DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 7/11

Confidence in Prior Knowledge: Calibration and Impact on Portfolio Performance

by

Tobias Wickern

November 30, 2011

DISKUSSIONSBEITRÄGE ZUR STATISTIK UND ÖKONOMETRIE

SEMINAR FÜR WIRTSCHAFTS- UND SOZIALSTATISTIK
UNIVERSITÄT ZU KÖLN

Albertus-Magnus-Platz, D-50923 Köln, Deutschland

DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 7/11

Confidence in Prior Knowledge: Calibration and Impact on Portfolio Performance[1]

by

Tobias Wickern[2]

November 30, 2011

Abstract

The specification of prior parameters is a common practical problem when implementing Bayesian approaches to portfolio optimization. The precision parameter of the prior on the expected asset returns reflects the confidence of the investor in the prior knowledge. Within the framework of the normal-inverse-Wishart model, this paper investigates which factors drive this confidence in order to deduce reasonable values of the precision parameter. We recommend that the investor concentrates on the specification of the precision parameter. By contrast, experts should assess the values of the prior location and dispersion parameter. In the second part of the paper, the impact of the investor's confidence on the performance of investment strategies is examined by a simulation study. The study focusses less on detecting superior portfolio strategies, and more on providing a sensitivity analysis for different levels of confidence. In addition, we show how the posterior distribution of the normal-inverse-Wishart model can be used as a starting point of a simulation process.

Keywords: Bayesian portfolio optimization, Conjugate prior, Confidence parameter, Normal-inverse-Wishart model, Tangency portfolio, Markowitz, Sharpe ratio, Sensitivity analysis, Simulation study.

JEL: C11, C15, G11.

[1]Many thanks go to Karl Mosler for his very important comments on the manuscript.

[2]University of Cologne, Department of Economic and Social Statistics, Chair for Statistics and Econometrics, Albertus-Magnus-Platz, D-50923 Cologne, tel: +49 (0)221 470 4519, email: wickern@wiso.uni-koeln.de.

Confidence in Prior Knowledge: Calibration and Impact on Portfolio Performance☆

Tobias Wickern*

University of Cologne, Graduate School of Risk Management, Albertus-Magnus-Platz, D-50923 Cologne, Germany

Abstract

The specification of prior parameters is a common practical problem when implementing Bayesian approaches to portfolio optimization. The precision parameter of the prior on the expected asset returns reflects the confidence of the investor in the prior knowledge. Within the framework of the normal-inverse-Wishart model, this paper investigates which factors drive this confidence in order to deduce reasonable values of the precision parameter. We recommend that the investor concentrates on the specification of the precision parameter. By contrast, experts should assess the values of the prior location and dispersion parameter. In the second part of the paper, the impact of the investor's confidence on the performance of investment strategies is examined by a simulation study. The study focusses less on detecting superior portfolio strategies, and more on providing a sensitivity analysis for different levels of confidence. In addition, we show how the posterior distribution of the normal-inverse-Wishart model can be used as a starting point of a simulation process.

Keywords:
Bayesian portfolio optimization, Conjugate prior, Confidence parameter, Normal-inverse-Wishart model, Tangency portfolio, Markowitz, Sharpe ratio, Sensitivity analysis, Simulation study
JEL: C11, C15, G11

☆Many thanks go to Karl Mosler for his very important comments on the manuscript.
*Available by phone on +49 221 470-4519, or by fax on +49 221 470-5084
Email address: `wickern@wiso.uni-koeln.de` (Tobias Wickern)

1. Introduction

Modern portfolio theory seeks for mean-variance efficient portfolios and Markowitz (1952) provides a theoretical framework for identifying portfolios that are optimal with respect to an investor-specific risk aversion. Since the parameters of the return distribution are not known to the investor, they must be estimated on a stage preceding the process of portfolio optimization.

Consequently, the effect of parameter uncertainty on the optimal portfolio choice, or – in other words – the sensitivity of the optimal portfolio weights to small changes in the input parameters is extensively analyzed (see, e.g., Jobson and Korkie, 1980; Michaud, 1989). It turns out that the true optimal portfolio can be assumed to be far from its estimate if the latter is obtained by replacing the true input parameters with their sample counterparts in the optimization problem.

Chopra and Ziemba (1993) as well as Best and Grauer (1991) clarify that it is the vector of expected returns, μ that causes the main part of the estimation error. Hence, μ is the main factor of sub-optimality. Furthermore, and in contrast to the covariance matrix Σ of the returns, the sample-based estimation of μ can not be improved by increasing the sample frequency (see Merton, 1980).

Accordingly, one approach to account for the estimation risk is the consideration of portfolios that do not need an explicit estimator for the expected returns. Some authors propose investing in the global minimum variance portfolio (see, e.g., Jagannathan and Ma, 2003; Ledoit and Wolf, 2003). Others suggest consisting of portfolio optimization at all and investing in the equally-weighted portfolio (see Jobson and Korkie, 1981, and recently DeMiguel et al., 2009b). We will come back to these benchmark-forming portfolios later on.

On the other hand, it seems natural to incorporate the parameter uncertainty in a direct way during the estimation of μ and Σ. The Bayesian framework proves to be suitable for this purpose since it allows us to combine information from historical data with prior knowledge. This can lead to robust parameter estimates and to a well-diversified portfolio. Mao and Särndal (1966) and Kalymon (1971) introduce the Bayesian calculus to parameter estimation in the context of modern portfolio theory. Klein and Bawa (1976), and independently Brown (1976), propose using a diffuse prior for μ and Σ originally derived by Jeffreys (1961). This approach continues to rely on the sample estimators for μ and Σ. For this reason, Jeffreys' prior is sometimes referred to as non-informative.

By contrast, Raiffa and Schlaifer (1961) introduce the concept of conjugate priors, which enjoys great popularity among both practitioners and researchers since it combines the benefits of an analytically tractable posterior distribution with the ability to suitably model real problems. Applied to a normal market model, the approach of Ando and Kaufman (1965) is used to tackle the problem of a poor sample estimator for the expected returns by shrinking it toward the assessment of an expert. This involves the determination of the prior mean μ_0 and covariance matrix Σ_0. Finally, it is the role of the investor to express his/her confidence in the expert's views by specifying the prior precision parameter τ which has an influence on the scale of the dispersion.

The main contribution of this paper to the related literature is to provide answers to the following questions:

1. What does the confidence parameter τ depend on?
2. To what extent does the confidence in prior knowledge influence the performance of investment strategies?
3. Which are reasonable values for τ?

In Section 2 we introduce conjugate priors for the market model and discuss some approaches to parameter estimation that are derived from the general model. We develop an analytical expression for the confidence parameter τ in order to respond to question one above. Section 3 is mainly devoted to examining the direct and indirect influence of the portfolio dimension d on the confidence parameter τ. This serves to calibrate the confidence of the investor.

Section 4 addresses our second and third research question by evaluating the results of a simulation study. The performance of well-known portfolio strategies including the global minimum variance portfolio (MVP) and the equally-weighted portfolio (EWP) is investigated for reasonable levels of confidence. This is done in the standard framework of modern portfolio theory and in a more realistic framework that includes short-selling constraints and a risk-free asset. In contrast to many empirical studies that evaluate the out-of-sample performance of portfolio models, we do not aim at finding an outperforming strategy, but rather at conducting a sensitivity analysis for *reasonable* values of τ. Section 5 concludes our work.

2. Confidence in A Normal Market Model

2.1. Standard Model and Existing Estimation Approaches

Throughout this paper we assume that the d-dimensional vector of asset excess returns[1] R_t is multivariate normally distributed, viz.

$$R_t \sim \mathcal{N}(\mu, \Sigma) \ . \tag{1}$$

[1]We will always refer to excess returns, i.e., asset returns minus the corresponding risk-free interest rate. Therefore, in the following we will drop the prefix *excess* for convenience.

Here, μ is the vector of expected returns and Σ is the covariance matrix of the returns. Market information is reflected by a random sample $\mathcal{R} = \{r_1, \ldots, r_n\}$ of past return realizations. This information can be sufficiently[2] summarized by the sample mean

$$\hat{\mu} = \frac{1}{n} \sum_{t=1}^{n} r_t \,, \tag{2}$$

and the sample covariance matrix

$$\widehat{\Sigma} = \frac{1}{n} \sum_{t=1}^{n} (r_t - \hat{\mu})\,(r_t - \hat{\mu})' \,. \tag{3}$$

According to Theorem 7.1.4 and 7.1.5 in Press (1972) it holds that

$$\hat{\mu} \sim \mathcal{N}\,(\mu, \Sigma/n) \quad \text{and} \quad \widehat{\Sigma} \sim W\,(n-1, \Sigma/n)\,, \tag{4}$$

where $W(\nu, \Psi)$ denotes a Wishart distributed random matrix with ν degrees of freedom and scale parameter Ψ (cf. also Appendix A). As we have already mentioned, it is meaningful to focus on the estimation of the vector of expected returns. To start with, assume that the covariance matrix Σ is known or can at least be estimated with sufficient precision. A conjugate prior for μ with respect to the distribution of the sample mean – see Eq. (4) – is given by

$$\mu \sim \mathcal{N}\,(\mu_0, \tau\Sigma_0)\,. \tag{5}$$

The prior parameters μ_0 and Σ_0 as well as the confidence parameter τ must be specified in order to carry out the Bayesian analysis. In our view, the investor should draw on expert knowledge to specify the location parameter μ_0 and the dispersion parameter Σ_0. By contrast, s/he should calibrate the parameter τ to express his/her confidence on the expert assessments. Note that small values of τ correspond to a high level of confidence. Following Gelman et al. (2004, p. 85f), the posterior distribution of μ is normal, too, with the posterior parameters

$$\begin{aligned} \mu_1 &= ((\tau\Sigma_0)^{-1} + n\Sigma^{-1})^{-1}((\tau\Sigma_0)^{-1}\mu_0 + n\Sigma^{-1}\hat{\mu})\,, \\ \Sigma_1^{-1} &= (\tau\Sigma_0)^{-1} + n\Sigma^{-1}\,. \end{aligned} \tag{6}$$

Then, the predictive return distribution reads $R_{n+1}|\mathcal{R} \sim \mathcal{N}(\mu_1, \Sigma + \Sigma_1)$.

[2]A proof that $\hat{\mu}$ and $\widehat{\Sigma}$ are *sufficient statistics* can be found in Press (1972, Theorem 7.1.1)

In the following, we look at some well-known estimation procedures which can be derived within the standard model as presented. We focus on how the parameters of the prior distribution are specified, especially the confidence parameter τ.

Since the seminal work of Stein (1956) it is well-known that the sample mean $\hat{\mu}$ is dominated by the so-called James-Stein estimator $\phi_{\mathrm{JS}}\mu_0 + (1-\phi_{\mathrm{JS}})\hat{\mu}$ in terms of a quadratic loss function. The James-Stein estimator belongs to the class of shrinkage estimators which have been used in portfolio optimization for many years (see, e.g., Jobson and Korkie, 1979). Jorion (1986) introduces a shrinkage estimator which can be derived within a Bayesian framework. More precisely, consider the following specification of the prior parameters in Eq. (5):

$$\mu_0 \equiv \hat{\mu}_{\mathrm{MVP}}\mathbf{1} = \frac{\mathbf{1}'\Sigma^{-1}\hat{\mu}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}\mathbf{1} \quad \text{and} \quad \Sigma_0 = \Sigma\,, \tag{7}$$

where $\mathbf{1}$ denotes a column vector of ones. Interestingly, the scale factor $\hat{\mu}_{\mathrm{MVP}}$ equals the expected return of the global minimum variance portfolio in the conventional setting which is examined in greater detail later on. With this specific informative prior, the predictive return distribution is easily shown to be normal with mean

$$\mathrm{E}\left(R_{n+1}|\mathcal{R}\right) = \phi_{\mathrm{BS}}\,\frac{\mathbf{1}'\Sigma^{-1}\hat{\mu}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}\,\mathbf{1} + \left(1-\phi_{\mathrm{BS}}\right)\hat{\mu}\,, \quad \phi_{\mathrm{BS}} = \frac{1}{1+\tau n}\,. \tag{8}$$

Consequently, the latter expression is referred to as the *Bayes-Stein* estimator. Jorion (1986, Eq. 17) offers a way to estimate the shrinkage weight ϕ_{BS} directly from the data. His approach corresponds to a specific choice of confidence, viz.

$$\tau = \frac{\left(\hat{\mu}-\hat{\mu}_{\mathrm{MVP}}\mathbf{1}\right)'\Sigma^{-1}\left(\hat{\mu}-\hat{\mu}_{\mathrm{MVP}}\mathbf{1}\right)}{d+2}\,. \tag{9}$$

Due to the data-driven estimation process, Jorion (1986) refers to it as the *empirical* Bayes approach. He claims that 'this approach will outperform the classical sample mean because it relies on a richer model'. Even though this has been proven to be true in many empirical studies, it is still a matter of debate whether this assertion leads to portfolio strategies which can outperform the naive diversification rule.

Although it is clear that experts eventually fall back on market data in order to generate forecasts, we should keep in mind that this empirical Bayes approach uses the *same* data to specify the prior parameters and to incorporate the market information. Strictly speaking, this course of action contradicts the principle of Bayesian analysis. More importantly, the confidence parameter τ cannot be chosen in an explicit way by the investor but is rather a by-product of the determination of the shrinkage weight ϕ_{BS}.

The well-known approach of Black and Litterman (1991, 1992) can also be viewed in terms of the Bayesian standard model described above. Instead of market information, the Black-Litterman model combines prior knowledge with subjective *views* on μ. Note that Black and Litterman (1991) proceed on the prior distribution in (5) and set $\Sigma_0 = \Sigma$, following Jorion (1986). The parameter μ_0 is replaced by the vector of implicit returns π, which is the solution of a reverse mean-variance optimization problem, i.e., $\pi = \lambda \Sigma w_{\text{opt}}$ and $w_{\text{opt}} = \arg\max_w w'\mu - \lambda/2 \cdot w'\Sigma w$.

Several specifications of the confidence parameter τ are proposed by the literature on the Black-Litterman approach. Some authors – including Fusai and Meucci (2003), Meucci (2010) as well as Satchell and Scowcroft (2000) – set τ to a value of one. This often goes along with the so-called *Alternative Reference Model* in which the prior distribution is defined without an explicit precision parameter. Neglecting the confidence parameter τ contrasts with the original work of Black and Litterman (1991) who state that τ will be close to zero. This suggestion is further promoted by Idzorek (2007). He and Litterman (2002) set $\tau = 1/n$, considering it as the ratio of the sampling variance to the distribution variance.

Walters (2010, 2011) discusses some empirical ways to calibrate τ, mostly ending in values close to zero. However, many approaches are built around the implicit return π and cannot be used in general for the prior parameter μ_0. Other contributions concern rules of thumb and suffer from their lack of theoretical foundation. Nevertheless, it seems to be widely accepted that τ is in practical applications much less than one, reflecting the fact that the uncertainty in the mean of the return distribution is much smaller than the (co-)variances of the returns (see also Walters, 2011).

In contrast to Jorion (1986) and Black and Litterman (1991), Kempf et al. (2002) propose to model the estimation risk[3] independently of the innovation risk Σ. To be precise, they replace Σ_0 by the identity matrix I and set $\mu_0 = \bar{\mu}_0 \mathbf{1}$. In other words, a simplified structure of the prior mean is assumed, indicating that the expected returns are identical across all d assets. In this setting, the parameters of the posterior distribution can be expressed as follows (see Memmel, 2004, p. 83f):

$$
\begin{aligned}
\mu_1 &= (I - K)\, \bar{\mu}_0 \mathbf{1} + K \hat{\mu}\,, \\
\Sigma_1 &= \tau\, (I - K)^{-1}\ ,
\end{aligned} \tag{10}
$$

[3]Note that Kempf et al. (2002) consider the prior distribution in Eq. (5) as a model of estimation risk which can thus be controlled by the dispersion parameter Σ_0.

Table 1: Optimal strategies for extreme constellations of n and τ

	homogeneous returns $\tau = 0$	inhomogeneous returns $\tau \to \infty$
no history $n = 0$	global minimum variance portfolio	equally-weighted portfolio
long history $n \to \infty$	global minimum variance portfolio	sample-based approach

where the matrix-valued shrinkage weight is defined by $K = (\Sigma/(\tau n) + I)^{-1}$. Based on the specification in Eq. (10), Kempf et al. (2002) derive optimal portfolio strategies depending on the values τ and n (see Table 1). Taking a closer look at this, we consider the tangency portfolio (TP),

$$w_{\text{TP}} = \frac{\Sigma^{-1}\mu}{\mathbf{1}'\Sigma^{-1}\mu}\,, \tag{11}$$

which is the unique efficient portfolio consisting of risky assets solely. According to Tobin's two-fund separation theorem (Tobin, 1958) the optimal portfolio $w_{\text{opt}} = \Sigma^{-1}\mu/\lambda$ can be assessed by the combination of the risk-free asset with the tangency portfolio.[4] By contrast, the global minimum variance portfolio (MVP), defined by the weight vector

$$w_{\text{MVP}} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}\,, \tag{12}$$

is the solution of the optimization problem $\min_w w'\Sigma w$ s.t. $w'\mathbf{1} = 1$ and yields the lowest portfolio variance if *no risk-free asset* is assessable. Note that both the global minimum variance portfolio and the tangency portfolio lie on the efficient frontier in the case without a risk-free asset whereas the tangency portfolio is also part of the capital market line. Finally, the equally-weighted portfolio (EWP) is simply defined by $w_{\text{EWP}} = \mathbf{1}/d$.

In the standard setting without constraining shortsales and with the opportunity to invest in a risk-free asset, the investor's main task is to determine the tangency portfolio (see the explanations above). Kempf et al. (2002) basically show that the tangency portfolio coincides with the MVP if the investor highly trusts the expert's

[4]More generally, the separation theorem holds for *every* efficient portfolio.

assessment. This is due to the simple structure of the prior parameter μ_0. Mathematically speaking, it holds that $\lim_{\tau\to 0} K = 0_d$, $\lim_{\tau\to 0}\mu_1 = \bar{\mu}_0\mathbf{1}$, $\lim_{\tau\to 0}\Sigma_1 = 0_d$ and, consequently,

$$\lim_{\tau\to 0}\hat{w}_{\mathrm{TP}} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}} \equiv w_{\mathrm{MVP}}\,, \tag{13}$$

where we define $\hat{w}_{\mathrm{TP}} := (\Sigma_1+\Sigma)^{-1}\mu_1/(\mathbf{1}'(\Sigma_1+\Sigma)^{-1}\mu_1)$, and 0_d denotes the $d\times d$ zero matrix. By contrast, $\lim_{\substack{\tau\to\infty\\ n\to 0}} K = 0_d$, $\lim_{\substack{\tau\to\infty\\ n\to 0}} \tau\,(\Sigma_1+\Sigma)^{-1} = I$ and

$$\lim_{\substack{\tau\to\infty\\ n\to 0}}\hat{w}_{\mathrm{TP}} = \frac{\mathbf{1}}{d} \equiv w_{\mathrm{EWP}}\,, \tag{14}$$

if we additionally assume $\tau n\to 0$. Thus, the TP corresponds to the EWP if the investor is suspicious of the expert's knowledge and no market information is available. Intuitively, if it is impossible to assess the asset returns ex ante, one should better pass on portfolio optimization and follow the naive diversification rule. On the other hand, if the investor can rely on a very large return sample, the sample estimator for the expected returns comes into play favoring the traditional Markowitz approach of portfolio optimization.[5] Then, $\lim_{\tau,n\to\infty} K = I$, $\lim_{\tau,n\to\infty}\mu_1 = \hat{\mu}$, $\lim_{\tau,n\to\infty}\Sigma_1 = 0_d$ and

$$\lim_{\tau,n\to\infty}\hat{w}_{\mathrm{TP}} = \frac{\Sigma^{-1}\hat{\mu}}{\mathbf{1}'\Sigma{-1}\hat{\mu}}\,. \tag{15}$$

A natural question is how far away typical values of τ are from their limits. In Section 4 we evaluate the performance of the afore-mentioned special portfolios for different values of τ by simulation. Our special emphasis will be on deducing *realistic* values of the confidence parameter. The theoretical findings of Kempf et al. (2002) might be less relevant from a practical point of view if the performance of portfolio strategies is very sensitive to changes in the level of confidence.

2.2. Determinants of Confidence in the Normal-Inverse-Wishart Model

The basic advantages of the conjugate prior idea can be extended to a situation in which the covariance matrix Σ is assumed to be unknown. To this end, Ando and Kaufman (1965) introduced the normal-inverse-Wishart conjugate prior

$$\begin{aligned}\mu|\Sigma &\sim \mathcal{N}\left(\mu_0,\tau\Sigma\right)\,,\\ \Sigma &\sim \mathrm{IW}\left(\nu_0+d+1,\nu_0\Sigma_0\right)\,,\end{aligned} \tag{16}$$

[5]In contrast to the seminal work of Markowitz (1952), the problem of parameter estimation is limited here to the vector of expected returns.

i.e., in addition to the prior for μ we assume that the covariance matrix Σ is a priori distributed as a $d \times d$ random matrix following the inverse Wishart distribution with $\nu_0 + d + 1$ degrees of freedom, precision parameter ν_0 and dispersion parameter Σ_0 (cf. Appendix A for notational convenience).[6]

The extended model in Eq. (16) can also be seen as a response to the critique of Kempf et al. (2002) since it explicitly takes into account the risk inherent in the estimation of Σ. Note that in the normal-inverse-Wishart (NIW) model, the prior parameter Σ_0 plays a role similar to that in the basic model. This is reflected by the fact that $\mathrm{E}(\Sigma^{-1}) = \Sigma_0^{-1}$ (see Appendix A for the relation between the Wishart and the inverse Wishart distribution).[7] The posterior distributions are given by

$$\begin{aligned} \mu|\Sigma, \mathcal{R} &\sim \mathcal{N}\left(\mu_1, \tau w_\tau \Sigma\right) , \\ \Sigma|\mathcal{R} &\sim \mathrm{IW}\left(\nu_1 + d + 1, \nu_1 \Sigma_1\right) , \end{aligned} \tag{17}$$

where

$$\begin{aligned} w_\tau &= 1/(1+\tau n) , \\ \mu_1 &= w_\tau \mu_0 + (1 - w_\tau)\,\hat{\mu} , \\ \nu_1 &= \nu_0 + n , \\ \Sigma_1 &= \left(n\widehat{\Sigma} + \nu_0 \Sigma_0 + n \cdot w_\tau \left(\hat{\mu} - \mu_0\right)\left(\hat{\mu} - \mu_0\right)'\right) / \nu_1 . \end{aligned} \tag{18}$$

According to Brown (1976, p. 145) the predictive return distribution reads

$$R_{n+1}|\mathcal{R} \sim \mathrm{t}\left(\nu_1, \mu_1, (1 + \tau w_\tau)\,\Sigma_1\right) , \tag{19}$$

where $\mathrm{t}(\nu, \mu, \Sigma)$ denotes the multivariate noncentral t distribution with location parameter μ, dispersion parameter Σ and ν degrees of freedom (cf. Kotz and Nadarajah, 2004, p. 1). The parameters μ and Σ are replaced in the optimization problem by the first two moments of the distribution of R_{n+1}, viz.

$$\begin{aligned} \mathrm{E}\left(R_{n+1}|\mathcal{R}\right) &= \mu_1 , \\ \mathrm{Var}\left(R_{n+1}|\mathcal{R}\right) &= \frac{\nu_1}{\nu_1 - 2}\left(1 + \tau w_\tau\right)\Sigma_1 . \end{aligned} \tag{20}$$

[6] Note that the prior for μ is now modeled conditional on Σ. The corresponding *unconditional* prior is multivariate t distributed; see, e.g., Meucci (2005, p. 371) for details.

[7] By contrast, it holds that $\mathrm{E}(\Sigma) = \nu_0/(\nu_0 - d - 1) \cdot \Sigma_0 \neq \Sigma_0$.

Note that we have to deal with a further prior parameter ν_0 which controls the investor's confidence in the expert knowledge concerning the dispersion matrix Σ_0. However, we focus on modeling the confidence parameter τ for three reasons:

1. In terms of the *prior* distributions, we can completely separate the *direct* influence of τ on the expected returns and of ν_0 on the covariance matrix of the returns. Remember that covariance matrices can generally be estimated more accurately than expected returns (see Merton, 1980). Even more seriously, the vector of optimal portfolio weights is much more sensitive to changes in the expected returns (see Best and Grauer, 1991; Chopra and Ziemba, 1993; Kempf and Memmel, 2002). We therefore believe that choosing a small value for ν_0 is adequate to express the investor's confidence in that regard.
2. The normal-inverse Wishart prior in (16) is constructed as follows. In a first step, the covariance matrix Σ is generated. Then, conditional on the realized covariance matrix, the vector of expected returns is modeled. For this purpose, ν_0 works like a normalization constant and can afterwards be canceled out if the value for τ is suitably chosen, since both parameters affect the *scale* of Σ.
3. The confidence parameter τ has an influence on *both* the first and the second moment of the predictive return distribution while the parameter ν_0 only impacts on the second moment (cf. Eq. (18) and (20)). Moreover, τ controls the intensity with which the sample mean is shrunk towards the prior mean μ_0.[8]

Applying Theorem 1.3.4 of Muirhead (1982) we conclude from Eq. (16) that

$$\mu' \left(\tau\Sigma\right)^{-1} \mu \,|\, \Sigma \sim \chi^2_d(\delta)\,, \tag{21}$$

where $\chi^2_d(\delta)$ denotes the non-central chi-squared distribution with d degrees of freedom and non-centrality parameter δ. In our case, it holds that

$$\delta = \mu_0' \left(\tau\Sigma\right)^{-1} \mu_0\,. \tag{22}$$

Note that the expected return and the variance of the tangency portfolio are given by $\mu_{\mathrm{TP}} = \mu'\Sigma^{-1}\mu/(\mathbf{1}'\Sigma^{-1}\mu)$ and $\sigma^2_{\mathrm{TP}} = \mu'\Sigma^{-1}\mu/(\mathbf{1}'\Sigma^{-1}\mu)^2$, respectively. Hence, the Sharpe ratio of the tangency portfolio is defined as follows:

$$\mathrm{Sh}_{\mathrm{TP}} := \frac{\mu_{\mathrm{TP}}}{\sigma_{\mathrm{TP}}} = \sqrt{\mu'\Sigma^{-1}\mu}\,. \tag{23}$$

[8]The following holds: $\tau\uparrow\Rightarrow\mu_{1,i}\downarrow$ if $\mu_{0,i} > \hat{\mu}_i$ and $\tau\uparrow\Rightarrow\mu_{1,i}\uparrow$ if $\mu_{0,i} < \hat{\mu}_i$ $(i = 1,\ldots,d)$.

Consequently, we can write $\mathrm{Sh}_{\mathrm{TP}}^2|\Sigma \sim \tau\chi^2(\delta)$ and it holds that

$$\mathrm{E}\left(\mathrm{Sh}_{\mathrm{TP}}^2|\Sigma\right) = \tau\left(d+\delta\right) . \tag{24}$$

Applying the law of iterated expectations, we get

$$\mathrm{E}\left(\mathrm{Sh}_{\mathrm{TP}}^2\right) = \mathrm{E}\left\{\mathrm{E}\left(\mathrm{Sh}_{\mathrm{TP}}^2|\Sigma\right)\right\} = \mathrm{E}\left(\tau(d+\delta)\right) = \tau d + \mu_0'\mathrm{E}\left(\Sigma^{-1}\right)\mu_0 . \tag{25}$$

As already mentioned, it holds that $\mathrm{E}(\Sigma^{-1}) = \Sigma_0^{-1}$ and thus

$$\tau = \frac{\mathrm{E}\left(\mathrm{Sh}_{\mathrm{TP}}^2\right) - \mu_0'\Sigma_0^{-1}\mu_0}{d} . \tag{26}$$

The key observation is that we are able to model τ by undertaking reasonable values for the Sharpe ratio of the tangency portfolio, e.g., from the relevant literature. One should not be concerned about the fact that the Sharpe ratio of the TP comes into play only through its expectation since this reflects the Bayesian point of view.

Note that ν_0 is the only prior parameter which does not affect the value of τ. This confirms the argument in point 1 above that the two confidence parameters τ and ν_0 can be discriminated according to their direct influence on the prior distribution of μ and Σ, respectively. The investor's confidence in the prior mean increases as the true Sharpe ratio of the TP is approximated with increasing accuracy by that one following from the expert assessments μ_0 and Σ_0. Furthermore, the value of τ is inversely proportional to the number of assets d, implying that the investor's confidence increases the more assets are on the market. We discuss this point in greater detail in Section 3.

Frost and Savarino (1986) apply an empirical Bayes approach (cf. Jorion, 1986) to the normal-inverse-Wishart model. More precisely, the model parameters are obtained via maximum-likelihood (ML) estimation assuming the following structure of the location and dispersion parameter, respectively:

$$\mu_0 \equiv \bar{\mu}_0\mathbf{1} \quad \text{and} \quad \Sigma_0 \equiv \bar{\sigma}_0^2\left\{\rho_0\mathbf{1}\mathbf{1}' + (1-\rho_0)I\right\} . \tag{27}$$

In other words, it is a priori assumed that the expected returns as well as the return variances are equal. On top of that, the returns are supposed to be equicorrelated. These simplifications might be motivated by the fact that the estimation error for the parameter of a particular asset increases the more its sample estimate differs from the average for all assets (cf. Frost and Savarino, 1986).

We want to analyze the input parameters in (26) against the backdrop of the simplifications proposed by Frost and Savarino (1986). Note that the following holds due to the equicorrelation structure introduced in (27) (cf. Press, 1972, p. 23):

$$\Sigma_0^{-1} = \frac{1}{\bar{\sigma}_0^2}\left(\frac{I}{1-\rho_0} - \frac{\rho_0 \mathbf{1}\mathbf{1}'}{(1-\rho_0)\,(1+(d-1)\rho_0)}\right). \tag{28}$$

In the special case of (28), the expression for τ in (26) can be simplified to

$$\tau = \frac{\mathrm{E}\left(\mathrm{Sh}_{\mathrm{TP}}^2\right)}{d} - \frac{\bar{\mu}_0^2}{\bar{\sigma}_0^2\,(1+(d-1)\rho_0)}. \tag{29}$$

According to Jobson and Korkie (1982), we can treat $\mathrm{E}(\mathrm{Sh}_{\mathrm{TP}}^2)$ as an indicator for the potential performance of the d asset set since the slope of the capital market line is equal to $\mathrm{Sh}_{\mathrm{TP}}$. The set of feasible portfolios increases the higher the Sharpe ratio of the TP and thus, better risk-return combinations are attainable for the investor. Furthermore, the prior parameters $\bar{\mu}_0$, $\bar{\sigma}_0^2$ and ρ_0 can economically be interpreted as indicators for certain market variables which are assessed by an expert. More precisely, $\bar{\mu}_0$ and $\bar{\sigma}_0^2$ refer to the expert view on the market potential and the market variation, respectively, while ρ_0 can be seen as an expert assessment on the average return correlation. Indeed, the confidence parameter τ itself has an economic interpretation. If we assume $\mu_0 = \bar{\mu}_0\mathbf{1}$, i.e., all expected returns are a priori equal, it can be interpreted as a measure of return diversity (cf. also Kempf et al., 2002).

Note that the subtrahend in (29) is a multiple of the Sharpe ratio of the TP implied by the expert assessment of the prior parameters. Hence, the value of the confidence parameter τ in Eq. (29) depends on how well the Sharpe ratio of the tangency portfolio is approximated by that one following from the expert assessments; see also our explanations on the general formula in Eq. (26). According to the presumed split of roles, the investor uses $\mathrm{E}(\mathrm{Sh}_{\mathrm{TP}}^2)$ in order to calibrate his/her confidence in the expert views. Consequently, in Section 3 we aim at providing reasonable values for the Sharpe ratio of the tangency portfolio as an aid to decision-making for the investor.

3. Assessing the Potential Performance of Asset Sets

3.1. Stock Indices as Proxies

First of all, statements about the potential performance of assets, i.e., the Sharpe ratio of the tangency portfolio, are manifold in the relevant literature. Frahm (2010a) considers the range of 0.2 and 0.5 as typical values of the *annualized* Sharpe ratio of

Table 2: Estimates for the Sharpe ratio of the tangency portfolio

reference	data source	period	Sh^2_{TP}/mo.
MacKinlay (1995)	CRSP value-weighted index[a]	1963/07-1991/12	0.006
	S&P 500 Index	1981/01-1992/06	0.009
Cogley and Sargent (2008)	S&P Composite Index[b]	1872/01-2003/12	0.004
		1934/01-2003/12	0.014
Dimson et al. (2003)	world stock portfolio[c]	1900/12-2002/12	0.010
Jorion (1991)	CRSP value-weighted index[a]	1931/01-1987/12	0.014
	CRSP equal-weighted index[a]	1931/01-1987/12	0.020
Brennan et al. (1998)	3 Fama-French factor portfolios[d]	1963/01-1995/12	0.067
Kan and Zhou (2007)	10 NYSE size-ranked portfolios[d]	1926/01-2003/12	0.025
	25 Fama-French size and book-to-market portfolios[d]	1932/01-2003/12	0.118
Gospodinov et al. (2010)	25 Fama-French size and book-to-market portfolios[d]	1952/06-2000/12	0.168
	10 NYSE size-ranked portfolios & 12 FF industry portfolios[d]	1952/06-2000/12	0.093

We list the monthly squared Sharpe ratios of the tangency portfolio, estimated using various indices and portfolios. Note that in most of the references, the Sharpe ratio is estimated on a quarterly or yearly basis. Thus, the data above are converted appropriately.

[a] The CRSP value-weighted index uses all issues listed on the NYSE, AMEX and NASDAQ.

[b] This sample basically corresponds to the S&P 500 Index (prior to March 1957: S&P 90 Index).

[c] Dimson et al. (2002) explain in detail the composition of the world stock portfolio.

[d] The data are available on Ken French's Web site at `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/`. The Fama-French portfolios include all NYSE, AMEX and NASDAQ stocks.

the tangency portfolio.[9] This would approximately coincide with squared monthly values of 0.003 and 0.021. By contrast, Kan and Robotti (2008) are confident that values between 0.2 and 0.4 cover a reasonably wide range of *monthly* Sharpe ratios corresponding to squared values of 0.040 and 0.160. MacKinlay (1995) asserts that a reasonable value for the squared monthly Sharpe ratio is 0.031 if a perfect capital market is assumed.

Assessing the theoretical value of the Sharpe ratio is typically founded on empirical data. More precisely, the Sharpe ratio is often approximated by the empirical Sharpe ratio of well-diversified stock indices, using long estimation horizons (up to 100 years). Table 2 gives an overview of the estimates of selected indices and

[9] The Sharpe ratio of the *tangency portfolio* is referred to as the Sharpe ratio in this section.

portfolios. Having a closer look to the data, we notice that the Sharpe ratio estimates mainly differ due to the following reasons:

- *Asset characteristics*: It makes a difference whether the proxy for the estimated Sharpe ratio reproduces the performance of a weighted average of single shares or of stock portfolios. Among the latter, Fama and French's portfolios are the standard test assets in recent empirical studies and yield higher Sharpe ratios than the CRSP indices or the S&P index.

- *Observation period*: The squared Sharpe ratio for the CRSP value-weighted index from the years 1931 to 1987 is more than twice that from 1963 to 1991. Similarly, the S&P index yields a squared Sharpe ratio of about 0.4% in the years 1872 to 2003. The respective value for the years 1981 to 1992 is 0.9% while being more than three times as high in the period from 1934 to 2002 (1.4%). However, the Sharpe ratio level is not primarily influenced by the sample size but, rather, by the stock market situation and extremal events like stock market crashes.

3.2. The Impact of the Number of Assets

In addition to these factors, which must be considered carefully by an investor for a *specific* asset market, Kan and Zhou (2007) come up with another factor which *generally* impacts on the Sharpe ratio. They state that the Sharpe ratio increases with more assets leading to a reduced or even reversed effect of the number of assets d on the level of confidence τ (cf. Eq. (26)). Therefore, we investigate the relationship between d and the Sharpe ratio in greater detail.

To that end, consider a market with d_1 assets to which d_2 assets are added. After the addition, there are $d = d_1 + d_2$ assets on the market. According to the partitioning of the d assets, we define

$$\mu = \begin{bmatrix} \mu_{d_1} \\ \mu_{d_2} \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} \Sigma_{d_1} & \Sigma_{d_{1\text{-}2}} \\ \Sigma_{d_{2\text{-}1}} & \Sigma_{d_2} \end{bmatrix} . \tag{30}$$

Jobson and Korkie (1984) consider the multivariate regression of the returns from the d_2 new assets on the returns from the original d_1 assets, viz.

$$R_{d_2,t} = \alpha + \beta R_{d_1,t} + u_t \, , \tag{31}$$

where $R_t = \begin{bmatrix} R_{d_1,t} & R_{d_2,t} \end{bmatrix}'$, i.e., the return vector of the d assets is partitioned accordingly. The matrix of coefficients is given by $\beta = \Sigma_{d_{2\text{-}1}} \Sigma_{d_1}^{-1}$. The vector of intercept terms $\alpha = \mu_{d_1} - \beta\mu_{d_2}$ can be thought of as the vector of generalized Jensen measures. For the error term u_t it holds that $\mathrm{E}(u_t) = \mathbf{0}$, $\mathrm{Cov}(u_t, R_{d_1,t}) = 0$ as well as

$$\mathrm{Cov}\,(u_t) := \Sigma_u = \left(\Sigma_{d_2} - \Sigma_{d_{2\text{-}1}}\Sigma_{d_1}^{-1}\Sigma_{d_{1\text{-}2}}\right) . \tag{32}$$

Note that

$$\Sigma^{-1} = \begin{bmatrix} \Sigma_{d_1}^{-1} + \beta'\,\Sigma_u^{-1}\beta & -\beta'\,\Sigma_u^{-1} \\ -\Sigma_u^{-1}\beta & \Sigma_u^{-1} \end{bmatrix} . \tag{33}$$

The squared Sharpe ratio of the new set of assets can be calculated as follows:

$$\begin{aligned} \mathrm{Sh}^2_{\mathrm{TP}} = \mu\Sigma^{-1}\mu &= \mu'_{d_1}\left(\Sigma_{d_1}^{-1} + \beta'\Sigma_u^{-1}\beta\right)\mu_{d_1} - \mu'_{d_1}\beta'\Sigma_u^{-1}\mu_{d_2} - \mu'_{d_2}\Sigma_u^{-1}\beta\mu_{d_1} + \mu'_{d_2}\Sigma_u^{-1}\mu_{d_2} \\ &= \mu'_{d_1}\Sigma_{d_1}^{-1}\mu_{d_1} + (\beta\mu_{d_1} - \mu_{d_2})'\,\Sigma_u^{-1}\,(\beta\mu_{d_1} - \mu_{d_2}) = \mathrm{Sh}^2_{\mathrm{TP}}(d_1) + \alpha'\Sigma_u^{-1}\alpha\,, \end{aligned} \tag{34}$$

where $\mathrm{Sh}^2_{\mathrm{TP}}(d_1)$ is the squared Sharpe ratio of the original set of assets. Due to the positive definiteness of Σ_u it holds that $\mathrm{Sh}^2_{\mathrm{TP}} \geq \mathrm{Sh}^2_{\mathrm{TP}}(d_1)$. The change in squared Sharpe ratios is equal to the inner product of the vector of Jensen's alphas weighted by the inverse of the covariance matrix of the error terms. Hence, we may write $\mathrm{Sh}^2_{\mathrm{TP}}(d)$ in order to clarify that the Sharpe ratio depends on the number of assets. Furthermore, define the quotient

$$q(d, d_1) := \frac{\mathrm{Sh}^2_{\mathrm{TP}}(d)}{\mathrm{Sh}^2_{\mathrm{TP}}(d_1)} . \tag{35}$$

In general, it is difficult to assess the *strength* of the influence of d on the Sharpe ratio. However, we investigate this question at least for the case in which μ and Σ are simply structured. More precisely, assume that the returns are equicorrelated and have equal means and variances, i.e., $\mu = \bar{\mu}\mathbf{1}$ and $\Sigma = \bar{\sigma}^2\{\rho\mathbf{1}\mathbf{1}' + (1-\rho)I\}$ (cf. Frost and Savarino, 1986). Then, it follows that

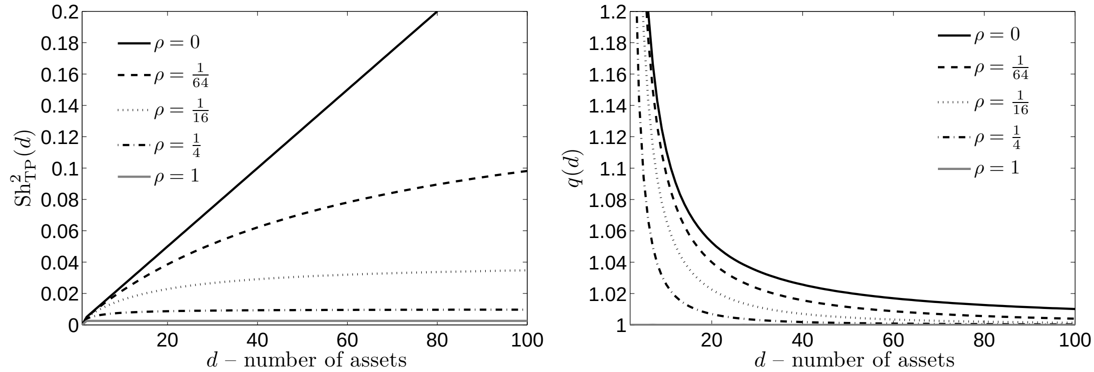
$$\mathrm{Sh}^2_{\mathrm{TP}}(d) = \frac{d\,\bar{\mu}^2}{\bar{\sigma}^2\,(1 + (d-1)\rho)} \tag{36}$$

and

$$q(d, d_1) = \frac{d}{d_1}\frac{1 + (d_1 - 1)\rho}{1 + (d-1)\rho} . \tag{37}$$

In Figure 1, we plot on the left hand side the true squared Sharpe ratio as a function of d according to Eq. (36). In particular, we consider a range between one and 100 assets and assume that all returns exhibit a monthly mean of 0.5% and a monthly variance of 1%. Obviously, the fraction $\bar{\mu}^2/\bar{\sigma}^2$ only affects the steepness of the squared

Figure 1: True Sharpe ratio (quotient)



Depicted are the true squared Sharpe ratio (on the left) and the true Sharpe ratio quotient (on the right) as a function of d if it is assumed that returns are equicorrelated and have equal means and variances ($\bar{\mu} = .005$ and $\bar{\sigma}^2 = .01$). We consider different levels of return correlation.

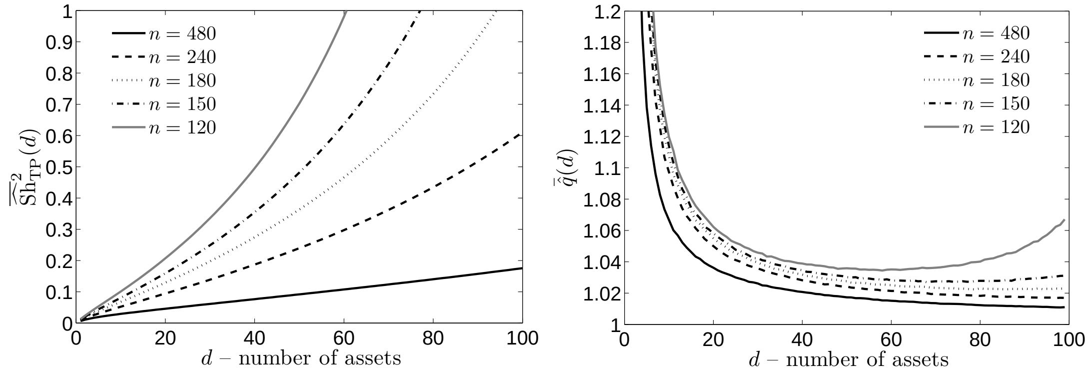
Sharpe ratio, while the level of return correlation controls the curvature of $\mathrm{Sh}^2_{\mathrm{TP}}$ (left hand side of Figure 1). If the asset returns are uncorrelated ($\rho = 0$), the relationship between the squared Sharpe ratio and the number of assets is linear. On the contrary, if the returns are perfectly correlated, the squared Sharpe ratio is constant in d. For $0 < \rho < 1$, $\mathrm{Sh}^2_{\mathrm{TP}}$ is a concave function in d indicating that the squared Sharpe ratio typically increases less than proportionately to the increase in the number of assets.

On the right hand side of Figure 1, we plot the quotient of two squared Sharpe ratios if one asset is added to the market, i.e., $q(d) = d/(d-1) \cdot (1 - \rho/(1 + (d-1)\rho))$ (cf. also Eq. (37)). The marginal contribution of an additional asset decreases with an increasing number of assets, as can be seen from the shape of the graphs. By contrast, the lower the asset correlation, the larger is the quotient $q(d)$.

The question is whether the results from the analysis of true Sharpe ratios under the restriction of simplified structures for μ and Σ hold also for real market conditions. In Figure 2 we display estimates for $\mathrm{Sh}^2_{\mathrm{TP}}(d)$ and $q(d)$ which are implied by the sample counterparts of μ and Σ using a sample of monthly returns from the CRSP data set containing 480 observations and 283 assets. The data come from the monthly stock file of the Center for Research in Security Prices (CRSP) including stocks from NYSE, AMEX and NASDAQ. We consider all stocks with *complete* return history between January 1969 and December 2008. The estimation procedure is as follows:

1. A set of 100 assets is randomly drawn from the whole sample.
2. The sample mean and the sample covariance matrix are estimated in order to calculate the squared Sharpe ratio.

Figure 2: Sample estimates of the Sharpe ratio (quotient)



The estimated Sharpe ratio is pictured on the left as the average over 50,000 draws from a real return sample, whereas the right figure shows the respective estimated Sharpe ratio quotient. The effect of different sample sizes is clearly demonstrated.

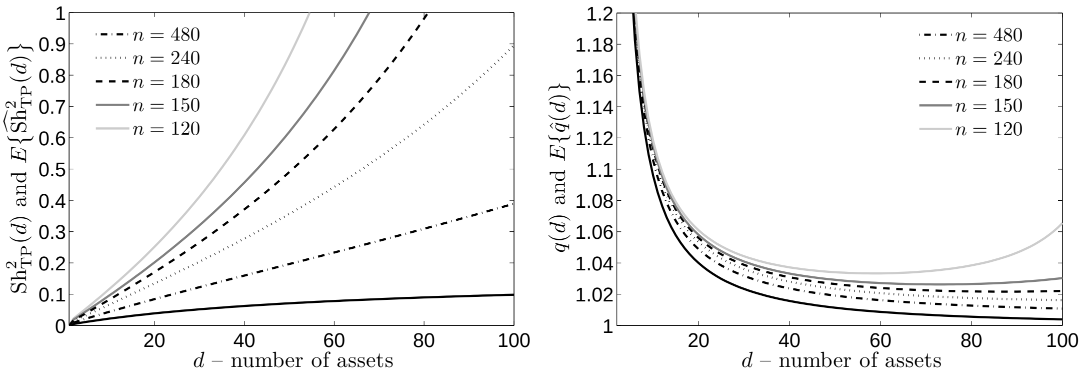
3. One randomly chosen asset is deleted from the set. The sample mean and sample covariance matrix as well as the squared Sharpe ratio of the diminished asset set is computed. Finally, the Sharpe ratio quotient is formed. This step is repeated until only one asset remains in the set. Then, return to step 1.

Altogether, the steps are $S = 50{,}000$ times run through. At the end, the squared Sharpe ratio and the Sharpe ratio quotient are averaged over all repetitions for a given d, i.e., $\overline{\widehat{\mathrm{Sh}}}^2_{\mathrm{TP}}(d) = 1/S \sum_{s=1}^{S} \widehat{\mathrm{Sh}}^2_{\mathrm{TP},s}(d)$ and similarly $\bar{q}(d) = 1/S \sum_{s=1}^{S} \hat{q}_s(d)$. After that, the procedure is slightly modified at the first step. Instead of considering all 480 observations of the assets, we randomly determine a starting point from which the subsequent 120 (150, 180, 240) monthly observations are used. Figure 2 illustrates the influence of different sample sizes on the estimation of $\mathrm{Sh}^2_{\mathrm{TP}}(d)$ and $q(d)$. An estimation horizon of 480 months provides results similar to those in Figure 1. With decreasing n, both $\mathrm{Sh}^2_{\mathrm{TP}}(d)$ and $q(d)$ are more and more overestimated. Moreover, the squared Sharpe ratio becomes a *convex* function in d and the Sharpe ratio quotient is U-shaped, thus contrasting with our findings in the theoretical analysis.

Kan and Zhou (2007) derive the exact distribution of the estimated squared Sharpe ratio implied by the sample counterparts of μ and Σ if it is assumed that the underlying asset returns are multivariate normally distributed. They state that

$$\widehat{\mathrm{Sh}}^2_{\mathrm{TP}} = \hat{\mu}'\widehat{\Sigma}^{-1}\hat{\mu} \sim \frac{d}{n-d}\,\mathrm{F}_{d,n-d}\left(n \cdot \mathrm{Sh}^2_{\mathrm{TP}}\right), \tag{38}$$

Figure 3: Effect of estimation risk on the Sharpe ratio (quotient)



These figures illustrate the influence of estimation error on the squared Sharpe ratio (left hand side) and the Sharpe ratio quotient (right hand side). The black solid lines represent true values of $\mathrm{Sh}^2_{\mathrm{TP}}(d)$ and $q(d)$, respectively.

where $\mathrm{F}_{d_1,d_2}(\delta)$ denotes the noncentral F distribution with d_1 and d_2 degrees of freedom and noncentrality parameter δ (cf. Muirhead, 1982, p. 24ff). In particular,

$$\mathrm{E}\left\{\widehat{\mathrm{Sh}}^2_{\mathrm{TP}}(d)\right\} = \frac{d}{n-d-2} + \frac{n}{n-d-2}\mathrm{Sh}^2_{\mathrm{TP}}(d)\,. \tag{39}$$

By means of the latter expression, we are able to analyze the process of derivation between the estimated and the true squared Sharpe ratio – at least in terms of the expected value. This is graphically illustrated on the left hand side of Figure 3. To set the values of the true squared Sharpe ratio, we use the relationship in (37) and specify $\bar{\mu} = 0.5\%$, $\bar{\sigma}^2 = 1\%$ as well as $\rho = 1/64$. The expected Sharpe ratio quotient on the right side of Figure 3 is defined by $\mathrm{E}\{\hat{q}(d)\} := \mathrm{E}\{\widehat{\mathrm{Sh}}^2_{\mathrm{TP}}(d)\}/\mathrm{E}\{\widehat{\mathrm{Sh}}^2_{\mathrm{TP}}(d-1)\}$.

Concerning the curvature of the graphs, Figure 2 complies with Figure 3. Hence, we can proceed from the assumption, that the theoretical Sharpe ratio depicted in Figure 1 matches the reality, i.e., $\mathrm{Sh}^2_{\mathrm{TP}}$ is a strictly increasing, concave function in d. Consequently, the overall effect of the number of assets in Eq. (26) and (29) is weaker compared to the *direct* effect of d but the *indirect* effect does not completely reverse the direct effect. Furthermore, the true value of the expected squared Sharpe ratio must be considered very carefully since any sample-based estimation of $\mathrm{Sh}^2_{\mathrm{TP}}$ suffers from its strong positive bias.

4. Simulation Study

4.1. Portfolio Strategies

In Table 3, we list the investment strategies that are considered in the simulation study. The choice of strategies is motivated by the findings of Kempf et al. (2002). Focussing on the sample-based approach, the minimum variance portfolio and the equally-weighted portfolio, they show that each of these strategies can serve as an optimal allocation rule for extreme constellations of the sample size n and the level of confidence τ. We aim at sharpening the analysis by differentiating between two settings: in the first one, shortsales are completely unconstrained while the second one restricts short-selling of assets. The latter is reflected by two additional constraints in the Markowitz objective function, viz.

$$\max_w w'\mu - \frac{\lambda}{2} w'\Sigma w \quad \text{s.t.} \quad w \geq \mathbf{0}\,, \quad w'\mathbf{1} \leq 1\,. \tag{40}$$

The constraint $w'\mathbf{1} \leq 1$ represents the fact that borrowing is not allowed. Together with the second constraint, $w \geq \mathbf{0}$, this assures that neither the risky assets nor the risk-free asset is sold short. The impact of additional constraints on portfolio performance is extensively studied in the literature (see, e.g., Frahm, 2010b; Frost and Savarino, 1988; Jagannathan and Ma, 2003). DeMiguel et al. (2009b) find that constraining shortsales leads to a much better portfolio performance than any unconstrained policy. It is thus worthwhile studying the impact of τ on the portfolio performance in the case of the restricted optimization problem given by (40) in addition to conducting an analysis in the conventional setting.

The concept of minimum variance portfolios is not affected by these considerations since the MVP has its own unique optimization problem (cf. the remarks in Section 2). Nevertheless, Frahm et al. (2011) provide a way to extend this concept by noting that

$$\min_w w'\Sigma w \quad \text{s.t.} \quad w'\mathbf{1} = 1 \quad \Leftrightarrow \quad \max_w w'\bar{\mu}\mathbf{1} - \frac{\lambda}{2} w'\Sigma w \quad \text{s.t.} \quad w'\mathbf{1} = 1\,, \tag{41}$$

i.e., the minimum variance portfolio can be attained via the Markowitz objective function if the returns are assumed to have equal means, shortsales are unconstrained and no risk-free asset is available as an investment alternative. This idea yields

$$\widetilde{w}_{\text{MVP}-1} = \arg\max_w w'\hat{\bar{\mu}}\mathbf{1} - \frac{\lambda}{2} w'\widehat{\Sigma} w = \frac{\hat{\bar{\mu}}}{\lambda}\widehat{\Sigma}^{-1}\mathbf{1}\,, \tag{42}$$

if shortsales are unconstrained but a risk-free asset is available. Here, the ML esti-

Table 3: Overview of the portfolio strategies considered in the simulation study

No.	Strategy	Abbreviation	Formula
1	Optimal portfolio with shortsales unconstrained	OPT-U	$\Sigma^{-1}\mu/\lambda$
2	Optimal portfolio with shortsales constrained[a]	OPT-C	—
3	Sample-based approach with shortsales unconstrained	SBA-U	$\widehat{\Sigma}^{-1}\hat{\mu}/\lambda$
4	Sample-based approach with shortsales constrained[a]	SBA-C	—
5	Bayesian portfolio with shortsales unconstrained[b]	BAY-U	$(c_1/\lambda)\Sigma_1^{-1}\mu_1$
6	Bayesian portfolio with shortsales constrained[a]	BAY-C	—
7	Minimum variance portfolio (estimated)	MVP-E	$\widehat{\Sigma}^{-1}\mathbf{1}/(\mathbf{1}'\widehat{\Sigma}^{-1}\mathbf{1})$
8	Minimum variance portfolio (extension 1)	MVP-1	$(\hat{\bar{\mu}}/\lambda)\widehat{\Sigma}^{-1}\mathbf{1}$
9	Minimum variance portfolio (extension 2)[a]	MVP-2	—
10	Minimum variance portfolio (extension 3)[a]	MVP-3	—
11	Equally-weighted portfolio	EWP	$\mathbf{1}/d$

[a] If short-selling is constrained, no closed-form expression is available for any portfolio since the optimization problem includes an inequality restriction (cf. Eq. (40)). Numerical methods must then be applied to find a maximum.

[b] In particular, we apply the empirical Bayes approach of Jorion (1986). Furthermore, we define the constant $c_1 = (1+\tau n)(\nu_1 - 2)/(\nu_1(1+\tau n+\tau))$ (see Eq. (20)).

mate of $\bar{\mu}$ is given by $\hat{\bar{\mu}} = \mathbf{1}'\widehat{\Sigma}\hat{\mu}/(\mathbf{1}'\widehat{\Sigma}\mathbf{1})$. In the case of an additional short-selling constraint, we extend the minimum variance concept to

$$\widetilde{w}_{\mathrm{MVP-2}} = \arg\max_{w} w'\hat{\bar{\mu}}\mathbf{1} - \frac{\lambda}{2}w'\widehat{\Sigma}w \quad \text{s.t.} \quad w \geq \mathbf{0}\,, \quad w'\mathbf{1} \leq 1\,. \tag{43}$$

By contrast, Jagannathan and Ma (2003) consider the following adaption of the minimum variance concept to a setting with short-selling constraint:

$$\widetilde{w}_{\mathrm{MVP-3}} = \arg\min_{w} w'\widehat{\Sigma}w \quad \text{s.t.} \quad w \geq \mathbf{0}\,, \quad w'\mathbf{1} = 1\,. \tag{44}$$

We include all three extensions in our simulation study. Furthermore, we implement a Bayesian strategy, i.e., the sample estimators in the respective objective function are replaced by the first two moments of the predictive return distribution. In particular, we use the specifications of Jorion (1986) (see Section 2 for details). Finally, we also apply the naive diversification rule to the data, i.e., $w_{\mathrm{EWP}} = \mathbf{1}/d$.

4.2. Simulation Procedure

Our simulation study is based on stock returns from the CRSP data set that covers price information of common stocks traded on the NYSE, AMEX and NASDAQ.

Our sample incorporates monthly returns between January 1969 and December 2008. The returns of the risk-free asset are proxied by 3-month treasury bill rates provided by the U.S. Federal Reserve Bank.[10]

For each year, beginning in 1978, we consider the subsample consisting of those stocks that exhibit complete return data for the last 10 years. Altogether, 31 such asset sets are generated for the years 1978-2008. The number of stocks in the respective set ranges from 1,290 assets in the period 1969-1978 to 3,225 assets in the period 1999-2008. We proceed in this way to ensure that the input data are not affected by any survivor bias.

For the sensitivity analysis to be robust, we proceed as follows. Before each simulation run, we randomly draw a ten-year subsample of CRSP stock returns from the 31 asset sets described above. From this subsample, we independently draw d stocks. The sample mean and sample covariance matrix of these assets together with the prior parameters τ, μ_0, Σ_0 and ν_0 are used to generate an a posteriori realization of μ and Σ. This is done using the normal-inverse-Wishart model from Section 2 (see Appendix B for details). According to the normal market model given in Eq. (1) we simulate a return history with a sample size of n to which the portfolio strategies described in Table 3 are applied.

In order to run the simulation we have to specify the remaining input parameters. The prior parameter Σ_0 is structured according to the proposition of Frost and Savarino (1986), i.e., $\Sigma_0 = \bar{\sigma}_0^2(\rho_0 \mathbf{1}\mathbf{1} + (1 - \rho_0)I)$. Furthermore, we assume a priori that the expected returns are equal, i.e., $\mu_0 = \bar{\mu}_0 \mathbf{1}$. The prior parameters are specified using conservative values, or more specifically $\bar{\sigma}_0^2 = 1\%$, $\bar{\mu}_0 = 0.5\%$ and $\rho_0 = 0.25$. The confidence parameter ν_0 is chosen to equal $d + 4$. This corresponds to the smallest possible value if we want to assure the existence of the first two moments of the distribution of Σ_0.[11]

In our simulation study, the influence of the number of assets is investigated by varying the portfolio dimension, i.e., $d = \{5, 30, 100\}$. Furthermore, the simulation is conducted using different sample sizes, i.e., $n = \{12, 60, 120\}$ if $d = 5$, $n = \{60, 120, 240\}$ if $d = 30$ and $n = \{120, 240, 480\}$ if $d = 100$. Concerning the level of confidence τ, our analysis pursues different goals. First, the findings of Section 3 should be taken into account appropriately, i.e., (i) conservative values are assigned to $\mathrm{E}(\mathrm{Sh}^2_{\mathrm{TP}})$ bearing in mind the vast positive bias of the sample-based estimate $\widehat{\mathrm{Sh}}^2_{\mathrm{TP}}$,

[10] The data can be accessed online at `http://www.federalreserve.gov/releases/H15/data/Monthly/H15_TB_M3.txt`.

[11] For example, it holds that $\mathrm{var}(\sigma_i^2) = 2\bar{\sigma}_0^4/((\nu_0 - d - 1)^2(\nu_0 - d - 3))$ with $\Sigma = [\sigma_{ij}]$ and $\sigma_i^2 = \sqrt{\sigma_{ii}}$ $(i = 1, \ldots, d)$ if $\nu_0 \geq d + 4$ (cf. Press, 1972).

(ii) we account for the fact that $\mathrm{Sh}^2_{\mathrm{TP}}$ is strictly increasing, but concave in d. This assures us that the level of confidence is chosen conservatively and increases with the number of assets in a less than linear fashion.

Second, we want to reexamine the theoretical findings of Kempf et al. (2002) (cf. Table 1) who derived optimal strategies by considering the limits of n and τ, respectively. This gives rise to values of $\mathrm{E}(\mathrm{Sh}^2_{\mathrm{TP}})$ which are less conservative but still refer to reality. More precisely, we set $\mathrm{E}(\mathrm{Sh}^2_{\mathrm{TP}}) = \{0.03, 0.09, 0.15, 0.21, 0.27\}$ if $d = 30$ and divide each value by three if $d = 5$. For high-dimensional data ($d = 100$), the values of the expected squared Sharpe ratio are multiplied by a factor of two.

4.3. Performance Measurement

Given the vector of portfolio weights $\widetilde{w}_{m,s}$ of the mth strategy, the true mean and the true variance of the portfolio return for simulation run s are given by $\mu_{m,s} = \widetilde{w}'_{m,s}\mu$ and $\sigma^2_{m,s} = \widetilde{w}'_{m,s}\Sigma\widetilde{w}_{m,s}$. We measure the performance of the portfolio strategies by computing the true Sharpe ratio of strategy m in simulation run s, defined as

$$\mathrm{Sh}_{m,s} = \frac{\mu_{m,s}}{\sigma_{m,s}}\,. \tag{45}$$

Note that we are able to evaluate the *true* performance of each strategy since we know the true parameters of the return distribution. However, the true performances vary with the simulation runs owing to the design of our simulation study. Thus, we average the performances over all simulation runs, viz.

$$\overline{\mathrm{Sh}}_m = \frac{1}{S}\sum_{s=1}^{S}\mathrm{Sh}_{m,s}\,. \tag{46}$$

We are not permitted to *test* for the best strategy if the ranking of the strategies and, consequently, the benchmark strategy is identified using the simulated return observations. This is typically referred to as data mining. Instead of conducting hypothesis tests, we use pure computer power to diminish the standard error of the estimator for the expected value of the respective performance measure and, thus, to minimize the probability that the ranking of strategies could be incorrect. For this purpose, $S = 10{,}000$ simulation runs serve well.

4.4. Discussion of Results

Tables C.4 - C.6 in the appendix report the average true Sharpe ratios for the different portfolios and for various constellations of the number of assets, the sample size and the parameter of confidence. The major results are graphically represented in Figure 4. First, when looking at the pictures on the left-hand side of Figure 4, it becomes clear that an additional short-selling constraint lowers the maximum attainable Sharpe ratio (black solid lines) while it improves portfolio performance compared to unrestricted portfolio optimization. Second, we present our findings with regard to the propositions of Kempf et al. (2002) which can be summarized as follows:

1. The *equally-weighted portfolio* is preferable if no market information is available and the assets are very inhomogeneous.
2. One should rely on the *global minimum variance portfolio* whenever the assets are completely homogeneous.
3. The *sample-based approach* becomes optimal only in the theoretical case of an infinite sampling period and very inhomogeneous assets.

ad 1. Part (b) of Figure 4 illustrates the average portfolio Sharpe ratios if the market consists of five assets, twelve monthly return observations are available for each asset and shortsales are constrained. In this constellation, the EWP outperforms the other strategies if the confidence parameter is equal to or higher than 0.0048. Note that both for $d = 30$ / $n = 60$ and for $d = 100$ / $n = 120$, the performance of the EWP diminishes compared to the performance of the sample-based approach if τ increases (see Tables C.5 and C.6). By definition of the simulation study, the level of confidence is also an indicator for the inhomogeneity of the assets. Thus, we can state more precisely that the EWP is a preferable allocation rule if τ is large, d is small and $n/d \searrow 1$. It is noteworthy that the latter holds only for the standard setting without short-selling constraints. Otherwise, the EWP is less and less competitive for increasingly inhomogeneous assets.

ad 2. In accordance with the conclusions of Kempf et al. (2002), we find the MVP to be an outperforming strategy if the assets are very homogeneous, i.e., the investor has great confidence in the expert's prior assessment (see parts (b), (d) and (h) in Figure 4). However, there is an exception to this rule if d is large and $n/d \searrow 1$. Then, the EWP turns out to be the best strategy. If shortsales are constrained and τ is small, the MVP is in all cases the best strategy (cf. parts (a), (c), (e) and (g) of Figure 4). One should bear in mind that a setting with short-selling constraint involves an *extended* minimum variance concept. Under these extensions, the MVP-2 strategy proposed by Frahm et al. (2011) performs best and, consequently, was used as a benchmark.

ad 3. As may be seen from parts (d) and (h) of Figure 4, the SBA is found to be the best strategy if τ and n/d have large values. This coincides with the findings of Kempf et al. (2002). In addition, the SBA is highly competitive if τ and d are large and $n/d \searrow 1$ (cf. part (f) of Figure 4). If shortsales are constrained, similar results can be derived (see parts (c), (e) and (g) of Figure 4).

In general, we are convinced that the effect of the sample size n on portfolio performance should be measured relative to the number of assets d. This allows a more differentiated analysis and leads – in some constellations – to deviating results. Moreover, our analysis shows that it is not necessary to let the confidence parameter approach infinity in order to stress its influence on portfolio performance. On the contrary, relevant values of τ are identified to range between zero and a number much smaller than one. Note that the upper limit of relevant values for τ mainly depends on the number of assets (see Section 3).

Figure 4: Performance of the portfolio strategies

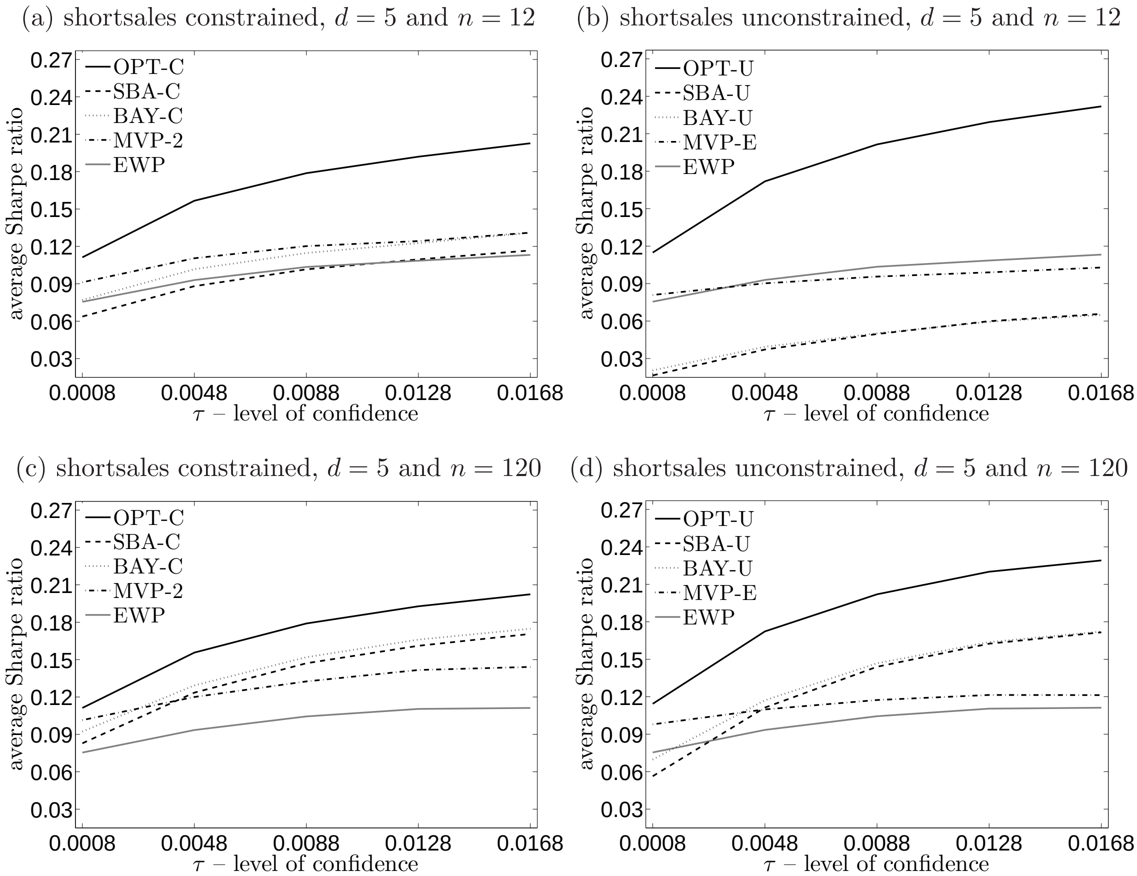
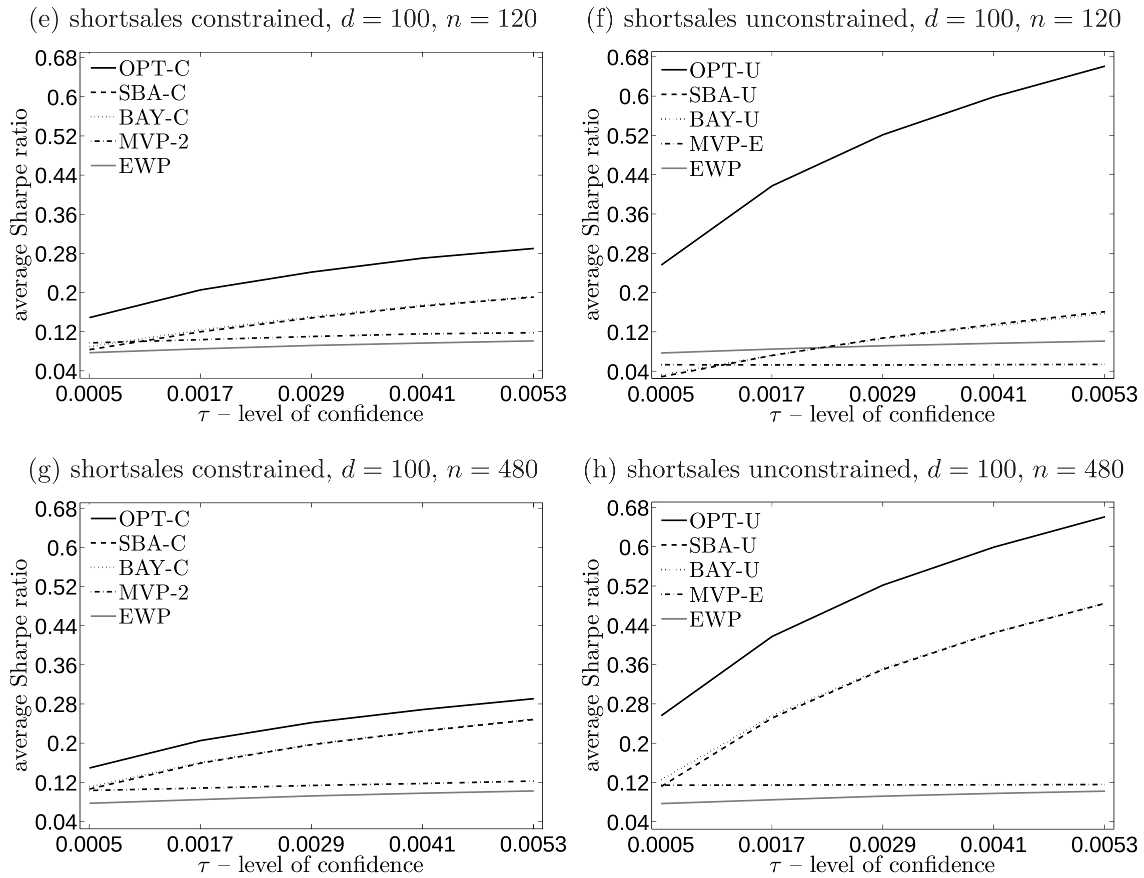


Figure 4: Performance of the portfolio strategies (cont'd)



The results of empirical studies in the field of portfolio performance (see, e.g., Frahm et al., 2011; DeMiguel et al., 2009a,b) may be used to deduce a range of *realistic* values for τ. These studies detect that the performance of the sample-based approach can be slightly improved by instead using the empirical Bayes approach of Jorion (1986). Furthermore, both strategies are outperformed by the minimum variance portfolio. While the former coincides with our results (cf. Figure 4), the latter finding is retrieved in our simulation study only for a sufficiently high level of confidence. Thus, an upper bound for a *realistic* value of the confidence parameter τ may be assessed as follows. Consider the results of our simulation study for a sample size of $n = 120$ months. This is a common estimation window used both by practitioners and researchers in the field of financial data. Next, find the smallest value of τ for which the performance of the sample-based approach exceeds the performance of the minimum variance portfolio. This value is the upper bound to be selected.

It turns out that this upper bound is in all cases the second smallest of the five τ values, irrespective of whether short-selling is constrained and how many assets the market contains. More precisely, it should be realistic to assume $\tau < 0.0048$ if the market consists of five assets (cf. Table C.4). A larger market produces an even higher level of confidence, i.e., if the number of assets is 30, it is realistic to assume that τ takes values below 0.0027, and we expect $\tau < 0.0017$ if $d = 100$ (see Tables C.5 and C.6).

Remember that the level of confidence is not a direct input parameter of our simulation study, but rather is calculated from the range of values for $\mathrm{E}(\mathrm{Sh}^2_{\mathrm{TP}})$. Accordingly, the upper bounds for τ found above imply realistic values for the Sharpe ratio of the tangency portfolio: for $d = 5$ it may be realistic to assume $\mathrm{E}(\mathrm{Sh}^2_{\mathrm{TP}}) < 0.03$; for $d = 30$ it should be the case that $\mathrm{E}(\mathrm{Sh}^2_{\mathrm{TP}}) < 0.09$ and for $d = 100$ it can be expected that $\mathrm{E}(\mathrm{Sh}^2_{\mathrm{TP}}) < 0.18$. This comes full circle back to our initial claim to consider possible values for $\mathrm{E}(\mathrm{Sh}^2_{\mathrm{TP}})$ very carefully; see the remarks in Section 3.

5. Conclusion

We derive an analytical expression for the confidence parameter τ of the normal-inverse-Wishart prior. By the specification of confidence, the investor is enabled to optimize his/her portfolio assuming that expectations and the covariance structure of the asset returns are a priori assessed by an expert. The confidence in prior knowledge increases the more assets are on the market and the better the true potential performance – indicated by the Sharpe ratio of the tangency portfolio – is approximated by that one following from the expert assessments on the prior parameters.

In the literature, the Sharpe ratio of the tangency portfolio is usually approximated by the empirical Sharpe ratio of a well-diversified stock index, using long estimation horizons. We propose to be cautious about relying on these approximations since any sample-based estimation of $\mathrm{Sh}^2_{\mathrm{TP}}$ suffers from a strong positive bias. Furthermore, we show that the Sharpe ratio of the TP increases with the number of assets. This, in turn, weakens the direct effect of d on the level of confidence τ.

The results of our simulation study show that the performance of investment strategies is generally very sensitive to changes in the confidence parameter. Additionally, the *relative* portfolio performance varies with τ. In our view a meaningful performance analysis also involves the number of assets d and the 'effective sample size' n/d. The sample-based strategy turns out to be most competitive in a situation in which it is appropriate to strongly mistrust the prior knowledge – when the effective sample size is large. If n/d and d are small, the equally-weighted portfolio is the best-performing allocation rule. On the contrary, the global minimum variance portfolio is preferable if it is advisable to place a high level of confidence in the expert's prior assessments.

We further contribute to the literature on confidence in prior knowledge in two ways: first we clarify that – in general – values close to zero should be assigned to the confidence parameter τ which is in accordance with some of the propositions in the context of the Black-Litterman model. Second, we derive upper limits for the confidence parameter depending on the number of assets, i.e., we propose to choose $\tau < 0.0048$ if $d = 5$, $\tau < 0.0027$ if $d = 30$ and $\tau < 0.0017$ if $d = 100$. Neglecting the confidence parameter, or – in other words – assuming that $\tau = 1$ would seriously reduce the influence of prior knowledge. As a result, the objective of a robust parameter estimation would clearly be missed.

Acknowledgement

The author gratefully acknowledges financial support by the German Research Foundation (DFG).

Appendix A. The Wishart and related distributions

Consider a set of d-dimensional random variables $\{X_1, \ldots, X_n\}$ that are independent and multivariate normally distributed with zero expected value and with the same covariance matrix:

$$X_t \sim \mathcal{N}(\mathbf{0}, \Sigma) , \quad t = 1 \ldots, n . \tag{A.1}$$

The *Wishart distribution* with n degrees of freedom is the distribution of the random matrix $W \equiv X_1 X_1' + \ldots + X_n X_n'$. We use the following notation to indicate that W is a Wishart-distributed matrix with n degrees of freedom and dispersion parameter Σ:

$$W \sim W(n, \Sigma) . \tag{A.2}$$

The expectations and cross-covariances, respectively, can be expressed as follows:

$$\mathrm{E}(W) = n\Sigma , \tag{A.3}$$

$$\mathrm{Cov}(W_{mn}, W_{pq}) = n(\Sigma_{mp}\Sigma_{nq} + \Sigma_{mq}\Sigma_{np}) . \tag{A.4}$$

For the inverse of W, it holds that

$$W^{-1} \sim \mathrm{IW}\left(n + d + 1, \Sigma^{-1}\right) . \tag{A.5}$$

The first- and second-order moments are given by

$$\mathrm{E}\left(W^{-1}\right) = \frac{\Sigma^{-1}}{n - d - 1} , \tag{A.6}$$

$$\mathrm{Cov}\left(W^{-1}_{mn}, W^{-1}_{pq}\right) = \frac{\frac{2}{n-d-1}\Sigma^{-1}_{mn}\Sigma^{-1}_{pq} + \Sigma^{-1}_{mp}\Sigma^{-1}_{nq} + \Sigma^{-1}_{mq}\Sigma^{-1}_{np}}{(n-d)(n-d-1)(n-d-3)} , \quad n - d > 3 . \tag{A.7}$$

Suppose a generic $(k \times d)$ matrix A with $k \leq d$. Then, if

$$W \sim W(n, \Sigma) \Rightarrow AWA' \sim W(n, A\Sigma A') . \tag{A.8}$$

Appendix B. Parameter Generation Using the NIW Model

Assume that the prior parameters ν_0, μ_0 and Σ_0 as well as the sample mean $\hat{\mu}$ and the sample covariance matrix $\widehat{\Sigma}$ are given inputs. The posterior parameters ν_1, μ_1 and Σ_1 are computed using Eq. (18). Let Z_t be a d-dimensional random vector following the multivariate standard normal distribution. We draw ν_1 independent copies from Z_t and calculate $\sum_{t=1}^{\nu_1} Z_t Z_t'$. With

$$\hat{\mu}_{Z_1} = \frac{1}{\nu_1}\sum_{t=1}^{\nu_1} Z_t \quad \text{and} \quad \widehat{\Sigma}_{Z_1} = \frac{1}{\nu_1}\sum_{t=1}^{\nu_1} \left(Z_t - \hat{\mu}_{Z_1}\right)\left(Z_t - \hat{\mu}_{Z_1}\right)' \tag{B.1}$$

it follows that $\sum_{t=1}^{\nu_1} Z_t Z_t' \equiv \nu_1 \widehat{\Sigma}_{Z_1} + \nu_1 \hat{\mu}_{Z_1}\hat{\mu}_{Z_1}'$. From Appendix A we know that this quantity is Wishart-distributed with ν_1 degrees of freedom and scale matrix I. Thus, we conclude that

$$\left(\nu_1 \widehat{\Sigma}_{Z_1} + \nu_1 \hat{\mu}_{Z_1}\hat{\mu}_{Z_1}'\right)^{-1} \sim \mathrm{IW}\left(\nu_1 + d + 1, I\right) . \tag{B.2}$$

Defining $\nu_1\Sigma_1 \equiv (\nu_1\Sigma_1)^{1/2}(\nu_1\Sigma_1)^{1/2}$, an a posteriori realization of Σ under the normal-inverse-Wishart model is given by

$$\Sigma = \Sigma_1^{1/2}\left(\widehat{\Sigma}_{Z_1} + \hat{\mu}_{Z_1}\hat{\mu}_{Z_1}'\right)^{-1}\Sigma_1^{1/2} . \tag{B.3}$$

This is due to the distribution law of transformations of Wishart matrices (see again Appendix A). Now we can draw from the conditional posterior distribution of μ by applying the previously generated a posteriori realization of Σ. At this point, it is useful to distinguish two cases:

(a) $n_1 = (1 + \tau n)/\tau$ is an integer. Then draw n_1 independent realizations from Z_t and calculate $\hat{\mu}_{Z_2} = 1/n_1 \sum_{t=1}^{n_1} Z_t$. An a posteriori realization of μ can be attained as follows:

$$\mu = \mu_1 + \Sigma^{1/2}\hat{\mu}_{Z_2} = w_\tau \bar{\mu}\mathbf{1} + \left(1 - w_\tau\right)\hat{\mu} + \Sigma^{1/2}\hat{\mu}_{Z_2} , \tag{B.4}$$

where we define $\Sigma = \Sigma^{1/2}\Sigma^{1/2}$.

(b) If $n_1 = (1 + \tau n)/\tau$ is not an integer, we draw n independent realizations from Z_t and calculate $\hat{\mu}_{Z_3} = 1/n \sum_{t=1}^{n} Z_t$. Recall that it holds $\hat{\mu}_{Z_3} \sim \mathcal{N}(\mathbf{0}, I/n)$ (cf. Press, 1972, p. 183). An a posteriori realization of μ is then calculated as

$$\mu = \mu_1 + \sqrt{\frac{n}{n_1}}\Sigma^{1/2}\hat{\mu}_{Z_3} = w_\tau \bar{\mu}\mathbf{1} + \left(1 - w_\tau\right)\hat{\mu} + \sqrt{1 - w_\tau}\Sigma^{1/2}\hat{\mu}_{Z_3} . \tag{B.5}$$

Appendix C. Tables

Table C.4: Portfolio Sharpe ratios, $d = 5$

τ	0.0008			0.0048			0.0088			0.0128			0.0168		
n	12	60	120	12	60	120	12	60	120	12	60	120	12	60	120
Panel A: Parameters μ and Σ known															
OPT-U	0.115	0.115	0.114	0.172	0.172	0.172	0.201	0.201	0.202	0.219	0.220	0.220	0.232	0.232	0.229
OPT-C	0.111	0.111	0.111	0.157	0.157	0.156	0.179	0.179	0.179	0.192	0.192	0.193	0.203	0.204	0.202
Panel B: Investor's perspective															
SBA-U	0.016	0.042	0.056	0.037	0.088	0.111	0.050	0.114	0.144	0.060	0.133	0.163	0.066	0.144	0.172
SBA-C	0.064	0.077	0.083	0.088	0.113	0.123	0.102	0.133	0.147	0.110	0.146	0.161	0.117	0.156	0.171
BAY-U	0.021	0.053	0.070	0.039	0.095	0.117	0.050	0.118	0.147	0.059	0.136	0.164	0.065	0.146	0.172
BAY-C	0.077	0.087	0.092	0.102	0.121	0.129	0.115	0.141	0.152	0.123	0.153	0.166	0.131	0.163	0.175
MVP-E	0.081	0.096	0.098	0.090	0.108	0.110	0.096	0.114	0.117	0.099	0.118	0.121	0.103	0.122	0.121
MVP-1	0.017	0.045	0.059	0.028	0.064	0.080	0.033	0.075	0.094	0.037	0.082	0.100	0.042	0.087	0.101
MVP-2	0.091	0.101	0.101	0.111	0.121	0.120	0.120	0.132	0.133	0.124	0.139	0.142	0.131	0.144	0.144
MVP-3	0.088	0.097	0.098	0.102	0.109	0.110	0.108	0.116	0.118	0.111	0.120	0.122	0.115	0.125	0.125
EWP	0.076	0.075	0.075	0.093	0.093	0.093	0.104	0.102	0.104	0.109	0.109	0.110	0.113	0.114	0.110

This table reports the monthly Sharpe ratio for the portfolio strategies listed in Table 3 as an average over the true Sharpe ratios of the $S = 10{,}000$ simulation runs. The market consists of $d = 5$ risky assets and one risk-free asset. In Panel A, we report the results of the optimal strategy both with shortsales unconstrained and constrained as well as the results of the true minimum variance portfolio. Note that these values only serve as a benchmark since the true parameters μ and Σ are not known to the investor. Panel B reflects the investor's perspective since here, the parameters μ and Σ are assumed to be unknown and hence must be estimated. According to Table 3, the strategies are abbreviated as follows: SBA – sample-based approach, BAY – Bayesian strategy, MVP – minimum variance portfolio, EWP – equally-weighted portfolio. With the 'C' and 'U' adjunct we indicate whether shortsales are constrained or unconstrained. For the Bayesian strategy, we use the specification of Jorion (1986). The sample estimate of the minimum variance portfolio is named with MVP-E, while the abbreviations MVP-1, MVP-2 and MVP-3 refer to the extended concept of minimum variance portfolios (see Section 4.1 for details).

Table C.5: Portfolio Sharpe ratios, $d = 30$

τ	0.0007			0.0027			0.0047			0.0067			0.0087		
n	60	120	240	60	120	240	60	120	240	60	120	240	60	120	240
Panel A: Parameters μ and Σ known															
OPT-U	0.197	0.197	0.196	0.302	0.302	0.301	0.367	0.368	0.367	0.412	0.412	0.412	0.446	0.447	0.448
OPT-C	0.159	0.158	0.158	0.214	0.214	0.215	0.247	0.247	0.247	0.271	0.271	0.270	0.288	0.288	0.288
Panel B: Investor's perspective															
SBA-U	0.038	0.063	0.090	0.085	0.137	0.184	0.122	0.190	0.249	0.150	0.229	0.294	0.171	0.261	0.330
SBA-C	0.085	0.098	0.111	0.120	0.141	0.162	0.146	0.171	0.196	0.164	0.195	0.220	0.180	0.211	0.239
BAY-U	0.049	0.081	0.110	0.092	0.147	0.193	0.126	0.197	0.254	0.152	0.233	0.297	0.172	0.264	0.332
BAY-C	0.100	0.109	0.120	0.132	0.150	0.168	0.157	0.178	0.201	0.174	0.201	0.224	0.190	0.218	0.243
MVP-E	0.098	0.119	0.127	0.099	0.122	0.131	0.102	0.124	0.134	0.103	0.126	0.135	0.103	0.127	0.135
MVP-1	0.051	0.079	0.103	0.060	0.091	0.111	0.066	0.097	0.116	0.067	0.100	0.119	0.073	0.104	0.121
MVP-2	0.118	0.122	0.125	0.130	0.133	0.135	0.137	0.140	0.142	0.142	0.147	0.146	0.147	0.151	0.152
MVP-3	0.116	0.120	0.124	0.123	0.128	0.132	0.129	0.132	0.136	0.132	0.137	0.139	0.135	0.139	0.143
EWP	0.090	0.089	0.089	0.103	0.104	0.103	0.114	0.113	0.114	0.120	0.122	0.121	0.127	0.126	0.126

This table reports the monthly Sharpe ratio for the portfolio strategies listed in Table 3 as an average over the true Sharpe ratios of the $S = 10{,}000$ simulation runs. The market consists of $d = 30$ risky assets and one risk-free asset. In Panel A, we report the results of the optimal strategy both with shortsales unconstrained and constrained as well as the results of the true minimum variance portfolio. Note that these values only serve as a benchmark since the true parameters μ and Σ are not known to the investor. Panel B reflects the investor's perspective since here, the parameters μ and Σ are assumed to be unknown and hence must be estimated. According to Table 3, the strategies are abbreviated as follows: SBA – sample-based approach, BAY – Bayesian strategy, MVP – minimum variance portfolio, EWP – equally-weighted portfolio. With the 'C' and 'U' adjunct we indicate whether shortsales are constrained or unconstrained. For the Bayesian strategy, we use the specification of Jorion (1986). The sample estimate of the minimum variance portfolio is named with MVP-E, while the abbreviations MVP-1, MVP-2 and MVP-3 refer to the extended concept of minimum variance portfolios (see Section 4.1 for details).

Table C.6: Portfolio Sharpe ratios, $d = 100$

τ	0.0005			0.0017			0.0029			0.0041			0.0053		
n	120	240	480	120	240	480	120	240	480	120	240	480	120	240	480
Panel A: Parameters μ and Σ known															
OPT-U	0.256	0.256	0.256	0.417	0.417	0.417	0.521	0.522	0.522	0.599	0.598	0.599	0.661	0.661	0.661
OPT-C	0.149	0.148	0.149	0.205	0.206	0.205	0.242	0.242	0.242	0.270	0.270	0.269	0.290	0.290	0.291
Panel B: Investor's perspective															
SBA-U	0.029	0.072	0.112	0.072	0.173	0.251	0.108	0.252	0.350	0.135	0.311	0.425	0.161	0.362	0.484
SBA-C	0.083	0.095	0.106	0.120	0.140	0.159	0.148	0.174	0.197	0.172	0.201	0.225	0.191	0.222	0.248
BAY-U	0.032	0.083	0.126	0.072	0.178	0.256	0.106	0.255	0.352	0.133	0.313	0.426	0.157	0.363	0.485
BAY-C	0.089	0.100	0.110	0.124	0.143	0.161	0.151	0.176	0.198	0.174	0.203	0.225	0.192	0.223	0.249
MVP-E	0.053	0.098	0.115	0.053	0.098	0.115	0.053	0.100	0.115	0.053	0.099	0.115	0.054	0.100	0.116
MVP-1	0.032	0.075	0.100	0.035	0.076	0.100	0.036	0.079	0.102	0.037	0.081	0.103	0.039	0.083	0.103
MVP-2	0.097	0.101	0.103	0.104	0.107	0.108	0.110	0.113	0.113	0.116	0.119	0.118	0.118	0.122	0.122
MVP-3	0.096	0.100	0.103	0.101	0.105	0.107	0.105	0.109	0.111	0.109	0.112	0.114	0.111	0.115	0.117
EWP	0.077	0.077	0.077	0.085	0.085	0.085	0.092	0.092	0.092	0.097	0.097	0.098	0.101	0.101	0.102

This table reports the monthly Sharpe ratio for the portfolio strategies listed in Table 3 as an average over the true Sharpe ratios of the $S = 10{,}000$ simulation runs. The market consists of $d = 100$ risky assets and one risk-free asset. In Panel A, we report the results of the optimal strategy both with shortsales unconstrained and constrained as well as the results of the true minimum variance portfolio. Note that these values only serve as a benchmark since the true parameters μ and Σ are not known to the investor. Panel B reflects the investor's perspective since here, the parameters μ and Σ are assumed to be unknown and hence must be estimated. According to Table 3, the strategies are abbreviated as follows: SBA – sample-based approach, BAY – Bayesian strategy, MVP – minimum variance portfolio, EWP – equally-weighted portfolio. With the 'C' and 'U' adjunct we indicate whether shortsales are constrained or unconstrained. For the Bayesian strategy, we use the specification of Jorion (1986). The sample estimate of the minimum variance portfolio is named with MVP-E, while the abbreviations MVP-1, MVP-2 and MVP-3 refer to the extended concept of minimum variance portfolios (see Section 4.1 for details).

References

Ando, A., Kaufman, G.M., 1965. Bayesian analysis of the independent multinormal process - neither mean nor precision known. Journal of the American Statistical Association 60, 347–358.

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. Review of Financial Studies 4, 315–342.

Black, F., Litterman, R.B., 1991. Asset allocation: Combining investor views with market equilibrium. Journal of Fixed Income 1, 7–18.

Black, F., Litterman, R.B., 1992. Global portfolio optimization. Financial Analysts Journal 48, 28–43.

Brennan, M.J., Chordia, T., Subrahmanyam, A., 1998. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. Journal of Financial Economics 49, 345–373.

Brown, S.J., 1976. Optimal Portfolio Choice Under Uncertainty: A Bayesian Approach. Ph.D. thesis. University of Chicago.

Chopra, V.K., Ziemba, W.T., 1993. The effect of errors in means, variances, and covariances on optimal portfolio choice. Journal of Portfolio Management 19, 6–11.

Cogley, T., Sargent, T.J., 2008. The market price of risk and the equity premium: A legacy of the Great Depression? Journal of Monetary Economics 55, 454–476.

DeMiguel, V., Garlappi, L., Nogales, F.J., Uppal, R., 2009a. A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. Management Science 55, 798–812.

DeMiguel, V., Garlappi, L., Uppal, R., 2009b. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? The Review of Financial Studies 22, 1915–1953.

Dimson, E., Marsh, P., Staunton, M., 2002. Triumph of the Optimists: 101 Years of Global Investment Returns. Princeton University Press, Princeton.

Dimson, E., Marsh, P., Staunton, M., 2003. Global evidence on the equity risk premium. Journal of Applied Corporate Finance 15, 27–38.

Frahm, G., 2010a. An analytical investigation of estimators for expected asset returns from the perspective of optimal asset allocation. Working Paper, University of Cologne.

Frahm, G., 2010b. Linear statistical inference for global and local minimum variance portfolios. Statistical Papers 51, 789–812.

Frahm, G., Wickern, T., Wiechers, C., 2011. Multiple tests for the performance of different investment strategies. Advances in Statistical Analysis, forthcoming.

Frost, P.A., Savarino, J.E., 1986. An empirical Bayes approach to efficient portfolio selection. Journal of Financial and Quantitative Analysis 21, 293–305.

Frost, P.A., Savarino, J.E., 1988. For better performance: Constrain portfolio weights. Journal of Portfolio Management 15, 29–34.

Fusai, G., Meucci, A., 2003. Assessing views. Risk 16, S18–S21.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. Bayesian Data Analysis. Chapman & Hall/CRC, Boca Raton.

Gospodinov, N., Kan, R., Robotti, C., 2010. On the Hansen-Jagannathan distance with a no-arbitrage constraint. Working Paper, Federal Reserve Bank of Atlanta.

He, G., Litterman, R., 2002. The intuition behind Black-Litterman model portfolios. Working Paper.

Idzorek, T.M., 2007. A step-by-step guide to the Black-Litterman model: Incorporating user-specified confidence levels, in: Satchell, S. (Ed.), Forecasting Expected Returns in the Financial Markets, Elsevier/AP, Amsterdam, pp. 17-38.

Jagannathan, R., Ma, T., 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. Journal of Finance 58, 1651–1684.

Jeffreys, H., 1961. Theory of Probability. Oxford University Press, London.

Jobson, J.D., Korkie, B., 1979. Improved estimation for Markowitz portfolios using James-Stein type estimators. Proceedings of the American Statistical Association (Business and Economic Statistics) 1, 279–284.

Jobson, J.D., Korkie, B., 1980. Estimation for Markowitz efficient portfolios. Journal of the American Statistical Association 75, 544–554.

Jobson, J.D., Korkie, B., 1981. Putting Markowitz theory to work. Journal of Portfolio Management 7, 70–74.

Jobson, J.D., Korkie, B., 1982. Potential performance and tests of portfolio efficiency. Journal of Financial Economics 10, 433–466.

Jobson, J.D., Korkie, B., 1984. On the Jensen measure and marginal improvements in portfolio performance: A note. Journal of Finance 39, 245–251.

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.

Kalymon, B.A., 1971. Estimation risk in the portfolio selection model. Journal of Financial and Quantitative Analysis 6, 559–582.

Kan, R., Robotti, C., 2008. The exact distribution of the Hansen-Jagannathan bound. Working Paper, University of Toronto and Federal Reserve Bank of Atlanta.

Kan, R., Zhou, G., 2007. Optimal portfolio choice with parameter uncertainty. Journal of Financial and Quantitative Analysis 42, 621–656.

Kempf, A., Kreuzberg, K., Memmel, C., 2002. How to incorporate estimation risk into Markowitz optimization, in: Chamoni, P., Leisten, R., Martin, A., Minnemann, J., Stadtler, H. (Eds.), Operations Research Proceedings 2001, Springer, Berlin. pp. 175–182.

Kempf, A., Memmel, C., 2002. Schätzrisiken in der Portfoliotheorie, in: Kleeberg, J.M., Rehkugler, H. (Eds.), Handbuch Portfoliomanagement, Uhlenbruch, Bad Soden. pp. 893–919.

Klein, R.W., Bawa, V.S., 1976. The effect of estimation risk on optimal portfolio choice. Journal of Financial Economics 3, 215–231.

Kotz, S., Nadarajah, S., 2004. Multivariate t Distributions and Their Applications. Cambridge University Press, Cambridge.

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.

MacKinlay, A.C., 1995. Multifactor models do not explain deviations from the CAPM. Journal of Financial Economics 38, 3–28.

Mao, J.C.T., Särndal, C.E., 1966. A decision theory approach to portfolio selection. Management Science 12, B–323–334.

Markowitz, H., 1952. Portfolio selection. Journal of Finance 7, 77–91.

Memmel, C., 2004. Schätzrisiken in der Portfoliotheorie: Auswirkungen und Möglichkeiten der Reduktion. Ph.D. thesis. University of Cologne.

Merton, R.C., 1980. On estimating the expected return on the market: An exploratory investigation. Journal of Financial Economics 8, 323–361.

Meucci, A., 2005. Risk and Asset Allocation. Springer, Berlin.

Meucci, A., 2010. The Black-Litterman approach, in: Cont, R. (Ed.), The Encyclopedia of Quantitative Finance. John Wiley and Sons, Chichester.

Michaud, R.O., 1989. The Markowitz optimization enigma: Is 'optimized' optimal? Financial Analysts Journal 45, 31–42.

Muirhead, R.J., 1982. Aspects of Multivariate Statistical Theory. John Wiley and Sons, New Jersey.

Press, J.S., 1972. Applied Multivariate Analysis. Holt, Rinehart and Winston, New York.

Raiffa, H., Schlaifer, R., 1961. Applied Statistical Decision Theory. MIT Press, Cambridge.

Satchell, S., Scowcroft, A., 2000. A demystification of the Black-Litterman model: Managing quantitative and traditional portfolio construction. Journal of Asset Management 1, 138–150.

Stein, C., 1956. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, in: Proceedings of the 3rd Berkeley Symposium on Probability and Statistics, University of California Press, Berkeley. pp. 197–206.

Tobin, J., 1958. Liquidity preference as behavior towards risk. The Review of Economic Studies 25, 65–86.

Walters, J., 2010. The factor tau in the Black-Littterman model. Working Paper.

Walters, J., 2011. The Black-Litterman model in detail. Working Paper.