat{\omega}_{t-1}^{+}$ denotes the vector of portfolio weights at the end of month $t-1$ that result from holding the portfolio given by weights $\hat{\omega}_{t-1}$ through month $t-1$ (for the SP data sets we also account for firms leaving and entering the portfolio).

3.5. Empirical results

In this section, we present empirical backtesting results of the optimized portfolio strategies. First, we consider goodness-of-fit statistics of the estimated covariance matrix models. Second, we discuss the performance statistics and properties of unconstrained and constrained minimum variance portfolios.

Table 2
Out-of-sample goodness-of-fit statistics for predicted volatilities (VOL), correlations (COR), and covariances (COV) from the static and dynamic covariance matrix models for the FF data sets.

	RMSE			MAE		
	VOL	COR	COV	VOL	COR	COV
FF10						
NL	2.860	0.203	53.468	1.883	0.138	18.760
BEKK-NL	2.249	0.180	49.326	1.358	0.122	14.729
AFM-BEKK-NL	2.168	0.190	47.737	1.312	0.130	14.257
DCC-NL	2.151	0.181	47.385	1.290	0.126	**13.423**
JDCC-NL	**2.136**	0.180	**47.310**	1.287	0.125	13.488
AFM-DCC-NL	2.174	0.190	47.579	1.318	0.130	14.163
AFM-JDCC-NL	2.173	0.190	47.580	1.320	0.130	14.178
RDCC-NL	2.236	0.189	49.568	1.291	0.131	13.623
MHEx	2.144	**0.174**	48.140	**1.272**	**0.120**	13.777
FF30						
NL	3.175	0.219	60.096	2.075	0.162	21.136
BEKK-NL	2.645	0.204	55.428	1.676	0.150	17.781
AFM-BEKK-NL	2.402	0.208	51.415	1.479	0.155	15.784
DCC-NL	2.390	0.203	52.222	1.439	0.153	**15.084**
JDCC-NL	2.384	0.202	52.044	1.449	0.152	15.168
AFM-DCC-NL	2.372	0.206	51.232	1.452	0.153	15.691
AFM-JDCC-NL	**2.366**	0.204	**51.220**	1.451	0.151	15.692
RDCC-NL	2.499	0.208	55.030	1.464	0.157	15.580
MHEx	2.381	**0.192**	53.134	**1.420**	**0.143**	15.574
FF100						
NL	3.088	0.192	60.661	1.967	0.143	21.241
BEKK-NL	2.783	0.185	57.320	1.774	0.137	19.244
AFM-BEKK-NL	2.393	0.192	50.637	1.473	0.146	15.437
DCC-NL	2.351	0.178	51.522	1.393	0.136	14.873
JDCC-NL	2.312	0.178	51.080	1.369	0.135	**14.680**
AFM-DCC-NL	2.312	0.182	50.544	1.393	0.137	15.291
AFM-JDCC-NL	**2.304**	0.182	**50.470**	1.395	0.137	15.306
RDCC-NL	2.489	0.195	56.081	1.491	0.152	16.074
MHEx	2.323	**0.171**	52.676	**1.356**	**0.127**	15.351

Note: The table reports root mean squared errors (RMSE) and mean absolute errors (MAE) for the predictions of monthly realized volatilities, correlations, and covariances of percentage returns over the out-of-sample periods of the Fama–French data sets given in Table 1. Bold numbers indicate the lowest RMSE and MAE among the compared models.

3.5.1. Goodness-of-fit statistics

Tables 2 and 3 present goodness-of-fit measures for the static NL as well as the dynamic (AFM-)BEKK-NL, (J)DCC-NL, AFM-(J)DCC-NL, RDCC-NL, and MHEx covariance matrix models based on the portfolio (FF) and single stock (SP) data sets, respectively. We report RMSE and MAE for volatilities (VOL), correlations (COR), and covariances (COV), since the (J)DCC-NL, AFM-(J)DCC-NL, RDCC-NL and MHEx covariance matrix models are all estimated via two-step approaches. In terms of goodness-of-fit statistics, a distinction can be made between the FF (Table 2) and SP data sets (Table 3). The more pronounced unpredictable risk, i.e., noise, in single stock returns is reflected in overall higher RMSE and MAE for the SP data sets. For the FF data sets, the RMSE for VOL (COR) predictions is between 2.10 and 3.20 (0.17 and 0.22), while for the SP data sets, the RMSE lies between 4.85 and 7.22 (0.24 and 0.26). When comparing the different models, the results show fairly similar goodness-of-fit measures for the dynamic models and clear improvements over the static NL model. For VOL (COR) forecasts, the RMSE and MAE of the dynamic models are mostly between 15 and 30% (3 and 15%) lower than those of the static model. The improvements of the MHEx model over the static NL model for COR predictions are greater for the FF data sets (11 to 14% relative improvement) than for the SP data sets (5 to 6% relative improvement). For COV forecasts, where the realized covariances are presumably more driven by outliers (compared to volatilities or correlations), we find about 8 to 15% lower RMSE and 15 to 30% lower MAE for the dynamic models. Comparing the dynamic models, the MHEx model has the lowest MAE for VOL and COR predictions for all five data sets. It also has the lowest RMSE

Table 3
Out-of-sample goodness-of-fit statistics for predicted volatilities (VOL), correlations (COR), and covariances (COV) from the static and dynamic covariance matrix models for the SP data sets.

	RMSE			MAE		
	VOL	COR	COV	VOL	COR	COV
SP500						
NL	6.272	0.259	105.888	3.942	0.209	37.997
BEKK-NL	4.908	0.254	93.058	2.806	0.203	29.828
AFM-BEKK-NL	4.856	0.256	90.411	2.803	0.205	29.533
DCC-NL	5.125	0.249	96.348	2.927	0.202	29.474
JDCC-NL	4.955	0.249	93.408	2.835	0.201	**28.779**
AFM-DCC-NL	4.955	0.253	90.610	2.855	0.203	28.897
AFM-JDCC-NL	4.972	0.251	**90.236**	2.855	0.201	28.906
RDCC-NL	5.161	0.253	95.939	3.078	0.204	31.526
MHEx	**4.850**	**0.245**	93.313	**2.745**	**0.197**	29.099
SP1500						
NL	7.216	0.255	113.234	4.697	0.207	43.849
BEKK-NL	5.951	0.250	101.735	3.419	0.200	36.216
AFM-BEKK-NL	**5.851**	0.253	97.449	3.382	0.203	35.240
DCC-NL	6.261	0.244	103.861	3.669	0.198	36.301
JDCC-NL	6.106	0.245	102.637	3.537	0.198	35.160
AFM-DCC-NL	6.034	0.249	97.227	3.501	0.200	**34.264**
AFM-JDCC-NL	6.059	0.247	**97.170**	3.508	0.199	34.305
RDCC-NL	6.178	0.249	103.998	3.741	0.200	37.897
MHEx	5.867	**0.241**	101.802	**3.359**	**0.194**	35.516

Note: The table reports root mean squared errors (RMSE) and mean absolute errors (MAE) for the predictions of monthly realized volatilities, correlations, and covariances of percentage returns over the out-of-sample periods of the S&P 500 and S&P 1500 data sets given in Table 1. Bold numbers indicate the lowest RMSE and MAE among the compared models.

for COR predictions for all five data sets and is consistently in the top four for VOL predictions, where the differences are rather small and the ranking is more variable. For COV forecasts, either the (J)DCC-NL or AFM-(J)DCC-NL models attain the lowest RMSE and MAE. In summary, the MHEx model ranks overall first in terms of goodness-of-fit measures, followed by the (AFM-)JDCC-NL and (AFM-)DCC-NL models. Thus, the more flexible specification for modeling volatilities and correlations in the MHEx model proves to be beneficial. The potential advantage of direct monthly forecasting with the RDCC-NL model over the (J)DCC-NL and BEKK-NL models could not outweigh its drawback of being estimated based on less frequent (monthly) returns. The AFM-BEKK-NL model ranks second after the MHEx model in terms of RMSE and MAE for VOL predictions on the SP data sets, but yields mixed results for predicted volatilities, correlations and covariances over all data sets.

3.5.2. Results for unconstrained minimum variance portfolios

Tables 4 and 5 present out-of-sample performance measures for estimated minimum variance portfolios for the FF and SP data sets, respectively. The tables report results for unconstrained minimum variance portfolios as well as for minimum variance portfolios with weight constraints, which will be discussed in more detail in Section 3.5.3. In both tables, we focus on performance results based on daily returns as they provide more accurate estimates of out-of-sample standard deviations, which we consider to be the primary performance measure for minimum variance portfolios.

Across the five data sets, we find that all estimated minimum variance portfolios have significantly lower standard deviations than the equally-weighted portfolio (EW), indicating that the use of these covariance matrix estimates adds value within portfolio optimization. The annualized standard deviations of EW are between 16.0 and 17.5% for the FF data sets and between 20.0 and 21.5% for the single stock SP data sets, while the standard deviations for the minimum variance portfolios are between 8.1 and 12.2%. Most dynamic covariance matrix models, in particular the JDCC-NL, RDCC-NL, and MHEx models on all five data sets, lead to statistically significant improvements at the 1% level over the static NL model. The MHEx model significantly outperforms the NL model with absolute reductions in out-of-sample percentage standard deviations of 0.67, 0.80, 1.05, 1.27, and 1.78 across the five data sets. These improvements over the NL model, e.g. from 9.67 to 8.62% for the FF100 data set, are economically significant and become larger as the number of assets and the diversification opportunities increase.

Comparing the results for the dynamic models, we find that the MHEx model outperforms all benchmark models in terms of out-of-sample portfolio standard deviation on each data set. The outperformance versus each benchmark model is statistically significant at the 1% level for each data set. The JDCC-NL model ranks second overall across all data sets, followed by the AFM-JDCC-NL, RDCC-NL, and BEKK-NL models. The "joint" JDCC-NL model outperforms its "individual" DCC-NL counterpart in each data set (as does the AFM-JDCC-NL compared to the AFM-DCC-NL, except for the SP1500 data set). The reduction in portfolio standard deviation is 0.39 on average and is statistically significant at the 1% level, except for the SP1500 data set in Table 5. That is, imposing joint dynamics in the conditional variances across all assets in the DCC-NL model significantly improves the out-of-sample portfolio performance. The strongly restrictive scalar BEKK-NL model shows a good performance on the smaller data sets, e.g., the FF10 and FF30 data sets in Table 4, but does not perform as well on the larger SP data sets in Table 5, where it is outperformed by the more flexible JDCC-NL model. The reverse is true for the AFM-JDCC-NL and to some extent for the RDCC-NL model. We find that the AFM-DCC-NL outperforms the DCC-NL for the larger (FF100 and SP) data sets, as in De Nard et al. (2022), but for the more restrictive JDCC-NL and BEKK-NL models, the AFM-JDCC-NL and AFM-BEKK-NL variants show diminished outperformance or even underperformance.

The MHEx model improves upon the JDCC-NL (RDCC-NL) model with reductions in standard deviations of 0.17 (0.36), 0.31 (0.34), 0.44 (0.66) for the FF data sets in Table 4 and 0.67 (0.72), 0.93 (0.68) for the SP data sets in Table 5. That is, for the unconstrained minimum variance portfolios, the gains over the JDCC-NL (RDCC-NL) model are typically greater for a larger number of assets. In addition to their statistical significance (at the 1% level), the reductions in out-of-sample standard deviations obtained with the MHEx model are economically meaningful. The results show that the use of the MHEx model is beneficial for selecting from predefined portfolios as well as individual stocks, confirming the advantages of direct estimation on and prediction of monthly variables, i.e., volatilities and correlations, as well as flexibility in the specification of their lag structure.

Remark 3.1. Another way to enhance the compared models would be to incorporate intraday data, which could be OHLC data as in the IDR-DCC-NL model in De Nard et al. (2022) or high-frequency intraday returns as in the RDCC model in Bollerslev et al. (2020b) or the HExp model in Bollerslev et al. (2018). While for large individual stocks, OHLC data and intraday returns are available for recent subperiods, we use widely available daily returns as the common basis for comparison in this paper. In the same way that the IDR-DCC-NL model can yield improvements over the DCC-NL model, which could be of similar magnitude as the improvements with the MHEx model, we may expect improvements for our MHEx model by incorporating OHLC or intraday return data. Because our focus in this paper is on models based on daily returns, we reserve these extensions and analyses for future research.

With regard to out-of-sample Sharpe ratios, we find that the estimated minimum variance portfolios outperform the equally-weighted portfolio for all data sets in Tables 4 and 5. A statistically significant outperformance relative to EW is achieved with the MHEx model for the FF100 portfolios (at the 1% level), for the SP500 data set (at the 5% level), and for the SP1500 data set (at the 1% level). Comparing the dynamic MHEx model with the static NL model, the MHEx outperforms the static model at the 1% significance level for the SP data sets. The resulting out-of-sample Sharpe ratios are 0.92 (MHEx), 0.60 (NL),

Table 4
Out-of-sample performance measures of estimated unconstrained and 130/30 minimum variance portfolios based on different covariance matrix models for the FF data sets.

	Unconstr. portfolios			130/30 portfolios		
	AV	SD	SR	AV	SD	SR
FF10						
EW	11.92	16.07	0.47	11.92	16.07	0.47
NL	12.02	12.20	0.63	12.92	13.04	0.66
BEKK-NL	11.94	11.72	0.65	12.02	12.69	0.61
AFM-BEKK-NL	12.62	12.22	0.68	12.69	13.12	0.64
DCC-NL	11.75	11.86	0.63	12.08	12.83	0.61
JDCC-NL	11.82	11.70	0.64	11.88	12.74	0.59
AFM-DCC-NL	11.87	12.14	0.62	12.64	13.03	0.64
AFM-JDCC-NL	11.99	12.06	0.64	12.62	13.06	0.64
RDCC-NL	11.18	11.88	0.58	11.85	12.77	0.59
MHEx	11.54	**11.53*****	0.63	11.52	**12.68**	0.57
FF30						
EW	11.93	16.83	0.45	11.93	16.83	0.45
NL	10.66	11.45	0.55	11.78	11.91	0.63
BEKK-NL	9.99	10.90	0.52	11.10	11.46	0.59
AFM-BEKK-NL	9.99	11.33	0.50	11.31	11.89	0.59
DCC-NL	11.37	11.31	0.62	11.79	11.75	0.64
JDCC-NL	10.84	10.96	0.60	11.58	11.56	0.63
AFM-DCC-NL	9.94	11.23	0.50	11.36	11.76	0.60
AFM-JDCC-NL	10.26	11.09	0.54	11.57	11.74	0.62
RDCC-NL	9.94	10.99	0.51	10.54	11.57	0.54
MHEx	10.27	**10.65*****	0.56	10.90	**11.36*****	0.58
FF100						
EW	12.87	17.43	0.49	12.87	17.43	0.49
NL	17.83	9.67	1.40	16.41	12.31	0.98
BEKK-NL	17.86	9.25	1.46	16.33	11.74	1.02
AFM-BEKK-NL	17.35	9.43	1.38	15.92	12.16	0.96
DCC-NL	16.48	9.49	1.28	15.34	11.59	0.95
JDCC-NL	17.11	9.06	1.41	15.98	11.36	1.03
AFM-DCC-NL	17.55	9.21	1.44	16.21	12.06	0.99
AFM-JDCC-NL	17.62	9.11	1.46	16.28	12.03	1.00
RDCC-NL	16.56	9.27	1.32	15.58	11.51	0.98
MHEx	16.10	**8.62*****	1.37	15.23	**11.10*****	0.98

Note: The table reports the annualized means (AV), standard deviations (SD), and Sharpe ratios (SR) of daily out-of-sample percentage returns of estimated unconstrained and 130/30 minimum variance portfolios for the Fama–French data sets given in Table 1. Bold numbers indicate the lowest portfolio standard deviations for each data set. *, **, and *** indicate significant differences in standard deviations between the MHEx model and the benchmark models at the 10%, 5%, and 1% levels.

Table 5
Out-of-sample performance measures of estimated unconstrained and 130/30 minimum variance portfolios based on different covariance matrix models for the SP data sets.

	Unconstr. portfolios			130/30 portfolios		
	AV	SD	SR	AV	SD	SR
SP500						
EW	13.04	20.08	0.54	13.04	20.08	0.54
NL	8.88	11.30	0.60	10.17	11.99	0.67
BEKK-NL	9.35	11.14	0.65	10.63	11.34	0.75
AFM-BEKK-NL	9.24	11.49	0.62	10.52	11.89	0.71
DCC-NL	11.67	11.50	0.83	10.45	11.88	0.70
JDCC-NL	12.27	10.70	0.95	11.14	11.25	0.80
AFM-DCC-NL	10.47	10.86	0.77	10.84	11.62	0.75
AFM-JDCC-NL	9.70	10.66	0.71	9.84	11.58	0.67
RDCC-NL	10.32	10.74	0.76	9.99	11.32	0.70
MHEx	11.33	**10.03*****	0.92	10.82	**10.87*****	0.80
SP1500						
EW	15.20	21.44	0.61	15.20	21.44	0.61
NL	10.77	9.88	0.88	10.21	11.06	0.73
BEKK-NL	10.89	9.68	0.91	11.01	10.19	0.87
AFM-BEKK-NL	10.68	9.83	0.87	10.88	10.75	0.82
DCC-NL	11.94	9.22	1.07	11.40	10.43	0.89
JDCC-NL	12.24	9.03	1.12	10.97	9.88	0.90
AFM-DCC-NL	10.87	8.45	1.04	9.95	10.18	0.77
AFM-JDCC-NL	10.50	8.85	0.95	9.85	10.40	0.74
RDCC-NL	11.05	8.78	1.02	10.16	9.75	0.82
MHEx	11.40	**8.10*****	1.15	10.30	**8.92*****	0.92

Note: The table reports the annualized means (AV), standard deviations (SD), and Sharpe ratios (SR) of daily out-of-sample percentage returns of estimated unconstrained and 130/30 minimum variance portfolios for the S&P 500 and S&P 1500 data sets given in Table 1. Bold numbers indicate the lowest portfolio standard deviations for each data set. *, **, and *** indicate significant differences in standard deviations between the MHEx model and the benchmark models at the 10%, 5%, and 1% levels.

and 0.54 (EW) on the SP500 data set and 1.15 (MHEx), 0.88 (NL), and 0.61 (EW) on the SP1500 data set. The MHEx model significantly outperforms (at least at the 5% level) the (AFM-)BEKK-NL, AFM-JDCC-NL, and RDCC-NL models in terms of out-of-sample Sharpe ratios on the SP data sets. Between the MHEx and the other dynamic benchmark models, especially the JDCC-NL model, the Sharpe ratios are quite similar and there are no significant differences.

To further analyze the robustness of the empirical results in Tables 4 and 5, we break down the aggregate performance measures into rolling subperiods of the out-of-sample periods. Fig. 2 shows the annualized three-year rolling standard deviations for the estimated minimum variance portfolios using the NL, JDCC-NL, RDCC-NL, or MHEx models for the investment universes of predefined portfolios, as for FF100, and individual stocks, as for SP1500. The MHEx model consistently outperforms the static (NL) and dynamic (JDCC-NL, RDCC-NL) benchmark models over time. This holds true for various periods of low market volatility as well as for periods of high volatility, such as the early 2000s, the 2007/2008 financial crisis, or the 2020 stock market crash. Fig. 2 shows that for practically all three-year subperiods between 1970 (1995) and 2022 for the FF100 (SP1500) data set, the MHEx model yields lower portfolio standard deviations than the NL, JDCC-NL, or RDCC-NL models. In general, the RDCC-NL and JDCC-NL models follow each other quite closely, with occasional shifts in dominance over time. However, both are always, i.e., for all three-year subperiods, dominated by the MHEx model.

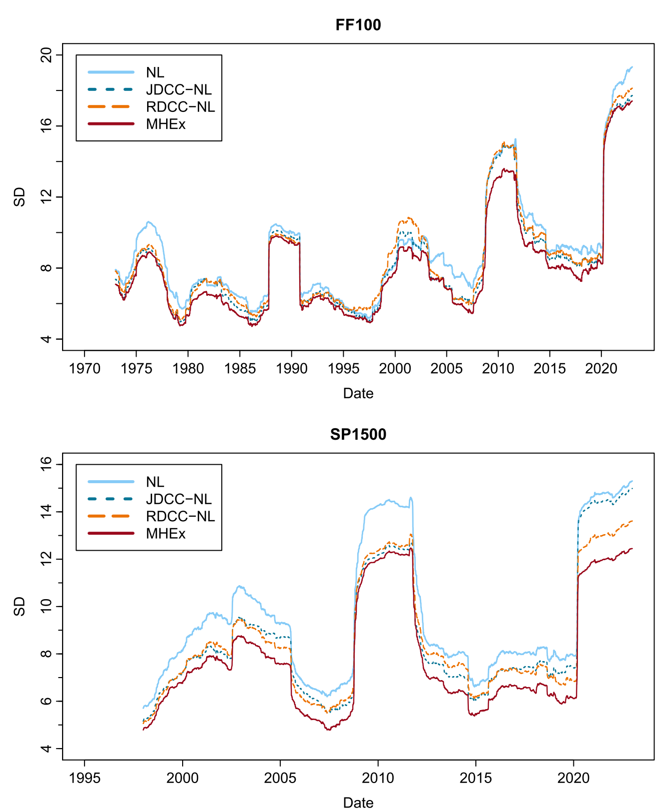

Fig. 2. Annualized three-year rolling standard deviations of daily out-of-sample percentage portfolio returns of estimated minimum variance portfolios based on different covariance matrix models. The rolling window ends on the date indicated on the x-axis.

From the out-of-sample analysis, we conclude that the MHEx model is the preferred (high-dimensional) covariance matrix model. It generally outperforms the static and dynamic benchmark models for unconstrained minimum variance portfolio optimization in different investment universes and for arbitrary subperiods of the respective out-of-sample periods. The next section examines the robustness of the performance statistics under reasonable portfolio weight constraints.

3.5.3. Results for constrained minimum variance portfolios

In terms of accuracy of dynamic covariance matrix estimation, the MHEx model compares favorably with the benchmark models. This is based on goodness-of-fit statistics and the out-of-sample performance of estimated unconstrained minimum variance portfolios. However, the unconstrained optimized portfolios have a large number of short positions, large absolute portfolio weights, high gross leverage, and high portfolio turnover, especially with the dynamic models, as discussed in more detail in Section 3.5.4. These characteristics render portfolio strategies relatively unattractive for practical implementation. To re-examine the application of the models in a more practical setting, we introduce constraints into the portfolio optimization problem as formulated in (17) and compare the out-of-sample performance of the resulting estimated minimum variance optimal portfolios.

In this application, we limit the gross leverage to a maximum of 1.6, which represents a 130/30 long-short equity strategy, i.e., the portfolio then holds no more than 30% of its initial capital in short positions. This type of strategy is quite popular among the long-short strategies of investment funds, as discussed in Lo and Patel (2008). We further impose restrictions on the size of individual portfolio positions, which is common among investment funds. In our empirical application, portfolio positions are limited to 30% in absolute value for the FF10 and FF30 industry portfolios, 10% for the FF100 stock characteristic portfolios, and 5% for the single stock SP data sets.[11]

The right-hand side columns of Tables 4 and 5 present the out-of-sample performance statistics for the constrained 130/30 minimum variance portfolios for the FF and SP data sets, respectively. The standard deviations of the estimated minimum variance portfolios collectively rise as a result of the constraints on portfolio weights. Focusing on the MHEx model, the smallest increase is from 10.65 to 11.36 for the FF30 data set, i.e., 0.71 in terms of annualized percentage standard deviation, while the largest increase is from 8.62 to 11.10, or 2.48, for the FF100 data set. These figures imply that the restrictions on the portfolio leverage and the size of individual portfolio positions have an economically relevant impact.

The estimated minimum variance portfolios show significantly lower out-of-sample standard deviations than the equally-weighted portfolio. The improvements in terms of portfolio standard deviations with the MHEx model over the static NL model reduce slightly for the FF10 and FF30 industry portfolios, but remain robust for the FF100 and SP data sets, when comparing the constrained with the unconstrained portfolios. The improvements over the NL model range from 0.36 to 2.14 and are statistically significant at the 1% level for all five data sets in Tables 4 and 5.

Comparing the dynamic models for the constrained portfolios, our MHEx model outperforms the benchmark models for all five data sets, with significant outperformance of the MHEx model at the 1% level only lacking for the FF10 data set. The improvements in terms of portfolio standard deviations, in particular over the JDCC-NL and RDCC-NL models (which rank second and third), are monotonically increasing with the number of assets, i.e., they are highest for the single stock data sets. Compared to the unconstrained minimum variance

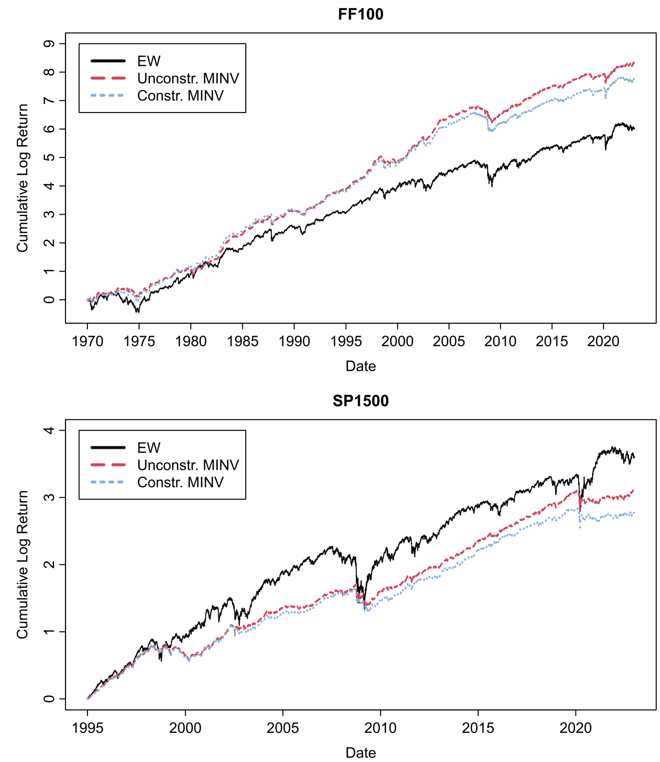

Fig. 3. Cumulative out-of-sample performance paths based on daily log returns of estimated unconstrained and 130/30 minimum variance portfolios. "EW" is the equally-weighted portfolio, "Unconstr. MINV" and "Constr. MINV" refer to the unconstrained and 130/30 minimum variance portfolios based on the MHEx covariance matrix model.

portfolios, these differences between the JDCC-NL (RDCC-NL) and the MHEx model have decreased slightly, i.e., on average by 0.14. That is, for the FF30 and FF100 data sets, the MHEx model improves upon the JDCC-NL (RDCC-NL) model by 0.20 (0.21) and 0.26 (0.41), while for the SP500 and SP1500 data sets, the MHEx model yields improvements of 0.38 (0.45) and 0.96 (0.83) in terms of annualized percentage portfolio standard deviations. In addition to their statistical significance, these reductions in out-of-sample standard deviations are economically meaningful, especially for the single stock portfolios. For the SP1500 data set, the annualized standard deviations are 11.06% for NL, 9.88% for JDCC-NL, 9.75% for RDCC-NL, and 8.92% for MHEx. These results show that the MHEx model is the preferred high-dimensional covariance matrix model for minimum variance portfolios under portfolio constraints.

Fig. 3 plots the cumulative out-of-sample performance paths from daily realized returns of the unconstrained ("Unconstr. MINV") and 130/30 ("Constr. MINV") minimum variance portfolios based on the MHEx model for the FF100 and SP1500 data sets. For comparison with a simple benchmark index, we also show the paths of the equally-weighted portfolio. In both charts, the performance paths of the estimated unconstrained and constrained minimum variance portfolios show stable performance over time and are visibly less volatile than EW, implying superior risk-adjusted performance, e.g., in terms of Sharpe ratio. The 130/30 minimum variance portfolio slightly underperforms the unconstrained minimum variance portfolio since the early 2000s. For the FF100 characteristics portfolio data set, the estimated minimum variance portfolios – with exceptions in the early 1970s and early 1980s – consistently outperform the equally-weighted portfolio over the 53-year out-of-sample period. For the SP1500 single stock data set, the estimated minimum variance portfolios end up below the equally-weighted portfolio, but show a stable performance and achieve their goal of producing a low volatility strategy with lower drawdown during periods of distress.

[11] We have also considered limiting gross portfolio leverage to 2.0 or 1.0, i.e., 150/50 long-short equity portfolios and long-only portfolios, with the same bounds on portfolio weights. The conclusions from these results are similar to those discussed in this section.

Table 6
Portfolio weight statistics for estimated unconstrained and 130/30 minimum variance portfolios based on different covariance matrix models for the FF data sets.

	Unconstr. portfolios					130/30 portfolios				
	MAX	MIN	NEG	LEV	TO	MAX	MIN	NEG	LEV	TO
FF10										
EW	0.10	0.10	0.00	1.00	0.02	0.10	0.10	0.00	1.00	0.02
NL	0.77	−0.26	0.43	2.11	0.11	0.30	−0.16	0.29	1.53	0.07
BEKK-NL	0.77	−0.30	0.44	2.25	0.77	0.30	−0.18	0.26	1.54	0.38
AFM-BEKK-NL	0.79	−0.32	0.41	2.30	0.79	0.30	−0.18	0.26	1.53	0.31
DCC-NL	0.71	−0.26	0.42	2.08	1.13	0.30	−0.18	0.27	1.55	0.67
JDCC-NL	0.73	−0.26	0.43	2.08	1.09	0.30	−0.18	0.27	1.54	0.67
AFM-DCC-NL	0.79	−0.30	0.42	2.18	0.80	0.30	−0.18	0.26	1.52	0.32
AFM-JDCC-NL	0.78	−0.30	0.42	2.22	0.85	0.30	−0.18	0.26	1.52	0.34
RDCC-NL	0.74	−0.20	0.44	1.82	0.63	0.30	−0.16	0.30	1.54	0.43
MHEx	0.76	−0.25	0.44	2.09	1.07	0.30	−0.17	0.28	1.53	0.63
FF30										
EW	0.03	0.03	0.00	1.00	0.03	0.03	0.03	0.00	1.00	0.03
NL	0.53	−0.19	0.46	2.84	0.21	0.29	−0.13	0.20	1.60	0.11
BEKK-NL	0.54	−0.21	0.45	2.97	0.71	0.29	−0.12	0.20	1.60	0.34
AFM-BEKK-NL	0.52	−0.20	0.44	2.86	0.73	0.29	−0.12	0.21	1.60	0.35
DCC-NL	0.50	−0.17	0.47	2.71	1.71	0.29	−0.12	0.22	1.60	1.03
JDCC-NL	0.51	−0.17	0.47	2.69	1.67	0.29	−0.11	0.22	1.60	1.04
AFM-DCC-NL	0.54	−0.20	0.44	2.77	0.91	0.29	−0.13	0.21	1.60	0.46
AFM-JDCC-NL	0.53	−0.19	0.45	2.75	0.90	0.29	−0.13	0.21	1.60	0.45
RDCC-NL	0.58	−0.15	0.47	2.52	0.91	0.29	−0.11	0.22	1.60	0.56
MHEx	0.51	−0.15	0.47	2.66	1.56	0.29	−0.11	0.22	1.60	0.95
FF100										
EW	0.01	0.01	0.00	1.00	0.02	0.01	0.01	0.00	1.00	0.02
NL	0.32	−0.15	0.48	5.01	0.50	0.10	−0.09	0.06	1.60	0.13
BEKK-NL	0.32	−0.15	0.48	5.18	1.08	0.10	−0.09	0.06	1.60	0.25
AFM-BEKK-NL	0.31	−0.14	0.48	4.97	1.06	0.10	−0.09	0.07	1.60	0.26
DCC-NL	0.29	−0.11	0.50	4.58	3.26	0.10	−0.09	0.07	1.60	1.20
JDCC-NL	0.30	−0.11	0.50	4.44	3.10	0.10	−0.09	0.07	1.60	1.23
AFM-DCC-NL	0.33	−0.14	0.50	4.86	1.72	0.10	−0.09	0.07	1.60	0.42
AFM-JDCC-NL	0.33	−0.14	0.50	4.86	1.66	0.10	−0.09	0.07	1.60	0.42
RDCC-NL	0.30	−0.12	0.50	4.63	2.34	0.10	−0.09	0.07	1.60	0.77
MHEx	0.29	−0.10	0.51	4.46	2.70	0.10	−0.09	0.07	1.60	1.07

Note: The table reports the average maximum (MAX) and minimum (MIN) weights, proportion of negative weights (NEG), gross leverage (LEV), and turnover (TO) of estimated unconstrained and 130/30 minimum variance portfolios for the Fama–French data sets.

3.5.4. Portfolio weight statistics for unconstrained and constrained portfolios

To analyze the empirical properties of the optimized portfolio weights for the FF and SP data sets in more detail, Tables 6 and 7 present statistics on the weights for the estimated unconstrained and 130/30 minimum variance portfolios, i.e., the average maximum and minimum portfolio weights, proportion of negative portfolio weights, gross leverage, and portfolio turnover, as defined in (18).

For the unconstrained minimum variance portfolios, the portfolio positions are quite large (especially for the smaller portfolios), e.g., the largest long position is over 70% on average for the FF10 data set and over 9% on average for the SP500 data set. The largest short positions are about three times smaller than the largest long positions, but the proportion of negative portfolio weights is consistently between 42 and 50%, meaning that about half of the positions in the portfolios are short positions. Gross leverage for the unconstrained portfolios is considerably higher for the FF100 characteristic portfolios and the SP single stock data sets than for those based on the FF10 and FF30 data sets. For the former, the gross portfolio leverage ranges from 3.7 to 5.8, while for the latter it ranges from 1.8 to 3.0. The portfolios based on the static covariance matrix model exhibit similar leverage to those based on the dynamic models, while the (AFM-)BEKK-NL model yields higher leverage than the other benchmark models, and the AFM-type models yield higher leverage than their respective counterparts. Using the MHEx model, average gross portfolio leverage is lowest among all covariance matrix models for the SP data sets in Table 7, and second lowest for the FF30 and FF100 data sets in Table 6.

In terms of portfolio turnover, the equally-weighted portfolio has the lowest average turnover, as its portfolio positions remain roughly the same from period to period. The unconstrained minimum variance portfolios based on the static NL model display significantly lower turnover than those based on the dynamic models. The average monthly portfolio turnover generally increases with the number of assets, from 0.11 to 2.70 for the NL model and from 1.07 to 3.03 for the MHEx model across the five data sets. Among the dynamic models, BEKK-NL yields the lowest turnover for the portfolio-based FF30 and FF100 data sets in Table 6 (across all five data sets, it ranges from 0.71 to 3.30). The MHEx model yields the lowest turnover for the SP1500 data set and second lowest turnover for the SP500 data set in Table 7. The DCC-NL and JDCC-NL models yield the highest portfolio turnovers for all data sets, while their AFM-type counterparts yield lower turnover.

The statistics for the unconstrained minimum variance portfolio weights show an average gross leverage above 1.6 for all optimized portfolios, so the leverage constraint is almost always binding for the constrained 130/30 portfolios. In the same vein, the average maximum portfolio weights for the constrained portfolios are at the specified upper bounds for the portfolio weights. The absolute average minimum weights decrease due to the leverage constraint and thus are always greater than the lower bounds. In terms of the number of short positions, the average proportion of negative portfolio weights is greatly reduced by the leverage constraint, from about 50% for the unconstrained portfolios to less than 10% for the constrained portfolios for the FF100 and SP data sets, and to 20%–30% for the FF10 and FF30 data sets. Likewise, portfolio turnover is reduced as a result of the portfolio constraints. The average turnover decreases by 35%–75% for the estimated minimum variance portfolios. For example, for the SP1500 data set, the average monthly turnover with the MHEx model

Table 7
Portfolio weight statistics for estimated unconstrained and 130/30 minimum variance portfolios based on different covariance matrix models for the SP data sets.

	Unconstr. portfolios					130/30 portfolios				
	MAX	MIN	NEG	LEV	TO	MAX	MIN	NEG	LEV	TO
SP500										
EW	0.00	0.00	0.00	1.00	0.07	0.00	0.00	0.00	1.00	0.07
NL	0.05	−0.04	0.44	4.69	1.53	0.05	−0.04	0.08	1.60	0.57
BEKK-NL	0.06	−0.05	0.45	5.32	2.63	0.05	−0.04	0.06	1.60	1.34
AFM-BEKK-NL	0.07	−0.05	0.45	5.75	2.73	0.05	−0.04	0.06	1.60	1.14
DCC-NL	0.13	−0.04	0.49	4.20	3.41	0.05	−0.03	0.09	1.60	1.64
JDCC-NL	0.09	−0.04	0.47	4.46	3.16	0.05	−0.03	0.09	1.60	1.50
AFM-DCC-NL	0.10	−0.05	0.45	4.77	2.68	0.05	−0.04	0.07	1.60	1.03
AFM-JDCC-NL	0.09	−0.05	0.46	4.90	2.41	0.05	−0.04	0.08	1.60	0.94
RDCC-NL	0.09	−0.04	0.47	4.64	2.85	0.05	−0.03	0.08	1.60	1.18
MHEx	0.09	−0.03	0.48	3.74	2.62	0.05	−0.03	0.10	1.60	1.43
SP1500										
EW	0.00	0.00	0.00	1.00	0.09	0.00	0.00	0.00	1.00	0.09
NL	0.02	−0.01	0.42	4.58	2.70	0.03	−0.02	0.06	1.60	1.12
BEKK-NL	0.02	−0.02	0.43	5.10	3.30	0.04	−0.02	0.05	1.60	1.60
AFM-BEKK-NL	0.02	−0.02	0.44	5.47	3.54	0.04	−0.03	0.04	1.60	1.54
DCC-NL	0.10	−0.01	0.48	3.99	3.78	0.05	−0.02	0.07	1.60	1.98
JDCC-NL	0.04	−0.01	0.48	4.45	3.63	0.05	−0.02	0.07	1.60	1.80
AFM-DCC-NL	0.06	−0.02	0.45	4.69	3.51	0.05	−0.03	0.07	1.60	1.50
AFM-JDCC-NL	0.04	−0.02	0.46	4.86	3.38	0.05	−0.02	0.07	1.60	1.43
RDCC-NL	0.05	−0.02	0.47	4.48	3.35	0.05	−0.02	0.07	1.60	1.54
MHEx	0.06	−0.01	0.48	3.81	3.03	0.05	−0.01	0.09	1.60	1.69

Note: The table reports the average maximum (MAX) and minimum (MIN) weights, proportion of negative weights (NEG), gross leverage (LEV), and turnover (TO) of estimated unconstrained and 130/30 minimum variance portfolios for the S&P 500 and S&P 1500 data sets.

is 1.69 for the weight-constrained portfolio instead of 3.03 for the unconstrained portfolio. These statistics show that the practical implementability of the portfolio strategies is greatly enhanced by the introduction of portfolio constraints. The constraints lead to a modest overall decrease in portfolio performance, however, using the MHEx model remains beneficial in this more practical setting.[12]

4. Conclusion

We propose a new framework for modeling and forecasting large dynamic covariance matrices for optimal portfolio selection. Our model produces direct forecasts of monthly covariance matrices using daily returns and enables smooth and separate dynamics for volatilities and correlations. Specifically, we model monthly volatilities and correlations based on exponentially-weighted heterogeneous autoregressive components and estimate their dynamics via standard least squares methods. This results in a parsimonious yet flexible covariance matrix model that is computationally efficient and simple to implement. Incorporating shrinkage estimators into our model for correlations makes our framework applicable in high-dimensional settings, which is important for the selection of individual stock portfolios.

Empirical backtesting results confirm the benefits of our proposed MHEx model. We use the model for monthly portfolio selection and consider asset universes ranging from ten to 1500, where the assets are predefined portfolios or individual stocks. Minimum variance portfolios based on covariance matrix forecasts from our MHEx model deliver favorable out-of-sample performance statistics compared to (R)DCC-NL, AFM-DCC-NL, or BEKK-NL benchmark models. The differences in portfolio standard deviations are statistically significant and economically meaningful for all empirical data sets. For the 100 portfolios formed on size and book-to-market, annualized portfolio standard deviations improve from 9.49% (9.27%) using the (R)DCC-NL model to 8.62% using our MHEx model. The AFM-DCC-NL and BEKK-NL perform similarly to the RDCC-NL model, i.e., they still produce higher portfolio standard deviations than the MHEx model, with values of 9.21 and 9.25%, respectively.

As unconstrained optimized portfolios often exhibit large portfolio positions, high portfolio leverage and high turnover, we consider a more practical setting with weight-constrained 130/30 long-short minimum variance portfolios. The differences in out-of-sample standard deviations remain statistically significant and economically meaningful for all but our smallest data set. For example, considering the universe of S&P 1500 constituents, the annualized portfolio standard deviation improves to 8.92% using our MHEx model compared to 10.43% (9.75%) using the (R)DCC-NL model. Likewise, we find increasing outperformance of the MHEx model over the AFM-DCC-NL, BEKK-NL, and a restricted DCC-NL model. For large single stock data sets, the minimum variance portfolios based on our MHEx model significantly outperform portfolios based on a static covariance matrix model and the equally-weighted portfolio in terms of out-of-sample Sharpe ratio.

Our paper offers various directions for future research. To further enhance its forecasting ability, our model could be extended to incorporate semi-(co)variance and realized measures based on high-frequency intraday data. For example, Patton and Sheppard (2015) use intraday-based semi-variances in a HAR model, Bollerslev et al. (2020b) use intraday-based semi-covariances in a DCC model, and De Nard et al. (2022) use range-based variance estimators in a DCC-NL model. In addition, for larger stock based data sets, more sophisticated panel methods could be used for model specification and estimation.

CRediT authorship contribution statement

Igor Honig: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Felix Kircher:** Writing – review & editing, Writing – original draft, Software, Methodology, Formal analysis, Data curation, Conceptualization.

[12] We have also evaluated the portfolio performance with returns net of transaction costs (similar to Tables 4 and 5), assuming constant transaction costs of 10 or 50 bps. The impact of transaction costs on portfolio standard deviations is usually small and does not alter the ranking of the models and our conclusions. However, the impact on average returns can be large, depending on the level of transaction costs. To get an indication, annualized average trading costs can be estimated as twelve times the average monthly turnover (TO in Tables 6 and 7) times the constant transaction cost. To further reduce the turnover of optimized portfolios, turnover constraints should be implemented at the portfolio optimization stage, which may be relevant for practical portfolio management but is beyond the scope of this paper.

Data availability

The authors do not have permission to share data.

References

Alves, R.P., de Brito, D.S., Medeiros, M.C., Ribeiro, R.M., 2024. Forecasting large realized covariance matrices: The benefits of factor models and shrinkage. J. Financ. Econom. 22 (3), 696–742.

Andersen, T.G., Li, Y., Todorov, V., Zhou, B., 2023. Volatility measurement with pockets of extreme return persistence. J. Econometrics 237 (2, Part C), 105048.

Ang, A., Hodrick, R.J., Xing, Y., Zhang, X., 2009. High idiosyncratic volatility and low returns: International and further U.S. evidence. J. Financ. Econ. 91 (1), 1–23.

Barndorff-Nielsen, O.E., Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. J. R. Stat. Soc. Ser. B Stat. Methodol. 64 (2), 253–280.

Barndorff-Nielsen, O.E., Shephard, N., 2004. Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. Econom. 72 (3), 885–925.

Bauwens, L., Xu, Y., 2023. DCC- and DECO-HEAVY: Multivariate GARCH models based on realized variances and correlations. Int. J. Forecast. 39 (2), 938–955.

Bekaert, G., Bergbrant, M., Kassa, H., 2025. Expected idiosyncratic volatility. J. Financ. Econ. 167, 104023.

Bekaert, G., Engstrom, E.C., Xu, N.R., 2022. The time variation in risk appetite and uncertainty. Manag. Sci. 68 (6), 3975–4004.

Bekaert, G., Wu, G., 2000. Asymmetric volatility and risk in equity markets. Rev. Financ. Stud. 13 (1), 1–42.

Bollerslev, T., Hood, B., Huss, J., Pedersen, L.H., 2018. Risk everywhere: Modeling and managing volatility. Rev. Financ. Stud. 31 (7), 2729–2773.

Bollerslev, T., Li, J., Patton, A.J., Quaedvlieg, R., 2020a. Realized semicovariances. Econom. 88 (4), 1515–1551.

Bollerslev, T., Meddahi, N., Nyawa, S., 2019. High-dimensional multivariate realized volatility estimation. J. Econometrics 212 (1), 116–136.

Bollerslev, T., Patton, A.J., Quaedvlieg, R., 2016. Exploiting the errors: A simple approach for improved volatility forecasting. J. Econometrics 192 (1), 1–18.

Bollerslev, T., Patton, A.J., Quaedvlieg, R., 2020b. Multivariate leverage effects and realized semicovariance GARCH models. J. Econometrics 217 (2), 411–430.

Bollerslev, T., Patton, A.J., Quaedvlieg, R., 2022. Realized semibetas: Disentangling "good" and "bad" downside risks. J. Financ. Econ. 144 (1), 227–246.

Brailsford, T.J., Faff, R.W., 1996. An evaluation of volatility forecasting techniques. J. Bank. Financ. 20 (3), 419–438.

Corsi, F., 2009. A simple approximate long-memory model of realized volatility. J. Financ. Econom. 7 (2), 174–196.

Creal, D.D., Tsay, R.S., 2015. High dimensional dynamic stochastic copula models. J. Econometrics 189 (2), 335–345.

De Nard, G., Engle, R.F., Ledoit, O., Wolf, M., 2022. Large dynamic covariance matrices: Enhancements based on intraday data. J. Bank. Financ. 138, 106426.

De Nard, G., Ledoit, O., Wolf, M., 2021. Factor models for portfolio selection in large dimensions: The good, the better and the ugly. J. Financ. Econom. 19 (2), 236–257.

DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy? Rev. Financ. Stud. 22 (5), 1915–1953.

Engle, R., 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. J. Bus. Econ. Stat. 20 (3), 339–350.

Engle, R.F., Kroner, K.F., 1995. Multivariate simultaneous generalized ARCH. Econometric Theory 11 (1), 122–150.

Engle, R.F., Ledoit, O., Wolf, M., 2019. Large dynamic covariance matrices. J. Bus. Econ. Stat. 37 (2), 363–375.

French, K.R., Schwert, G.W., Stambaugh, R.F., 1987. Expected stock returns and volatility. J. Financ. Econ. 19 (1), 3–29.

Ghysels, E., Plazzi, A., Valkanov, R., Rubia, A., Dossani, A., 2019. Direct versus iterated multiperiod volatility forecasts. Annu. Rev. Financ. Econ. 11, 173–195.

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.

Jagannathan, R., Ma, T., 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. J. Finance 58 (4), 1651–1683.

Ledoit, O., Wolf, M., 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. J. Empir. Finance 10 (5), 603–621.

Ledoit, O., Wolf, M., 2004a. Honey, I shrunk the sample covariance matrix. J. Portf. Manag. 30 (4), 110–119.

Ledoit, O., Wolf, M., 2004b. A well-conditioned estimator for large-dimensional covariance matrices. J. Multivar. Anal. 88 (2), 365–411.

Ledoit, O., Wolf, M., 2011. Robust performance hypothesis testing with the variance. Wilmott 2011 (55), 86–89.

Ledoit, O., Wolf, M., 2012. Nonlinear shrinkage estimation of large-dimensional covariance matrices. Ann. Stat. 40 (2), 1024–1060.

Ledoit, O., Wolf, M., 2017. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks. Rev. Financ. Stud. 30 (12), 4349–4388.

Ledoit, O., Wolf, M., 2020. Analytical nonlinear shrinkage of large-dimensional covariance matrices. Ann. Stat. 48 (5), 3043–3065.

Ledoit, O., Wolf, M., 2022. Quadratic shrinkage for large covariance matrices. Bernoulli 28 (3), 1519–1547.

Lo, A.W., Patel, P.N., 2008. 130/30: The new long-only. J. Portf. Manag. 34 (2), 12–38.

Markowitz, H., 1952. Portfolio selection. J. Finance 7 (1), 77–91.

Merton, R.C., 1980. On estimating the expected return on the market. J. Financ. Econ. 8 (4), 323–361.

Moreira, A., Muir, T., 2017. Volatility-managed portfolios. J. Finance 72 (4), 1611–1644.

Oh, D.H., Patton, A.J., 2016. High-dimensional copula-based distributions with mixed frequency data. J. Econometrics 193 (2), 349–366.

Oh, D.H., Patton, A.J., 2023. Dynamic factor copula models with estimated cluster assignments. J. Econometrics 237 (2, Part C), 105374.

Opschoor, A., Lucas, A., Barra, I., van Dijk, D., 2021. Closed-form multi-factor copula models with observation-driven dynamic factor loadings. J. Bus. Econ. Stat. 39 (4), 1066–1079.

Pakel, C., Shephard, N., Sheppard, K., Engle, R.F., 2021. Fitting vast dimensional time-varying covariance models. J. Bus. Econ. Stat. 39 (3), 652–668.

Patton, A.J., Sheppard, K., 2015. Good volatility, bad volatility: Signed jumps and the persistence of volatility. Rev. Econ. Stat. 97 (3), 683–697.