Asset Allocation with Factor-Based Covariance Matrices[ⱡ]

Thomas Conlon[a], John Cotter[b], and Iason Kynigakis[c*]

Abstract

We examine whether a factor-based framework to construct the covariance matrix can enhance the performance of minimum-variance portfolios. We conduct a comprehensive comparative analysis of a wide range of factor models, which can differ based on the dimensionality reduction approach used to construct the latent factors and whether the covariance matrix is static or dynamic. The results indicate that factor models exhibit superior predictive accuracy compared to several covariance benchmarks, which can be attributed to the reduced degree of over predictions. Factor-based portfolios generate statistically significant outperformance with respect to standard deviation and Sharpe ratio relative to multiple portfolio benchmarks. After accounting for transaction costs strategies based on static covariance matrices outperform dynamic specifications in terms of risk-adjusted returns.

JEL classifications: C38, C58, G11

EJOR Keyword: Finance

Keywords: Covariance matrix; Dimensionality reduction; Factor models; Minimum-variance portfolio.

[ⱡ] An earlier version of this paper circulated under the title "Machine Learning and Factor-Based Portfolio Optimization". The authors gratefully acknowledge the support of Science Foundation Ireland under Grants Numbers 16/SPP/3347, 13/RC/2106 P2, 17/SPP/5447 and of the Irish Research Council under grant number GOIPD/2022/721. We would like to thank three anonymous referees and the Editor, Emanuele Borgonovo, for helpful advice and suggestions. We also acknowledge the comments of Andrea Barbon, Wolfgang Bessler, Martin Brown, Gregory Connor, Harald Lohre, and seminar participants at the University of Canterbury, the University of Hamburg, the University of Otago, the University of St. Gallen, the Victoria University of Wellington, the 3rd Frontiers of Factor Investing Conference, the 2022 European, Annual Meeting and Asia Pacific Conferences - Financial Management Association.

[a] Smurfit Graduate Business School, University College Dublin, Dublin, Ireland. Email: conlon.thomas@ucd.ie

[b] Smurfit Graduate Business School, University College Dublin, Dublin, Ireland. Email: john.cotter@ucd.ie

[c] Smurfit Graduate Business School, University College Dublin, Dublin, Ireland. Email: iason.kynigakis@ucd.ie

[*] Corresponding author.

1. Introduction

The classic mean-variance framework of Markowitz (1952) requires knowledge of the mean and the covariance matrix of the investment opportunity set, which are unknown quantities that need to be estimated. In practice, sample moments are commonly used to replace their true counterparts, which often carry considerable estimation risk, leading to suboptimal portfolios with extreme weights that fluctuate considerably over time and perform poorly out-of-sample (Broadie, 1993; Kan and Zhou, 2007). Among the many approaches proposed to deal with the challenges surrounding the Markowitz portfolio optimization problem, Kolm, Tütüncü and Fabozzi (2014), include risk-based allocation approaches, which require a risk model but no return model. Therefore, we focus instead on minimum-variance portfolios, which correspond to a risk averse investor who aims to minimize portfolio variance, without a need for estimates of expected returns. This framework is well motivated by Merton (1980), since it only requires estimates of the covariance matrix, which are often considered to be more accurate than estimates of the mean (Best and Grauer, 1991a,b; Chopra and Turner, 1993) and has been shown to outperform simpler benchmarks (see e.g., Chan, Karceski, and Lakonishok, 1999) and mean-variance portfolios (see e.g., Jagannathan and Ma, 2003). However, the covariance matrix estimates are still subject to estimation error (DeMiguel, Garlappi, Nogales and Uppal, 2009), which is exacerbated for larger portfolio sizes, where the asset covariances to be estimated increase with the number of assets.

In this paper, we overcome the problem of covariance misspecification by imposing a factor structure on the covariance matrix.[1] Factor models reduce the dimensionality of the problem by describing the dependence structure of N asset returns using $K \ll N$ factors, which has the effect of reducing the number of parameters to be estimated. Factor models assume that asset returns are driven by a set of observed or latent factors typically constructed from a large number of variables using principal component analysis (PCA) and partial least squares (PLS). There is also a growing strand of literature that improves upon the traditional PCA estimates. For example, Kelly, Pruitt and Su (2019) propose instrumented PCA, where factors are latent, and the time-varying loadings depend on characteristics. Lettau and Pelger (2020) introduce risk-premia PCA, which identifies factors with small time-series variation that are useful in the cross-section of returns. Huang, Jiang, Li, Tong and Zhou (2022) suggest scaled PCA, which assigns greater weights to predictors with more forecasting power, by scaling each predictor with its slope on the response variable. These recent developments focus on forecasting the conditional mean or explaining the cross-section of asset returns. In contrast, we perform a comprehensive analysis of the ability of a variety of dimensionality reduction techniques and factor

[1] Alternative solutions involve imposing short-selling constraints (Hui, Kwan and Lee, 1993; Jagannathan and Ma, 2003), limiting turnover via norm constraints (DeMiguel, Garlappi, Nogales and Uppal, 2009) or penalizing the objective function (Olivares-Nadal and DeMiguel, 2018). Another approach uses either shrinkage estimators (Ledoit and Wolf, 2004; Ledoit and Wolf, 2017; Bodnar, Parolya and Schmid, 2018), which tend to shrink the covariance matrix towards a specific target covariance or sparse estimators that derive a regularized version of the precision matrix (Friedman, Hastie and Tibshirani, 2008). Using higher frequency data can also reduce estimation error (Jagannathan and Ma, 2003; Palczewski and Palczewski, 2014).

model specifications to accurately estimate the covariance matrix and whether they add value to minimum-variance portfolios.

Furthermore, while factor models are commonly used in finance, there are fewer studies that explore their benefits in portfolio optimization and specifically minimum-variance portfolios. Green and Hollifield (1992) and Chan, Karceski and Lakonishok (1999) show that introducing a factor structure to the covariance matrix can improve portfolio performance. Moskowitz (2003) examined the covariance structure of returns with respect to various factors and finds that the size factor can better explain covariance risk, both in and out of sample, while the book-to-market factor exhibits a weaker association and the momentum factor appears unrelated to return second moments. The benefits of using the factor model-based approach to estimate the covariance matrix have also been investigated by Fan, Fan and Lv (2008) and Fan, Liao and Mincheva (2011) who propose covariance estimators for exact and approximate factor models respectively. More recently, De Nard, Ledoit and Wolf (2021), use a factor framework and evaluate portfolios for different estimates of the error covariance matrix. Factor models have also been used by Lassance and Vrins (2021) and Lassance, DeMiguel and Vrins (2022), who extract uncorrelated risk factors based on independent component analysis to improve risk-parity and higher-moment strategies.

Latent factor models are appealing due to their capacity to combine information from a large number of variables in a simple and parsimonious way. Particularly, a drawback of PCA and PLS is that factor weights are non-zero, which leads to estimation difficulties in high dimensional settings. Another disadvantage of this framework is that it is confined to a linear relation between the variables. To address these issues, we apply a variety of dimensionality reduction methods from the machine learning literature to the estimation of factor-based covariance matrices in a portfolio allocation context.[2] The statistical and economic value of a variety of machine learning algorithms in finance has been examined by Krauss, Do, and Huck (2017), who generate profitable trading signals based on a classification framework, and Gu, Kelly and Xiu (2020) who formulate a regression problem to measure asset risk premiums and find that nonlinear methods lead to the best performance.[3] In contrast, our focus is on improving the estimates of factor-based covariance matrices through methods that produce modified latent factors with sparse weights, such that each latent factor is a linear combination of only a few of the original variables. In addition, we introduce non-linearities to the reduced representation of the variables, by constructing factors generated by autoencoders; a type of unsupervised neural network used for dimensionality reduction. Autoencoders have been employed in the recent literature

[2] Machine learning has been shown to be well suited to many theoretical and empirical problems in finance. A thorough survey of machine learning approaches used for optimization problems such as regression, classification, clustering, deep learning, and adversarial learning has been conducted by Gambella, Ghaddar and Naoum-Sawaya (2021).

[3] The asset universe of these studies is comprised of stocks, however, a comprehensive comparison of machine learning methodologies in terms of return prediction is provided by Bianchi, Büchner and Tamoni (2021), Sermpinis, Theofilatos, Karathanasopoulos, Georgopoulos and Dunis (2013), Wu, Chen, Yang and Tindall (2020) for bond, foreign exchange rate and hedge fund markets, respectively.

by Huck (2019) in portfolio management, while Gu, Kelly, and Xiu (2021) use autoencoders in an asset pricing setting. An additional contribution of this study is to conduct a systematic evaluation of static and dynamic specifications of the covariance matrix based on latent factors. Specifically, the structure of a dynamic covariance matrix can differ based on whether the factor loadings, the factor covariance matrix or the residual covariance matrix are allowed to vary over time.

We first investigate the predictive accuracy of the factor-based covariance matrices, which in the baseline case are derived using monthly observations for the largest 100 stocks from the CRSP database for the period from 1960 to 2022. As the proxy for the true covariance matrix, we consider the sample estimator based on 12-month ahead daily data and rely on several loss functions for the evaluation of the results. We also compare the performance of the factor models to several covariance benchmarks frequently employed by the literature, which include the sample estimator, the linear (Ledoit and Wolf, 2002) and non-linear (Ledoit and Wolf, 2017) shrinkage estimators and the Wishart stochastic covariance matrix (Moura, Santos and Ruiz, 2020). The results based on mean squared error and mean absolute error, which are symmetric loss functions, indicate that the majority of the factor-based covariance matrices outperform the covariance benchmarks, while according to two asymmetric loss functions the improved performance of the factor models is because they tend to overestimate the target covariance matrix less than the benchmarks.

The evaluation of the performance of the minimum-variance portfolios is conducted over the same sample. Specifically, the factor models lead to portfolios that significantly outperform the equally weighted and value weighted portfolio benchmarks and economically outperform portfolios based on the alternative optimal portfolio benchmarks. The best-performing methods to generate the covariance matrix are unsupervised learning methods, which can lead to portfolios that exhibit 25% higher risk-adjusted returns and 22% lower volatility than the equally weighted one or 10% and 4% respectively from the non-linear shrinkage estimator. Portfolios based on dimensionality reduction approaches also have weights that are smaller, vary less over time and are more diversified, than those based on the alternative covariance matrix benchmarks, with portfolios based on the static factor covariance specification and linear dimensionality reduction methods exhibiting lower turnover and thus reduced sensitivity to transaction costs. Overall, factor-based portfolios generate statistically significant outperformance after transaction costs with respect to standard deviation and Sharpe ratio relative to a series of benchmarks. When comparing the results across alternative specifications of the factor-based covariance matrix, the differences become less pronounced. Approaches based on unsupervised learning methods that allow the loadings or the residual covariance matrix to vary over time yield lower portfolio risk. However, after transaction costs are taken into account, strategies based on static factor covariance matrices outperform the dynamic specifications in terms of Sharpe ratio, which aligns with the conclusion by DeMiguel, Nogales and Uppal (2014) that dynamic strategies outperform static strategies only for transaction costs below ten basis points. The performance of the latent factor portfolios is amplified during periods of high volatility. Finally, the factor-based portfolios continue to outperform the equally weighted benchmark after increasing the number of assets, however, shrinkage estimators outperform factor-based estimators for a higher number of assets.

The remainder of this study is organized as follows. Section 2 describes our refined approach for imposing a factor structure to the covariance matrix, which involves constructing factors based on dimensionality reduction methods that induce sparsity or introduce non-linearities and different specifications for the covariance matrix that allow its components to vary over time. Section 3 provides details on the data, sample splitting and the approach to hyperparameter tuning that is based on economically motivated criteria. Section 4 examines the predictive accuracy of the covariance matrices and the economic value and properties of the factor-based minimum-variance portfolios. Section 5 concludes.

2. Methodology

In this section we introduce the methods for dimensionality reduction used to construct the latent factors, we then describe the different specifications under which the factor-based covariance matrices are estimated and finally, present the optimization framework used to derive the portfolios.

2.1. Factor Models

We consider models with factors that are latent quantities, which are derived from the data using dimensionality reduction techniques. When factors are latent, principal component analysis is a very common approach to reduce dimensionality. The studies of Chamberlain and Rothschild (1983) and Connor and Korajczyk (1988) are among the first to use latent factors in applications of the APT. A factor model for the returns of every asset, $r_{i,t}$, with $i = 1, \dots, N$ assets, $t = 1, \dots, T$ observations and $k = 1, \dots, K$ latent factors, takes the following general form

$$r_{i,t} = a_i + \beta_i(\mathrm{R}_t\mathrm{W}) + u_{i,t} = a_i + \beta_i\mathrm{F}_t + u_{i,t}, \tag{1}$$

where $\mathrm{R}_t = (r_{1,t}, \dots, r_{p,t})$ is the $T \times N$ matrix of asset returns and $\mathrm{W} = (w_1, \dots, w_K)$ is the $N \times K$ matrix of weights, with $K \ll N$, and $u_{i,t}$ is the error term for asset i at date t and $E(u_{i,t}|\mathrm{F}_t) = 0$. Each w_k is the vector of weights used to construct the k^{th} latent factor, f_k. The $T \times K$ matrix of latent factors is given by $\mathrm{F}_t = \mathrm{R}_t\mathrm{W}$, for factors $\mathrm{F}_t = (f_{t,1}, \dots, f_{t,K})$. The time-invariant factor loadings, $\beta_i = (\beta_{i,1}, \dots, \beta_{i,K})$, and the intercept, a_i, can be estimated by ordinary least squares (OLS) using the different factor representations.

Two commonly used dimensionality reduction techniques are principal component analysis and partial least squares. They are both designed to uncover a lower dimensional linear combination of the original dataset; however, the methods differ in the way the latent factor matrix, F_t, is extracted. PCA derives the latent factors in an unsupervised way, by producing the weight matrix W to reflect only the covariance structure between asset returns. In contrast, PLS derives the factors in a supervised way by constructing K linear combinations of R_t that have maximum correlation with the target.

The latent factors generated from PCA and PLS are linear combinations of all the original variables, with the elements of the weight matrix W being non-zero, which does not automatically lead to the selection of the most important variables to construct the factors. To address this issue, we consider methods that produce modified latent factors with sparse weights, such that each latent factor is a linear combination of only a few of the original variables. Specifically, we use sparse principal component analysis (SPCA) and sparse partial least squares (SPLS). Both methods impose a penalty based on the combination of the l_1 and l_2 norms allowing for the construction of sparse latent factors.

Finally, we construct latent factors using autoencoders which are a type of unsupervised neural network. Autoencoders are nonlinear generalizations of PCA. The goal of PCA and autoencoders is to learn a parsimonious representation of the original input data, R_t, through a bottleneck structure. The autoencoder behaves differently from PCA and SPCA, which reduce the dimensionality by mapping the original N inputs into $K \ll N$ factors in a linear way, while the autoencoder uses non-linear activation functions to discover non-linear representations of the data. To reduce estimation error, we use two types of autoencoders; sparse autoencoders (AEN) that add a penalty to the loss function and denoising autoencoders (DAE) that attempt to reconstruct the original dataset after it has been corrupted by random noise. We consider a shallow network with a single hidden layer.[4] Autoencoders have previously been used in a financial context by Huck (2019) as part of a prediction framework used to enhance the performance of statistical arbitrage strategies, while Gu, Kelly and Xiu (2021) propose a model for the cross-section of stock returns, where factors are latent, and the time-varying loadings depend on characteristics. Alternative network architectures that have been used in financial modeling and prediction include multilayer perceptrons (Gu, Kelly and Xiu, 2020) and recurrent neural networks (Fischer and Krauss, 2018). However, since our objective is to improve the estimates of the covariance matrix by reducing dimensionality through the creation of latent factors, it warrants the use of autoencoders as an appropriate solution for this task.

Further details on the machine learning approaches and related literature are provided in the Online Appendix.

2.2. Factor-based Covariance Matrices

After the factor model is estimated from equation (1) or (2) the covariance matrix of returns, Σ_r, is obtained by its decomposition into two components: the first is based on the factor loadings and the factor covariance matrix, while the second is the covariance matrix of the errors. The time-invariant covariance matrix of the returns $\mathrm{R} = (r_1, \dots, r_N)$ is given by:

[4] In unreported results we also examine the performance of portfolios of factors based on neural networks with two to four hidden layers. Similar to recent studies (e.g., Gu, Kelly and Xiu, 2020), the results indicate that shallow learning outperforms deeper learning. This decline in portfolio performance is potentially associated with the high degree of turnover of strategies based on autoencoders with more hidden layers. The results are available from the authors upon request.

$$\Sigma_r = \mathrm{B}'\Sigma_f\mathrm{B} + \Sigma_u, \quad (2)$$

where B is a $K \times N$ matrix with the ith column containing the vector of time-invariant factor loadings β_i and Σ_f and Σ_u denote the time-invariant covariance matrices of the factors and the errors respectively. We focus on exact factor models (EFM) by Fan, Fan and Lv (2008), where the covariance matrix of the residuals u_t is diagonal, $\Sigma_u \equiv \mathrm{diag}(\Sigma_u)$.

The models presented so far rely on a static factor covariance (SFC) specification. In this study we also consider dynamic factor models, which we define as a model that allows the factor loadings to be time varying (see e.g., Bali, Engle and Tang, 2017) or models in which either the factor or residual covariance matrix varies over time (Engle, Ng and Rothchild, 1990). A dynamic factor model is one in which at least one of the following three generalizations holds true: (i) the intercept and factor loadings are time-varying (dynamic beta covariance, DBC), (ii) the covariance matrix of the factors is time-varying (dynamic factor covariance, DFC) or (iii) the covariance matrix of the errors is time-varying (dynamic error covariance, DEC).

In the static case the betas of the assets remain constant over the estimation period. This assumption may not be plausible since betas typically vary over time. To this end we consider a time-varying estimator of the factor loadings. When the intercepts a_i and factor loadings β_i are allowed to be time-varying the conditional dynamic factor model takes the following form

$$r_{i,t} = a_{i,t} + \beta_{i,t}\mathrm{F}_t + u_{i,t}. \quad (3)$$

The estimates of the time-varying regression coefficients are then obtained by $\hat{\beta}_{i,t} = \Sigma_{f,t}^{-1}\sigma_{fr_i,t}$. The coefficients, $\hat{\beta}_{i,t}$, of this expression are the dynamic conditional betas and are based on time-varying estimates of the factor covariance matrix $\Sigma_{f,t}$ and the vector of covariances, $\sigma_{fr_i,t}$ between the returns of asset i, r_i and factor f_k, with $k = 1, \dots, K$. The intercept can be obtained by $\hat{a}_{i,t} = \bar{r}_i - \hat{\beta}_{i,t}\bar{\mathrm{F}}$. The time-varying covariance matrix of R_t is given by:

$$\Sigma_{r,t} = \mathrm{B}_t'\Sigma_f\mathrm{B}_t + \Sigma_u, \quad (4)$$

where B_t is a $K \times N$ matrix with the ith column containing the vector of time-varying factor loadings $\beta_{i,t}$.

The unconditional dynamic factor model under generalization (ii) and (iii) takes a form similar to equations (1) or (2), but with time-varying conditional covariance matrices for f_t and u_t respectively. If Σ_f is time-varying, then the covariance matrix of R_t is given by

$$\Sigma_{r,t} = \mathrm{B}'\Sigma_{f,t}\mathrm{B} + \Sigma_u. \quad (5)$$

Otherwise, if Σ_u is assumed to be time-varying, then

$$\Sigma_{r,t} = \mathrm{B}'\Sigma_f\mathrm{B} + \Sigma_{u,t}. \quad (6)$$

The factor covariance, $\Sigma_{f,t}$ is estimated by the dynamic conditional correlation (DCC) model (Engle, 2002) and the diagonal elements of $\Sigma_{u,t}$ are estimated by univariate GARCH models.

2.3. Minimum-Variance Portfolios

To investigate the economic value of the different estimates of the covariance matrix, $\hat{\Sigma}_r$, from the factor models we focus on the minimum-variance framework, which has frequently been used in the portfolio optimization literature (see e.g., Maillet, Tokpavi and Vaucher, 2015; Carroll, Conlon, Cotter and Salvador, 2017). Assuming there are N assets in the investment universe and $r_t = (r_{1,t}, \dots, r_{N,t})$ is a vector of asset returns, the portfolio objective functions we consider are

$$\underset{\omega}{\operatorname{argmin}}\, \omega' \hat{\Sigma}_r \omega, \quad \text{s.t.} \quad \omega' \mathrm{i}_N = 1, \ \ \omega_i \geq 0, \tag{7}$$

$$\underset{\omega}{\operatorname{argmin}}\, \omega' \hat{\Sigma}_r \omega, \quad \text{s.t.} \quad \omega' \mathrm{i}_N = 1, \tag{8}$$

$$\underset{\omega}{\operatorname{argmin}}\, \omega' \hat{\Sigma}_r \omega + \kappa \|\omega - \omega_0\|_1, \quad \text{s.t.} \quad \omega' \mathrm{i}_N = 1, \tag{9}$$

for $i = 1, \dots, N$, where $\omega = (\omega_1, \dots, \omega_N)$ is the portfolio weight vector and i_N is a $N \times 1$ unit vector. The return of the portfolio can then be calculated as $r_{p,t+1} = \hat{\omega}' r_{t+1}$. All portfolios include a leverage constraint, by imposing that the sum of the weights is equal to unity. In the baseline case we consider minimum-variance portfolios with short-selling constraints (Equation 7), by setting the lower bound of the portfolio weights to zero. The additional non-negativity constraint on minimum variance portfolios has been shown (Jagannathan and Ma, 2003) to be equivalent to shrinking the elements of the covariance matrix. We also consider global minimum-variance portfolios (Equation 8), where short-selling is allowed. Finally, we consider minimum-variance portfolios (Equation 9), that explicitly take account of transaction costs during the portfolio formation process (Olivares-Nadal and DeMiguel, 2018) by adding a penalization term based on the portfolio turnover to the portfolio's objective function. Specifically, $\kappa = 10$ bps is the transaction cost parameter that controls for the degree to which portfolio turnover is penalized and ω_0 are the weights of the portfolio from the previous period before rebalancing.

2.4. Benchmark Models

We consider several alternative strategies, whose performance is compared to that of factor-based allocations. The equally weighted portfolio (EW), with weights $\omega_i = 1/N$, for $i = 1, \dots, N$, while another scheme that requires no parameter estimation is the value-weighted (VW) portfolio, whose weights are based on the average market capitalization over the formation period.[5] The remaining benchmarks are minimum-variance portfolios relying on four different estimators of the covariance matrix. We employ the sample covariance (Sample) estimator, the linear shrinkage (LINS) covariance

---

[5] DeMiguel, Garlappi and Uppal (2009) and Bianchi and Guidolin (2014) show the equally weighted portfolio to be a very stringent benchmark to outperform. In contrast, Platanakis, Sutcliffe and Ye (2021) show that while EW is preferable for stock selection, optimal portfolios can be beneficial for asset allocation, which takes place in smaller dimensions.

by Ledoit and Wolf (2002), the non-linear shrinkage (NLS) covariance by Ledoit and Wolf (2017) and the Wishart stochastic covariance (Wishart) by Moura, Santos and Ruiz (2020).

3. Data and Sample Splitting

The data set consists of monthly total individual stock returns from the Center for Research in Security Prices (CRSP) starting on January 1960 to December 2022 or $T = 756$ monthly observations. Our approach regarding the backtest and the restrictions we impose on the data set is similar to that of Ledoit and Wolf (2017) and De Nard, Ledoit and Wolf (2019), but adapted to a monthly frequency. We restrict our data set to stocks listed on the NYSE, AMEX, and NASDAQ stock exchanges, to ordinary common shares whose price is greater than $1.

We adopt a rolling window approach to examine the out-of-sample (OOS) economic performance of our models. The size of the rolling window is set to $T_0 = 240$ monthly observations, with the initial window spanning the period from January 1960 to December 1979. The rolling window moves across the full sample by one monthly observation at a time, leading to an out-of-sample size of $T_{OOS} = T - T_0 = 505$ monthly observations, from January 1980 to January 2022. The portfolios are constructed in each iteration of the rolling window, based on stocks that have at least 97.5% history of returns available in the rolling window (missing values are replaced by the mean of the series) and are also not missing the return observation for the following month after the end of the rolling window. This forward-looking restriction is commonly applied to allow for the out-of-sample evaluation of portfolios, which are based on in-sample estimates of the covariance matrix.

In the baseline case, the latent factors, covariance matrices and portfolio weights are estimated based on the $N = 100$ stocks with the highest market capitalization within each iteration of the rolling window, before we expand the analysis to larger portfolios. In each iteration of the rolling window, we cross-sectionally transform the asset returns, R_t. Specifically, we calculate the rank of a stock based on the return and then divide the ranks by the number of observations and subtract 0.5 to map the features into the $[-0.5,0.5]$ interval. This rank-transformation focuses on the ordering of the data and is insensitive to outliers and has been applied in several studies, for example, Gu, Kelly and Xiu (2020).

The dimensionality reduction approaches used to derive the latent factors rely on hyperparameter tuning. The choice of hyperparameters controls the amount of model complexity and is critical for the performance of the model. We adopt the validation sample approach to select the optimal set of hyperparameters. Further details on the hyperparameter tuning can be found in Appendix 1.

4. Empirical Results

4.1. Forecast Evaluation

The true covariance matrix S_t is unobservable, and therefore the predictive accuracy of the models has to be measured with respect to some ex-post estimator, $\hat{\mathrm{S}}_t$. In each iteration, we approximate S_t using the sample covariance matrix based on daily returns from the following 12-month period, $t + 1, \ldots, t + 12$. We use the estimated target covariance to compare the predictive performance of the alternative covariance matrices based on four loss functions previously employed by Laurent, Rombouts and Violante (2013) and Becker, Clements, Doolan and Hurn (2015). First, we consider two symmetric loss functions that do not penalize differently underpredictions and overpredictions. The mean squared error (MSE) is a symmetric loss that is derived as the mean squared distance between the covariance forecast, $\hat{\Sigma}_t$, and the target covariance proxy, $\hat{S}_t$: $\mathrm{MSE} = 1/N^2 \operatorname{vec}(\hat{\Sigma}_t - \hat{S}_t)' \operatorname{vec}(\hat{\Sigma}_t - \hat{S}_t)$, where $\operatorname{vec}(\cdot)$ represents the column stacking operator. As an alternative to squared errors, absolute errors can be measured using the mean absolute error (MAE) loss, computed as: $\mathrm{MAE} = 1/N^2 \mathrm{i}' \operatorname{abs}(\hat{\Sigma}_t - \hat{S}_t)\, \mathrm{i}$, where $\operatorname{abs}(\cdot)$ is the absolute operator. The next two loss functions we consider are asymmetric with respect to under and over predictions. The quasi-likelihood function (QLK): $\mathrm{QLK} = \log|\hat{\Sigma}_t| + \mathrm{i}'(\hat{\Sigma}_t \odot \hat{S}_t)\mathrm{i}$ heavily penalizes under predictions, while a loss function that penalizes over predictions instead is: $\mathrm{ASYM} = 1/[b(b-1)]\operatorname{tr}(\hat{S}_t^b - \hat{\Sigma}_t^b) - 1/(b-1)\operatorname{tr}[\hat{\Sigma}_t^{b-1}(\hat{\Sigma}_t - \hat{S}_t)]$, where $\operatorname{tr}(\cdot)$ is the trace operator and following Laurent, Rombouts and Violante (2013) we set $b = 3$, indicating a mild degree of asymmetry. For all measures, a lower value is preferable. We also examine whether the alternative covariance matrix forecasts are statistically significantly different relative to the linear shrinkage covariance benchmark. The two-sided p-value is adjusted for autocorrelation up to 12-month lags. The results for the predictive accuracy of the covariance matrices are reported in Table 1.

Focusing on the results for the MSE and MAE loss functions (Panel 1), we observe that the majority of the factor-based covariance matrices generate lower values for both measures compared to the remaining covariance matrices and yield results that are statistically significant at the 5% or 1% level relative to the LINS benchmark. Specifically, unsupervised learning methods (PCA, SPCA and autoencoders), are better at predicting the 12-month ahead covariance matrix than supervised methods (PLS and SPLS). The covariance matrices that generate the lowest MSE and MAE values are those based on the static factor covariance specification or those that allow the factor covariance or error covariance to vary over time, with the DEC based on AEN factors being the only case of a factor model yielding higher MSE than the covariance benchmarks.

Turning to the predictive performance according to the QLK and ASYM loss functions (Panel B), the findings suggest that the predictive gains exhibited by the factor-based covariances in terms of MSE and MAE can be attributed to the lower degree of over predictions, evidenced by the lower ASYM values of the factor models compared to the covariance matrix benchmarks. The SFC and DFC specifications, in addition to a dynamic beta covariance based on PCA and dynamic error covariances based on linear dimensionality reduction approaches, yield results that are significantly different than those of the LINS in terms of the ASYM loss, while none of the remaining covariance benchmarks are significant. In contrast, the sample covariance, the shrinkage estimators and the Wishart stochastic covariance yield lower QLK values than the majority of the factor-based covariances, however, the results for the Sample, NLS and Wishart covariances are not statistically significant.

Table 1 Forecast evaluation of the covariance matrices of exact factor models

This table reports the ability of the alternative covariance matrices to predict the out-of-sample realized covariance matrix based on four loss functions. Panel A reports the mean squared error (MSE) and mean absolute error (MAE), whereas the quasi-likelihood function (QLK) and asymmetric loss function (ASYM) can be found in Panel B. The average value of each measure over the out-of-sample period from January 1980 to December 2022 is reported. The results are presented for the sample estimator (Sample), linear shrinkage (LINS) and non-linear shrinkage (NLS) estimators, Wishart stochastic covariance (Wishart) and for four exact factor covariance specifications: static factor covariance (SFC), dynamic factor covariance (DFC), dynamic beta covariance (DBC) and dynamic error covariance (DEC). The factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS), autoencoder (AEN) and denoising autoencoder (DAE). The statistical significance of the alternative covariance matrices compared to the LINS covariance benchmark is denoted by: *, **, and *** for significance at the 10%, 5%, and 1% level, respectively.

Panel A Mean squared error and mean absolute error

	MSE	MAE
Sample	51.666	18.777
LINS	51.972	19.139
NLS	48.217	18.306
Wishart	54.911	19.143

	SFC		DFC		DBC		DEC	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
PCA	14.900**	6.540***	16.752***	6.989***	19.561***	6.397***	17.381***	6.384***
PLS	18.471**	8.071**	20.665***	8.353***	26.225**	7.941***	20.301***	7.926**
SPCA	14.639**	6.463**	15.977***	6.806***	23.382***	6.309***	16.802***	6.308**
SPLS	18.394**	8.022**	20.645***	8.317***	26.060**	7.883***	21.000***	7.878**
AEN	14.729**	6.467**	17.828***	7.201**	19.708***	6.184**	71.469	6.316**
DAE	14.730**	6.492*	17.550***	7.181**	19.337***	6.240**	24.374***	6.336**

Panel B Quasi-likelihood function and asymmetric loss function

	QLK	ASYM
Sample	21.810	3.488
LINS	23.599	3.558
NLS	24.965	3.247
Wishart	23.328	3.818

	SFC		DFC		DBC		DEC	
	QLK	ASYM	QLK	ASYM	QLK	ASYM	QLK	ASYM
PCA	29.000	0.444**	29.004	0.628***	29.178**	1.433**	24.086	0.907***
PLS	28.677*	0.663**	28.675*	0.984***	28.862***	2.685	23.868	0.942***
SPCA	29.022	0.422**	29.023	0.556***	29.187*	4.406	24.136	0.850***
SPLS	28.697*	0.660**	28.696*	0.977***	28.879***	2.698	23.887	1.140***
AEN	29.004	0.434**	29.013	0.715**	29.190	1.667	24.053	98.997
DAE	28.984*	0.430**	28.992*	0.687***	29.174**	1.512*	24.046	6.638

4.2. Portfolio Performance

To examine the economic value of factor-based covariance matrices we rely on portfolio objective functions that are designed to minimize variance rather than maximize the expected return. Therefore, similar to Ledoit and Wolf (2017) and De Nard, Ledoit and Wolf (2019), we primarily compare the economic value of the alternative covariance matrices using the standard deviation, followed by the

Sharpe ratio.[6] In Table 2 we report the monthly performance of the portfolios over the out-of-sample period, T_{OOS}, based on the standard deviation (SD) of the 505 out-of-sample portfolio returns in excess of the risk-free rate and the Sharpe ratio (SR) of the portfolio calculated as $(\bar{r}_p - \bar{r}_f)/\text{SD}$, where $\bar{r}_p$ is the average value of the portfolio returns and $\bar{r}_f$ is the average value of the risk-free rate.[7]

Table 2 Portfolio performance of exact factor models based on standard deviation and Sharpe ratio
This table documents monthly portfolio performance measured using the standard deviation (SD) and Sharpe ratio (SR), over the out-of-sample period from January 1980 to December 2022. The results are presented for the equally weighted portfolio (EW), value-weighted portfolio (VW) and minimum-variance portfolios with short-selling constraints based on the sample estimator (Sample), linear shrinkage (LINS) and non-linear shrinkage (NLS) estimators, Wishart stochastic covariance (Wishart) and for four exact factor model (EFM) covariance specifications: static factor covariance (SFC), dynamic factor covariance (DFC), dynamic beta covariance (DBC) and dynamic error covariance (DEC). The factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS), autoencoder (AEN) and denoising autoencoder (DAE).

	SD	SR
EW	4.275	0.196
VW	4.137	0.188
Sample	3.501	0.212
LINS	3.510	0.218
NLS	3.450	0.219
Wishart	3.736	0.210

	SFC		DFC		DBC		DEC	
	SD	SR	SD	SR	SD	SR	SD	SR
PCA	3.413	0.241	3.412	0.241	3.347	0.245	3.323	0.243
PLS	3.422	0.229	3.429	0.228	3.348	0.233	3.320	0.241
SPCA	3.424	0.241	3.427	0.240	3.346	0.246	3.340	0.243
SPLS	3.425	0.230	3.431	0.228	3.352	0.233	3.306	0.241
AEN	3.404	0.243	3.401	0.243	3.343	0.249	3.329	0.244
DAE	3.408	0.241	3.405	0.241	3.362	0.245	3.317	0.244

The results indicate that optimal portfolios consistently outperform the EW and VW portfolios in terms of standard deviation and Sharpe ratio by a wide margin. The VW portfolio outperforms the EW portfolio in terms of risk but not Sharpe ratio. From the benchmark covariance matrices, NLS has the lowest SD and highest Sharpe ratio, of 3.45% and 0.219 respectively, indicating it is the hardest benchmark to outperform. Using a factor-based covariance matrix, can lead to a decrease in out-of-sample standard deviation of up to 22% and an increase in Sharpe ratio of over 25% relative to the EW portfolio. Compared to the NLS estimator, using a factor model can lead to a decrease in portfolio risk of up to 4% and an increase in Sharpe ratio of over 10%. Factors based on unsupervised methods (PCA, SPCA, AEN and DAE) are found to yield the best performance. For the SFC, DFC and DBC the autoencoder yields the lowest standard deviation, 3.343% to 3.404% depending on the specification,

[6] The standard deviation and Sharpe ratio for the global minimum-variance portfolio and turnover-constrained minimum-variance portfolio are reported in Tables A1 and A2 in the Online Appendix.

[7] Hwang, Xu, and In (2018) emphasize the importance of tail risk when comparing the performance of optimal strategies with that of the $1/N$ rule. Therefore, we examine the portfolio performance using several alternative risk measures, including the value-at-risk (VaR) and conditional value-at-risk (CVaR) and the ratios of portfolio excess returns with VaR or CVaR. The results are reported in Table A3 in the Appendix.

except for DEC where SPLS yields the lowest standard deviation with a value of 3.306%. The two autoencoders have the highest Sharpe ratio, with values between 0.243 and 0.249. AEN has the highest Sharpe ratio in the dynamic beta covariance specification, while SPCA is the best performing model for the remaining specifications. Overall, the best performing portfolios are based on the dynamic error or dynamic beta covariance specifications, while portfolios based on static or dynamic factor covariance generate comparable performance.

4.3. Properties of Portfolio Weights

In this Section we explore how the weighing structure of the portfolios differs across different estimates of the covariance matrix. We start by analyzing the properties of the portfolio weights, $\widehat{\omega}$, using the minimum non-zero weight (MIN), maximum weight (MAX), the standard deviation of the portfolio weights (SD) and in line with DeMiguel, Garlappi and Uppal (2009), we report the average monthly portfolio turnover (TO) computed as the average absolute change of the portfolio weights over the T_{OOS} rebalancing periods across the N assets. The turnover at time $t+1$ is given by $\|\omega_{t+1} - \omega_t\|_1$, where ω_{t+1} is the vector of portfolio weights at time $t+1$ and ω_t are the portfolio weights at the time before rebalancing. Furthermore, we examine the concentration of the portfolio using the Herfindahl-Hirschman index (HHI) computed as $\sum_{i=1}^{N} \widehat{\omega}_i^2$, with a lower HHI implying a more diversified portfolio. Finally, we report the percentage of non-zero weights (NZ). Table 3 reports the average value of each weight characteristic over the out-of-sample period.[8]

Overall, portfolios based on latent factors are more diversified and tend to produce weights which are smaller and less volatile than portfolios based on the covariance benchmarks. Specifically, the value of maximum weight varies between 14.1% and 21.4% for the sample, shrinkage and Wishart covariance matrices, while for the latent factor models the value of the maximum weight is from 4.8% to 8.6%, with portfolios based on the dynamic error covariance specification generating higher weights. Furthermore, portfolios based on the static factor covariance or the dynamic factor covariance specifications have the lowest weight standard deviation. Turnover varies depending on the type of covariance specification and factor considered. The lowest turnover is produced by portfolios based on PLS and SPLS factors for a static factor covariance specification (approximately 4%). Comparing across different covariance specifications, portfolios based on the DBC, DEC and autoencoders exhibit higher turnover than the remaining factor specifications. Finally, the Herfindahl-Hirschman index and percentage of non-zero weights measures indicate that strategies based on latent factors generated from unsupervised learning methods are less concentrated on a small number of stocks than other minimum-variance portfolios.

[8] The properties of the portfolio weight vectors for minimum-variance portfolios that allow short-selling and portfolios with a turnover penalty are reported in Tables A4 and A5 respectively, found in the Online Appendix.

Table 3 Characteristics of the portfolio weight vectors of exact factor models

This table presents the monthly characteristics of the portfolio weight vectors. Panel A reports the minimum non-zero weight (MIN), maximum weight (MAX) and standard deviation of the weights (SD), whereas the portfolios turnover (TO), Herfindahl-Hirschman index (HHI) and percentage of non-zero weights (NZ) can be found in Panel B. The average value of each weight characteristic over the out-of-sample period from January 1980 to December 2022 is reported. MIN, MAX, TO and NZ are reported as a percentage. The results are presented for the equally weighted portfolio (EW), value-weighted portfolio (VW) and minimum-variance portfolios with short-selling constraints based on the sample estimator (Sample), linear shrinkage (LINS) and non-linear shrinkage (NLS) estimators, Wishart stochastic covariance (Wishart) and for four exact factor model (EFM) covariance specifications: static factor covariance (SFC), dynamic factor covariance (DFC), dynamic beta covariance (DBC) and dynamic error covariance (DEC). The factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS), autoencoder (AEN) and denoising autoencoder (DAE).

Panel A Minimum non-zero weight, maximum weight and standard deviation of the weights

	MIN	MAX	SD
EW	1.000	1.000	0.346
VW	0.310	6.710	0.307
Sample	0.186	21.414	0.595
LINS	0.173	16.893	0.559
NLS	0.165	14.160	0.525
Wishart	0.205	19.486	0.641

	SFC			DFC			DBC			DEC		
	MIN	MAX	SD	MIN	MAX	SD	MIN	MAX	SD	MIN	MAX	SD
PCA	0.028	5.022	0.389	0.028	5.048	0.390	0.032	4.914	0.404	0.020	6.766	0.457
PLS	0.029	6.411	0.391	0.027	6.430	0.392	0.033	6.034	0.420	0.024	8.603	0.452
SPCA	0.036	5.169	0.387	0.038	5.193	0.389	0.034	5.041	0.399	0.026	6.974	0.455
SPLS	0.031	6.416	0.388	0.029	6.435	0.390	0.033	6.019	0.418	0.026	8.560	0.449
AEN	0.034	5.023	0.392	0.037	5.091	0.396	0.040	4.814	0.406	0.027	6.865	0.463
DAE	0.031	4.977	0.390	0.032	5.034	0.393	0.039	4.806	0.407	0.027	6.863	0.460

Panel B Portfolio turnover, Herfindahl-Hirschman index and percentage of non-zero weights

	TO	HHI	NZ
EW	1.030	0.01	100
VW	0.880	4.288	100
Sample	8.247	21.127	8.109
LINS	8.004	16.515	9.307
NLS	7.957	13.366	10.233
Wishart	10.575	19.358	8.342

	SFC			DFC			DBC			DEC		
	TO	HHI	NZ	TO	HHI	NZ	TO	HHI	NZ	TO	HHI	NZ
PCA	5.929	4.358	30.012	6.178	4.394	29.802	13.125	4.275	30.448	21.800	5.405	29.002
PLS	3.933	5.768	23.582	4.090	5.793	23.579	12.263	5.470	24.993	20.156	7.392	22.272
SPCA	9.260	4.426	29.966	9.439	4.456	29.903	15.551	4.350	30.326	24.424	5.489	29.119
SPLS	4.037	5.754	23.715	4.198	5.781	23.679	12.515	5.455	24.997	20.142	7.364	22.300
AEN	12.674	4.353	29.935	13.253	4.430	29.634	17.708	4.192	30.667	26.823	5.431	28.906
DAE	10.795	4.342	29.903	11.168	4.408	29.643	16.787	4.209	30.569	25.247	5.420	28.900

4.4. Portfolio Performance After Transaction Costs

The portfolio's return is modified to account for transaction costs based on portfolio turnover. Given a transaction cost level of c, the trading cost of the entire portfolio is $c\|\omega_{t+1} - \omega_t\|_1$. The return of the portfolio after transaction costs becomes $r_{p,t+1}^{TC} = \left(1 + r_{p,t+1}\right)(1 - c\|\omega_{t+1} - \omega_t\|_1) - 1$. Portfolio performance after transaction costs is reported in Table 4 for transaction costs of $c = 5$ bps and 20 bps.[9]

[9] The performance after transaction costs for minimum-variance portfolios that allow short-selling and portfolios with a turnover penalty are reported in Tables A6 and A7 respectively in the Online Appendix.

Table 4 Portfolio performance of exact factor models after transaction costs

This table presents monthly portfolio performance measured using the standard deviation (SD) and Sharpe ratio (SR), after transaction costs are taken into account. In this setting transaction costs would arise from changes to the stock universe from one month to the next and from the change in weights of stocks that remain in the stock universe for multiple iterations. The portfolio's return is modified to account for transaction costs based on portfolio turnover. Panel A reports the results for transaction costs of c=5 bps, while Panel B presents the results for transaction costs of c=20 bps. The out-of-sample period is from January 1980 to December 2022. The results are presented for the equally weighted portfolio (EW), value-weighted portfolio (VW) and minimum-variance portfolios with short-selling constraints based on the sample estimator (Sample), linear shrinkage (LINS) and non-linear shrinkage (NLS) estimators, Wishart stochastic covariance (Wishart) and for four exact factor model (EFM) covariance specifications: static factor covariance (SFC), dynamic factor covariance (DFC), dynamic beta covariance (DBC) and dynamic error covariance (DEC). The factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS), autoencoder (AEN) and denoising autoencoder (DAE).

Panel A Transaction costs of c = 5 bps

	SD	SR
EW	4.275	0.196
VW	4.137	0.188
Sample	3.502	0.211
LINS	3.510	0.217
NLS	3.450	0.218
Wishart	3.736	0.209

	SFC		DFC		DBC		DEC	
	SD	SR	SD	SR	SD	SR	SD	SR
PCA	3.413	0.240	3.412	0.240	3.347	0.243	3.322	0.239
PLS	3.422	0.229	3.429	0.227	3.348	0.231	3.320	0.238
SPCA	3.424	0.240	3.427	0.239	3.346	0.244	3.340	0.240
SPLS	3.425	0.229	3.431	0.227	3.352	0.231	3.306	0.237
AEN	3.404	0.242	3.400	0.241	3.343	0.246	3.329	0.240
DAE	3.407	0.239	3.405	0.239	3.362	0.243	3.317	0.240

Panel B Transaction costs of c = 20 bps

	SD	SR
EW	4.275	0.196
VW	4.137	0.188
Sample	3.502	0.208
LINS	3.511	0.213
NLS	3.451	0.215
Wishart	3.738	0.204

	SFC		DFC		DBC		DEC	
	SD	SR	SD	SR	SD	SR	SD	SR
PCA	3.413	0.237	3.412	0.237	3.346	0.237	3.322	0.229
PLS	3.423	0.227	3.429	0.225	3.347	0.226	3.319	0.229
SPCA	3.425	0.236	3.428	0.235	3.346	0.237	3.341	0.229
SPLS	3.425	0.227	3.431	0.225	3.352	0.225	3.305	0.228
AEN	3.403	0.236	3.400	0.235	3.342	0.238	3.328	0.228
DAE	3.407	0.234	3.405	0.234	3.361	0.235	3.316	0.229

The performance in terms of out-of-sample standard deviation after transaction costs of 5 or 20 bps remains qualitatively unchanged from the no-transaction cost case. Optimal strategies consistently outperform the EW and VW allocation approaches and factor-based portfolios exhibit similar or lower risk than the portfolio based on the NLS estimator. In terms of Sharpe ratio, outperformance of the optimal strategies relative to the EW and VW, while still relevant, is diminished when transaction costs are introduced. The strategies that are most affected are those that exhibit the highest turnover, which include portfolios based on the Wishart covariance, non-linear factors or those that allow the betas or error covariance to be time-varying. Overall, while dynamic strategies result in lower portfolio risk, their performance in terms of Sharpe ratio after transaction costs are taken into account is worse than that based on static specifications.

4.5. Statistical Significance of Portfolio Performance

We also consider the question whether one portfolio delivers improved out-of-sample performance relative to another portfolio at a level that is statistically significant. DeMiguel, Garlappi, and Uppal (2009) provide persuasive evidence that the simple equally weighted portfolio should serve as a natural benchmark to assess the performance of more sophisticated strategies. We also consider seven alternative benchmarks, namely the VW portfolio, and optimal allocations based on the Sample, LINS, NLS, Wishart covariance matrices, as well as minimum-variance portfolios based on the static factor covariance specification using PCA and PLS latent factors (PCA-SFC and PLS-SFC). For each case, the test-statistics and two-sided p-values are obtained by the Ledoit and Wolf (2011) test for the null hypothesis of equal standard deviations and by the Opdyke (2007) test for the null hypothesis of equal Sharpe ratios. The test statistics based on portfolio returns after transaction costs of 20 bps are reported in Table 5, with a positive value indicating economic outperformance for the respective measure. [10]

Overall, factor-based portfolios provide statistical outperformance relative to the series of benchmarks proposed after transaction costs. When the EW is the benchmark the results for the standard deviation indicate a significance at the 1% level for all strategies, however, significant outperformance for the Sharpe ratios is observed only for factor-based allocations. Turning to the results for when the VW portfolio is the benchmark, we observe that the optimal strategies outperform the value-weighted portfolio in terms of risk at the 1% significance level, while in terms of Sharpe ratio the two shrinkage estimators exhibit significant outperformance at the 5% level and the majority of the factor-based allocations significantly outperform the VW benchmark at the 1% level. For the sample estimator benchmark, portfolios based on the NLS estimator and the DBC and DEC factor specifications lead to significant outperformance at the 5% or 10% level for standard deviation, while in terms of Sharpe ratio all factor-based portfolios are significant at the 1% level. The results for the LINS estimator are similar to those of Sample, with portfolios based on the NLS covariance matrix outperforming at the 1% level and those based on specifications that allow the factor loadings or error covariance to vary over time showing outperformance at the 5% level. In contrast, the NLS estimator is a more difficult benchmark to significantly outperform in terms of risk, with seven portfolios based on either DBC or DEC specifications showing significant results at the 10% level. Comparatively, when the Wishart covariance is the benchmark, all optimal strategies have statistically significant results for standard deviation at the 1% level, with latent factor models generating significant Sharpe ratio against the benchmark at the 1%

[10] The statistical significance of the performance for minimum-variance portfolios that allow short-selling and portfolios with a turnover penalty are reported in Tables A8 and A9 respectively in the Online Appendix.

level. Finally, setting the benchmark to be a static factor model based on PCA or PLS latent factors, the strategies that yield significant standard deviation are those based on DBC and DEC specifications, while only a few strategies based on unsupervised dimensionality reduction methods and the DBC specification lead to significant Sharpe ratios at the 1% level.

Table 5 Statistical evaluation of the portfolio standard deviation and Sharpe ratio after transaction costs of 20 bps of exact factor models

This table reports the test statistics for the standard deviation and Sharpe ratio of the alternative strategies against several benchmarks. The out-of-sample period is from January 1980 to December 2022. The out-of-sample period is from January 1980 to December 2022. A positive test statistic indicates economic outperformance of the alternative portfolio (rows) against the benchmark (columns) for the respective performance measure. The results are presented for the equally weighted portfolio (EW), value-weighted portfolio (VW) and minimum-variance portfolios with short-selling constraints based on the sample estimator (Sample), linear shrinkage (LINS) and non-linear shrinkage (NLS) estimators, Wishart stochastic covariance (Wishart) and for four exact factor model (EFM) covariance specifications: static factor covariance (SFC), dynamic factor covariance (DFC), dynamic beta covariance (DBC) and dynamic error covariance (DEC). The factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS), autoencoder (AEN) and denoising autoencoder (DAE). The statistical significance of the alternative strategies compared to the benchmark strategy is denoted by: *, **, and *** for significance at the 10%, 5%, and 1% level, respectively.

	EW		VW		Sample		LINS	
	SD	SR	SD	SR	SD	SR	SD	SR
EW	-	-	-3.580***	0.183	-5.758***	-0.264	-6.098***	-0.395
VW	3.580***	-0.183	-	-	-4.712***	-0.447	-4.998***	-0.578
Sample	5.758***	0.264	4.712***	0.447	-	-	0.373	-0.131
LINS	6.098***	0.395	4.998***	0.578**	-0.373	0.131	-	-
NLS	6.603***	0.419	5.463***	0.602**	2.182**	0.155	3.201***	0.024
Wishart	4.163***	0.188	3.098***	0.371	-4.711***	-0.077	-5.077***	-0.207
Static Factor Covariance								
PCA	9.025***	0.930***	7.577***	1.113***	1.103	0.666***	1.290	0.535***
PLS	7.220***	0.702**	5.934***	0.885***	1.053	0.438***	1.214	0.307***
SPCA	8.797***	0.895***	7.329***	1.078***	0.933	0.631***	1.110	0.500***
SPLS	7.127***	0.706**	5.845***	0.889***	1.002	0.442***	1.152	0.311***
AEN	8.829***	0.902***	7.459***	1.085***	1.241	0.638***	1.427	0.507***
DAE	9.250***	0.867***	7.754***	1.050***	1.197	0.603***	1.391	0.472***
Dynamic Factor Covariance								
PCA	9.032***	0.932***	7.603***	1.115***	1.104	0.668***	1.295	0.537***
PLS	7.213***	0.661**	5.936***	0.844***	0.962	0.397***	1.125	0.266**
SPCA	8.792***	0.872***	7.329***	1.055***	0.889	0.608***	1.064	0.477***
SPLS	7.103***	0.656**	5.835***	0.839***	0.914	0.392***	1.067	0.261**
AEN	8.820***	0.879***	7.466***	1.062***	1.268	0.614***	1.458	0.484***
DAE	9.184***	0.863***	7.705***	1.046***	1.212	0.598***	1.410	0.467***
Dynamic Beta Covariance								
PCA	8.806***	0.917***	7.072***	1.100***	2.143**	0.652***	2.296**	0.521***
PLS	7.442***	0.671**	6.023***	0.854***	2.196**	0.407***	2.295**	0.276***
SPCA	8.876***	0.92***	7.147***	1.103***	2.165**	0.656***	2.330**	0.525***
SPLS	7.268***	0.656**	5.882***	0.839**	2.072**	0.392***	2.161**	0.261**
AEN	8.693***	0.955***	6.972***	1.138***	2.204**	0.691***	2.332**	0.560***
DAE	8.789***	0.885***	6.979***	1.068***	1.915*	0.621***	2.060**	0.490***
Dynamic Error Covariance								
PCA	8.613***	0.752***	7.071***	0.935***	2.252**	0.488***	2.417**	0.357***
PLS	7.205***	0.749**	5.944***	0.932**	2.216**	0.485***	2.352**	0.354**
SPCA	8.463***	0.733***	6.975***	0.916***	2.052**	0.469***	2.18**	0.338***
SPLS	7.191***	0.729**	5.953***	0.912**	2.367**	0.465***	2.48**	0.334**
AEN	8.442***	0.712***	6.951***	0.895***	2.170**	0.448***	2.328**	0.317**
DAE	8.724***	0.743***	7.208***	0.926***	2.333**	0.479***	2.504**	0.348***

Table 5 (Continued)								
	NLS		Wishart		PCA-SFC		PLS-SFC	
	SD	SR	SD	SR	SD	SR	SD	SR
EW	-6.603***	-0.419	-4.163***	-0.188	-9.025***	-0.930	-7.220***	-0.702
VW	-5.463***	-0.602	-3.098***	-0.371	-7.577***	-1.113	-5.934***	-0.885
Sample	-2.182**	-0.155	4.711***	0.077	-1.103	-0.666	-1.053	-0.438
LINS	-3.201***	-0.024	5.077***	0.207	-1.290	-0.535	-1.214	-0.307
NLS	-	-	6.343***	0.232	-0.535	-0.511	-0.427	-0.283
Wishart	-6.343***	-0.232	-	-	-4.095***	-0.742	-3.923***	-0.515
Static Factor Covariance								
PCA	0.535	0.511***	4.095***	0.742***	-	-	0.206	0.228
PLS	0.427	0.283***	3.923***	0.515***	-0.206	-0.228	-	-
SPCA	0.355	0.475***	3.916***	0.707***	-0.706	-0.035	-0.060	0.192
SPLS	0.382	0.287***	3.848***	0.519***	-0.253	-0.224	-0.606	0.004
AEN	0.682	0.483***	4.223***	0.715***	0.799	-0.028	0.457	0.200
DAE	0.634	0.448***	4.182***	0.680***	0.316	-0.063	0.330	0.165
Dynamic Factor Covariance								
PCA	0.546	0.513***	4.083***	0.744***	0.459	0.002	0.228	0.230
PLS	0.334	0.242***	3.828***	0.473***	-0.342	-0.269	-1.898*	-0.041
SPCA	0.312	0.453***	3.866***	0.685***	-0.862	-0.058	-0.123	0.170
SPLS	0.290	0.237***	3.757***	0.469***	-0.389	-0.274	-1.713*	-0.046
AEN	0.721	0.459***	4.243***	0.691***	1.023	-0.051	0.523	0.176
DAE	0.659	0.443***	4.182***	0.675***	0.406	-0.067	0.370	0.160
Dynamic Beta Covariance								
PCA	1.620	0.497***	5.145***	0.729***	2.009**	-0.013	1.496	0.214***
PLS	1.631	0.252***	5.048***	0.483***	1.260	-0.259	2.529**	-0.031
SPCA	1.643	0.501***	5.140***	0.733***	1.988**	-0.010	1.632	0.218
SPLS	1.505	0.237***	4.878***	0.469***	1.136	-0.274	2.389**	-0.046
AEN	1.668*	0.536***	5.048***	0.767***	2.273**	0.025	1.653*	0.253***
DAE	1.365	0.466***	4.875***	0.698***	1.284	-0.045	1.146	0.183***
Dynamic Error Covariance								
PCA	1.785*	0.333***	4.826***	0.565***	2.797***	-0.178	2.319**	0.050
PLS	1.741*	0.330***	4.256***	0.561***	1.612	-0.181	2.907***	0.047
SPCA	1.534	0.314***	4.661***	0.546***	2.151**	-0.197	1.959*	0.031
SPLS	1.890*	0.310***	4.373***	0.542***	1.786*	-0.201	3.276***	0.027
AEN	1.692*	0.292***	4.699***	0.524***	2.625***	-0.218	2.354**	0.009
DAE	1.881*	0.324***	4.886***	0.555***	3.054***	-0.187	2.349**	0.041

4.6. Subperiod Analysis

In this section we examine portfolio performance during different subperiods as defined by market volatility.[11] The impact of different market regimes on asset allocation has been well documented in the literature (see Buckley, Saunders and Seco, 2008; Guidolin and Timmermann, 2008; Guidolin and Hyde, 2012; Bae, Kim, and Mulvey, 2014). During high volatility periods (Table 6, Panel A) all optimal portfolios outperform the EW and VW schemes in terms of both standard deviation and Sharpe ratio. Strategies based on the DBC and DEC specifications exhibit lower risk and higher Sharpe ratio than

[11] The regimes are determined based on the filtered probabilities of the following two-state Markov Switching model: $r_{m,t} = \mu_s + e_{t,s}$, with $e_{t,s} \sim N(0, \sigma_{t,s}^2)$, where r_m is the market factor return, s represents the latent state and μ_s and σ_s^2 denote the state dependent mean and variance. When the filtered probability of the low volatility state is lower than 0.5 the market is in a high-volatility period, while observations where the filtered probability of the low volatility state is higher than 0.5 are low-volatility periods. The market factor was obtained from Kenneth French's Data Library.

the remaining covariance estimators. Specifically, the best performing strategies are those based on PLS and SPLS for the DEC specification with monthly Sharpe ratio of 0.173 and 0.172 respectively. For low volatility periods (Panel B) portfolios based on latent factors generate risk and Sharpe ratio that are similar to the EW portfolio, with factor-based allocations outperforming the remaining benchmarks.[12]

Table 6 Portfolio performance of exact factor models during different volatility regimes

In this table, we document the monthly portfolio performance measured using the standard deviation (SD) and Sharpe ratio (SR), during high (Panel A) and low (Panel B) volatility periods based on the filtered probabilities of a Markov-switching model estimated using the market factor. Observations where the filtered probability of the low volatility regime is above 0.5 are considered low-volatility periods, and observations where the filtered probability of the low volatility regime is below 0.5 are considered high-volatility periods. The results are presented for the equally weighted portfolio (EW), value-weighted portfolio (VW) and minimum-variance portfolios with short-selling constraints based on the sample estimator (Sample), linear shrinkage (LINS) and non-linear shrinkage (NLS) estimators, Wishart stochastic covariance (Wishart) and for four exact factor model (EFM) covariance specifications: static factor covariance (SFC), dynamic factor covariance (DFC), dynamic beta covariance (DBC) and dynamic error covariance (DEC). The factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS), autoencoder (AEN) and denoising autoencoder (DAE).

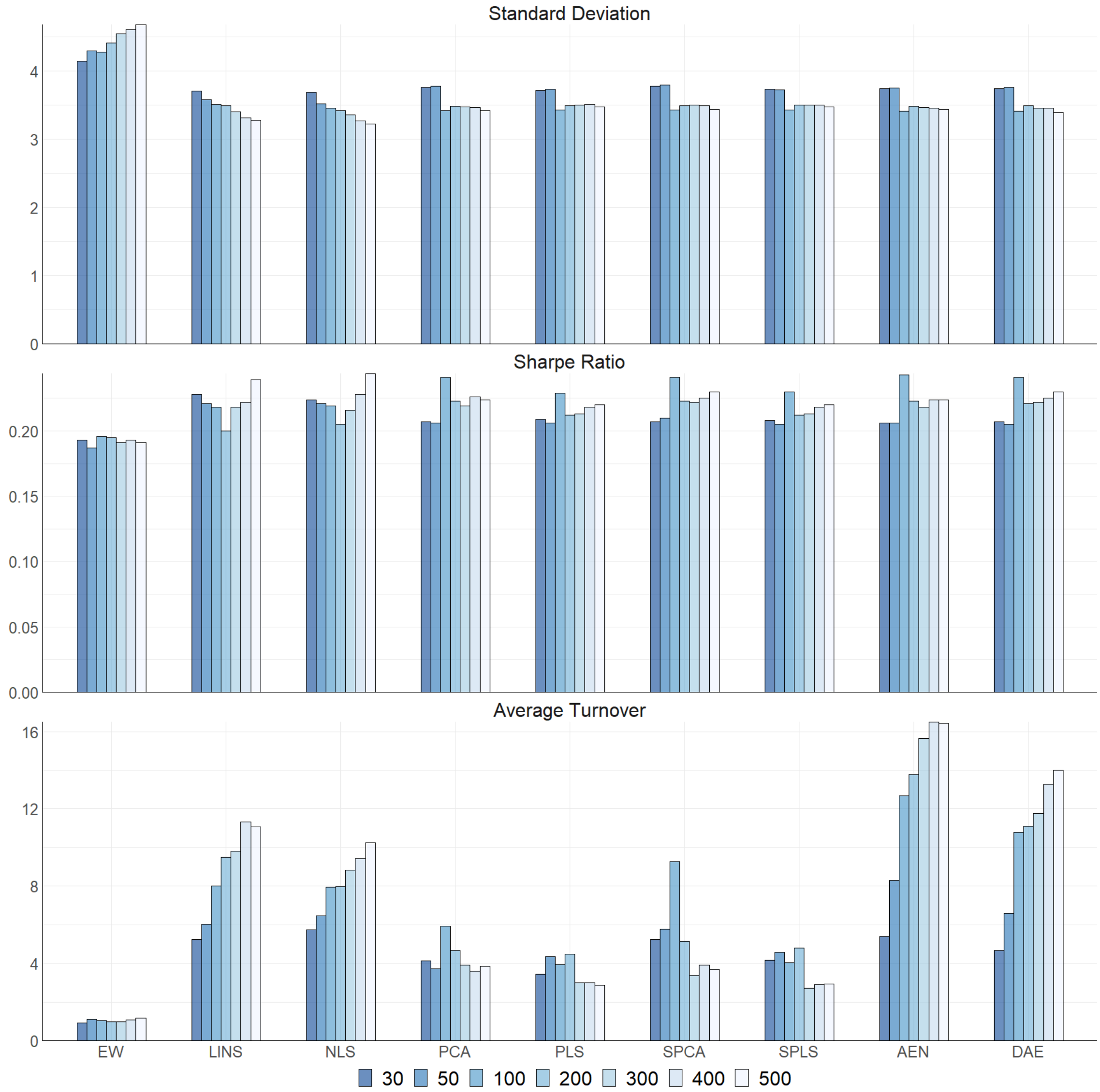
Panel A High volatility regime								
	SD	SR						
EW	5.291	0.123						
VW	5.080	0.120						
Sample	4.045	0.158						
LINS	4.082	0.158						
NLS	4.019	0.147						
Wishart	4.375	0.151						
	SFC		DFC		DBC		DEC	
	SD	SR	SD	SR	SD	SR	SD	SR
PCA	4.080	0.159	4.078	0.159	3.982	0.158	3.945	0.160
PLS	4.060	0.154	4.068	0.152	3.955	0.152	3.893	0.173
SPCA	4.096	0.160	4.100	0.159	3.979	0.161	3.965	0.160
SPLS	4.064	0.154	4.072	0.151	3.962	0.151	3.878	0.172
AEN	4.058	0.161	4.053	0.160	3.970	0.163	3.938	0.162
DAE	4.062	0.161	4.057	0.160	3.999	0.163	3.923	0.162
Panel B Low volatility regime								
	SD	SR						
EW	2.375	0.457						
VW	2.415	0.412						
Sample	2.638	0.334						
LINS	2.589	0.356						
NLS	2.524	0.384						
Wishart	2.692	0.352						
	SFC		DFC		DBC		DEC	
	SD	SR	SD	SR	SD	SR	SD	SR
PCA	2.265	0.463	2.266	0.462	2.260	0.471	2.265	0.457
PLS	2.347	0.423	2.348	0.422	2.327	0.435	2.382	0.405
SPCA	2.267	0.461	2.268	0.461	2.266	0.469	2.280	0.457
SPLS	2.346	0.425	2.347	0.424	2.327	0.437	2.367	0.407
AEN	2.285	0.461	2.284	0.461	2.277	0.471	2.304	0.450
DAE	2.288	0.454	2.293	0.454	2.278	0.460	2.297	0.450

[12] The performance during different volatility regimes for minimum-variance portfolios that allow short-selling and portfolios with a turnover penalty are reported in Tables A10 and A11 respectively in the Appendix.

4.7. Varying Number of Assets

Here we examine how performance is affected according to the number of assets in the portfolio. Along with the baseline case for $N = 100$, the results for portfolios for $N = \{30, 50, 200, 300, 400, 500\}$ largest stocks by market capitalization[13] are presented in Figure 1 for the case of the static factor covariance specification.[14]

Figure 1 Portfolio performance for a different number of stocks: Static Factor Covariance
This figure shows the monthly portfolio performance for a varying number of assets. Performance is based on the standard deviation, Sharpe ratio and average turnover. The out-of-sample period is from January 1980 to December 2022. The results are presented for the equally weighted portfolio (EW), for the linear shrinkage (LINS) and non-linear shrinkage (NLS) estimators and for the static factor covariance (SFC) based on an exact factor model (EFM). The factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS), autoencoder (AEN) and denoising autoencoder (DAE). The standard deviation and average turnover are reported as a percentage.



[13] The maximum fixed number of assets available throughout the out-of-sample period is 500.
[14] The results for the remaining covariance specifications exhibit a similar pattern to that of the static case and are presented in Figures A1, A2 and A3 in the Appendix, for the cases when B, Σ_f and Σ_u are dynamic, respectively.

When the number of assets changes, the $1/N$ portfolio is still consistently outperformed by the alternative strategies, with standard deviation increasing with the size of the portfolio, and Sharpe ratio remaining relatively flat. Decreasing the number of assets to $N = \{30, 50\}$, strategies based on shrinkage estimators tend to outperform those using latent factors in terms of risk and Sharpe ratio. The volatility of the portfolios based on shrinkage covariance matrices consistently decreases as the size of the portfolio increases, however, the results for the Sharpe ratio are mixed, decreasing for $N = \{200, 300\}$ and then increasing again for $N = \{400, 500\}$. The volatility of latent factor models is high for $N = \{30, 50\}$ and then becomes lower and stabilizes across different portfolio sizes. The Sharpe ratio for latent factors is highest for $N = 100$, decreases when $N = \{200, 300\}$ and then increases again for $N = \{400, 500\}$. Average monthly turnover steadily increases with the number of stocks in the portfolio, with shrinkage methods generating higher turnover than factor models based on linear dimensionality reduction methods and lower turnover than portfolios based on autoencoders.

4.8. Performance of Approximate Factor Models

In the analysis so far, we have considered exact factor models, where the residual covariance matrix is diagonal. In this Section we examine the forecasting accuracy and economic value of covariance matrices based on approximate factor models (AFM). In the three specifications that the residual covariance matrix is static (SFC, DFC and DBC) the residual covariance is estimated using the linear shrinkage estimator by Ledoit and Wolf (2004), while for the dynamic error covariance specification, the residual covariance is estimated using the DCC-NL by Engle, Ledoit and Wolf (2019) dynamic covariance estimator.

Table 7 Forecast evaluation of the covariance matrices of approximate factor models
This table reports the ability of the alternative covariance matrices to predict the out-of-sample realized covariance matrix based on four loss functions. Panel A reports the mean squared error (MSE) and mean absolute error (MAE), whereas the quasi-likelihood function (QLK) and asymmetric loss function (ASYM) can be found in Panel B. The average value of each measure over the out-of-sample period from January 1980 to December 2022 is reported. The results are presented for the sample estimator (Sample), linear shrinkage (LINS) and non-linear shrinkage (NLS) estimators, Wishart stochastic covariance (Wishart) and for four approximate factor model (AFM) covariance specifications: static factor covariance (SFC), dynamic factor covariance (DFC), dynamic beta covariance (DBC) and dynamic error covariance (DEC). The factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS), autoencoder (AEN) and denoising autoencoder (DAE). The statistical significance of the alternative covariance matrices compared to the LINS covariance benchmark is denoted by: *, **, and *** for significance at the 10%, 5%, and 1% level, respectively.

Panel A Mean squared error and mean absolute error

	MSE	MAE
Sample	51.666	18.777
LINS	51.972	19.139
NLS	48.217	18.306
Wishart	54.911	19.143

	SFC		DFC		DBC		DEC	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
PCA	45.613**	17.461**	48.314	17.813**	55.477	17.707**	41.483	13.958**
PLS	45.477**	17.414**	48.205	17.696**	61.138	17.686	42.454	14.732**

	SFC		DFC		DBC		DEC	
SPCA	45.596**	17.458**	47.555*	17.710**	59.741	17.657**	41.194	13.946**
SPLS	45.469**	17.413**	48.277	17.709**	61.033	17.671	43.223	14.712**
AEN	45.614**	17.460**	50.095	18.036*	55.987	17.628	95.780	13.863**
DAE	45.599**	17.457**	49.744	18.000*	55.494	17.635**	48.536	13.899*

Panel B Quasi-likelihood function and asymmetric loss function

	QLK	ASYM
Sample	21.810	3.488
LINS	23.599	3.558
NLS	24.965	3.247
Wishart	23.328	3.818

	SFC		DFC		DBC		DEC	
	QLK	ASYM	QLK	ASYM	QLK	ASYM	QLK	ASYM
PCA	23.811*	2.898**	23.815	3.273	23.922	5.493	19.726	3.521
PLS	24.017	2.901**	24.014	3.394	24.105**	8.040	19.944	3.269
SPCA	23.800	2.894**	23.801	3.173	23.910	9.400	19.762	3.492
SPLS	24.018	2.901**	24.016	3.391	24.105**	8.056	19.943	3.525
AEN	23.799	2.896**	23.808	3.476	23.913	6.011	19.697	104.522
DAE	23.800	2.894**	23.808	3.434	23.916	5.848	19.713	9.982

Table 7 reports the results for the predictive accuracy of the approximate factor models. Comparing the performance of the AFMs with that of the EFMs (Table 1), we observe that the predictive accuracy in terms of MSE, MAE and ASYM worsens. There are some gains in terms of the QLK loss, however, they are not significant enough to lead to consistent significant outperformance relative to the LINS benchmark. Static factor covariance matrices significantly outperform the benchmark in terms of MSE and MAE, while the results for the dynamic specifications primarily exhibit significant outperformance for MAE when the factor covariance or error covariance are time varying. Dynamic error covariance matrices generate lower but insignificant QLK values than the benchmarks, while static factor covariances exhibit significant outperformance in terms of the ASYM measure.

Table 8 Portfolio performance of approximate factor models based on standard deviation and Sharpe ratio
This table documents monthly portfolio performance measured using the standard deviation (SD) and Sharpe ratio (SR), over the out-of-sample period from January 1980 to December 2022. The results are presented for the equally weighted portfolio (EW), value-weighted portfolio (VW) and minimum-variance portfolios with short-selling constraints based on the sample estimator (Sample), linear shrinkage (LINS) and non-linear shrinkage (NLS) estimators, Wishart stochastic covariance (Wishart) and for four approximate factor model (AFM) covariance specifications: static factor covariance (SFC), dynamic factor covariance (DFC), dynamic beta covariance (DBC) and dynamic error covariance (DEC). The factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS), autoencoder (AEN) and denoising autoencoder (DAE).

	SD	SR
EW	4.275	0.196
VW	4.137	0.188
Sample	3.501	0.212
LINS	3.510	0.218
NLS	3.450	0.219
Wishart	3.736	0.210

	SFC		DFC		DBC		DEC	
	SD	SR	SD	SR	SD	SR	SD	SR
PCA	3.457	0.219	3.456	0.218	3.430	0.220	3.361	0.229
PLS	3.449	0.220	3.454	0.218	3.420	0.221	3.468	0.226
SPCA	3.457	0.219	3.458	0.217	3.429	0.221	3.442	0.229
SPLS	3.449	0.220	3.454	0.218	3.420	0.221	3.422	0.240
AEN	3.454	0.219	3.451	0.219	3.421	0.221	3.407	0.219

DAE	3.455	0.219	3.454	0.219	3.437	0.218	3.360	0.242

The economic value of approximate factor models in terms of standard deviation and Sharpe ratio is reported in Table 8. Overall, exact factor models consistently outperform their approximate factor model counterparts in terms of both performance measures. The factor-based covariance matrices continue to outperform the EW and VW benchmarks, however, their performance becomes closer to that of the other covariance benchmarks, with the exception of portfolios based on dynamic error covariance matrices with latent factors estimated using SPLS and DAE.

5. Conclusion

Minimum-variance portfolios are frequently advocated by both academics and financial professionals, since they avoid the high estimation error associated with expected returns and have been shown to outperform competing strategies. Nevertheless, this investment strategy remains crucially dependent on the quality of the estimates of the covariance matrix, which are exacerbated for high-dimensional opportunity sets. In this paper we address this issue by imposing a factor structure to the covariance matrix and conduct a systematic evaluation of the performance and properties of factor-based minimum-variance portfolios. Furthermore, we enhance factor-based covariance matrix estimation, by considering latent factors derived from dimensionality reduction methods that induce sparsity or introduce non-linearities and specifications of the factor-based covariance matrix that allow its components to be time-varying.

Overall, the results for the predictive accuracy based on two symmetric loss functions indicate that the majority of the factor models outperform several covariance benchmarks, while according to two asymmetric loss functions the improved performance of the factor models is due to the reduced degree of over predictions. From the economic evaluation of the covariance matrices based on the minimum-variance framework, we find that using factor-based covariance matrices can translate into economic gains for optimal portfolios. We find that the proposed models can lead to a statistically significant reduction in portfolio volatility and a significant increase in the Sharpe ratios relative to the EW and VW portfolios. When the factor-based allocations are compared to minimum-variance allocations based on alternative covariance benchmarks, the alternative strategies outperform the sample, linear shrinkage and Wishart stochastic covariance benchmarks, with the non-linear shrinkage estimator being a more difficult benchmark to outperform in terms of portfolio standard deviation. Factor-based allocations yield portfolios that require less frequent rebalancing, with weights that are less volatile and more diversified relative to other covariance matrix benchmarks.

References

Ardia, D., Bolliger, G., Boudt, K., & Gagnon-Fleury, J. P. (2017). The impact of covariance misspecification in risk-based portfolios. *Annals of Operations Research*, 254, 1-16.

Bae, G. I., Kim, W. C., & Mulvey, J. M. (2014). Dynamic asset allocation for varied financial markets under regime switching framework. *European Journal of Operational Research*, 234(2), 450-458.

Bali, T. G., Engle, R. F., & Tang, Y. (2017). Dynamic conditional beta is alive and well in the cross section of daily stock returns. *Management Science*, 63(11), 3760-3779.

Becker, R., Clements, A. E., Doolan, M. B., & Hurn, A. S. (2015). Selecting volatility forecasting models for portfolio allocation purposes. *International Journal of Forecasting*, 31(3), 849-861.

Best, M. J., & Grauer, R. R. (1991a). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results. *The Review of Financial Studies*, 4(2), 315-342.

Best, M. J., & Grauer, R. R. (1991b). Sensitivity analysis for mean-variance portfolio problems. *Management Science*, 37(8), 980-989.

Bianchi, D., & Guidolin, M. (2014). Can long-run dynamic optimal strategies outperform fixed-mix portfolios? Evidence from multiple data sets. *European Journal of Operational Research*, 236(1), 160-176.

Bianchi, D., Büchner, M., & Tamoni, A. (2021). Bond risk premiums with machine learning. *The Review of Financial Studies*, 34(2), 1046-1089.

Bodnar, T., Parolya, N., & Schmid, W. (2018). Estimation of the global minimum variance portfolio in high dimensions. *European Journal of Operational Research*, 266(1), 371-390.

Broadie, M. (1993). Computing efficient frontiers using estimated parameters. *Annals of Operations Research*, 45(1), 21-58.

Buckley, I., Saunders, D., & Seco, L. (2008). Portfolio optimization when asset returns have the Gaussian mixture distribution. *European Journal of Operational Research*, 185(3), 1434-1461.

Carroll, R., Conlon, T., Cotter, J., & Salvador, E. (2017). Asset allocation with correlation: A composite trade-off. *European Journal of Operational Research*, 262(3), 1164-1180.

Chamberlain, G., & Rothschild, M. (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica: Journal of the Econometric Society*, 1281-1304.

Chan, L. K., Karceski, J., & Lakonishok, J. (1999). On portfolio optimization: Forecasting covariances and choosing the risk model. *The Review of Financial Studies*, 12(5), 937-974.

Chopra, V. K., Hensel, C. R., & Turner, A. L. (1993). Massaging mean-variance inputs: Returns from alternative global investment strategies in the 1980s. *Management Science*, 39(7), 845-855.

Connor, G., & Korajczyk, R. A. (1988). Risk and return in an equilibrium APT: Application of a new test methodology. *Journal of Financial Economics*, 21(2), 255-289.

De Nard, G., Ledoit, O., & Wolf, M. (2021). Factor models for portfolio selection in large dimensions: The good, the better and the ugly. *Journal of Financial Econometrics*, 19(2), 236-257.

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *The Review of Financial Studies*, 22(5), 1915-1953.

DeMiguel, V., Garlappi, L., Nogales, F. J., & Uppal, R. (2009). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science*, 55(5), 798-812.

DeMiguel, V., Nogales, F. J., & Uppal, R. (2014). Stock return serial dependence and out-of-sample portfolio performance. *The Review of Financial Studies*, 27(4), 1031-1073.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3), 339-350.

Engle, R. F., Ledoit, O., & Wolf, M. (2019). Large dynamic covariance matrices. *Journal of Business & Economic Statistics*, 37(2), 363-375.

Engle, R. F., Ng, V. K., & Rothschild, M. (1990). Asset pricing with a factor-ARCH covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics*, 45(1-2), 213-237.

Fan, J., Fan, Y., & Lv, J. (2008). High dimensional covariance matrix estimation using a factor model. *Journal of Econometrics*, 147(1), 186-197.

Fan, J., Liao, Y., & Mincheva, M. (2011). High dimensional covariance matrix estimation in approximate factor models. *Annals of Statistics*, 39(6), 3320-3356.

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654-669.

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 9(3), 432-441.

Gambella, C., Ghaddar, B., & Naoum-Sawaya, J. (2021). Optimization problems for machine learning: A survey. *European Journal of Operational Research*, 290(3), 807-828.

Green, R. C., & Hollifield, B. (1992). When will mean-variance efficient portfolios be well diversified? *The Journal of Finance*, 47(5), 1785-1809.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5), 2223-2273.

Gu, S., Kelly, B., & Xiu, D. (2021). Autoencoder asset pricing models. *Journal of Econometrics*, 222(1), 429-450.

Guidolin, M., & Hyde, S. (2012). Can VAR models capture regime shifts in asset returns? A long-horizon strategic asset allocation perspective. *Journal of Banking & Finance*, 36(3), 695-716.

Guidolin, M., & Timmermann, A. (2008). International asset allocation under regime switching, skew, and kurtosis preferences. *The Review of Financial Studies*, 21(2), 889-935.

Huang, D., Jiang, F., Li, K., Tong, G., & Zhou, G. (2022). Scaled PCA: A new approach to dimension reduction. *Management Science*, 68(3), 1678-1695.

Huck, N. (2019). Large data sets and machine learning: Applications to statistical arbitrage. *European Journal of Operational Research*, 278(1), 330-342.

Hui, T. K., Kwan, E. K., & Lee, C. B. (1993). Optimal portfolio diversification: Empirical Bayes versus classical approach. *Journal of the Operational Research Society*, 44(11), 1155-1159.

Hwang, I., Xu, S., & In, F. (2018). Naive versus optimal diversification: Tail risk and performance. *European Journal of Operational Research*, 265(1), 372-388.

Jagannathan, R., & Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, 58(4), 1651-1683.

Kan, R., & Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, 621-656.

Kelly, B. T., Pruitt, S., & Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, 134(3), 501-524.

Kolm, P. N., Tütüncü, R., & Fabozzi, F. J. (2014). 60 Years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, 234(2), 356-371.

Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research*, 259(2), 689-702.

Lassance, N., & Vrins, F. (2021). Portfolio selection with parsimonious higher comoments estimation. *Journal of Banking & Finance*, 126, 106115.

Lassance, N., DeMiguel, V., & Vrins, F. (2022). Optimal portfolio diversification via independent component analysis. *Operations Research*, 70(1), 55-72.

Laurent, S., Rombouts, J. V., & Violante, F. (2013). On loss functions and ranking forecasting performances of multivariate volatility models. *Journal of Econometrics*, 173(1), 1-10.

Ledoit, O., & Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2), 365-411.

Ledoit, O., & Wolf, M. (2011). Robust performances hypothesis testing with the variance. *Wilmott*, 2011(55), 86-89.

Ledoit, O., & Wolf, M. (2017). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *The Review of Financial Studies*, 30(12), 4349-4388.

Lettau, M., & Pelger, M. (2020). Factors that fit the time series and cross-section of stock returns. The *Review of Financial Studies*, 33(5), 2274-2325.

Maillet, B., Tokpavi, S., & Vaucher, B. (2015). Global minimum variance portfolio optimisation under some model risk: A robust regression-based approach. *European Journal of Operational Research*, 244(1), 289-299.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 77-91.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8(4), 323-361.

Moskowitz, T. J. (2003). An analysis of covariance risk and pricing anomalies. *The Review of Financial Studies*, 16(2), 417-457.

Moura, G. V., Santos, A. A., & Ruiz, E. (2020). Comparing high-dimensional conditional covariance matrices: Implications for portfolio selection. *Journal of Banking & Finance*, 118, 105882.

Olivares-Nadal, A. V., & DeMiguel, V. (2018). A robust perspective on transaction costs in portfolio optimization. *Operations Research*, 66(3), 733-739.

Opdyke, J. D. (2007). Comparing Sharpe ratios: so where are the p-values? *Journal of Asset Management*, 8, 308-336.

Palczewski, A., & Palczewski, J. (2014). Theoretical and empirical estimates of mean–variance portfolio sensitivity. *European Journal of Operational Research*, 234(2), 402-410.

Platanakis, E., Sutcliffe, C., & Ye, X. (2021). Horses for courses: Mean-variance for asset allocation and 1/N for stock selection. *European Journal of Operational Research*, 288(1), 302-317.

Sermpinis, G., Theofilatos, K., Karathanasopoulos, A., Georgopoulos, E. F., & Dunis, C. (2013). Forecasting foreign exchange rates with adaptive neural networks using radial-basis functions and particle swarm optimization. *European Journal of Operational Research*, 225(3), 528-540.

Wu, W., Chen, J., Yang, Z., & Tindall, M. L. (2021). A cross-sectional machine learning approach for hedge fund return prediction and selection. *Management Science*, 67(7), 4577-4601.