

Flexible shrinkage in portfolio selection

Vasyl Golosnoy^a, Yarema Okhrin^{b,*}

^a *Institute of Statistics and Econometrics, University of Kiel, Germany*

^b *Department of Economics, University of Bern, Switzerland*

A B S T R A C T

How to quantify estimation risk is important in portfolio selection. For this purpose we derive the flexible shrinkage estimator for the optimal portfolio weights, which allows dynamic adjustments of model structure. Our estimator is based on grouping the assets in order to capture non-homogeneity of estimation risk. The assets are assigned to groups using a clustering procedure with the number of groups determined from the data. The proposed flexible shrinkage approach exhibits sound and robust performance compared to the popular portfolio selection alternatives.

JEL classification:

C13

D81

G11

Keywords:

Estimation risk and model uncertainty

k-means clustering

Model structure amount

Multivariate shrinkage estimator

1. Introduction

The destructive effect of estimation risk on portfolio performance has been recognized since Klein and Bawa (1976). Its influence can be reduced either by diminishing the risk and/or by correctly accounting for it (for an overview see Michaud, 1998; Brandt, 2005). The most popular methods for dealing with estimation risk are various types of factor models (Young and Lenk, 1998; Brandt et al., 2008), constrained portfolio weights (Frost and Savarino, 1988; Garlappi et al., 2007), different shrinkage estimators (Jorion, 1986; Ledoit and Wolf, 2003; Golosnoy and Okhrin, 2007), cluster analysis (Tola et al., 2008) or simple investment rules (DeMiguel et al., 2008). All these methods mitigate the impact of estimation risk by introducing some additional amount of structure in a portfolio selection model. Keeping the amount of structure fixed, however, may also be disadvantageous, because its misspecification can lead to suboptimal investment decisions. This introduces an additional model uncertainty into portfolio selection procedures (see, e.g., Wang, 2005). The rigid model framework may cause losses if it is no longer valid, so its adjustments should be possible. In this paper we propose a methodology for overcoming structural rigidity and better accounting for estimating risk. Our approach allows adjustment of structure by exploiting the quality of the newly incoming information. Thus we simultaneously control for estimation risk and model uncertainty in portfolio choice. Moreover, it requires no additional assumptions and is directly related to the investor's objective function.

A shrinkage approach, initiated by Stein (1955), suggests a sound framework for improving estimation procedure by using additional information. It can be related to the procedures of Bayesian analysis, however, the assumptions and interpretation are different. The shrinkage estimator for portfolio weights finds an optimal balance between the

* Corresponding author. Tel.: +41 31 6314792; fax: +41 31 6313783.

E-mail address: yarema.okhrin@vwi.unibe.ch (Y. Okhrin).

non-stochastic target weights and estimates from some portfolio procedure. It linearly combines the target, representing the long-term proportions, and estimated weights, representing the recent market information. Extending the approach of Longford (1999), Golosnoy and Okhrin (2007) suggest multivariate shrinkage for the mean-variance investor, providing an individual shrinkage factor for each asset. Each shrinkage factor measures an amount of estimation risk in the corresponding portfolio weight.

The approach of Golosnoy and Okhrin (2007) would be superior if the optimal shrinkage factors were known exactly. However, the shrinkage vector is a function of unknown distribution parameters, thus it is also subjected to estimation risk. Taking more distinct shrinkage coefficients increases the overall estimation risk. Hence, there exists model uncertainty about the optimal number of shrinkage coefficients. Consequently, it is necessary to adopt the shrinkage portfolio procedure to model uncertainty as well. This paper resolves this problem by characterizing the amount of model structure by a number of different shrinkage factors, determined by statistical considerations. Each shrinkage factor represents a group of assets with homogenous level of estimation risk.

We group the assets with similar shrinkage coefficients with a dynamic procedure determining the optimal number of groups from the data. This procedure is combined with the clustering algorithm to assign assets to the groups. We differentiate between a certain group composition, where each asset refers to one group only, and uncertain groupings, where assets belong to the groups with a certain probability. Based on the obtained groupings we provide the explicit solutions for the optimal shrinkage factors and corresponding portfolio proportions. Our approach can be seen as augmenting the initial asset space with artificial state variables, constructed by grouping the assets with similar amounts of estimation risk. State variables serve as a hedging opportunity, reflecting the useful sample information (cf. Brandt and Santa-Clara, 2006). This procedure improves estimation risk quantification by removing unfavorable structure rigidity and, consequently, accounting for model uncertainty.

The suggested methodology is validated using empirical data. We construct the portfolio of seven risky assets and compare the performance of our flexible shrinkage with the most popular benchmarks for different estimation windows, investment horizons and investor's risk aversion coefficients. The flexible shrinkage estimator selects either one or two groups both for certain and uncertain composition of groups. The empirical evidence shows that our approach provides sound and robust results for different performance criteria in the majority of the considered situations. Its flexibility allows the use of more market information when it is reliable, and reduction of estimation risk in more uncertain cases. The reported results justify the use of a time-varying amount of model structure for portfolio selection.

The remainder of the paper is organized as follows. Section 2 introduces the portfolio problem and multivariate shrinkage for the portfolio weights. Section 3 provides the flexible shrinkage estimator and the methodology for determining the optimal number of groups. Section 4 discusses the issues concerning implementation of our approach. The empirical illustration is presented in Section 5, while Section 6 concludes. The proofs of major results are given in the Appendix.

2. Shrinkage for portfolio weights

2.1. Portfolio problem

The considered investor allocates his wealth among k risky assets, where the k -dimensional return vector \mathbf{r} has expectation $E(\mathbf{r}) = \boldsymbol{\mu}$ and covariance matrix $\text{Cov}(\mathbf{r}) = \boldsymbol{\Sigma}$. He selects the optimal portfolio weights \mathbf{w} by maximizing the mean-variance objective function of portfolio return $r_p = \mathbf{w}'\mathbf{r}$

$$\max_{\mathbf{w}} EU = E(r_p) - \frac{\gamma}{2} V(r_p) \quad \text{w.r.t. } \mathbf{w}'\mathbf{1} = 1, \quad (1)$$

where $E(\cdot)$, $V(\cdot)$ stand for the expectation and variance of portfolio return, $\gamma > 0$ is the risk aversion coefficient, $\mathbf{1}$ is a vector of ones. The Markowitz weights \mathbf{u} are the solution of the task (1)

$$\mathbf{u} = \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}} + \frac{\mathbf{R}\boldsymbol{\mu}}{\gamma} \quad \text{where } \mathbf{R} = \boldsymbol{\Sigma}^{-1} - \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}'\boldsymbol{\Sigma}^{-1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}. \quad (2)$$

The unknown true moments $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ should be replaced by their estimators $(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$, thus the estimated weights $\hat{\mathbf{u}} = f(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$ are subjected to estimation risk. An increase in the number of assets k leads to an increase in total estimation error. Then the estimated Markowitz weights $\hat{\mathbf{u}}$ may contain a small amount of valuable information. Introducing additional structure into the portfolio selection model is a remedy for this estimation risk problem.

2.2. Multivariate shrinkage approach

The shrinkage procedure in portfolio theory allows one to quantify and reduce estimation risk (Jorion, 1986). It introduces structure into the portfolio problem using statistical considerations. The Markowitz weights are shrunk to a non-stochastic shrinkage target in order to find an optimal trade-off between the bias caused by the introduced structure and estimation risk. The shrinkage in portfolio selection is used to improve estimation of the means (Jorion, 1986) and of the covariance matrix of asset returns (Ledoit and Wolf, 2003). Golosnoy and Okhrin (2007) derive single and multivariate

shrinkage estimators directly for the portfolio weights. Multivariate shrinkage constructs a new portfolio weights vector $\hat{\mathbf{w}}$ by shrinking the estimated Markowitz weights $\hat{\mathbf{u}}$ to the target vector \mathbf{c} :

$$\hat{\mathbf{w}} = \mathbf{a} \circ \hat{\mathbf{u}} + (\mathbf{1} - \mathbf{a}) \circ \mathbf{c}, \quad (3)$$

where \circ is a Hadamard product. The vector \mathbf{a} denotes the shrinkage factors, reflecting the amount of estimation risk in the corresponding Markowitz weights. Then the optimization task (1) should be solved with respect to \mathbf{a} . Shrinkage factors close to one denote low estimation risk and strong belief in the Markowitz proportions $\hat{\mathbf{u}}$. Factors close to zero make the investor to give greater weight to the shrinkage target due to little useful information in the estimated weights $\hat{\mathbf{u}}$. The target weights \mathbf{c} can be seen as long-term robust investment alternatives, while the estimated vector $\hat{\mathbf{u}}$ captures the recent sample information about the optimal portfolio proportions.

The multivariate estimator $\mathbf{a} = (a_1, \dots, a_k)'$ provides all distinct shrinkage factors, while its single shrinkage counterpart $\mathbf{a} = a\mathbf{1}$ is a special case with equal shrinkages for all assets. The former is theoretically more advantageous than the latter, as it quantifies the amount of estimation risk for each single asset. However, the investor could fully exploit these advantages only if the shrinkage factors were known with certainty. Since the shrinkage factors should be estimated as well, an additional estimation risk arises. The estimation of \mathbf{a} causes an increase of the overall estimation error and explains why the single shrinkage approach may be preferred to the multivariate one. Reducing the number of distinct shrinkage coefficients decreases estimation risk but limits the model flexibility. A similar problem in the context of robust regression is considered by Wang et al. (2007). The unknown optimal number of the distinct shrinkage factors introduces model uncertainty. In this paper we propose a flexible shrinkage approach for determining the optimal number of distinct coefficients. It allows timely adjustment of model structure based on the newly incoming information. This allows a simultaneous accounting for estimation risk and model uncertainty.

3. Flexible shrinkage estimator

Now the flexible shrinkage estimator is introduced. First we derive the optimal shrinkage factors for arbitrary asset groupings. Then a dynamic rule for choosing the optimal number of groups with equal shrinkage factors is suggested. Finally, we interpret our approach from the portfolio theoretical viewpoint.

3.1. Composition of groups

Assume p different groups. The assets in each group have equal shrinkage factors. We differentiate between cases with known and uncertain group composition.

3.1.1. Certain composition

Each asset is assigned to one of p disjoint groups with $1 \leq p \leq k$. All assets in group i possess the same shrinkages a_i , with $1 \leq i \leq p$. Thus we unite the assets with similar amounts of estimation risk in their portfolio weights. To formalize the grouping let

$$\mathbf{a} = \mathbf{J}\tilde{\mathbf{a}}, \quad \text{where } \tilde{\mathbf{a}} = (a_1, \dots, a_p)' \text{ and } \mathbf{J}_{k \times p} = (J_{lm}), \quad \text{with } 1 \leq l \leq k, \quad 1 \leq m \leq p, \quad (4)$$

$$J_{lm} = \begin{cases} 1 & \text{if the } l\text{-asset belongs to the } m\text{-th group,} \\ 0 & \text{otherwise.} \end{cases}$$

The columns of \mathbf{J} are orthogonal since each asset can belong only to one group out of p , thus $\sum_l \sum_m J_{lm} = k$. It holds $\mathbf{J} = \mathbf{I}_k$ for all different and $\mathbf{J} = \mathbf{1}'$ for all equal shrinkage factors. The optimization task (1) is now solved with respect to the shrinkage coefficients $\tilde{\mathbf{a}}$ with a technical constraint $\mathbf{1} - \mathbf{c}'\mathbf{1} = v \geq 0$ for fixed and known matrix \mathbf{J} .

Proposition 1. *Let the structure of the shrinkage coefficients be introduced as in Eq. (4). Then the solution of the maximization problem (1) is given by*

$$\tilde{\mathbf{a}}^* = \gamma^{-1}(\mathbf{J}'\mathbf{Q})^{-1}\mathbf{J}'(\mathbf{q} - \pi(\mathbf{E}(\hat{\mathbf{u}}) - \mathbf{c})), \quad \pi = \frac{\mathbf{q}'\mathbf{Q}^{-1}(\mathbf{E}(\hat{\mathbf{u}}) - \mathbf{c}) - v\gamma}{(\mathbf{E}(\hat{\mathbf{u}}) - \mathbf{c})'\mathbf{Q}^{-1}(\mathbf{E}(\hat{\mathbf{u}}) - \mathbf{c})}$$

with $\mathbf{q} = (\boldsymbol{\mu} - \gamma\boldsymbol{\Sigma}\mathbf{c}) \circ (\mathbf{E}(\hat{\mathbf{u}}) - \mathbf{c})$, $\mathbf{Q} = \text{Cov}(\hat{\mathbf{u}}) \circ (\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}') + \boldsymbol{\Sigma} \circ (\mathbf{E}(\hat{\mathbf{u}}) - \mathbf{c})(\mathbf{E}(\hat{\mathbf{u}}) - \mathbf{c})'$.

Proposition 1 is proven in the Appendix. The sample moments of the optimal weights $\mathbf{E}(\hat{\mathbf{u}})$ and $\text{Cov}(\hat{\mathbf{u}})$ assuming iid log asset returns are derived by Okhrin and Schmid (2006). The composition of groups, given in matrix \mathbf{J} , is determined in a clustering procedure, described in Section 4.

3.1.2. Uncertain group composition

There arises an additional uncertainty about assigning k assets to p different groups. As earlier, we assume exactly p groups, but now the composition of groups remains uncertain. Hence the deterministic 0/1 matrix \mathbf{J} is replaced by the transition matrix $\mathbf{P} = (p_{ij})$, $i = 1, \dots, k$, $j = 1, \dots, p$. The entry $p_{ij} \geq 0$ denotes the probability of the i th asset belonging to the

j th group. Note that $\sum_{j=1}^p p_{ij} = 1$ for each i . The matrix \mathbf{P} should be substituted for \mathbf{J} in Eq. (4). The corresponding optimal shrinkages are given by

$$\tilde{\mathbf{a}}^* = \gamma^{-1}(\mathbf{P}'\mathbf{Q}\mathbf{P})^{-1}\mathbf{P}'(\mathbf{q} - \pi(\mathbf{E}(\hat{\mathbf{u}}) - \mathbf{c})), \quad (5)$$

where $(\mathbf{q}, \mathbf{Q}, \pi)$ are the same as in Proposition 1. The estimation of \mathbf{P} is discussed in Section 4.

Both matrices \mathbf{J} and \mathbf{P} contain a number of additional parameters. However, these parameters serve merely for averaging and smoothing of the shrinkage factors $\tilde{\mathbf{a}}$, e.g., the component $\mathbf{J}\mathbf{Q}\mathbf{J}$ in Proposition 1 mitigates the impact of fluctuation in the matrix \mathbf{Q} on the shrinkage vector $\tilde{\mathbf{a}}$.

3.2. Determining the number of groups

Now we describe the dynamic procedure for determining the number of groups, i.e., the number of distinct shrinkage factors. The matrix \mathbf{J}_t , introduced above, is used for calculating the optimal shrinkages $\tilde{\mathbf{a}}_t^*$ at each t . Our flexible shrinkage approach makes the number of different groups not constant, but time-varying. For the number of groups p_t it holds that $p_t \in \mathcal{M} = \{1, 2, \dots, k\}$. Let $p_t \in \mathcal{M}^{(m)}$, where $\mathcal{M}^{(m)} \subseteq \mathcal{M}$ and $\mathcal{M}^{(m)}$ consists of m elements. The number of groups is chosen based on the utility, realized in the previous holding period of length s . The optimal number of groups \hat{p}_{t-s} , realized in the period $(t-s, t]$, can be ex-post determined at t from maximizing the realized utility RU

$$\max_{p_{t-s} \in \mathcal{M}^{(m)}} RU(\hat{\mathbf{w}}_{t-s}) \quad (6)$$

with

$$\hat{\mathbf{w}}_{t-s} = \hat{\mathbf{a}}_{t-s}^*(p_{t-s}) \circ \hat{\mathbf{u}}_{t-s} + (\mathbf{1} - \hat{\mathbf{a}}_{t-s}^*(p_{t-s})) \circ \mathbf{c}_{t-s}.$$

The utility $RU(\hat{\mathbf{w}}_{t-s})$ is calculated from intraperiod portfolio returns rp_{t-h} by

$$RU_{t-s,t} = \sum_{h=0}^{s-1} rp_{t-h} - \frac{\gamma}{2} \sum_{h=0}^{s-1} rp_{t-h}^2, \quad (7)$$

where $rp_{t-h} = \hat{\mathbf{w}}_{t-s}' \mathbf{r}_{t-h}$ for $h = 0, \dots, s-1$. Then we assume $p_t = \hat{p}_{t-s}$, i.e., the number of groups optimal in the previous period $(t-s, t]$ stays optimal for the coming period $(t, t+s]$, as well. Our procedure resembles the method of Aiolfi and Timmermann (2006) and can be seen as a frequentistic analogy to the Bayesian approach of Raftery and Dean (2006), where the variable selection problem is reformulated as the model selection task. However, contrary to Bayesian inferences, we do not require prior assumptions for implementing our approach. The number of groups can be similarly determined in case with uncertain group composition, characterized by matrix \mathbf{P} . Note that for computational convenience it is useful to restrict the set $\mathcal{M}^{(m)}$ to the most important alternatives.

3.3. Measuring portfolio performance

The performance is measured using a realized utility approach based on the intraperiod portfolio returns. We transform the time scale and denote the end points of τ th investment period by τ with $\tau = [1, \dots, T]$. The realized utilities for intraperiod portfolio returns $rp_{\tau-1+m/s}$ with $m = 1, \dots, s$ are calculated by (7) for each period $(\tau-1, \tau]$ corresponding to $(t-s, t]$. The approaches are compared using the average realized utilities

$$ARU = T^{-1} \sum_{\tau=1}^T RU_{\tau-1;\tau}, \quad (8)$$

with $\tau = 1, \dots, T$ denoting investment periods and $m = 1, \dots, s$ time points of the intraperiod returns. Additionally, we use the realized Sharpe ratio and the average realized Sharpe ratios as a conventional performance measure:

$$SR_{\tau} = \frac{\sum_{m=1}^s rp_{\tau-1+m/s}}{\left(\sum_{m=1}^s rp_{\tau-1+m/s}^2\right)^{1/2}}, \quad ASR = T^{-1} \sum_{\tau=1}^T SR_{\tau}. \quad (9)$$

3.4. Interpreting flexible shrinkage approach

The flexible shrinkage methodology can be interpreted from the portfolio theoretical point of view. The shrinkage portfolio return $r_p = \hat{\mathbf{w}}' \mathbf{r}$ can be alternatively written as

$$r_p = \hat{\mathbf{w}}' \mathbf{r} = \mathbf{a}'((\hat{\mathbf{u}} - \mathbf{c}) \circ \mathbf{r}) + \mathbf{c}' \mathbf{r} = \tilde{\mathbf{a}}' \mathbf{J}'((\hat{\mathbf{u}} - \mathbf{c}) \circ \mathbf{r}) + \mathbf{c}' \mathbf{r}. \quad (10)$$

The original space of k returns is augmented by p state variables with returns $\mathbf{J}'((\hat{\mathbf{u}} - \mathbf{c}) \circ \mathbf{r})$. The shrinkage factors $\tilde{\mathbf{a}}$ serve as the weights for these state variables. Then the portfolio return r_p can be seen as a sum of returns on a target $\mathbf{c}' \mathbf{r}$ and state variables $\tilde{\mathbf{a}}' \mathbf{J}'((\hat{\mathbf{u}} - \mathbf{c}) \circ \mathbf{r})$ portfolios. The target portfolio return $\mathbf{c}' \mathbf{r}$ is the return on a benchmark portfolio, while the state

variables portfolio return $\tilde{\mathbf{a}}' \mathbf{J}' ((\hat{\mathbf{u}} - \mathbf{c}) \circ \mathbf{r})$ could be interpreted as a hedge fund, constructed based on sample information (Brandt and Santa-Clara (2006)). If $\mathbf{c}' \mathbf{1} = 1$, the hedge fund weights sums to zero $\tilde{\mathbf{a}}' \mathbf{1} = 0$.

State variables average information from k Markowitz weights $(\hat{\mathbf{u}} - \mathbf{c}) \circ \mathbf{r}$ with the matrix $\mathbf{J}_{k \times p}$. The matrix \mathbf{J} (or \mathbf{P}) describes the composition of asset groups with similar amounts of estimation risk. Such grouping of original assets is equivalent to summing together their portfolio fractions. This reduces estimation risk for assets within each group. The time varying number of state variables p is determined from the data as the optimal number of groups. High homogeneity in estimation risk reduces the number of state variables, accordingly high diversity in estimation risk increases their number. The number of state variables represents the conditional amount of model uncertainty. Hence the state variables serve for capturing useful sample information from the Markowitz procedure. Thus we amend conventional approaches for mitigating estimation risk by introducing flexibility in the amount of model structure.

3.5. Extensions

The suggested methodology can be extended in various directions. We provide the flexible shrinkage estimator for the class of unconstrained portfolio weights. It may also be implemented, however, for the important class of constrained portfolio weights. The moments of the constrained portfolio weights can be obtained under the normality assumption from a Monte Carlo simulation study. If the normality assumption is relaxed, the moments required in Proposition 1 could be calculated using the non-parameter bootstrap procedure. These moments could be used for further asset groupings and calculating optimal shrinkage factors.

The extension of our approach for the multiperiod horizon is possible as well. The multiperiod problem with flexible shrinkages and the portfolio return for the H -period problem are given by

$$\max E_{\tau}(rp_{\tau \rightarrow \tau+H}) - \frac{\gamma}{2} V_{\tau}(rp_{\tau \rightarrow \tau+H}), \quad rp_{\tau \rightarrow \tau+H} = \prod_{h=1}^H \mathbf{w}'_{t+h-1} (\mathbf{1} + \mathbf{r}_{t+h}). \quad (11)$$

The optimal portfolio weights for the multiperiod mean-variance portfolio problem without shrinkage are derived by Li and Ng (2000). Using their results, the moments of the estimated portfolio weights can be obtained from Monte-Carlo simulations. Note that the mean-variance objective function is not time separable, making the backward induction procedure more complicated than in the power utility case. The portfolio return for the shrinkage estimator can be written by

$$\mathbf{w}'_{t+h-1} (\mathbf{1} + \mathbf{r}_{t+h}) = \mathbf{a}'_{t+h-1} [(\hat{\mathbf{u}}_{t+h-1} - \mathbf{c}_{t+h-1}) \circ (\mathbf{1} + \mathbf{r}_{t+h})] + \mathbf{c}'_{t+h-1} (\mathbf{1} + \mathbf{r}_{t+h}) \quad (12)$$

for all h values. It has the same structure as the multiperiod problem without shrinkage. Therefore, the same technique may be used for estimating \mathbf{a}_{t+h-1} for $h = 1, \dots, H$. The matrices \mathbf{J}_{t+h-1} (or \mathbf{P}_{t+h-1}) can again be determined from the clustering procedure. This approach requires the assumption about the optimal number of groups over the whole horizon H . Unfortunately, it does not allow the elegant time separation for weights, proposed by Brandt and Santa-Clara (2006). This significantly complicates its practical implementation in the multiperiod setting.

4. Implementation issues

For implementing our approach, we provide a statement about the first two moments of the estimated Markowitz weights. Then we describe the clustering procedure for asset grouping.

4.1. Estimating optimal portfolio weights

Okhrin and Schmid (2006) derive the moments of the optimal portfolio weights for equal sampling frequency for the means and covariance matrix of portfolio returns. However, Merton (1980) has already shown that the estimation precision for the means can be improved by taking a longer historical period, while the covariance estimator is ameliorated by increasing the sampling frequency. To implement Merton's idea we estimate the unknown parameters $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ using historical returns sampled at different frequencies and over different horizons. Let the superscript κ refer to quantities based on the frequency κ . The daily frequency is chosen as a base one with $\kappa = 1$ and no superscript. The distribution parameters are estimated by

$$\hat{\boldsymbol{\mu}}_t^{(\kappa)} = \frac{1}{n_1} \sum_{i=0}^{n_1-1} \mathbf{X}_{t-ik}^{(\kappa)}, \quad \hat{\boldsymbol{\Sigma}}_t = \frac{1}{n_2 - 1} \sum_{i=0}^{n_2-1} (\mathbf{X}_{t-i} - \hat{\boldsymbol{\mu}})(\mathbf{X}_{t-i} - \hat{\boldsymbol{\mu}})', \quad (13)$$

where $\mathbf{X}_{t-ik}^{(\kappa)} = \sum_{j=0}^{\kappa-1} \mathbf{X}_{t-ik-j}$ and $\hat{\boldsymbol{\mu}} = (1/n_2) \sum_{i=0}^{n_2-1} \mathbf{X}_{t-i}$. The estimator of the portfolio weights is computed by inserting the introduced estimators $\hat{\mathbf{u}} = \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1} / \mathbf{1}' \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1} + \hat{\mathbf{R}} \hat{\boldsymbol{\mu}}^{(\kappa)} / (\gamma \kappa)$. Assuming $\mathbf{X}_i \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ implies that $\mathbf{X}_i^{(\kappa)} \sim \mathcal{N}_k(\boldsymbol{\mu}^{(\kappa)}, \boldsymbol{\Sigma}^{(\kappa)})$ with $\boldsymbol{\mu}^{(\kappa)} = \kappa \boldsymbol{\mu}, \boldsymbol{\Sigma}^{(\kappa)} = \kappa \boldsymbol{\Sigma}$.

Proposition 2. Let $\mathbf{X}_{t-n_1\kappa}, \dots, \mathbf{X}_t$ be independent random vectors with $\mathbf{X}_i \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ for $i = t - n_1\kappa, \dots, t$. Assuming that $n_2 > k + 2$ it follows that

$$\begin{aligned} E(\hat{\mathbf{u}}) &= \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}} + \frac{n_2 - 1}{n_2 - k - 1} \gamma^{-1} \mathbf{R}\boldsymbol{\mu}. \\ \text{Cov}(\hat{\mathbf{u}}) &= \frac{\mathbf{R}/\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}{n_2 - k - 1} + \frac{c_1}{\gamma^2} \mathbf{R}\boldsymbol{\mu}\boldsymbol{\mu}'\mathbf{R} + \frac{c_2}{\gamma^2} \boldsymbol{\mu}'\mathbf{R}\boldsymbol{\mu}\mathbf{R} \\ &\quad + \frac{\gamma^{-2}}{\kappa n_1} \left(c_1 + c_2(k-1) + \frac{(n_2 - 1)^2}{(n_2 - k - 1)^2} \right) \mathbf{R}, \end{aligned}$$

where

$$c_1 = \frac{(n_2 - 1)^2(n_2 - k + 1)}{(n_2 - k)(n_2 - k - 1)^2(n_2 - k - 3)}, \quad c_2 = \frac{(n_2 - 1)^2}{(n_2 - k)(n_2 - k - 1)(n_2 - k - 3)}.$$

This proposition allows estimation of Markowitz weights $\hat{\mathbf{u}}$ using the longer time period for the mean and the shorter time period for the covariance matrix.

4.2. Determining composition of groups

The matrices \mathbf{J} and \mathbf{P} determine the composition of groups with equal shrinkage factors. Cluster analysis suggests methods for building such groups of assets. The most commonly used approaches are hierarchical and k -means clusterings, see [Witten and Frank \(2005\)](#). The hierarchical method is a deterministic procedure creating a tree structure of assets from the data by either successive agglomeration of k objects into groups or by separation of a single group into subgroups. Since we determine the number of groups in advance, we adopt for our purposes the k -means procedure extended by the expectation maximization (EM) algorithm. Our clustering algorithm for estimating the matrix \mathbf{J} consists of the following steps:

1. Randomly assign k different shrinkage factors into p groups.
2. Using the k -means method, repeatedly reassign the shrinkage factors to other groups to achieve the smallest sum of distances from the shrinkage factors to their cluster centers.
3. Consider all combinations of assigning k elements into p groups for small p and k . Otherwise, repeat steps 1 and 2 until the number of repetitions is sufficient to consider the obtained minimal sum of distances close to the global minimum.
4. Take the clustering scheme with the overall smallest sum of distances between the shrinkage factors and their cluster centers as the final clustering solution.

The Euclidian norm measures distances between shrinkage factors and cluster centers, computed with the arithmetic mean. Note that each group has at least one asset.

Estimation of the matrix \mathbf{P} is more demanding computationally. First the vector of distribution parameters $\boldsymbol{\theta} = (\boldsymbol{\mu}, \text{vech}(\boldsymbol{\Sigma}))'$ is estimated from the data and used as true for generating Monte-Carlo returns. The vectors $\hat{\boldsymbol{\theta}}_\ell$ and $\hat{\mathbf{a}}_\ell$ with $\ell = 1, \dots, M$ are estimated for each of M replications. Then the k -dimensional vector $\hat{\mathbf{a}}_\ell$ is clustered into p groups with the algorithm described above. Set $d_{\ell,ij} = 1$ if the i -th shrinkage coefficient in the ℓ -th replication gets into the j -th group and $d_{\ell,ij} = 0$ otherwise. Then the probabilities are estimated by $\hat{p}_{ij} = (1/M) \sum_{\ell=1}^M d_{\ell,ij}$, forming the matrix $\hat{\mathbf{P}}$. Note that the estimated shrinkage vector $\hat{\mathbf{a}}^* = \hat{\mathbf{P}}\hat{\mathbf{a}}^*$ may consist of k different variables. However, this does not contradict the idea of introducing structure by p shrinkage groups.

5. Empirical evidence

5.1. Methodological issues

The analyzed portfolio consists of $k = 7$ assets: the MSCI country indices for Canada, France, Germany, Japan, UK, US and gold as an additional risky asset. The choice of $k \leq 10$ is typical for portfolio selection, see [Fleming et al. \(2001\)](#). The daily data in US\$ are taken from `DataStream` for the period 01 July 1980–01 January 2006. The monthly data for 01 January 1970–01 July 1980 are used for the pre-run estimation. Following [Härdle et al. \(2003\)](#), we take $n = 150$ daily returns for the covariance matrix, and $n = 150$ monthly returns for the mean-vector estimation. Other choices of n deliver similar results. The chosen investment horizons are {50, 100, 150} days. The risk aversions are $\gamma = \{5, 10, 25, 50\}$. The values $\gamma = 5, 10$ imply greater influence of portfolio return on the objective function (1). $\gamma = 50$ causes greater influence of portfolio variance, while $\gamma = 25$ leads to a roughly equal impact of portfolio return and variance on investment decisions.

5.2. Empirical results

First we consider the family of shrinkage estimators. The market capitalizations are chosen as targets with the capitalization of gold taken to be zero. The estimation is conducted as described in Section 4. We consider the multivariate, single and two-group shrinkages with both certain and uncertain composition of groups. The certain composition of groups presumes that each asset refers to one group only at each point of time. The composition of groups is determined via the clustering procedure introduced in Section 4.2. These approaches are compared to the flexible shrinkage with $\mathcal{M}^{(2)} = \{1, 2\}$, i.e., switching between one and two groups with certain and uncertain group composition. For $p = 2$ and $k = 7$ it is sufficient to make 10^3 repetitions in step 3 of the clustering algorithm, as doing more does not alter significantly the groupings obtained. The shrinkage procedures are used for improving the unconstrained Markowitz or global minimum variance portfolio (GMVP) compositions. The average realized utilities and average realized Sharpe ratios for the shrinkage-based approaches are reported in Table 1. The performance measures are normalized in order to enable comparison of different investment horizons.

Flexible shrinkage increases the realized utility compared to single- and multivariate shrinkages of the unconstrained Markowitz weights with capitalization proportions as the target. Dynamic shrinkage with uncertain groups is the overall best strategy for this class of estimators. However, the Markowitz weights give a rather poor performance compared to the other shrinkage-based approaches. Thus we examine the flexible shrinkage to the GMVP weights. In this case the dynamic composition of groups is also the best choice in the majority of cases. There is no significant difference in certain or uncertain group composition. The superiority of dynamic shrinkage is even more pronounced when using the average realized Sharpe ratio as a performance criterion. Hence, introducing flexibility into shrinkage procedure is justified, because it improves portfolio performance.

Table 1
The performance of shrinkage approaches for Markowitz and GMVP weights

γ	s	Mark	Multi	$p = 1$	Certain groups			Uncertain groups			GMV	Multi	$p = 1$	Certain groups		Uncertain groups	
					$p = 2$	Flex	$p = 2$	Flex	$p = 2$	Flex				$p = 2$	Flex		
<i>Panel A. Average realized utilities ($\times 1000$)</i>																	
5	50	-1.0782	-0.3535	-0.2554	-0.3394	-0.2870	-0.3194	-0.2803	0.1371	0.1369	0.1416	0.1380	0.1421	0.1392	0.1425		
5	100	-1.0443	-0.3235	-0.2293	-0.3121	-0.2731	-0.3003	-0.2419	0.1259	0.1266	0.1303	0.1267	0.1286	0.1300	0.1293		
5	150	-1.2058	-0.4585	-0.3416	-0.4227	-0.3510	-0.4054	-0.3484	0.1328	0.1322	0.1368	0.1350	0.1362	0.1346	0.1348		
10	50	-0.5750	-0.2211	-0.1878	-0.2022	-0.2089	-0.2024	-0.1863	0.0501	0.0508	0.0548	0.0516	0.0553	0.0527	0.0545		
10	100	-0.5712	-0.2141	-0.1848	-0.2290	-0.2058	-0.2043	-0.1856	0.0344	0.0359	0.0389	0.0359	0.0370	0.0387	0.0370		
10	150	-0.6349	-0.2674	-0.2243	-0.2702	-0.2378	-0.2377	-0.2085	0.0384	0.0387	0.0427	0.0412	0.0433	0.0406	0.0428		
25	50	-0.4818	-0.3741	-0.3905	-0.3792	-0.3849	-0.3865	-0.3856	-0.2109	-0.2076	-0.2053	-0.2078	-0.2052	-0.2069	-0.2047		
25	100	-0.5072	-0.3861	-0.3980	-0.3812	-0.3877	-0.3992	-0.4042	-0.2404	-0.2362	-0.2354	-0.2366	-0.2338	-0.2351	-0.2343		
25	150	-0.5189	-0.4034	-0.4096	-0.3998	-0.3987	-0.4084	-0.4060	-0.2448	-0.2419	-0.2396	-0.2400	-0.2416	-0.2399	-0.2382		
50	50	-0.7988	-0.7624	-0.7827	-0.7606	-0.7677	-0.7698	-0.7697	-0.6458	-0.6382	-0.6390	-0.6401	-0.6385	-0.6408	-0.6396		
50	100	-0.8521	-0.7975	-0.8263	-0.7943	-0.8041	-0.8209	-0.8181	-0.6983	-0.6897	-0.6926	-0.6907	-0.6879	-0.6912	-0.6908		
50	150	-0.8579	-0.7984	-0.8266	-0.8026	-0.8150	-0.8139	-0.8119	-0.7169	-0.7096	-0.7102	-0.7087	-0.7112	-0.7074	-0.7055		
<i>Panel B. Average realized Sharpe ratios</i>																	
5	50	1.395	1.787	1.922	1.783	1.925	1.844	1.936	3.307	3.307	3.374	3.322	3.383	3.339	3.387		
5	100	2.584	3.046	3.225	3.079	3.127	3.055	3.314	4.424	4.441	4.514	4.442	4.479	4.508	4.491		
5	150	2.192	2.502	2.796	2.656	2.901	2.706	2.925	5.575	5.564	5.674	5.631	5.659	5.624	5.625		
10	50	2.008	2.583	2.658	2.673	2.593	2.698	2.749	3.307	3.307	3.374	3.322	3.379	3.341	3.369		
10	100	3.333	4.022	4.099	3.870	3.947	4.088	4.129	4.424	4.441	4.514	4.442	4.470	4.507	4.480		
10	150	3.269	4.000	4.254	4.006	4.160	4.378	4.565	5.575	5.564	5.674	5.631	5.687	5.610	5.665		
25	50	2.909	3.306	3.202	3.298	3.218	3.273	3.258	3.307	3.307	3.374	3.322	3.363	3.348	3.375		
25	100	4.334	4.967	4.820	5.037	4.961	4.835	4.727	4.424	4.441	4.514	4.442	4.521	4.503	4.518		
25	150	4.881	5.583	5.517	5.645	5.643	5.616	5.461	5.575	5.564	5.674	5.631	5.590	5.616	5.643		
50	50	3.247	3.387	3.380	3.450	3.400	3.461	3.431	3.307	3.307	3.374	3.322	3.354	3.340	3.336		
50	100	4.608	4.856	4.830	4.891	4.829	4.880	4.848	4.424	4.441	4.514	4.442	4.531	4.509	4.510		
50	150	5.479	5.894	5.830	5.927	5.733	5.817	5.915	5.575	5.564	5.674	5.631	5.566	5.622	5.644		

Comparison of shrinkage approaches for average realized utility (Panel A) and Sharpe ratio (Panel B) performance criteria for different risk aversions γ and holding periods s . We consider the classical Markowitz (Mark, left) and GMVP (right) weights; multivariate (mult), two ($p = 2$) and one ($p = 1$) group shrinkages as well as flexible shrinkages (flex) to capitalization target. The estimation period is $n = 150$ days.

Next we compare the flexible shrinkage to the approaches widely used for portfolio selection purposes. The benchmarks are the classical Markowitz, no short sales Markowitz, the sample as well as Ledoit and Wolf GMVP, the portfolio based on the one-factor model. The non-stochastic equal weight and capitalization strategies and the strategy combining the estimators of Ledoit and Wolf for the covariance matrix with those of Jorion for the means are considered as well. Tables 2 and 3 compare the flexible shrinkage with these benchmarks. We report both the average realized utilities and average realized Sharpe ratios in Table 2, while Table 3 provides the average portfolio returns and variances. Note that the average realized Sharpe ratios cannot be simply obtained from the average realized portfolio returns and variances. Flexible shrinkage for Markowitz weights provides much worse results than for the GMVP ones. Flexible GMVP shrinkage outperforms all benchmarks in the majority of cases in terms of the average realized utility. Additionally, it provides a robust performance for different holding periods s and risk aversions γ . The Sharpe ratio criterion supports the evidence in favor of shrinking the GMVP weights. We find no significant difference between flexible shrinkage with certain and uncertain groups. This may happen because the situations with large unimodal probabilities for all assets in \mathbf{P} provide compositions, similar to the case with 0/1 matrix \mathbf{J} . We recommend a certain grouping with matrix \mathbf{J} , since it requires much less computational effort.

The graphical illustration reveals an intuition behind our methodology. Fig. 1 presents the optimal number of groups for different risk aversions and $\mathcal{M}^{(2)} = \{1, 2\}$. We observe time periods with sequences of one (or two) optimal groups. These clusters justify taking the optimal number of groups from the previous period. The length of these periods increases with the increase of the risk aversion γ . Fig. 2 shows shrinkage factors for the flexible estimator. Single group denotes an equal amount of estimation risk for all assets, while two groups emerge when an asset subgroup exhibits a higher amount of estimation risk. Flexible shrinkage allows for any asset to change its group affiliation from one period to another.

Table 2

The performance of flexible shrinkage approaches vs. benchmark strategies: average realized utilities and Sharpe ratios

γ	s	Benchmark strategies								Certain groups		Uncertain groups	
		EW	Caps	Mark	Mark-c	GMV	J&LW	LW	1-factor	Flex1	Flex2	Flex1	Flex2
<i>Panel A. Average realized utilities ($\times 1000$)</i>													
5	50	0.1129	0.1231	-1.0782	-0.0517	0.1371	0.0943	0.1233	0.0709	-0.2870	0.1421	-0.2803	0.1425
5	100	0.1129	0.1231	-1.0443	-0.0413	0.1259	0.1138	0.1119	0.0945	-0.2731	0.1286	-0.2419	0.1293
5	150	0.1129	0.1245	-1.2058	-0.0722	0.1328	0.0446	0.1156	0.0465	-0.3510	0.1362	-0.3484	0.1348
10	50	-0.0382	-0.0498	-0.575	-0.161	0.0501	0.0179	0.0358	-0.0118	-0.2089	0.0553	-0.1863	0.0545
10	100	-0.0382	-0.0496	-0.5712	-0.1451	0.0344	0.0166	0.0203	-0.0032	-0.2058	0.037	-0.1856	0.0370
10	150	-0.0382	-0.0486	-0.6349	-0.1792	0.0384	-0.0159	0.0203	-0.024	-0.2378	0.0433	-0.2085	0.0428
25	50	-0.4915	-0.5685	-0.4818	-0.4635	-0.2109	-0.2377	-0.2265	-0.2988	-0.3849	-0.2052	-0.3856	-0.2047
25	100	-0.4915	-0.5677	-0.5072	-0.4572	-0.2404	-0.2615	-0.2546	-0.3103	-0.3877	-0.2338	-0.4042	-0.2343
25	150	-0.4915	-0.5679	-0.5189	-0.4786	-0.2448	-0.2809	-0.2657	-0.3236	-0.3987	-0.2416	-0.406	-0.2382
50	50	-1.2470	-1.4330	-0.7988	-0.8980	-0.6458	-0.6727	-0.6637	-0.7901	-0.7677	-0.6385	-0.7697	-0.6396
50	100	-1.2470	-1.4313	-0.8521	-0.9079	-0.6983	-0.7207	-0.7126	-0.8266	-0.8041	-0.6879	-0.8181	-0.6908
50	150	-1.2470	-1.4334	-0.8579	-0.9156	-0.7169	-0.7505	-0.7423	-0.8524	-0.815	-0.7112	-0.8119	-0.7055
<i>Panel B. Average realized Sharpe ratios</i>													
5	50	2.956	3.098	1.395	1.767	3.307	2.746	3.101	2.522	1.925	3.383	1.936	3.387
5	100	4.180	4.381	2.584	2.660	4.424	4.226	4.139	3.970	3.127	4.479	3.314	4.491
5	150	5.119	5.393	2.192	2.838	5.575	3.877	5.149	3.998	2.901	5.659	2.925	5.625
10	50	2.956	3.098	2.008	2.527	3.307	3.060	3.101	2.800	2.593	3.379	2.749	3.369
10	100	4.180	4.381	3.333	3.832	4.424	4.435	4.139	4.233	3.947	4.470	4.129	4.480
10	150	5.119	5.393	3.269	4.223	5.575	4.702	5.149	4.748	4.160	5.687	4.565	5.665
25	50	2.956	3.098	2.909	3.198	3.307	3.137	3.101	2.850	3.218	3.363	3.258	3.375
25	100	4.180	4.381	4.334	4.864	4.424	4.345	4.139	4.174	4.961	4.521	4.727	4.518
25	150	5.119	5.393	4.881	5.653	5.575	5.057	5.149	5.002	5.643	5.590	5.461	5.643
50	50	2.956	3.098	3.247	3.385	3.307	3.129	3.101	2.833	3.400	3.354	3.431	3.336
50	100	4.180	4.381	4.608	5.087	4.424	4.258	4.139	4.098	4.829	4.531	4.848	4.510
50	150	5.119	5.393	5.479	5.994	5.575	5.120	5.149	5.023	5.733	5.566	5.915	5.644

Flexible shrinkage strategies for Markowitz (flex1) and GMVP (flex2) weights for certain and uncertain groups compared to benchmarks for average realized utility (Panel A) and Sharpe ratio (Panel B) performance criteria for different risk aversions γ and holding periods s . The benchmarks are equal weights (EW), capitalizations (caps), Markowitz (Mark), no short sales Markowitz (Mark-c), GMVP weights, Jorion and Ledoit/Wolf (J&LW), Ledoit/Wolf (LW), one factor model (1-factor). The estimation period is $n = 150$ days.

Table 3

The performance of flexible shrinkage approaches vs. benchmark strategies: average portfolio returns and variances

γ	s	Benchmark strategies								Certain groups		Uncertain groups	
		EW	Caps	Mark	Mark-c	GMV	J&LW	LW	1-factor	Flex1	Flex2	Flex1	Flex2
<i>Panel A. Average portfolio return ($\times 1000$)</i>													
5	50	0.2640	0.2960	0.3883	0.2062	0.2241	0.2446	0.2107	0.2341	0.3542	0.2287	0.3505	0.2291
5	100	0.2640	0.2959	0.5262	0.2216	0.2175	0.2799	0.2035	0.2708	0.4244	0.2198	0.4236	0.2217
5	150	0.2640	0.2976	0.3645	0.1944	0.2272	0.2077	0.2109	0.2224	0.3138	0.2304	0.3329	0.2287
10	50	0.2640	0.2960	0.3062	0.2693	0.2241	0.2276	0.2107	0.2182	0.3007	0.2284	0.3160	0.2275
10	100	0.2640	0.2959	0.3719	0.2906	0.2175	0.2417	0.2035	0.2397	0.3328	0.2192	0.3446	0.2208
10	150	0.2640	0.2976	0.2959	0.2636	0.2272	0.2093	0.2109	0.2200	0.2811	0.2314	0.3080	0.2310
25	50	0.2640	0.2960	0.2569	0.2842	0.2241	0.2175	0.2107	0.2086	0.2670	0.2273	0.2704	0.2281
25	100	0.2640	0.2959	0.2793	0.3095	0.2175	0.2188	0.2035	0.2211	0.2987	0.2217	0.2910	0.2213
25	150	0.2640	0.2976	0.2547	0.2950	0.2272	0.2102	0.2109	0.2185	0.2748	0.2270	0.2676	0.2291
50	50	0.2640	0.2960	0.2405	0.2653	0.2241	0.2141	0.2107	0.2054	0.2492	0.2267	0.2522	0.2260
50	100	0.2640	0.2959	0.2484	0.2855	0.2175	0.2112	0.2035	0.2149	0.2554	0.2221	0.2609	0.2212
50	150	0.2640	0.2976	0.2410	0.2743	0.2272	0.2106	0.2109	0.2180	0.2480	0.2261	0.2567	0.2290
<i>Panel B. Average portfolio variance ($\times 1000$)</i>													
5	50	0.0604	0.0692	0.5866	0.1032	0.0348	0.0601	0.0350	0.0653	0.2565	0.0346	0.2527	0.0346
5	100	0.0604	0.0691	0.6282	0.1052	0.0366	0.0664	0.0366	0.0705	0.2790	0.0365	0.2727	0.0367
5	150	0.0604	0.0692	0.6281	0.1067	0.0378	0.0652	0.0381	0.0704	0.2659	0.0377	0.2684	0.0377
10	50	0.0604	0.0692	0.1762	0.0861	0.0348	0.0419	0.0350	0.0460	0.1019	0.0346	0.1010	0.0347
10	100	0.0604	0.0691	0.1886	0.0871	0.0366	0.0450	0.0366	0.0486	0.1077	0.0364	0.1064	0.0367
10	150	0.0604	0.0692	0.1862	0.0886	0.0378	0.0450	0.0381	0.0488	0.1038	0.0376	0.1039	0.0374
25	50	0.0604	0.0692	0.0591	0.0598	0.0348	0.0364	0.0350	0.0406	0.0522	0.0346	0.0524	0.0346
25	100	0.0604	0.0691	0.0629	0.0613	0.0366	0.0384	0.0366	0.0425	0.0549	0.0364	0.0552	0.0365
25	150	0.0604	0.0692	0.0619	0.0619	0.0378	0.0393	0.0381	0.0434	0.0539	0.0375	0.0540	0.0374
50	50	0.0604	0.0692	0.0416	0.0465	0.0348	0.0355	0.0350	0.0398	0.0407	0.0346	0.0409	0.0345
50	100	0.0604	0.0691	0.0440	0.0477	0.0366	0.0373	0.0366	0.0417	0.0424	0.0364	0.0431	0.0366
50	150	0.0604	0.0692	0.0440	0.0476	0.0378	0.0384	0.0381	0.0428	0.0425	0.0375	0.0427	0.0374

Flexible shrinkage strategies for Markowitz (flex1) and GMVP (flex2) weights for certain and uncertain groups compared to benchmarks for average portfolio return (Panel A) and average portfolio variance (Panel B) performance criteria for different risk aversions γ and holding periods s . The benchmarks are equal weights (EW), capitalizations (caps), Markowitz (Mark), no short sales Markowitz (Mark-c), GMVP weights, Jorion and Ledoit/Wolf (J&LW), Ledoit/Wolf (LW), one factor model (1-factor). The estimation period is $n = 150$ days.

This procedure captures time-varying diversity of estimation risk and accounts for model uncertainty. Relaxing model structure rigidity is the main advantage of our approach.

5.3. Discussion

The introduced flexible shrinkage allows timely portfolio adjustments for the changing market situation caused by the incoming informational shocks. It suggests to alter portfolio positions which are perceived as uncertain at the momentum. Such situation may arise only when a subgroup of risky assets is influenced by some striking news. This justifies the two group approach with respect to the amount of estimation risk. The ability to account for new conditions is the main advantage of the flexible shrinkage approach.

Our shrinkage estimator suggests an efficient trade-off between the long-term financial alternatives, represented by the shrinkage target, and the recent sample information, represented by some estimated portfolio weights, e.g., Markowitz, GMVP or some other proportions. Of course, the moments of these estimated weights are required for the implementation of the approach. The time-varying number of different shrinkage coefficient captures estimation risk, and, simultaneously, accounts for model uncertainty by exploiting the newly arriving information. The whole procedure aims to determine the optimal amount of structure in portfolio choice in the statistically efficient way.

A shrinkage approach in portfolio selection can be based on Bayesian inferences as well (cf. Jorion, 1986; Wang, 2005). A single shrinkage procedure for portfolio weights with a stochastic target is considered by Okhrin and Schmid (2008). However, the additional assumptions about the priors are required there. The flexible shrinkage does not require any prior

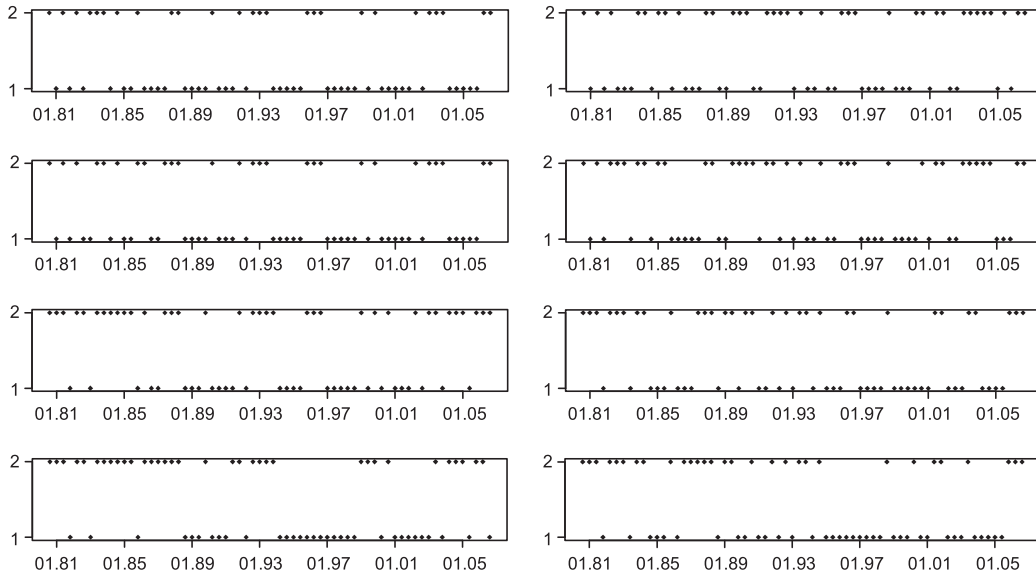


Fig. 1. The optimal number of groups $p_t \in \{1, 2\}$ for flexible shrinkage as a function of time (MM.YY). Optimal number of groups for flexible shrinkage with different risk aversion coefficient $\gamma = \{5, 10, 25, 50\}$, from top to bottom. The cases with certain groups are plotted left, while the cases with uncertain groups are right. The estimation period is $n = 150$, the holding period is $s = 100$ days.

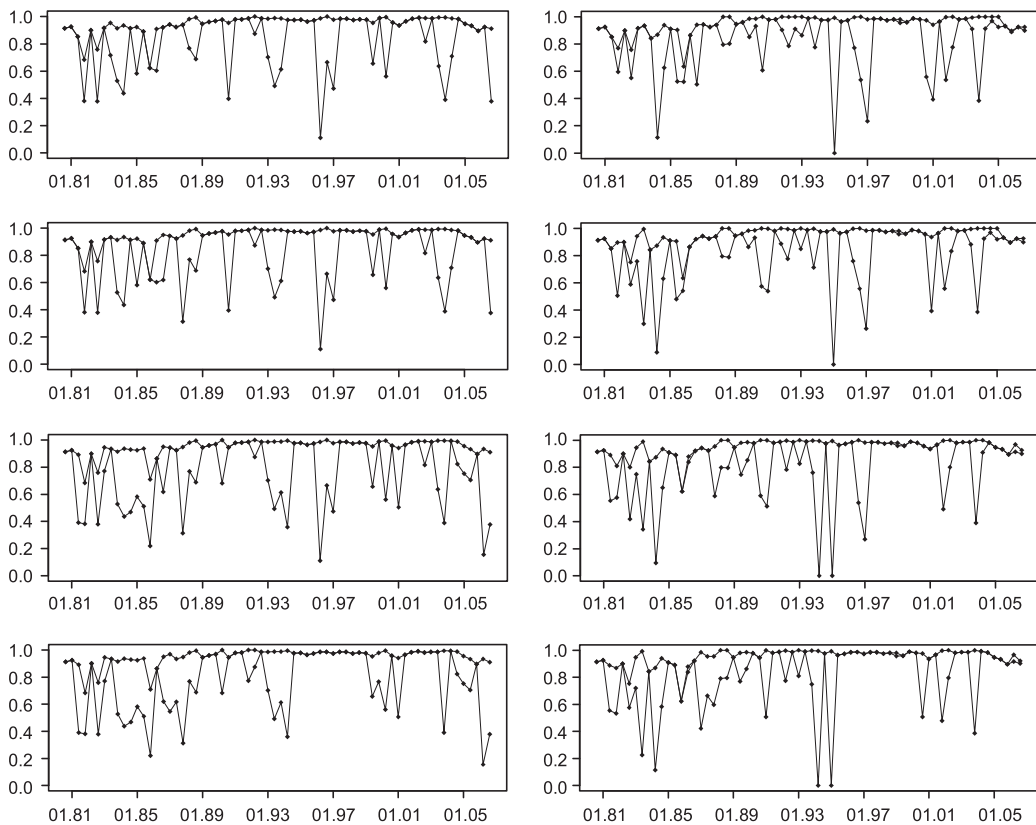


Fig. 2. The values of the flexible shrinkage factors $a_1, a_2 \in [0, 1]$ as a function of time (MM.YY). The flexible shrinkage factors for different risk aversion coefficient $\gamma = \{5, 10, 25, 50\}$, from top to bottom. The cases with certain groups are plotted left, while the cases with uncertain groups are right. The estimation period is $n = 150$, the holding period is $s = 100$ days.

specifications and is directly related to the objective function of the investor. Moreover, because it takes model uncertainty into account by structure adjustment, it stays suitable for the high dimensional portfolio problems as well. Shrinkage factors can be interpreted as an amount of estimation risk in each group of asset and hedge portfolio in the spirit of [Brandt and Santa-Clara \(2006\)](#).

6. Summary

The issue of estimation risk is of immense importance for the portfolio selection. Estimation risk can be mitigated by finding the optimal trade-off between the model structure and bias due to suboptimality of the introduced restrictions. However, the misspecified structure may lead to significant losses as well. This requires accounting for model uncertainty using incoming information.

This paper considers an investor with a mean-variance objective function. Non-homogeneity of estimation risk among assets and time requires its exact quantification. We propose a flexible shrinkage estimator for portfolio weights based on the grouping of assets exhibiting similar amounts of estimation risk. The number of groups corresponds to the number of distinct shrinkages, dynamically adjusted based on market information. The assets are assigned to groups using the clustering procedure with both certain and uncertain group composition. This allows for a more correct simultaneous accounting for both estimation risk and model uncertainty in portfolio choice.

The empirical application of the proposed flexible estimator shows its robustness, compared to alternative approaches, for different performance criteria. We allow dynamic switching between one and two groups of assets with similar amounts of estimation risk. The observed time clusters with the same optimal number of groups validate the necessity to adjust the model structure amount. Our findings apply to different risk aversions and portfolio holding periods. Thus portfolio performance can be improved by relaxing the model structure rigidity and by timely accounting for model uncertainty.

Acknowledgements

We are grateful to the editor, Carl Chiarella, the associate editor, and anonymous referees for valuable comments and suggestions. Any remaining errors are of our own responsibility.

Appendix

Proof of Proposition 1. The structure of the shrinkage coefficient $\mathbf{a} = \mathbf{J}\tilde{\mathbf{a}}$ implies that $\hat{\mathbf{w}} = \mathbf{J}\tilde{\mathbf{a}} \circ \hat{\mathbf{u}} + (\mathbf{1} - \mathbf{J}\tilde{\mathbf{a}}) \circ \mathbf{c}$. Thus the moments of portfolio return can be expressed as

$$\begin{aligned} E(\hat{\mathbf{w}}'\mathbf{r}) &= [\boldsymbol{\mu} \circ (E(\hat{\mathbf{u}} - \mathbf{c}))]' \mathbf{J}\tilde{\mathbf{a}} + \boldsymbol{\mu}'\mathbf{c}, \\ \text{Cov}(\hat{\mathbf{w}}'\mathbf{r}) &= \text{tr}[(\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}') \circ \text{Cov}(\hat{\mathbf{w}})] + E(\hat{\mathbf{w}}'\boldsymbol{\Sigma}E(\hat{\mathbf{w}})). \end{aligned}$$

The covariance matrix of the estimated portfolio weights is $\text{Cov}(\hat{\mathbf{w}}) = (\mathbf{J}\tilde{\mathbf{a}}\tilde{\mathbf{a}}'\mathbf{J}') \circ \text{Cov}(\hat{\mathbf{u}})$. Then the task of constraint optimization is given by

$$\max_{\tilde{\mathbf{a}}} \mathbf{q}'\mathbf{J}\tilde{\mathbf{a}} - \frac{\gamma}{2} \tilde{\mathbf{a}}'\mathbf{J}'\mathbf{Q}\tilde{\mathbf{a}} \quad \text{w.r.t. } \mathbf{1} - \mathbf{c}'\mathbf{1} = \nu.$$

with $\mathbf{q} = (\boldsymbol{\mu} - \gamma\boldsymbol{\Sigma}\mathbf{c}) \circ (E(\hat{\mathbf{u}}) - \mathbf{c})$, $\mathbf{Q} = \text{Cov}(\hat{\mathbf{u}}) \circ (\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}') + \boldsymbol{\Sigma} \circ (E(\hat{\mathbf{u}}) - \mathbf{c})(E(\hat{\mathbf{u}}) - \mathbf{c})'$. The technical constraint ν follows from $E(\hat{\mathbf{w}})\mathbf{1} = \mathbf{1}$. Then the solution is given by

$$\tilde{\mathbf{a}}^* = \frac{(\mathbf{J}'\mathbf{Q}\mathbf{J})^{-1}}{\gamma} \mathbf{J}'(\mathbf{q} - \pi(E(\hat{\mathbf{u}}) - \mathbf{c})), \quad \text{where } \pi = \frac{\mathbf{q}'\mathbf{Q}^{-1}(E(\hat{\mathbf{u}}) - \mathbf{c}) - \nu\gamma}{(E(\hat{\mathbf{u}}) - \mathbf{c})'\mathbf{Q}^{-1}(E(\hat{\mathbf{u}}) - \mathbf{c})}. \quad \square$$

Proof of Proposition 2. Note that $(n_2 - 1)\hat{\boldsymbol{\Sigma}} \sim \mathcal{W}_k(n_2 - 1, \boldsymbol{\Sigma})$ and $\hat{\boldsymbol{\mu}}^{(\kappa)} = \kappa\hat{\boldsymbol{\mu}} \sim \mathcal{N}_k(\kappa\boldsymbol{\mu}, \kappa\boldsymbol{\Sigma}/n_1)$. Because $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are independent, the distribution of $\hat{\mathbf{u}}|\hat{\boldsymbol{\mu}} = \boldsymbol{\mu}$ is equal to the distribution of the random vector

$$\mathbf{u}(\boldsymbol{\mu}, \hat{\boldsymbol{\Sigma}}) = \frac{\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}}{\mathbf{1}'\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}} + \gamma^{-1}\hat{\mathbf{R}}\boldsymbol{\mu}. \quad (14)$$

Following Okhrin and Schmid (2006) we obtain the conditional moments of the portfolio weights as

$$\begin{aligned} E(\hat{\mathbf{u}}|\hat{\boldsymbol{\mu}}) &= \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}} + \frac{n_2 - 1}{n_2 - k - 1} \gamma^{-1}\mathbf{R}\hat{\boldsymbol{\mu}}, \\ \text{Cov}(\hat{\mathbf{u}}|\hat{\boldsymbol{\mu}}) &= \frac{1}{n - k - 1} \frac{\mathbf{R}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}} + \frac{c_1\mathbf{R}\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}'\mathbf{R} + c_2\hat{\boldsymbol{\mu}}'\mathbf{R}\hat{\boldsymbol{\mu}}\mathbf{R}}{\gamma^2}. \end{aligned}$$

It holds that $E(\hat{\mathbf{u}}) = E(E(\hat{\mathbf{u}}|\hat{\boldsymbol{\mu}}))$ and $\text{Cov}(\hat{\mathbf{u}}) = E(\text{Cov}(\hat{\mathbf{u}}|\hat{\boldsymbol{\mu}})) + \text{Cov}(E(\hat{\mathbf{u}}|\hat{\boldsymbol{\mu}}))$. Because $\hat{\boldsymbol{\mu}} \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}/(\kappa n_1))$, $\mathbf{R}\boldsymbol{\Sigma}\mathbf{R} = \mathbf{R}$, and $\text{tr}(\mathbf{R}\boldsymbol{\Sigma}) = k - 1$, it follows that

$$E(\mathbf{R}\hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}'\mathbf{R}) = \mathbf{R}\boldsymbol{\mu}\boldsymbol{\mu}'\mathbf{R} + \mathbf{R}/(\kappa n_1) \quad \text{and} \quad E(\hat{\boldsymbol{\mu}}'\mathbf{R}\hat{\boldsymbol{\mu}}) = \boldsymbol{\mu}'\mathbf{R}\boldsymbol{\mu} + (k - 1)/(\kappa n_1). \quad (15)$$

Tedious substitution completes the proof of the proposition. \square

References

- Aiolfi, M., Timmermann, A., 2006. Persistence in forecasting performance and conditional combination strategies. *Journal of Econometrics* 127, 31–53.
- Brandt, M., 2005. Portfolio choice problems. In: Ait-Sahalia, Y., Hansen, L.P. (Eds.), *Handbook of Financial Econometrics*. Elsevier, New York.
- Brandt, M., Santa-Clara, P., 2006. Dynamic portfolio selection by augmenting the asset space. *Journal of Finance* 61, 2187–2217.
- Brandt, M., Santa-Clara, P., Valkanov, R., 2008. Parametric portfolio policies: exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies*, in press.
- DeMiguel, V., Garlappi, L., Uppal, R., 2008. Optimal versus naive diversification: how inefficient is the $1/N$ portfolio strategy? *Review of Financial Studies*, in press.
- Fleming, J., Kirby, C., Ostdiek, B., 2001. The economic value of volatility timing. *Journal of Finance* 56, 329–352.
- Garlappi, L., Uppal, R., Wang, T., 2007. Portfolio selection with parameter and model uncertainty: a multi-prior approach. *Review of Financial Studies* 20, 41–81.
- Golosnoy, V., Okhrin, Y., 2007. Multivariate shrinkage for optimal portfolio weights. *European Journal of Finance* 13, 441–458.
- Härdle, W., Herwartz, H., Spokoiny, V., 2003. Time inhomogenous multiple volatility modeling. *Journal of Financial Econometrics* 1, 55–95.
- Jorion, P., 1986. Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis* 21, 279–292.
- Klein, R., Bawa, V., 1976. The effect of estimation risk on optimal portfolio choice. *Journal of Financial Economics* 3, 215–231.
- Ledoit, O., Wolf, M., 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance* 10, 603–621.
- Li, D., Ng, W.-L., 2000. Optimal dynamic portfolio selection: multiperiod mean-variance formulation. *Mathematical Finance* 10, 387–406.
- Longford, N., 1999. Multivariate shrinkage estimation of small area means and proportions. *Journal of the Royal Statistical Society: Series A* 162, 227–245.
- Merton, R., 1980. On estimating the expected return on the market. *Journal of Financial Economics* 8, 323–361.
- Michaud, O., 1998. *Efficient Asset Management*. Harvard Business School Press, Boston.
- Okhrin, Y., Schmid, W., 2006. Distributional properties of optimal portfolio weights. *Journal of Econometrics* 134, 235–256.
- Okhrin, Y., Schmid, W., 2008. Estimation of optimal portfolio weights. *International Journal of Theoretical and Applied Finance* 11, 249–276.
- Raftery, A., Dean, N., 2006. Variable selection for model-based clustering. *Journal of American Statistical Association* 101, 168–178.
- Stein, C., 1955. Inadmissability of the usual estimator for the mean of a multivariate normal distribution. In: *Proceedings of Third Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 197–206.
- Tola, V., Lillo, F., Gallegati, M., Mantegna, R., 2008. Cluster analysis for portfolio optimization. *Journal of Economic Dynamics and Control* 32, 235–258.
- Wang, Z., 2005. A shrinkage approach to model uncertainty and asset allocation. *Review of Financial Studies* 18, 673–705.
- Wang, H., Li, G., Jiang, G., 2007. Robust regression shrinkage and consistent variable selection through the LAD-Lasso. *Journal of Business and Economic Statistics* 25, 347–355.
- Witten, I.H., Frank, E., 2005. *Data Mining: Practical Machine Learning Tools and Techniques*, second ed. Morgan Kaufmann, San Francisco, CA.
- Young, M., Lenk, P., 1998. Hierarchical Bayes methods for multifactor model estimation and portfolio selection. *Management Science* 44 (11), 111–124.