Mathematics and Financial Economics (2024) 18:151–170
https://doi.org/10.1007/s11579-024-00366-y

Are minimum variance portfolios in multi-factor models long in low-beta assets?

Ansgar Steland[1]

Received: 23 October 2023 / Accepted: 11 April 2024 / Published online: 29 June 2024

Abstract
Within the one-factor capital asset pricing model (CAPM), the minimum-variance portfolio (MVP) is known to have long positions in those assets of the underlying investment universe whose betas are less than a well-defined long-short threshold beta. We study the structure of MVPs in more general multi-factor asset pricing models and clarify the low-beta puzzle for multi-factor models: For multi-factor models we derive a similar criterion in terms of the betas with explicit closed-form formulas. But the structural relationship is now more involved and the long-short threshold turns out to be asset-specific. The results rely on recursive inverse-free formulas for the precision matrix, which hold for multi-factor models and allow quick computation of that inverse matrix without the need to invert matrices going beyond diagonal ones. We illustrate our findings by analyzing S&P 500 asset returns. Our empirical results of the S&P 500 constituents between 2019 and 2022 confirm the theoretical findings and shows that the minimum variance portfolio is long in low-beta assets when applying estimates of the established asset-specific thresholds.

Keywords Asset pricing models · Factor models · Minimum-variance portfolio · PCA · Portfolio optimization · Long-short strategies

JEL Classification G11 · G12

1 Introduction

Investment strategies related to the minimum variance portfolio (MVP) have a couple of unique features interesting for investors. By definition, a minimum variance portfolio for a financial market consisting of a universe of risky assets weights the assets in such a way that the variance of the portfolio return is minimized. That portfolio is unique if the variance-covariance matrix of the d asset returns is positive definite. It coincides with the Markowitz-optimal portfolio in the case that mean returns are zero, and thus it is located on the efficient frontier in the mean-variance space. Concretely, it is located at the very left tip of the mean-variance efficient frontier of feasible portfolios. It is the only efficient portfolio which does not require to know or estimate resp. forecast mean returns. This is a beneficial feature,

✉ Ansgar Steland
steland@stochastik.rwth-aachen.de

[1] Institute of Statistics and AI Center, RWTH Aachen University, Aachen, Germany

since estimating or predicting future mean returns is highly challenging, especially for small time horizons and high dimensions, even for i.i.d. returns. It is also well known and easily observable in practice that feasible portfolios aiming at realizing high mean returns assign high weights to a small number of highly volatile assets, namely those possessing the largest predicted future returns. But such assets often tend to show local trends and bubble-like behaviour followed by local bull-like mean-reverting phases. Thus such portfolios tend to show trend- or momentum-following behaviour, such that estimates of the mean returns may be already outdated when forming the portfolio. Investing in the MVP avoids this. It is diversified, invests in low-beta assets, is agnostic to the mean returns and hedges each asset by optimally selected positions in all other assets to minimize the remaining risk, [21]. Starting with [6] empirical studies have shown that portfolios investing in low-beta assets with low idiosyncratic variances tend to outperform the market. These phenomena are called low-beta anomaly and idiosyncratic volatility or low-risk anomaly. For a recent study dealing with the Euro area and a literature review we refer to [2] and [22], respectively.

Although the MVP only depends on the covariance matrix of the asset returns and not on their mean returns, its computation is still challenging, since without further structural assumptions one needs to calculate a high-dimensional inverse covariance matrix. In general, the latter is a computationally expensive and error-prone task. However, under the capital asset pricing model (CAPM), which explains return fluctuations by the market as a single factor and idiosyncratic noise of the shares, simple closed-form expressions for the precision matrix and optimal portfolio weights are known, see [5, 6, 15]. These formulas analytically reveal that the MVP holds long positions in low-beta assets and short positions in high-beta shares, and the associated long-short threshold beta is explicitly known. For more sophisticated models the structure of the MVP has not yet been studied in detail.

This paper contributes by elaborating inverse-free formulas for the precision matrix for general multi-factor asset models, which substantially simplify the calculation of MVPs. We derive a recursion, which makes it easy to study the MVP step by step as the number of factors increases. More importantly, we clarify the low-beta anomaly for multi-factor models: For two-factor models we derive a closed-form criterion which shows that the sign of each position of the MVP can be inferred by comparing the position's beta with a threshold. However, it turns out that the threshold is now asset-specific and has a more involved form. Although the long-short thresholds may vary and there is a lack of an unique threshold applying to all assets, the novel results clarify the structural relations of the beta factors and the optimal weights, and they uniquely identify long and short positions. For general factor models with more factors a similar result can be obtained, although there are no closed-form formulas.

We demonstrate our findings by analyzing asset returns from the S&P 500. Following the general conception that daily returns of exchange-traded assets can be reasonably well explained by five factors, a five-factor asset pricing model was considered with the S&P 500 index as observable factor and four unobservable factors spanning a subspace in the orthogonal complement. In addition, two-factor models are examined with two unobservable factors and the S&P 500 as a observable factor, respectively. Unobservable factors are estimated by principal component analysis (PCA) and thus extracted from the estimated covariance matrix. The empirical results show that the returns of the estimated market portfolio (corresponding to the leading eigenvector of the estimated two-factor model covariance matrix of the asset returns) are strongly correlated to the returns of the S&P 500 index. The empirical results essentially confirm the theoretical findings. Especially, the results for the five-factor model demonstrate that the estimated minimum variance portfolio is long in low-beta assets and the estimated thresholds uniquely determine the sign of each position.

The paper is organized as follows. Section 2 recalls the general definition of the MVP, basic facts and its low-beta structure when assuming a one-factor model such as the CAPM. Section 3 reviews multi-factor asset pricing models and establishes the recursion for the computation of the precision matrix of the returns. The structure of the MVP is studied in Sect. 4 for two-factor models, where closed-form formulas can be derived, and general multi-factor models. Section 5 illustrates the findings for the S&P 500 investment universe.

2 The minimum-variance portfolio in one factor models

For the problem at hand we can and will confine ourselves to a one-period setting, but clearly the results carry over to multi-period settings. Thus, suppose that $\boldsymbol{R}_t = (R_{t1}, \ldots, R_{td})'$ is a vector of d mean zero (excess) asset returns with positive definite covariance matrix $\boldsymbol{\Sigma}$. Consider the global minimum-variance portfolio (MVP) $\boldsymbol{w}^* \in \mathbb{R}^d$ defined by the minimization problem

$$\min_{\boldsymbol{w}\in\mathbb{R}^d} \boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}, \quad \text{such that} \quad \boldsymbol{w}'\mathbf{1} = 1,$$

where $\mathbf{1}$ is a d-vector of ones. The constraint is a budget constraint which allows for long as well as short position. It is well known that there is a closed-form formula for the solution, namely

$$\boldsymbol{w}^* = \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}.$$

If $\boldsymbol{\mu}$ denotes the vector of mean returns, the portfolio $\boldsymbol{w}$ earns the mean return $\boldsymbol{w}'\boldsymbol{\mu}$. One may aim at constructing a portfolio earning a target mean return μ_0 with minimal variance. The Markowitz approach to portfolio opimization, [13], thus adds the constraint $\boldsymbol{w}'\boldsymbol{\mu} = \mu_0$ and hence considers the optimization problem

$$\min_{\boldsymbol{w}\in\mathbb{R}^d} \boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}, \quad \text{such that} \quad \boldsymbol{w}'\boldsymbol{\mu} = \mu_0, \boldsymbol{w}'\mathbf{1} = 1.$$

If $\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \cdot \mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1} - (\mathbf{1}'\boldsymbol{\Sigma}\mathbf{1})^2 \neq 0$, the optimal solution is given by the portfolio vector

$$\boldsymbol{\theta}^* = \lambda_1^*\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \lambda_2^*\boldsymbol{\Sigma}^{-1}\mathbf{1}$$

where $\boldsymbol{\lambda}^* = (\lambda_1^*, \lambda_2^*)' = \mathbf{A}^{-1}(\mu_0, 1)'$ with a 2×2 matrix $\mathbf{A}$ with diagonal elements $\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}$ and $\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ and off-diagonal element $-\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$. Any rational investor in the Markowitz sense aiming at a target mean return μ_0 holds a linear combination of the portfolio $\boldsymbol{\Sigma}^{-1}\mathbf{1}$ related to the MVP portfolio and the portfolio $\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$. We see that the inverse of the covariance matrix is essential to calculate optimal portfolios. This carries over to problems such as dollar neutral optimal long-short porfolios, see [10]. For further interpretations in terms of optimal portfolios and hedging of assets see [21]. For simplicity, however, we confine ourselves to a discussion of the MVP.

If short sales are excluded, one could invest only in the long positions, i.e. set negative entries of $\boldsymbol{w}^*$ to zero. Usually this leads to suboptimal portfolios. The associated formal optimization problem determines the optimal long-only minimum-variance portfolio (LMVP) as the solution of

$$\min_{\boldsymbol{w}\in\mathbb{R}^d} \boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}, \quad \text{such that } w_i \geq 0, 1 \leq i \leq d, \ \boldsymbol{w}'\mathbf{1} = 1,$$

which adds the long-only constraints to the optimization problem. In general, there are no explicit solutions so that one has to rely on numerical algorithms. If the covariance matrix of the underlying assets has a block structure with sufficiently small common between-asset correlation and between-block correlations, the MVP has no short positions, see [7]. But that assumption is doubtful for asset returns.

In the single factor model, a simple closed-form solution for the MVP can be derived which carries over to a semi-closed form solution for the LMVP, see [5] and [15]. Recall that the single factor model with the market portfolio as factor, i.e., the CAPM, assumes that the time t excess return, R_{ti}, is given by

$$R_{ti} = \beta_i (R_{Mt} - r) + \epsilon_{ti}, \qquad 1 \leq i \leq d, 1 \leq t \leq n,$$

where r is the risk-free rate of return, $\beta_1, \ldots, \beta_d$ are the beta factors and ϵ_{ti} the mean zero idiosyncratic errors with variances σ_i^2, uncorrelated across i, and n is the size of the available sample. R_{Mt} stands for the time t return of the market portfolio. Here, in practice, one often uses an index serving as a proxy of the market. For example, the S&P 500 for the US market or, more generally, a capitalization-weighted average of the investable universe under consideration. $\sigma_M^2 = \text{Var}(R_{Mt} - r)$ will denote the market's variance. In the following it is assumed that all beta factors are nonnegative. Put $\boldsymbol{b} = (\beta_1, \ldots, \beta_d)'$. Then the covariance matrix takes the form

$$\boldsymbol{\Sigma} = \sigma_M^2 \boldsymbol{b}\boldsymbol{b}' + \mathbf{S}, \qquad \mathbf{S} = \text{diag}(\sigma_1^2, \ldots, \sigma_d^2). \tag{1}$$

If all idiosyncratic risks, σ_i^2, are positive, which we assume throughout the paper, $\mathbf{S}$ is invertible and then the inverse covariance matrix is given by

$$\boldsymbol{\Sigma}^{-1} = \mathbf{S}^{-1} - \frac{\boldsymbol{b}_r \boldsymbol{b}_r'}{\sigma_M^{-2} + \boldsymbol{b}_r' \boldsymbol{b}}.$$

Here $\boldsymbol{b}_r = (\beta_1/\sigma_1^2, \ldots, \beta_d/\sigma_d^2)'$ is the vector of risk-adjusted beta factors and $\mathbf{S}^{-1} = \text{diag}(\sigma_1^{-2}, \ldots, \sigma_d^{-2})$. Thus, in a single-factor model the precision matrix $\boldsymbol{\Sigma}^{-1}$ is easy to obtain and there is no need to compute an inverse matrix going beyond the simple case of inverting a diagonal matrix. As a result, one gets the explicit closed-form optimal solution for the MVP weights

$$w_i^* = \frac{\sigma_{MVP}^2}{\sigma_i^2} \left(1 - \frac{\min(\beta_i, \beta_{LS})}{\beta_{LS}} \right), \qquad 1 \leq i \leq d. \tag{2}$$

Here $\beta_{LS} = \frac{\sigma_{MVP}^{-2} + \sum_{i=1}^d \beta_i/\sigma_i^2}{\sum_{i=1}^d \beta_i/\sigma_i^2}$ is the *long-short threshold beta*. One buys only assets with betas smaller than this threshold. Assets with beta factors exceeding β_{LS} are shorted. The long-short threshold β_{LS} depends on the idiosyncratic variances but also on all beta factors, so that changing a single β_i changes the threshold. Therefore, the characterization of the signs of all position in terms of the beta factors is a *structural* property of the MVP in terms of the beta factors. It is worth mentioning that it has been conjectured in [6] and shown in [15] that the long-only minimum-variance portfolio attains the same formula with β_{LS} replaced by the long-only threshold beta given by the smallest solution of the equation

$$\beta_{LO} = \frac{\sigma_{MVP}^{-2} + \sum_{\beta_i < \beta_{LO}} \beta_i/\sigma_i^2}{\sum_{\beta_i < \beta_{LO}} \beta_i/\sigma_i^2}.$$

Therefore, the LMVP weights are given $w^*_{L,i} = \frac{1}{\sigma_i^2}\left(1 - \frac{\min(\beta_i, \beta_{LO})}{\beta_{LO}}\right)$. The optimal long-only portfolio has positions in all investable assets with beta factors not exceeding β_{LO}, i.e. in low-beta assets.

In [1] an alternative criterion has been elaborated for the case $d = 2$, which is based on the general representation $\mathbf{\Sigma} = \mathbf{P}^\delta e^{\mathbf{\Lambda}} e^{-\mathbf{\Gamma}} \mathbf{P}^\delta$, where $\mathbf{\Lambda}$ is the diagonal matrix of the positive eigenvalues $\lambda_1, \ldots, \lambda_d$, $\mathbf{P}$ is the block-diagonal matrix with upper left block $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and lower right block $\mathbf{I}_{d-2}$, $\delta = 0$ or $\delta = 1$, and $\mathbf{\Gamma}$ is a skew-symmetric matrix, i.e. $\mathbf{\Gamma}' = -\mathbf{\Gamma}$. Here, $\mathbf{I}_k$ stands for the unit matrix of dimension k. In this representation the entries p_{ij}, $1 \le j < i$, $1 \le i \le d$, parameterize $\mathbf{\Sigma}$. In dimension $d = 2$ the MVP weights are then given by

$$w_1 = \frac{\lambda_1 + (\lambda_2 - \lambda_1)(\cos v + \sin v)\cos v}{\lambda_1 + \lambda_2 + 2(\lambda_2 - \lambda_1)\sin v \cos v}, \qquad w_2 = 1 - w_1,$$

and the numerator determines the sign of the position. However, the results of this paper hold for arbitrary d.

3 Multi-factor models

To set the stage for our results, let us now consider and review multi-factor models for asset returns. For further reading we refer to [3, 4, 8] and the references given therein. Our discussion starts with the case of observed factors and then proceeds to the case that a some or all of the factors are unobserved.

Due to its specific role, we assume that one factor is the market portfolio which is augmented by K additional factors assumed to be uncorrelated among each other and with the market. Let us first consider the case of observable factors, i.e., in addition to the market returns R_{Mt} we are given K time series F_{kt}, $1 \le t \le n$, for $1 \le k \le K$. Then the time t excess return, R_{ti}, of asset i is explained by the model

$$R_{ti} = \beta_i(R_{Mt} - r) + \sum_{k=1}^{K} l_i^{(k)} F_{tk} + \epsilon_{ti},$$

where the first component is as above, F_{tk} is the time t observation of factor k, assumed to be uncorrelated with the market and across k, with variance σ_{Fk}^2. The unknown coefficient $l_i^{(k)}$ is the factor-beta (or loading) of asset i with respect to factor k, thus measuring the influence of the kth factor on the mean excess return of asset i. The errors ϵ_{ti} are again assumed to be mean zero random variables with variances σ_i^2, independent of the market and factor returns. Since usually the market plays a specific role, we do not subsume it under the factors. For a data set of size n, the model can be compactly written in matrix notation

$$\underbrace{\mathbf{R}}_{n\times d} = \underbrace{\mathbf{F}}_{n\times K}\underbrace{\mathbf{L}'}_{K\times d} + \underbrace{\boldsymbol{e}}_{n\times d} \tag{3}$$

where $\mathbf{R}$ is the data matrix of asset excess returns, matrix $\boldsymbol{F} = (F_{tk})_{\substack{1\le t\le n\\1\le k\le K}}$ collects the factors as columns with first column given by $(F_{t0})_{t=1}^n = (R_{Mt} - r)_{t=1}^n$. $\mathbf{L} = (l_i^{(k)})_{\substack{1\le i\le d\\1\le k\le K}}$ is the $d \times K$ loadings matrix with entries $l_i^{(k)}$ and first column given by $\boldsymbol{b}$, and $\boldsymbol{e}$ is the matrix of error terms. We omit intercept terms, α_i, since we are not interested in analyzing pricing errors. However, in our empirical analysis we included intercept terms in the CAPM regressions to estimate the betas.

The covariance matrix $\mathbf{\Sigma}_K$ of the d risky assets in the presence of K additional factors attains the form

$$\mathbf{\Sigma}_K = \sigma_M^2 \boldsymbol{b}\boldsymbol{b}' + \sum_{k=1}^{K} \sigma_{Fk}^2 \boldsymbol{l}_k \boldsymbol{l}_k' + \mathbf{S}_0, \tag{4}$$

where $\boldsymbol{l}_k = (l_1^{(k)}, \dots, l_d^{(k)})'$ is the kth column of $\mathbf{L}$ and $\mathbf{S}_0 = \mathrm{diag}(\sigma_1^2, \dots, \sigma_d^2)$. Consequently, the covariance matrix $\mathbf{\Sigma}_K$ can be calculated from the model parameters $\beta_1, \dots, \beta_d, l_1^{(k)}, \dots, l_d^{(k)}, 1 \le k \le K$, and $\sigma_1^2, \dots, \sigma_d^2$.

If some of the factors are observable, say, the first $q \le K$, the model takes the general form

$$R_{ti} = \sum_{j=1}^{q} l_{obs,i}^{(j)} X_{tj} + \sum_{k=1}^{K-q} l_i^{(k)} F_{tk} + \epsilon_{ti}, \qquad 1 \le i \le d, 1 \le t \le n,$$

where X_{tj}, $1 \le t \le n$, $1 \le j \le q$, are the observed factor series. In matrix notation we have the compact representation

$$\underbrace{\mathbf{R}}_{n\times d} = \underbrace{\mathbf{X}}_{n\times q} \underbrace{\mathbf{L}_{obs}'}_{q\times d} + \underbrace{\mathbf{F}}_{n\times(K-q)} \underbrace{\mathbf{L}'}_{(K-q)\times d} + \underbrace{\boldsymbol{e}}_{n\times d}, \tag{5}$$

where $\mathbf{X} = (X_{ti})$ is the $n \times q$ data matrix of the observable factors, $\mathbf{L}_{obs}$ their loadings (regression coefficients), and $\mathbf{F}$ and $\mathbf{L}$ have the same meaning as above. The full structure is given by $\mathbf{L}^* = (\mathbf{L}_{obs}, \mathbf{L})$ and $\mathbf{F}^* = (\mathbf{X}, \mathbf{F})$. Clearly, if $q = K$ we are given a multivariate regression model. Otherwise, the returns are explained by q external factors, which can be observed, and additional $K - q$ factors which explain the structure of $\mathbf{R} - \mathbf{X}\mathbf{L}_{obs}'$, the returns corrected by the influence of the observables. Such extended models are of particular interest, if one beliefs that a market index such as the S&P 500 represents a good proxy for the market portfolio or if one wants to include other factors in the asset pricing model, such as size, profitability, momentum or book-to-market ratio. The latter factors are not returns.

3.1 Estimation by PCA

If the factors are not observable, then both the factor matrix $\mathbf{F}$ and the loadings matrix $\mathbf{L}$ need to be estimated. The above formula (4) for $\mathbf{\Sigma}_K$ is the starting point, as it links the columns of $\mathbf{L}$ to the spectral representation of $\mathbf{\Sigma}_K$. A common approach to estimate $\boldsymbol{b}$ and $\boldsymbol{l}_k$, $1 \le k \le K$, is to apply PCA. One calculates all eigenvalues and eigenvectors of the sample covariance matrix (or of a more sophisticated estimator of the covariance matrix) $\hat{\mathbf{\Sigma}}_n$. This results in ordered eigenvalue-eigenvector pairs $(\hat{\lambda}_0, \hat{\boldsymbol{u}}_0), \dots, (\hat{\lambda}_{d-1}, \hat{\boldsymbol{u}}_{d-1})$ where $\hat{\lambda}_0 \ge \hat{\lambda}_1 > \cdots > \hat{\lambda}_{d-1}$. Now one estimates $\boldsymbol{b}$ by the leading eigenvector $\hat{\boldsymbol{u}}_0$ and $\boldsymbol{l}_k$ by $\hat{\boldsymbol{u}}_k$ for $1 \le k \le K \le d - 1$. This means, the loadings matrix is estimated by the matrix $\hat{\mathbf{L}}$ consisting of the first K eigenvectors. Further, $\hat{\sigma}_M^2 = \hat{\lambda}_0$ and $\hat{\sigma}_{Fk}^2 = \hat{\lambda}_k$, $1 \le k \le K$. The idiosyncratic variances $\sigma_1^2, \dots, \sigma_d^2$ are estimated as the diagonal elements of $\hat{\mathbf{\Sigma}}_n - \sum_{k=1}^{K} \hat{\lambda}_k \hat{\boldsymbol{u}}_k \hat{\boldsymbol{u}}_k'$. Lastly, the factors (also called principal components) can then be obtained by regressing the returns on the estimated loadings which gives $\hat{\mathbf{F}} = \mathbf{R}\hat{\mathbf{L}}$. The first column of $\hat{\mathbf{F}}$, i.e. $\hat{\boldsymbol{r}}_M = \mathbf{R}\hat{\boldsymbol{u}}_0$, gives the estimated returns of the market portfolio.

In order to make the estimated market returns comparable to a market proxy such as the S&P 500, one can rescale the estimated market index appropriately and thus calculate

$$\hat{\boldsymbol{r}}_M^* = \frac{\mathrm{sd}_{SP500}}{\mathrm{sd}(\hat{\boldsymbol{r}}_M)} \hat{\boldsymbol{r}}_M,$$

where sd_{SP500} and $\mathrm{sd}(\hat{\boldsymbol{r}}_M)$ are the standard deviations of the S&P 500 returns and $\hat{\boldsymbol{r}}_M$, respectively. In the factor model equation this rescaling is then compensated by introducing the suitably rescaled betas

$$\hat{\beta}_i^* = \frac{\mathrm{sd}(\hat{\boldsymbol{r}}_M)\beta_i}{\mathrm{sd}_{SP500}}.$$

This scaling makes the estimated betas comparable to the given market proxy. Based on $\hat{\boldsymbol{r}}_M^*$ we may define the corresponding price process $\hat{\boldsymbol{p}}_M^*$ by $\hat{p}_i = P_{SP500,0} \prod_{t=1}^{i} (1 + r_{M,i}^*)$, $1 \le i \le n$, where $P_{SP500,0}$ stands for the inital quote of the market proxy (S&P 500).

Behind this rescaling is the fact that factors and loadings are not uniquely determined: For any full-rank $K \times K$ matrix $\mathbf{A}$ we have $(\mathbf{FA})(\mathbf{A}^{-1}\mathbf{L}') = \mathbf{FL}'$. This indeterminacy can be eliminated by imposing K^2 restrictions. PCA imposes the normalization $\mathbf{F}'\mathbf{F} = \mathbf{I}$, $\mathbf{L}'\mathbf{L}$ diagonal with distinct entries, but any transformation by a full-rank matrix $\mathbf{A}$ yields a valid model for the unobserved component as well.

In view of the wide acceptance of capitalization-weighted indices (such as the S&P 500) as proxies for the market portolio, and the importance of observable factors (such as the Fama and French factors), the question arises how to determine the $d - q$ unobserved factors in the presence of q observed factors with associated $n \times q$ data matrix $\boldsymbol{X}$. Clearly, in case of a market proxy, $q = 1$ and $\mathbf{X}$ is the $n \times 1$ matrix of the excess returns of the index. This can be achieved by the following two-step approach, [17]. In the first step one runs a least-squares regression of the asset excess returns, $\mathbf{R}$, on $\boldsymbol{X}$. This gives regression coefficients, $[\hat{\mathbf{L}}_{obs}]^i = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'[\mathbf{R}]^i$, $1 \le i \le d$, where $[\mathbf{A}]^i$ is the ith column of a matrix $\mathbf{A}$. In case of a market index, $\hat{\beta}_i = [\hat{\mathbf{L}_{obs}}]^i$ is a scalar, the asset's beta, and then one puts $\hat{\boldsymbol{b}} = (\hat{\beta}_1, \ldots, \hat{\beta}_d)'$. Next, one calculates the $n \times q$ matrix $\hat{\mathbf{E}}$ of regression residuals. In the second step, one determines $K - q$ factors in the orthogonal complement $\mathrm{rg}(\boldsymbol{X})^{\perp}$ of the column space of $\boldsymbol{X}$ by applying the above PCA procedure with $\mathbf{R}$ replaced by $\hat{\mathbf{E}}$ yielding a $d \times (K - q)$ matrix $\hat{\mathbf{L}}$ of estimated eigenvectors, associated principal components (factors) $\hat{\mathbf{F}} = \hat{\mathbf{E}}\hat{\mathbf{L}}$, and estimates of the idiosyncratic variances $\hat{\sigma}_1^2, \ldots, \hat{\sigma}_d^2$. One may puts things together by letting $\hat{\mathbf{L}}^* = (\mathbf{L}_{obs}, \mathbf{L}^{\perp})$ and $\hat{\mathbf{F}}^* = (\mathbf{X}, \hat{\mathbf{F}})$.

Classical PCA as reviewed above extracts factors and loadings from the sample covariance matrix. Here, one may also use more sophisticated estimators of the asset returns' variance-covariance matrix $\boldsymbol{\Sigma}$. Especially for high-dimensional settings, shrinkage estimators may be used. For example, Ledoit and Wolf [11] proposed shrinkage estimators shrinking the sample covariance, which is a nonparametric estimator, towards a parametric target such as a multiple of the identity matrix. Sancetta [16] studies consistency in high dimension for dependent time series, Steland and von Sachs [19, 20] provide distributional approximations. For a recent more general approach where one shrinks towards a linear combination of typical nonparametric targets, namely Toeplitz and banded covariance matrices, respectively, we refer to [14]. Nonlinear shrinkage has been proposed by Ledoit and Wolf [12].

3.2 Inverse-free computations of inverses and recursions

Let us return to the factor model. If one ignores the idiosyncratic noise and takes all components, i.e. $K = d - 1$, the covariance structure has a pure factor structure and the associated precision matrix, needed for model inference and calculation of optimal portfolios, can be easily computed by $\mathbf{\Sigma}_K^{-1} = \frac{1}{\sigma_M^2}\boldsymbol{b}\boldsymbol{b}' + \sum_{k=1}^K \frac{1}{\sigma_{Fk}^2}\boldsymbol{l}_k\boldsymbol{l}_k'$ assuming that $\boldsymbol{b}, \boldsymbol{l}_1, \ldots, \boldsymbol{l}_K$ are d unit vectors, and similarly for the estimated inverse matrix. But the realistic assumption of the presence of idiosyncratic terms ϵ_{ti} representing asset-specific fluctuations of the returns invalidates this simple formula for the inverse. However, the following result shows that $\mathbf{\Sigma}_K^{-1}$ can be calculated for any value of K without the need to compute inverse matrices going beyond the inverse of a diagonal matrix. We shall name such a formula *inverse-free*. Further, the sequence of precision matrices $\mathbf{\Sigma}_k^{-1}$, $0 \le k \le K$, can be calculated recursively.

Theorem 1 *In the multi-factor asset return model it holds*

$$\mathbf{\Sigma}_K^{-1} = (\mathbf{I} - \mathbf{\Sigma}_0^{-1}(\mathbf{I} + \mathbf{B}_K\mathbf{\Sigma}_0^{-1})^{-1}\mathbf{B}_K)\mathbf{\Sigma}_0^{-1},$$

where $\mathbf{\Sigma}_0^{-1} = \mathbf{S}_0^{-1} - \frac{\boldsymbol{b}_r\boldsymbol{b}_r'}{1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b}}$, $\mathbf{B}_K = \sum_{k=1}^K \sigma_{Fk}^2\boldsymbol{l}_k\boldsymbol{l}_k'$ *is the contribution of the K factors to the covariance matrix, and* $\boldsymbol{b}_r = (\beta_1/\sigma_1^2, \ldots, \beta_d/\sigma_d^2)'$ *is the vector of the beta factors risk-adjusted by the idiosyncratic variances from the K-factor regression. Especially,*

$$\mathbf{\Sigma}_1^{-1} = \left(\mathbf{I} - \mathbf{\Sigma}_0^{-1}\left[\mathbf{I} - \frac{\sigma_{F1}^2\boldsymbol{l}_1\boldsymbol{g}_1'}{1 + \sigma_{F1}^2\boldsymbol{l}_1'\boldsymbol{g}_1}\right]\sigma_{F1}^2\boldsymbol{l}_1\boldsymbol{l}_1'\right)\mathbf{\Sigma}_0^{-1}$$

with $\boldsymbol{g}_1 = (\mathbf{S}_0^{-1}\boldsymbol{l}_1) - \frac{\sigma_{F1}^2\boldsymbol{l}_1'\boldsymbol{b}_r}{1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b}}\boldsymbol{b}_r$. *Further, for any* $k \ge 2$ *we have the recursive formula*

$$\mathbf{\Sigma}_k^{-1} = \left(\mathbf{I} - \mathbf{\Sigma}_{k-1}^{-1}\left[\mathbf{I} - \frac{\sigma_{Fk}^2}{1 + \sigma_{Fk}^2\boldsymbol{l}_k'\boldsymbol{g}_k}\boldsymbol{l}_k\boldsymbol{g}_k'\right]\sigma_{Fk}^2\boldsymbol{l}_k\boldsymbol{l}_k'\right)\mathbf{\Sigma}_{k-1}^{-1}, \qquad k \ge 1,$$

with $\boldsymbol{g}_k = \mathbf{\Sigma}_{k-1}^{-1}\boldsymbol{l}_k$, *where* $\mathbf{\Sigma}_{k-1}^{-1}$ *is the inverse matrix using factors* $\boldsymbol{l}_1, \ldots, \boldsymbol{l}_{k-1}$ *and the idiosyncratic variances* $\sigma_1^2, \ldots, \sigma_d^2$ *from the K-factor regression model.*

When using the recursion to compute the inverse covariance matrices, it is important to note that in all steps one needs to use the idiosyncratic variances from the full factor model. Therefore, the calculations start with computing $\mathbf{\Sigma}_0^{-1}$ using the matrix $\mathbf{S}_0 = \operatorname{diag}(\sigma_1^2, \ldots, \sigma_d^2)$ of the idiosyncratic variances of the factor model employing all K factors. Then one computes for $k = 1, \ldots, K$

$$\mathbf{\Sigma}_k^{-1} = \left(\mathbf{I} - \mathbf{\Sigma}_{k-1}^{-1}\left[\mathbf{I} - \frac{\sigma_{Fk}^2}{1 + \sigma_{Fk}^2\boldsymbol{l}_k'\boldsymbol{g}_k}\boldsymbol{l}_k\boldsymbol{g}_k'\right]\sigma_{Fk}^2\boldsymbol{l}_k\boldsymbol{l}_k'\right)\mathbf{\Sigma}_{k-1}^{-1}.$$

The recursive formula for the inverse covariance matrix, which is needed to calculate optimal porfolios, allows to calculate these portfolios step-by-step starting with the CAPM (i.e. no additional factor) and then adding factors successively. In each step one determines the factor betas $\boldsymbol{l}_k$, the factor variances σ_{Fk}^2 and the new idiosyncratic risks σ_i^2. Then one calculates $\boldsymbol{g}_k$ and uses the inverse-free update formula to calculate $\mathbf{\Sigma}_k^{-1}$.

Lastly, note that all these formulas hold for the estimated precision matrix associated to the estimators $\hat{\mathbf{\Sigma}}_K = \hat{\sigma}_M^2\hat{\boldsymbol{b}}\hat{\boldsymbol{b}}' + \sum_{k=1}^K \hat{\sigma}_{Fk}^2\hat{\boldsymbol{l}}_k\hat{\boldsymbol{l}}_k' + \hat{\mathbf{\Sigma}}_0$, where $\hat{\mathbf{\Sigma}}_0 = \operatorname{diag}(\hat{\sigma}_1^2, \ldots, \hat{\sigma}_d^2)$, using arbitrary estimators $\hat{\boldsymbol{b}}, \hat{\boldsymbol{l}}_1, \ldots, \hat{\boldsymbol{l}}_K, \hat{\sigma}_M^2, \hat{\sigma}_{F1}^2, \ldots, \hat{\sigma}_{FK}^2, \hat{\sigma}_1^2, \ldots, \hat{\sigma}_d^2$.

4 Structure of the minimum variance portfolio

Recall from above that the formula for the global minimum variance portfolio is given by $\boldsymbol{w}^* = \sigma^2_{MVP,K}\boldsymbol{\Sigma}_K^{-1}\mathbf{1}$ where $\sigma^2_{MVP,K} = 1/\mathbf{1}'\boldsymbol{\Sigma}_K^{-1}\mathbf{1}$. Denote by $\boldsymbol{w}^*_{1F} = \boldsymbol{\Sigma}_0^{-1}\mathbf{1}/\mathbf{1}'\boldsymbol{\Sigma}_0^{-1}\mathbf{1}$ the minimum variance portfolio using the idiosyncratic risks of the full factor model and the associated value $\sigma^2_{MVP,0} = 1/\mathbf{1}'\boldsymbol{\Sigma}_0^{-1}\mathbf{1}$ instead of σ^2_{MVP}. Then $\boldsymbol{\Sigma}_0^{-1}\mathbf{1} = \boldsymbol{w}^*_{1F}/\sigma^2_{MVP,0}$ and therefore the formula for $\boldsymbol{\Sigma}_K^{-1}$ of Theorem 1 yields

$$\begin{aligned}
\boldsymbol{w}^* &= \sigma^2_{MVP,K}(\mathbf{I} - \boldsymbol{\Sigma}_0^{-1}(\mathbf{I} + \mathbf{B}_K\boldsymbol{\Sigma}_0^{-1})^{-1}\mathbf{B}_K)\boldsymbol{\Sigma}_0^{-1}\mathbf{1} \\
&= \frac{\sigma^2_{MVP,K}}{\sigma^2_{MVP,0}}(\mathbf{I} - \boldsymbol{\Sigma}_0^{-1}(\mathbf{I} + \mathbf{B}_K\boldsymbol{\Sigma}_0^{-1})^{-1}\mathbf{B}_K)\boldsymbol{w}^*_{1F} \\
&= \frac{\sigma^2_{MVP,K}}{\sigma^2_{MVP,0}}\left(\boldsymbol{w}^*_{1F} - \boldsymbol{\Sigma}_0^{-1}(\mathbf{I} + \mathbf{B}_K\boldsymbol{\Sigma}_0^{-1})^{-1}\mathbf{B}_K\boldsymbol{w}^*_{1F}\right).
\end{aligned}$$

The first formula shows that each K-factor model MVP can be obtained from the one-factor MVP by means of a linear transformation. The second formula shows that the one-factor optimal portfolio $\boldsymbol{w}^*_{1F}$ is corrected for the influence of the additional factors and then rescaled by the minimum-variance ratio $\frac{\sigma^2_{MVP,K}}{\sigma^2_{MVP,0}}$ to yield the K-factor MVP.

Let us first confine our study to the case of one additional factor. Then

$$\boldsymbol{w}^* = \frac{\sigma^2_{MVP,K}}{\sigma^2_{MVP,0}}\left(\boldsymbol{w}^*_{1F} - \boldsymbol{\Sigma}_0^{-1}\left(\mathbf{I} - \frac{\sigma^2_{F1}\boldsymbol{l}_1\boldsymbol{g}_1'}{1 + \sigma^2_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1}\right)\sigma^2_{F1}\boldsymbol{l}_1\boldsymbol{l}_1'\boldsymbol{w}^*_{1F}\right).$$

Define the following expressions.

$$\begin{aligned}
A &= \sigma^2_{F1}\pi_1^*, \\
B &= \frac{\sigma^4_{F1}\pi_1^*\boldsymbol{l}_1'\boldsymbol{g}_1}{1 + \sigma^2_{F1}\boldsymbol{l}'\boldsymbol{g}_1}, \\
C &= \frac{\sigma^4_{F1}\boldsymbol{b}_r'\boldsymbol{l}_1\boldsymbol{g}_1'\boldsymbol{l}_1\pi_1^*}{(1 + \sigma^2_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1)(1/\sigma^2_M + \boldsymbol{b}_r'\boldsymbol{b})}, \\
D &= \frac{\sigma^2_{F1}\pi_1^*\boldsymbol{b}_r'\boldsymbol{l}_1}{1/\sigma^2_M + \boldsymbol{b}_r'\boldsymbol{b}}.
\end{aligned}$$

Theorem 2 *In the two-factor model consisting of the market portfolio and an additional factor the minimum variance portfolio weights are obtained from the optimal one-factor weights, $\mathbf{w}^*_{1F}$, using the two-factor idiosyncratic variances by*

$$\boldsymbol{w}^* = \frac{\sigma^2_{LMV,K}}{\sigma^2_{MVP,0}}(\boldsymbol{w}^*_{1F} - \boldsymbol{w}^*_c)$$

where the correction term, $\mathbf{w}^*_c = (w^*_{ci})_{i=1}^d$, has entries

$$w^*_{ci} = \left(\sigma^2_{1F}\pi_1^* - \frac{\sigma^4_{F1}\pi_1^*(\boldsymbol{l}_1'\boldsymbol{g}_1)}{1 + \sigma^2_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1}\right)l_{1ri} + \left(\frac{\sigma^4_{F1}(\boldsymbol{b}_r'\boldsymbol{l}_1)(\boldsymbol{g}_1'\boldsymbol{l}_1)\pi_1^*}{(1 + \sigma^2_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1)(1/\sigma^2_M + \boldsymbol{b}_r'\boldsymbol{b})} - \frac{\sigma^2_{F1}\pi_1^*(\boldsymbol{b}_r'\boldsymbol{l}_1)}{1/\sigma^2_M + \boldsymbol{b}_r'\boldsymbol{b}}\right)\frac{\beta_i}{\sigma_i^2}.$$

Here, $\pi_1^* = \boldsymbol{l}_1' \mathbf{w}_{1F}^*$ *is the length of the projection of the optimal one-factor portfolio* $\mathbf{w}_{1F}^*$ *(onto the subspace spanned by) the additional factor* $\boldsymbol{l}_1$, $l_{1ri} = l_{1i}/\sigma_i^2$ *are the risk-adjusted factor betas,* $\boldsymbol{g}_1 = (g_{1i})_{i=1}^d$ *with* $g_{1i} = \frac{l_{1i}}{\sigma_i^2} - \frac{\sigma_{F1}^2 \sum_{j=1}^d l_{1j}\beta_j/\sigma_j^2}{1/\sigma_M^2 + \sum_{j=1}^d \beta_j/\sigma_j^2} \frac{\beta_i}{\sigma_i^2}$, $1 \le i \le d$, *and* $\boldsymbol{b}_r'\boldsymbol{b} = \sum_{j=1}^d \frac{\beta_j^2}{\sigma_j^2}$.

If $\beta_i \ge 0$ *for all* $1 \le i \le n$, *then we have the following equivalent characterizations:*

(i) The portfolio holds a long position in asset i, *if and only if* $w_{1Fi}^* > w_{ci}^*$.

(ii) If $C > D$, *then asset* i *is long, if and only if*

$$\beta_i < \beta_{LS,i},$$

and if $C < D$, *then asset* i *is long, if*

$$\beta_i < |\beta_{LS,i}|, \qquad \beta_{LS,i} < 0,$$

where the asset-specific long-short threshold betas are given by

$$\beta_{LS,i} = \frac{\sigma_i^2}{C-D}\left(w_{1F,i}^* - (A+B)l_{1ri}\right)$$

for $1 \le i \le d$.

The long-short thresholds are proportional to the idiosyncratic variances and thus indicate that the two-factor MVP prefers assets with small idiosyncratic risks. However, the result also shows that in the two-factor case with one factor augmenting the market portfolio the structure of the MVP is more complex than in the one-factor case. There is no long-short threshold beta applicable to all assets. Instead, each asset has its own long-short threshold and the threshold can be positive or negative. The assumption of nonnegative betas is empirically justified as shown by many empirical analyses including the example discussed in the next section. Real financial markets of exchange-traded companies have this property. It is important to note that the derived inequalities are a structural statement about the minimum variance portfolio, since the thresholds depend on all inputs. For example, they change if the idiosyncratic variances changes.

Theorem 2 also reveals that the optimal porfolio, $\boldsymbol{w}^*$, can be expressed in terms of the generally non-orthogonal spanning vectors $\boldsymbol{w}_{1F}^*, \boldsymbol{l}_{1r}$ and $\boldsymbol{b}_r$ with coefficients which are nonlinear functions. Specifically,

$$\boldsymbol{w}^* = \vartheta_0 \boldsymbol{w}_{1F}^* + \vartheta_1(\boldsymbol{l}_1)\boldsymbol{l}_{1r} + \vartheta_2(\boldsymbol{l}_1, \boldsymbol{b}_r)\boldsymbol{b}_r,$$

with smooth functions

$$\begin{aligned}
\vartheta_0 &= \frac{\sigma_{LMV,K}^2}{\sigma_{MVP,0}^2},\\
\vartheta_1(\boldsymbol{l}_1) &= -\frac{\sigma_{LMV,K}^2}{\sigma_{MVP,0}^2}\left(\sigma_{1F}^2\pi_1^* - \frac{\sigma_{F1}^4\pi_1^*(\boldsymbol{l}_1'\boldsymbol{g}_1)}{1+\sigma_{F1}^2\boldsymbol{l}_1'\boldsymbol{g}_1}\right),\\
\vartheta_2(\boldsymbol{l}_1, \boldsymbol{b}_r) &= \frac{\sigma_{LMV,K}^2}{\sigma_{MVP,0}^2}\left(\frac{\sigma_{F1}^2\pi_1^*(\boldsymbol{b}_r'\boldsymbol{l}_1)}{1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b}} - \frac{\sigma_{F1}^4(\boldsymbol{b}_r'\boldsymbol{l}_1)(\boldsymbol{g}_1'\boldsymbol{l}_1)\pi_1^*}{(1+\sigma_{F1}^2\boldsymbol{l}_1'\boldsymbol{g}_1)(1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b})}\right).
\end{aligned}$$

This means, the optimal portfolio is embedded in a threedimensional subspace spanned by the one-factor optimal portfolio, the risk-adjusted factor betas and the risk-adjusted market betas.

In general factor models one can also derive a long-short criterion in terms of a condition on the betas. However, the expressions need matrix calculus for their formulation and require to invert a $d \times d$ matrix.

Theorem 3 *Consider a general multi-factor model with K factors and $\beta_i \geq 0$ for all $1 \leq i \leq d$. Let $a = \boldsymbol{b}_r'(I + \mathbf{B}_K \boldsymbol{\Sigma}_0^{-1})^{-1}(-\mathbf{B}_K \boldsymbol{w}_{1F}^*)$. If $a > 0$, then the MVP is long in asset i, if and only if*

$$\beta_i < \beta_{LS,i} = \sigma_i^2 \frac{1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b}}{\boldsymbol{b}_r'(I + \mathbf{B}_K \boldsymbol{\Sigma}_0^{-1})^{-1}(-\mathbf{B}_K \boldsymbol{w}_{1F}^*)} \left(w_{1F,i}^* - \frac{1}{\sigma_i^2}[(I + \mathbf{B}_K \boldsymbol{\Sigma}_0^{-1})^{-1}]_i \mathbf{B}_K \boldsymbol{w}_{1F}^* \right)$$

for $1 \leq i \leq d$, where $[\mathbf{A}]_i$ denotes the ith row vector of a matrix $\mathbf{A}$, $\mathbf{B}_K = \sum_{k=1}^K \sigma_{Fk}^2 \boldsymbol{l}_k \boldsymbol{l}_k'$ is the contribution of the K additional factors to the covariance matrix of the asset returns and $\boldsymbol{w}_{1F}^*$, $\boldsymbol{b}_r$ and $\sigma_1^2, \ldots, \sigma_d^2$ are as in the previous result.

If $a < 0$, then the MVP is long in asset i, if and only if

$$\beta_i < |\beta_{LS,i}|, \qquad \beta_{LS,i} < 0,$$

for $1 \leq i \leq d$.

5 Empirical analysis

To illustrate the results, we analyze daily asset returns of S&P 500 stocks using data from 1,008 trading days for the years 2019 to 2022 downloaded from yahoo finance. One may criticize that this period includes the COVID-19 crash. Below we provide an analysis of the period from 2016 to 2019 to indicate and highlight the differences. Generally, the asset returns were demeaned using a running mean before estimating the covariance matrix. Clearly, now all quantities such as the beta factors are calculated from the data. Basically, two analyses were conducted:

Firstly, a two-factor model without observable factors, named pure factor model in what follows, estimated by PCA. Here the market portfolio is estimated by the leading principal component. This means, the leading eigenvector, $\hat{\boldsymbol{u}}_0$, of the estimated covariance matrix plays the role of the market. One flips the sign of the leading eigenvector if the majority of entries is negative and then estimates $\boldsymbol{b}$ by $\hat{\boldsymbol{u}}_0$. The associated principal component, $\mathbf{R}\hat{\boldsymbol{u}}_0$, is the return series of the estimated market and $\mathbf{P}\hat{\boldsymbol{u}}_0$, where $\mathbf{P}$ denotes the $n \times d$ matrix of asset prices, yields the associated price series, i.e. the estimated prices of the market portfolio. For the data under investigation, this results in an estimator with positive entries. Figure 1 demonstrates that the estimated market index closely resembles the S&P 500 index commonly used as a proxy. The sample coefficient of correlation between $\mathbf{R}\hat{\boldsymbol{u}}_0$ and the S&P 500 returns is 0.929.

Secondly, a factor model was considered with the S&P 500 as observable factor and four additional unobserved factors determined by PCA within the orthogonal complement. In the following, we discuss in parallel the results for both analyses.

The left panel of Fig. 2 shows a plot of the beta factors of all assets for the pure factor model approach, and the right panel for the extended factor model. Assets which are long in the MVP are marked blue and shorted assets are in red. Although there is no clear cut as in a one-factor model, there is a visible tendency that the MVP is long in low-beta assets. For the factor model with five factors, there is still a tendency that the MVP is long in low-beta assets, but the point clouds of long and short assets have a large overlap.

Our theoretical results resolve this structure: The asset-specific long-short thresholds allow to identify long and short positions. Figure 3 shows betas and the associated long-short

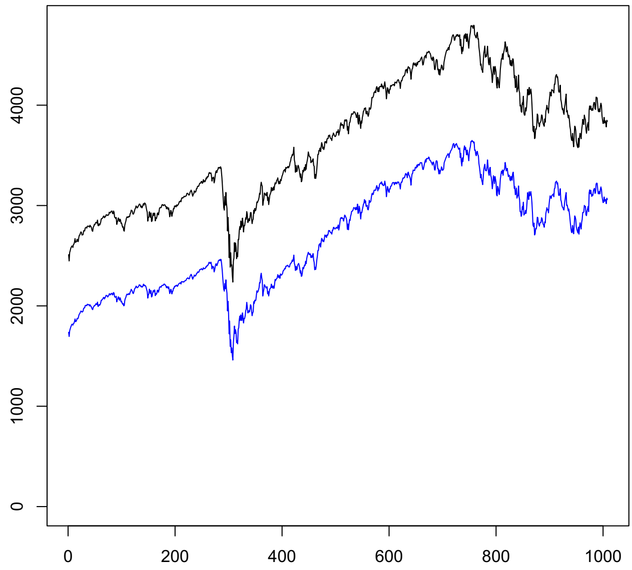

Fig. 1 S&P 500 (black) and market index estimated by PCA (blue)

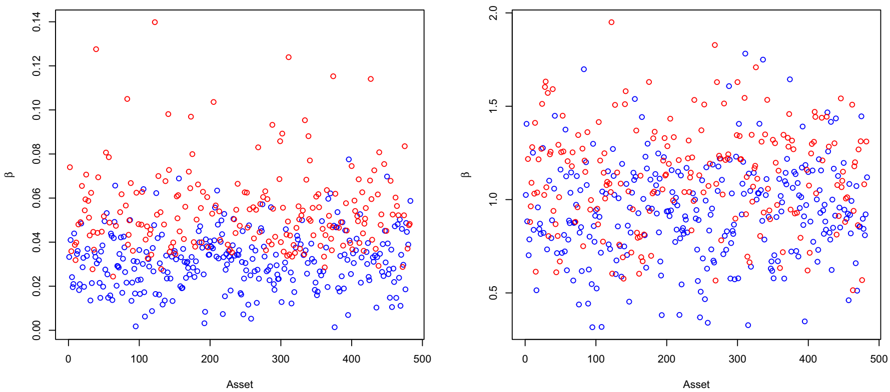

Fig. 2 Shown are estimated betas of all assets. Long positions of the minimum-variance portfolio are marked in blue, short positions are red. Left: Full factor model. Right: Five-factor model, first factor is S&P 500

thresholds $|\hat{\beta}_{LS,i}|$ (connected by a line) for all long positions of the full two-factor MVP where the market portfolio is estimated by the leading principal component. 56.73% of the MVP positions are long. There are some cases where the characterization is violated (marked in red in the plot). A possible explanation is estimation error of the idiosyncratic variances: Contrary to a five-factor model, a two-factor model cannot fully explain the dependence structure of the assets, such that the covariance matrix of the error terms is not well approximated by a diagonal matrix. It is puzzling that the empirical thresholds are large compared to the betas. Thus, they seem to provide only loose bounds for the betas, which complicates their interpretation in practice. The analysis was repeated based on nonlinear shrinkage estimator of the assets' covariance matrix, see [12], but the result were almost the same. However, the structural relationship between the betas, $\hat{\beta}_i$, and their long-short threshods, $\hat{\beta}_{LS,i}$, explains

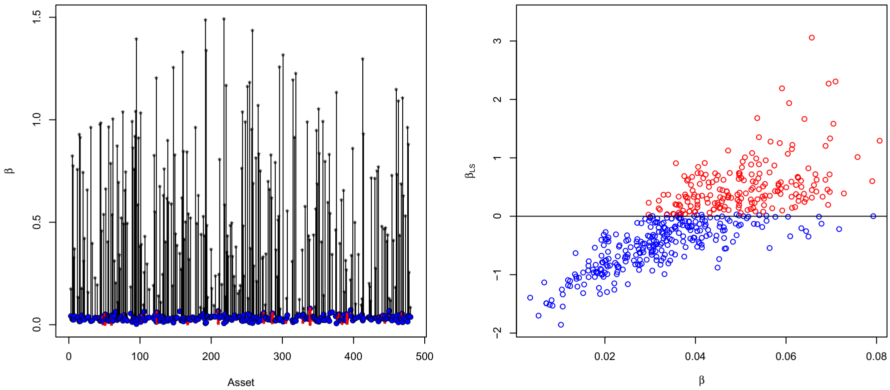

Fig. 3 Pure two-factor model betas and long-short threshold betas. Left: Estimated betas, $\hat{\beta}_i$, and their long-short thresholds $\hat{\beta}_{LS,i}$ for long positions. Cases violating the rule are marked in red. Right: Plot of estimated betas and $\hat{\beta}_{LS,i}$. Long positions are characterized by negative thresholds, short positions by positive ones

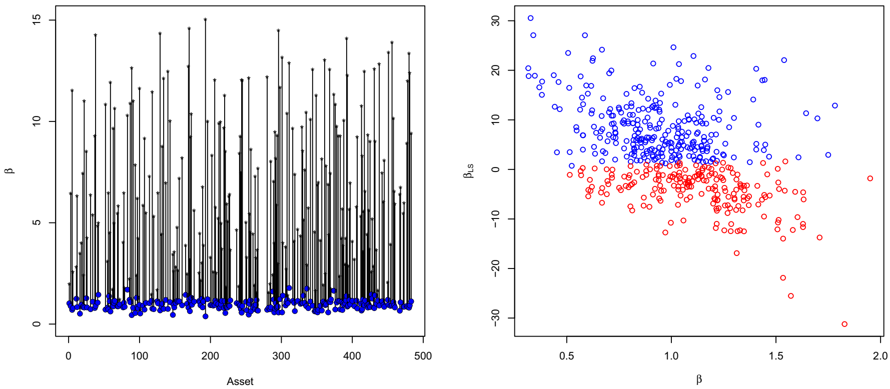

Fig. 4 Five-factor model betas and long-short-threshold betas. Left: Estimated betas, $\hat{\beta}_i$, and their long-short thresholds $\hat{\beta}_{LS,i}$ for long positions. Cases violating the rule are marked in red. Right: Plot of estimated betas and $\hat{\beta}_{LS,i}$. Long positions in blue, short positions in red

the long-short structure of the MVP. The second illustration in Fig. 3 plots the betas, $\hat{\beta}_i$, against their thresholds, $\hat{\beta}_{LS,i}$, for all assets. One can see that the sign of the thresholds is highly informative and determines the sign (long or short) of the MVP position.

Figure 4 provides the corresponding plots for the five-factor model where the S&P 500 serves as market proxy. Now there is a 1–1 relationship between the sign of a position (long/short) and the ordering of the assets' betas and their long-short thresholds.

The plot in Fig. 5 provides an alternative representation. It depicts the estimates for the differences $\hat{\beta}_i - \hat{\beta}_{LS,i}$ betweeen the betas and their thresholds for all assets i. Such a plot can help in asset monitoring by identifying those securities whose betas are close to the threshold. Then, a small change in their linkage to the market may change the sign of the position. Contrary, assets with betas far away from the threshold are likely to keep their sign even if there are small frictions resulting in changing betas. However, because the asset-specific long-short threholds depend on all betas, this interpretation needs some care.

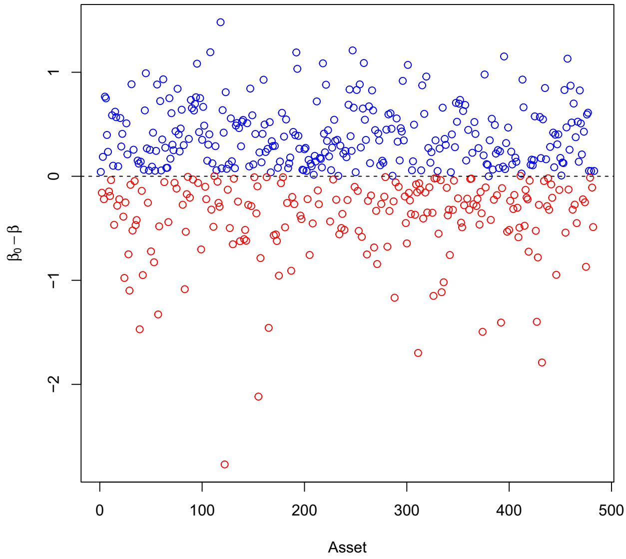

Fig. 5 Estimates of the distances, $\hat{\beta}_i - \hat{\beta}_{LS,i}$ of all betas from their respective long-short thresholds using a two-factor model estimated by PCA. Long positions are blue, short positions red

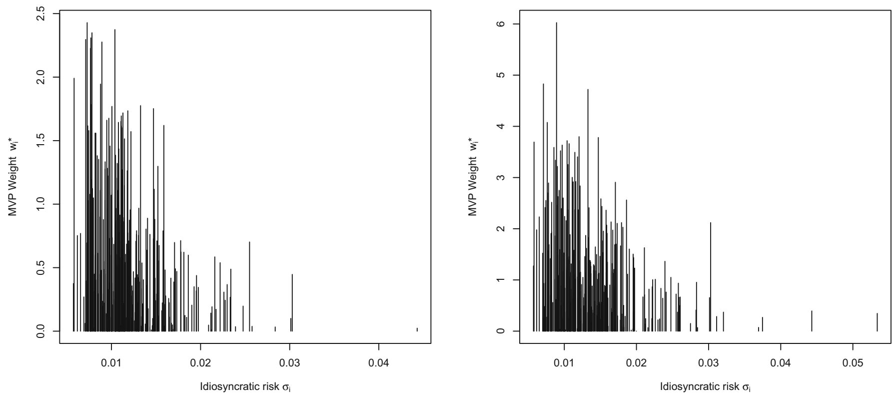

Fig. 6 The MVP is long in low risk assets: Plot of the idiosyncratic standard deviations against the assets' weights in the minimum variance portfolio (in %). Left: Pure two-factor model. Right: Five-factor model with S&P 500 as observed factor

Figure 6 plots the idiosyncratic risks, expressed as the standard deviations σ_i, against the assets' weights in the two-factor MVP. The analysis is in agreement with the obtained bounds, $|\hat{\beta}_{LS,i}|$, which are proportional to σ_i^2 (ceteris paribus): The MVP is long in low-risk assets.

Figure 7 analyzes for the five-factor model the relationship between the weigth of an asset in the minimum variance portfolio and the distance of its beta, $\hat{\beta}_i$, from the applicable long-short threshold, $\hat{\beta}_{LS,i}$. For both long and short positions there is quite strong empirical relationship between the distance and the portofolio weight. Simple linear regression for long and short positions, respecitvely, with distances not exceeding 30 yield adjusted R^2 values of 0.45 and 0.57.

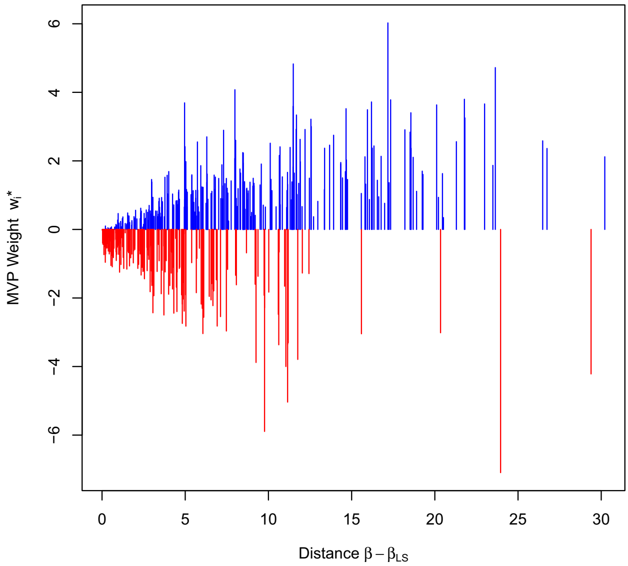

Fig. 7 Plot of the distance, $|\hat{\beta}_i - \hat{\beta}_{LS,i}|$, of asset betas from their long-short tresholds against the weight in the five-factor minimum variance portfolio. Long positions are in blue, short positions in red

One may criticize that the time span of the data includes the COVID-19 crash and the market turmoils of the following period. Indeed, Bours and Steland [4] found a break (change-point) in the loadings of the Fama/French factors dated between February 21 and February 28, 2020, and Steland [18] provides evidence for a significant change in the time-frequency representation of the S&P 500 investment universe. Thus, the five-factor model including the S&P 500 index was refitted to the shorter time period from 2016–2019 where markets were relatively stable, at least compared to the period 2020–2022. Figure 8 provides the crucial plots of the estimated betas and long-short thresholds. Now, the long-short threshold are much smaller in magnitude and show a notable smaller variation, and the thresholds provide considerably tighter bounds for the beta factors.

6 Conclusions

It is shown that in a multi-factor models for asset pricing the global minimum-variance portfolio is long in low-beta assets similar as in a one-factor model where explicit formulas are known. Although now the long-short threshold is asset-specific, this sheds new light on this question. For general multi factor asset pricing models a criterion in terms of the beta factors can also be developed but not in closed-form. In addition, recursive formulas are derived which simplify the calculation of inverse covariance matrices in multi factor models, which are a basic ingredient of optimal porfolios related to the classical and widely used Markowitz framework. Empirical analyses of the constituents of the S&P 500 confirm the theoretical findings. For a five-factor asset pricing model with the S&P 500 index return as observable market proxy the derived criteria uniquely determine the long and short positions of the minimum variance portfolio. However, the analysis reveals a quite unexpected behaviour of the estimated long-short thresholds, which could motivate future research on this question.

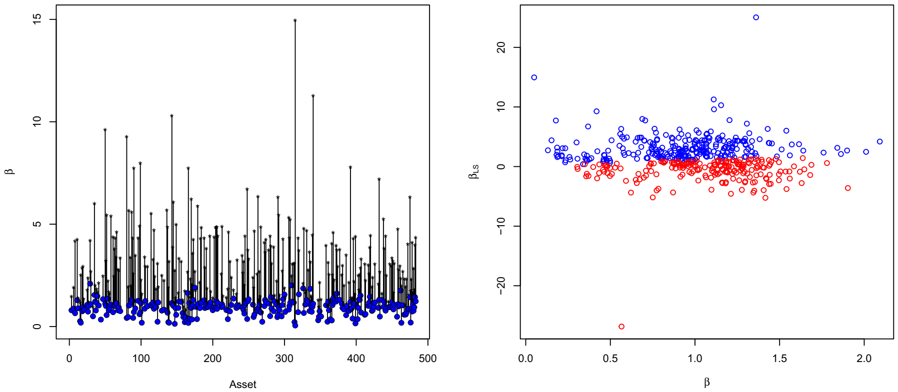

Fig. 8 Five-factor model betas and long-short-threshold betas for the period 2016–2019. Left: Estimated betas, $\hat{\beta}_i$, and their long-short thresholds $\hat{\beta}_{LS,i}$ for long positions. Cases violating the rule are marked in red. Right: Plot of estimated betas and $\hat{\beta}_{LS,i}$. Long positions in blue, short positions in red

Appendix: Proofs

In the derivations, we make frequent use of the following known result, which we prove for completeness.

Lemma A *Let $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ be vectors, $x \in \mathbb{R}$, and let $\mathbf{D}$ be an invertible $n \times n$ matrix with real-valued entries and inverse $\mathbf{E} = \mathbf{D}^{-1}$. Then*

$$(\mathbf{D} + x\boldsymbol{a}'\boldsymbol{b})^{-1} = \mathbf{E} - \frac{x}{1 + x\boldsymbol{a}'\mathbf{E}\boldsymbol{b}}\boldsymbol{a}'\boldsymbol{b}$$

Proof of Lemma A: Make the ansatz $(\mathbf{D} + x\boldsymbol{a}'\boldsymbol{b})^{-1} = \mathbf{E} - y\boldsymbol{a}'\boldsymbol{b}$. Then, after some algebra,

$$(\mathbf{D} + x\boldsymbol{a}'\boldsymbol{b})(\mathbf{E} - y\boldsymbol{a}'\boldsymbol{b}) = \mathbf{I} + (x - y - xy\boldsymbol{a}'\mathbf{E}\boldsymbol{b})\boldsymbol{a}(\mathbf{E}\boldsymbol{b})'$$

Solving $x - y - xy\boldsymbol{a}'\mathbf{E}\boldsymbol{b} = 0$ for y establishes the assertion. □

Proof of Theorem 1 To show the first general formula, observe that we may write $\boldsymbol{\Sigma}_K$ as

$$\boldsymbol{\Sigma}_K = \boldsymbol{\Sigma}_0 + \mathbf{B}_K,$$

where

$$\mathbf{B}_K = \sum_{k=1}^{K} \sigma_{Fk}^2 \boldsymbol{l}_k \boldsymbol{l}_k'$$

and $\boldsymbol{\Sigma}_0 = \sigma_M^2 \boldsymbol{b}\boldsymbol{b}' + \mathbf{S}_0$ with $\mathbf{S}_0 = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_d^2)$. Thus, $\boldsymbol{\Sigma}_0$ is a 1-factor-model covariance matrix, but now the idiosyncratic risks are those from the multi-factor regression. The representation of $\boldsymbol{\Sigma}_K$ as the sum of the 1-factor model matrix $\boldsymbol{\Sigma}_0$ and a perturbation matrix $\mathbf{B}_K$ allows to make use of explicit formulas for the inverse $\boldsymbol{\Sigma}^{-1}$ of an invertible matrix of this form. By [9, formula (23)]

$$\boldsymbol{\Sigma}_K^{-1} = (\mathbf{I} - \boldsymbol{\Sigma}_0^{-1}(\mathbf{I} + \mathbf{B}_K\boldsymbol{\Sigma}_0^{-1})^{-1}\mathbf{B}_K)\boldsymbol{\Sigma}_0^{-1}.$$

Here, $\boldsymbol{\Sigma}_0^{-1} = \left(\mathbf{S}_0^{-1} - \frac{\boldsymbol{b}_r\boldsymbol{b}_r'}{1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b}}\right)$, where $\boldsymbol{b}_r = (\beta_1/\sigma_1^2, \ldots, \beta_d/\sigma_d^2)'$.

Let us now consider $\boldsymbol{\Sigma}_1$ such that $\mathbf{B}_1 = \sigma_{F1}^2 \boldsymbol{l}_1 \boldsymbol{l}_1'$. We have

$$\begin{aligned}\mathbf{I} + \mathbf{B}_1 \boldsymbol{\Sigma}_0^{-1} &= \mathbf{I} + \sigma_{F1}^2 \boldsymbol{l}_1 \boldsymbol{l}_1' \left(\mathbf{S}_0^{-1} - \frac{\boldsymbol{b}_r \boldsymbol{b}_r'}{1/\sigma_M^2 + \boldsymbol{b}_r' \boldsymbol{b}} \right) \\ &= \mathbf{I} + \sigma_{F1}^2 \boldsymbol{l}_1 (\mathbf{S}_0^{-1} \boldsymbol{l}_1)' - \frac{\sigma_{F1}^2 \boldsymbol{l}_1' \boldsymbol{b}_r}{1/\sigma_M^2 + \boldsymbol{b}_r' \boldsymbol{b}} \boldsymbol{l}_1 \boldsymbol{b}_r' \\ &= \mathbf{I} + \sigma_{F1}^2 \boldsymbol{l}_1 \boldsymbol{g}_1'\end{aligned}$$

where

$$\boldsymbol{g}_1 = (\mathbf{S}_0^{-1} \boldsymbol{l}_1) - \frac{\sigma_{F1}^2 \boldsymbol{l}_1' \boldsymbol{b}_r}{1/\sigma_M^2 + \boldsymbol{b}_r' \boldsymbol{b}} \boldsymbol{b}_r$$

with entries $g_{1i} = \frac{l_{1i}}{\sigma_i^2} - \frac{\sigma_{F1}^2 \sum_{j=1}^d l_{1j} \beta_j / \sigma_j^2}{1/\sigma_M^2 + \sum_{j=1}^d \beta_j / \sigma_j^2} \frac{\beta_i}{\sigma_i^2}$, $1 \le i \le d$. The vector $\boldsymbol{g}_1$ is given by the scaled factor loadings corrected by a scaled projection of the factor loadings $\boldsymbol{l}_1$ onto the risk-adjusted beta factors $\boldsymbol{b}_r$. It follows that

$$(\mathbf{I} + \mathbf{B}_1 \boldsymbol{\Sigma}_0^{-1})^{-1} = \mathbf{I} - \frac{\sigma_{F1}^2 \boldsymbol{l}_1 \boldsymbol{g}_1'}{1 + \sigma_{F1}^2 \boldsymbol{l}_1' \boldsymbol{g}_1}.$$

Thus, we arrive at

$$\boldsymbol{\Sigma}_1^{-1} = \left(\mathbf{I} - \boldsymbol{\Sigma}_0^{-1} \left[\mathbf{I} - \frac{\sigma_{F1}^2 \boldsymbol{l}_1 \boldsymbol{g}_1'}{1 + \sigma_{F1}^2 \boldsymbol{l}_1' \boldsymbol{g}_1} \right] \sigma_{F1}^2 \boldsymbol{l}_1 \boldsymbol{l}_1' \right) \boldsymbol{\Sigma}_0^{-1}.$$

For $k \ge 2$ observe that

$$\boldsymbol{\Sigma}_k = \boldsymbol{\Sigma}_{k-1} + \mathbf{E}_k, \qquad \mathbf{E}_k = \sigma_{Fk}^2 \boldsymbol{l}_k \boldsymbol{l}_k'$$

where $\boldsymbol{\Sigma}_{k-1}$ is the covariance matrix using factors $\boldsymbol{l}_1, \ldots, \boldsymbol{l}_{k-1}$ and the idiosyncratic variances from the full factor model with all K factors. By [9, formula (23)]

$$\boldsymbol{\Sigma}_k^{-1} = (\mathbf{I} - \boldsymbol{\Sigma}_{k-1}^{-1} (\mathbf{I} + \mathbf{E}_k \boldsymbol{\Sigma}_{k-1}^{-1})^{-1} \mathbf{E}_k) \boldsymbol{\Sigma}_{k-1}^{-1}. \tag{6}$$

Write

$$\mathbf{I} + \mathbf{E}_k \boldsymbol{\Sigma}_{k-1}^{-1} = \mathbf{I} + \sigma_{Fk}^2 \boldsymbol{l}_k \boldsymbol{l}_k' \boldsymbol{\Sigma}_{k-1}^{-1} = \mathbf{I} + \sigma_{Fk}^2 \boldsymbol{l}_k \boldsymbol{g}_k'$$

with

$$\boldsymbol{g}_k = \boldsymbol{\Sigma}_{k-1}^{-1} \boldsymbol{l}_k.$$

Then

$$(\mathbf{I} + \mathbf{E}_k \boldsymbol{\Sigma}_{k-1}^{-1})^{-1} = \mathbf{I} - \frac{\sigma_{Fk}^2}{1 + \sigma_{Fk}^2 \boldsymbol{l}_k' \boldsymbol{g}_k} \boldsymbol{l}_k \boldsymbol{g}_k'.$$

Pluging this into the formula for $\boldsymbol{\Sigma}_k^{-1}$ yields the inverse-free recursion

$$\boldsymbol{\Sigma}_k^{-1} = \left(\mathbf{I} - \boldsymbol{\Sigma}_{k-1}^{-1} \left[\mathbf{I} - \frac{\sigma_{Fk}^2}{1 + \sigma_{Fk}^2 \boldsymbol{l}_k' \boldsymbol{g}_k} \boldsymbol{l}_k \boldsymbol{g}_k' \right] \sigma_{Fk}^2 \boldsymbol{l}_k \boldsymbol{l}_k' \right) \boldsymbol{\Sigma}_{k-1}^{-1}, \qquad k \ge 1.$$

This completes the proof. □

Proof of Theorem 2

$$\begin{aligned}
\boldsymbol{w}^* &= \frac{\sigma^2_{MVP,K}}{\sigma^2_{MVP,0}}\left(\boldsymbol{w}^*_{1F} - \left[\mathbf{S}_0^{-1} - \frac{\boldsymbol{b}_r\boldsymbol{b}_r'}{1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b}}\right]\left[\mathbf{I} - \frac{\sigma^2_{F1}\boldsymbol{l}_1\boldsymbol{g}_1'}{1+\sigma^2_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1}\right]\sigma^2_{F1}\boldsymbol{l}_1\boldsymbol{l}_1'\boldsymbol{w}^*_{1F}\right)\\
&= \frac{\sigma^2_{MVP,K}}{\sigma^2_{MVP,0}}\left(\boldsymbol{w}^*_{1F} - \boldsymbol{w}^*_c\right)
\end{aligned}$$

where the optimal portfolio weights $\boldsymbol{w}^*_{1F}$ are corrected by

$$\boldsymbol{w}^*_c = \left[\mathbf{S}_0^{-1} - \frac{\boldsymbol{b}_r\boldsymbol{b}_r'}{1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b}}\right]\left[\mathbf{I} - \frac{\sigma^2_{F1}\boldsymbol{l}_1\boldsymbol{g}_1'}{1+\sigma^2_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1}\right]\sigma^2_{F1}\boldsymbol{l}_1\boldsymbol{l}_1'\boldsymbol{w}^*_{1F}$$

where $\boldsymbol{g}_1$ is as in the previous proof. Put $w_1^* = \boldsymbol{l}_1'\boldsymbol{w}^*_{1F}$ and let $\boldsymbol{l}_{1r} = \mathbf{S}_0^{-1}\boldsymbol{l}_1 = (\frac{l_{1i}}{\sigma_i^2})_{i=1}^d$ be the risk-adjusted factor betas. The correction term can be simplified as follows.

$$\begin{aligned}
\boldsymbol{w}^*_c &= \left[\mathbf{S}_0^{-1} - \frac{\boldsymbol{b}_r\boldsymbol{b}_r'}{1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b}}\right]\left[\mathbf{I} - \frac{\sigma^2_{F1}\boldsymbol{l}_1\boldsymbol{g}_1'}{1+\sigma^2_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1}\right]\sigma^2_{F1}\boldsymbol{l}_1\boldsymbol{l}_1'\boldsymbol{w}^*_{1F}\\
&= \left[\mathbf{S}_0^{-1} - \frac{\boldsymbol{b}_r\boldsymbol{b}_r'}{1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b}} - \mathbf{S}_0^{-1}\frac{\sigma^2_{F1}\boldsymbol{l}_1\boldsymbol{g}_1'}{1+\sigma^2_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1} + \frac{\boldsymbol{b}_r\boldsymbol{b}_r'}{1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b}}\frac{\sigma^2_{F1}\boldsymbol{l}_1\boldsymbol{g}_1'}{1+\sigma^2_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1}\right]\sigma^2_{F1}\boldsymbol{l}_1\boldsymbol{l}_1'\boldsymbol{w}^*_{1F}\\
&= \left[\mathbf{S}_0^{-1}\sigma^2_{F1}\boldsymbol{l}_1\boldsymbol{l}_1' - \frac{\boldsymbol{b}_r'\boldsymbol{l}_1\sigma^2_{F1}\boldsymbol{b}_r\boldsymbol{l}_1'}{1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b}} - \mathbf{S}_0^{-1}\frac{\sigma^4_{F1}\boldsymbol{g}_1'\boldsymbol{l}_1\boldsymbol{l}_1\boldsymbol{l}_1'}{1+\sigma^2_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1} + \frac{\sigma^4_{F1}\boldsymbol{b}_r(\boldsymbol{b}_r'\boldsymbol{l}_1)(\boldsymbol{g}_1'\boldsymbol{l}_1)\boldsymbol{l}_1'}{(1+\sigma^2_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1)(1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b})}\right]\boldsymbol{w}^*_{1F}\\
&= \sigma^2_{F1}\pi_1^*\boldsymbol{l}_{1r} - \frac{\sigma^4_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1\pi_1^*}{1+\sigma^2_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1}\boldsymbol{l}_{1r} + \boldsymbol{b}_r\left(\frac{\sigma^4_{F1}\boldsymbol{b}_r'\boldsymbol{l}_1\boldsymbol{g}_1'\boldsymbol{l}_1\pi_1^*}{(1+\sigma^2_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1)(1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b})} - \frac{\sigma^2_{F1}\boldsymbol{b}_r'\boldsymbol{l}_1\pi_1^*}{1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b}}\right)\\
&= (A-B)\boldsymbol{l}_{1r} + (C-D)\boldsymbol{b}_r,
\end{aligned}$$

where the expressions A, B, C, D are given in the theorem. Especially, the entries w^*_{ci} of $\boldsymbol{w}^*_c$ are given by

$$w^*_{ci} = \left(\sigma^2_{1F}\pi_1^* - \frac{\sigma^4_{F1}\pi_1^*(\boldsymbol{l}_1'\boldsymbol{g}_1)}{1+\sigma^2_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1}\right)l_{1ri} + \left(\frac{\sigma^4_{F1}(\boldsymbol{b}_r'\boldsymbol{l}_1)(\boldsymbol{g}_1'\boldsymbol{l}_1)\pi_1^*)}{(1+\sigma^2_{F1}\boldsymbol{l}_1'\boldsymbol{g}_1)(1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b})} - \frac{\sigma^2_{F1}\pi_1^*(\boldsymbol{b}_r'\boldsymbol{l}_1)}{1/\sigma_M^2 + \boldsymbol{b}_r'\boldsymbol{b}}\right)\frac{\beta_i}{\sigma_i^2},$$

where l_{1ri} are the entries of $\boldsymbol{l}_{1r}$ and $g_{1i} = \frac{l_{1i}}{\sigma_i^2} - \frac{\sigma^2_{F1}\sum_{j=1}^d l_{1j}\beta_j/\sigma_j^2}{1/\sigma_M^2 + \sum_{j=1}^d \beta_j/\sigma_j^2}\frac{\beta_i}{\sigma_i^2}$, $1 \le i \le d$. We can conclude that the position i is a long position, if and only if $w^*_{1Fi} > w^*_{ci}$.

Lastly, let us show (ii) and thus assume $\beta_i \ge 0$.

First, consider the case $C > D$: Then solving the inequality $w^*_{F1,i} \ge w^*_{ci}$ for β_i, immediately yields the long-short criterion $\beta_i < \beta_{LS,i}$.

Next consider the case $C < D$: By definition of $\beta_{LS,i} = \frac{\sigma_i^2}{C-D}(w^*_{1F} - (A-B)l_{1ri})$ we have the equality

$$w_i^* = \frac{\sigma^2_{MVP,K}}{\sigma^2_{MVP,0}}(C-D)(\beta_{LS,i} + \beta_i)$$

Thus, $w_i^* > 0$ if and only if $\beta_{LS,i} + \beta_i < 0$. This implies $\beta_{LS,i} < 0$, since otherwise $\beta_{LS,i} + \beta_i \ge 0$ (because $\beta_i \ge 0$), which in turn implies $w_i^* < 0$, in view of $C < D$, a contradiction. But $\beta_{LS,i} + \beta_i < 0$ and $\beta_{LS,i} < 0$ holds if and only if

$$\beta_i < -\beta_{LS,i} = |\beta_{LS,i}|, \qquad \beta_{LS,i} < 0.$$

□

Proof of Theorem 3 In view of the formula for $\mathbf{\Sigma}_0^{-1}$ we have

$$\begin{aligned} \boldsymbol{w}^* &= \frac{\sigma^2_{MVP,K}}{\sigma^2_{MVP,0}} \left(\boldsymbol{w}^*_{1F} - \mathbf{\Sigma}_0^{-1}(\mathbf{I} + \mathbf{B}_K \mathbf{\Sigma}_0^{-1})^{-1} \mathbf{B}_K \boldsymbol{w}^*_{1F} \right) \\ &= \frac{\sigma^2_{MVP,K}}{\sigma^2_{MVP,0}} \left(\boldsymbol{w}^*_{1F} - (\mathbf{S}_0^{-1} - \frac{\boldsymbol{b}_r \boldsymbol{b}_r'}{1/\sigma_M^2 + \boldsymbol{b}_r' \boldsymbol{b}})(I + \mathbf{B}_K \mathbf{\Sigma}_0^{-1})^{-1} \mathbf{B}_K \boldsymbol{w}^*_{1F} \right) \\ &= \frac{\sigma^2_{MVP,K}}{\sigma^2_{MVP,0}} \left(\boldsymbol{w}^*_{1F} - \mathbf{S}_0^{-1}(I + \mathbf{B}_K \mathbf{\Sigma}_0^{-1})^{-1} \mathbf{B}_K \boldsymbol{w}^*_{1F} - \frac{\boldsymbol{b}_r'(I + \mathbf{B}_K \mathbf{\Sigma}_0^{-1})^{-1}(-\mathbf{B}_K \boldsymbol{w}^*_{1F})}{1/\sigma_M^2 + \boldsymbol{b}_r' \boldsymbol{b}} \boldsymbol{b}_r \right). \end{aligned}$$

Therefore, $w_i^* > 0$ if and only if

$$w^*_{1F,i} - \frac{1}{\sigma_i^2}[(I + \mathbf{B}_K \mathbf{\Sigma}_0^{-1})^{-1}]_i \mathbf{B}_K \boldsymbol{w}^*_{1F} > \frac{\boldsymbol{b}_r'(I + \mathbf{B}_K \mathbf{\Sigma}_0^{-1})^{-1}(-\mathbf{B}_K \boldsymbol{w}^*_{1F})}{1/\sigma_M^2 + \boldsymbol{b}_r' \boldsymbol{b}} \frac{\beta_i}{\sigma_i^2} \tag{7}$$

If

$$a = \boldsymbol{b}_r'(I + \mathbf{B}_K \mathbf{\Sigma}_0^{-1})^{-1}(-\mathbf{B}_K \boldsymbol{w}^*_{1F}) > 0,$$

then, in view of (7), $w_i^* > 0$ if and only if

$$\beta_i < \beta_{LS,i} = \sigma_i^2 \frac{1/\sigma_M^2 + \boldsymbol{b}_r' \boldsymbol{b}}{\boldsymbol{b}_r'(I + \mathbf{B}_K \mathbf{\Sigma}_0^{-1})^{-1}(-\mathbf{B}_K \boldsymbol{w}^*_{1F})} \left(w^*_{1F,i} - \frac{1}{\sigma_i^2}[(I + \mathbf{B}_K \mathbf{\Sigma}_0^{-1})^{-1}]_i \mathbf{B}_K \boldsymbol{w}^*_{1F} \right).$$

If $a < 0$, then we may argue as in the previous proof. The equation

$$w_i^* = \frac{\sigma^2_{MVP,K}}{\sigma^2_{MVP,0}} a(\beta_{i,LS} + \beta_i)$$

again leads to $\beta_{i,LS} < 0$ and $\beta_i < |\beta_{i,LS}|$, since $\beta_i \geq 0$, $1 \leq i \leq n$, by assumption. □

Funding Open Access funding enabled and organized by Projekt DEAL.

References

1. Abadir, K.M.: Explicit minimal representation of variance matrices, and its implication for dynamic volatility models. Econom. J. **26**(1), 88–104 (2022)
2. Annaert, J., De Ceuster, M., Van Doninck, F.: Decomposing the idiosyncratic volatility anomaly among euro area stocks. Financ. Res. Lett. **47**, 102672 (2022)
3. Bai, J., Wang, P.: Econometric analysis of large factor models. Annu. Rev. Econ. **8**(1), 53–80 (2016)
4. Bours, M., Steland, A.: Large-sample approximations and change testing for high-dimensional covariance matrices of multivariate linear time series and factor models. Scand. J. Stat. **48**(2), 610–654 (2021)

5. Clarke, R.G., de Silva, H., Thorley, S.: Minimum Variance Portfolio Composition. J. Portf. Manag. **2**, 31–45 (2011)
6. Clarke, R.G., de Silva, H., Thorley, S.: Minimum-variance portfolios in the U.S. equity market. J. Portf. Manag. **33**(1), 10–24 (2006)
7. Disatnik, D., Katz, S.: Portfolio optimization using a block structure for the covariance matrix. J. Bus. Finance Account. **39**(5–6), 806–843 (2012)
8. Fama, E.F.: Foundations of Finance: Portfolio Decisions and Securities Prices. Basic Books, New York (1976)
9. Henderson, H.V., Searle, S.R.: On deriving the inverse of a sum of matrices. SIAM Rev. **23**(1), 53–60 (1981)
10. Jacobs, B.I., Levy, K.N., Starer, D.: On the optimality of long-short strategies. Financ. Anal. J. **54**(2), 40–51 (1998)
11. Ledoit, O., Wolf, M.: A well-conditioned estimator for large-dimensional covariance matrices. J. Multivar. Anal. **88**(2), 365–411 (2004)
12. Ledoit, O., Wolf, M.: Nonlinear shrinkage estimation of large-dimensional covariance matrices. Ann. Stat. **40**(2), 1024–1060 (2012)
13. Markowitz, H.: Portfolio selection. J. Finance **7**, 77–91 (1952)
14. Mies, F., Steland, A.: Projection inference for high-dimensional covariance matrices with structured shrinkage targets (2022)
15. Qi, H.-D.: On the long-only minimum variance portfolio under single factor model. Oper. Res. Lett. **49**(5), 795–801 (2021)
16. Sancetta, A.: Sample covariance shrinkage for high dimensional dependent data. J. Multivar. Anal. **99**(5), 949–967 (2008)
17. Hwang, H.: Two-step estimation of a factor model in the presence of observable factors. Econ. Lett. **105**(3), 247–249 (2009)
18. Steland, A.: Flexible nonlinear inference and change-point testing of high-dimensional spectral density matrices. J. Multivar. Anal. **199**, 105245 (2024)
19. Steland, A., von Sachs, R.: Large-sample approximations for variance-covariance matrices of high-dimensional time series. Bernoulli **23**(4A), 2299–2329 (2017)
20. Steland, A., von Sachs, R.: Asymptotics for high-dimensional covariance matrices and quadratic forms with applications to the trace functional and shrinkage. Stoch. Process. Appl. **128**(8), 2816–2855 (2018)
21. Stevens, G.V.G.: On the inverse of the covariance matrix in portfolio analysis. J. Financ. **53**(5), 1821–1827 (1998)
22. Traut, J.: What we know about the low-risk anomaly: a literature review. Fin. Mark. Portf. Mgmt. **37**(3), 297–324 (2023)