Machine Learning and Factor-Based Portfolio Optimization*

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

Current Version: March 9, 2022

* The authors gratefully acknowledge the support of Science Foundation Ireland under grant number 16/SPP/3347 and 17/SPP/5447. We also acknowledge the comments of Andrea Barbon, Martin Brown, Gregory Connor, and seminar participants at the University of St. Gallen.

[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

Machine Learning and Factor-Based Portfolio Optimization

Abstract

We adopt a factor-based framework to construct the covariance matrix by using latent factors based on machine learning, with the goal of enhancing minimum-variance portfolio optimization. We find that factors based on autoencoder neural networks exhibit a weaker relationship with commonly used characteristic-sorted portfolios than popular dimensionality reduction techniques. Machine learning also leads to covariance and portfolio weight structures that diverge from simpler estimators. Portfolios using latent factors derived from autoencoders and sparse methods outperform simpler benchmarks in terms of risk minimization. The improved performance is amplified for investors with increased sensitivity to risk and during high volatility periods.

JEL classifications: C38, C4, C45, C5, C58, G1, G11

Keywords: Covariance matrix; Dimensionality reduction; Factor models; Machine learning; Minimum-variance; Portfolio optimization.

I. Introduction

In this paper we synthesize the areas of covariance matrix estimation and minimum-variance optimization with that of machine learning. Specifically, we examine the characteristics and benefits of minimum-variance portfolios based on factor-implied covariance matrices when the latent factors are generated from machine learning dimensionality reduction techniques.

The frequently described issues associated with expected mean estimation in the mean-variance rule by Markowitz (1952) have led many researchers to instead focus on covariance estimation applied to the minimum-variance framework (see e.g., Merton, 1980; Best and Grauer, 1991; Kan and Zhou, 2007). Minimum-variance portfolios have been frequently advocated in the finance literature. Chan, Karcesky, and Lakonishok (1999) show that minimum-variance portfolios based on forecasted variances and covariances yield lower risk than simpler benchmarks such as the equally weighted portfolio (EW). Jagannathan and Ma (2003) find that minimum-variance portfolios generate better out-of-sample performance than mean-variance portfolios. Kempf, Korn and Saßning (2015), also show that estimating covariance matrices relying on forward-looking information leads to minimum-variance portfolios that outperform index investing and the EW portfolio, experiencing significant gains during recessions. Ledoit and Wolf (2017) introduce a nonlinear shrinkage estimator for the covariance matrix and show that minimum-variance portfolios based on the new estimator outperform those based on linear shrinkage, sample moments or the EW portfolio. More recently, Shi, Shu, Yang and He (2020), propose a structure-free and computationally efficient estimator of the covariance matrix that regularizes the sample eigenvalues and yields improved performance over alternative portfolio strategies in various settings.

Although the minimum-variance framework avoids the problem of estimation error associated with expected returns, its performance remains crucially dependent on the quality of the estimated covariance matrix (DeMiguel, Garlappi, Nogales and Uppal, 2009). The approach we follow to lessen the impact of covariance misspecification on the optimal weights, is to impose a factor structure on the covariance matrix (Chan, Karceski and Lakonishok, 1999), which reduces the number of parameters to be estimated.[4] Factor models assume that asset returns are driven by a set of observed or latent factors. It has been shown that introducing a factor structure to the covariance can improve portfolio performance (Green and Hollifield, 1992; Chan, Karceski and Lakonishok, 1999). The benefits of using the factor model-based approach have also been investigated by Fan, Fan and Lv (2008) and Fan, Liao and Mincheva (2011; 2013) who propose estimators of the covariance for exact and approximate factor models respectively. More recently, De Nard, Ledoit and Wolf (2019), use a factor framework and evaluate portfolios for different estimates of the error covariance matrix.

Machine learning has been shown to be well suited for risk premium predictability problems (e.g., Gu, Kelly and Xiu, 2020) or constructing factors that explain the cross-section of stock returns (e.g., Kozak, Nagel and Santosh, 2020). However, there exists fewer papers that focus on modelling the structure of the covariance matrix or examining the performance in a factor-based minimum-variance framework. In terms of modelling the covariance matrix using machine learning, Callot, Caner, Önder, and Ulasan (2019), use nodewise regression and the lasso to directly estimate the sparse precision matrix. Minimum-variance portfolios based on their proposed approach exhibit lower variances and higher Sharpe ratios compared to commonly used covariance estimators. Turning to factor-based portfolio optimization, Lassance, DeMiguel and Vrins (2020) estimate factor-risk-parity portfolios by choosing a set of uncorrelated factors using independent component analysis. They show that portfolios based

[4] The literature proposes several other approaches to reduce the impact of covariance misspecification on the impact of the optimal weights. One approach is to impose restrictions on the weights of the portfolios, either by introducing short-selling constraints (Jagannathan and Ma, 2003), or by limiting turnover via additional constraints (DeMiguel, Garlappi, Nogales and Uppal, 2009) or penalizing the objective function (Olivares-Nadal and DeMiguel, 2018). Another approach uses either shrinkage estimators of the covariance matrix (Ledoit and Wolf, 2004), 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 (see e.g., Jagannathan and Ma, 2003).

on independent components provide greater diversification benefits and outperform those using principal components and other benchmarks. We also relate to papers that use autoencoder neural networks in financial applications. Gu, Kelly and Xiu (2021) propose a model for the cross-section of stock returns based on autoencoders, where factors are latent, and the time-varying loadings depend on characteristics.

Our contribution stems from the framework through which we bridge the gap between machine learning and finance. The aforementioned studies improve covariance estimation via shrinkage methods, focus on factor-risk-parity portfolios or explain the cross-section of returns using machine learning. We differ by using machine learning in a factor-based framework to model the covariance matrix and focusing on improving minimum-variance portfolio optimization. Specifically, we construct latent factors using machine learning, explore their impact on the structure of factor-based covariances and on the composition and performance of minimum-variance portfolios. We examine the economic value of latent factors generated using a variety of supervised and unsupervised dimensionality reduction methods and their relation to popular characteristics. In addition to classical approaches, such as principal component analysis (PCA) and partial least squares (PLS), we further consider their respective regularized versions that induce sparsity through a penalty in the objective function. We also investigate the performance of factors generated by autoencoders; a type of unsupervised neural network used for dimensionality reduction. Another contribution to the literature of factor-based portfolio optimization arises from conducting a comparative analysis of a static and several dynamic specifications of the covariance matrix, based on observed or latent factors. The structure of a dynamic covariances can differ based on whether the factor loadings, the factor covariance matrix or the residual covariance matrix are allowed to vary over time.

To determine the effects of using factors based on machine learning in covariance estimation and portfolio optimization, we first explore the potential links between the latent factors and popular factor proxies and examine the structure of the factor-implied covariance matrices. Moskowitz (2003) examines the covariance structure of returns with respect to various factors and finds that size can better explain covariance risk, both in and out of sample, while the book-to-market factor exhibits a weaker association and momentum factors appear unrelated to return second moments. Our findings indicate that latent factors produced by PCA and PLS based methods exhibit a stronger connection with well-known factors (such as those from the Fama and French (2015) five-factor model) throughout the out-of-sample period, compared to factors based on autoencoders. Furthermore, machine learning yields factors that cause the covariances to diverge from those based on commonly used estimators. Specifically, the covariance matrices whose structure deviates most from the sample estimator are based on unsupervised methods or allow the residual covariance matrix to be time-varying.

In the baseline case, machine learning leads to portfolios that significantly outperform the equally weighted portfolio benchmark, which DeMiguel, Garlappi and Uppal (2009) show to be a very stringent benchmark and improves upon portfolios based on the sample estimator or observed factors. The best-performing methods to generate the covariance matrix are autoencoders and sparse principal component analysis, which can lead to portfolios with higher risk-adjusted return and standard deviation that is 3% lower per annum than the equally weighted one. Certainty equivalent returns are also improved relative to the benchmark as the investor's aversion to risk increases, which is in keeping with the optimization goal. In particular, investors with moderate or conservative risk preferences using machine learning factors would realize significant utility gains that are between 2.5% and 4.5% higher than those of the EW portfolio on an annual basis. The performance of portfolios based on machine learning is amplified during periods of high volatility.

Portfolios based on machine learning also have weights that are smaller, vary less over time and are more diversified, than those based on observed factors. Covariances based on unsupervised methods also lead to portfolios with lower turnover and thus reduced sensitivity to transaction costs. Specifically, all machine learning portfolios generate positive breakeven transaction costs, indicating they outperform EW. The factor-based portfolios continue to outperform the benchmark after increasing the number of assets. Turnover increases considerably with the size of the portfolio, while the breakeven transaction costs decrease. Machine learning models still generate the highest performance across larger portfolio sizes.

Similar to recent studies, the results indicate that shallow learning outperforms deeper learning, which can be attributed to the small size of the data set and the low signal-to-noise ratio. When we consider factors based on neural networks with one to four hidden layers, we find that for the baseline optimization framework the shallowest network outperforms those with deeper architectures. This decline in portfolio performance is potentially associated with the high degree of turnover of strategies based on autoencoders with more hidden layers. Additionally, unsupervised methods tend to perform better than supervised methods. When we compare PLS and sparse PLS with PCA and sparse PCA, the results show that the PCA based approaches significantly outperform the EW benchmark more often. Otherwise, the ranking among factors persists across specifications of the covariance matrix. When comparing the results across the alternative specifications of the factor-based covariance matrix, the differences become less pronounced, with approaches that allow the loadings or the residual covariance matrix to vary over time yielding higher risk-adjusted performance. Overall, the results indicate that machine learning improves factor-based portfolio optimization.

The remainder of this study is organized as follows. Section II describes the methodology. Section III provides details on the data, sample splitting and hyperparameter tuning. Section IV and V examine the properties and economic value of the portfolios using the alternative covariance estimates, while Section VI concludes.

II. Methodology

In this section we introduce the machine learning 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.

A. Latent Factors via Machine Learning

The presence of a factor structure in asset returns has been widely accepted in the economic literature. The capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965) implies a simple factor structure. Other early studies that support the presence of a factor structure in asset returns include the APT of Ross (1976) and the intertemporal CAPM of Merton (1973). Factors can be observable quantities, such as macroeconomic indicators (Chen, Roll and Ross, 1986) or observable proxies, where factors are the returns of portfolios constructed by sorting stocks based on firm characteristics, such as the single index model of Sharpe (1963) or the three-factor model of Fama and French (1993) and its variations.

A factor model for the returns of every asset, $r_{i,t}$, with $i = 1, \ldots, N$ assets, $t = 1, \ldots, T$ observations and $k = 1, \ldots, K$ observed factors takes the form

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

where $\beta_i = (\beta_{i,1}, \ldots, \beta_{i,K})$ are the time-invariant factor loadings for factors $\mathrm{F}_t = (f_{t,1}, \ldots, f_{t,K})$, a_i is the intercept and $u_{i,t}$ is the error term for asset i at date t. The intercepts and loadings can be estimated by ordinary least squares (OLS) using the different factor representations.

Factors can also be 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. The general form of a latent factor model is given by

$$r_{i,t} = a_i + \beta_i(\mathrm{R}_t\mathrm{W}) + u_{i,t} = a_i + \beta_i\mathrm{F}_t + u_{i,t}, \quad (2)$$

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 $K \ll N$. 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 $\mathrm{F}_t = \mathrm{R}_t\mathrm{W}$.

The two most 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, PCA derives the latent factors in an unsupervised way, while in the case of PLS the factors are constructed in a supervised way. The methods differ in the way the latent factor matrix, F_t, is extracted, since PCA produces the weight matrix W reflecting the covariance structure between assets, while PLS computes weights so as the latent factors have maximal correlation with the target.

PCA and PLS have the drawback that for each latent factor the weights are typically non-zero, which leads to difficulties in high dimensional settings. 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), developed by Zou, Hastie and Tibshirani (2006) and sparse partial least squares (SPLS) by Chun and Keles (2010). Both methods impose a penalty based on the combination of the l_1 and l_2 allowing for the construction of sparse latent factors.

Finally, we construct latent factors using autoencoders (see e.g., Hinton and Zemel, 1994; Gu, Kelly and Xiu, 2021) 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. We consider four different network architectures based on the depth of the network. First, we construct a shallow autoencoder (AEN1) with a single hidden layer. The other three models include additional hidden layers to network representation, up to a maximum of three layers (AEN2, AEN3 and AEN4). Recent applications in finance (see e.g., Gu, Kelly and Xiu, 2020) find that shallower networks generate better performance. Further details on the machine learning approaches and related literature are provided in the Appendix.

B. Factor-based Covariance Estimation

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:

$$\Sigma_r = \mathrm{B}'\Sigma_f \mathrm{B} + \Sigma_u, \qquad (3)$$

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 (Fan, Fan and Lv, 2008), where the covariance matrix of the residuals u_t is diagonal, $\Sigma_u \equiv diag(\Sigma_u)$, by assuming cross-sectional independence.

The models presented so far rely on a static specification. However, Sargent and Sims (1977) and Geweke (1977) introduce dynamic factor models (DFM) as an extension. There are various definitions of DFMs (see Stock and Watson, 2011), the one we follow in this study is a model that allows the factor loadings to be time varying (Avramov and Chordia, 2006; Engle, 2016 and Bali, Engle and Tang, 2017) or models in which either the factor or residual covariance matrix varies over time (Engle, Ng and Rothchild, 1990). In the description below, the factors can be observed quantities or latent factors.

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, (ii) the covariance matrix of the factors is time-varying or (iii) the covariance matrix of the errors is time-varying.

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

$$r_{i,t} = a_{i,t} + \beta_{i,t}\mathrm{F}_t + u_{i,t}. \tag{4}$$

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:

$$\Sigma_{r,t} = \mathrm{B}_t'\Sigma_f\mathrm{B}_t + \Sigma_u, \tag{5}$$

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

$$\Sigma_{r,t} = \mathrm{B}'\Sigma_{f,t}\mathrm{B} + \Sigma_u. \tag{6}$$

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

$$\Sigma_{r,t} = \mathrm{B}'\Sigma_f\mathrm{B} + \Sigma_{u,t}. \tag{7}$$

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.

C. Minimum-Variance Portfolios

The sensitivity of portfolio weights to estimates of asset means is well documented (Best and Grauer, 1991; Chopra and Turner, 1993). To this end, we focus on minimum-variance portfolios, which have frequently been used in the portfolio optimization literature (see e.g. Carroll, Conlon, Cotter, Salvador, 2017; Moura, Santos and Ruiz, 2020), thus avoiding the issue of estimation error in expected returns.

Specifically, the different estimates of the covariance matrix, $\hat{\Sigma}_r$, from the factor models are evaluated through the minimum-variance framework with short-selling constraints, where the goal is to minimize portfolio risk. 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 objective is

$$\underset{\omega}{\operatorname{argmin}}\, \omega'\hat{\Sigma}_r\omega, \quad \text{s.t.} \quad \omega'\mathrm{i}_N = 1, \;\; \omega_i \geq 0, \text{ for } i = 1, \dots, N, \tag{8}$$

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 short-selling and leverage constraints to avoid implausible positions, by imposing a lower bound of zero on all weights and that the sum of the weights does not exceed one. 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.

III. Data, Sample Splitting and Hyperparameter Tuning

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 2019, for a period of 60 years (or $T = 720$ monthly observations). Our approach regarding the backtest and the restrictions we impose on the data set is similar to that of Engle, Ledoit and Wolf (2019) 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 (exchange codes 1, 2 or 3) and to ordinary common shares (share codes 10 or 11).

We adopt a rolling window approach to examine the out-of-sample (OOS) performance of our models. The size of the rolling window is set to $T_0 = 240$ monthly observations (or 20 years), 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 $T_{OOS} = T - T_0 = 480$ monthly observations (or 40 years), from January 1980 to December 2019. The portfolios are constructed in each iteration of the rolling window, based on stocks that have at least 97.5% history of returns available over the past $T_0 = 240$ months (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. Finally, we consider stocks whose price is greater than $5 within each iteration. 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.[5] In each iteration of the rolling window, we cross-sectionally transform the asset returns, R_t, before estimating the latent factors. Specifically, we calculate the rank of a stock based on the return and then divide the ranks by the number of observations. We subtract 0.5, to map the features into the $[-0.5, 0.5]$ interval. This transformation focuses on the ordering of the data and is insensitive to outliers.

The machine learning models 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, in which the optimal set of values for the tuning parameters is selected in the validation sample. One of the

[5] Figure A1 in the Appendix displays the number of stocks in the sample for each month for the full sample period, while the number of stocks in each iteration of the rolling window for the out-of-sample period is displayed in Figure A2.

advantages of using this approach over k-fold cross validation is that we maintain the temporal ordering of the data. Specifically, in each iteration of the rolling window, the in-sample ($T_0 = 240$) is split into two disjointed periods, the training subsample, $T_0^{\mathcal{T}}$, consisting of 80% of the observations, with the remaining observations belonging to the validation subsample, $T_0^{\mathcal{V}}$. In the training subsample the model is estimated for several sets of values of the tuning parameters. The second subsample is used to select the optimal set of tuning parameters, by using the latent factor weight and loading estimates for each set of hyperparameters from the training sample, forecasts are constructed for the observations in the validation sample.

Specifically, the factor model is first estimated using information only from the training sample, from the following regression equation

$$r_{i,t} = a_i^{\mathcal{T}} + \beta_i^{\mathcal{T}}(\mathrm{R}_t \mathrm{W}^{\mathcal{T}}) + u_{i,t}, \quad \text{for } t = 1, \dots, T_0^{\mathcal{T}}, \tag{9}$$

where $\mathrm{W}^{\mathcal{T}}$ is the factor weight matrix estimated by one of the dimensionality reduction methods using a specific set of hyperparameters, $\beta_i^{\mathcal{T}}$ is one of the N columns of the matrix of factor loadings, $B^{\mathcal{T}}$, estimated by OLS using data only from $\mathcal{T}$ and $a_i^{\mathcal{T}}$ is the intercept. The covariance over the training subsample is then calculated by

$$\Sigma_r^{\mathcal{T}} = B^{\mathcal{T}\prime}\Sigma_f^{\mathcal{T}} B^{\mathcal{T}} + \Sigma_u^{\mathcal{T}}, \tag{10}$$

where $\Sigma_f^{\mathcal{T}}$ is the covariance of the factors $\mathrm{R}_t \mathrm{W}^{\mathcal{T}}$ and $\Sigma_u^{\mathcal{T}}$ is the covariance of the errors u_t for $t = 1, \dots, T_0^{\mathcal{T}}$. The covariance over the training set is compared to that of the validation set, which is derived using the estimated matrices $\mathrm{W}^{\mathcal{T}}$ and $B^{\mathcal{T}}$. The covariance using data from the validation subsample is estimated as

$$\Sigma_r^{\mathcal{V}} = B^{\mathcal{T}\prime}\Sigma_f^{\mathcal{V}} B^{\mathcal{T}} + \Sigma_u^{\mathcal{V}}, \tag{11}$$

where $\Sigma_f^{\mathcal{V}}$ is the covariance of the factors $\mathrm{R}_t \mathrm{W}^{\mathcal{T}}$, for $t = 1, \dots, T_0^{\mathcal{V}}$ and $\Sigma_u^{\mathcal{V}}$ is the covariance of the errors, derived by $u_{i,t} = r_{i,t} - \left(a_i^{\mathcal{T}} + \beta_i^{\mathcal{T}}(\mathrm{R}_t \mathrm{W}^{\mathcal{T}})\right)$, for $t = 1, \dots, T_0^{\mathcal{V}}$.

The covariance matrices based on different sets of hyperparameters are evaluated by employing a measure of economic performance related to the given portfolio application. Following Engle and Colacito (2006) and Becker, Clements, Doolan and Hurn (2015), the optimal set of hyperparameters is chosen to minimize the portfolio variance:

$$\mathcal{L}_{\text{MVP}}\left(\widehat{\omega}, \widehat{\Sigma}_r\right) = \widehat{\omega}' \widehat{\Sigma}_r \widehat{\omega}. \tag{12}$$

The vector of portfolio weights, ω, is obtained by solving the minimum-variance portfolio problem from equation (8), based on the covariance over the training set, $\Sigma_r^{\mathcal{T}}$. The portfolio variance, $\mathcal{L}_{\text{MVP}}$, is calculated using the weight estimates and setting $\widehat{\Sigma}_r = \Sigma_r^{\mathcal{V}}$, for each set of hyperparameters. The set of optimal tuning parameters is the one that yields the lowest portfolio variance over the validation sample. The weights used to construct the latent factors and the factor loadings are re-estimated using all observations in the rolling window for the optimal set of tuning parameters. Finally, the true out-of-sample performance is evaluated by constructing the return of the portfolio using the asset returns one month ahead from the end of each rolling window, which are not included in the validation procedure or parameter estimation.

IV. Characteristics of Latent Factors and Covariance Matrices

In this Section we investigate how the factors based on the dimensionality reduction approaches relate to those of popular factor models and long-short anomaly portfolios and examine the structure of the alternative covariance matrix specifications.

A. Links to Popular Factors

Here we examine the links that the estimated factors have with popular characteristic-based factors from the literature. We investigate the links of the latent factors, first with the five-factors from Fama and French (2015) [6] and then with a larger dataset which consists of the long-short anomaly portfolios constructed by Chen and Zimmermann (2020).

[6] Data on the Fama-French factors were downloaded from Kenneth French's Data Library.

For each dimensionality reduction method, we regress each of the estimated latent factors $f_{t,k}$, for $k \in [1, 5]$, on the factors from the Fama and French (2015) five-factor model using OLS.[7] To compare how well the Fama and French model explains the latent factors across dimensionality reduction approaches, we report the adjusted R^2, averaged over the 438 estimation windows, in Figure 1. A higher average adjusted R^2 indicates that the Fama and French five-factor model is well suited to explain the variability of that specific latent factor throughout the out-of-sample period.

[Insert Figure 1 about here]

The factors based on supervised methods tend to have higher R^2 than those of unsupervised methods, while the R^2 of factors based on autoencoders varies less with K compared to the remaining approaches. Specifically, the results for PCA show that the Fama and French five-factor model is better at explaining the first and third principal components, with R^2 between 34% and 36%, while the R^2 for the remaining components is between 8% and 15%, decreasing as K increases. The pattern for sparse principal components is similar, but the average R^2 is smaller for the first and third factor, decreasing to about 33% and 27%, respectively, while for the remaining components it increases to between 14% and 22%. The Fama and French model explains almost 50% of the variability of the first factor by PLS or SPLS. For the remaining four factors the average R^2 varies between approximately 13% and 23%. The average R^2 of latent factors based on autoencoders remains relatively stable across different values of K and number of hidden layers, with values between 18% and 25%.

[7] Information on the five-factors by Fama and French (2015) is available from July 1963. Given a rolling window size of 240 observations, we have an out-of-sample size of 438 observations, from July 1983 to December 2019. The OLS regressions are re-estimated in each iteration of the rolling window for all combinations of dimensionality reduction techniques and fixed number of factors K. The OLS regressions include an intercept, but since we focus on the potential relationship between the latent factors with the observed factors, we omit the intercept from the Figures for the sake of brevity.

Following Gu, Kelly and Xiu (2020), we quantify the influence of the Fama-French factors as the change in R^2 from setting the observations of a factor proxy to zero within each estimation window. The values are averaged to obtain a single variable importance measure for each of the five Fama and French factors and then scaled to sum to 100. The variable importance is presented in Figure 2.

[Insert Figure 2 about here]

According to the change in R^2, the results for PCA indicate that the market factor (MKTRF) is most influential for the first principal component, $K = 1$, while the value factor (HML) relates mainly to the third latent factor, $K = 3$. The profitability (RMW) and investment (CMA) factors are better at explaining the $K = 4$ and $K = 5$ components, respectively. In contrast, none of the proxies particularly dominates in terms of explaining the second factor, $K = 2$. The characteristics that are influential to the corresponding sparse principal components are similar to those for PCA. However, their importance is decreased with the remaining proxies contributing more to the explanation of the latent factors. The fourth sparse principal component poses an exception since it relates more to the size factor (SMB) than RMW.

Both PLS methods exhibit a similar pattern in their relation to the Fama-French five factor model. The first latent factor of both approaches exhibits a connection to the market, which is much stronger than the one observed for the respective principal component. The value characteristic can explain the second, third and fourth latent factors of PLS and SPLS, while RMW and CMA equally explain the fifth factor. The results for the autoencoders remain relatively unchanged when comparing across different values of K and number of hidden layers. The value and, to a lesser extent, the market factors are those which exhibit the strongest relation to the latent factors.

We further examine the connection of the latent factors to the Fama and French five-factor model by aggregating the distributional properties of the t-statistics, estimated throughout the out-of-sample period, using boxplots.[8] We consider that a latent factor of a specific dimensionality reduction technique is linked to one of the observed factor proxies, when the median t-statistic is non-zero and when the size of the box (IQR) or the distance between the ends of the two whiskers (range) is large, indicating that the majority of the t-statistics throughout the rolling window iterations are significantly different from zero. On the contrary, the relationship of a latent factor with a particular Fama-French factor will tend to be weak if the box falls within the lines depicting the Student's t critical values at the 5% level.

[Insert Figure 3 about here]

The results based on the five factors by Fama and French (2015) are presented in Figure 3 and corroborate those from the variable importance analysis. Overall, there is substantial significance for the latent factors based on PCA and PLS methods with some of the Fama-French factors, and especially the market factor. When comparing PCA with SPCA, the relationship of the latent factors with the observed factors is diminished for the latter. In contrast, both PLS approaches generate latent factors that exhibit a similar pattern. Autoencoders do not display any strong links to a particular factor, indicating that latent factors based on neural networks cannot be adequately explained by the Fama-French factors.

Focusing on the results for PCA, the first principal component, $K = 1$, is primarily related to the market factor (MKTRF), with the middle 50% of the t-statistics exceeding 5 in absolute value and to the size factor (SMB) where most of the t-statistics are around three standard errors from zero. For both factors the distribution of t-statistics is right skewed since the median is located towards the left and the left whisker is shorter, indicating a negative relationship. The second factor, $K = 2$, based on PCA does not exhibit any strong links to any

[8] Boxplots of the t-statistics provide a simple five-number summary of their distribution, which consists of the median (marked by the line within the box), first and third quartiles (the edges of the box, with its length representing the interquartile range, IQR), and the minimum and maximum individual t-statistics (depicted by the two lines or whiskers, with the distance from the end of each line representing the range).

of the observed proxies, since the t-statistics are less dispersed, as evidenced by the interquartile range and values that are clustered within the 5% critical values bands. The third principal component, $K = 3$, has a strong relation to the value (HML) factor, with t-statistics in the range of -11 to 11. The fourth latent factor, $K = 4$, relates to the size and profitability (RMW) characteristics, while the fifth factor, $K = 5$, is primarily linked to the investment (CMA) factor. The pattern of the t-statistics for SPCA is similar to that of PCA, however the connection of the sparse principal components to the five factors is overall weaker, as evidenced by the smaller dispersion of the t-statistics.

Turning to the results for supervised methods, the patterns observed when comparing the boxplots of both PLS-type approaches appear similar. The first factor of PLS and SPLS is positively related to the market factor and, to a lesser extent, the size factor according to the median, with the middle 50% of the t-statistics lying approximately between -9 and 8 for MKTRF and from approximately -2 to 5 for SMB. The remaining latent factors are all positively related to the market, albeit to a lesser degree than the first factor. The second latent factor of both methods is also related to HML, according to the negatively skewed boxplot, with the lower quartile being approximately equal to -1 and the third quartile to 5. The third factor is also linked to the value factor, although the distribution of the t-statistics is more symmetrical. The fourth factor displays a weak relationship with the size and value factors, while the fifth factor is weakly related to RMW and CMA. Finally, the results for the latent factors based on autoencoders do not reveal any significantly strong links with a particular proxy. For the majority of the autoencoder factors the boxes are symmetrical and smaller, with the middle 50% of the t-statistics being within the 5% critical value bands and concentrated around the median, however, the whiskers are longer indicating a greater range of potential values over the out-of-sample period.[9]

We extend the variable importance analysis to a larger dataset, by examining the relationship between the latent factors and long-short anomaly portfolios constructed by Chen and Zimmermann (2020). The equity portfolios are based on a comprehensive reconstruction of firm-level characteristics, which have been replicated using the same data and methods as the original papers. Due to the large number of variables, the influence of each feature to the respective latent factor is derived based on the lasso (Tibshirani, 1996). The penalized model is estimated using a rolling window approach and the results are averaged throughout the out-of-sample period, from January 1980 to December 2019. The variable importance is estimated as the change in R^2 from setting the observations of a feature to zero within each iteration of the rolling window. The results are further aggregated by summing the variable importance of the characteristics-based portfolios belonging in the same group. We use the same categorization as Chen and Zimmermann (2020) to assign a variable to a group.[10] Figure 4 presents the results of the variable importance.

[Insert Figure 4 about here]

Across all different models the group comprised of portfolios sorted by accounting characteristics is the most influential according to the lasso. Important variables within this group include portfolios sorted on gross profits/total assets, dividend yield, book-to-market and leverage. The second most influential group is that of portfolios sorted on price characteristics. Individual portfolios with high variable importance values are based on characteristics such as

[9] We also examine the relationship of the latent factors with the Hou, Xue and Zhang (2015) q-factor model augmented by the expected growth factor (Hou, Mo, Xue and Zhang, 2021). The results based on OLS regressions of each of the five latent factors on factors from the augmented q-factor model are reported in Figures A3, A4 and A5 in the Appendix.
[10] Data on the long-short anomaly portfolios are obtained from Andrew Y. Chen's website. We consider portfolios that have a full history of returns from January 1960 to December 2019, which reduces the initial number of available series from 205 to 110. Details on which anomaly portfolios belong in each of the six groups can be found in Table A1 in the Appendix.

CAPM beta, earnings-to-price ratio, tail risk beta, coskewness and idiosyncratic risk. Portfolios based on trading characteristics, such as liquidity and volume indicators, also relate significantly to the majority of the latent factors. The portfolios from the remaining categories contribute a relatively small part to the total variable importance. The variable importance pattern remains similar across autoencoders with different number of hidden layers, while the results for PCA and PLS methods are less homogenous.[11]

B. Structure of the Covariance Matrices

Given the dependence of the minimum-variance portfolio on the estimates of the covariance, it is informative to investigate how different factors affect the structure of the covariance matrix. We compare the structure of four factor-implied covariance specifications using latent factors based on machine learning and observed factors such as the single index (Market) model (Sharpe, 1963) or the three-factor model (FF3) by Fama and French (1993). The differences between factor-implied covariances are examined by comparing the structure of the alternative covariance, Σ, to that of the sample estimator (Sample), S.

Following the analysis of Moskowitz (2003) we consider three measures. The first measure examines the similarity between two matrices and is given by $Eig_t = \sqrt{\mathrm{tr}(\Sigma_t'\Sigma_t)/\mathrm{tr}(S_t'S_t)}$, where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix. This metric represents the sum of the eigenvalues of the factor-implied covariance matrix as a fraction of the sum of the eigenvalues of the sample covariance matrix. The matrices are squared to capture the absolute amount of covariation. The sum of the eigenvalues provides a measure of total covariation represented by the matrix. The second measure compares the factor-implied covariance with the sample covariance in terms of magnitude and is calculated as $Mag_t = (\mathrm{i}'|S_t - \Sigma_t|\mathrm{i})/(\mathrm{i}'|S_t|\mathrm{i})$. This measure sums the absolute value of all the elements of the

[11] As an alternative we consider the macroeconomic dataset by McCracken and Ng (2015). The results based on the McCracken and Ng (2015) macroeconomic dataset can be found in Figure A6 in the Appendix, while details on the eight groups and the variables within each group are reported in Table A2.

difference between the two matrices and scales this sum with the sum of the absolute value of all the elements of the sample covariance matrix. Finally, we compare the alternative covariance estimates with the benchmark in terms of the direction of the covariances, by determining the fraction of the covariances from the alternative models that have the same sign with those of the sample covariance which is given by $Dir_t = (\mathrm{i}'\mathrm{sign}(\mathrm{S}_t \circ \Sigma_t)\mathrm{i})/\mathrm{rank}(\mathrm{S}_t)^2$.

We report the average value of the measures across the out-of-sample period in Table 1. A high Eig ratio indicates that the alternative covariance matrix is close to the sample estimator. Additionally, if the factor-implied covariance captures similar information to the sample covariance then the magnitude, Mag, should be close to zero and the direction, Dir, should be close to unity. We also examine whether the difference from the sample estimator of a factor-implied covariance is the same as that of a covariance matrix based on the market factor. The two-sided bootstrapped p-value is adjusted for autocorrelation up to 12-month lags.

[Insert Table 1 about here]

Overall, covariance matrices based on latent factors and especially those where the factors are derived from unsupervised methods, exhibit the greatest differences from the sample estimator, for all three measures. The structure of the covariances based on SPCA and shallow autoencoders are the ones that deviate most from the benchmark. In contrast, the structure of covariances using observed factors is closest to the sample estimator. When comparing across covariance specifications, the differences become more pronounced in the covariance structure when the residual covariance is dynamic. Furthermore, the information captured by the covariance matrices based on latent factors relative to the sample estimator is significantly different at the 1% level to that of the corresponding matrix based on the market factor.

Specifically, according to the Eig measure, the covariance matrices based on latent factors differ considerably from the sample estimator, by a range of 40% to 52%. In contrast, covariances based on the three Fama-French factors or the market factor, differ to the sample covariance by only 1% to 7% or 7% to 9%, respectively. The results under the magnitude and direction measures, remain consistent, with covariances based on latent factors deviating considerably from the sample covariance, while those of observed factors remain very close to the benchmark. In particular, the values of the latent factors for the Mag measure are closer to unity (0.60 to 0.77) further highlighting the difference from the sample estimator, while the results for the observed factors are much closer to zero (0.01 to 0.1). For Dir the results show that on average, 36% to 65% of the covariances have a different sign from those of the benchmark, compared to observed factors where there is only a 2.5% to 5% difference. The values for the direction measure do not significantly change across covariance specifications.

V. Asset Allocation

In this Section we first explore the out-of-sample performance of the portfolios based on the different factor and covariance specifications using a variety of performance measures, analyze the properties of the portfolio weight vectors and investigate the effects of transaction costs on portfolio performance. We also analyze the behavior of the portfolios during high and low volatility subperiods and for a different number of assets.

The buy-and-hold portfolio returns are calculated for the period of one month and the portfolio is rebalanced monthly until the end of the evaluation period (January 1980 to December 2019). The portfolios based on latent factors are compared to the sample estimator and covariances based on observed factors and to the equally weighted portfolio, a scheme which requires no parameter estimation, since the weights are $\omega_i = 1/N$, for $i = 1, \ldots, N$.

A. Portfolio Performance

We focus on evaluating the performance of the portfolios based on measures of risk and risk-adjusted returns, since equation (8) is designed to minimize variance rather than maximize the expected return.[12] 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.[13] 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 480 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.

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. To avoid a multiple-testing problem and since one of the major goals of this study is to outperform the $1/N$ rule, we restrict our focus to the comparison of the EW with the alternative portfolios. For each case, a two-sided p-value is obtained by the prewhitened HAC_{PW} test proposed by Ledoit and Wolf (2011) for the null hypothesis of equal standard deviations and by Ledoit and Wolf (2008) for the null hypothesis of equal Sharpe ratios.

[Insert Table 2 about here]

The results indicate that portfolios using machine learning consistently outperform the EW benchmark in terms of standard deviation and Sharpe ratio by a wide margin. Using a covariance matrix based on machine learning factors, can lead to a statistically significant

[12] The baseline results are for minimum-variance portfolios with short-selling constraints. We also consider two alternative strategies based on portfolios in the absence of short-selling constraints and turnover-constrained portfolios (Olivares-Nadal and DeMiguel, 2018). The standard deviation and Sharpe ratio of the two alternative strategies are reported in Tables A3 and A4 in the Appendix.

[13] Tail risk is of particular importance for portfolios during periods of financial distress. Therefore, we examine the portfolio performance using several alternative risk measures, including the mean absolute deviation (MAD), value-at-risk (VaR) and conditional value-at-risk (CVaR) of the out-of-sample portfolio returns in excess of the risk-free rate. The results are reported in Table A5 in the Appendix. Results for the alternative risk criteria also point towards the benefits of using machine learning latent factors for covariance estimation, with all models outperforming the EW benchmark and offering a measure of protection to investors concerned with tail risk.

decrease in out-of-sample standard deviation of up to 29% and a significant increase in Sharpe ratio of over 25% relative to the $1/N$ portfolio. Factors based on SPCA and autoencoders are found to yield the best performance, with shallower neural networks outperforming those with more hidden layers. Additionally, estimators that allow the loadings or error covariance to be time-varying outperform the static or dynamic factor covariance specifications.

Unsupervised learning approaches used to derive the latent factors tend to produce portfolios with lower standard deviation. The best performing model is AEN1, with monthly standard deviation of 3.216% and 3.318% depending on covariance specification, except for dynamic beta covariance where SPLS yields the lowest standard deviation with a value of 3.261%. The outperformance of the shallow autoencoder (AEN1) from the EW strategy is between 2.9% and 3.3% per annum, varying based on the covariance specification. The shallow autoencoder and SPCA have the highest SR, with values between 0.82 and 0.85 in annual terms. AEN1 has the highest 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.

The outperformance over $1/N$ portfolio across factors and model specifications in terms of standard deviation is statistically significant at the 1% level. The outperformance in terms of Sharpe ratio is, however, statistically significant only for latent factor models. For the Sharpe ratio, latent factors generate statistically significant outperformance, compared to observed factors and the sample estimator that yield insignificant results. Specifically, SPCA consistently outperforms the $1/N$ portfolio at the 1% level, while AEN1 is significant at the 5% or 1% level depending on the covariance specifications.[14]

[14] The results throughout the paper are based on models using a validation window consisting of the last 20% of the observations in the rolling window. We also examine the performance of the portfolios when the size of the validation window is reduced to 10% or increased to 30% of the observations in the rolling window. The results

We further investigate the economic value of the covariance matrix estimates using the certainty equivalent return (CER), defined as: $\text{CER} = \bar{r}_P - 0.5\gamma\bar{\sigma}_P^2$, where $\bar{r}_P$ and $\bar{\sigma}_P^2$ are the mean and variance of the portfolio returns over the out-of-sample period. The CER can be interpreted as the risk-free return that a mean-variance investor with a coefficient of relative risk aversion γ is willing to accept instead of investing in the risky portfolio. We consider values of the risk aversion parameter γ, that roughly correspond to an aggressive ($\gamma = 2$), moderate ($\gamma = 5$) or conservative ($\gamma = 10$) investor.[15] Given that the optimization objective is to minimize variance, the portfolios would be of interest to investors more sensitive to risk, which makes CER with higher values of γ a more representative performance measure.

Following Neely, Rapach, Tu and Zhou (2014), we report the difference in monthly CER (ΔCER), which is equivalent to the percentage CER generated by the alternative portfolio minus that of the EW benchmark. ΔCER can be interpreted as the performance fee that the investor would be willing to pay to use the information of each alternative covariance estimator instead of the benchmark. To test the statistical significance of the CER gains, relative to the EW allocation approach, we use the two-sided p-value obtained by the test developed by DeMiguel, Garlappi and Uppal (2009), for the null hypothesis of equal CER.

[Insert Table 3 about here]

The results reported in Table 3 show that portfolios based on machine learning outperform simpler benchmarks, leading to statistically significant utility gains that exceed those of the EW by 2.5% to 4.5% on an annual basis, for investors with moderate or conservative risk preferences, respectively. Ideally the results would remain statistically significant across all values of γ. However, since the objective of this asset allocation exercise is to minimize variance, the portfolios would primarily be of interest to investors that are not

are reported in Table A6 in the Appendix and are qualitatively similar to the baseline case, with methods such as SPCA and autoencoders favoring the longer validation window.

[15] Similar values for the parameter of risk aversion have been used in DeMiguel, Garlappi and Uppal (2009).

willing to undertake a high degree of risk, while more aggressive investors would opt for an objective function that seeks to maximize return or Sharpe ratio. The CER for the $1/N$ strategy decreases as the parameter of risk aversion increases, turning negative when $\gamma = 10$. All portfolios generate positive ΔCER, with the outperformance from the EW portfolio increasing as the parameter of risk aversion, γ, increases in value. The best performing model in the case of the static factor covariance is SPCA with monthly ΔCER between 0.109% and 0.346% depending on the degree of risk aversion. In the case of dynamic estimators, autoencoders yield the highest ΔCER for $\gamma = 10$. For $\gamma = 5$, AEN1 outperforms the other models for dynamic beta covariance, while SPCA performs best when $\Sigma_{f,t}$ and $\Sigma_{u,t}$ are time varying.

B. Properties of Portfolio Weights

We now explore how the weighing structure of the portfolios based on machine learning differs compared to the naïve allocation, which assigns equal weights to all assets, and to that of the sample estimator and the observed factors.

We start by analyzing the properties of the portfolio weights, $\widehat{\omega}$, using the maximum weight (MAX), the standard deviation of the portfolio weights (SD_{ω}) 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. The similarities of each strategy with the $1/N$ are examined by using the mean absolute deviation from the equally weighted portfolio ($\mathrm{MAD}_{\mathrm{EW}}$) calculated as

$1/N \sum_{i=1}^{N} |\widehat{\omega}_i - 1/N|$ and the percentage of weights greater than $1/N$ ($\omega_i > 1/N$). Table 4 reports the average value of each weight characteristic over the out-of-sample period.[16]

[Insert Table 4 about here]

Overall, the portfolios based on machine learning methods tend to produce weights which are smaller, less volatile, and closer to those of the EW portfolio, than models based on observed factors. In addition, covariance matrices based on latent factors lead to more diversified portfolios that require less frequent rebalancing, which is a positive indicator regarding the effects of transaction costs on portfolio performance. The lowest turnover is produced by portfolios based on PCA and SPCA factors (approximately 30%), followed by portfolios using autoencoder factors with turnover varying from 31% to 50%. Neural networks with more hidden layers have consistently higher rebalancing requirements than shallower networks. Comparing across different covariance specifications, portfolios based on the dynamic error covariance exhibit higher maximum weights, SD_{ω} and turnover, are more concentrated and diverge more from the EW, than the static factor covariance.

We also examine how the quantiles of the portfolio weight vectors vary throughout the out-of-sample period. The results for ten quantiles $\tau \in [0.1,1]$ are presented in Figure 5 for the case of the static factor covariance specification.[17]

[Insert Figure 5 about here]

The weights of portfolios based on latent factors are more varied across quantiles, while those of the sample estimator and observed factors are zero for quantiles below 0.8. The weights of the Market and FF3 exhibit similar behavior, while unsupervised methods lead to portfolio weights that are more dissimilar throughout time.

---

[16] The properties of the portfolio weight vectors for minimum-variance portfolios that allow short-selling and portfolios with a turnover penalty are reported in Table A5 and Table A6 respectively, found in the Appendix.

[17] The results for the remaining covariance specifications are presented in Figures A7, A8 and A9 in the Appendix, for the cases when B, Σ_f and Σ_u are dynamic, respectively. Additionally, the results for autoencoders with two, three and four hidden layers were similar to those of an autoencoder with a single hidden layer and are not presented for the sake of brevity.

C. Effects of Transaction Costs

Following Han (2006), we consider the monthly breakeven transaction costs in basis points that would cause an investor to be indifferent between a certain strategy and the benchmark strategy.[18] The breakeven transaction costs in terms of Sharpe ratio of a portfolio relative to the equally weighted portfolio, c_{ew}, are calculated as $\Delta\text{SR}/\Delta\text{TO}$, where $\Delta\text{SR} = SR_p - SR_{ew}$ is the difference in Sharpe ratios of the alternative portfolio from the equally weighted benchmark and $\Delta\text{TO} = TO_p - TO_{ew}$ is the difference in the respective average turnover. Breakeven transaction costs become important in the absence of reliable estimates of transaction costs, with positive values of c_{ew} indicating that the alternative model outperforms the benchmark.

[Insert Table 5 about here]

Table 5 shows that all models generate positive c_{ew}, indicating outperformance from the benchmark. However, an investor would realize greater economic benefits over the EW portfolio by choosing machine learning portfolios, since they exhibit higher breakeven transaction costs by a factor of three from portfolios based on observed factors. The breakeven transaction costs are higher for unsupervised methods than the sample estimator or the remaining factor models. Furthermore, static and dynamic factor covariance strategies produce the highest breakeven transaction costs across all specifications. Portfolios with factors estimated by SPCA have monthly breakeven transaction costs between 13.9 and 20.1 bps, while a shallow autoencoder would realize c_{ew} from 13.3 to 18.5 bps. On the contrary, the worst performing factor-based portfolios are those of the market factor, which yield the lowest breakeven transaction costs (4 to 6.4 basis points) and the sample estimator with breakeven transaction costs of 6.4 basis points. This concurs with the earlier evidence from portfolio

[18] We also examine the performance of the portfolio for specific levels of transaction costs. The results are reported in Table A9 in the Appendix, for transaction costs of $c \in \{5, 20\}$ basis points. The value of 5 bps may be low by academic standards, where values as high as 50 bps (Kirby and Ostdiek, 2012) have been used. Other studies are less conservative and use a range of values. Ledoit and Wolf (2017) consider values of $c \in \{3, 50\}$ bps, pointing out that 3 bps is representative for liquid stocks, while Moura, Santos and Ruiz (2020), use $c \in \{5, 10\}$ bps.

turnover, which shows that these strategies have lower rebalancing requirements. The results for average turnover and breakeven transaction costs suggest that rebalancing frequency is a key contributor to portfolio performance. This is in keeping with the literature (see Han, 2006), that finds increasing breakeven transaction costs for reduced rebalancing frequency.[19]

D. Subperiod Analysis

In this section we examine portfolio performance during different subperiods as defined by market volatility. The impact of different market regimes on asset allocation has been well documented in the literature (see Ang and Bekaert, 2002; Guidolin and Timmermann, 2008; Pettenuzzo and Timmermann, 2011). To analyze the results under different market regimes, we derive the periods of high and low market volatility using a Markov Switching model.[20]

[Insert Table 6 about here]

During high volatility periods (Panel A) all portfolios outperform the EW benchmark in terms of both standard deviation and Sharpe ratio.[21] Strategies based on machine learning, especially unsupervised methods, generate the highest statistically significant results. Specifically, the best performing strategy is that of a shallow autoencoder with monthly SR ranging from 0.197 to 0.205 and significant at the 1% level. An exception is the case of covariances with loadings allowed to vary over time, which leads to the FF3 model having the highest performance with a ratio of 0.213 that is significant at the 5% level. For low volatility periods (Panel B) machine learning portfolios are less risky by up to 6%, albeit statistically indistinguishable from the $1/N$ portfolio. Furthermore, portfolios based on machine learning

[19] The performance in terms of c_{ew} for portfolios without short-selling constraints or with a turnover penalty are reported in Tables A10 and A11 in the Appendix.

[20] 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.

[21] The properties of the portfolio weight vectors during high and low volatility subperiods can be found in Tables A12 and A13 in the Appendix.

are statistically indistinguishable from the EW portfolio also in terms of Sharpe ratio. In contrast, portfolios based on observed factors are more negatively affected than those based on latent factors and exhibit significantly reduced performance from the benchmark.[22]

E. 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 are presented in Figure 6 for the case of the static factor covariance specification.[23]

[Insert Figure 6 about here]

When the number of assets changes, the $1/N$ portfolio is still consistently outperformed by the alternative strategies, with SD increasing with the size of the portfolio, and SR remaining relatively flat. Decreasing the number of assets to $N = \{30, 50\}$, simpler strategies based on observed factors tend to outperform those using latent factors in terms of SR and breakeven transaction costs, with the Market outperforming FF3. The pattern for observed factors is mixed for larger sizes, first increasing for $N = \{200, 300\}$ and then decreasing slightly for $N = \{400, 500\}$. The volatility of latent factor models slightly increases for $N = 200$ and then remains relatively stable across different portfolio sizes. The SR for latent factors is highest for $N = 100$, decreases when $N = \{200, 300\}$ and then increases again for $N = \{400, 500\}$, with latent factor models producing higher ratios than observed factors for larger portfolios.

Average monthly turnover steadily increases with the number of stocks in the portfolio, with observed factor models generating higher turnover than latent factor models consistently

[22] Asset allocation is also affected by shifts in the economic environment due to inflation and credit spread, where the presence of separate regimes has been detected. To this end, we also investigate changes in portfolio performance during subperiods of high and low inflation or credit spread. The results can be found in Table A14 and Table A15 in the Appendix.

[23] The results for the remaining covariance specifications exhibit a similar pattern to that of the static case and are presented in Figures A10, A11 and A12 in the Appendix, for the cases when B, Σ_f and Σ_u are dynamic, respectively. Furthermore, deeper autoencoders yielded similar results to those of a shallow autoencoder and are not presented for the sake of brevity.

across all portfolio sizes, indicating that portfolios based on the Market and FF3 factors are more sensitive to transaction costs. The results show that increasing the size of the portfolios has a considerable negative impact on breakeven transaction costs, with all models experiencing a sharp decrease in c_{ew}, indicating larger portfolios are more sensitive to transaction costs. Comparing across factor specifications, unsupervised methods generate the highest c_{ew} for values of $N \geq 100$ and observed factor perform best when $N < 100$.

VI. Conclusion

In this paper we explore the performance and properties of machine learning factor-based portfolios. Specifically, we examine whether factor-implied covariance matrices based on machine learning dimensionality reduction techniques can benefit minimum-variance portfolios comprised of individual stocks. Overall, our findings indicate that machine learning can help improve factor-based portfolio optimization.

When exploring the characteristics of machine learning portfolios, we find that factors based on PCA and PLS exhibit a stronger relationship with commonly used factor proxies than autoencoders. Furthermore, the structure of covariance matrices that are dynamic or based on unsupervised methods diverges the most from that of the sample estimator. Our analysis also shows that methods which induce sparsity and autoencoder neural networks tend to be the best performing models. We find that the proposed models can lead to a statistically significant reduction in portfolio volatility and an increase in the Sharpe ratios relative to the $1/N$ portfolio. These findings become more acute as an investor's sensitivity to risk increases. Unsupervised learning methods yield portfolios that require less frequent rebalancing, with weights that are less volatile and more diversified relative to their supervised counterparts or observed factors. Finally, when comparing across factor model specifications, the results indicate that models which allow the error component of the covariance matrix to vary over time can deliver increased performance but at the cost of higher portfolio turnover.

References

Ang, A., & Bekaert, G. (2002). International asset allocation with regime shifts. *The Review of Financial Studies, 15*(4), 1137-1187.

Avramov, D., & Chordia, T. (2006). Asset pricing models and financial market anomalies. *The Review of Financial Studies, 19*(3), 1001-1040.

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. (1991). 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.

Callot, L., Caner, M., Önder, A. Ö., & Ulaşan, E. (2019). A nodewise regression approach to estimating large portfolios. *Journal of Business & Economic Statistics*, 1-12.

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.

Chen, A. Y., & Zimmermann, T. (2020). Open source cross-sectional asset pricing. *Available at SSRN 3604626,* Chen, N., Roll, R., & Ross, S. A. (1986). Economic forces and the stock market. *Journal of Business, ,* 383-403.

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.

Chun, H., & Keleş, S. (2010). Sparse partial least squares regression for simultaneous dimension reduction and variable selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 72*(1), 3-25.

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)

De Nard, G., Ledoit, O., & Wolf, M. (2018). Factor models for portfolio selection in large dimensions: The good, the better and the ugly. *Journal of Financial Econometrics,*

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., 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.

Engle, R. F. (2016). Dynamic conditional beta. *Journal of Financial Econometrics, 14*(4), 643-667.

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.

Engle, R., & Colacito, R. (2006). Testing and valuing dynamic correlations for asset allocation. *Journal of Business & Economic Statistics, 24*(2), 238-253.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3-56.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics, 116*(1), 1-22.

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.

Fan, J., Liao, Y., & Mincheva, M. (2013). Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society.Series B, Statistical Methodology, 75*(4), 603-680.

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

Geweke, J. (1977). The dynamic factor analysis of economic time series. *Latent Variables in Socio-Economic Models,*

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. (2021). Autoencoder asset pricing models. *Journal of Econometrics,*

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

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

Han, Y. (2006). Asset allocation with a high dimensional latent factor stochastic volatility model. *The Review of Financial Studies, 19*(1), 237-271.

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science (New York, N.Y.), 313*(5786), 504-507.

Hou, K., Mo, H., Xue, C., & Zhang, L. (2021). An augmented q-factor model with expected growth. *Review of Finance, 25*(1), 1-41.

Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies, 28*(3), 650-705.

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.

Kempf, A., Korn, O., & Saßning, S. (2015). Portfolio optimization using forward-looking information. *Review of Finance, 19*(1), 467-490.

Kirby, C., & Ostdiek, B. (2012). It's all in the timing: Simple active portfolio strategies that outperform naïve diversification. *Journal of Financial and Quantitative Analysis,* 437-467.

Kozak, S., Nagel, S., & Santosh, S. (2020). Shrinking the cross-section. *Journal of Financial Economics, 135*(2), 271-292.

Lassance, N., DeMiguel, V., & Vrins, F. D. (2020). Optimal portfolio diversification via independent component analysis. *Available at SSRN 3285156,* Ledoit, O., & Wolf, M. (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance, 15*(5), 850-859.

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.

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The Journal of Finance*, *20*(4), 587-615.

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

McCracken, M. W., & Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics, 34*(4), 574-589.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society*, 867-887.

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.

Neely, C. J., Rapach, D. E., Tu, J., & Zhou, G. (2014). Forecasting the equity risk premium: The role of technical indicators. *Management Science, 60*(7), 1772-1791.

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

Pettenuzzo, D., & Timmermann, A. (2011). Predictability of stock returns and asset allocation under structural breaks. *Journal of Econometrics, 164*(1), 60-78.

Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13*(3), 341-360.

Sargent, T. J., & Sims, C. A. (1977). Business cycle modeling without pretending to have too much a priori economic theory. *New Methods in Business Cycle Research, 1*, 145-168.

Sharpe, W. F. (1963). A simplified model for portfolio analysis. *Management Science, 9*(2), 277-293.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*(3), 425-442.

Shi, F., Shu, L., Yang, A., & He, F. (2020). Improving minimum-variance portfolios by alleviating overdispersion of eigenvalues. *Journal of Financial and Quantitative Analysis, 55*(8), 2700-2731.

Stock, J. H., & Watson, M. (2011). Dynamic factor models. *Oxford Handbooks Online,*

Zou, H., Hastie, T., & Tibshirani, R. (2006). Sparse principal component analysis. *Journal of Computational and Graphical Statistics, 15*(2), 265-286.

Tables

TABLE 1

Comparison of factor-based covariance matrices

This table reports monthly measures that compare the factor-implied covariance matrices to the sample estimator, based on total covariation (Eig), Magnitude (Mag) and direction (Dir), over the out-of-sample period from January 1980 to December 2019. The results are presented for four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). The significant deviation from the sample estimator of the alternative factor-implied covariance matrices from a covariance matrix based on the market factor is denoted by * for significance at the 1% level.

	Static Factor Covariance			Dynamic Beta Covariance			Dynamic Factor Covariance			Dynamic Error Covariance		
	Eig	Mag	Dir	Eig	Mag	Dir	Eig	Mag	Dir	Eig	Mag	Dir
Market	0.913	0.076	0.973	0.933	0.105	0.954	0.926	0.070	0.973	0.908	0.082	0.973
FF3	0.932	0.041	0.976	0.990	0.062	0.951	0.963	0.015	0.973	0.926	0.047	0.976
PCA	0.501*	0.742*	0.393*	0.517*	0.754*	0.384*	0.523*	0.725*	0.394*	0.485*	0.750*	0.393*
PLS	0.584*	0.616*	0.619*	0.603*	0.634*	0.580*	0.597*	0.601*	0.620*	0.573*	0.624*	0.619*
SPCA	0.487*	0.763*	0.479*	0.513*	0.757*	0.462*	0.489*	0.764*	0.474*	0.469*	0.771*	0.479*
SPLS	0.574*	0.626*	0.648*	0.595*	0.637*	0.610*	0.577*	0.619*	0.648*	0.563*	0.634*	0.648*
AEN1	0.480*	0.761*	0.370*	0.501*	0.770*	0.363*	0.501*	0.745*	0.368*	0.461*	0.769*	0.370*
AEN2	0.498*	0.744*	0.389*	0.501*	0.760*	0.380*	0.532*	0.720*	0.388*	0.479*	0.753*	0.389*
AEN3	0.505*	0.735*	0.404*	0.501*	0.759*	0.397*	0.551*	0.701*	0.404*	0.487*	0.743*	0.404*
AEN4	0.507*	0.734*	0.405*	0.501*	0.757*	0.396*	0.551*	0.702*	0.406*	0.489*	0.742*	0.405*

TABLE 2
Portfolio performance 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 2019. The results are presented for the equally weighted portfolio (EW) and minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). The significant outperformance of the alternative strategies from the equally weighted strategy is denoted by: *, **, and *** for significance at the 10%, 5%, and 1% level, respectively.

	SD	SR
EW	4.159	0.183
Sample	3.469***	0.209

	Static Factor Covariance		Dynamic Beta Covariance		Dynamic Factor Covariance		Dynamic Error Covariance	
	SD	SR	SD	SR	SD	SR	SD	SR
Market	3.630***	0.209	3.347***	0.218	3.640***	0.210	3.597***	0.205
FF3	3.522***	0.218	3.329***	0.241	3.538***	0.213	3.469***	0.223
PCA	3.361***	0.235**	3.290***	0.239**	3.360***	0.235**	3.268***	0.241**
PLS	3.330***	0.229	3.276***	0.232*	3.335***	0.227	3.217***	0.238*
SPCA	3.373***	0.241***	3.307***	0.243***	3.378***	0.240***	3.263***	0.246***
SPLS	3.339***	0.233*	3.261***	0.242**	3.347***	0.231*	3.225***	0.242*
AEN1	3.318***	0.239***	3.285***	0.243***	3.315***	0.239***	3.216***	0.244**
AEN2	3.352***	0.232**	3.297***	0.235**	3.350***	0.232**	3.259***	0.237**
AEN3	3.338***	0.238**	3.283***	0.240**	3.334***	0.238**	3.235***	0.246**
AEN4	3.357***	0.225*	3.301***	0.227**	3.356***	0.224*	3.256***	0.227*

TABLE 3
Portfolio performance based on certainty equivalent return

This table reports monthly portfolio performance measured using the difference in certainty equivalent return (ΔCER) for various levels of risk aversion, γ, over the out-of-sample period from January 1980 to December 2019. For the EW portfolio the monthly CER is reported as a percentage, while for the remaining portfolios the percentage ΔCER is provided, which is calculated as the difference in monthly CER between the alternative strategies and the equally weighted strategy. The results are presented for the equally weighted portfolio (EW) and minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). The significant outperformance of the alternative strategies from the equally weighted strategy is denoted by: *, **, and *** for significance at the 10%, 5%, and 1% level, respectively.

	γ=2	γ=5	γ=10
EW	0.589	0.330	-0.103
Sample	0.016	0.095	0.226*

	Static Factor Covariance			Dynamic Beta Covariance			Dynamic Factor Covariance			Dynamic Error Covariance		
	γ=2	γ=5	γ=10	γ=2	γ=5	γ=10	γ=2	γ=5	γ=10	γ=2	γ=5	γ=10
Market	0.036	0.098	0.201	0.029	0.120	0.273*	0.043	0.104	0.205	0.017	0.082	0.191
FF3	0.056	0.129	0.252	0.101	0.194	0.350***	0.040	0.111	0.231	0.065	0.144	0.276*
PCA	0.087	0.177**	0.327***	0.089	0.186**	0.348***	0.087	0.177**	0.327***	0.093	0.192*	0.358***
PLS	0.061	0.155	0.310***	0.065	0.163	0.327***	0.057	0.150	0.304***	0.074	0.178	0.352***
SPCA	0.109	0.198**	0.346***	0.104	0.200**	0.359***	0.108	0.196**	0.343***	0.106	0.206**	0.372***
SPLS	0.077	0.169	0.323***	0.092	0.192*	0.359***	0.071	0.162	0.315***	0.086	0.190	0.362***
AEN1	0.095	0.189**	0.346***	0.101	0.199**	0.362***	0.093	0.188**	0.346***	0.092	0.197**	0.371***
AEN2	0.075	0.166*	0.317***	0.077	0.173*	0.334***	0.074	0.165*	0.317***	0.075	0.176*	0.343***
AEN3	0.092	0.185**	0.339***	0.090	0.188**	0.351***	0.093	0.185**	0.340***	0.102	0.204**	0.375***
AEN4	0.052	0.143	0.294***	0.051	0.147	0.307***	0.050	0.140	0.291***	0.043	0.143	0.311***

TABLE 4
Characteristics of the portfolio weight vectors

This table presents the monthly characteristics of the portfolio weight vectors. Panel A reports the standard deviation of the weights (SD_{ω}), maximum weight (MAX) and portfolio turnover (TO), whereas the Herfindahl-Hirschman index (HHI), mean absolute deviation from the equally weighted benchmark ($\mathrm{MAD}_{\mathrm{EW}}$) and percentage of weights greater than $1/N$ ($\omega_i > 1/N$) can be found in Panel B. The average value of each weight characteristic over the out-of-sample period from January 1980 to December 2019 is reported. TO, SD_{ω}, $\mathrm{MAD}_{\mathrm{EW}}$ and $\omega_i > 1/N$ are reported as a percentage. The results are presented for the equally weighted portfolio (EW) and minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN).

Panel A Standard deviation of the weights, maximum weight and portfolio turnover

	SD_{ω}	MAX	TO
EW	0.000	0.010	1.081
Sample	3.146	0.383	41.269

	Static Factor Covariance			Dynamic Beta Covariance			Dynamic Factor Covariance			Dynamic Error Covariance		
	SD_{ω}	MAX	TO	SD_{ω}	MAX	TO	SD_{ω}	MAX	TO	SD_{ω}	MAX	TO
Market	2.824	0.302	41.297	2.452	0.296	60.745	2.769	0.433	43.957	3.433	0.445	56.045
FF3	2.738	0.311	42.785	2.473	0.312	71.744	2.669	0.450	47.423	3.342	0.420	57.039
PCA	1.091	0.070	29.737	1.048	0.078	35.064	1.098	0.073	30.017	1.323	0.139	46.098
PLS	1.419	0.095	33.408	1.334	0.092	39.694	1.423	0.094	33.496	1.700	0.157	49.269
SPCA	1.067	0.071	29.931	1.058	0.087	36.287	1.064	0.071	29.972	1.288	0.134	46.138
SPLS	1.440	0.100	33.372	1.359	0.108	40.552	1.439	0.098	33.331	1.727	0.194	49.203
AEN1	1.053	0.066	31.196	1.012	0.075	36.924	1.059	0.067	31.396	1.281	0.119	46.615
AEN2	1.078	0.066	31.060	1.030	0.079	35.694	1.089	0.068	31.367	1.305	0.123	46.855
AEN3	1.107	0.074	33.136	1.047	0.076	37.462	1.121	0.073	33.577	1.349	0.166	48.857
AEN4	1.110	0.072	33.347	1.052	0.076	37.276	1.124	0.078	33.885	1.344	0.132	49.029

Panel B Herfindahl-Hirschman index, mean absolute deviation from the $1/N$ portfolio and percentage of weights greater than $1/N$

	HHI	$\mathrm{MAD}_{\mathrm{EW}}$	$\omega_i > 1/N$
EW	0.010	0.000	0.000
Sample	0.109	1.621	16.617

	Static Factor Covariance			Dynamic Beta Covariance			Dynamic Factor Covariance			Dynamic Error Covariance		
	HHI	$\mathrm{MAD}_{\mathrm{EW}}$	$\omega_i > 1/N$	HHI	$\mathrm{MAD}_{\mathrm{EW}}$	$\omega_i > 1/N$	HHI	$\mathrm{MAD}_{\mathrm{EW}}$	$\omega_i > 1/N$	HHI	$\mathrm{MAD}_{\mathrm{EW}}$	$\omega_i > 1/N$
Market	0.091	1.616	16.227	0.072	1.535	20.654	0.090	1.582	18.000	0.129	1.711	13.171
FF3	0.085	1.582	18.454	0.072	1.520	20.792	0.084	1.546	19.242	0.124	1.669	14.640
PCA	0.022	0.819	37.398	0.021	0.791	38.342	0.022	0.824	37.277	0.028	0.928	32.421
PLS	0.030	1.050	34.058	0.028	0.997	35.352	0.030	1.052	34.083	0.039	1.165	28.840
SPCA	0.022	0.772	35.688	0.022	0.777	36.985	0.022	0.771	35.708	0.027	0.882	32.027
SPLS	0.031	1.033	33.035	0.029	0.993	33.875	0.031	1.032	33.050	0.040	1.153	27.915
AEN1	0.022	0.802	37.927	0.021	0.771	38.362	0.022	0.806	37.881	0.027	0.909	32.904
AEN2	0.022	0.815	37.533	0.021	0.779	38.388	0.022	0.822	37.413	0.027	0.920	32.467
AEN3	0.023	0.835	37.217	0.021	0.793	38.304	0.023	0.844	37.042	0.029	0.943	32.206
AEN4	0.023	0.836	37.367	0.022	0.795	38.417	0.023	0.846	37.192	0.029	0.943	32.310

TABLE 5
Portfolio performance based on breakeven transaction costs

This table presents monthly portfolio performance measured breakeven transaction costs (c_{ew}) with respect to the equally weighted portfolio (EW), over the out-of-sample period from January 1980 to December 2019. The results are presented for the minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). The breakeven transaction costs are reported in basis points and a positive value indicates that the alternative portfolio outperforms the EW.

Sample	6.470			
	Static Factor Covariance	Dynamic Beta Covariance	Dynamic Factor Covariance	Dynamic Error Covariance
Market	6.465	5.866	6.297	4.003
FF3	8.392	8.208	6.474	7.148
PCA	18.146	16.479	17.971	12.884
PLS	14.230	12.690	13.574	11.414
SPCA	20.104	17.043	19.729	13.983
SPLS	15.484	14.948	14.883	12.261
AEN1	18.595	16.739	18.473	13.397
AEN2	16.345	15.023	16.179	11.797
AEN3	17.158	15.668	16.925	13.187
AEN4	13.017	12.156	12.498	9.177

TABLE 6

Portfolio performance 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) and minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). The significant outperformance of the alternative strategies from the equally weighted strategy is denoted by: *, **, and *** for significance at the 10%, 5%, and 1% level, respectively.

Panel A High volatility regime

	SD	SR
EW	5.037	0.128
Sample	3.994***	0.179

	Static Factor Covariance		Dynamic Beta Covariance		Dynamic Factor Covariance		Dynamic Error Covariance	
	SD	SR	SD	SR	SD	SR	SD	SR
Market	4.075***	0.177	3.732***	0.197	4.153***	0.173	3.958***	0.180
FF3	3.978***	0.192	3.792***	0.213**	4.048***	0.175	3.846***	0.205
PCA	3.968***	0.181**	3.874***	0.183**	3.965***	0.181**	3.820***	0.187**
PLS	3.882***	0.180	3.820***	0.179	3.888***	0.178	3.700***	0.196*
SPCA	3.988***	0.188**	3.893***	0.188**	3.994***	0.187**	3.832***	0.192**
SPLS	3.896***	0.186*	3.798***	0.191*	3.907***	0.183	3.714***	0.200*
AEN1	3.904***	0.198***	3.859***	0.198***	3.899***	0.197***	3.759***	0.205***
AEN2	3.952***	0.178*	3.878***	0.179*	3.949***	0.178*	3.819***	0.184*
AEN3	3.930***	0.189**	3.857***	0.188**	3.923***	0.189**	3.776***	0.197**
AEN4	3.956***	0.179**	3.882***	0.178*	3.953***	0.179*	3.801***	0.182*

Panel B Low volatility regime

	SD	SR
EW	2.283	0.411
Sample	2.493*	0.298*

	Static Factor Covariance		Dynamic Beta Covariance		Dynamic Factor Covariance		Dynamic Error Covariance	
	SD	SR	SD	SR	SD	SR	SD	SR
Market	2.846***	0.285	2.677***	0.269**	2.705***	0.307	2.986***	0.258*
FF3	2.705***	0.288	2.485*	0.319	2.602**	0.315	2.820***	0.267*
PCA	2.157	0.414	2.14*	0.423	2.160	0.414	2.199	0.410
PLS	2.268	0.377	2.229	0.393	2.271	0.377	2.319	0.357
SPCA	2.150*	0.423	2.151*	0.423	2.151*	0.423	2.148	0.420
SPLS	2.269	0.378	2.230	0.396	2.269	0.378	2.316	0.361
AEN1	2.171	0.381	2.163*	0.392	2.173	0.381	2.167	0.372
AEN2	2.163	0.409	2.154*	0.415	2.165	0.409	2.169	0.402
AEN3	2.174	0.399	2.158	0.408	2.177	0.399	2.195	0.398
AEN4	2.176	0.377	2.160	0.388	2.181	0.376	2.207	0.366

Figures

FIGURE 1

Average R^2_{adj} of the regressions of the latent factors on the Fama and French (2015) factors

This figure shows the R^2_{adj} as a percentage based on OLS estimation results for regressions of the latent factors on factors from the Fama and French (2015) five-factor model. The average over the out-of-sample period from July 1983 to December 2019 is given. The latent factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN).

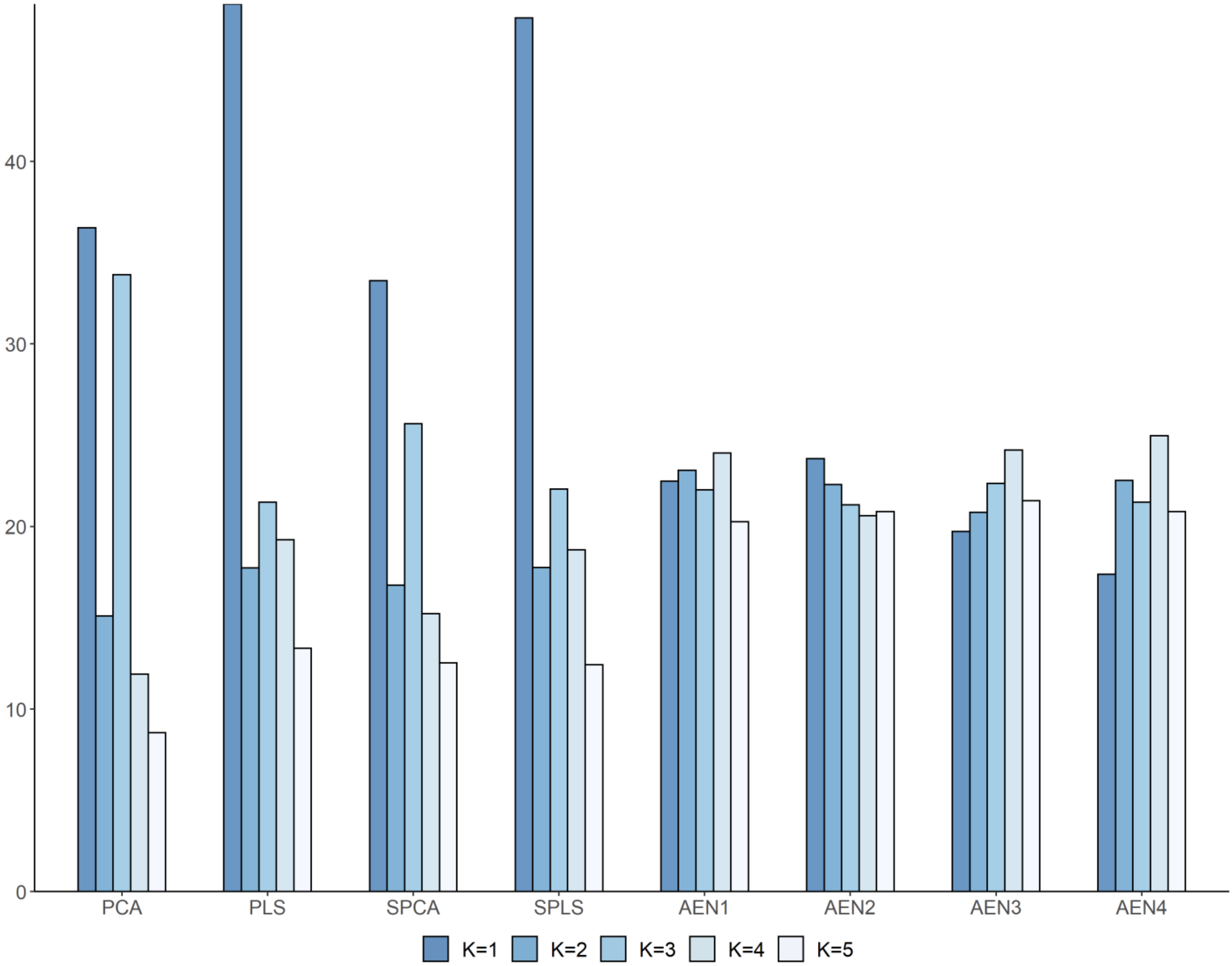


FIGURE 2

Variable importance of the Fama and French (2015) factors

This figure shows the variable importance based on OLS estimation results for regressions of the latent factors on factors from the Fama and French (2015) five-factor model. The measure of variable importance is calculated as the change in R^2 from setting the observations of a factor proxy to zero within each estimation window. The average over the out-of-sample period from July 1983 to December 2019 is given. The variable importance measures for each latent factor are scaled to sum to 100. The latent factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). MKTRF, SMB, HML, RMW, and CMA are the Fama and French (2015) excess returns of the market from the risk-free rate, size, value, profitability, and investment factors, respectively.

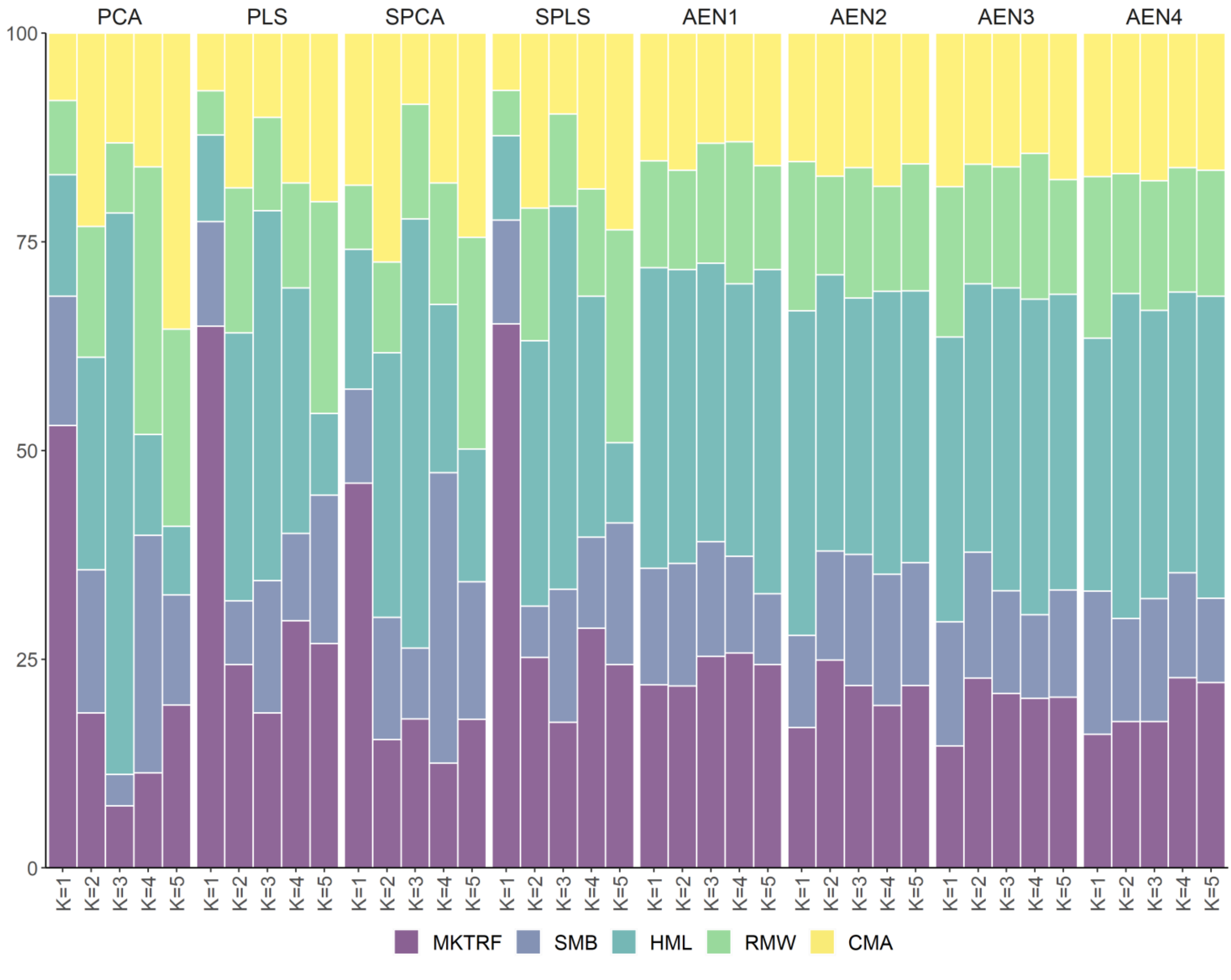


FIGURE 3

Explaining the latent factors based on the Fama and French (2015) five-factor model

This figure shows boxplots of the t-statistics based on OLS regressions of each of the five latent factors on factors from the Fama and French (2015) five-factor model. The horizontal axis reports t-statistics values ranging from -10 to 10, whereas the vertical axis reports the latent factors, $K = 1, \ldots, 5$. The sample period is from July 1963 to December 2019. The median is marked by the line within the box, the edges of the box denote the first and third quartiles, while the minimum and maximum t-statistics are depicted by the end of the lines outside the box. The latent factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). MKTRF, SMB, HML, RMW, and CMA are the Fama and French (2015) excess returns of the market from the risk-free rate, size, value, profitability, and investment factors, respectively. The t-statistics are computed using heteroskedasticity and autocorrelation-robust standard errors (Newey and West, 1987). The red lines depict the Student's t critical values at the 5% level.

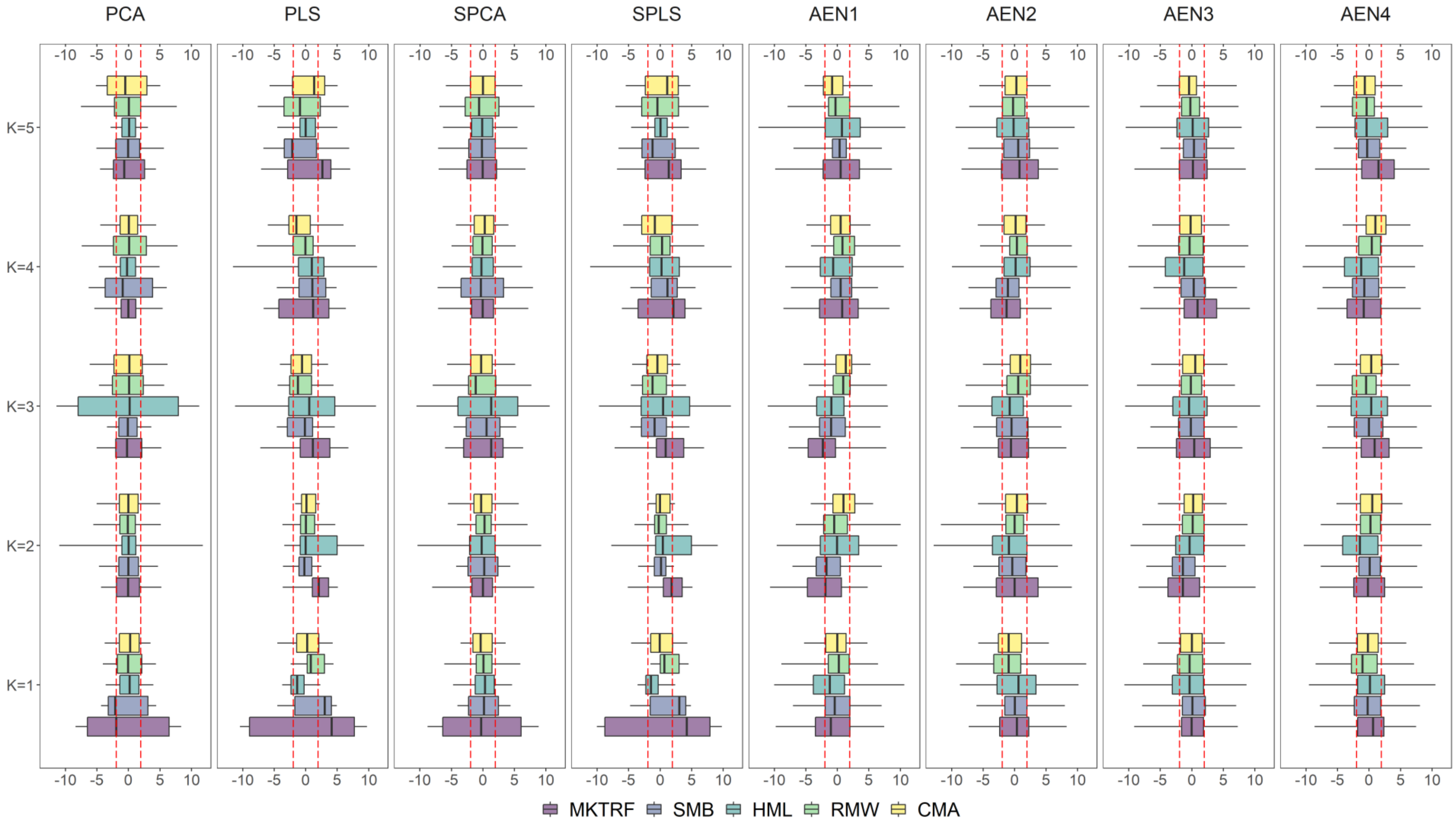


FIGURE 4

Variable importance of the long-short anomaly portfolio returns from the Chen and Zimmermann (2020) dataset

This figure shows the variable importance based on lasso regressions of the latent factors on long-short anomaly portfolio returns. The measure of variable importance is calculated as the change in R^2 from setting the observations of a feature to zero within each estimation window. The results are aggregated by summing the variable importance of the characteristics-based portfolios belonging in the same group. Details on the portfolios within each group can be found in Table A1 in the Appendix. The average over the out-of-sample period from January 1980 to December 2019 is given. The variable importance measures for each group are scaled to sum to 100. The latent factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). The explanatory variables are 110 anomaly portfolios from the open-source asset pricing dataset by Chen and Zimmermann (2020), that have no missing values over the full sample period, from January 1960 to December 2019.

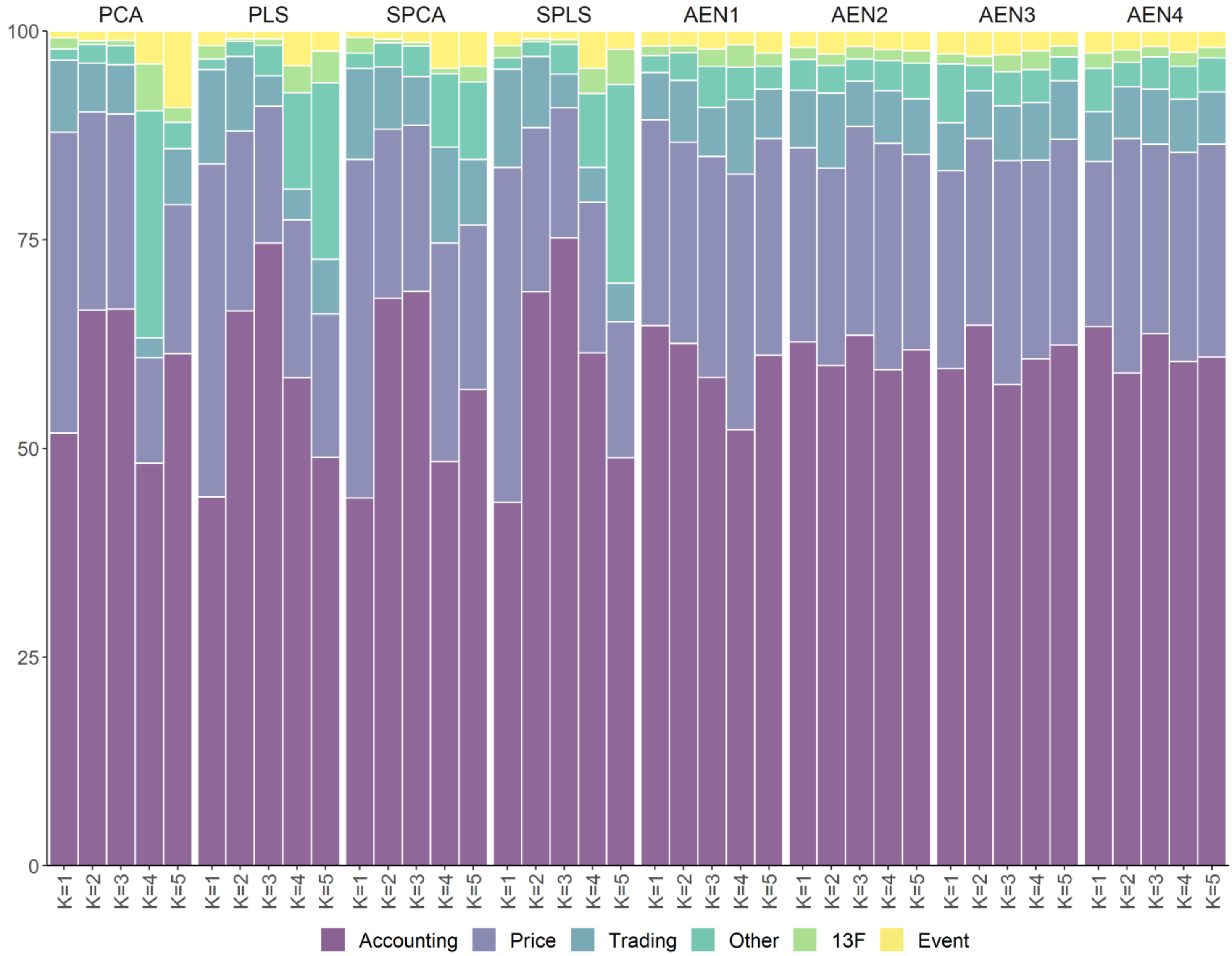


FIGURE 5

Quantiles of portfolio weight vectors: Static Factor Covariance

This figure shows the quantiles of the portfolio weight vectors across the out-of-sample period, from January 1980 to December 2019. The quantiles for $\tau \in [0.1,1]$ are depicted. The results are presented for the sample estimator (Sample) and for the static factor covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and an autoencoders with 1 hidden layer (AEN1).

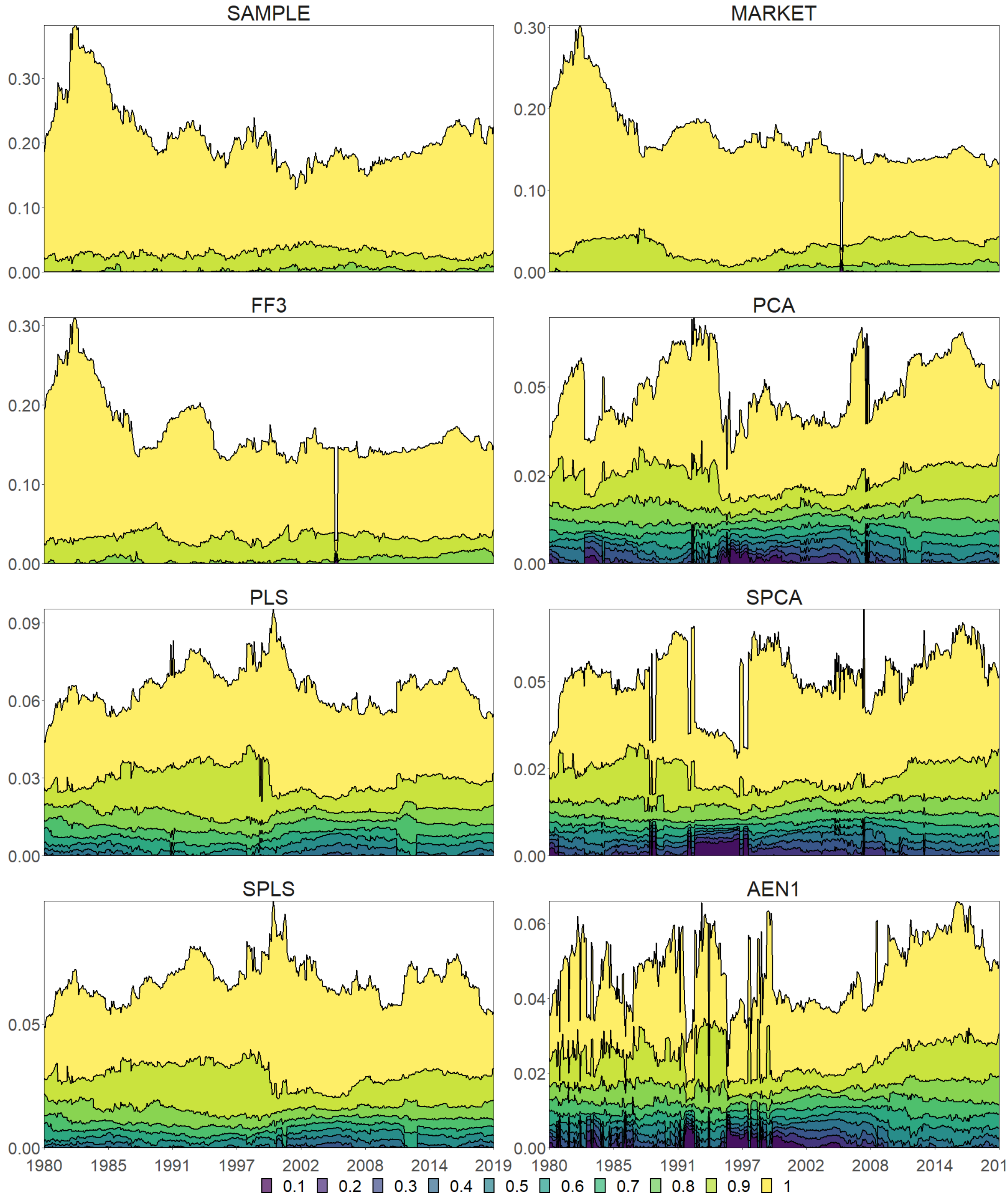
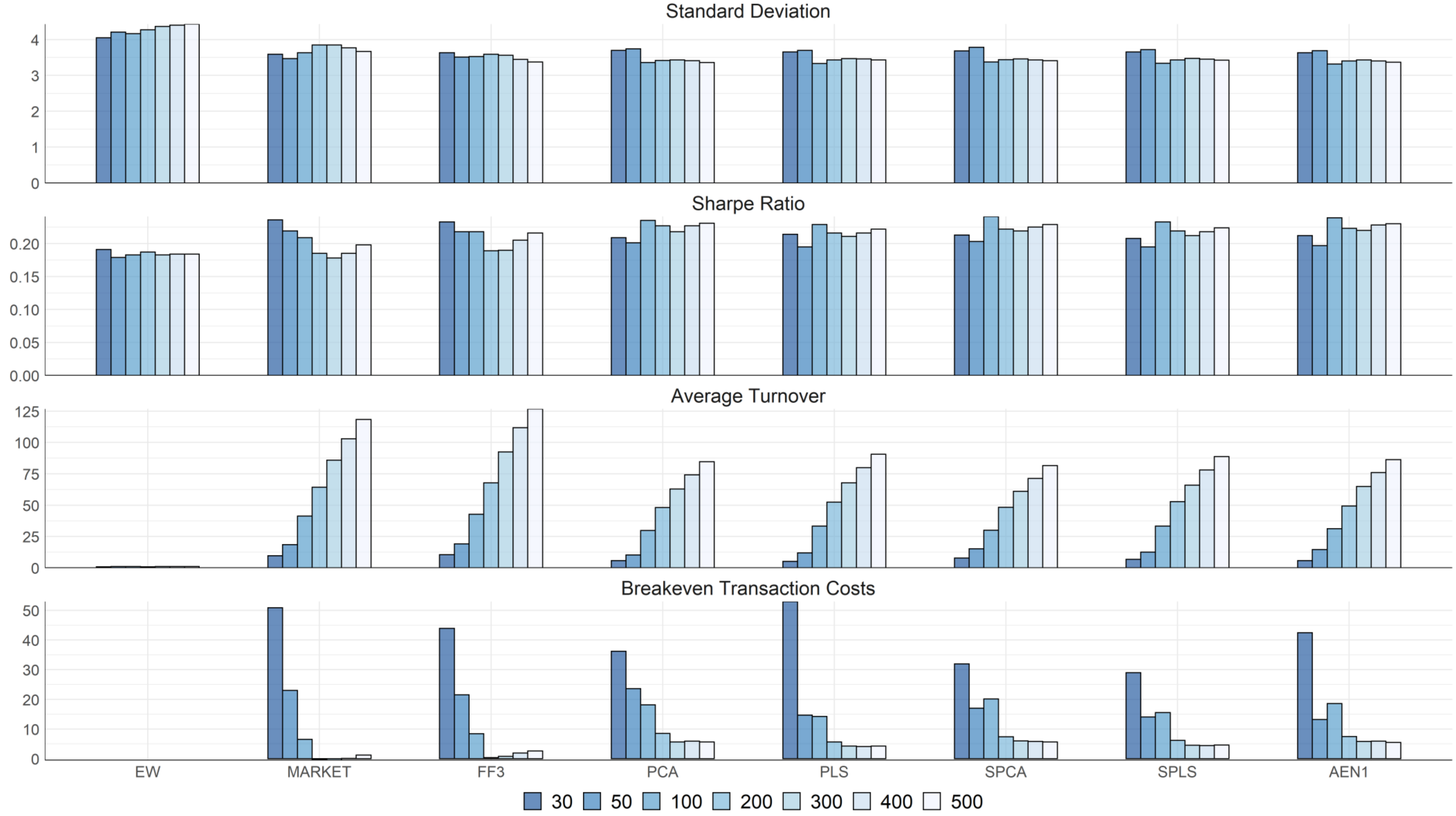


FIGURE 6

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, average turnover and breakeven transaction costs with respect to the EW portfolio. The out-of-sample period is from January 1980 to December 2019. The results are presented for the equally weighted portfolio (EW) and for the static factor covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and an autoencoder with 1 hidden layer (AEN1). The standard deviation and average turnover are reported as a percentage. The breakeven transaction costs are reported in basis points and a positive value indicates that the alternative portfolio outperforms the EW.



Appendix for
"Machine Learning and Factor-Based Portfolio Optimization"

I. Description of Machine Learning Dimensionality Reduction Methods

In this part of the Appendix we expand on the machine learning methods for dimensionality reduction used to construct the latent factors.

Principal component analysis can be viewed as a regression-type problem where the goal is to find the first K principal component weight vectors by minimizing:

$$\underset{\mathrm{W}}{\operatorname{argmin}} \|\mathrm{R}_t - \mathrm{R}_t \mathrm{W}\mathrm{W}'\|^2, \quad \text{s.t.} \quad \mathrm{W}'\mathrm{W} = \mathrm{I}_K, \tag{13}$$

where I_K is a $K \times K$ identity matrix. The solution to this problem is most often obtained via singular value decomposition: $\mathrm{R}_t = \mathrm{UDV}'$, by setting $\mathrm{W} = \mathrm{V}$. The columns of $\mathrm{V} = (v_1, \dots, v_K)$ are the principal components weights. Each v_j is used to derive the k^{th} principal component, $f_k = \mathrm{R}_t v_k$, thus, $\mathrm{F}_t\mathrm{V}$ is the dimension reduced version of the original data. The derived variable f_1 is the first principal component of R_t and has the largest sample variance amongst all linear combinations of the columns of R_t.

Partial least squares, introduced by Wold (1966), identifies the features in a supervised way, by constructing K linear combinations of R_t that have maximum correlation with the target. In order to find the PLS component matrix F_t, the columns of the weight matrix W need to be obtained through consecutive optimization problems. The criterion to find the k^{th} estimated weight vector w_k is

$$\underset{w}{\operatorname{argmax}}[w'\mathrm{M}w], \quad \text{s.t.} \quad w'w = 1, \;\; w'\Sigma_{\mathrm{RR}} w_k = 0, \tag{14}$$

where Σ_{RR} is the covariance of R_t and $\mathrm{M} = \mathrm{R}_t'\mathrm{R}_t\mathrm{R}_t'\mathrm{R}_t$. The latent factor matrix is then given by $\mathrm{F}_t = \mathrm{R}_t\mathrm{W}$. The version of PLS we employ is SIMPLS proposed by de Jong (1993).

Sparse principal component analysis (SPCA), developed by Zou, Hastie and Tibshirani (2006), is based on the regression/reconstruction property of PCA and produces modified principal components with sparse weights, such that each principal component is a linear combination of only a few of the original variables. They show how PCA can be viewed in terms of a ridge regression problem and by adding the l_1 penalty, they convert it to an elastic net regression, which allows for the estimation of sparse principal components. The following regression criterion is proposed to derive the sparse principal component weights

$$\underset{\mathrm{W,C}}{\operatorname{argmin}}[\|\mathrm{R}_t - \mathrm{R}_t \mathrm{W}\mathrm{C}'\|^2 + \lambda_1 \|w\|_1 + \lambda_2 \|w\|^2], \quad \text{s.t.} \quad \mathrm{W}'\mathrm{W} = \mathrm{I}_K, \tag{15}$$

where W and C are both $N \times K$. If $\lambda_1 = \lambda_2 = 0$, $T > N$ and we restrict $\mathrm{C} = \mathrm{W}$, then the minimizer of the objective function is exactly the first K weight vectors of ordinary PCA. When $N \gg T$, in order to obtain a unique solution, $\lambda_2 > 0$ is required. The l_1 penalty on c_k induces sparseness of the weights, with larger values of λ_1 leading to sparser solutions.

Sparse partial least squares (SPLS) is an extension of PLS that imposes the l_1 penalty to promote sparsity onto a surrogate weight vector c instead of the original weight vector w, while keeping w and c close to each other (Chun and Keles, 2010). The first weight vector solves

$$\underset{w,c}{\operatorname{argmin}}\left[-\frac{1}{2}w'\mathrm{M}w + \frac{1}{2}(c-w)'\mathrm{M}(c-w) + \lambda_1\|c\|_1 + \lambda_2\|c\|^2\right], \quad \text{s.t.} \quad w'w = 1, \tag{16}$$

where $\mathrm{M} = \mathrm{R}_t'\mathrm{R}_t\mathrm{R}_t'\mathrm{R}_t$, λ_1 and λ_2 are non-negative tuning parameters. To solve SPLS a large λ_2 value is usually required and setting $\lambda_2 = \infty$ yields a solution that has the form of the soft threshold estimator by Zou and Hastie (2005). This reduces the number of tuning parameters to two, the tuning parameter λ_1 and the number of latent factors K.

Another approach we use to construct the latent factors is based on autoencoders (Bourlard and Kamp, 1988; LeCun, Boser, Denker, Henderson, Howard, Hubbard, Jackel, 1989; Hinton and Zemel, 1994), which are a type of unsupervised neural network that can be used for dimensionality reduction. Autoencoders have a similar structure to feed-forward neural networks, which have been shown to be universal approximators for any continuous function (Hornik, Stinchcombe and White, 1989; Cybenko, 1989). However, an autoencoder differs in that the number of inputs is the same as the number of outputs and that it is used in an unsupervised context. Autoencoders have also been shown to be 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, which reduces 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 (Japkowicz, Hanson and Gluck, 2000).

An autoencoder is trying to learn an approximation to the identity function so as the output $\widehat{\mathrm{R}}_t$ is similar to the input R_t. The network consists of two parts: an encoder and a decoder. The encoder creates a compressed representation of R_t when the input variables pass through the units in the hidden layers, which are then decompressed to the output layer through the decoder. By placing constraints on the network, such as limiting the number of hidden units, it is forced to learn a compressed representation of the input, potentially uncovering an interesting structure of the data. Most often the encoding and decoding parts of an autoencoder are symmetrical, in that they both feature the same number of hidden layers with the same number of hidden units per layer. The output of the decoder is most commonly used to validate information loss, while the smallest hidden layer of the encoder (or code, at the bottleneck of the network) corresponds to the dimension-reduced data representation.

Let L denote the number of hidden layers and $K^{(l)}$ denote the number of hidden units in each layer, for $l = 1, \dots, L$, while the output of unit k in layer l is defined as the vector $z_k^{(l)}$ and the output of layer l as the matrix $\mathrm{Z}^{(l)} = \left(z_1^{(l)}, \dots, z_{K^{(l)}}^{(l)}\right)$. The original data, R_t, enters the network through the input layer ($l = 0$), while in each hidden layer inputs from the previous layer are transformed through nonlinear activation functions $h(\cdot)$ before being passed as inputs onto the next layer. The output of each hidden unit k in layer l is based on the function

$$z_k^{(l)} = h\big(\mathrm{Z}^{(l-1)}\mathrm{W}^{(l-1)} + b^{(l-1)}\big), \tag{17}$$

where $\mathrm{W}^{(l-1)}$ is a $K^{(l-1)} \times K^{(l)}$ weight matrix and $b^{(l-1)}$ is a $1 \times K^{(l)}$ bias vector. For the first hidden layer the matrix of asset returns is used as input, $\mathrm{Z}^{(0)} = \mathrm{R}_t$, such that $z_k^{(1)} = h\big(\mathrm{R}_t\mathrm{W}^{(0)} + b^{(0)}\big)$. We use the hyperbolic tangent (tanh) activation function defined as $h(x) = 2/(1 + e^{-2x}) - 1$, which is a zero-centered function whose range lies between -1 to 1. The results from each hidden layer are aggregated in the output layer

$$\widehat{\mathrm{R}}_t = h\big(\mathrm{Z}^{(L-1)}\mathrm{W}^{(L-1)} + b^{(L-1)}\big). \tag{18}$$

Since an autoencoder tries to approximate R_t the dimensions of the input and the output layer are identical, $K^{(0)} = N = K^{(L)}$. We consider four different network architectures based on the depth of the network. First, we construct a shallow autoencoder (AEN1) with a single hidden layer (the code). The other three models include additional hidden layers to the encoder and decoder representation, up to a maximum of three layers (AEN2, AEN3 and AEN4). The number of hidden nodes in each layer is selected according to the geometric pyramid rule by Masters (1993).

It can also be shown that linear autoencoders are equivalent to PCA (Baldi and Hornik, 1989; Karhunen and Joutsensalo 1995). Specifically, when the autoencoder has a single hidden layer, so the network representation becomes $\mathrm{R}_t \to \mathrm{Z}^{(1)} \to \widehat{\mathrm{R}}_t$ and all activation functions are linear, it can be shown that the $K^{(1)}$ latent variables at the bottleneck correspond to the first K principal components of the data. Hinton and Salakhutdinov (2006) show that deep autoencoders outperform shallow or linear autoencoders in image recognition tasks, however, recent applications in finance (see e.g., Gu, Kelly and Xiu, 2020) find that shallower networks generate better performance.

The parameters of the neural network are estimated by minimizing the square loss of the form

$$\underset{b,\mathrm{W}}{\operatorname{argmin}}\, \mathcal{L}\big(\mathrm{R}_t, \widehat{\mathrm{R}}_t\big) = \underset{b,\mathrm{W}}{\operatorname{argmin}} \big\|\mathrm{R}_t - \widehat{\mathrm{R}}_t\big\|^2, \quad \text{s.t.} \quad \big\|\mathrm{W}^{(l)}\big\|^2 = \mathrm{i}_{K^{(l)}}. \tag{19}$$

The estimates of the parameters of a neural network are solutions of a non-convex optimization problem. The neural network is trained using stochastic gradient descent (SGD), which evaluates the gradient from a random subset of the data and iteratively minimizes the objective function through back propagation. The version of SGD we implement is the adaptive moment estimation algorithm (Adam), introduced by Kingma and Ba (2015). Adam computes individual adaptive learning rates for the model parameters using estimates of first and second moments of the gradients.

Training a neural network can be challenging due to the large number of parameters to be estimated and the nonconvexity of the objective function. To alleviate those concerns we modify the loss function by adding a penalty on the output of the layers (activations), encouraging the activations of the nodes to be sparse (Goodfellow, Bengio and Courville, 2016). We consider activity regularization based on the elastic net penalty that shrinks the output of the bottleneck layer. Following papers such as Gu, Kelly and Xiu (2020), we implement early stopping, which prevents overfitting and significantly speeds up the training process. Specifically, the optimization process halts when the maximum number of iterations is reached or if the validation error has not improved for a certain number of consecutive iterations. In both cases the parameter estimates of the best performing model are retrieved.

References

Baldi, P., & Hornik, K. (1989). Neural networks and principal component analysis: Learning from examples without local minima. *Neural Networks, 2*(1), 53-58.

Bourlard, H., & Kamp, Y. (1988). Auto-association by multilayer perceptrons and singular value decomposition. *Biological Cybernetics, 59*(4), 291-294.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems, 2*(4), 303-314.

De Jong, S. (1993). SIMPLS: An alternative approach to partial least squares regression. *Chemometrics and Intelligent Laboratory Systems, 18*(3), 251-263.

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning*. MIT press Cambridge.

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

Hinton, G. E., & Zemel, R. S. (1994). Autoencoders, minimum description length, and helmholtz free energy. *Advances in Neural Information Processing Systems, 6,* 3-10.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks, 2*(5), 359-366.

Japkowicz, N., Hanson, S. J., & Gluck, M. A. (2000). Nonlinear autoassociation is not equivalent to PCA. *Neural Computation, 12*(3), 531-545.

Karhunen, J., & Joutsensalo, J. (1995). Generalizations of principal component analysis, optimization problems, and neural networks. *Neural Networks, 8*(4), 549-562.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *ArXiv Preprint arXiv:1412.6980.*

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation, 1*(4), 541-551.

Masters, T. (1993). *Practical neural network recipes in C*. Morgan Kaufmann.

Wold, H. (1966). Estimation of principal components and related models by iterative least squares. *Multivariate Analysis*, 391-420.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67*(2), 301-320.

II. Supplementary Tables and Figures

TABLE A1

Constituents of each group for the Chen and Zimmermann (2020) dataset

Information on the six groups of variables for the Chen and Zimmermann (2020) dataset.

Group	# of Variables	Variables
13F	2	RIO_Turnover, RIO_Volatility
Accounting	46	Accruals, AM, AssetGrowth, BMdec, BookLeverage, CashProd, CF, ChAssetTurnover, ChInv, ChInvIA, ChNNCOA, ChNWC, CompEquIss, CompositeDebtIssuance, DelCOA, DelCOL, DelFINL, DelLTI, DelNetFin, DivYieldST, EarningsConsistency, EntMult, GP, grcapx, grcapx3y, GrLTNOA, GrSaleToGrInv, GrSaleToGrOverhead, IntanCFP, IntanEP, IntanSP, Investment, InvestPPEInv, InvGrowth, MeanRankRevGrowth, OPLeverage, OrgCap, RD, RDAbility, ShareIss1Y, ShareIss5Y, SP, tang, Tax, TotalAccruals, VarCF
Event	5	ConvDebt, DivInit, DivSeason, Spinoff, SurpriseRD
Other	5	FirmAge, Herf, HerfAsset, HerfBE, sinAlgo
Price	41	Beta, BetaFP, BetaTailRisk, Coskewness, EP, High52, IdioRisk, IdioVol3F, IdioVolAHT, IndMom, IndRetBig, IntMom, Leverage, LRreversal, MaxRet, Mom12m, Mom12mOffSeason, Mom6m, MomOffSeason, MomOffSeason06YrPlus, MomOffSeason11YrPlus, MomOffSeason16YrPlus, MomRev, MomSeason, MomSeason06YrPlus, MomSeason11YrPlus, MomSeason16YrPlus, MomSeasonShort, MomVol, MRreversal, NetPayoutYield, PayoutYield, Price, PriceDelayRsq, PriceDelaySlope, PriceDelayTstat, ResidualMomentum, ReturnSkew, ReturnSkew3F, Size, STreversal
Trading	11	BidAskSpread, DolVol, Illiquidity, ShareVol, std_turn, VolMkt, VolSD, VolumeTrend, zerotrade, zerotradeAlt1, zerotradeAlt12

Table A2

Constituents of each group for the McCracken and Ng (2015) dataset

Information on the eight groups of variables for the McCracken and Ng (2015) dataset.

Group	# of Variables	Variables
Consumption, orders, and inventories	7	DPCERA3M086SBEA, CMRMTSPLx, RETAILx, AMDMNOx, AMDMUOx, BUSINVx, ISRATIOx
Housing	5	HOUST, HOUSTNE, HOUSTMW, HOUSTS, HOUSTW
Interest and exchange rates	19	FEDFUNDS, TB3MS, TB6MS, GS1, GS5, GS10, AAA, BAA, TB3SMFFM, TB6SMFFM, T1YFFM, T5YFFM, T10YFFM, AAAFFM, BAAFFM, EXSZUSx, EXJPUSx, EXUSUKx, EXCAUSx
Labor market	31	HWI, HWIURATIO, CLF16OV, CE16OV, UNRATE, UEMPMEAN, UEMPLT5, UEMP5TO14, UEMP15OV, UEMP15T26, UEMP27OV, CLAIMSx, PAYEMS, USGOOD, CES1021000001, USCONS, MANEMP, DMANEMP, NDMANEMP, SRVPRD, USTPU, USWTRADE, USTRADE, USFIRE, USGOVT, CES0600000007, AWOTMAN, AWHMAN, CES0600000008, CES2000000008, CES3000000008
Money and credit	14	M1SL, M2SL, M2REAL, BOGMBASE, TOTRESNS, NONBORRES, BUSLOANS, REALLN, NONREVSL, CONSPI, MZMSL, DTCOLNVHFNM, DTCTHFNM, INVEST
Output and income	16	RPI, W875RX1, INDPRO, IPFPNSS, IPFINAL, IPCONGD, IPDCONGD, IPNCONGD, IPBUSEQ, IPMAT, IPDMAT, IPNMAT, IPMANSICS, IPB51222S, IPFUELS, CUMFNS
Prices	20	WPSFD49207, WPSFD49502, WPSID61, WPSID62, OILPRICEx, PPICMM, CPIAUCSL, CPIAPPSL, CPITRNSL, CPIMEDSL, CUSR0000SAC, CUSR0000SAD, CUSR0000SAS, CPIULFSL, CUSR0000SA0L2, CUSR0000SA0L5, PCEPI, DDURRG3M086SBEA, DNDGRG3M086SBEA, DSERRG3M086SBEA
Stock market	4	SP500, SPindust, SPdivyield, SPPEratio

TABLE A3

Portfolio performance when short-selling is allowed

This table documents monthly portfolio performance, based on minimum-variance portfolios without short-selling constraints, measured using the standard deviation (SD) and Sharpe ratio (SR), over the out-of-sample period from January 1980 to December 2019. In this case the portfolio objective becomes

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

The results are presented for the equally weighted portfolio (EW) and minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). The significant outperformance of the alternative strategies from the equally weighted strategy is denoted by: *, **, and *** for significance at the 10%, 5%, and 1% level, respectively.

	SD	SR
EW	4.159	0.183
Sample	4.380	0.163

	Static Factor Covariance		Dynamic Beta Covariance		Dynamic Factor Covariance		Dynamic Error Covariance	
	SD	SR	SD	SR	SD	SR	SD	SR
Market	4.109	0.211	3.639**	0.231	4.136	0.207	4.332	0.214
FF3	4.120	0.214	3.515***	0.242	4.163	0.202	4.441	0.206
PCA	3.387***	0.231**	3.305***	0.238**	3.386***	0.231**	3.292***	0.240**
PLS	3.41***	0.224	3.289***	0.234*	3.411***	0.223	3.308***	0.237
SPCA	3.425***	0.235**	3.342***	0.241***	3.426***	0.235**	3.305***	0.241**
SPLS	3.432***	0.223	3.279***	0.239*	3.435***	0.222	3.326***	0.234
AEN1	3.328***	0.239***	3.283***	0.244***	3.324***	0.239***	3.226***	0.245**
AEN2	3.375***	0.229**	3.309***	0.235**	3.373***	0.229**	3.28***	0.234**
AEN3	3.362***	0.237**	3.298***	0.241**	3.359***	0.237**	3.258***	0.246**
AEN4	3.374***	0.225*	3.316***	0.226*	3.373***	0.225*	3.278***	0.228*

TABLE A4

Portfolio performance using a penalized minimum-variance objective function

This table documents monthly portfolio performance, on minimum-variance portfolios with a turnover penalty, measured using the standard deviation (SD) and Sharpe ratio (SR), over the out-of-sample period from January 1980 to December 2019. The new constrained optimization problem becomes

$$\underset{\omega}{\operatorname{argmin}}\, \omega'\hat{\Sigma}_r\omega + \kappa\|\omega - \omega_0\|_1, \quad \text{s.t.} \quad \omega' \mathrm{i}_N = 1, \;\; \omega_i \geq 0, \text{for } i = 1, \dots, N$$

where $\kappa = 5$ 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. The results are presented for the equally weighted portfolio (EW) and minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). The significant outperformance of the alternative strategies from the equally weighted strategy is denoted by: *, **, and *** for significance at the 10%, 5%, and 1% level, respectively.

	SD	SR
EW	4.159	0.183
Sample	3.626***	0.206

	Static Factor Covariance		Dynamic Beta Covariance		Dynamic Factor Covariance		Dynamic Error Covariance	
	SD	SR	SD	SR	SD	SR	SD	SR
Market	3.522***	0.228	3.331***	0.243*	3.59***	0.221	3.456***	0.225
FF3	3.554***	0.232	3.33***	0.242*	3.591***	0.217	3.517***	0.229
PCA	3.645***	0.207	3.527***	0.222**	3.623***	0.209	3.593***	0.214*
PLS	3.547***	0.231**	3.444***	0.225**	3.553***	0.230**	3.458***	0.238***
SPCA	3.648***	0.216**	3.539***	0.212*	3.656***	0.215**	3.606***	0.218**
SPLS	3.551***	0.227**	3.426***	0.221*	3.563***	0.223**	3.484***	0.231**
AEN1	3.647***	0.203	3.532***	0.215*	3.626***	0.206	3.607***	0.208
AEN2	3.655***	0.204	3.538***	0.219*	3.633***	0.206	3.615***	0.213*
AEN3	3.606***	0.210*	3.514***	0.224**	3.564***	0.213*	3.565***	0.219**
AEN4	3.621***	0.206	3.524***	0.215*	3.599***	0.206	3.567***	0.218**

TABLE A5

Portfolio performance based on alternative risk measures

In this table, we present the monthly portfolio performance measured using the mean absolute deviation (MAD), value-at-risk (VaR) and conditional value-at-risk (CVaR), over the out-of-sample period from January 1980 to December 2019. The VaR and CVaR are calculated at the 95% confidence level. The results are presented for the equally weighted portfolio (EW) and minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN).

	MAD	VaR	CVaR
EW	3.693	6.079	7.817
Sample	3.111	4.981	6.431

	Static Factor Covariance			Dynamic Beta Covariance			Dynamic Factor Covariance			Dynamic Error Covariance		
	MAD	VaR	CVaR	MAD	VaR	CVaR	MAD	VaR	CVaR	MAD	VaR	CVaR
Market	3.220	5.215	6.732	3.187	4.775	6.173	3.106	5.223	6.745	3.370	5.181	6.684
FF3	3.168	5.024	6.495	3.065	4.674	6.065	3.178	5.066	6.545	3.254	4.932	6.381
PCA	2.819	4.739	6.143	2.775	4.625	6.000	2.826	4.737	6.141	2.917	4.586	5.952
PLS	2.882	4.715	6.107	2.886	4.627	5.995	2.911	4.728	6.121	2.798	4.525	5.869
SPCA	2.773	4.736	6.146	2.696	4.637	6.019	2.788	4.745	6.157	2.767	4.565	5.929
SPLS	2.963	4.715	6.111	2.897	4.577	5.940	2.952	4.734	6.132	2.883	4.526	5.874
AEN1	2.774	4.664	6.050	2.699	4.605	5.978	2.780	4.661	6.046	2.785	4.504	5.848
AEN2	2.841	4.737	6.137	2.867	4.648	6.025	2.847	4.735	6.135	2.874	4.590	5.952
AEN3	2.772	4.697	6.092	2.818	4.613	5.984	2.806	4.691	6.084	2.831	4.526	5.878
AEN4	2.893	4.767	6.169	2.853	4.681	6.061	2.894	4.768	6.171	2.878	4.618	5.978

TABLE A6

Portfolio performance for different validation window size

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 2019. The hyperparameters are selected using two alternative validation window sizes to the 20% of the baseline results: 10% (Panel A) and 30% (Panel B). The results are presented for the four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). The significant outperformance of the alternative strategies from the equally weighted strategy is denoted by: *, **, and *** for significance at the 10%, 5%, and 1% level, respectively.

Panel A Validation subsample set to 10%

	Static Factor Covariance		Dynamic Beta Covariance		Dynamic Factor Covariance		Dynamic Error Covariance	
	SD	SR	SD	SR	SD	SR	SD	SR
PCA	3.361***	0.235**	3.29***	0.239**	3.36***	0.235**	3.268***	0.241**
PLS	3.33***	0.229	3.276***	0.233*	3.335***	0.227	3.217***	0.238*
SPCA	3.387***	0.235**	3.322***	0.236**	3.392***	0.234**	3.273***	0.244***
SPLS	3.328***	0.232*	3.264***	0.236*	3.337***	0.23*	3.21***	0.241*
AEN1	3.333***	0.232**	3.288***	0.237**	3.331***	0.232**	3.235***	0.239**
AEN2	3.356***	0.23**	3.294***	0.236**	3.357***	0.229**	3.271***	0.235**
AEN3	3.353***	0.234**	3.305***	0.238**	3.351***	0.234**	3.245***	0.244**
AEN4	3.361***	0.225*	3.304***	0.228**	3.361***	0.225*	3.265***	0.234**

Panel B Validation subsample set to 30%

	Static Factor Covariance		Dynamic Beta Covariance		Dynamic Factor Covariance		Dynamic Error Covariance	
	SD	SR	SD	SR	SD	SR	SD	SR
PCA	3.367***	0.236**	3.312***	0.237**	3.368***	0.235**	3.265***	0.242**
PLS	3.349***	0.226	3.296***	0.231*	3.354***	0.225	3.231***	0.237*
SPCA	3.386***	0.237**	3.309***	0.238**	3.395***	0.235**	3.292***	0.243**
SPLS	3.338***	0.232*	3.262***	0.24**	3.346***	0.230	3.218***	0.242*
AEN1	3.34***	0.238***	3.297***	0.24***	3.338***	0.239***	3.23***	0.243**
AEN2	3.36***	0.23**	3.316***	0.229**	3.36***	0.23**	3.268***	0.234**
AEN3	3.338***	0.226*	3.302***	0.231**	3.336***	0.226*	3.241***	0.236**
AEN4	3.363***	0.228**	3.322***	0.229**	3.366***	0.229*	3.257***	0.236**

TABLE A7

Characteristics of the portfolio weight vectors when short-selling is allowed

This table presents the monthly characteristics of the portfolio weight vectors when short-selling is allowed. Panel A reports the standard deviation of the weights (SD_ω), maximum weight (MAX) and portfolio turnover (TO), whereas the Herfindahl-Hirschman index (HHI), mean absolute deviation from the equally weighted benchmark (MAD_EW) and percentage of weights greater than $1/N$ ($\omega_i > 1/N$) can be found in Panel B. The average value of each weight characteristic over the out-of-sample period from January 1980 to December 2019 is reported. TO, SD_ω, MAD_EW and $\omega_i > 1/N$ are reported as a percentage. The results are presented for the equally weighted portfolio (EW) and minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN).

Panel A Standard deviation of the weights, maximum weight and portfolio turnover

	SD_ω	MAX	TO
EW	0.000	0.010	1.081
Sample	6.317	0.378	220.992

	Static Factor Covariance			Dynamic Beta Covariance			Dynamic Factor Covariance			Dynamic Error Covariance		
	SD_ω	MAX	TO	SD_ω	MAX	TO	SD_ω	MAX	TO	SD_ω	MAX	TO
Market	2.855	0.226	78.726	2.412	0.176	96.922	2.803	0.253	79.290	3.319	0.297	117.430
FF3	3.242	0.237	93.821	2.628	0.201	121.381	3.132	0.278	96.111	3.929	0.309	144.916
PCA	1.051	0.064	30.216	1.014	0.065	35.686	1.056	0.063	30.437	1.261	0.132	47.821
PLS	1.320	0.078	35.856	1.255	0.072	42.399	1.322	0.078	35.928	1.556	0.140	55.090
SPCA	1.010	0.060	29.738	1.014	0.073	36.120	1.009	0.060	29.798	1.216	0.128	47.115
SPLS	1.316	0.082	35.123	1.266	0.077	42.530	1.316	0.081	35.075	1.554	0.163	54.283
AEN1	1.024	0.062	31.497	0.984	0.064	37.252	1.028	0.062	31.683	1.234	0.114	47.918
AEN2	1.042	0.058	31.598	0.997	0.068	36.269	1.049	0.059	31.860	1.251	0.112	48.511
AEN3	1.070	0.063	33.660	1.016	0.066	38.092	1.078	0.063	33.954	1.288	0.130	50.529
AEN4	1.071	0.067	34.001	1.018	0.069	37.986	1.080	0.067	34.410	1.283	0.128	50.839

Panel B Herfindahl-Hirschman index, mean absolute deviation from the $1/N$ portfolio and percentage of weights greater than $1/N$

	HHI	MAD_EW	$\omega_i > 1/N$
EW	0.010	0.000	0.000
Sample	0.410	4.873	47.229

	Static Factor Covariance			Dynamic Beta Covariance			Dynamic Factor Covariance			Dynamic Error Covariance		
	HHI	MAD_EW	$\omega_i > 1/N$	HHI	MAD_EW	$\omega_i > 1/N$	HHI	MAD_EW	$\omega_i > 1/N$	HHI	MAD_EW	$\omega_i > 1/N$
Market	0.095	2.122	42.087	0.071	1.864	43.906	0.093	2.082	42.044	0.125	2.327	39.610
FF3	0.117	2.383	42.885	0.081	2.001	45.460	0.113	2.313	42.231	0.169	2.765	40.913
PCA	0.021	0.785	40.904	0.021	0.763	41.294	0.021	0.789	40.919	0.026	0.894	35.206
PLS	0.028	0.979	40.656	0.026	0.940	40.817	0.028	0.981	40.683	0.034	1.095	34.835
SPCA	0.020	0.727	39.717	0.021	0.743	40.158	0.020	0.726	39.698	0.025	0.839	34.673
SPLS	0.028	0.947	40.110	0.026	0.927	40.108	0.028	0.947	40.098	0.034	1.069	34.110
AEN1	0.021	0.776	40.354	0.020	0.748	40.773	0.021	0.780	40.348	0.026	0.883	35.202
AEN2	0.021	0.783	40.633	0.020	0.752	41.179	0.021	0.788	40.640	0.026	0.888	35.069
AEN3	0.022	0.803	40.590	0.021	0.765	41.040	0.022	0.810	40.600	0.027	0.911	34.952
AEN4	0.022	0.804	40.746	0.021	0.766	41.190	0.022	0.810	40.727	0.027	0.911	35.233

TABLE A8

Characteristics of the portfolio weight vectors when using a penalized minimum-variance objective function

This table presents the monthly characteristics of the portfolio weight vectors when using a penalized minimum-variance objective function. Panel A reports the standard deviation of the weights (SD_{ω}), maximum weight (MAX) and portfolio turnover (TO), whereas the Herfindahl-Hirschman index (HHI), mean absolute deviation from the equally weighted benchmark (MAD_{EW}) and percentage of weights greater than $1/N$ ($\omega_i > 1/N$) can be found in Panel B. The average value of each weight characteristic over the out-of-sample period from January 1980 to December 2019 is reported. TO, SD_{ω}, MAD_{EW} and $\omega_i > 1/N$ are reported as a percentage. The results are presented for the equally weighted portfolio (EW) and minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN).

Panel A Standard deviation of the weights, maximum weight and portfolio turnover

	SD_{ω}	MAX	TO
EW	0.000	0.010	1.081
Sample	2.912	0.330	19.862

	Static Factor Covariance			Dynamic Beta Covariance			Dynamic Factor Covariance			Dynamic Error Covariance		
	SD_{ω}	MAX	TO	SD_{ω}	MAX	TO	SD_{ω}	MAX	TO	SD_{ω}	MAX	TO
Market	2.647	0.258	24.110	2.466	0.296	30.240	2.873	0.392	25.008	3.305	0.351	25.953
FF3	2.659	0.271	22.064	2.456	0.289	33.572	2.819	0.354	23.963	3.228	0.336	24.011
PCA	1.209	0.073	13.250	1.298	0.085	16.390	1.259	0.079	14.116	1.446	0.145	14.634
PLS	1.447	0.081	15.203	1.453	0.092	18.306	1.487	0.101	15.552	1.726	0.150	16.912
SPCA	1.188	0.072	14.039	1.290	0.087	16.826	1.212	0.076	14.307	1.447	0.168	15.759
SPLS	1.465	0.085	15.538	1.503	0.107	18.814	1.476	0.095	15.526	1.726	0.193	16.999
AEN1	1.144	0.071	13.546	1.259	0.090	17.031	1.183	0.076	14.056	1.390	0.128	15.416
AEN2	1.202	0.070	13.954	1.304	0.085	17.089	1.262	0.083	14.812	1.456	0.137	15.078
AEN3	1.225	0.074	13.808	1.299	0.080	16.851	1.286	0.085	14.660	1.478	0.139	15.326
AEN4	1.246	0.075	14.253	1.309	0.083	16.691	1.299	0.082	15.240	1.476	0.147	15.577

Panel B Herfindahl-Hirschman index, mean absolute deviation from the $1/N$ portfolio and percentage of weights greater than $1/N$

	HHI	MAD_{EW}	$\omega_i > 1/N$
EW	0.010	0.000	0.000
Sample	0.096	1.514	20.517

	Static Factor Covariance			Dynamic Beta Covariance			Dynamic Factor Covariance			Dynamic Error Covariance		
	HHI	MAD_{EW}	$\omega_i > 1/N$	HHI	MAD_{EW}	$\omega_i > 1/N$	HHI	MAD_{EW}	$\omega_i > 1/N$	HHI	MAD_{EW}	$\omega_i > 1/N$
Market	0.081	1.547	19.375	0.073	1.522	20.815	0.096	1.572	18.850	0.121	1.624	16.198
FF3	0.082	1.513	20.858	0.072	1.501	21.363	0.092	1.539	20.121	0.116	1.602	16.306
PCA	0.025	0.982	39.885	0.027	1.042	37.392	0.026	1.017	37.765	0.031	1.071	34.485
PLS	0.031	1.148	33.844	0.031	1.144	34.288	0.032	1.168	33.529	0.040	1.238	30.908
SPCA	0.024	0.984	40.862	0.027	1.051	37.565	0.025	0.990	39.317	0.031	1.054	35.823
SPLS	0.031	1.169	35.465	0.032	1.173	33.117	0.032	1.172	34.985	0.040	1.227	32.221
AEN1	0.023	0.931	39.710	0.026	1.005	38.185	0.024	0.950	38.756	0.030	1.014	37.190
AEN2	0.024	0.981	39.419	0.027	1.050	36.200	0.026	1.020	38.473	0.032	1.077	35.052
AEN3	0.025	0.999	39.288	0.027	1.048	36.367	0.027	1.039	37.442	0.033	1.097	35.677
AEN4	0.026	1.001	38.469	0.027	1.048	35.979	0.027	1.037	37.281	0.032	1.072	35.019

Table A9

Portfolio performance for different levels of 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. Given a transaction cost 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 $r_{p,t+1}^{TC} = (1 + r_{p,t+1})(1 - c\|\omega_{t+1} - \omega_t\|_1) - 1$. 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 2019. The results are presented for the equally weighted portfolio (EW) and minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). The significant outperformance of the alternative strategies from the equally weighted strategy is denoted by: *, **, and *** for significance at the 10%, 5%, and 1% level, respectively.

Panel A Transaction costs of 5 bps

	SD	SR
EW	4.159	0.183
Sample	3.470***	0.203

	Static Factor Covariance		Dynamic Beta Covariance		Dynamic Factor Covariance		Dynamic Error Covariance	
	SD	SR	SD	SR	SD	SR	SD	SR
Market	3.630***	0.203	3.346***	0.209	3.640***	0.204	3.596***	0.197
FF3	3.521***	0.212	3.327***	0.23	3.538***	0.206	3.468***	0.215
PCA	3.361***	0.230**	3.290***	0.234**	3.359***	0.230**	3.267***	0.234**
PLS	3.329***	0.224	3.276***	0.226	3.334***	0.222	3.216***	0.231
SPCA	3.373***	0.236**	3.306***	0.237**	3.377***	0.236**	3.262***	0.239**
SPLS	3.339***	0.228	3.260***	0.236*	3.346***	0.226	3.224***	0.234
AEN1	3.317***	0.235**	3.284***	0.238***	3.314***	0.234**	3.215***	0.237**
AEN2	3.351***	0.227**	3.296***	0.230**	3.349***	0.227**	3.258***	0.229*
AEN3	3.337***	0.233**	3.282***	0.234**	3.333***	0.233**	3.234***	0.238**
AEN4	3.356***	0.22	3.300***	0.221*	3.355***	0.219	3.255***	0.219

Panel B Transaction costs of 20 bps

	SD	SR
EW	4.159	0.183
Sample	3.471***	0.185

	Static Factor Covariance		Dynamic Beta Covariance		Dynamic Factor Covariance		Dynamic Error Covariance	
	SD	SR	SD	SR	SD	SR	SD	SR
Market	3.629***	0.186	3.344***	0.182	3.641***	0.186	3.595***	0.173
FF3	3.520***	0.194	3.323***	0.197	3.537***	0.186	3.464***	0.19
PCA	3.361***	0.217	3.291***	0.218	3.359***	0.217	3.266***	0.213
PLS	3.328***	0.209	3.276***	0.208	3.333***	0.207	3.213***	0.208
SPCA	3.371***	0.223*	3.305***	0.221*	3.376***	0.222*	3.26***	0.217
SPLS	3.337***	0.213	3.260***	0.217	3.345***	0.211	3.222***	0.211
AEN1	3.315***	0.220*	3.282***	0.221*	3.312***	0.220*	3.213***	0.215
AEN2	3.349***	0.213	3.295***	0.213	3.347***	0.213	3.256***	0.208
AEN3	3.335***	0.218	3.281***	0.217	3.331***	0.218	3.232***	0.216
AEN4	3.355***	0.205	3.299***	0.204	3.354***	0.204	3.253***	0.196

TABLE A10

Portfolio performance based on breakeven transaction costs when short-selling is allowed

This table presents monthly portfolio performance measured breakeven transaction costs (c_{ew}) with respect to equally weighted portfolio (EW), over the out-of-sample period from January 1980 to December 2019. The results are presented for the minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). The breakeven transaction costs are reported in basis points and a positive value indicates that the alternative portfolio outperforms the EW.

Sample	-0.909			
	Static Factor Covariance	Dynamic Beta Covariance	Dynamic Factor Covariance	Dynamic Error Covariance
Market	3.606	5.008	3.069	2.664
FF3	3.343	4.904	1.999	1.599
PCA	16.476	15.894	16.351	11.981
PLS	11.790	12.343	11.478	9.998
SPCA	18.146	16.553	18.108	12.599
SPLS	11.750	13.511	11.473	9.586
AEN1	18.411	16.864	18.299	13.237
AEN2	15.074	14.778	14.945	10.964
AEN3	16.575	15.671	16.427	12.741
AEN4	12.758	11.652	12.602	9.044

Table A11

Portfolio performance based on breakeven transaction costs when using a penalized minimum-variance objective function

This table presents monthly portfolio performance measured breakeven transaction costs (c_{ew}) with respect to the equally weighted portfolio (EW), over the out-of-sample period from January 1980 to December 2019. The results are presented for the minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). The breakeven transaction costs are reported in basis points and a positive value indicates that the alternative portfolio outperforms the EW.

Sample	12.246			
	Static Factor Covariance	Dynamic Beta Covariance	Dynamic Factor Covariance	Dynamic Error Covariance
Market	19.540	20.575	15.881	16.885
FF3	23.351	18.158	14.858	20.495
PCA	19.721	25.475	19.945	22.873
PLS	33.987	24.382	32.477	34.738
SPCA	25.467	18.419	24.195	23.845
SPLS	30.433	21.428	27.689	30.153
AEN1	16.045	20.063	17.725	17.440
AEN2	16.313	22.489	17.479	21.432
AEN3	21.213	25.999	22.091	25.974
AEN4	17.460	20.498	16.243	24.145

TABLE A12
Characteristics of the portfolio weight vectors during high volatility subperiods

This table presents the monthly characteristics of the portfolio weight vectors during high volatility subperiods 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. Panel A reports the standard deviation of the weights (SD_{ω}), maximum weight (MAX) and portfolio turnover (TO), whereas the Herfindahl-Hirschman index (HHI), mean absolute deviation from the equally weighted benchmark (MAD_{EW}) and percentage of weights greater than $1/N$ ($\omega_i > 1/N$) can be found in Panel B. The average value of each weight characteristic over the out-of-sample period from January 1980 to December 2019 is reported. TO, SD_{ω}, MAD_{EW} and $\omega_i > 1/N$ are reported as a percentage. The results are presented for the equally weighted portfolio (EW) and minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN).

Panel A Standard deviation of the weights, maximum weight and portfolio turnover

	SD_{ω}	MAX	TO
EW	0.000	0.010	1.208
Sample	3.184	0.383	47.160

	Static Factor Covariance			Dynamic Beta Covariance			Dynamic Factor Covariance			Dynamic Error Covariance		
	SD_{ω}	MAX	TO	SD_{ω}	MAX	TO	SD_{ω}	MAX	TO	SD_{ω}	MAX	TO
Market	2.896	0.302	47.216	2.395	0.296	66.893	3.050	0.433	51.436	3.333	0.445	60.647
FF3	2.791	0.311	48.873	2.415	0.312	78.049	2.928	0.450	55.147	3.179	0.381	61.934
PCA	1.067	0.065	33.122	1.076	0.078	39.918	1.077	0.073	33.592	1.279	0.129	48.594
PLS	1.421	0.095	38.636	1.391	0.092	46.131	1.427	0.094	38.707	1.674	0.157	53.274
SPCA	1.111	0.066	33.836	1.140	0.087	41.577	1.112	0.065	33.932	1.310	0.132	48.559
SPLS	1.444	0.100	38.499	1.419	0.108	46.859	1.443	0.098	38.444	1.705	0.194	52.944
AEN1	1.039	0.066	35.783	1.054	0.075	43.273	1.046	0.067	35.998	1.253	0.117	49.684
AEN2	1.078	0.066	34.409	1.079	0.079	40.743	1.092	0.067	34.865	1.289	0.123	49.069
AEN3	1.070	0.069	35.464	1.065	0.076	41.728	1.088	0.072	35.978	1.292	0.141	50.060
AEN4	1.073	0.069	36.310	1.066	0.076	42.017	1.091	0.076	37.049	1.290	0.124	50.665

Panel B Herfindahl-Hirschman index, mean absolute deviation from the $1/N$ portfolio and percentage of weights greater than $1/N$

	HHI	MAD_{EW}	$\omega_i > 1/N$
EW	0.010	0.000	0.000
Sample	0.112	1.624	16.542

	Static Factor Covariance			Dynamic Beta Covariance			Dynamic Factor Covariance			Dynamic Error Covariance		
	HHI	MAD_{EW}	$\omega_i > 1/N$	HHI	MAD_{EW}	$\omega_i > 1/N$	HHI	MAD_{EW}	$\omega_i > 1/N$	HHI	MAD_{EW}	$\omega_i > 1/N$
Market	0.094	1.629	16.125	0.069	1.516	21.594	0.104	1.661	15.069	0.123	1.693	13.875
FF3	0.088	1.592	18.201	0.070	1.503	21.556	0.097	1.621	16.837	0.113	1.649	15.451
PCA	0.022	0.806	37.503	0.022	0.820	38.122	0.022	0.813	37.354	0.027	0.899	32.358
PLS	0.030	1.052	34.000	0.030	1.042	34.524	0.031	1.055	34.052	0.038	1.153	28.878
SPCA	0.022	0.799	35.594	0.023	0.835	36.736	0.023	0.800	35.538	0.027	0.894	31.573
SPLS	0.031	1.038	32.545	0.030	1.040	33.094	0.031	1.036	32.583	0.039	1.141	27.740
AEN1	0.021	0.791	37.868	0.022	0.807	38.083	0.021	0.796	37.826	0.026	0.890	33.024
AEN2	0.022	0.814	37.476	0.022	0.819	38.017	0.022	0.824	37.278	0.027	0.906	32.250
AEN3	0.022	0.810	37.389	0.022	0.811	38.028	0.022	0.822	37.132	0.027	0.906	32.434
AEN4	0.022	0.813	37.385	0.022	0.813	38.226	0.022	0.825	37.142	0.027	0.908	32.455

TABLE A13

Characteristics of the portfolio weight vectors during low volatility subperiods

This table presents the monthly characteristics of the portfolio weight vectors during low volatility subperiods 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. Panel A reports the standard deviation of the weights (SD_{ω}), maximum weight (MAX) and portfolio turnover (TO), whereas the Herfindahl-Hirschman index (HHI), mean absolute deviation from the equally weighted benchmark (MAD_{EW}) and percentage of weights greater than $1/N$ ($\omega_i > 1/N$) can be found in Panel B. The average value of each weight characteristic over the out-of-sample period from January 1980 to December 2019 is reported. TO, SD_{ω}, MAD_{EW} and $\omega_i > 1/N$ are reported as a percentage. The results are presented for the equally weighted portfolio (EW) and minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN).

Panel A Standard deviation of the weights, maximum weight and portfolio turnover

	SD_{ω}	MAX	TO
EW	0.000	0.010	0.890
Sample	3.089	0.345	32.387

	Static Factor Covariance			Dynamic Beta Covariance			Dynamic Factor Covariance			Dynamic Error Covariance		
	SD_{ω}	MAX	TO	SD_{ω}	MAX	TO	SD_{ω}	MAX	TO	SD_{ω}	MAX	TO
Market	2.717	0.258	32.372	2.538	0.274	51.475	2.347	0.243	32.679	3.583	0.429	49.105
FF3	2.658	0.261	33.606	2.560	0.287	62.236	2.280	0.235	35.777	3.586	0.420	49.657
PCA	1.128	0.070	24.632	1.005	0.073	27.745	1.130	0.069	24.627	1.389	0.139	42.335
PLS	1.415	0.080	25.524	1.249	0.079	29.987	1.418	0.081	25.639	1.739	0.152	43.231
SPCA	1.001	0.071	24.042	0.937	0.064	28.310	0.994	0.071	24.001	1.255	0.134	42.486
SPLS	1.435	0.082	25.642	1.270	0.081	31.043	1.435	0.085	25.622	1.759	0.146	43.562
AEN1	1.074	0.066	24.280	0.949	0.073	27.352	1.079	0.067	24.456	1.323	0.119	41.986
AEN2	1.079	0.066	26.011	0.955	0.064	28.081	1.084	0.068	26.093	1.330	0.123	43.516
AEN3	1.163	0.074	29.626	1.022	0.073	31.028	1.171	0.073	29.957	1.435	0.166	47.043
AEN4	1.167	0.072	28.879	1.030	0.070	30.128	1.174	0.078	29.115	1.425	0.132	46.562

Panel B Herfindahl-Hirschman index, mean absolute deviation from the 1/N portfolio and percentage of weights greater than 1/N

	HHI	MAD_{EW}	$\omega_i > 1/N$
EW	0.010	0.000	0.000
Sample	0.105	1.617	16.729

	Static Factor Covariance			Dynamic Beta Covariance			Dynamic Factor Covariance			Dynamic Error Covariance		
	HHI	MAD_{EW}	$\omega_i > 1/N$	HHI	MAD_{EW}	$\omega_i > 1/N$	HHI	MAD_{EW}	$\omega_i > 1/N$	HHI	MAD_{EW}	$\omega_i > 1/N$
Market	0.085	1.597	16.380	0.075	1.564	19.245	0.068	1.464	22.396	0.140	1.737	12.115
FF3	0.081	1.568	18.833	0.076	1.545	19.646	0.064	1.432	22.849	0.141	1.699	13.422
PCA	0.023	0.839	37.240	0.020	0.747	38.672	0.023	0.841	37.161	0.030	0.971	32.516
PLS	0.030	1.047	34.146	0.026	0.929	36.594	0.030	1.048	34.130	0.040	1.184	28.781
SPCA	0.020	0.732	35.828	0.019	0.690	37.359	0.020	0.727	35.964	0.026	0.864	32.708
SPLS	0.031	1.026	33.771	0.026	0.922	35.047	0.031	1.026	33.750	0.041	1.170	28.177
AEN1	0.022	0.818	38.016	0.019	0.718	38.781	0.022	0.821	37.964	0.028	0.937	32.724
AEN2	0.022	0.816	37.620	0.020	0.719	38.943	0.022	0.819	37.615	0.028	0.940	32.792
AEN3	0.024	0.872	36.958	0.021	0.766	38.719	0.024	0.878	36.906	0.031	0.998	31.865
AEN4	0.024	0.872	37.339	0.021	0.768	38.703	0.024	0.877	37.266	0.031	0.996	32.094

TABLE A14
Portfolio performance during different inflation regimes

In this table, we document the monthly portfolio performance measured using the standard deviation (SD) and Sharpe ratio (SR), during periods of high (Panel A) and low (Panel B) inflation. Periods of high (low) inflation are those when inflation for the specific month is higher (lower) than the median over the out-of-sample period. Inflation is based on the year-on-year change of Consumer Price Index for all urban consumers, retrieved from FRED. The results are presented for the equally weighted portfolio (EW) and minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). The significant outperformance of the alternative strategies from the equally weighted strategy is denoted by: *, **, and *** for significance at the 10%, 5%, and 1% level, respectively.

Panel A High inflation regime

	SD	SR
EW	4.515	0.072
Sample	3.800***	0.178**

	Static Factor Covariance		Dynamic Beta Covariance		Dynamic Factor Covariance		Dynamic Error Covariance	
	SD	SR	SD	SR	SD	SR	SD	SR
Market	3.696**	0.224**	3.385***	0.261***	3.730**	0.217*	3.521***	0.258**
FF3	3.657***	0.222**	3.412***	0.261***	3.693**	0.209**	3.548***	0.258***
PCA	3.529***	0.186***	3.488***	0.200***	3.519***	0.187***	3.420***	0.206***
PLS	3.552***	0.193***	3.510***	0.203***	3.558***	0.191***	3.335***	0.237***
SPCA	3.570***	0.184***	3.502***	0.201***	3.574***	0.182***	3.450***	0.206***
SPLS	3.564***	0.199***	3.488***	0.223***	3.572***	0.195***	3.347***	0.240***
AEN1	3.552***	0.185***	3.513***	0.198***	3.547***	0.186***	3.417***	0.210***
AEN2	3.539***	0.180***	3.510***	0.193***	3.530***	0.181***	3.438***	0.204***
AEN3	3.513***	0.190***	3.498***	0.201***	3.500***	0.191***	3.404***	0.212***
AEN4	3.547***	0.175***	3.518***	0.183***	3.543***	0.175***	3.417***	0.195***

Panel B Low inflation regime

	SD	SR
EW	3.926	0.259
Sample	3.271***	0.231

	Static Factor Covariance		Dynamic Beta Covariance		Dynamic Factor Covariance		Dynamic Error Covariance	
	SD	SR	SD	SR	SD	SR	SD	SR
Market	3.598*	0.199	3.327***	0.193	3.594*	0.206	3.642	0.175
FF3	3.448**	0.216	3.285***	0.228	3.453**	0.215	3.426***	0.202
PCA	3.264***	0.265	3.176***	0.264	3.268***	0.264	3.182***	0.263
PLS	3.200***	0.252	3.140***	0.252	3.205***	0.250	3.153***	0.239
SPCA	3.258***	0.277	3.193***	0.270	3.263***	0.277	3.155***	0.271
SPLS	3.209***	0.255	3.129***	0.254	3.216***	0.253	3.159***	0.243
AEN1	3.179***	0.275	3.151***	0.272	3.178***	0.273	3.099***	0.266
AEN2	3.243***	0.264	3.173***	0.262	3.246***	0.263	3.157***	0.257
AEN3	3.237***	0.267	3.157***	0.264	3.239***	0.267	3.140***	0.267
AEN4	3.246***	0.256	3.174***	0.255	3.248***	0.255	3.165***	0.246

TABLE A15

Portfolio performance during different credit spread regimes

In this table, we document the monthly portfolio performance measured using the standard deviation (SD) and Sharpe ratio (SR), during periods of high (Panel A) and low (Panel B) credit spread. Periods of high (low) credit spread are those when the spread for the specific month is higher (lower) than the median over the out-of-sample period. Credit spread is based on the year-on-year change of Consumer Price Index for all urban consumers, retrieved from FRED. Periods of high (low) credit spread are defined by whether credit spread for a specific month is higher (lower) than the median over the out-of-sample period. The results are presented for the equally weighted portfolio (EW) and minimum-variance portfolios based on the sample estimator (Sample) and four factor-implied covariance specifications: static factor covariance, dynamic beta covariance, dynamic factor covariance and dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). The significant outperformance of the alternative strategies from the equally weighted strategy is denoted by: *, **, and *** for significance at the 10%, 5%, and 1% level, respectively.

Panel A High credit spread regime

	SD	SR
EW	5.343	0.127
Sample	3.843***	0.176

	Static Factor Covariance		Dynamic Beta Covariance		Dynamic Factor Covariance		Dynamic Error Covariance	
	SD	SR	SD	SR	SD	SR	SD	SR
Market	3.966***	0.178	3.580***	0.223	4.112***	0.167	3.850***	0.195
FF3	3.782***	0.192	3.555***	0.247**	3.964***	0.168	3.662***	0.232
PCA	4.050***	0.177	3.773***	0.192	4.053***	0.176	3.774***	0.201*
PLS	3.868***	0.185	3.701***	0.198	3.883***	0.181	3.624***	0.224*
SPCA	4.090***	0.171	3.844***	0.188	4.106***	0.169	3.769***	0.204*
SPLS	3.901***	0.186	3.674***	0.211	3.926***	0.181	3.621***	0.228*
AEN1	3.931***	0.188*	3.781***	0.195	3.923***	0.186	3.702***	0.214**
AEN2	3.988***	0.170	3.736***	0.184	3.989***	0.169	3.759***	0.196
AEN3	4.009***	0.180	3.770***	0.189	4.004***	0.181	3.737***	0.206*
AEN4	4.002***	0.184	3.764***	0.193	4.002***	0.183	3.755***	0.205*

Panel B Low credit spread regime

	SD	SR
EW	3.701	0.213
Sample	3.344**	0.222

	Static Factor Covariance		Dynamic Beta Covariance		Dynamic Factor Covariance		Dynamic Error Covariance	
	SD	SR	SD	SR	SD	SR	SD	SR
Market	3.520	0.220	3.272**	0.216	3.479	0.227	3.516	0.208
FF3	3.438	0.228	3.256***	0.238	3.394	0.230	3.409	0.220
PCA	3.109***	0.261*	3.122***	0.258	3.106***	0.262*	3.091***	0.259
PLS	3.140***	0.247	3.131***	0.246	3.141***	0.247	3.078***	0.244
SPCA	3.110***	0.273**	3.117***	0.266*	3.110***	0.273**	3.085***	0.263
SPLS	3.141***	0.253	3.120***	0.254	3.141***	0.252	3.091***	0.247
AEN1	3.098***	0.262*	3.112***	0.263**	3.098***	0.262*	3.046***	0.257
AEN2	3.122***	0.259*	3.146***	0.256	3.119***	0.259*	3.084***	0.254
AEN3	3.094***	0.264*	3.113***	0.261*	3.090***	0.264*	3.060***	0.263
AEN4	3.124***	0.243	3.141***	0.241	3.123***	0.243	3.082***	0.236

FIGURE A1

Number of stocks per month after the filters have been applied to the CRSP dataset

This figure shows the monthly number of stocks that are listed to the NYSE, AMEX, and NASDAQ stock exchanges (exchange codes 1, 2 or 3) and are ordinary common shares (share codes 10 or 11), over the full sample period from January 1960 to December 2019 (720 observations).

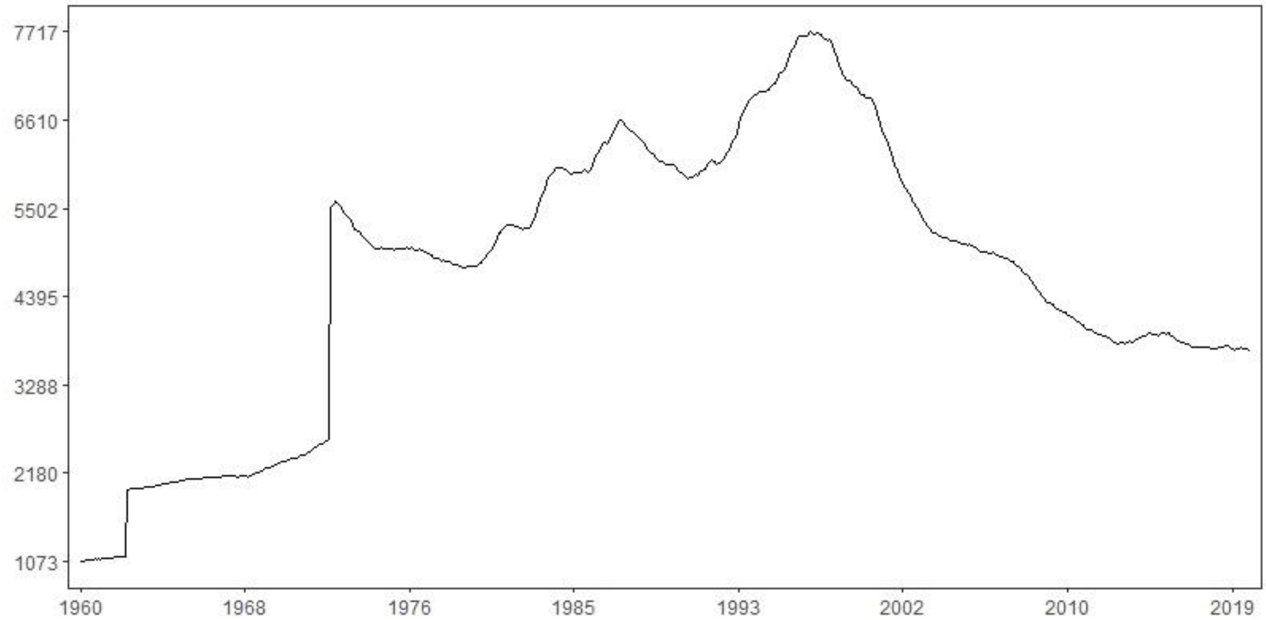


Figure A2

Number of stocks per month that fulfil the conditions of the rolling window

This figure shows the number of stocks in each iteration of the rolling window. Stocks are considered if they have at least 97.5% of the in-sample observations available, if they are not missing a return observation for the next month after the end of the rolling window and have a price greater than $5. The out-of-sample period is from December 1980 to December 2019 (480 observations).

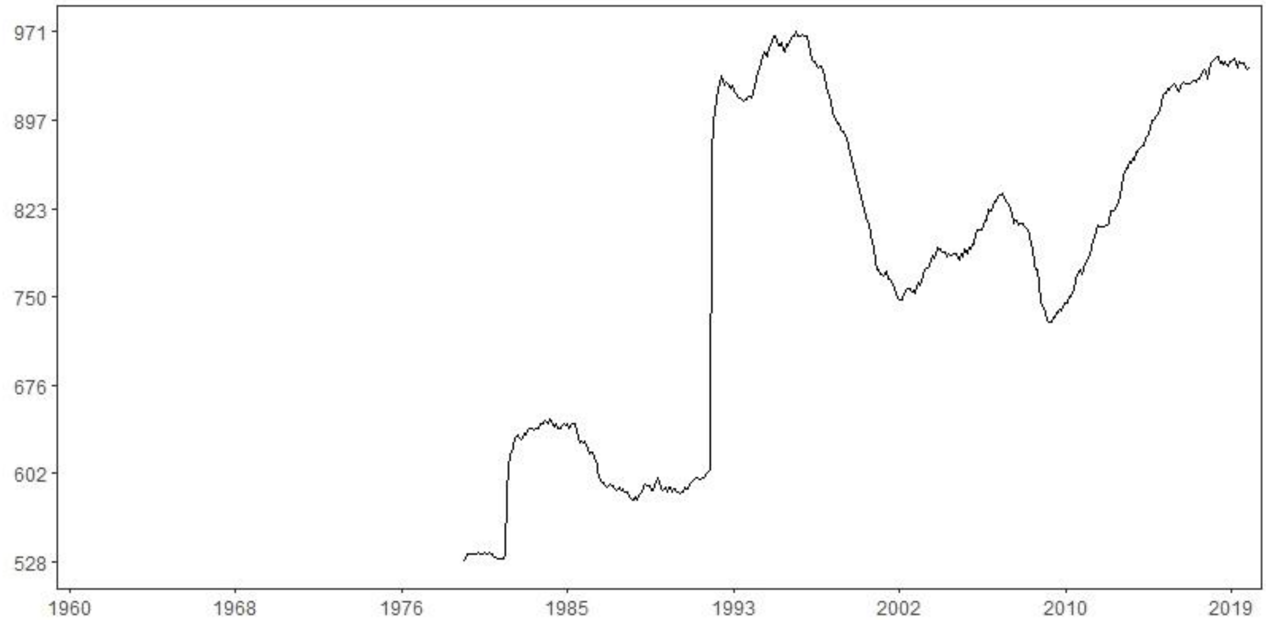


FIGURE A3

Average R^2_{adj} of the regressions of the latent factors on factors from the augmented q-factor model

This figure shows the R^2_{adj} as a percentage based on OLS estimation results for regressions of the latent factors on factors from the augmented q-factor model. The average over the out-of-sample period from January 1987 to December 2019 is given. The latent factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN).

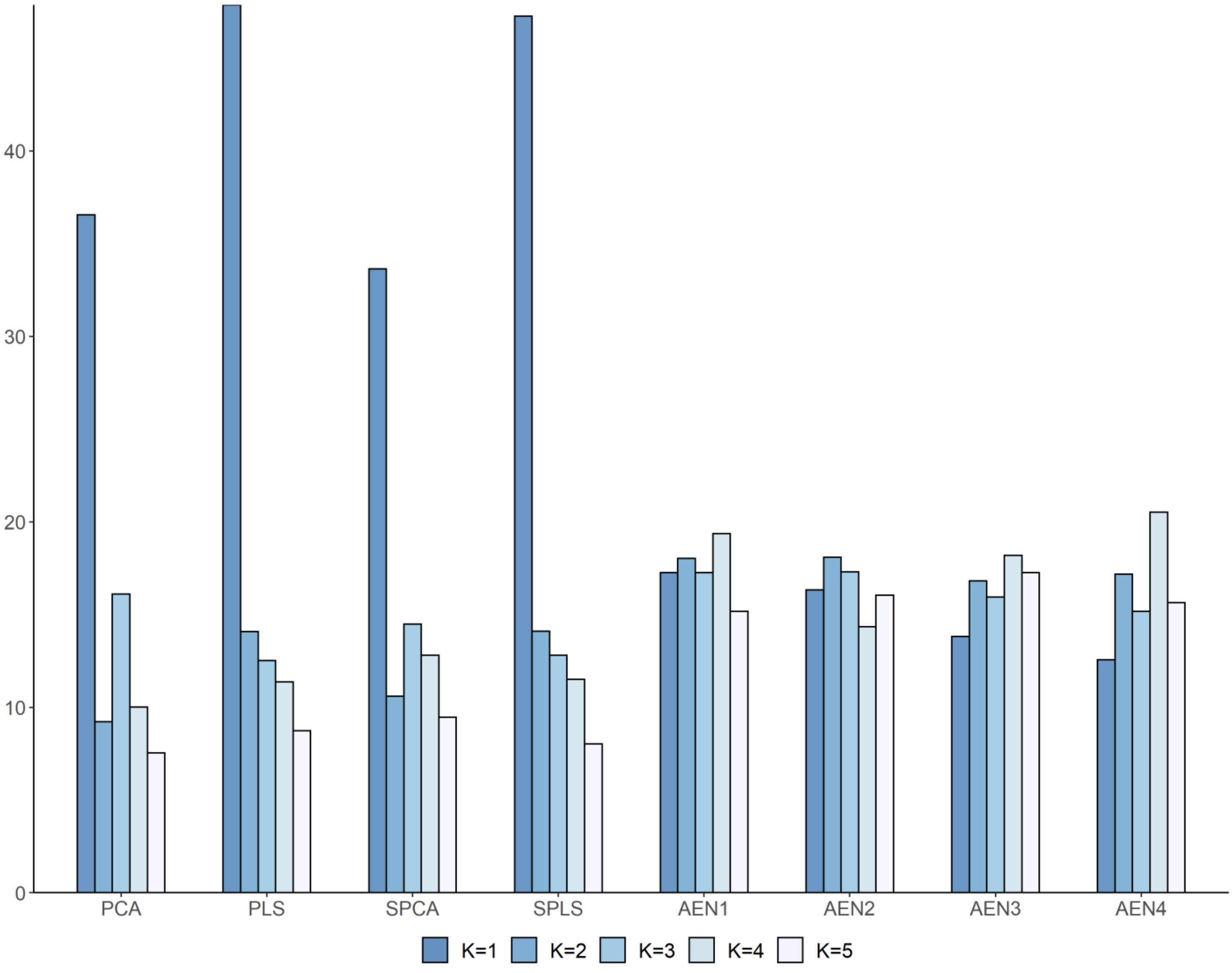


FIGURE A4

Variable importance based on the augmented q-factor model

This figure shows the variable importance based on OLS estimation results for regressions of the latent factors on factors from the augmented q-factor model. The measure of variable importance is calculated as the change in R^2 from setting the observations of a factor proxy to zero within each estimation window. The average over the out-of-sample period from January 1987 to December 2019 is given. The variable importance measures for each latent factor are scaled to sum to 100. The latent factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). MKTRF, ME, IA, ROE and EG are the Fama and French excess returns of the market from the risk-free rate, the Hou, Xue and Zhang (2015) size, investment, return-on-equity factors and the Hou, Mo, Xue and Zhang (2021) expected growth factor, respectively.

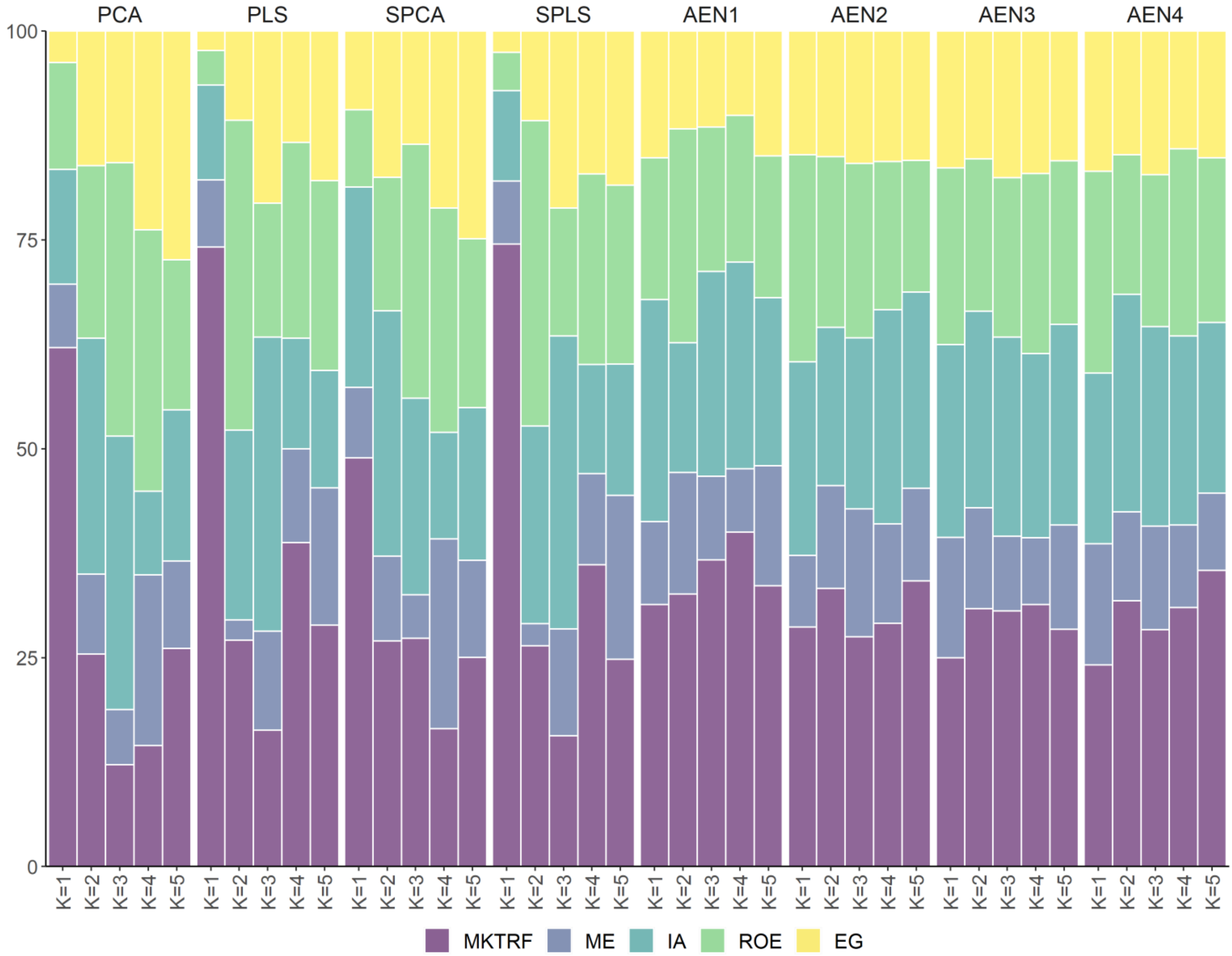


FIGURE A5

Explaining the latent factors based on the augmented q-factor model

This figure shows boxplots of the t-statistics based on OLS regressions of each of the five latent factors on factors from the augmented q-factor model. The horizontal axis reports t-statistics values ranging from -10 to 10 whereas, the vertical axis reports the latent factors, $K = 1, \ldots, 5$. The sample period is from January 1967 to December 2019. The median is marked by the line within the box, the edges of the box denote the first and third quartiles, while the minimum and maximum t-statistics are depicted by the end of the lines outside the box. The latent factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). MKTRF, ME, IA, ROE and EG are the Fama and French excess returns of the market from the risk-free rate, the Hou, Xue and Zhang (2015) size, investment, return-on-equity factors and the Hou, Mo, Xue and Zhang (2021) expected growth factor, respectively. The t-statistics are computed using heteroskedasticity and autocorrelation-robust standard errors (Newey and West, 1987). The red lines depict the Student's t critical values at the 5% level.

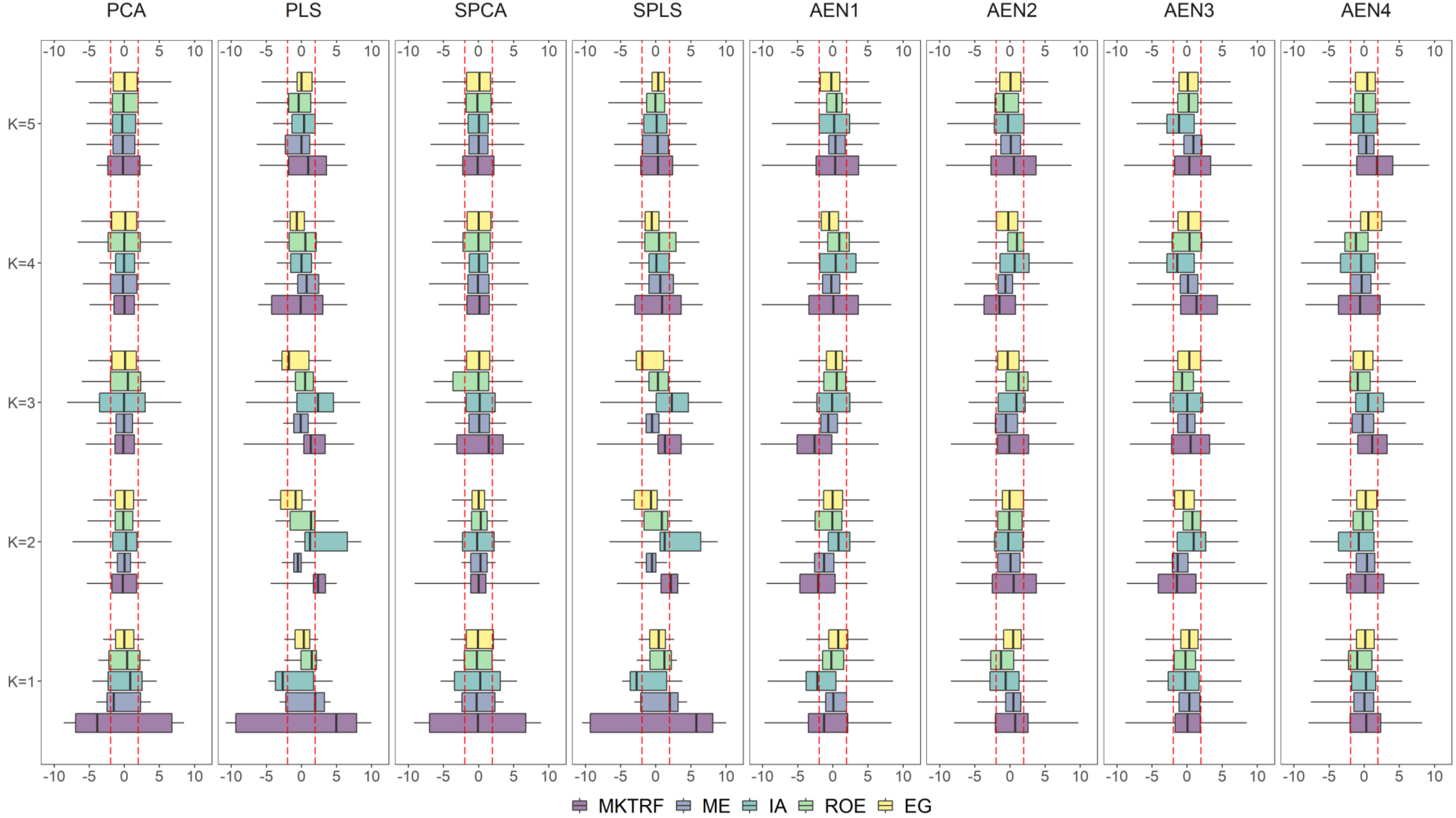


FIGURE A6

Variable importance of the economic indicators from the McCracken and Ng (2015) dataset

This figure shows the variable importance based on lasso regressions of the latent factors on economic indicators. The measure of variable importance is calculated as the change in R^2 from setting the observations of a feature to zero within each estimation window. The results are aggregated by summing the variable importance of the economic indicators belonging in the same group. Details on the variables within each group can be found in Table A2 in the Appendix. The average over the out-of-sample period from January 1980 to December 2019 is given. The variable importance measures for each group are scaled to sum to 100. The latent factor specifications are based on principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and autoencoders with 1, 2, 3 and 4 hidden layers (AEN). The explanatory variables are the 116 lagged economic indicators from the FRED-MD dataset by McCracken and Ng (2015) that have no missing values over the full sample period, from January 1960 to December 2019.

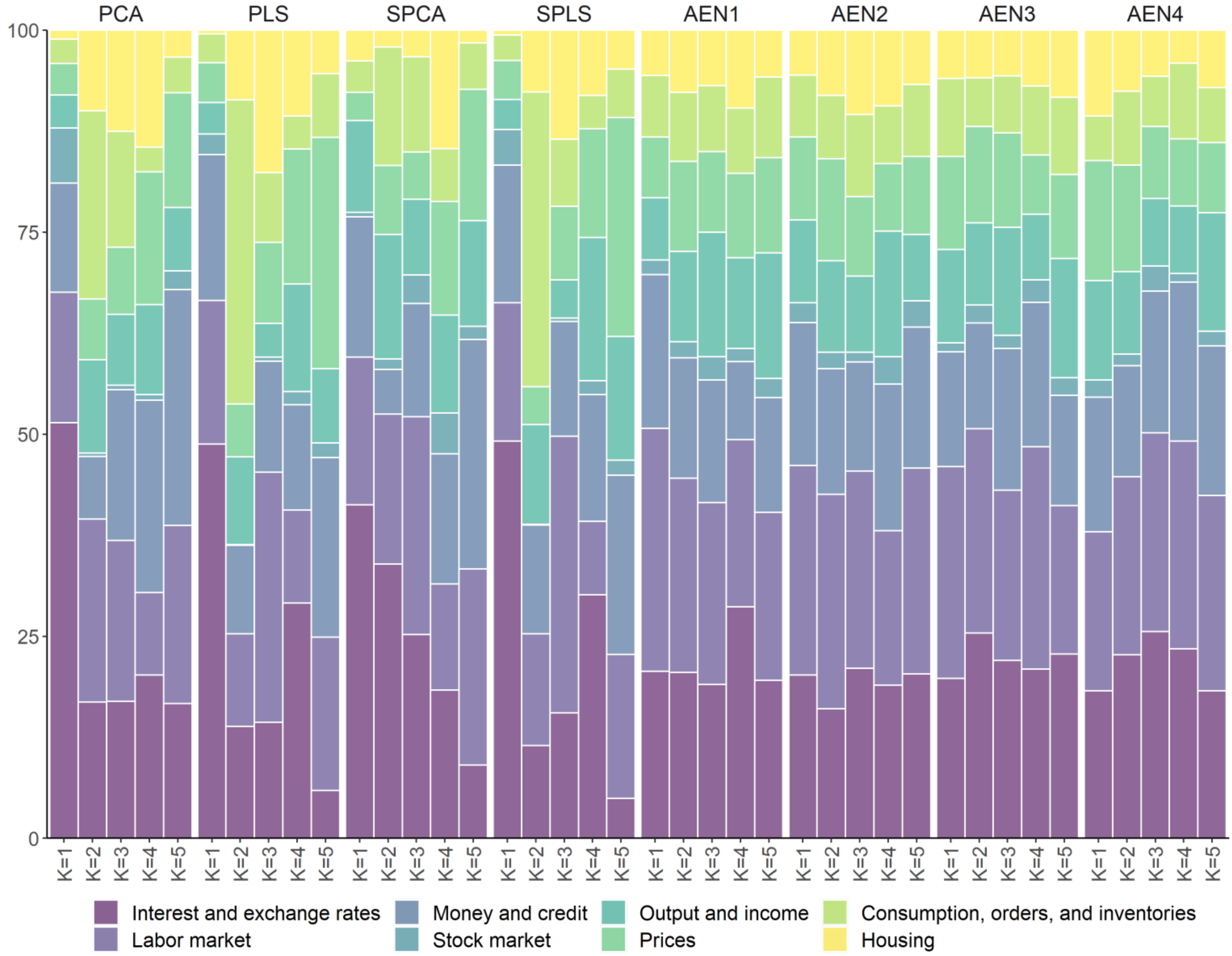


FIGURE A7

Quantiles of portfolio weight vectors: Dynamic Beta Covariance

This figure shows the quantiles of the portfolio weight vectors across the out-of-sample period, from January 1980 to December 2019. The quantiles for $\tau \in [0.1,1]$ are depicted. The results are presented for the sample estimator (Sample) and for the dynamic beta covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and an autoencoders with 1 hidden layer (AEN1).

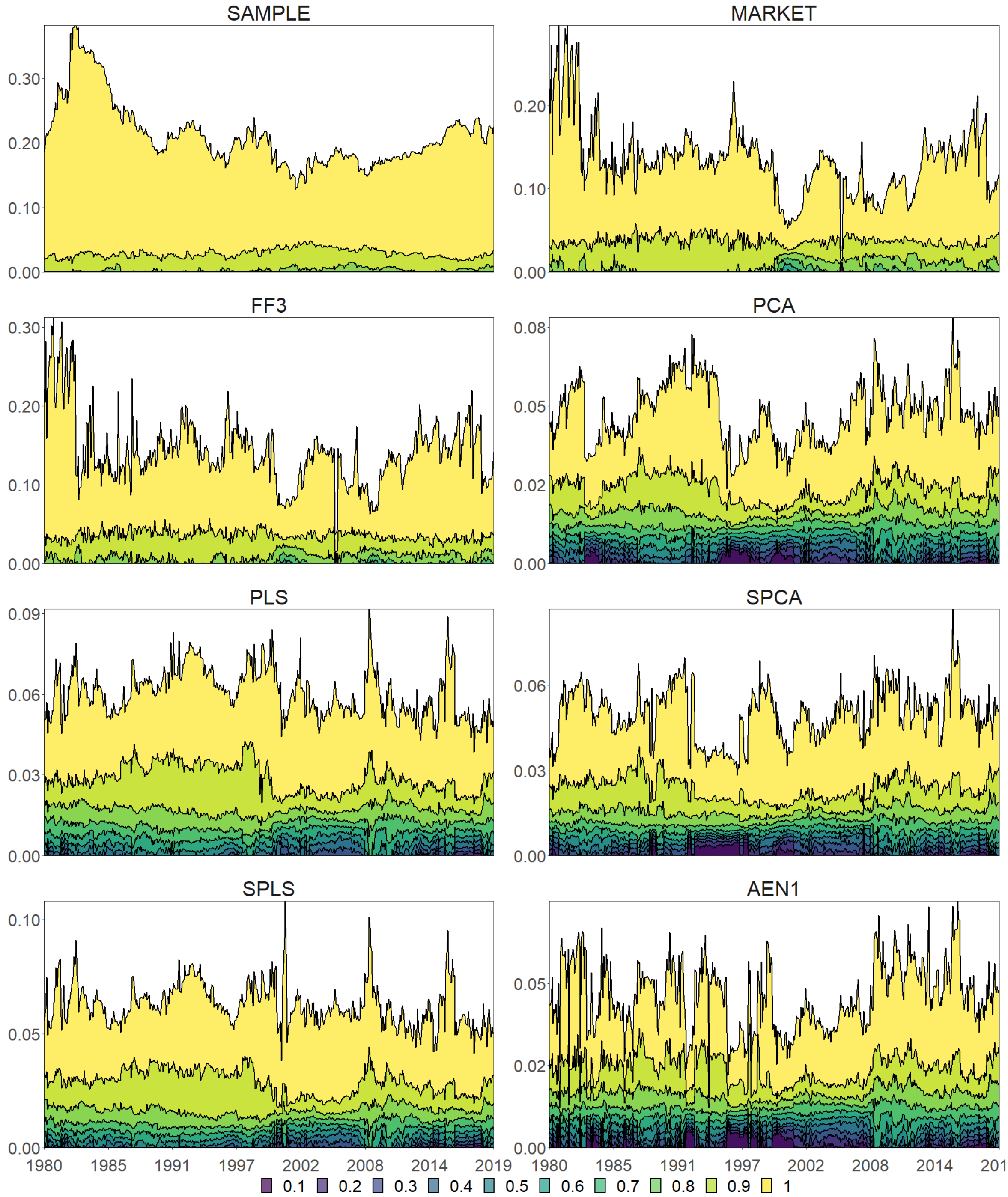


FIGURE A8

Quantiles of portfolio weight vectors: Dynamic Factor Covariance

This figure shows the quantiles of the portfolio weight vectors across the out-of-sample period, from January 1980 to December 2019. The quantiles for $\tau \in [0.1,1]$ are depicted. The results are presented for the sample estimator (Sample) and for the dynamic factor covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and an autoencoders with 1 hidden layer (AEN1).

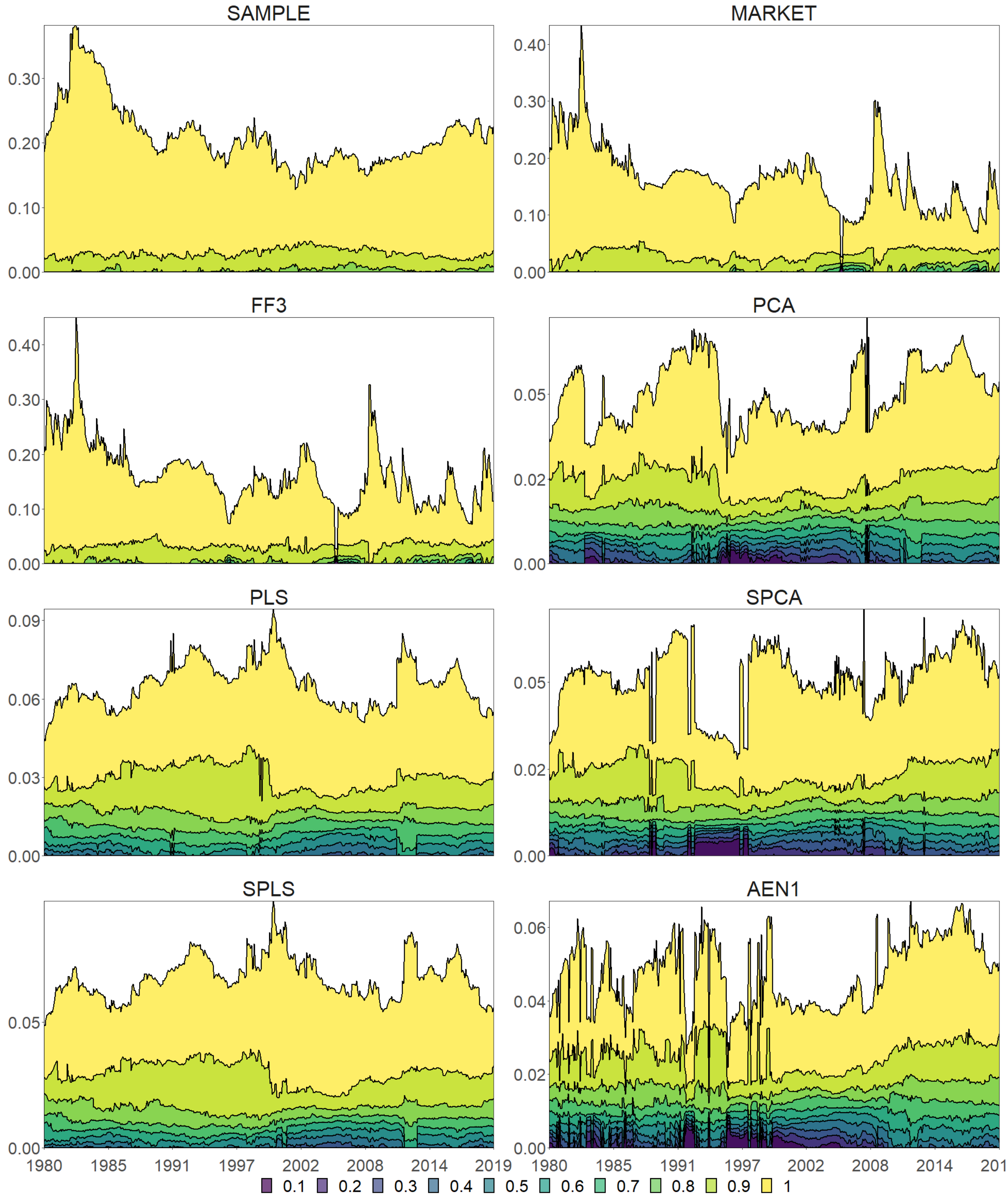


FIGURE A9

Quantiles of portfolio weight vectors: Dynamic Error Covariance

This figure shows the quantiles of the portfolio weight vectors across the out-of-sample period, from January 1980 to December 2019. The quantiles for $\tau \in [0.1,1]$ are depicted. The results are presented for the sample estimator (Sample) and for the dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and an autoencoders with 1 hidden layer (AEN1).

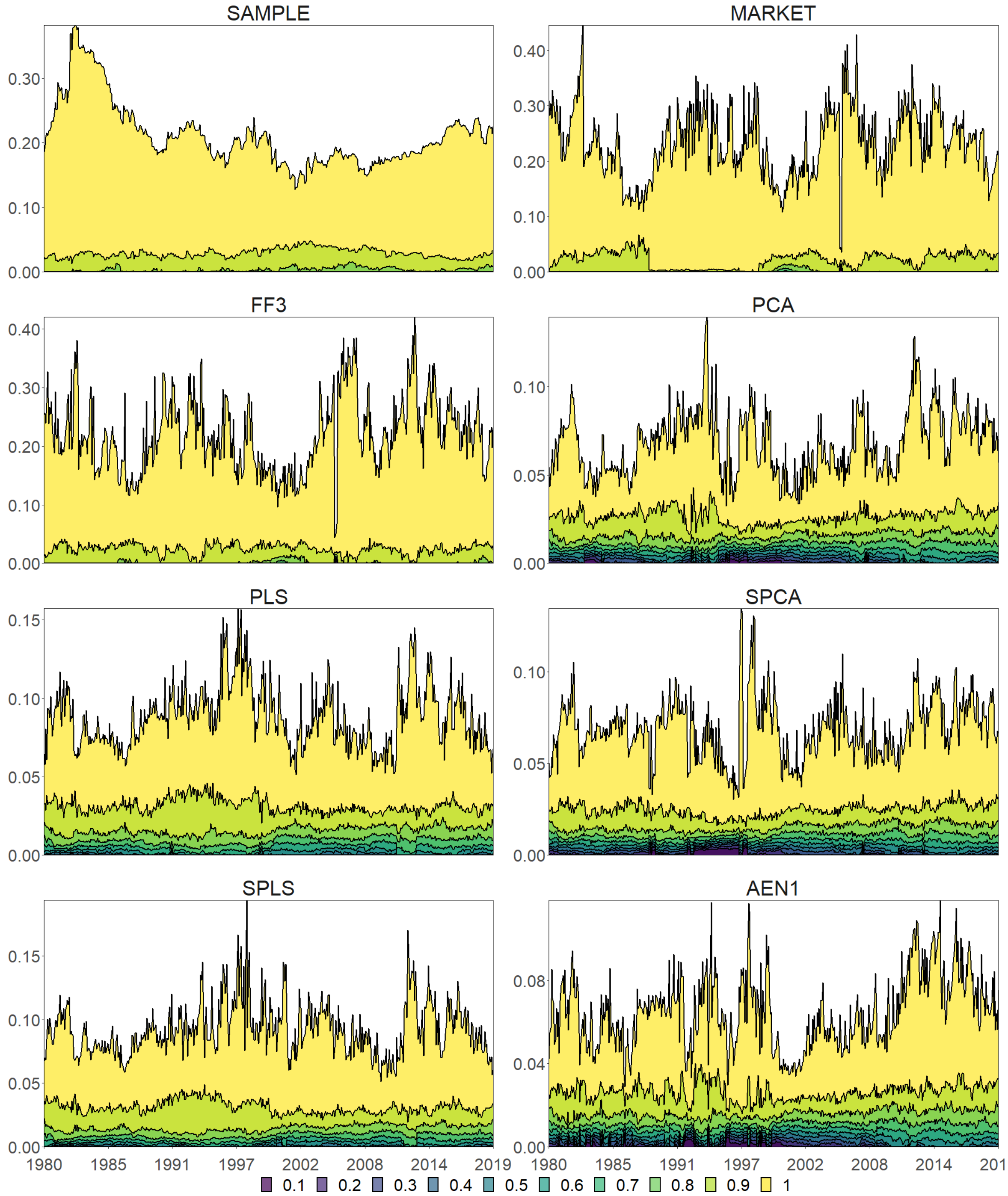


FIGURE A10

Portfolio performance for a different number of stocks: Dynamic Beta Covariance

This figure shows the monthly portfolio performance for a varying number of assets. Performance is based on the standard deviation, Sharpe ratio, average turnover and breakeven transaction costs with respect to the EW portfolio. The out-of-sample period is from January 1980 to December 2019. The results are presented for the equally weighted portfolio (EW) and for the dynamic beta covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and an autoencoder with 1 hidden layer (AEN1). The standard deviation and average turnover are reported as a percentage. The breakeven transaction costs are reported in basis points and a positive value indicates that the alternative portfolio outperforms the EW.

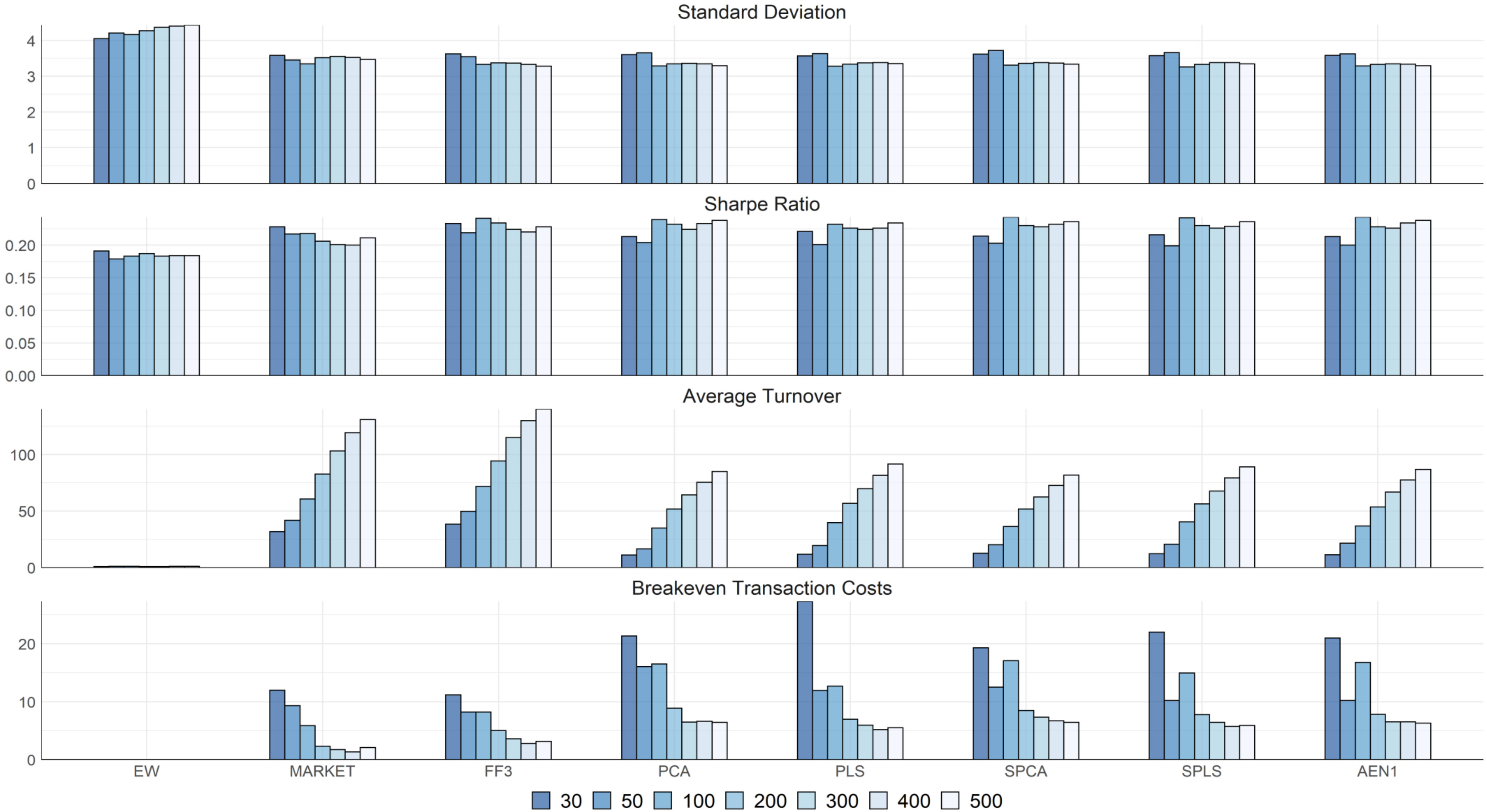


FIGURE A11

Portfolio performance for a different number of stocks: Dynamic Factor Covariance

This figure shows the monthly performance for a varying number of assets. Performance is based on the standard deviation, Sharpe ratio, average turnover and breakeven transaction costs with respect to the EW portfolio. The out-of-sample period is from January 1980 to December 2019. The results are presented for the equally weighted portfolio (EW) and for the dynamic factor covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and an autoencoder with 1 hidden layer (AEN1). The standard deviation and average turnover are reported as a percentage. The breakeven transaction costs are reported in basis points and a positive value indicates that the alternative portfolio outperforms the EW.

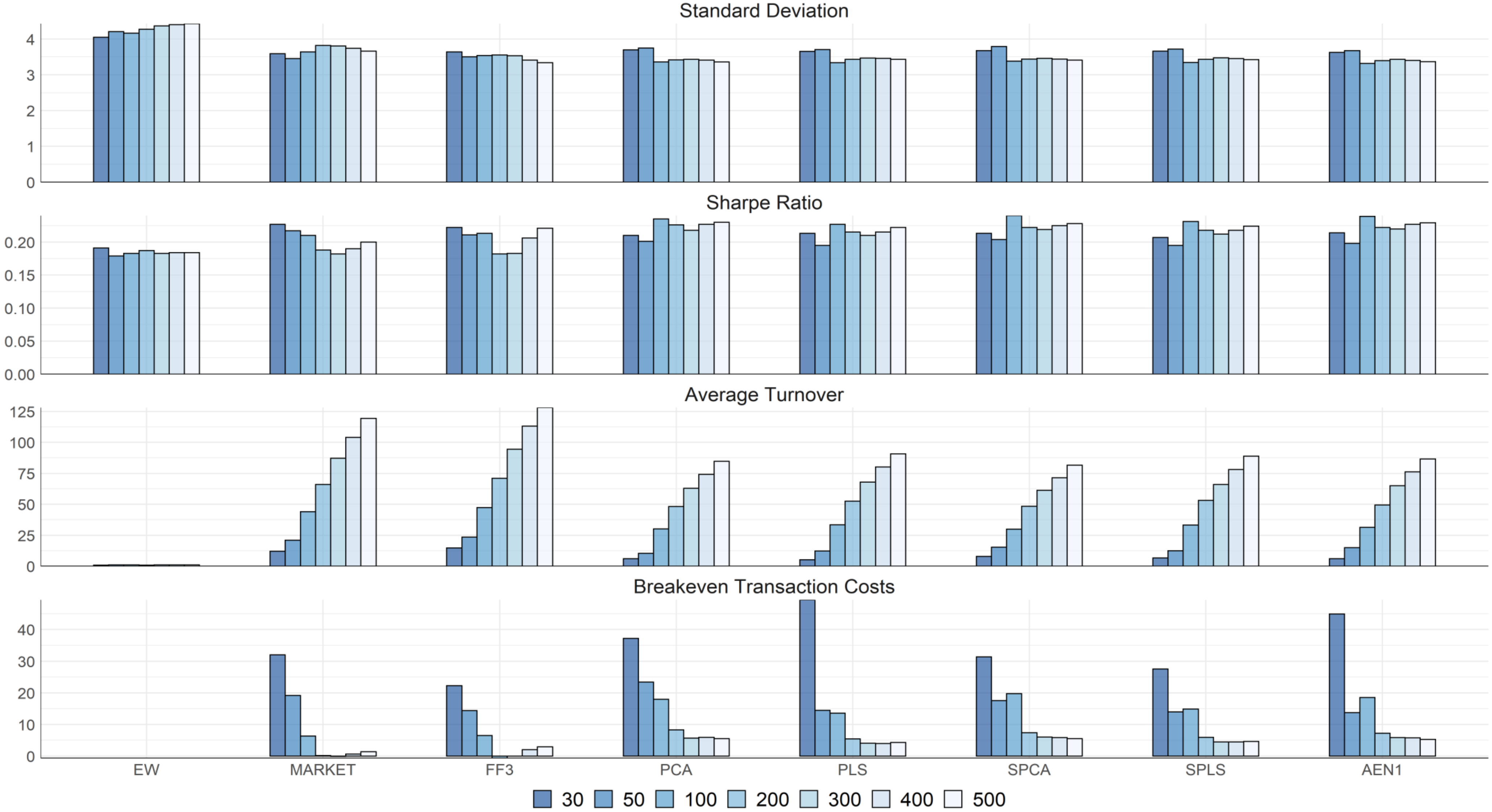


FIGURE A12

Portfolio performance for a different number of stocks: Dynamic Error Covariance

This figure shows the monthly performance for a varying number of assets. Performance is based on the standard deviation, Sharpe ratio, average turnover and breakeven transaction costs with respect to the EW portfolio. The out-of-sample period is from January 1980 to December 2019. The results are presented for the equally weighted portfolio (EW) and for the dynamic error covariance. The factor specifications are based on the single factor model (Market), the Fama-French 3-factor model (FF3), principal component analysis (PCA), partial least squares (PLS), sparse principal component analysis (SPCA), sparse partial least squares (SPLS) and an autoencoder with 1 hidden layer (AEN1). The standard deviation and average turnover are reported as a percentage. The breakeven transaction costs are reported in basis points and a positive value indicates that the alternative portfolio outperforms the EW.