

# Comparison of different estimation techniques for portfolio selection

Yarema Okhrin · Wolfgang Schmid

**Summary** The main problem in applying the mean-variance portfolio selection consists of the fact that the first two moments of the asset returns are unknown. In practice the optimal portfolio weights have to be estimated. This is usually done by replacing the moments by the classical unbiased sample estimators. We provide a comparison of the exact and the asymptotic distributions of the estimated portfolio weights as well as a sensitivity analysis to shifts in the moments of the asset returns. Furthermore we consider several types of shrinkage estimators for the moments. The corresponding estimators of the portfolio weights are compared with each other and with the portfolio weights based on the sample estimators of the moments. We show how the uncertainty about the portfolio weights can be introduced into the performance measurement of trading strategies. The methodology explains the bad out-of-sample performance of the classical Markowitz procedures.

**Keywords** Portfolio analysis · Mean-variance analysis · Estimation of portfolio weights · Shrinkage estimation

## 1 Introduction

The classical asset allocation theory of Markowitz (1952) is appealing both from theoretical and practical point of view. In the last few decades numerous extensions and generalizations have been proposed in literature. As a solution to the classical setup of Markowitz, an explicit expression for the optimal portfolio weights in terms of the

---

Y. Okhrin (✉) · W. Schmid

Department of Statistics, European University Viadrina, 15230 Frankfurt (Oder), Germany  
e-mail: okhrin@euv-frankfurt-o.de

W. Schmid

e-mail: schmid@euv-frankfurt-o.de

vector of the expected asset returns and the covariance matrix of the asset returns is obtained. Despite its simplicity and clearness, the mean-variance approach often fails in practical applications by leading to portfolios with unrealistic portfolio weights, for example, highly negative or much larger than one. The reason for such a behavior is that the true moments of the asset returns are unknown and should be estimated from past data. The estimation of the mean and the covariance matrix of multivariate distributions was recently intensively discussed from the perspective of portfolio selection.

The precise estimation of the optimal portfolio weights is of great importance. First, note that large errors in the estimated portfolio weights lead to a substantial but incorrect reallocation of the wealth. For small investors this leads to additional transaction costs. Moreover, significant changes in the portfolio without any visible changes in the fundamentals cause investors' distrust to the mean-variance analysis. Second, the large attraction of the classical mean-variance analysis is due to the fact that the optimal portfolio weights are given explicitly and we do not require any numerical optimization procedures. The price to be paid for this is that we do not demand non-negative portfolio weights. With the estimation of the moments of the asset returns this problem becomes even more pressing. Even if the true portfolio weight is 0.5, the estimated expected quadratic utility portfolio weight for moderate sample size (for example 60 monthly observations and the risk aversion coefficient set to 10) can lie outside the unit interval with probability more than 30%. Moreover, with rather high probability the weights can take extreme and unacceptable values of less than  $-0.5$  or more than  $1.5$ . Third, several testing procedures that rely on the portfolio weights become useless. This refers, for example, to tests for the equality of two portfolios. This is the case in tests for international diversification or tests for the efficiency of a given portfolio. In case of market or volatility timing such tests can be useful to assess the significance of portfolio adjustment as well (Barberis 2000, Fleming et al. 2001).

Despite the large importance of the estimated portfolio weights, there are only a few papers that discuss the behaviour of these estimators from empirical or theoretical perspective. Jobson and Korkie (1980) provide asymptotic approximations for the mean and the variance of the estimated weights that maximize the Sharpe ratio of a portfolio. Britten-Jones (1999) derives the exact distribution of the portfolio weights within a regression framework. Under the assumption of multivariate Gaussian asset returns Okhrin and Schmid (2006a) calculate the exact moments for four types of estimated portfolio weights. A detailed discussion of these results is given in Sect. 2. The problem of the robustness of the portfolio weights and the moments of the portfolio return to shifts in the moments of the asset returns is of special importance. This problem has been raised in the papers of Best and Grauer (1991) and Chopra and Ziemba (1993). Gouriéroux and Monfort (2005) provide a general theory of testing portfolio efficiency based on the asymptotic distribution of the optimal portfolio weights.

To improve the portfolio selection two procedures are available. On the one hand, we can restrict the optimal portfolio weights to belong to the interval  $[0, \bar{w}]$ , where  $\bar{w}$  is some fixed upper bound. This method is discussed by Frost and Savarino (1988). On the other hand we may try to improve the estimators of the moments of the asset returns. A popular approach in this direction is the shrinkage methodology pro-

posed by Stein (1956). Jorion (1986) was the first to apply the shrinkage estimator to portfolio selection by using a shrinkage estimator of the expected asset returns. A shrinkage estimator of the covariance matrix has been proposed and applied to portfolio selection by Ledoit and Wolf (2003, 2004). Both methodologies appear to be very successful in reducing the estimation uncertainty. Okhrin and Schmid (2006b) and Golosnoy and Okhrin (2006) proposed a shrinkage estimator applied directly to the portfolio weights. Both papers argue that shrinkage estimators of the portfolio weights lead to a decrease of the variance of the portfolio weights and to an increase of the utility.

The aim of this paper is to illustrate and to assess the impact of estimating the moments of the asset returns on the portfolio weights. An extensive simulation study is performed to obtain the sample densities and the moments of the portfolio weights using classical and shrinkage estimators for the moments of the asset returns. The results stress the need for new and precise estimators of the portfolio weights. Special attention is drawn to the Sharpe ratio optimal weights, where the moments of order greater or equal than one do not exist. This questions the usefulness of such a measure for practical applications. We also provide a comparison for of the out-of-sample performance of Markowitz-based trading strategies by introducing the uncertainty in portfolio weights. Additionally, the robustness of the portfolio weights with respect to changes in the mean and in the variance of the asset returns is analyzed. It appears that changes in the covariances of the asset returns have a very strong impact on the mean and the covariance matrix of the portfolio weights.

The paper is structured as follows. In Sect. 2 various optimal portfolio weights are discussed and the estimation methodology is described. In Sect. 3 the results of the empirical study are given.

## 2 Portfolio selection

### 2.1 Choice of the portfolio weights

In this paper we consider four classical types of optimal portfolio weights related to the Markowitz theory. Let  $\mathbf{X}$  denote the  $k$ -dimensional vector of the asset returns. It is common to assume in academic literature that  $\mathbf{X}$  follows a multivariate normal distribution, i.e.,  $\mathbf{X} \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ , where  $\boldsymbol{\Sigma}$  is positive definite. This assumption, despite being heavily criticized (see Fama 1965, Mittnik and Rachev 1993, etc.), is a standard benchmark and brings enormous technical advantages. Moreover, there are several papers that argue that the asset returns over longer sampling periods (monthly, quarterly data, etc.) can be seen as a realization of a Gaussian random variable (cf. Fama 1976). Let  $\mathbf{w}$  denote the vector of portfolio weights, i.e., relative fractions of the individual assets in the portfolio.

At each moment of time the investor is confronted with the problem of the expected utility maximization

$$\max_{\mathbf{w}} E U(R_p), \quad \text{s.t. } \mathbf{w}'\mathbf{1} = 1,$$

where  $R_p$  denotes the portfolio return. It is equal to  $R_p = \mathbf{w}'\mathbf{X}$  in case if no risk-free asset is available and to  $R_p = \mathbf{w}'(\mathbf{X} - r_f\mathbf{1}) + r_f$  if a risk-free asset with fixed return  $r_f$  is present. Here,  $\mathbf{1}$  stands for the vector of ones, i.e.,  $\mathbf{1} = (1, \dots, 1)'$ . The power, exponential, log and quadratic utilities are the most widely used utility functions in the economic literature. However, in general any concave and monotone increasing function can be used for the purposes of portfolio selection. Assuming this type of the utility function and Gaussian asset returns, the problem of maximization of the utility is equivalent to the mean-variance or the expected quadratic utility maximization problem (cf. Ingersoll 1987) given by

$$\max_{\mathbf{w}} E(R_p) - \frac{\gamma}{2} \text{Var}(R_p), \quad \text{s.t. } \mathbf{w}'\mathbf{1} = 1,$$

where  $\gamma > 0$  is called the risk-aversion coefficient. It measures the investors attitude to risk (see Ingersoll 1987). In case if no risk-free asset is available the problem is given by

$$\max_{\mathbf{w}} \mathbf{w}'\boldsymbol{\mu} - \frac{\gamma}{2} \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}, \quad \text{s.t. } \mathbf{w}'\mathbf{1} = 1. \quad (1)$$

The solution of (1) is denoted by  $\mathbf{w}_{EU}$  and it is given by

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

Note that  $\mathbf{w}_{EU}$  is a linear function of the asset means, and, therefore, a precise estimation of the expected stock returns is crucial for the estimation of  $\mathbf{w}_{EU}$ . The impact diminishes with the increase of the risk aversion coefficient. The limit as  $\gamma$  tends to infinity leads to the global minimum-variance portfolio weights  $\mathbf{w}_{GMV}$  with

$$\mathbf{w}_{GMV} = \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}. \quad (3)$$

In the presence of a risk-free asset with return  $r_f$  the mean-variance optimization problem is given by

$$\max_{\mathbf{w}} \mathbf{w}'(\boldsymbol{\mu} - r_f\mathbf{1}) - \frac{\gamma}{2} \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}.$$

This leads to the tangency portfolio. The optimal weights of the risk assets are equal to

$$\mathbf{w}_{TP} = \gamma^{-1}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1}). \quad (4)$$

The weight of the risk-free asset is given by  $w_{TP,0} = 1 - \mathbf{1}'\mathbf{w}_{TP}$ .

Another popular measure of the portfolio performance is the Sharpe ratio

$$SR = \frac{E(R_p)}{\sqrt{\text{Var}(R_p)}}.$$

If no riskless asset is available, then maximizing the Sharpe ratio leads to the portfolio weights

$$\mathbf{w}_{SR} = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}, \quad (5)$$

provided that  $\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \neq 0$ . These portfolio weights depend on the inverse of a linear form in the expected asset returns. This implies that all classical estimation procedures of  $\boldsymbol{\mu}$  would lead to results that are difficult to interpret since they provide estimators that do not bound  $\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$  away from zero. We illustrate this point below and Okhrin and Schmid (2006a) assessed it more rigorously.

Note that the components of  $\mathbf{w}$  usually take values between zero and one; however, negative values or values larger than one are allowed. The ‘short selling’ corresponds to the situation, when the investor sells an asset which he does not own. This leads to negative portfolio weights. Usually it is implemented by borrowing the asset from another institution and selling it on the market. A value of the weight higher than one implies that the investor sold other assets short in order to achieve more than 100% stake of this asset in the portfolio. An extensive discussion of this point can be found in Farrell (1997). If we pose the restriction that the portfolio weights belong to the interval between zero and one, then the formulae for the optimal portfolio weights cannot be derived explicitly and numerical Kuhn–Tucker optimization is required.

## 2.2 Sample estimators of the portfolio weights

The expressions for the optimal portfolio weights in the last section are infeasible, since the true distribution parameters of the asset returns are unknown. In practice  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$  are estimated by historical data. It is assumed that a sample  $\mathbf{X}_1, \dots, \mathbf{X}_n$  is available with  $\mathbf{X}_i \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ . Next we consider several estimation methods.

The classical approach is based on the unbiased sample estimators of the expected returns and the covariance matrix of the risky assets. Let  $\mathcal{X} = (\mathbf{X}_1, \dots, \mathbf{X}_k)'$ . Then the classical estimators of  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$  are given by

$$\hat{\boldsymbol{\mu}}_{cl} = \frac{1}{n}\mathcal{X}'\mathbf{1}, \quad \hat{\boldsymbol{\Sigma}}_{cl} = \frac{1}{n-1}\mathcal{X}'(\mathbf{I} - \mathbf{1}\mathbf{1}'/n)\mathcal{X}.$$

Thus the estimated portfolio weights are given by

$$\begin{aligned} \hat{\mathbf{w}}_{EU}^{(cl)} &= \frac{\hat{\boldsymbol{\Sigma}}_{cl}^{-1}\mathbf{1}}{\mathbf{1}'\hat{\boldsymbol{\Sigma}}_{cl}^{-1}\mathbf{1}} + \gamma^{-1}\hat{\mathbf{R}}_{cl}\hat{\boldsymbol{\mu}}_{cl}, & \hat{\mathbf{w}}_{GMV}^{(cl)} &= \frac{\hat{\boldsymbol{\Sigma}}_{cl}^{-1}\mathbf{1}}{\mathbf{1}'\hat{\boldsymbol{\Sigma}}_{cl}^{-1}\mathbf{1}}, \\ \hat{\mathbf{w}}_{TP}^{(cl)} &= \gamma^{-1}\hat{\boldsymbol{\Sigma}}_{cl}^{-1}(\hat{\boldsymbol{\mu}}_{cl} - r_f\mathbf{1}), & \hat{\mathbf{w}}_{SR}^{(cl)} &= \frac{\hat{\boldsymbol{\Sigma}}_{cl}^{-1}\hat{\boldsymbol{\mu}}_{cl}}{\mathbf{1}'\hat{\boldsymbol{\Sigma}}_{cl}^{-1}\hat{\boldsymbol{\mu}}_{cl}}. \end{aligned}$$

Under the assumption of Gaussian returns  $(n-1)\hat{\boldsymbol{\Sigma}}_{cl}$  follows a Wishart distribution with  $n-1$  degrees of freedom. Thus the analysis of the estimated portfolio weights involves the analysis of inverted Wishart matrices, linear and quadratic forms in Wishart matrices and their ratios. Based on such results, Okhrin and Schmid (2006a, b) provide properties of the exact and asymptotic distributions of the estimated portfolio weights. In case of a finite sample they assumed that the variables  $\mathbf{X}_1, \dots, \mathbf{X}_n$  are independent, while for asymptotic considerations it is demanded that  $\{\mathbf{X}_i\}$  is a stationary sequence. They proved that  $\hat{\mathbf{w}}_{EU}^{(cl)}$ ,  $\hat{\mathbf{w}}_{GMV}^{(cl)}$  and  $\hat{\mathbf{w}}_{TP}^{(cl)}$  are asymptotically unbiased and consistent as well. However, the finite sample bias and the sample

variance can be extremely large for the values of  $n$  usually used in practical applications. For the estimated Sharpe ratio optimal weight it is shown that the first moment does not exist at all. This implies in practice that different sample sizes used for estimating higher moments may provide absolutely opposite results. Moreover, relying on the moments of the asymptotic is not eligible since they are useless if the finite moments do not exist.

Two simple modifications of the classical estimators are directly obtained. First, the estimators of the portfolio weights are corrected for the bias. Nevertheless, it can be shown that the mean squared error for the unbiased estimators is higher than for the biased estimators. The second possibility is to consider the maximum-likelihood estimator of the covariance matrix. In this case the estimator for the portfolio weights is a maximum-likelihood estimator of the true portfolio weights since the latter are continuous functions of  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$ . However, the performance of the maximum-likelihood estimators is very similar. This implies, that to obtain a less volatile estimator other estimation procedures are required.

### 2.3 Shrinkage estimators

There are several alternatives to estimate the moments of the asset returns. Recently, the shrinkage technique is widely applied. Shrinkage estimators are constructed as a weighted sum of the usual estimators and estimators based on other models or assumptions. The idea of shrinkage was stimulated by the seminal work of Stein (1956), who argues that the sample mean is inadmissible in terms of the average risk function in dimensions higher than two. Efron and Morris (1976) extended the result of Stein. For  $\mathbf{X} \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  with  $k \geq 3$  they obtained the class of estimators of the form  $\hat{\boldsymbol{\mu}}_j = (1 - \omega)\hat{\boldsymbol{\mu}}_{cl} + \omega\boldsymbol{\mu}_0$  that has uniformly lower risk than the sample mean.  $\omega$  denotes the shrinkage coefficient and  $\boldsymbol{\mu}_0$  an arbitrary constant to which the sample mean is shrunk. The optimal choice of the weighting parameter is obtained by minimizing the mean squared error (MSE) of the shrinkage estimator. Thus, the shrinkage methodology provides a solution to the trade-off between bias and estimation error. In other words the new estimator is biased, but achieves a lower MSE compared with the classical estimator.

Jorion (1986, 1991) was the first to apply this method to the expected asset returns with the estimators given by

$$\hat{\boldsymbol{\mu}}_j = (1 - \hat{\omega})\hat{\boldsymbol{\mu}}_{cl} + \hat{\omega}\mathbf{1}\hat{\boldsymbol{\mu}}_0, \quad \hat{\boldsymbol{\Sigma}}_j = \hat{\boldsymbol{\Sigma}}_{cl} \left( 1 + \frac{1}{n + \lambda} \right) + \frac{\lambda}{n(n + 1 + \lambda)} \frac{\mathbf{1}\mathbf{1}'}{\mathbf{1}'\hat{\boldsymbol{\Sigma}}_{cl}^{-1}\mathbf{1}}, \quad (6)$$

where

$$\hat{\boldsymbol{\mu}}_0 = \frac{\mathbf{1}'\hat{\boldsymbol{\Sigma}}_{cl}^{-1}\hat{\boldsymbol{\mu}}_{cl}}{\mathbf{1}'\hat{\boldsymbol{\Sigma}}_{cl}^{-1}\mathbf{1}}, \quad \lambda = \frac{\hat{\omega}n}{1 - \hat{\omega}}. \quad (7)$$

The optimal shrinkage intensity  $\omega$  is chosen to minimize the quadratic loss function

$$L(\boldsymbol{\mu}, \hat{\boldsymbol{\mu}}_j) = (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}_j)' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}_j).$$

The result of Stein can be alternatively derived using a Bayesian setup with conjugate informative prior for the mean (see Jorion 1986). From this approach and using empirical Bayes techniques the estimator of  $\omega$  used in (6) and (7) can be given by

$$\hat{\omega} = \frac{k+2}{k+2+n(\hat{\boldsymbol{\mu}}-\hat{\boldsymbol{\mu}}_0)'\boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\mu}}-\hat{\boldsymbol{\mu}}_0)}.$$

In practice  $\boldsymbol{\Sigma}$  is unknown and it is replaced as in Zellner and Chetty (1965) by  $(n-1)\hat{\boldsymbol{\Sigma}}_{cl}/(n-k-2)$ .  $\hat{\boldsymbol{\mu}}_0$  is called the grand mean. In our framework it is equal to the average return on the global minimum-variance portfolio, if the parameters are replaced by classical estimators. Using these estimators the following portfolio weights can be constructed

$$\hat{\mathbf{w}}_{EU}^{(j)} = \frac{\hat{\boldsymbol{\Sigma}}_j^{-1}\mathbf{1}}{\mathbf{1}'\hat{\boldsymbol{\Sigma}}_j^{-1}\mathbf{1}} + \gamma^{-1}\hat{\mathbf{R}}_j\hat{\boldsymbol{\mu}}_j, \quad \hat{\mathbf{w}}_{TP}^{(j)} = \gamma^{-1}\hat{\boldsymbol{\Sigma}}_j^{-1}(\hat{\boldsymbol{\mu}}_j - r_f\mathbf{1}), \quad \hat{\mathbf{w}}_{SR}^{(j)} = \frac{\hat{\boldsymbol{\Sigma}}_j^{-1}\hat{\boldsymbol{\mu}}_j}{\mathbf{1}'\hat{\boldsymbol{\Sigma}}_j^{-1}\hat{\boldsymbol{\mu}}_j}.$$

Recently, Ledoit and Wolf (2003) proposed a shrinkage estimator of the covariance matrix under the assumption that the asset returns are Gaussian. It is a common practice in finance to use linear regression models (or factor models) for modeling asset returns or for testing asset pricing theories. More precisely, assume that the asset returns are generated by the following model

$$X_{ij} = \alpha_i + \beta_i Z_{0j} + \varepsilon_{ij}, \quad i = 1, \dots, k, \quad j = 1, \dots, n,$$

where  $Z_{0j}$  denotes a single stochastic regressor. Then the covariance matrix of  $\mathbf{X}_j = (X_{1j}, \dots, X_{kj})'$  can be estimated by  $\hat{\mathbf{F}} = s_{00}^2\hat{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}}' + \hat{\mathbf{D}}$ . Here  $\hat{\boldsymbol{\beta}}$  denotes the least-squares estimators (LSE) of the slope parameters  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_k)'$ .  $\hat{\mathbf{D}}$  is the LSE of the covariance matrix of the residuals and  $s_{00}^2$  is the sample variance of the single stochastic regressor. The shrinkage estimator of  $\boldsymbol{\Sigma}$  is then given by

$$\hat{\boldsymbol{\Sigma}}_{lw} = \hat{\theta}\hat{\mathbf{F}} + (1-\hat{\theta})\hat{\boldsymbol{\Sigma}}_{cl}.$$

A useful property of this estimator is that it is also feasible if the number of assets is larger than the length of the estimation period, i.e., if the sample covariance matrix is singular. The optimal shrinkage intensity minimizes as before the MSE of the weighted estimator. Ledoit and Wolf (2003) provide a detailed discussion on the consistent estimation of  $\theta$ . We use this approach to estimate the portfolio weights. The covariance matrix is estimated by  $\hat{\boldsymbol{\Sigma}}_{lw}$  and the expected returns are estimated by  $\hat{\boldsymbol{\mu}}_{cl}$  as above. We denote the corresponding portfolio weights by  $\hat{\mathbf{w}}^{(lw)}$ . For comparison purposes we also consider portfolio weights with the expected portfolio return estimated as in Jorion (1986) and the covariance matrix estimated according to Ledoit and Wolf (2003). These weights we write with the superscript  $\hat{\mathbf{w}}^{(lwj)}$ .

## 2.4 Out-of-sample utility

The fact that the optimal portfolio weights are estimated is crucial for the investor. This can be illustrated in the best way by analyzing the out-of-sample performance of the portfolios. The empirical and theoretical analysis of the out-of-sample performance has been an intriguing issue in portfolio selection starting from Frankfurter et al. (1971) and Jobson and Korkie (1981) and up to the recent contributions of Kan and Zhou (2006) and Kan and Smith (2006). In practice the investor estimates the portfolio weights at time point  $t$  and keeps the portfolio over the holding period  $s$ . For simplicity we discuss only the case of the expected quadratic utility. Let  $\hat{\mathbf{w}}_{EU,t}$  denote any estimator of  $\mathbf{w}_{EU}$  calculated at time point  $t$  by the independent random sample  $\mathbf{X}_{t-n+1}, \dots, \mathbf{X}_t$  with  $\mathbf{X}_i \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  for  $i = t-n+1, \dots, t$ . At time point  $t+s$  the investor is willing to assess the goodness of his strategy. To do so he uses the former portfolio weights  $\hat{\mathbf{w}}_{EU,t}$  and the moments of the asset returns valid at the end of the holding period  $t+s$ . We denote these moments by  $\boldsymbol{\mu}^*$  and  $\boldsymbol{\Sigma}^*$ . The direct way of computing the utility at time point  $t+s$  by

$$\hat{\mathbf{w}}'_{EU,t} E(\mathbf{X}_{t+s}) - \frac{\gamma}{2} \hat{\mathbf{w}}'_{EU,t} \text{Var}(\mathbf{X}_{t+s}) \hat{\mathbf{w}}_{EU,t} = \hat{\mathbf{w}}'_{EU,t} \boldsymbol{\mu}^* - \frac{\gamma}{2} \hat{\mathbf{w}}'_{EU,t} \boldsymbol{\Sigma}^* \hat{\mathbf{w}}_{EU,t} \quad (8)$$

does not take into account the uncertainty caused by estimating the portfolio weights, i.e., it treats the portfolio weights as known quantities. However, as it is illustrated in the empirical part this uncertainty may lead to substantial deviations in the efficient frontier. The correct way to compute the out-of sample utility which incorporates the uncertainty in  $\mathbf{w}_{EU}$  is, therefore, the following

$$E(\hat{\mathbf{w}}'_{EU,t} \mathbf{X}_{t+s}) - \frac{\gamma}{2} \text{Var}(\hat{\mathbf{w}}'_{EU,t} \mathbf{X}_{t+s}) = E(\hat{\mathbf{w}}_{EU,t})' \boldsymbol{\mu}^* - \frac{\gamma}{2} [\text{tr}(\boldsymbol{\Sigma}^* \text{Var}(\hat{\mathbf{w}}_{EU,t})) + \boldsymbol{\mu}^{*'} \text{Var}(\hat{\mathbf{w}}_{EU,t}) \boldsymbol{\mu}^* + E(\hat{\mathbf{w}}_{EU,t})' \boldsymbol{\Sigma}^* E(\hat{\mathbf{w}}_{EU,t})]. \quad (9)$$

This utility can be explicitly computed for the classical estimators using the results of Okhrin and Schmid (2006a). However, for the shrinkage estimators discussed in the previous section no explicit expression can be given.

As we will see there can be a substantial difference between the in-sample and the out-of-sample utilities if the variance of the portfolio weights is large. This fact explains the commonly reported evidence about the bad out-of-sample performance of Markowitz procedures.

## 3 Analysis and comparison of the portfolio estimators

The aim of the section is threefold. First, since the unbiased estimators are still the most common estimators of the distribution parameters, we want to assess the finite sample properties of the classical estimators of the portfolio weights. Second, the sensitivity of the moments of these weights with respect to changes in the means and the covariance matrices of the asset returns is examined. Third, we compare the impact of other estimation strategies on the distribution of the optimal portfolio weights. The



analysis is performed for different values of the risk aversion coefficient and different sample sizes.

For the empirical study we use monthly data from Morgan Stanley Capital International (MSCI) for the equity markets of 10 developed countries (U.K., Germany, France, the Netherlands, U.S., Canada, Japan, Italy, Spain, Switzerland) for the period from January 1970 to April 2004 (413 observations for each country). The MSCI World Index is taken as the single factor for the shrinkage estimation proposed by Ledoit and Wolf (2003). The whole sample is used to estimate the mean and the covariance matrix of the returns. In the following these values are taken as our model parameters. We choose this procedure to get realistic values for the moments of the returns.

For the classical estimators of the weights, we have formulas for their moments. Note that the explicit density is only available for the global minimum-variance portfolio. However, we are not aware of similar results for the shrinkage estimators. This is first of all due to the complicated structure of the estimated shrinkage intensities. In order to discuss the performance of the portfolio characteristics, it is possible to estimate the relevant quantities within a simulation study. For a given sample size  $n$  we simulate independent multivariate normal observations with the parameters given by the classical estimators as described above. Using the simulated data, we estimate all portfolio weights using the methods described above. This procedure is repeated  $10^7$  times. The sample of the estimated portfolio weights is used for further computations.

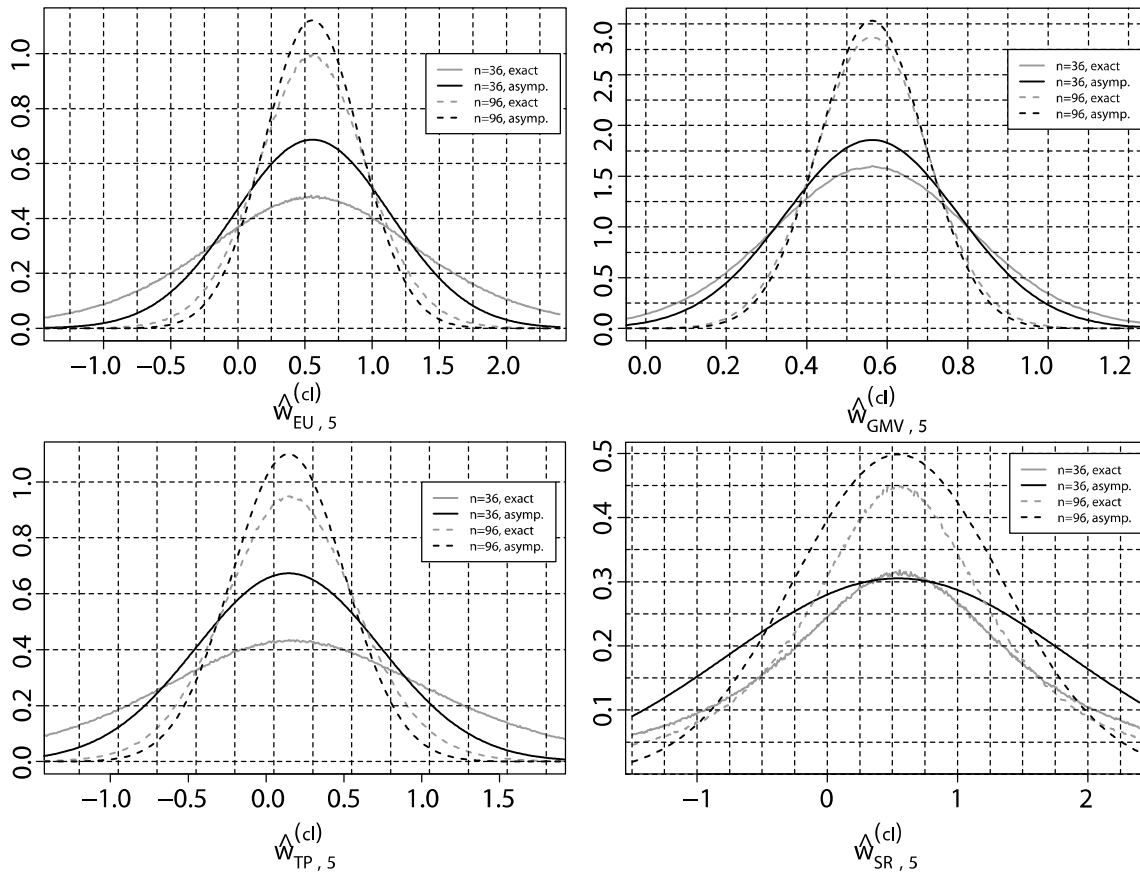
In our study the risk-free rate of return is set equal to 0.02. In financial literature a risk aversion coefficient equal to 2 reflects a low risk aversion and a value of 25 or 50 reflects a highly averse risk behaviour (cf. Elton and Gruber 1999). We use in our study an investor with a modest aversion to risk and set  $\gamma$  equal to 10.

### 3.1 Behaviour of the classical portfolio estimators

Figure 1 plots the exact and the asymptotic densities for different sample sizes  $n$ . According to Okhrin and Schmid (2006a) the estimated optimal portfolio weights are asymptotically normally distributed. These results are used for the plotting of the asymptotic densities. The exact finite sample densities are estimated by histograms. This approach is justified due to the large number of repetitions ( $10^7$ ).

First, note that the spread of the optimal portfolio weights is extremely large. For  $n = 96$  only about 70% of the expected quadratic utility optimal portfolio weight for the U.S. market lies in the interval between zero and one. This is especially striking since the estimated weight is approximately 0.55 and lies in the middle of the interval. For the weights of the other countries the situation is even worse. This problem seems to be of special relevance for the portfolio weights which depend on the estimated expected asset returns, i.e., the expected quadratic utility, the tangency and the Sharpe ratio optimal portfolio weights. The global minimum variance portfolio weights are independent on the estimated expected asset returns and the probability of obtaining negative or values higher than one is almost zero.

Second, note that the exact distributions appear to have much heavier tails than the asymptotic counterparts. This implies, that relying on the asymptotic results would underestimate the probabilities of extreme portfolio weights. However, with increas-

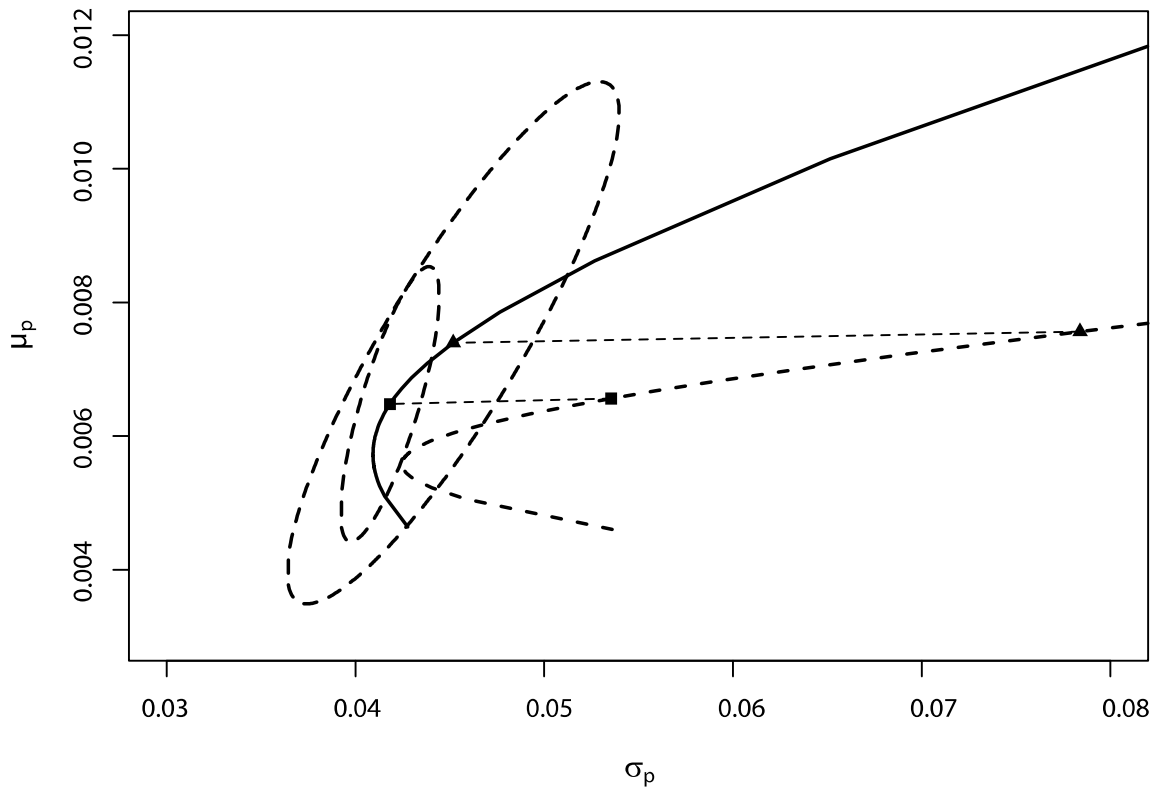


**Fig. 1** The densities of the portfolio weights of the U.S. market as a function of the sample size. The black lines correspond to the asymptotic densities and the grey lines to the exact ones. The number of replications is set to  $10^7$ , the risk aversion coefficient is equal to 10 and the risk-free rate to 0.02.

ing sample size, the exact distributions become more peaked and their tails decrease. For  $\hat{w}_{GMV}^{(cl)}$  the exact density appears to be closer to the asymptotic density compared with the other types of weights. For the other portfolio weights and for usual estimation periods of 4–5 years, the difference between the exact and asymptotic densities is still too large to rely on the asymptotic results.

The distribution of the Sharpe ratio portfolio weights requires a separate discussion. Okhrin and Schmid (2006a) showed that the moments of  $\hat{w}_{SR}^{(cl)}$  do not exist for orders greater or equal one. This implies that the average portfolio weight for different replications within the simulation study does not converge to the true portfolio weight as it is motivated by the weak law of large numbers. Furthermore the variance of the estimated portfolio weights may increase with the number of replications. Another consequence is the fact that minor changes in the observations, i.e., adding a few new observations, may lead to dramatic changes in the portfolio weights. Moreover, the tails of the exact distribution are extremely heavy, what is also illustrated on Fig. 1. Since  $\hat{w}_{SR}^{(cl)}$  is asymptotically normally distributed, the difference between the exact and asymptotic densities diminishes with larger sample sizes. However, the rate of convergency is extremely slow compared with the other types of portfolio weights.

Figure 2 illustrates the impact of uncertainty in the portfolio weights on the expected utility. The solid line shows the efficient frontier for the estimated portfolio



**Fig. 2** The picture shows the true (solid line) and the out-of-sample (dashed line) efficient frontiers given by equations (8) and (9), respectively. On both frontiers the portfolios with  $\gamma = 5$  and 10 are marked with a triangle and a square respectively. The ellipsoids are the asymptotic 95%-confidence intervals for given portfolios due to the uncertainty in the portfolio weights

weights taken as certain quantities, i.e., as in equation (8). The extreme left-side point corresponds to the global minimum variance portfolio with an infinite risk aversion coefficient. A decrease of the risk aversion moves the optimal portfolio along the efficient frontier in the direction of higher returns and volatility.

To assess the impact of the estimated portfolio weights we construct an asymptotic joint confidence interval for the portfolio return and the standard deviation of the portfolio return. Let  $g(\mathbf{w}_{EU}) = (\mathbf{w}'_{EU}\boldsymbol{\mu}^*, \sqrt{\mathbf{w}'_{EU}\boldsymbol{\Sigma}^*\mathbf{w}_{EU}})'$ . Following the Delta-method the asymptotic distribution of  $g(\hat{\mathbf{w}}_{EU})$  is given by

$$\sqrt{n}(g(\hat{\mathbf{w}}_{EU}) - g(\mathbf{w}_{EU})) \xrightarrow{d} N_2(\mathbf{0}, \boldsymbol{\Sigma}_g),$$

$$\text{where } \boldsymbol{\Sigma}_g = \frac{\partial g(\mathbf{w}_{EU})}{\partial \mathbf{w}'_{EU}} \left[ \lim_{n \rightarrow \infty} n \text{Var}(\hat{\mathbf{w}}_{EU}) \right] \frac{\partial g(\mathbf{w}_{EU})'}{\partial \mathbf{w}_{EU}}.$$

Then the asymptotic confidence intervals constitute such values of  $\boldsymbol{\theta} \in \mathbb{R}^2$  that satisfy the inequality

$$(g(\hat{\mathbf{w}}_{EU}) - \boldsymbol{\theta})' \boldsymbol{\Sigma}_g^{-1} (g(\hat{\mathbf{w}}_{EU}) - \boldsymbol{\theta}) \leq \chi_{2,1-\alpha}^2$$

where  $\chi_{2,1-\alpha}^2$  stands for the quantile of the  $\chi^2$  distribution with two degrees of freedom.

The two marked points on the efficient frontier correspond to the portfolios with the risk aversion coefficient equal to 5 (for the upper point) and 10 (for the lower



### 3.2 Sensitivity analysis for the classical portfolio estimators

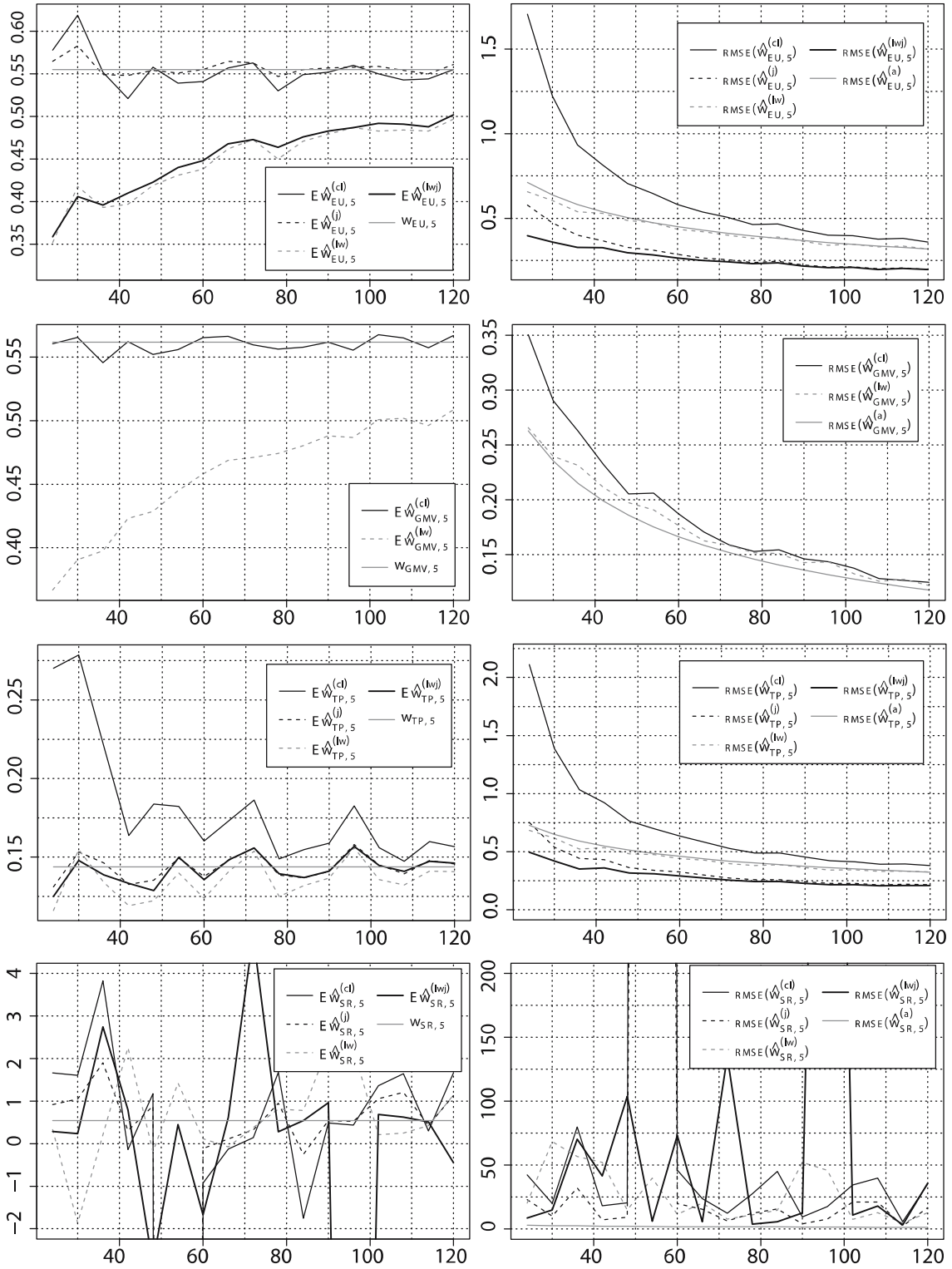
Table 1 contains the results of a robustness analysis for the exact moments of  $\hat{\mathbf{w}}_{EU}^{(cl)}$  and of  $\hat{\mathbf{w}}_{TP}^{(cl)}$  with respect to changes in the mean and the variance of the fifth asset (U.S.) for different values of the risk aversion coefficient. Block A contains the turnover which arises due to the shift. We compute it as the sum of the absolute deviations between the portfolio weights with and without the shift. This measure is of special importance since the transaction costs payed by the investor are in most cases proportional to it. Block B provides the largest absolute change in the portfolio weights and block C the largest relative change in the covariance matrix of the portfolio weights.

As it has been widely reported,  $\hat{\mathbf{w}}_{EU}^{(cl)}$  and  $\hat{\mathbf{w}}_{TP}^{(cl)}$  react sensitive already to small changes in  $\boldsymbol{\mu}$  (see Best and Grauer 1991). Note that in contrary to Best and Grauer (1991), we are able to compute the shifts in the moments of portfolio weights analytically and not via a simulation study. For moderate changes of the mean for the U.S. market (here:  $\xi = 10\%$ ) and a risk aversion of 2 the highest absolute change in the weights is 0.36, what implies a large reallocation of assets. The turnover provides even a more dramatic illustration. The investor falsely reallocates 73% of the expected quadratic utility portfolio (or 79% for tangency portfolio). Note that due to the large uncertainty in the estimation of the expected asset returns an error of 10% can still be seen as modest. The influence on the covariances decreases very fast for larger values of the risk aversion. The global minimum variance portfolio does not depend in the limit on the mean vector at all. The tangency portfolio is almost insensitive to any shifts in the mean of asset returns. The mean and the turnover depend symmetrically on the shifts in the mean; however, the variance reacts stronger to negative shifts than to positive ones.

The second part of the table presents the results subject to shifts in the variance of the U.S. market. All quantities of interest react strongly even to small changes and this reaction is very robust with respect to the risk aversion coefficient. This implies that investing to the global minimum variance portfolio reduces the systematic risk, but the estimation risk is kept at the same level as for the other portfolios on the efficient frontier. Here a decrease of the standard deviation of the asset leads to a stronger change in the quantities of interest than an increase. We do not report for which weight the strongest reaction was observed, but it is worth noting that mostly quantities related to the shifted asset are subject to the largest change. Thus estimation errors in the mean and the variance of an asset have the largest impact on the weight of the same asset. The results on the sensitivity of the correlation coefficients are not presented here, but the magnitude of the change was in most cases close to that for the variance.

### 3.3 Shrinkage estimators for the portfolio weights

Next we compare the classical estimators of the portfolio weights with the estimators based on the shrinkage technique. In general by its idea the shrinkage estimation should outperform the underlying strategy. This is because by shrinking we replace a fraction of a volatile quantity by a deterministic or a less volatile quantity. This leads to an overall reduction in the volatility but also to a higher finite sample bias. To



**Fig. 3** Mean and root mean-square error of the optimal portfolio weight for the U.S. market as a function of the sample size using different estimation methods. The number of replications is equal to  $10^6$ , the risk aversion coefficient is equal to 10 and the risk-free rate to 0.02

assess this bias and the decrease in the mean square error we plot the mean and the root mean-square error (RMSE) for each estimation method as a function of the sample size. The results are presented in Fig. 3. We consider the following estimators for

the portfolio weights: *a*)  $\hat{\mathbf{w}}^{(cl)}$  ( $\hat{\boldsymbol{\mu}}_{cl}$  and  $\hat{\boldsymbol{\Sigma}}_{cl}$  are used as the estimators of the moments of asset returns ); *b*)  $\hat{\mathbf{w}}^{(j)}$  ( $\hat{\boldsymbol{\mu}}_j$  and  $\hat{\boldsymbol{\Sigma}}_j$ ); *c*)  $\hat{\mathbf{w}}^{(lw)}$  ( $\hat{\boldsymbol{\mu}}_{cl}$  and  $\hat{\boldsymbol{\Sigma}}_{lw}$ ); *d*)  $\hat{\mathbf{w}}^{(jlw)}$  ( $\hat{\boldsymbol{\mu}}_j$  and  $\hat{\boldsymbol{\Sigma}}_{lw}$ ). The results for the mean are compared with the true optimal portfolio weights. The results for the RMSE are compared with the standard deviation of the asymptotic distribution which we denote by  $RMSE(\hat{\mathbf{w}}_{,5}^{(a)})$ . These quantities were derived in Okhrin and Schmid (2006a).

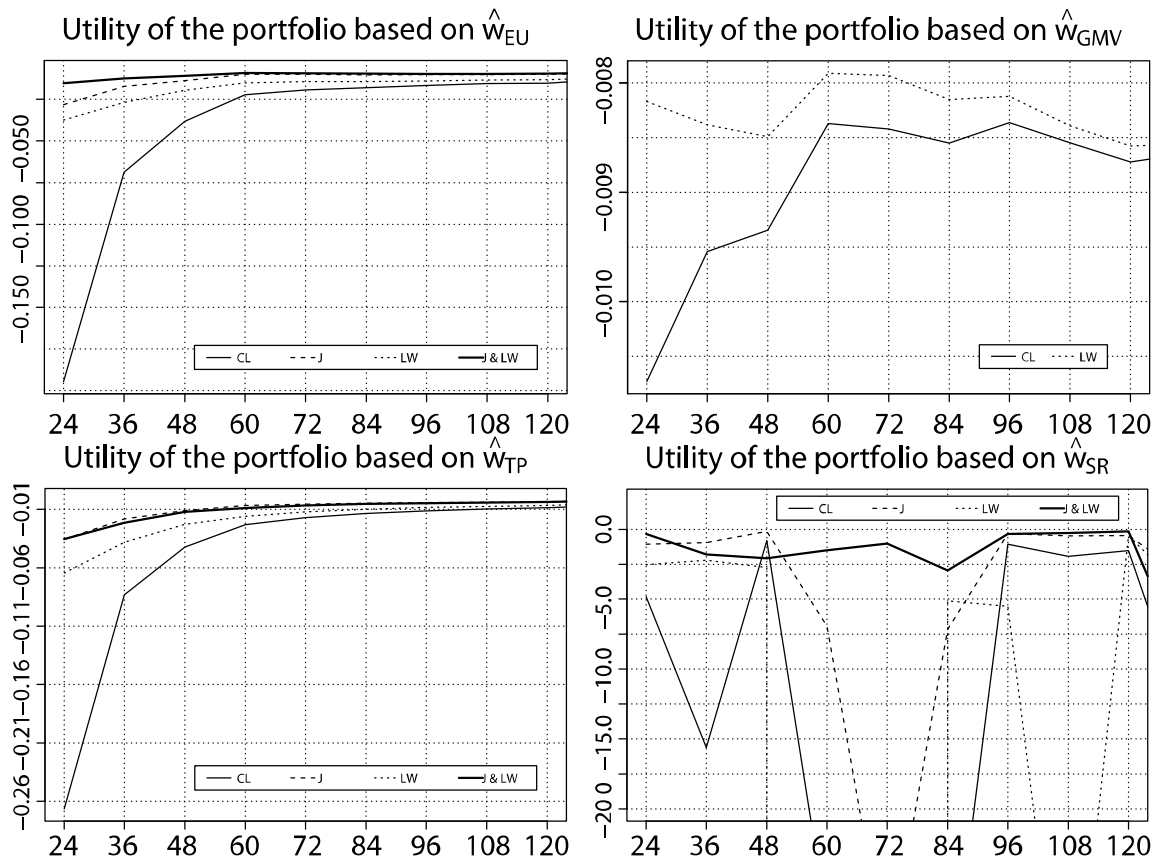
For  $\mathbf{w}_{EU}$  and  $\mathbf{w}_{GMV}$  the methodology which involves the shrinkage for the covariance matrix as proposed by Ledoit and Wolf (2003) leads to a substantial bias in the weights even for larger sample sizes (10 years). On the contrary, the shrinkage of Jorion (1986) is close to the asymptotic mean, for all weights except of  $\mathbf{w}_{SR}$ . In case of the RMSE the methods *c*) and *d*) show the best performance beating even the asymptotic RMSE with known  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$ . Thus it seems that minimization of the MSE for the mean and variance of the asset returns also substantially decreases the MSE of the estimated portfolio weights for small sample sizes. However, for larger sample sizes (5 years and more) the difference in the performance is not that drastic. The behaviour of the mean and of the RMSE of  $\hat{\mathbf{w}}_{SR}$  reflects the fact that the first two moments of the Sharpe ratio optimal portfolio weights do not exist.

### 3.4 Expected utility in case of estimated portfolio weights

In this section we assess the impact of the sample size on the out-of-sample utility, and we show the difference between the out-of-sample and the usual utility on the example of real trading.

Figure 4 illustrates the out-of-sample expected quadratic utilities given in (9) for different estimation methods as a function of the sample size. As already discussed above the out-of-sample utility can be computed explicitly for the classical estimation, but we use the portfolio returns from the simulation study to compute the moments of the portfolio weights for the shrinkage estimation. The number of replications is set to  $10^6$ . The moments  $\boldsymbol{\mu}^*$  and  $\boldsymbol{\Sigma}^*$  are estimated as described in Sect. 3.1. The holding period is set to 60 months. This approach allows us to assess the impact of the sample size on the trading strategy performance. For all types of portfolio weights the shrinkage estimation leads to a higher utility than the classical estimation. The outperformance is strong for small estimation periods up to 5 years and diminishes for longer estimation periods. These results imply that the reduction of the estimation uncertainty for the moments of the asset returns leads also to a significant reduction in the estimation uncertainty of the portfolio weights as well as of the moments of the portfolio return.

The last illustration helps to determine the impact of the sample size, however, does not lead to any conclusions about the impact of estimated weights. To assess this impact, we consider an investor with the following trading strategy. At each moment of time  $t$  the optimal portfolio weights are estimated using the last  $n$  observations. The length of the holding period is set  $s$ . At time point  $t + s$  the investor assesses the utility of the investment ones using the utility as given in (8) and ones as in (9). In the case of the classical estimators for the portfolio weights we use the results of Okhrin and Schmid (2006a). For the shrinkage estimator, we compute the moments of the estimated portfolio weights via a simulation study. The asset returns are generated

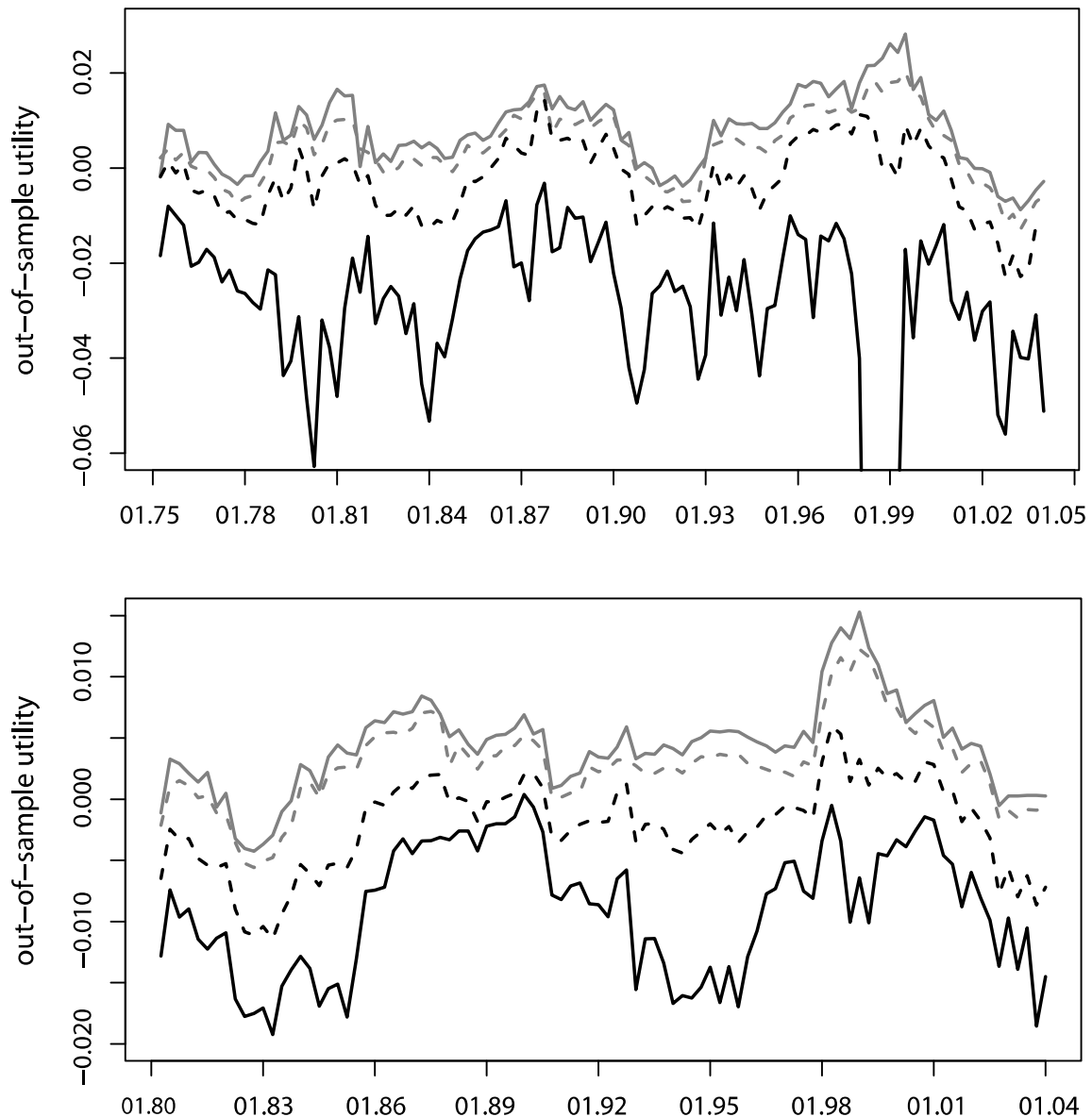


**Fig. 4** Out-of-sample expected quadratic utility as a function of the sample size of the portfolios based on: *a*) classical estimation of moments ('CL'); *b*) estimation as in Jorion (1986) ('J'); *c*) estimation as in Ledoit and Wolf (2003) ('LW'); *d*) combined estimation as *b*) and *c*) ('J & LW'). The risk aversion coefficient is equal to 10 and the risk free rate to 0.02. The number of replications is set to  $10^6$

from the normal distribution with the sample estimators taken as true values and using these returns we estimate the shrinkage portfolio weights. We estimate their moments using  $10^5$  replications.

In Fig. 5 we plot both types of the expected utilities for a holding period of three months and an estimation period of 60 and 120 months. The grey lines correspond to the utility without uncertainty in the portfolio weights and the black lines correspond to the utilities with uncertainty. The solid lines are used for the classical estimation and the dashed lines for the J & LW shrinkage estimations. We skip the other shrinkage estimators for a better visual perception; however, the results for them are similar. If the investor does not take the uncertainty in the portfolio weights into account then the classical portfolio weights outperform the shrinkage estimation over the whole estimation period. This is due to the fact that for short holding periods the estimated portfolio weights are close to the true optimal weights and, therefore, outperform other estimators in terms of the quadratic utility. However, if the uncertainty in the portfolio weights is taken into account, the classical estimators are inferior to their shrinkage counterparts. Note that there is a rather small difference between the two types of utility for shrinkage estimators. This supports the evidence from the previous subsection that the precision of these portfolio weights is high.





**Fig. 5** The expected utilities with taking the estimation of portfolio weights into account (equation (9), *black lines*) and without (equation (8), *grey lines*). *Solid lines* correspond classical estimators, *dashed lines* to Jorion and Ledoit & Wolf estimators. The holding period is set to three months, the estimation period to 60 (*above*) and 120 (*below*) months. The risk aversion coefficient is equal to 10. The number of replications is set to  $10^5$

From the above results we conclude that for moderate sample sizes (up to 100 observations) the shrinkage estimation really provides a significant improvement to portfolio selection. However, for larger samples this improvement becomes marginal. For example, for three years of weekly data (156 observations) there would be almost no improvement in terms of the MSE and the utility, but still a substantial bias. Thus we may conclude that despite of its drawbacks the classical estimation still provides acceptable results in case of larger sample sizes compared to more sophisticated estimation techniques.

## 4 Conclusions

The paper discusses different estimation procedures for portfolio selection. The standard mean-variance portfolio strategies are infeasible in practice due to the unknown parameters of the distribution of the asset returns. Here we consider the classical sample unbiased estimators as well as several shrinkage techniques. The empirical study shows that the classical approach leads to portfolio weights that are negative or larger than one with very high probability. Moreover, we cannot rely on the asymptotic results even for sample sizes commonly used in practice. The sensitivity analysis shows that the moments of the portfolio weights are very sensitive not only to shifts in the expected asset returns, but also to shifts in their covariance matrix. Moreover, upward shifts have a larger impact than downward shifts. As an alternative to the classical estimation we consider shrinkage estimators for the mean and for the covariance matrix of the asset returns. The improvement in the properties of the estimated portfolio weights due to the application of shrinkage techniques appears to be crucial only for short samples. For large sample sizes, the improvement is marginal. We also provide a technique for computing the out-of-sample utility and explain the bad performance of simple Markowitz strategies. To summarize, we conclude that for small sample sizes the current estimation procedures do not lead to satisfactory estimators of the weights and further research on new methods is required.

## References

- Barberis, N. (2000) Investing for the long run when returns are predictable. *The Journal of Finance* **55**, 225–264
- Best, M.J., Grauer, R.R. (1991) On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *The Review of Financial Studies* **4**, 315–342
- Britten-Jones, M. (1999) The sampling error in estimates of mean-variance efficient portfolio weights. *The Journal of Finance* **54**(2), 655–671
- Chopra, V.K., Ziemba, W.T. (1993) The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management*, Winter 1993, 6–11
- Efron, B., Morris, C. (1976) Families of minimax estimators of the mean of a multivariate normal distribution. *The Annals of Statistics* **4**, 11–21
- Elton, E.J., Gruber, M.J. (1999) *Investments*, MIT Press
- Fama, E.F. (1965) The behavior of stock market prices. *Journal of Business* **38**, 34–105
- Fama, E.F. (1976) *Foundations of Finance*, Basic Books, New York
- Farrell, J.L. (1997) *Portfolio Management*, McGraw-Hill
- Fleming, J., Kirby, C., Ostdiek, B. (2001) The economic value of volatility timing. *The Journal of Finance* **56**, 329–352
- Frankfurter, G.M., Phillips, H.E., Seagle, J.P. (1971) Portfolio selection: the effects of uncertain means, variances, and covariances. *The Journal of Financial and Quantitative Analysis* **6**(5), 1251–1262
- Frost, P.A., Savarino, J.E. (1988) For better performance: constrain portfolio weights. *Journal of Portfolio Management* **5**, 29–34
- Golosnoy, V., Okhrin, Y. (2006) Multivariate shrinkage for optimal portfolio weights. *The European Journal of Finance*, forthcoming
- Gourieroux, C., Monfort, A. (2005) The econometrics of efficient portfolios. *Journal of Empirical Finance* **12**, 1–41
- Ingersoll, J.E. (1987) *Theory of Financial Decision Making*, Rowman & Littlefield Publishers
- Jobson, J.D., Korkie, B. (1980) Estimation of Markowitz efficient portfolios. *Journal of the American Statistical Association* **75**(371), 544–554

- Jobson, J.D., Korkie, B. (1981) Putting Markowitz theory to work. *Journal of Portfolio Management* **7**, 70–74
- Jorion, P. (1986) Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis* **21**, 279–292
- Jorion, P. (1991) Bayesian and CAPM estimators of the means: implications for portfolio selection. *Journal of Banking and Finance* **15**, 717–727
- Kan, R., Smith, D. (2006) The distribution of the sample minimum-variance frontier. Working paper
- Kan, R., Zhou, G. (2006) Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, forthcoming
- 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
- Ledoit, O., Wolf, M. (2004) A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis* **88**, 365–411
- Markowitz, H. (1952) Portfolio selection. *The Journal of Finance* **7**, 77–91
- Mittnik, S., Rachev, S.T. (1993) Modeling asset returns with alternative stable distributions. *Econometric Reviews* **12**
- Muirhead, R.J. (1982) *Aspects of Multivariate Statistical Theory*, Wiley
- Okhrin, Y., Schmid, W. (2006a) Distributional properties of optimal portfolio weights. *Journal of Econometrics* **134**, 235–256
- Okhrin, Y., Schmid, W. (2006b) Estimation of portfolio weights, submitted for publication
- Stein, C. (1956) Inadmissibility of the usual estimator of the mean of a multivariate normal distribution. In: Neyman, J. (Ed.), *Proceedings of the Third Berkeley Symposium on Mathematical and Statistical Probability*, University of California, Berkeley, 197–206
- Zellner, A., Chetty, V.K. (1965) Prediction and decision problems in regression models from the Bayesian point of view. *Journal of the American Statistical Association* **60**, 608–615