Portfolio Optimization under Parameter Uncertainty

Daniel Niedermayer

Credit Suisse, Zurich and
University of Basel
Department of Finance
Holbeinstrasse 12
4051 Basel, Switzerland.
Tel: +41 78 666 52 30
daniel.niedermayer@unibas.ch

Portfolio Optimization under Parameter Uncertainty

Daniel Niedermayer*

September 1, 2008

Abstract

In this paper we introduce a numerical method that constructs portfolio weights while dealing with the problem of estimation error in sample means and the sample covariance matrix. We generalize the method for obtaining optimal three-fund separated portfolios proposed in Kan and Zhou (2007) in a way that allows to calculate the optimal weighting of any number of portfolios. Using this approach we construct different multi-fund separated portfolios. We show in a numerical simulation that such portfolios have higher expected out-of-sample performance than the three-fund separated portfolio and other alternative portfolios discussed in Kan and Zhou. Moreover, for low sample size where estimation error is particularly large, multi-fund separated portfolios generally outperform the global minimum variance portfolio.

*Credit Suisse, Zurich and University of Basel, Department of Finance, Holbeinstrasse 12, 4051 Basel, Switzerland. Tel: +41 78 666 52 30. Email: daniel.niedermayer@unibas.ch.
I thank seminar participants at the 41st Euro Working Group on Financial Modelling and the MFA 2008 for helpful discussions. In particlar, I thank Ferenc Niedermayer and Heinz Zimmermann for valuable comments.

I Introduction

In traditional mean-variance portfolio theory of Markowitz (1952), portfolio weights are constructed in a way which maximizes expected utility where expected returns and returns' covariance matrix are assumed to be known. However, in practice true expected returns and (co-)variances are unknown and, therefore, must be estimated. It is well known that resulting uncertainty in the input parameters used by a mean-variance optimizer magnifies estimation error and distorts weights from their optimal values (see e.g. (Michaud 1989)). In this paper we develop a flexible numerical method for portfolio construction under parameter uncertainty and show that it reduces the distorting effect of estimation error. We further show that under the assumption of normality, such portfolios lead to higher expected performance than competing methods described in Kan and Zhou (2007).

The Bayesian literature on parameter uncertainty in portfolio theory was pioneered by Zellner and Chetty (1965)[1] and applied for portfolio construction in Jorion (1986). Jorion examines a framework with known covariance matrix and an informative prior while Frost and Savarino (1986) relax the assumption of known covariance matrix. These studies generally proceed in

[1]Other early works include Barry (1974), Brown (1976), Brown (1979), Klein and Bawa (1976). For a survey on the literature see Bagasheva, Rachev, Hsu, and Fabozzi (2007).

two steps. First means (or the covariance matrix) are estimated and then used for optimizing portfolio weights.

Kan and Zhou (2007) demonstrate that the two steps procedure can be, however, suboptimal compared to methods which estimate portfolio weights directly. They show that under parameter uncertainty in means and covariances two-fund separation is no longer optimal.[2] Instead investing into three 'funds' - the riskless asset, the tangency portfolio and the global minimum variance portfolio - yields higher expected out-of-sample performance than Bayesian approaches (under a diffuse prior) and the Jorion (1986) method. While for known expected returns and covariance matrix the optimal combination of the three funds is derived analytically, under parameter uncertainty resulting portfolio weights are biased. To reduce the bias Kan and Zhou (2007) include a correction and demonstrate numerically the strong performance of the resulting portfolio.

This paper generalizes the method of Kan and Zhou (2007) from two perspectives. First, we develop a numerical simulation for optimally combining any number of funds. Second, because under parameter uncertainty the optimal weighting of funds is biased such as in Kan and Zhou (2007), we introduce a Monte Carlo method for reducing the bias. As both generaliza-

[2] A similar study is carried out by TerHorst, DeRoon, and Werker (2006) without uncertainty in the covariance matrix.

tions are flexible, they can be used to obtain optimal *multi-fund separated portfolios* while the complexity of the simulation is not affected by the structure of funds. Using empirical estimates of means and the covariance matrix we simulate the expected out-of-sample performance of differently specified multi-fund separated portfolios. The results indicate that for low sample sizes - thus, where estimation error is particularly large - some specifications of multi-fund separated portfolios outperform the three-fund separated portfolio in Kan and Zhou (2007) and the global minimum variance portfolio. (For higher sample sizes the outperformance diminishes.)

The rest of the paper is structured as follows. In section one, we set up the problem and review the results in Kan and Zhou (2007). Section two, develops a Monte Carlo method for obtaining optimal combinations of funds. As this method is biased when sample estimates are used, in the third section we introduce a numerical simulation for reducing the bias. In the fourth section we run performance tests on different data and report the results.

II The Problem

Consider a universe of one riskless asset and N risky assets. Assume excess returns (the returns of the N assets over the riskless return) follow a multivariate normal distribution $\boldsymbol{r}_t \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\mu}$ is an N-vector of expected excess returns and $\boldsymbol{\Sigma}$ is a positive definite $(N \times N)$ covariance matrix. An investor with relative risk aversion γ maximizes a quadratic objective function

$$U(\boldsymbol{w}, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \gamma) = \boldsymbol{w'\mu} - \frac{\gamma}{2}\boldsymbol{w'\Sigma w}\,, \tag{1}$$

by choosing an N-vector of portfolio weights $\boldsymbol{w}$. In this formulation weights do not have to sum to one as remaining wealth can be invested into the riskless asset. As no short selling constraints are imposed the sum of portfolio weights can also exceed unity. Note that a utility below zero implies that the investor would be better off investing solely into the riskless asset (which yields a utility of zero).

The optimal portfolio weight maximizing (1) is given by

$$\boldsymbol{w}^* = \frac{1}{\gamma}\boldsymbol{\Sigma^{-1}\mu}\,. \tag{2}$$

The basic difficulty in (1) and (2) is that the input variables $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are theoretical moments not known to investors. Because estimates of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are plugged into (2) the problem in (1) becomes stochastic due to randomness in the estimated parameters. While the traditional approach using the maximum likelihood estimators $\hat{\boldsymbol{\mu}} = 1/T\sum_{t=1}^{T}\boldsymbol{r_t}$ and $\hat{\boldsymbol{\Sigma}} = 1/T\sum_{t=1}^{T}(\boldsymbol{r}_t - \hat{\boldsymbol{\mu}})(\boldsymbol{r}_t - \hat{\boldsymbol{\mu}})'$ in (2) provides good results when the number of return observations T is large, it fails for small values of T (and large values of N).

Following Stein (1955), James and Stein (1961) construct an alternative estimator and show that it has lower expected loss than the maximum likelihood estimator. Implementing the result in James and Stein (1961) in a Bayesian framework (with an informative prior) Jorion (1986) derives a shrinkage estimator of sample means and shows that it outperforms the traditional approach leading to higher expected performance in (1). While the covariance matrix is assumed to be known in Jorion (1986), other work has been done on the estimation error reduction of the sample covariance matrix (e.g. Ledoit and Wolf (2003)) or the joint estimation of sample means and the covariance matrix (e.g. Frost and Savarino (1986)). While for a long return history expected performance is more sensitive to expected return estimation than to the estimation of the covariance matrix (compare e.g. Merton (1980) and Chopra and Ziemba (1993)), estimation error in the covariance matrix is not negligible when the number of observations is low and the number of assets is high (Kan and Zhou (2007)). For this reason we will not assume the covariance matrix to be known.

More recently, Kan and Zhou (2007) introduce an interesting technique for constructing portfolios that account for estimation error in expected returns and the covariance matrix. They show that an optimal combination of three funds, i.e. the riskless asset, tangency portfolio and the global minimum variance portfolio increases expected performance compared to the maximum likelihood estimator, a Bayesian estimator under a diffuse prior and an optimal two-fund portfolio strategy where the riskless asset is combined with the tangency portfolio. The portfolio weight $\hat{\boldsymbol{w}}$ in Kan and Zhou is given by

$$\hat{\boldsymbol{w}} = \frac{1}{\gamma}\left(c\hat{\boldsymbol{\Sigma}}^{-\mathbf{1}}\hat{\boldsymbol{\mu}} + d\hat{\boldsymbol{\Sigma}}^{-\mathbf{1}}\mathbf{1}\right), \tag{3}$$

where $\mathbf{1}$ is an N-vector of ones and c and d are scalars that we shall call *weight multipliers*. According to (3) the portfolio weight $\hat{\boldsymbol{w}}$ is given by the combination of two funds, i.e. the estimated tangency portfolio ($\hat{\boldsymbol{\Sigma}}^{-\mathbf{1}}\hat{\boldsymbol{\mu}}$) and the estimated global minimum variance portfolio ($\hat{\boldsymbol{\Sigma}}^{-\mathbf{1}}\mathbf{1}$). The remaining wealth is invested into the riskless asset. More generally, we will call a function that weights funds a *portfolio strategy*.

For known $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, Kan and Zhou show analytically that (3) yields higher expected performance (for $T > N + 4$) than competing methods.[3] The optimal (theoretical) weight multipliers are obtained by differentiating the expected performance of $\hat{\boldsymbol{w}}$ with respect to c and d. The solution has

[3] They compare the three-fund portfolio strategy with different strategies such as a Bayesian approach (with diffuse prior), a portfolio constructed by shrinkage (Jorion 1986) and a strategy with uncertainty aversion ((Garlappi, Uppal, and Wang 2007)).

the form

$$\begin{bmatrix} c^{**} \\ d^{**} \end{bmatrix} = f_{3F}\left(T, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right) , \tag{4}$$

where the subscript '$3F$' stands for three-fund portfolio strategy.[4]

When looking at (3) and (4) it is important to distinguish between the theoretical values $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ and estimated values $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$. In order to get an intuition for the working of (3) and (4) consider the following thought experiment. Assume there is an *advisor* and an *investor*. The advisor knows $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ and calculates the optimal weight multipliers c^{**}, d^{**} in (4) and gives them to the investor. Assume that the investor does not know the true values $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ but only the estimates $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$. Then the investor can maximize expected performance by constructing a portfolio weight

$$\hat{\boldsymbol{w}}^{**} = \frac{1}{\gamma}\left(c^{**}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}} + d^{**}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}\right) .$$

As in practice $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are unknown, the optimal multipliers c^{**} and d^{**} have to be estimated. Kan and Zhou (2007) replace $\boldsymbol{\mu}$ by $\hat{\boldsymbol{\mu}}$ and $\boldsymbol{\Sigma}$ by $\hat{\boldsymbol{\Sigma}}$ in (4) and include a correction to reduce the bias for low values of T. This allows for writing the estimated three-fund weight multipliers as a function[5] of the observed moments $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$

$$\begin{bmatrix} \hat{c}^{**} \\ \hat{d}^{**} \end{bmatrix} = g_{3F}\left(T, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}\right) . \tag{5}$$

[4] The function $f_{3F}(\cdot)$ can be derived by using some properties of the Wishart distribution described in e.g. Haff (1979) and Muirhead (1982).

[5] The exact form is provided in the appendix in (A-1) and (A-2).

The above equation is practically relevant. In the investor-advisor example we can thus drop the (unrealistic) assumption of an advisor knowing the true values $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Instead, the investor must rely on the estimates $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ and forms the portfolio according to

$$\hat{\boldsymbol{w}}^{**} = \frac{1}{\gamma}\left(\hat{c}^{**}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}} + \hat{d}^{**}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}\right) .$$

Applying the estimated values $\hat{c}^{**}$ and $\hat{d}^{**}$ in a numerical simulation, Kan and Zhou show that the estimated three-fund portfolio strategy outperforms competing strategies with respect to generating higher expected performance in (1) (for e.g. $N = 25$ and $T \geq 120$).[6]

It is, however, difficult to generalize the method of Kan and Zhou. For complicated portfolio strategies the calculation of the optimal weight multipliers (in (4)) and their estimation (in (5)) becomes non-trivial. In this paper we follow a numerical approach that avoids both difficulties. Using a simple Monte Carlo simulation we can obtain solutions for the theoretical weight multipliers and more importantly, also good approximations for estimated weight multipliers. The advantage of the numerical method is that it allows for easy generalization of the three-fund portfolio strategy to any portfolio strategy given by a function of $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$ and T. In the following we

[6] For shorter observation length investing into the global minimum variance portfolio sometimes yields higher expected performance than the three-fund strategy.

introduce a portfolio strategy with multiple funds and show how to obtain numerical solutions for optimal weight multipliers.

III Multi-Fund Separation

An important feature of the traditional Markowitz portfolio theory is the two-fund separation theorem; the combination of two 'portfolios', the riskless asset and the tangency portfolio, is an efficient portfolio. Hence, in order to maximize expected performance investors choose a combination of these two portfolios.

When including parameter uncertainty into the model, Kan and Zhou (2007) demonstrate that optimally combining the global minimum variance portfolio with the riskless asset and the tangency portfolio yields higher expected performance than two-fund separation. In this section we generalize this result to multi-fund separation and show that under parameter uncertainty combining more than three funds, by e.g. including the equally weighted portfolio, can increase expected performance even further.

The three-fund portfolio strategy in (3) can be written more generally as a multi-fund portfolio strategy given by the linear combination of M funds

$$\hat{\boldsymbol{w}}_{\text{MF}} = \frac{1}{\gamma} \sum_{i=1}^{M} c_i \hat{\boldsymbol{q}}_{\boldsymbol{i}} \,, \tag{6}$$

where c_i are scalar weight multipliers and $\boldsymbol{q_i}$ are N-vectors containing the weights of fund i. Note that (6) contains three-fund separation in (3) as special case $M = 2$, $\hat{\boldsymbol{q}}_{\mathbf{1}} = \hat{\boldsymbol{\Sigma}}^{-\mathbf{1}}\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{q}}_{\mathbf{2}} = \hat{\boldsymbol{\Sigma}}^{-\mathbf{1}}\mathbf{1}$.

Maximizing the expected performance in (1) with the weights in (6) yields the solution of the optimal weight multipliers[7]

$$\begin{aligned} \boldsymbol{c}^{\boldsymbol{*}} &= f_{MF}(T, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ &= \arg\max_{\boldsymbol{c}} \mathrm{E}U\left(\hat{\boldsymbol{w}}_{\mathrm{MF}}(\boldsymbol{c}), \boldsymbol{\mu}, \boldsymbol{\Sigma}\right) \qquad (7) \\ &= \left[\mathrm{E}\left(\hat{\boldsymbol{Q}}\boldsymbol{\Sigma}\hat{\boldsymbol{Q}}'\right)\right]^{-1} \mathrm{E}\left(\hat{\boldsymbol{Q}}\right)\boldsymbol{\mu}\,, \qquad (8) \end{aligned}$$

with an $(M \times N)$ matrix $\hat{\boldsymbol{Q}} = (\hat{\boldsymbol{q}}_{\mathbf{1}}\ \hat{\boldsymbol{q}}_{\mathbf{2}}\ \cdots\ \hat{\boldsymbol{q}}_{\boldsymbol{M}})'$ and an N-vector $\boldsymbol{c}^{\boldsymbol{*}} = (c_1^*\ c_2^*\ \cdots\ c_M^*)'$. Note that (8) is deterministic and can be numerically simulated as follows.

We draw K realizations $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ according to

$$\boldsymbol{r}_t^{(k)} \sim \mathcal{N}\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right), \qquad t = 1 \ldots T$$

$$\hat{\boldsymbol{\mu}}^{(k)} = \tfrac{1}{T}\textstyle\sum_{t=1}^{T} \boldsymbol{r}_t^{(k)}, \qquad \hat{\boldsymbol{\Sigma}}^{(k)} = \tfrac{1}{T-1}\textstyle\sum_{t=1}^{T}\left(\boldsymbol{r}_t^{(k)} - \hat{\boldsymbol{\mu}}^{(k)}\right)\left(\boldsymbol{r}_t^{(k)} - \hat{\boldsymbol{\mu}}^{(k)}\right)' \quad (9)$$

Choosing the number of iterations, K, sufficiently high gives a good approximation for the optimal weight multipliers in (8)

$$\boldsymbol{c}^{\boldsymbol{*}} \approx \left[\sum_{k=1}^{K} \hat{\boldsymbol{Q}}^{(k)}\boldsymbol{\Sigma}\hat{\boldsymbol{Q}}^{(k)'}\right]^{-1} \sum_{k=1}^{K} \hat{\boldsymbol{Q}}^{(k)}\boldsymbol{\mu}\,, \qquad (10)$$

[7]Note that due to the stochastic nature of $\hat{\boldsymbol{Q}}$ there is no restriction $M \leq N$ since $\mathrm{E}(\hat{\boldsymbol{Q}}\boldsymbol{\Sigma}\hat{\boldsymbol{Q}}')$ is positive definite.

where $\hat{\boldsymbol{Q}}^{(k)}$ contains the kth draw $\hat{\boldsymbol{\mu}}^{(k)}$ and $\hat{\boldsymbol{\Sigma}}^{(k)}$. In the investor-advisor example where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are known to the advisor, the advisor can obtain a good approximation for $\boldsymbol{c}^*$ by (10).

For $\hat{\boldsymbol{Q}} = (\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}} \;\; \hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1})'$, figure 1 shows how the numerically simulated values c_1^*, c_2^* from (10) approach their optimal values c^{**}, d^{**} analytically derived in Kan and Zhou.[8] For the example in figure 1, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are replaced by estimates using the 25 Fama-French size and book-to-market ranked, monthly portfolio returns from 1987.1 - 2006.12.

As in the numerical method suggested in this paper the complexity of the solution is not affected by the number of funds M, it is easy to include more than two funds into $\hat{\boldsymbol{Q}}$. More generally, we can write

$$\hat{\boldsymbol{Q}} = \begin{bmatrix} \hat{\boldsymbol{q}}_1' \\ \hat{\boldsymbol{q}}_2' \\ \vdots \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\mu}}'\hat{\boldsymbol{\Sigma}}^{-1} \\ \mathbf{1}'\hat{\boldsymbol{\Sigma}}^{-1} \\ \vdots \end{bmatrix}. \tag{11}$$

Constructing the portfolio according to (6), (10) and (11) ensures that the resulting expected performance in (1) is at least as high as for the three-fund separation case

$$\mathrm{E}(U(\hat{\boldsymbol{w}}^*_{\mathrm{MF}}, \boldsymbol{\mu}, \boldsymbol{\Sigma})) \geq \mathrm{E}(U(\hat{\boldsymbol{w}}^{**}, \boldsymbol{\mu}, \boldsymbol{\Sigma}))\,.$$

This result is clear because (11) contains three-fund separation as a special

[8]See eqs. (59) and (60) in Kan and Zhou (2007).

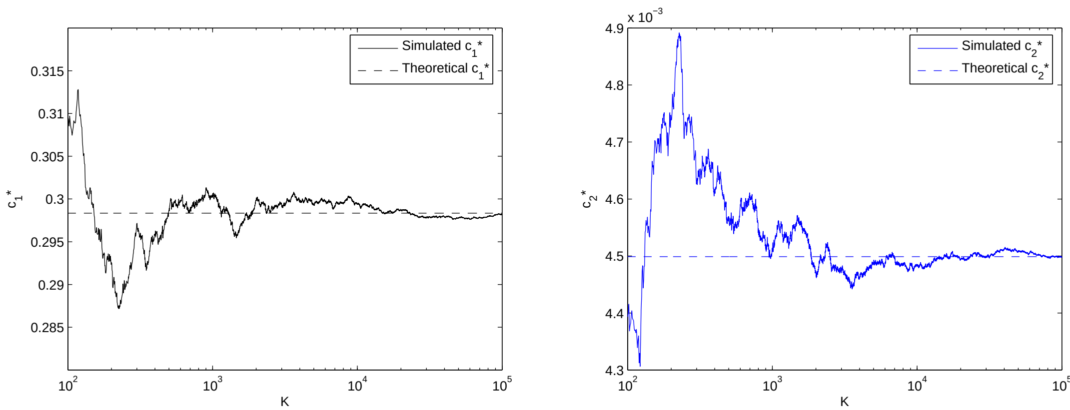

Figure 1: The analytically derived optimal weight multipliers in Kan and Zhou can be obtained by the Monte Carlo simulation given in (10). For this simulation data from the Fama-French 25 size and book-to-market ranked portfolio returns for the period 1987-2006 are used.

case where $c_i = 0$, $i = 3 \dots M$.

IV Estimating Weight Multipliers

The previous section has shown how to numerically approximate optimal weight multipliers by Monte Carlo simulation. For known $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ (i.e. known to the advisor) the Kan and Zhou (2007) results are generalized to more than three funds with an expected performance at least as high as the three-fund portfolio strategy.

However, the optimal expected performance cannot be achieved in practice as $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are unknown. When replacing $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ with $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ in (9) and (10), resulting weight multipliers will be biased. For a low sample size T the bias is particularly strong because of large estimation error in $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$. In order to reduce the bias, Kan and Zhou use an analytical correction. In this paper we obtain estimated weight multipliers by Monte Carlo simulation.

We proceed in three steps. First, we will assume that $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are drawn from a known multivariate normal distribution. Knowing this distribution allows for simulating the distribution of the optimal weight multipliers conditioned on a specific realization of $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$. In the second subsection we will show that even without the knowledge of the distribution of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ the conditional distribution of optimal multipliers can be quite well approximated. Finally we incorporate the information from the estimated weight multiplier distribution into a portfolio strategy.

IV.A The Conditional Distribution of Optimal Weight Multipliers

Assume that $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are drawn[9] K times from a known grand mean $\boldsymbol{\mu}^*$ and covariance matrix $\boldsymbol{\Sigma}^*$. Then for each realization of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ the optimal weight multipliers are calculated according to the simulation in (10). Next, for each realization of $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ we draw T number of return vectors from $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and calculate their mean $\hat{\boldsymbol{\mu}}$ and covariance matrix $\hat{\boldsymbol{\Sigma}}$. The

[9] According to (9) with an observation length T.

procedure is depicted in table 2.

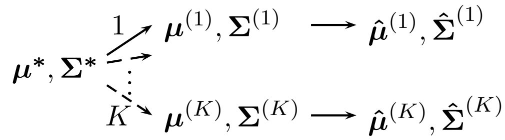

Figure 2: First, the moments $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ are drawn from known parameters $\boldsymbol{\mu}^{*}$, $\boldsymbol{\Sigma}^{*}$ (according to (9)). Then from each realization of $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ sample estimates $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are drawn the same way.

For each of the K realizations of $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ optimal weight multipliers are calculated according to (10). Moreover, for each realization of $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$ we store a scalar $\hat{s}$ given by a function of $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ and which is correlated with the optimal weight multipliers $\boldsymbol{c}^{*}$. The value $\hat{s}$ can be interpreted as a noisy signal of $\boldsymbol{c}^{*}$.

For the following experiment we specify the signal as[10]

$$\hat{s} = \hat{\boldsymbol{\mu}}'\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}} - \frac{\left(\hat{\boldsymbol{\mu}}'\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}\right)^2}{\mathbf{1}'\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}} \tag{12}$$

and the matrix of funds as $\boldsymbol{Q}' = [\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}} \;\; \hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}]$ which corresponds to the three-fund separated portfolio in Kan and Zhou. Each realization of $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ yields two optimal weight multipliers $\boldsymbol{c}^{*} = [c_1^{*} \;\; c_2^{*}]'$.[11] Figure 3 shows the relation between theoretically optimal weight multipliers $c_{1,i}^{*}$ and signals $\hat{s}_i$, $i = 1 \ldots K$. (For this simulation $\boldsymbol{\mu}^{*}$, $\boldsymbol{\Sigma}^{*}$ are replaced by estimates from

[10] This corresponds to the variable $\hat{\psi}^2$ in Kan and Zhou (2007, p.643).

[11] As for the three-fund separation the closed form solution is derived in Kan and Zhou, here there is no need to apply the numeric simulation in (10). We will use (10) for a more general portfolio strategy in the next section.

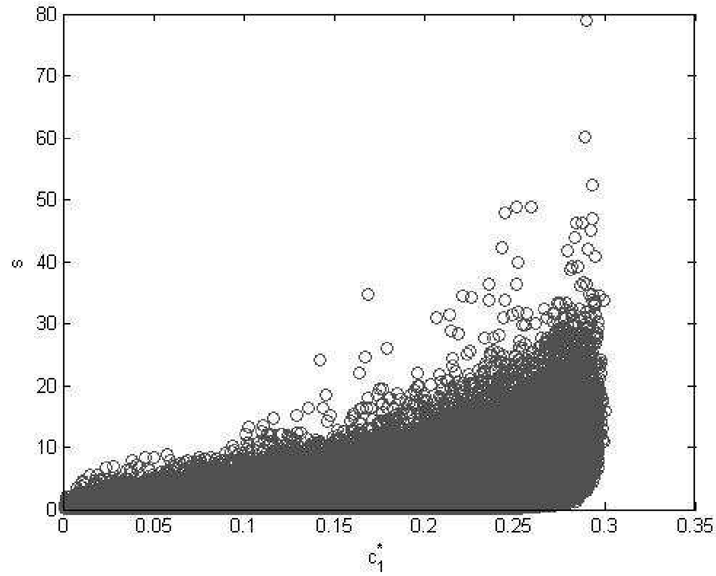

Figure 3: This figure shows the relation of theoretically optimal weight multipliers c_1^* and signals $\hat{s}$. The different values of c_1^* are obtained by drawing theoretical moments $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ from grand moments $\boldsymbol{\mu}^*$ and $\boldsymbol{\Sigma}^*$. The signals are given by (12) where $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are drawn from $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.

the 6 Fama-French portfolios for 1987.1-2006.12. Simulations with different data provide similar results.)

High values of the signal are generally associated with high values of multipliers. While for low values of $\hat{s}$ the possible multipliers c_1^* range from 0.0 to 0.3, for high values of $\hat{s}$ this range becomes narrower. In order to see the distribution of optimal weight multipliers conditioned on a realized signal $\hat{s}$ we define a vertical range $\hat{s} \pm \epsilon$. The optimal weight multipliers for which the corresponding signals lie within the range $\hat{s} \pm \epsilon$ have the conditional

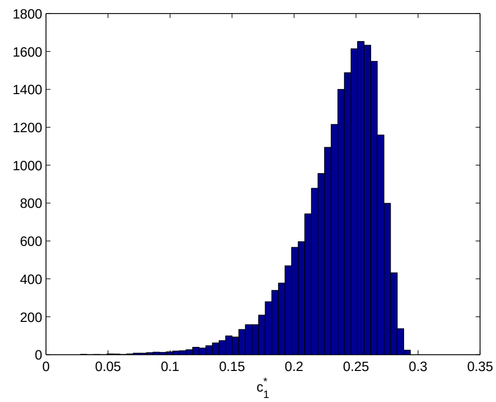

Figure 4: This figure shows the distribution of optimal weight multipliers conditioned on a signal $\hat{s} \in (5.0 - \epsilon, 5.0 + \epsilon)$.

distribution depicted in figure 4.[12]

Hence, given a realization of $\hat{s} \in (5.0 - \epsilon; 5.0 + \epsilon)$ it is likely that the optimal multiplier lies within the range of 0.15 and 0.28, with mean 0.234 and median 0.241. One can compare these results with the optimal weight multipliers obtained by Kan and Zhou. Therefore, for each realization of $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$ where the resulting signal is within the range, we calculate the unadjusted estimate $f_{3F}(T, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$ by simply plugging the sample moments in (4) and the bias corrected estimate $g_{3F}(T, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$ (in (5)). Interestingly, all unadjusted estimates lie in a range of 0.286 and 0.287 while all adjusted estimates are rounded to 0.273.

This shows that for the chosen prior distribution of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ with grand

[12] Here, ϵ should not be too large as this would widen the conditional distribution and not too small because this lowers the statistics obtained for the conditional distribution.

moments $\boldsymbol{\mu}^*$ and $\boldsymbol{\Sigma}^*$, both the unadjusted and adjusted estimates overestimate the average optimal weight multipliers. Moreover, the Kan and Zhou correction indeed reduces the bias by producing values closer to the true mean of 0.234.

Note that this method also allows for conditioning the multiplier distribution on multiple signals, preferably uncorrelated with each other and strongly correlated with optimal multipliers. While a multi-signal simulation has been implemented, there was no considerable improvement in the results because at least one of the two above properties were not given.[13]

IV.B Estimating the Conditional Distribution

In the above simulation of the conditional weight multiplier distribution, it was assumed that the grand moments $\boldsymbol{\mu}^*$, $\boldsymbol{\Sigma}^*$ and the realizations $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ were known. As however only $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are observed by an investor, the conditional distribution is unknown. In order to estimate the relation between the signal $\hat{s}$ and the optimal weight multiplier we replace the grand moments $\boldsymbol{\mu}^*$, $\boldsymbol{\Sigma}^*$ by $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$. Figure 5 shows the effect of this replacement.

While the circles depict the theoretical signal-multiplier relation (the same as reported in figure 3), the crosses show this relation when $\boldsymbol{\mu}^*$, $\boldsymbol{\Sigma}^*$ are replaced by some realization of $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$. As the unadjusted and adjusted es-

[13] As a signal is given by a function of $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ this observation is not surprising.

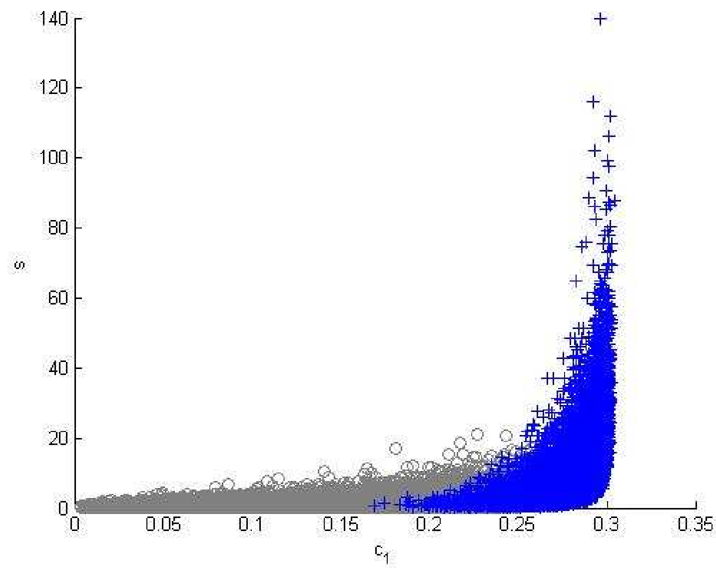

Figure 5: Replacing the grand moments $\boldsymbol{\mu}^*$ and $\boldsymbol{\Sigma}^*$ by $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ leads to a shift of the theoretic distribution (circles) to the estimated distribution (crosses).

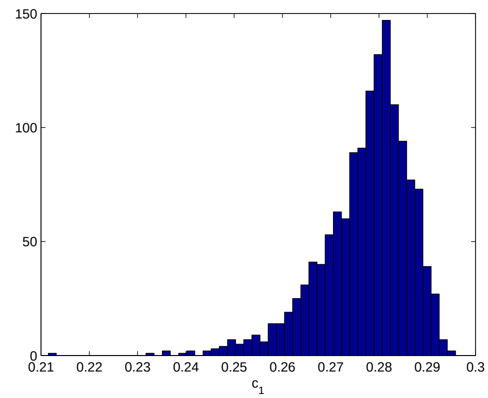

Figure 6: This figure shows the estimated optimal weight multiplier for signals $\hat{s} \in (5.0 - \epsilon, 5.0 + \epsilon)$ and where grand moments are replaced by estimated moments $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$.

timators of Kan and Zhou, the estimated signal-multiplier relation in figure 5 is upward biased. For low values of $\hat{s}$ theoretically optimal weight multipliers are here lower than the estimated ones. For the signal $\hat{s} \in (5.0-\epsilon, 5.0+\epsilon)$ the estimated conditional distribution of weight multipliers is shown in figure 6.

It is noteworthy that the weight multipliers for this signal no longer range from 0.07 to 0.29 but from 0.21 to 0.3, with a mean of 0.277 (and median of 0.279). As this value lies above the true mean of 0.234, it is upward biased such as the Kan and Zhou unadjusted (0.287) and adjusted mean (0.273). Note that the aim is, however, not to reduce the bias in the mean but eventually to have higher expected performance.

It is important to see that this estimation procedure is more general because it is not restricted to the Kan and Zhou three-fund separated portfolio case. For any portfolio strategy given by a function of $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ - and the corresponding matrix $\hat{\boldsymbol{Q}}$ - it provides estimates of the signal-conditioned weight multiplier distribution.

IV.C Feasible Portfolio Strategy

The simulation results in the previous subsections can be integrated into a portfolio strategy which is feasible - thus, does not have any knowledge of the true moments $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$.

Assume an investor observes a return history of size T and calculates sample moments $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$. The objective is to find a portfolio weight given by

$$\hat{\boldsymbol{w}} = w(T, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$$

that maximizes expected out-of-sample performance. This paper suggests a portfolio strategy that proceeds in five steps.

1. Specify the funds $\hat{\boldsymbol{q}}_i$ $(i = 1 \ldots M)$ in $\hat{\boldsymbol{Q}}$.

2. Replace grand moments $\boldsymbol{\mu}^*$, $\boldsymbol{\Sigma}^*$ by $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$ and run the simulation shown in figure 2. For each $\boldsymbol{\mu}^{(k)}$, $\boldsymbol{\Sigma}^{(k)}$ $(k = 1 \ldots K)$ calculate the optimal weight multipliers using the Monte Carlo method in (10). As for each realization[14] of $\boldsymbol{\mu}^{(k)}$, $\boldsymbol{\Sigma}^{(k)}$ in the simulation of 2 we further draw a

[14] Note that $\boldsymbol{\mu}^{(k)}$, $\boldsymbol{\Sigma}^{(k)}$ are moments drawn from $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ in the simulation of figure 2. In contrast, $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are the original inputs required by the portfolio strategy.

$\hat{\boldsymbol{\mu}}^{(k)}$, $\hat{\boldsymbol{\Sigma}}^{(k)}$, we can calculate a signal $\hat{s}^{(k)} = s(\hat{\boldsymbol{\mu}}^{(k)}, \hat{\boldsymbol{\Sigma}}^{(k)})$ given in (12). This procedure leads to a plot showing the relation between weight multipliers and signals (such as the crosses in figure 5).

3. Calculate the signal $\hat{s} = s(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$. Collect all weight multipliers from the weight multipliers-signal plot of step 2 where the corresponding signal is in the range $(\hat{s} - \epsilon, \hat{s} + \epsilon)$. This yields one signal-conditioned weight multiplier distribution for each fund specified in $\hat{\boldsymbol{Q}}$.

4. Obtain the weight multipliers $\hat{\boldsymbol{c}}^{*}$ by taking the median[15] of each signal-conditioned distribution.

5. Calculate the weights according to

$$\hat{\boldsymbol{w}}_{MF} = \frac{1}{\gamma} \sum_{i=1}^{M} \hat{c}_i^{*} \hat{\boldsymbol{q}}_i \,.$$

As it is not clear how well this procedure performs it has to be numerically tested. In this case, the proof of the pudding is in the eating!

V Performance Test

In order to assess the expected out-of-sample performance of different portfolio strategies we run a Monte Carlo simulation using $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ as input parameters. Here, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ can be arbitrarily generated (while making

[15] While using the mean produces good results as well we prefer the median since it is not sensitive to outliers.

sure the covariance matrix is positive definite). For practical purposes it is interesting to use empirically 'reasonable' values for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. In this paper we will use empirical estimates derived from a) Fama-French 6 size and book-to-market ratio ranked portfolios and b) Fama-French 25 size and book-to-market ratio ranked portfolios. Both data sets cover the period from 1987.1 to 2006.12.[16]

A simulation of a portfolio strategy's expected out-of-sample performance can be specified as follows.

Step 1. Draw $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ from

$$\begin{aligned}
\boldsymbol{r}_t^{(j)} &\sim \mathcal{N}\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right), \qquad t = 1 \ldots T \\
\hat{\boldsymbol{\mu}}^{(j)} &= \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{r}_t^{(j)}, \\
\hat{\boldsymbol{\Sigma}}^{(j)} &= \frac{1}{T-1} \sum_{t=1}^{T} \left(\boldsymbol{r}_t^{(j)} - \hat{\boldsymbol{\mu}}^{(j)}\right) \left(\boldsymbol{r}_t^{(j)} - \hat{\boldsymbol{\mu}}^{(j)}\right)', \qquad (13)
\end{aligned}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are known.

Step 2. Calculate the performance of a portfolio strategy

$$U^{(j)} = U\left(\boldsymbol{w}^{(j)}, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \gamma\right)$$

where $\boldsymbol{w}^{(j)} = w(T, \hat{\boldsymbol{\mu}}^{(j)}, \hat{\boldsymbol{\Sigma}}^{(j)})$ is the weight vector determined by some portfolio strategy.

[16] I am thankful to Kenneth French for making this data available on his website.

Step 3. For $j = 1 \ldots J$ repeat steps 1 to 2.

Step 4. Estimate the expected out-of-sample performance by

$$\bar{U} = \frac{1}{J} \sum_{j=1}^{J} U^{(j)} .$$

It is important to see that in step 2 the true moments $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are used for the calculation of $U^{(j)}$ whereas the weight vector $\boldsymbol{w}^{(j)}$ is determined by a historical sample. This ensures that the investor does not know the true moments.

V.A Specifications

For the multi-fund portfolio strategies presented in this paper we specify the matrix funds as

$$\hat{\boldsymbol{Q}} = \begin{bmatrix} \hat{\boldsymbol{\mu}}'\hat{\boldsymbol{\Sigma}}^{-1} \\ \mathbf{1}'\hat{\boldsymbol{\Sigma}}^{-1} \\ \mathrm{tr}(\hat{\boldsymbol{\Sigma}}^{-1})\mathbf{1} \end{bmatrix} , \tag{14}$$

where the third fund $\mathrm{tr}(\boldsymbol{\Sigma}^{-1})\mathbf{1}$ is a (stochastic) weight proportional to the equally weighted portfolio.

Although there are infinite many ways of constructing $\hat{\boldsymbol{Q}}$, this specification provided the most appealing results among those that were tested.[17] The specification in (14) is a four-fund separated portfolio (the riskless asset and the three weights in $\hat{\boldsymbol{Q}}$). It is also of interest how subsets of the

[17] Other specifications included $\mathbf{1}$, $\mathrm{tr}(\hat{\boldsymbol{\Sigma}})$ and $|\hat{\boldsymbol{\mu}}|$.

funds defined in (14) perform. Therefore, we number the funds in $\hat{\boldsymbol{Q}}$ as 1 for $\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}}$, 2 for $\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}$ and 3 for $\mathrm{tr}(\hat{\boldsymbol{\Sigma}}^{-1})\mathbf{1}$. Accordingly, the expected performance of the sub-specification of (14) containing e.g. fund 1 and 3 will be denoted as $\mathrm{E}(U_{MF}(1,3))$. This way we obtain $2^3 = 8$ possible multi-fund separated portfolio strategies.[18]

In the following we compare different sub-specifications of $\hat{\boldsymbol{Q}}$ with 5 different portfolio strategies used in Kan and Zhou (2007): the tangency portfolio, the global minimum variance portfolio, the portfolio of Jorion (1986) and the two- and three-fund separated portfolio strategies of Kan and Zhou.[19]

V.B Results

Table 1 reports simulation results where means and the covariance matrix are estimated from the Fama-French 6 size and book-to-market ratio ranked portfolios covering the period 1987.1-2006.12. In table 2 the expected performances using the 25 Fama-French portfolios for the same period are shown. The risk aversion coefficient γ was set to 1. This is, however, without loss of generality.[20]

[18] A specification where $\hat{\boldsymbol{Q}}$ is empty is allowed. It corresponds to investing solely into the riskless asset.

[19] The exact formulation of these strategies is provided in the appendix.

[20] The risk aversion here merely re-scales weights and expected performance, because when $\tilde{\boldsymbol{w}} \equiv \gamma\hat{\boldsymbol{w}} \longrightarrow \tilde{U} \equiv \gamma U(\hat{\boldsymbol{w}}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \tilde{\boldsymbol{w}}'\boldsymbol{\mu} - 1/2\tilde{\boldsymbol{w}}'\boldsymbol{\Sigma}\tilde{\boldsymbol{w}}$; thus, $\tilde{U}$ is independent from γ. This is only true as in our notation weights do not some to one and remaining wealth is

The first row contains the expected performance of the parameter certainty optimal portfolio. This performance can only be achieved when the true values of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are known (or where the sample size is very long and estimation error is negligible). These expected performances are reported as a benchmark. Note that negative numbers in tables 1 and 2 indicate that an investor would be better off investing into the riskless asset yielding an expected performance of 0. The highest values for each observation length T are highlighted in bold letters and standard deviations are reported in brackets.

The major simulation results in Kan and Zhou (2007) are supported by the two tables. The three-fund separated portfolio dominates the two-fund separated portfolio and the Jorion (1986) portfolio.

From tables 1 and 2 one can observe that $\mathrm{E}(U_{MF}(1,2,3))$ yields the highest expected performance (except for $T = 18$ in table 1 where the global minimum variance portfolio performs better). Moreover, the performance of $\mathrm{E}(U_{MF}(1,2))$ is noteworthy as it generally outperforms the Kan and Zhou (2007) three-fund separated portfolio. This is interesting as $\mathrm{E}(U_{MF}(1,2))$ is equivalent to the three-fund portfolio strategy when true moments $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are known. The difference in both portfolio strategies stems only from the

invested in to the riskless asset.

estimation of optimal weight multipliers. As the estimation method developed in section IV is not designed specifically to estimate weight multipliers of the three-fund portfolio strategy, it is surprising that despite its generality it seems to dominate the Kan and Zhou estimation with regard to yielding higher expected performance.

While $\mathrm{E}(U_{MF}(2))$, $\mathrm{E}(U_{MF}(2,3))$ and $\mathrm{E}(U_{MF}(3))$ do not perform particularly well, it is interesting to see that $\mathrm{E}(U_{MF}(3))$ is robust for low sample sizes, yielding positive performances with low standard deviations. As it does not contain the tangency portfolio ($\hat{\boldsymbol{\Sigma}}^{\mathbf{-1}}\hat{\boldsymbol{\mu}}$) as a fund, it does not perform well for higher sample size where estimation error is small. However, when mixing the fourth fund ($\mathrm{tr}(\hat{\boldsymbol{\Sigma}}^{\mathbf{-1}})\mathbf{1}$) with the first three funds (leading to $\mathrm{E}(U_{MF}(1,2,3))$), this increases expected performance for low sample sizes and does not harm when sample size is high.

VI Conclusion

Under parameter certainty traditional portfolio theory suggest two-fund separation - thus, combining the riskless asset with the tangency portfolio. However, under uncertainty in means and the covariance matrix, Kan and Zhou (2007) show that two-fund separation is no longer optimal. Instead, optimally combining three funds - the riskless asset, tangency portfolio and

Portfolio Strategy	$T = 12$	$T = 18$	$T = 24$	$T = 36$
Parameter Certainty Optimal Pf.	0.150	0.150	0.150	0.150
	(0.0)	(0.0)	(0.0)	(0.0)
Tangency Portfolio	-14.053	-1.229	-0.397	-0.089
	(106.669)	(2.771)	(0.643)	(0.242)
Global Minimum Variance Pf.	-0.021	**0.036**	0.053	0.071
	(0.768)	(0.109)	(0.060)	(0.047)
Jorion Portfolio	-1.125	-0.117	0.000	0.057
	(11.552)	(0.639)	(0.190)	(0.097)
Kan-Zhou 2 Fund Pf.	-0.038	0.010	0.038	0.061
	(0.734)	(0.177)	(0.069)	(0.051)
Kan-Zhou 3 Fund Pf.	-0.033	0.020	0.051	0.075
	(0.792)	(0.196)	(0.080)	(0.057)
E($U_{MF}(1,2)$)	-0.024	0.026	0.055	**0.080**
	(0.688)	(0.184)	(0.090)	(0.063)
E($U_{MF}(1,2,3)$)	**-0.016**	0.027	**0.058**	**0.080**
	(0.493)	(0.178)	(0.082)	(0.057)
E($U_{MF}(2)$)	-0.041	-0.014	0.006	0.043
	(1.004)	(0.274)	(0.165)	(0.093)
E($U_{MF}(2,3)$)	-0.035	-0.010	0.011	0.042
	(0.985)	(0.263)	(0.154)	(0.088)
E($U_{MF}(3)$)	**0.009**	0.023	0.028	0.030
	(0.005)	(0.004)	(0.003)	(0.004)

Table 1: This table reports simulated expected out-of-sample performances using the Fama-French 6 size and book-to-market ratio ranked portfolios for the period 1987.1.-2006.12. Numbers in brackets are standard deviations.

Portfolio Strategy	$T = 30$	$T = 45$	$T = 60$	$T = 120$
Parameter Certainty Optimal Pf.	0.413	0.413	0.413	0.413
	(0.0)	(0.0)	(0.0)	(0.0)
Tangency Portfolio	-615.731	-8.277	-1.964	-0.051
	(991.948)	(9.351)	(1.580)	(0.221)
Global Minimum Variance Pf.	0.020	0.084	0.138	0.195
	(0.021)	(0.071)	(0.047)	(0.036)
Jorion Portfolio	-3.980	-0.396	-0.005	0.227
	(7.994)	(1.039)	(0.296)	(0.078)
Kan-Zhou 2 Fund Pf.	0.015	0.077	0.145	0.251
	(0.020)	(0.076)	(0.065)	(0.041)
Kan-Zhou 3 Fund Pf.	0.021	**0.101**	**0.168**	**0.263**
	(0.023)	(0.091)	(0.066)	(0.038)
E($U_{MF}(1,2)$)	**0.030**	**0.100**	**0.164**	**0.264**
	(0.034)	(0.100)	(0.083)	(0.041)
E($U_{MF}(1,2,3)$)	**0.030**	**0.102**	**0.166**	**0.264**
	(0.033)	(0.097)	(0.080)	(0.040)
E($U_{MF}(2)$)	0.005	0.085	0.138	0.194
	(0.149)	(0.076)	(0.050)	(0.033)
E($U_{MF}(2,3)$)	0.005	0.087	0.137	0.196
	(0.150)	(0.071)	(0.049)	(0.034)
E($U_{MF}(3)$)	0.000	0.007	0.021	0.017
	(0.000)	(0.002)	(0.003)	(0.001)

Table 2: This table reports simulated expected out-of-sample performances using the Fama-French 25 size and book-to-market ratio ranked portfolios for the period 1987.1.-2006.12. Numbers in brackets are standard deviations.

the global minimum variance portfolio - generally leads to higher expected out-of-sample performance than two-fund separated portfolios (as well as Bayesian approaches under a diffuse prior and the method suggested in Jorion (1986)).

In this paper we generalize the three-fund separation results to multi-fund separation and develop a numerical method for finding the optimal combination of any number of funds in a portfolio. While the theoretically optimal combination (under known parameters) can be exactly simulated, under parameter uncertainty it has to be estimated. While the estimation procedure discussed in this paper is not explicitly designed for three-fund separation and works for any number of funds, it outperforms the Kan and Zhou (2007) method for the specific case of three-fund separation.

Conducting performance tests it is demonstrated that combining four funds - the three-fund portfolio of Kan and Zhou (2007) and a fund given by $\mathrm{tr}(\hat{\boldsymbol{\Sigma}}^{\mathbf{-1}})\mathbf{1}$ - yields higher expected out-of-sample performance than three-fund separated portfolios. The performance gain is especially large when sample size is low and estimation error is large. Then some specification of multi-fund separated portfolio strategies even outperform the global minimum variance portfolio.

This paper's primary contribution is methodological. We give good estimates for optimally combining different funds in a portfolio. However, it leaves the open question whether different specification of funds and signals used for estimation may improve results even further.

Appendix

The alternative portfolio strategies used in the performance tests are calculated as follows.

1. Tangency portfolio (TP)

$$\hat{\boldsymbol{w}}_{\text{TP}} = \frac{1}{\gamma}\hat{\boldsymbol{\Sigma}}^{\boldsymbol{-1}}\hat{\boldsymbol{\mu}}$$

2. Global minimum variance portfolio (GMVP) (compare Kan and Zhou (2007, p.648))[21]

$$\hat{\boldsymbol{w}}_{\text{GMVP}} = \frac{1}{\gamma}\frac{T}{T-1}\frac{(T-N-1)(T-N-4)}{T(T-2)}\mu_g\hat{\boldsymbol{\Sigma}}^{\boldsymbol{-1}}\boldsymbol{1}$$

where μ_g is the expected return of the global minimum variance portfolio given by

$$\hat{\mu}_g = \frac{\boldsymbol{1}'\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}}}{\boldsymbol{1}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{1}}\,.$$

3. Jorion (1986) Bayes-Stein portfolio (compare Jorion (1986, p.285) or Kan and Zhou (2007, p.644f))

$$\hat{\boldsymbol{w}}_{\text{Jr86}} = \frac{1}{\gamma}\hat{\boldsymbol{\Sigma}}_{\text{BS}}^{\boldsymbol{-1}}\hat{\boldsymbol{\mu}}_{\text{BS}} \qquad \text{with}$$

$$\hat{\boldsymbol{\Sigma}}_{\text{BS}} = \left(1+\frac{1}{T+\hat{\lambda}}\right)\frac{T-1}{T-N-2}\hat{\boldsymbol{\Sigma}} + \frac{\hat{\lambda}(T-1)}{T(T+1+\hat{\lambda})(T-N-2)}\frac{\boldsymbol{1}\boldsymbol{1}'}{\boldsymbol{1}'\hat{\boldsymbol{\Sigma}}^{\boldsymbol{-1}}\boldsymbol{1}}$$

$$\hat{\lambda} = \frac{T-1}{T-N-2}\frac{(N+2)}{(\hat{\boldsymbol{\mu}}-\hat{\mu}_g\boldsymbol{1})\hat{\boldsymbol{\Sigma}}^{\boldsymbol{-1}}(\hat{\boldsymbol{\mu}}-\hat{\mu}_g\boldsymbol{1})'}$$

[21] Note that the slight differences in the equations stem from the different definitions of $\hat{\boldsymbol{\Sigma}}$; Kan and Zhou (2007) use the maximum likelihood estimator $1/T\sum(\cdot)$ and we use the unbiased estimator $1/(T-1)\sum(\cdot)$.

4. Kan and Zhou (2007, p.634f) two-fund portfolio

$$
\begin{aligned}
\hat{\boldsymbol{w}}_{2F} &= \frac{1}{\gamma}\frac{T}{T-1}\hat{c}^{*}\hat{\boldsymbol{\Sigma}}^{\mathbf{-1}}\hat{\boldsymbol{\mu}} \\
\hat{c}^{*} &= c_3\left(\frac{\hat{\theta}_a^2}{\hat{\theta}_a^2+\frac{N}{T}}\right) \\
c_3 &= \frac{(T-N-1)(T-N-4)}{T(T-2)} \\
\hat{\theta}_a^2 &= \frac{(T-N-2)\hat{\theta}^2-N}{T}+\frac{2(\hat{\theta}^2)^{\frac{N}{2}}(1+\hat{\theta}^2)^{-\frac{T-2}{2}}}{TB_{\hat{\theta}^2/(1+\hat{\theta}^2)}(N/2,(T-N)/2)} \\
\hat{\theta}^2 &= \frac{T}{T-1}\hat{\boldsymbol{\mu}}'\hat{\boldsymbol{\Sigma}}^{\mathbf{-1}}\hat{\boldsymbol{\mu}} \qquad \text{and where} \\
B_x(a,b) &= \int_0^x y^{a-1}(1-y)^{b-1}\mathrm{d}y
\end{aligned}
$$

is the incomplete beta function.

5. Kan and Zhou (2007, p.642f) three-fund portfolio

$$
\begin{aligned}
\hat{\boldsymbol{w}}_{3F}^{**} &= \frac{c_3}{\gamma}\frac{T}{T-1}\left(\hat{c}^{**}\hat{\boldsymbol{\Sigma}}^{\mathbf{-1}}\hat{\boldsymbol{\mu}}+\hat{d}^{**}\hat{\boldsymbol{\Sigma}}^{\mathbf{-1}}\mathbf{1}\right) \\
\hat{c}^{**} &= \frac{\hat{\psi}_a^2}{\hat{\psi}_a^2+\frac{N}{T}} && \text{(A-1)} \\
\hat{d}^{**} &= \frac{\frac{N}{T}}{\hat{\psi}_a^2+\frac{N}{T}}\hat{\mu}_g && \text{(A-2)} \\
\hat{\psi}_a^2 &= \frac{(T-N-1)\hat{\psi}^2-(N-1)}{T}+\frac{2(\hat{\psi}^2)^{\frac{N-1}{2}}(1+\hat{\psi}^2)^{-\frac{T-2}{2}}}{TB_{\hat{\psi}^2/(1+\hat{\psi}^2)}((N-1)/2,(T-N+1)/2)} \\
\hat{\psi}^2 &= \frac{T}{T-1}(\hat{\boldsymbol{\mu}}-\hat{\mu}_g\mathbf{1})'\hat{\boldsymbol{\Sigma}}^{\mathbf{-1}}(\hat{\boldsymbol{\mu}}-\hat{\mu}_g\mathbf{1})\,.
\end{aligned}
$$

References

BAGASHEVA, B., S. RACHEV, J. HSU, AND F. J. FABOZZI (2007): *Bayesian Applications to the Investment Management Process, in: Handbook on Information Technology in Finance*, vol. forthcoming. Springer.

BARRY, C. (1974): "Portfolio Analysis under Uncertain Means, Variances, and Covariances," *Journal of Finance*, 29(2), 515–522.

BROWN, S. (1976): "Optimal Portfolio Choice under Uncertainty: a Bayesian approach," Ph.D. thesis, University of Chicago.

——— (1979): "The Effect of Estimation Risk on Capital Market Equilibrium," *Journal of Financial and Quantitative Analysis*, 14(2), 215–220.

CHOPRA, V., AND W. ZIEMBA (1993): "The Effect of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice," *Journal of Portfolio Management*, (Winter), 6–11.

FROST, P., AND J. SAVARINO (1986): "An Empirical Bayes Approach to Efficient Portfolio Selection," *Journal of Financial and Quantitative Analysis*, 21(3), 293–305.

Garlappi, L., R. Uppal, and T. Wang (2007): "Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach," *The Review of Financial Studies*, 20(1), 41–81.

Haff, L. (1979): "An Identity for the Wishart Distribution with Applications," *Journal of Multivariate Analysis*, 9, 531–544.

James, W., and C. Stein (1961): "Estimation with Quadratic Loss," *in: Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, 1, 361–379.

Jorion, P. (1986): "Bayes-Stein Estimation for Portfolio Analysis," *Journal of Financial and Quantitative Analysis*, 21(3), 279–292.

Kan, R., and G. Zhou (2007): "Optimal Portfolio Choice with Parameter Uncertainty," *Journal of Financial and Quantitative Analysis*, 42(3), 621–656.

Klein, R., and V. Bawa (1976): "The Effect of Estimation on Optimal Portfolio Choice," *Journal of Financial Economics*, 3, 215–231.

Ledoit, O., and M. Wolf (2003): "Improved Estimation of the Covariance Matrix of Stock Returns with and Application to Portfolio Selection," *Journal of Empirical Finance*, 10, 603–621.

Markowitz, H. M. (1952): "Portfolio Selection," *Journal of Finance*, 7(1), 77–91.

Merton, R. (1980): "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics*, 8, 323–361.

Michaud, R. O. (1989): "The Markowitz Optimization Enigma; Is 'Optimized' Optimal?," *Financial Analyst Journal*, 45(1), 31–42.

Muirhead, R. (1982): *Aspects of Multivariate Statistical Theory.* Wiley, New York.

Stein, C. (1955): "Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution," *in. Proceedings of the 3rd Berkely Symposium on Probability and Statistics*, 1, 197–206.

TerHorst, J., F. DeRoon, and B. Werker (2006): *Incorporating Estimation Risk in Optimal Portfolios, in: Advances in Corporate Finance and Asset Pricing.* Elsevier, Amsterdam.

Zellner, A., and V. Chetty (1965): "Prediction and decision problems in regression models from the Bayesian point of view," *Journal of the American Statistical Association*, 60, 608–616.