THE RIMINI CENTRE FOR ECONOMIC ANALYSIS

WP 47_13

Hao Liu
University of Konstanz, Germany

Winfried Pohlmeier
University of Konstanz, Germany
ZEW Centre for European Economic Research, Germany
The Rimini Centre for Economic Analysis (RCEA), Italy

RISK PREFERENCES AND ESTIMATION RISK IN PORTFOLIO CHOICE

Risk Preferences and Estimation Risk in Portfolio Choice

Hao Liu
University of Konstanz, CoFE

Winfried Pohlmeier*
University of Konstanz,
CoFE, ZEW, RCEA

August 3, 2013

Abstract

This paper analyzes the estimation risk of efficient portfolio selection. We use the concept of certainty equivalent as the basis for a well-defined statistical loss function and a monetary measure to assess estimation risk. For given risk preferences we provide analytical results for different sources of estimation risk such as sample size, dimension of the portfolio choice problem and correlation structure of the return process. Our results show that theoretically sub-optimal portfolio choice strategies turn out to be superior once estimation risk is taken into account. Since estimation risk crucially depends on risk preferences, the choice of the estimator for a given portfolio strategy becomes endogenous. We show that a shrinkage approach accounting for estimation risk in both, mean and covariance of the return vector, is generally superior to simple theoretically suboptimal strategies. Moreover, focusing on just one source of estimation risk, e.g. risk reduction in covariance estimation, can lead to suboptimal portfolios.

JEL classification: $G11, G12, G17$
Keywords: efficient portfolio, estimation risk, certainty equivalent, shrinkage

*Corresponding author: Department of Economics, University of Konstanz, Box 124, 78457 Konstanz, Germany, Phone: +49 7531 88-2660, Email: Winfried.Pohlmeier@uni-konstanz.de. Earlier versions of the paper have been presented at the conference "Recent Advances in Finance", Warwick Business School, the Constance-Strasbourg Workshop on Banking and Finance and the 4^{th} Quantitative Finance Workshop of the RCEA. Financial support by the Fritz Thyssen Foundation is gratefully acknowledged. For helpful comments we would like to thank Roxana Halbleib, Gabriel Frahm and Michael Wolf. The usual disclaimer applies.

1 Introduction

Empirical estimates of mean-variance efficient portfolio weights often turn out to be unrealistic and reveal considerable standard errors. The problem is well-known in empirical finance and has been documented in several previous studies (e.g. Black and Litterman (1992), Best and Grauer (1991) and Britten-Jones (1999)). The low precision of estimated portfolio weights coincides with many findings from horse races between different portfolio selection strategies showing that theoretically suboptimal portfolio choices do better in empirical applications than theoretically more efficient strategies. For instance, DeMiguel, Garlappi, and Uppal (2009) show that the equally weighted portfolio cannot be outperformed by the efficient portfolio and several other optimized portfolio strategies in terms of out of sample prediction performance for several performance measures.

This paper sheds more light on classical portfolio selection rules in the tradition of Markowitz (1952) when parameters of the underlying return distribution have to be estimated and estimation risk is taken into account. Our paper is the first to study analytically the performance of empirical efficient portfolio in relation to its benchmark, i.e. the theoretical efficient portfolio. In particular, we focus on the loss in performance due to estimation depending on (i) sample size, (ii) dimension of the portfolio choice problem, (iii) correlation structure of returns and (iv) the investor's risk preferences. By concentrating on the efficient portfolio, our results generalize previous findings for portfolio strategies also satisfying the budget constraint and incorporate those as special cases.

Our analysis of the role of risk preferences for estimation risk provides novel insights into the functioning of shrinkage strategies in portfolio analysis. In particular, we can show that shrinkage of parameters estimates or shrinkage of the estimated portfolio weights have a common representation in terms of the risk preference parameter. In terms of the certainty equivalent (CE) loss, reduction of estimation risk via shrinkage turns out to be equivalent to a redefinition of the theoretical portfolio choice problem for an investor with a higher level of risk aversion. Taking risk preferences as given, the choice of the portfolio strategy and the estimation approach become endogenous, so that theoretically sub-optimal strategies can outperform other portfolio strategies once estimation risk is taken into account.

Despite its relevance for practice, so far only few attempts have been made to theoretically understand the mechanism determining the poor empirical performance of portfolio strategies in order to derive appropriate strategies reducing estimation risks. Notable exceptions are Jagannathan and Ma (2003), who analyze the potentially beneficial impact of imposing false restrictions in portfolio optimization. Kan and Zhou (2007) and Frahm and Memmel (2010) explicitly study the estimation risk of various portfolio estimation strategies. However, their analytical results are either restricted to the tangency portfolio without an adding-up restriction or focus on the global minimum variance portfolio where the investor is assumed to be extremely risk averse.

In the following, we use the loss in certainty equivalent compared to the theoretical efficient portfolio as a well-defined statistical loss function. In the literature, several evaluation criteria have been proposed to evaluate the performance of estimated portfolios. DeMiguel, Garlappi, and Uppal (2009) use the Sharpe ratio, CE and the turnover rate to compare the performance of the MV portfolio and the equally weighted portfolio. However, among these evaluation rules, only the CE loss is a proper scoring rule which identifies the true optimal portfolio in the sense that an estimated portfolio can never dominate its theoretical counterpart. Therefore, a comparison of theoretical or empirical portfolio strategies with the theoretical CE of the efficient portfolio provides a clear ranking. The CE has the theoretical appeal of being a statistical loss function which assess the additional loss an investor faces if she relies on estimated rather than on the true parameter values of the return process. However, contrary to conventional statistical loss functions, the CE based loss expresses estimation loss in terms of monetary units.

We analyze the loss in CE due to estimation analytically for the case of i.i.d. multivariatenormally distributed asset returns and provide quantitative evidence for the extent of estimation risk of different portfolio strategies. In the presence of estimation risk, the global minimum variance portfolio, although theoretically inferior, can be shown to be the superior portfolio strategy even for an investor with a low level of risk aversion when estimation risk is high, e.g. in the presence of a high dimensional portfolio choice problem, for small or moderate sample sizes and/or in the presence of strong correlation dependencies in the theoretical return process. Unlike previous studies which ignore the role of risk preferences for the magnitude of the financial loss caused by estimation uncertainty, we show that risk preferences, besides determining the usual trade-off between risk and return, are decisive in determining the extent to which estimation risk with respect to mean and variance contributes to the overall estimation risk. Our findings have a rather intuitive explanation: A risk neutral investor only cares about expected returns and not about risk. Therefore, her estimation risk with respect to the variance-covariance matrix of the return vector does not matter at all. On the contrary, a highly risk averse investor cares a lot about how precisely the variance-covariance matrix of the return vector can be estimated. Her monetary loss due to estimation risk depends strongly on the quality of the estimation of the variance-covariance matrix. Therefore, the question regarding the superiority of various portfolio choice strategies taking financial and estimation risk into account can only be answered for given risk preferences.

The outline of the paper is as follows. In Section 2, we introduce the CE loss as a statistical loss function and monetary measure for suboptimal portfolio selection. We relate the CE loss and the expected CE loss for the case of parameter estimation of the efficient portfolio to their counterparts for the global minimum variance portfolio and the tangency portfolio. In Section 3, we give specific analytical results for the estimation risk based on the assumption of an iid normal return vector. Section 4, we propose an optimal shrinkage method tailored to the efficient portfolio with budget constraint, while in Section 5 some calibration results are presented to document the quantitative relevance of our analytical findings. Section 6 concludes and gives an outlook on future research.

2 Loss of certainty equivalent

2.1 Theoretical MV Efficient Portfolios

Suppose there are N risky assets and the investor can only allocate wealth to these assets. Let r_t denote the $N \times 1$ vector of returns of risky assets with mean $\mathrm{E}[r_t] = \mu$ and covariance matrix $\mathrm{V}[r_t] = \Sigma$. According to the standard mean-variance framework, the efficient frontier can be equivalently presented by the solution of the following optimization problem for the certainty equivalent with respect to the vector of portfolio weights $w = (w_1, w_2, \ldots, w_N)'$:

$$\max_{w,\, \iota' w = 1} CE(w) = \max_{w,\, \iota' w = 1} \{\mu' w - \frac{\gamma}{2} w' \Sigma w\}, \tag{2.1}$$

where the parameter $\gamma \in (0, \infty]$ reflects the investor's level of risk aversion and ι is a $N \times 1$ vector of ones. The objective function of the optimization problem (2.1) is the certainty equivalent (CE) of the investor. The closed form solution of (2.1) is given by w^*_{eff}, the weight vector of the efficient portfolio:

$$w^*_{eff} = w_{eff}(\mu, \Sigma) = \frac{\Sigma^{-1}\iota}{\iota' \Sigma^{-1} \iota} + \frac{1}{\gamma} \cdot A \cdot \mu, \tag{2.2}$$

where

$$A = \Sigma^{-1} - \frac{\Sigma^{-1}\iota\iota'\Sigma^{-1}}{\iota'\Sigma^{-1}\iota} \tag{2.3}$$

is a semi-positive definite matrix. As we will compare in the following w^*_{eff} with the plug-in estimate of the efficient weight vector, $w_{eff}(\hat{\mu}, \hat{\Sigma})$, we use the superscript $*$ to indicate that $w_{eff}(\cdot, \cdot)$ is evaluated at the true parameters of the return process. Note, that the first term on the right hand side of (2.2), $w_{gmv} = \Sigma^{-1}\iota / (\iota' \Sigma^{-1} \iota)$, is the vector of weights of the global minimum variance portfolio (GMVP) as the solution of

$$\min_{w} w' \Sigma w \quad s.t \quad \iota' w = 1.$$

Because $A \cdot \iota = 0$, the second term of (2.2), $w_z = \frac{1}{\gamma} \cdot A \cdot \mu$, is the weight vector of a zero-investment portfolio with weights summing up to zero, i.e. $\iota' w_z = 0$. Obviously, for the limiting case of an extremely risk averse investor ($\gamma \to \infty$), the weights of the efficient portfolio approach the weights of the GMVP, which solely depend on the variance of the return vector and are, thus, only exposed to estimation risk of Σ. Therefore, as risk aversion increases, exposure to estimation risk with respect to mean returns decreases. Since the efficient portfolio weight is the sum of the GMVP weight and the zero-investment portfolio weight, $w^*_{eff} = w_{gmv} + w_z$, and since w_{gmv} represents the optimal choice if the investor is not willing to trade any risks against returns, w_z contains all relevant information concerning the extent to which the investor is willing to trade risk against return given her preferences and the nature of the return process.

In a similar fashion, it is also helpful to formulate the efficient portfolio weights (2.2) in terms of a linear combination of the weight vector of the GMVP and the weight vector of the tangency portfolio, $w^*_{eff} = w_{tan} + (1 - \iota' w_{tan}) \cdot w_{gmv}$, where $w_{tan} = \frac{1}{\gamma}\Sigma^{-1}\mu$, is the weight of the tangency portfolio, which is the solution of the optimization problem (2.1) when the adding-up restriction $\iota' w = 1$ is ignored. Therefore, if the mean and the covariance matrix are both known, the optimal investment strategy for an investor with risk aversion level γ is to allocate $\iota' w_{tan}$ of wealth to the optimal tangency portfolio and $1 - \iota' w_{tan}$ to the GMVP. As shown in the Appendix, the CE of the GMVP takes on the form

$$CE(w_{gmv}) = \mu_{gmv} - \frac{\gamma}{2}\sigma^2_{gmv} = \frac{\iota'\Sigma^{-1}\mu}{\iota'\Sigma^{-1}\iota} - \frac{\gamma}{2} \cdot \frac{1}{\iota'\Sigma^{-1}\iota}, \tag{2.4}$$

where $\mu_{gmv} = \mu' w_{gmv} = \frac{\iota'\Sigma^{-1}\mu}{\iota'\Sigma^{-1}\iota}$ is the mean return and $\sigma^2_{gmv} = w'_{gmv}\Sigma w_{gmv} = 1/\iota'\Sigma^{-1}\iota$ the variance of the GMVP. Substituting the weight of the efficient portfolio w^*_{eff} into the objective function, we obtain the highest theoretical CE for a portfolio strategy satisfying the budget constraint.

Proposition 2.1 (Decomposition of the Efficient CE)

The CE of the efficient portfolio based on the weight vector w^*_{eff} *given in (2.2) is:*

$$CE(w^*_{eff}) = \mu' w^*_{eff} - \frac{\gamma}{2} w^{*\prime}_{eff} \cdot \Sigma \cdot w^*_{eff} = \frac{1}{2\gamma}\Delta_{SSR} + CE(w_{gmv}), \tag{2.5}$$

where Δ_{SSR} *is the difference between the squared Sharpe ratios of the tangency portfolio and the GMVP,*

$$\begin{aligned}\Delta_{SSR} = \mu' \cdot A \cdot \mu =& \frac{(\mu'\Sigma^{-1}\mu)(\iota'\Sigma^{-1}\iota) - (\iota'\Sigma^{-1}\mu)^2}{\iota'\Sigma^{-1}\iota} \\ =& \left(\frac{\mu' w_{tan}}{\sqrt{w'_{tan}\Sigma w_{tan}}}\right)^2 - \left(\frac{\mu' w_{gmv}}{\sqrt{w'_{gmv}\Sigma w_{gmv}}}\right)^2 \\ =& (\mu - \mu_{gmv}\cdot\iota)'\Sigma^{-1}(\mu - \mu_{gmv}\cdot\iota) > 0,\end{aligned} \tag{2.6}$$

$CE(w_{gmv})$ *is the CE of the GMVP defined in (2.4).*

Proof 2.1 *See Appendix.*

The decomposition of the efficient CE given in Proposition 2.1 helps to understand the performance of the efficient theoretical portfolio relative to the theoretical GMVP. For a portfolio choice based on the GMVP, the investor is assumed to be extremely risk averse and, thus, only cares about the risk of the investment. Therefore, by construction for the case of known parameters the GMVP always yields a lower CE than the optimal efficient portfolio based on w^*_{eff}. The extent of the theoretical dominance of the efficient portfolio over the GMVP in terms of the CE depends on Δ_{SSR}, which captures the additional return that the investor receives compared to the maximum risk averse investor.

Similar to Proposition 2.1, the CE of the efficient portfolio can also be expressed in relation to the CE of the tangency portfolio.

$$\begin{aligned}CE(w^*_{eff}) =& \frac{1}{2\gamma}\mu'\Sigma^{-1}\mu - \frac{1}{2\gamma}\frac{1}{\iota'\Sigma^{-1}\iota}\left(\iota'\Sigma^{-1}\mu - \gamma\right)^2 \\ =& CE(w_{tan}) - \frac{1}{2\gamma}\frac{1}{\iota'\Sigma^{-1}\iota}\left(\iota'\Sigma^{-1}\mu - \gamma\right)^2.\end{aligned}$$

By definition, the theoretical CE of the efficient portfolio is never larger than the CE of the tangency portfolio. Finally, consider the weights of the Maximum Sharpe Ratio portfolio (MaxSR):

$$w_{SR} = \frac{w_{tan}}{\iota' w_{tan}} = \frac{\Sigma^{-1}\mu}{\iota'\Sigma^{-1}\mu}.$$

For a given γ the CE of this portfolio is:

$$CE(w_{SR}) = \frac{\mu'\Sigma^{-1}\mu}{(\iota'\Sigma^{-1}\mu)^2}(\iota'\Sigma^{-1}\mu - \frac{\gamma}{2}).$$

The CE of the MaxSR portfolio and the CE of the GMVP are both positive if and only if $\iota'\Sigma^{-1}\mu > \frac{\gamma}{2}$. In this case the CE of the MaxSR portfolio is always larger than that of the GMVP. Theoretically the GMVP yields a lower CE than the efficient portfolio, the tangency portfolio, and the MaxSR portfolio do. However, as shown below, if the estimation risk is taken into account, the estimated GMVP can be much more reliable than the three other portfolio strategies.

2.2 Expected CE Loss

By definition, for a given risk preference, the efficient portfolio dominates any other portfolio satisfying the adding up restriction in terms of the CE for given risk preferences, i.e. $CE(w^*_{eff}) \geq CE(w)$, where w is a weight vector satisfying the budget constraint and obtained by some other arbitrary portfolio selection rule based on true or estimated parameters. In particular, the efficient portfolio always dominates any estimated efficient portfolio $CE(w^*_{eff}) \geq CE(\hat{w}_{eff})$. Thus,

$$\mathcal{L}(\hat{w}, w^*_{eff}) \equiv CE(w^*_{eff}) - CE(\hat{w}) \geq 0 \tag{2.7}$$

is a well defined statistical loss function with $\hat{w} = w(\hat{\mu}, \hat{\Sigma})$. In practice, when the mean and the covariance matrix are unknown and the optimized portfolio is based on estimated inputs, it is, by definition, inferior to its theoretical counterpart. Since $CE(\hat{w})$ is a random variable, estimation loss is also random. The expectation over the loss function (2.7) defines the risk function

$$\mathcal{R}(\hat{w}|w^*_{eff}) \equiv \mathrm{E}\left[\mathcal{L}(\hat{w}, w^*_{eff})\right] = CE(w^*_{eff}) - \mathrm{E}\left[CE(\hat{w})\right] > 0, \tag{2.8}$$

where $\mathcal{R}(\hat{w}|w^*_{eff})$ gives the expected loss in CE if an estimate of the portfolio weight is taken instead of the true efficient portfolio weight. The risk function can be interpreted as the additional expected return an investor doing portfolio choice on estimated parameters requires to be indifferent towards to a portfolio evaluated on the true parameters. Therefore, if two empirical portfolio strategies based on the estimated portfolio weights $\hat{w}$ and $\tilde{w}$ are compared, the comparison should be based on their expected CE measures. Hence, the portfolio based on $\hat{w}$ strictly dominates the portfolio based on $\tilde{w}$ if

$$\mathrm{E}\left[CE(\hat{w})\right] - \mathrm{E}\left[CE(\tilde{w})\right] = \mathcal{R}(\tilde{w}|w^*_{eff}) - \mathcal{R}(\hat{w}|w^*_{eff}) > 0. \tag{2.9}$$

Thus proving the dominance of an estimated portfolio over any other estimated portfolio in terms of the expected CE is equivalent to the comparison of their corresponding risk functions.

Cho (2010) uses the CE to define the economic loss and shows that the loss of a suboptimal portfolio can be approximated by:

$$\mathcal{R}(\hat{w}|w) = CE(w) - \mathrm{E}\left[CE(\hat{w})\right] \cong \frac{\gamma}{2} tr\left(\mathrm{Cov}[\hat{w}] \cdot \Sigma\right).$$

He argues that this approximation can be applied to all MV portfolio problems of any given constraint, although it only holds when the estimated portfolio weights are assumed to be unbiased. The plug-in estimators of the portfolio weights are, however, generally nonlinear functions of the estimated mean and the estimated covariance matrix. Even if these estimates are unbiased, the weights as nonlinear functions are generally biased. Kan and Zhou (2007) provide a formal proof for the tangency portfolio under iid normality of the return vector and derive the size of the finite sample bias depending on the sample size T and the dimension of the portfolio choice problem N. In a similar fashion, Okhrin and Schmid (2006) show that the plug-in estimated efficient portfolio weights are also biased but have a smaller mean squared error than their unbiased counterparts. In any case, the unbiasedness assumption turns out to be very restrictive and we will show below that the bias in the estimated weights can be large and can even dominate the variance-covariance matrix of the vector of weights. In order to identify the exact CE loss of a suboptimal portfolio relative to the true efficient portfolio, we provide the following proposition:

Proposition 2.2 (CE Loss and Expected CE Loss) *Let w^*_{eff} denote the solution of the MV-maximization problem (2.1) and let $\hat{w}$ be any portfolio weight vector satisfying $\iota'\hat{w} = 1$, then:*

$$\mathcal{L}(\hat{w}, w^*_{eff}) = CE(w^*_{eff}) - CE(\hat{w}) = \frac{\gamma}{2}(w^*_{eff} - \hat{w})'\Sigma(w^*_{eff} - \hat{w}).$$

with expected loss of CE given by:

$$\mathcal{R}(\hat{w}|w^*_{eff}) = CE(w^*_{eff}) - \mathrm{E}\left[CE(\hat{w})\right] = \frac{\gamma}{2}tr\left(\Sigma \cdot \left[\mathrm{Cov}[\hat{w}] + Bias(\hat{w})^2\right]\right),$$

where $Bias(\hat{w})^{2=} \left(\mathrm{E}\left[\hat{w}\right] - w^*_{eff}\right)\left(\mathrm{E}\left[\hat{w}\right] - w^*_{eff}\right)'$.

Proof 2.2 *See Appendix.*

For the GMVP, the CE loss due to estimation error is given by:[1]

$$\mathcal{L}(\hat{w}, w_{gmv}) = CE(w_{gmv}) - CE(\hat{w}) = (w_{gmv} - \hat{w})'\Sigma(w_{gmv} - \hat{w}),$$

while a similar result can be obtained for the tangency portfolio[2]:

$$\mathcal{L}(\hat{w}, w_{tan}) = CE(w_{tan}) - CE(\hat{w}) = \frac{\gamma}{2}(w_{tan} - \hat{w})'\Sigma(w_{tan} - \hat{w}).$$

[1]See Kempf and Memmel (2006) for a proof.
[2]see Frahm (2010)

Although Proposition 2.2 reveals some similarities to the corresponding loss functions for the GMVP and the tangency portfolio, considerable differences occur if the mean and the covariance matrix are replaced by their estimates. Proposition 2.3 relates the CE loss directly to the estimated mean.

Proposition 2.3 (Expected CE for known Variance) *If the true covariance matrix is known, then the CE loss of the efficient portfolio is:*

$$\mathcal{L}(w_{eff}(\hat{\mu},\Sigma), w^*_{eff}) = \frac{1}{2\gamma}(\mu-\hat{\mu})' \cdot A \cdot (\mu-\hat{\mu})$$

with the expected CE loss given by:

$$\mathcal{R}(w_{eff}(\hat{\mu},\Sigma)|w^*_{eff}) = \frac{1}{2\gamma} tr\left(A \cdot \left[\mathrm{Cov}[\hat{\mu}] + Bias(\hat{\mu})^2\right]\right),$$

where $\hat{\mu}$ is an arbitrary estimator for the mean returns.

Proof 2.3 *See Appendix.*

Since A is a positive semi-definite matrix, CE loss and expected CE loss are non-negative. Frahm (2010) derives the economic loss resulting from estimation for a given known covariance matrix for the case of the tangency portfolio:

$$\mathcal{L}(w_{tan}(\hat{\mu},\Sigma), w_{tan}) = \frac{1}{2\gamma}(\mu-\hat{\mu}) \cdot \Sigma^{-1} \cdot (\mu-\hat{\mu}).$$

Note that the estimation loss for the efficient portfolio is smaller than the estimation loss for the tangency portfolio because the former accounts for the budget constraint $\iota' w = 1$. In addition, the expected loss of the estimated efficient portfolio only depends on the differences between $(\mu_i - \hat{\mu}_i)$ and $(\mu_j - \hat{\mu}_j)$ for $i \neq j$. Therefore, systematical overestimation or underestimation of all means by a constant implies no economic loss in the case of efficient portfolio but can significantly reduce the value of the estimated tangency portfolio. Thus, the estimated efficient portfolio can outperform the estimated tangency portfolio in practice.

2.2.1 Within and Out-of-sample Measures

Aside from the theoretical CE defined in (2.1), the out-of-sample CE concept is often considered:

$$CE_{os}(w) = \mathrm{E}\left[w'r\right] - \frac{\gamma}{2}\mathrm{V}\left[w'r\right].$$

This concept is often used in practice in comparative empirical studies when μ and Σ are unknown. Obviously, if all input elements are known, the out-of-sample CE, $CE_{os}(w)$, is identical to the CE definition defined in (2.1). However, if estimated portfolio weights are used instead of the true ones, both weights and returns are random and, thus, the two CE concepts differ.

Given that returns r_t are stochastically independent of trading strategy $\hat{w}$ selected by the investor, the out-of-sample CE can be computed as follows:

$$\begin{aligned} CE_{os} &= \mathrm{E}\left[\hat{w}'r_t\right] - \frac{\gamma}{2}\mathrm{V}\left[\hat{w}'r_t\right] \\ &= \mathrm{E}\left[\hat{w}'\,\mathrm{E}\left[r_t|\,\hat{w}\right]\right] - \frac{\gamma}{2}\left(\mathrm{E}\left[\hat{w}'\,\mathrm{V}\left[r_t|\,\hat{w}\right]\hat{w}\right] + \mathrm{V}\left[\mathrm{E}\left[r_t|\,\hat{w}\right]'\cdot\hat{w}\right]\right) \\ &= \mathrm{E}\left[CE(\hat{w})\right] - \frac{\gamma}{2}\mu'\,\mathrm{Cov}[\hat{w}]\mu \end{aligned} \tag{2.10}$$

Therefore, if portfolio weights are estimated, CE_{os} is smaller than the theoretical CE which is based on the true mean and covariance matrix. In the following analysis, we use $\mathcal{R}_{os}(.|.)$ to denote the out-of-sample risk function (the out-of-sample expected loss) of estimated portfolio compared to the true optimal efficient portfolio. Based on Proposition 2.2, the out-of-sample risk function can be easily obtained by:

$$\mathcal{R}_{os}(\hat{w}|w^*_{eff}) = \frac{\gamma}{2}tr\left((\Sigma + \mu\mu')\cdot\mathrm{Cov}[\hat{w}] + \Sigma\cdot Bias(\hat{w})^2\right). \tag{2.11}$$

2.3 Implied Mean of a Portfolio

The exact risk function given in Subsection 2.2 was derived under the assumption that the true covariance matrix is known and that the estimation risk is solely due to estimation of mean returns. Using Proposition 2.4 given below, we can represent any theoretical or empirical portfolio weight in terms of an efficient portfolio weight with a known covariance matrix and an implied mean vector. Therefore, a comparison of any portfolio weight in terms of the CE loss can be reduced to a comparison of the equivalent representation of this portfolio weight with the efficient portfolio weight evaluated at the true mean and covariance. The CE differences are simply reflected by the differences between the true and the implied mean vector.

Proposition 2.4 (Equivalent Representation) *Let S be the subspace of $\mathbb{R}^N$ which is orthogonal to ι and let the $N \times (N-1)$ matrix V be the basis matrix of S, i.e. the column vectors of V construct a basis of S. Let $\hat{w}$ denote the weight vector of a given portfolio. A is the matrix defined in (2.3). Then, there is an implied mean vector $\hat{\mu}_{im} = c\iota + \hat{\mu}^0_{im}$ such that:*

$$\hat{w} \;=\; w_{eff}(\hat{\mu}_{im}, \Sigma) = \frac{1}{\gamma}\Sigma^{-1}\hat{\mu}_{im} + (1 - \frac{1}{\gamma}\iota'\Sigma^{-1}\hat{\mu}_{im})\frac{\Sigma^{-1}\iota}{\iota'\Sigma^{-1}\iota},$$

where c is any arbitrary constant and

$$\hat{\mu}^0_{im} = \gamma \cdot V(V'BV)^{-1}V' \cdot \Sigma \cdot \hat{w} \qquad \text{and} \qquad \iota'\hat{\mu}^0_{im} = 0.$$

with $B = \Sigma \cdot A$.

Proof 2.4 *See Appendix.*

The second term of the implied mean, $\hat{\mu}^0_{im}$, sums up to zero. For example, consider the case of the true efficient portfolio. Here, $\hat{\mu}^0_{im}$ measures the deviation from the average of the means for the single returns, i.e. $\mu - \bar{\mu}\iota$ with $\bar{\mu} = \iota'\mu/N$. Moreover, note there exists an infinite number of implied means $\hat{\mu}_{im}$, which generate the same portfolio and, hence, the same CE. Therefore, analyzing the $\hat{\mu}_{im}$ is fully equivalent to analyzing any other implied mean vector.

With the help of the equivalent representation given in Proposition 2.4 all errors in the suboptimal portfolio are contained in $\hat{\mu}_{im}$, and the difference between the true optimal portfolio and the estimated portfolio can be identified if the difference between the true population mean and the estimated (implied) mean is known. Thus the "best" implied mean can be defined as $\hat{\mu}^*_{im} = \hat{\mu}^0_{im} + \bar{\mu}\iota$, which has the smallest distance (measured by the Euclidean metric) to the overall true mean.

For given estimates of the mean and the covariance matrix, the implied mean $\hat{\mu}^0_{im}$ can be explicitly written as:

$$\hat{\mu}^0_{im} = V(V'BV)^{-1}V' \cdot \Sigma \cdot \hat{\Sigma}^{-1}(\gamma\iota + \hat{B}\hat{\mu}).$$

Even if the mean μ is estimated without error, the implied mean will differ from the true population mean if the covariance matrix is estimated with errors. Moreover, the difference will be large if the difference between the estimated inverse covariance matrix $\hat{\Sigma}^{-1}$ and the true Σ^{-1} is large. It is clear that, if the risk aversion parameter γ is large, the investor cares more about the variance of portfolio, and, therefore, the estimation risk in the covariance has a larger impact on the implied mean. In addition, the impact of the errors in the mean vector and the errors in the covariance matrix on the implied mean is not additive and the interaction might be large. This issue will be discussed below in more detail for cases where sample counterparts of the mean and the covariance matrix are used. With the help of Proposition 2.4, we can easily reformulate the CE loss and the expected CE loss of the efficient portfolio by replacing the estimated mean by the implied mean.

Example 2.1 *The theoretical GMVP has an implied mean equal to ι. Thus, the expected CE loss of the GMVP is given by*

$$\mathcal{R}(w_{gmv}|w^*_{eff}) = \mathcal{R}(w_{eff}(\iota, \Sigma)|w^*_{eff}) = \frac{1}{2\gamma}\mu \cdot A \cdot \mu = \frac{1}{2\gamma}\Delta_{SSR},$$

which is consistent with the result in Section 2.1.

Based on the equivalent representation by the implied mean Proposition 2.5 gives an upper bound for the CE loss.

Proposition 2.5 (Upper Bound of the CE loss) *The loss of CE is bounded by:*

$$\begin{aligned}\mathcal{L}(\hat{w}, w^*_{eff}) \leq & \frac{1}{2\gamma}(\sum_{i=1}^{N}\lambda_i^{-1}) \cdot (\sum_{i=1}^{N}\sigma_i^{-2}) \parallel (I - \frac{\Sigma^{-\frac{1}{2}}\iota\iota'\Sigma^{-\frac{1}{2}}}{\iota'\Sigma^{-1}\iota}) \parallel_2^2 \cdot \parallel \mu - \hat{\mu}_{im} \parallel_2^2 \\ = & \frac{N-1}{2\gamma}(\sum_{i=1}^{N}\lambda_i^{-1}) \cdot (\sum_{i=1}^{N}\sigma_i^{-2}) \left(\sum_{i=1}^{N}(\mu_i - \hat{\mu}_{im,i})^2\right),\end{aligned}$$

where λ_i, $i = 1, \ldots, N$ is the eigenvalue of the correlation matrix.

Proof 2.5 *See Appendix.*

The upper bound of CE loss depends on the risk aversion level γ, the number of assets, the level of variances, the collinearity between asset returns (i.e. the eigenvalues of Σ) and the estimation error (measured by $\parallel \mu - \hat{\mu}_{im} \parallel_2^2$). The upper bound gives the highest possible loss when a suboptimal portfolio strategy is used. It provides information regarding the outcome range of possible losses compared to the efficient portfolio strategy based on the true population parameters of the return process. As will be shown below, the risk function for cases of an unknown mean estimated by a sample mean and a given variance-covariance matrix turns out to be independent of the covariance matrix. This is because the squared estimation error of the first moment, $\mathrm{E}\left[(\hat{\mu} - \mu)(\hat{\mu} - \mu)'\right]$, is proportional to the true second moment Σ. Thus, the squared error in sample mean is to some extent compensated by Σ^{-1} in the loss function.

3 Expected CE Loss under Normality

Since the exact functional form of the expected CE loss depends on the underlying distributional properties of the return process, we assume i.i.d. multivariate normality for the return process:

a) $r_t \overset{\text{iid}}{\sim} \mathrm{N}(\mu, \Sigma)$ for $t = 1, \ldots, T$.

b) $T \geq N+4$ and $N \geq 3$.

Population mean and population covariance matrix are estimated by their sample counterparts:

$$\hat{\mu} = \bar{r} = \frac{1}{T}\sum_{t=1}^{T} r_t \qquad \text{and} \qquad \hat{\Sigma} = S = \frac{1}{T-1}\sum_{t=1}^{T}(r_t - \bar{r})(r_t - \bar{r})'.$$

Under the normality assumption from above, the two estimators are distributed as:

$$\bar{r} \sim \mathrm{N}(\mu, \frac{1}{\mathrm{T}}\Sigma) \qquad \text{and} \qquad \mathrm{S} \sim \mathrm{W_N}(\mathrm{T}-1, \Sigma)/\mathrm{T}-1,$$

where $W_N(T-1, \Sigma)$ denotes the Wishart distribution with $T-1$ degrees of freedom and covariance matrix Σ.

3.1 Expected CE Loss of the Efficient Portfolio

Case I: Sample Mean - True Covariance Matrix

Consider first the case where the covariance matrix is known but the mean vector is estimated with errors. The expected CE loss using the sample mean of the return vector and the true population covariance matrix is:

$$\mathcal{R}(w_{eff}(\bar{r}, \Sigma)|w^*_{eff}) = \frac{1}{2\gamma} tr\left(A \cdot \mathrm{V}\left[\bar{r}\right]\right) = \frac{1}{2\gamma}\frac{N-1}{T}. \tag{3.1}$$

Obviously, for this case estimation risk is negligible for the extremely risk averse investor as, for her, only the estimation risk concerning the covariance matrix matters. Moreover, a large sample size and a small number of assets in the portfolio reduces estimation risk as well. It should not be too surprising that the estimation risk of the tangency portfolio $\mathcal{R}(w_{tan}(\bar{r}, \Sigma)|w_{tan}(\mu, \Sigma)) = \frac{1}{2\gamma} \cdot \frac{N}{T}$ is larger than the one for the efficient portfolio given in (3.1), since the budget constraint is taken into account for the latter, which reduces estimation uncertainty.[3]

[3]See Kan and Zhou (2007) for a proof of the risk function of the tangency portfolio given Σ is known.

Consider now the out-of-sample CE loss caused by using the sample mean. Using (2.10) the difference between the out-of-sample CE and the within-sample CE is given by

$$\mathrm{E}\left[CE\left(w_{eff}(\bar{r},\Sigma)\right)-CE_{os}\left(w_{eff}(\bar{r},\Sigma)\right)\right] \;=\; \frac{1}{T}\cdot\frac{1}{2\gamma}\Delta_{SSR}.$$

Thus, the out-of-sample loss of CE due to estimation error in sample means is given by

$$\mathcal{R}_{os}(w_{eff}(\bar{r},\Sigma)|w^{*}_{eff}) = \frac{1}{2\gamma}\cdot\left(\frac{N-1+\Delta_{SSR}}{T}\right). \tag{3.2}$$

Case II: True Mean - Sample Covariance Matrix

Conditional on any given estimate of mean returns, the covariance of estimated portfolio weights using the sample covariance matrix S is (see Okhrin and Schmid (2006)):

$$\mathrm{Cov}[w_{eff}(\hat{\mu},S)|\hat{\mu}] = \frac{1}{T-N-1}\frac{A}{\iota'\Sigma^{-1}\iota}+\frac{1}{\gamma^2}\left(c_1A\cdot\hat{\mu}\hat{\mu}'\cdot A+c_2\hat{\mu}'\cdot A\cdot\hat{\mu}\cdot A\right),$$

where

$$c_1=\frac{(T-1)^2\,(T-N+1)}{(T-N)\,(T-N-1)^2\,(T-N-3)}\quad\text{and}\quad c_2=\frac{(T-1)^2}{(T-N)\,(T-N-1)\,(T-N-3)}.$$

In addition, the conditional expectation of estimated weights using sample covariance matrix is:

$$\mathrm{E}\left[w_{eff}(\hat{\mu},S)|\,\hat{\mu}\right] = \frac{\Sigma^{-1}\iota}{\iota'\Sigma^{-1}\iota}+\frac{T-1}{T-N-1}\cdot\frac{1}{\gamma}A\cdot\hat{\mu}.$$

Therefore, if the true mean μ is known, the expected CE loss due to estimation error in sample covariance matrix can be calculated as:

$$\begin{aligned}
&\mathcal{R}(w_{eff}(\mu,S)|w^{*}_{eff})\\
=\;&\frac{\gamma}{2}tr\left(\Sigma\cdot\left[\mathrm{Cov}[w_{eff}(\mu,S)]+Bias\left(w_{eff}(\mu,S)\right)^2\right]\right)\\
=\;&\frac{\gamma}{2}\frac{N-1}{T-N-1}\sigma^2_{gmv}+\frac{\Delta_{SSR}}{2\gamma}\left(c_1+c_2\,(N-1)+\left(\frac{N}{T-N-1}\right)^2\right).
\end{aligned} \tag{3.3}$$

The first term in (3.3), $\frac{\gamma}{2}\frac{N-1}{T-N-1}\sigma^2_{gmv}$, can be interpreted as the baseline risk component, as it occurs in both, the risk function for any estimated efficient portfolio and the risk function for the estimated GMVP derived below. Contrary to Case I, where μ has to be estimated, risk aversion has an ambivalent effect on the expected CE loss when Σ has to be estimated. A higher degree of risk aversion increases the impact of baseline risk of the return process represented by σ^2_{gmv}. However, higher risk aversion reduces the effect of the overall earnings potential of the return process. It is easy to show that the expected CE loss, as a function of the degree of risk aversion, has a unique minimum, i.e. investors with different degrees of risk aversion may face the same expected CE loss. The less risk averse investor faces less estimation risk compared to a more risky investor. However, she loses money in terms of CE by pursuing a less profitable strategy in terms of the theoretical CE.

Similar to (3.3), we can also compute the out-of-sample CE loss when the sample covariance matrix is used. Using (2.11), if the true mean is known, the out-of-sample loss of CE due to estimation error of sample covariances is:

$$\begin{aligned}
&\mathcal{R}_{os}(w_{eff}(\mu,S)|w^*_{eff}) \qquad (3.4)\\
=&\frac{\gamma}{2}tr\left((\Sigma+\mu\mu')\cdot \mathrm{Cov}[w_{eff}(\mu,S)] + \Sigma\cdot Bias(w_{eff}(\mu,S))^2\right)\\
=&\frac{\gamma}{2}\frac{N-1+\Delta_{SSR}}{T-N-1}\sigma^2_{gmv} + \frac{\Delta_{SSR}}{2\gamma}\left(c_1\left(1+\Delta_{SSR}\right)+c_2\left(N-1+\Delta_{SSR}\right)+\left(\frac{N}{T-N-1}\right)^2\right).
\end{aligned}$$

Case III: Sample Mean and Sample Covariance Matrix

The covariance of estimated weights using sample mean and sample covariances is given by:[4]

$$\begin{aligned}
\mathrm{Cov}[w_{eff}(\bar{r},S)] &= \frac{1}{T-N-1}\frac{A}{\iota'\Sigma^{-1}\iota}+\frac{1}{\gamma^2}\left(c_1 A\mu\mu' A + c_2\mu' A\mu A\right)\\
&+\frac{1}{T}\cdot\frac{A}{\gamma^2}\left(c_1+c_2\left(N-1\right)+\frac{\left(T-1\right)^2}{\left(T-N-1\right)^2}\right).
\end{aligned}$$

[4]see Okhrin and Schmid (2006) for a proof.

Using this result, we are able to derive the conditional loss of CE due to the estimation error when the sample mean and the sample covariance matrix are used:

$$
\begin{aligned}
&\mathcal{R}\left(w_{eff}(\bar{r},S)|w^*_{eff}\right)\\
=&\frac{\gamma}{2}tr\left(\Sigma\cdot\left[\mathrm{Cov}[w_{eff}(\bar{r},S)]+Bias(w_{eff}(\bar{r},S))^2\right]\right)\\
=&\frac{\gamma}{2}\frac{N-1}{T-N-1}\frac{1}{\iota'\Sigma^{-1}\iota}+\frac{\mu'\cdot A\cdot\mu}{2\gamma}\left(c_1+c_2\left(N-1\right)\right)\\
+&\frac{N-1}{T}\cdot\frac{1}{2\gamma}\left(c_1+c_2\left(N-1\right)+\frac{\left(T-1\right)^2}{\left(T-N-1\right)^2}\right)+\frac{1}{2\gamma}\cdot\left(\frac{N}{T-N-1}\right)^2\cdot\mu'\cdot A\cdot\mu\\
=&\mathcal{R}(w_{eff}(\mu,S)|w^*_{eff})+c_3\mathcal{R}(w_{eff}(\bar{r},\Sigma)|w^*_{eff}).
\end{aligned}
\tag{3.5}
$$

with

$$c_3=\frac{(T-1)^2(T-2)}{(T-N-1)(T-N)(T-N-3)}>1.$$

Note that the overall expected loss due to estimation is larger than the sum of the risks of estimating μ and Σ, given that the other parameters are known. The coefficient c_3 can be interpreted as an interaction effect which gives more weight to the estimation risk with respect to Σ if mean returns also have to be estimated. It gives an analytical explanation for the poor performance of the empirical efficient portfolio strategies. Note, that c_3 does not reflect the well-known estimation problem of mean returns but represents the additional estimation risk due to the estimation of Σ if μ is unknown. c_3 increases with the dimension of the portfolio choice problem and decreases with sample size. Table 1 below gives some values of c_3.

Table 1: Scale Effect due to Estimation of μ

$T \backslash N$	5	10	15	20	25	30
60	1.31	1.75	2.43	3.50	5.30	8.60
120	1.14	1.30	1.50	1.74	2.03	2.40
180	1.09	1.19	1.30	1.43	1.57	1.74
240	1.07	1.14	1.22	1.30	1.39	1.50
300	1.05	1.11	1.17	1.23	1.30	1.37

Value of c_3 for different T (number of observations) and N (number of assets).

Sample size matters particularly for large portfolios. The scale factor decreases by more than 84% if the sample size increases, e.g. from 5 years of monthly data (T = 60) to 25 years (T = 300), for a portfolio of 30 assets while the reduction due to an increase in sample size is only 20 % for portfolios of 5 assets.

For the out-of-sample case, we obtain very similar results. The difference between within-sample and out-of-sample CE is given by:

$$\begin{aligned} & \mathrm{E}\left[CE\left(w_{eff}(\bar{r},S)\right) - CE_{os}\left(w_{eff}(\bar{r},S)\right)\right] \\ = & \frac{\gamma}{2}\mu' \operatorname{Cov}[w_{eff}(\bar{r},S)]\mu \\ = & \frac{\gamma}{2}\frac{1}{T-N-1}\sigma^2_{gmv}\Delta_{SSR} + \frac{1}{2\gamma}\left(c_1+c_2\right)\cdot\Delta^2_{SSR} + c_3\cdot\frac{1}{T}\cdot\frac{\Delta_{SSR}}{2\gamma} \end{aligned}$$

Thus, the out-of-sample loss of CE due to the estimation error in the sample means and the sample covariances is:

$$\mathcal{R}_{os}\left(w_{eff}(\bar{r},S)|w^*_{eff}\right) = \mathcal{R}_{os}\left(w_{eff}(\mu,S)|w^*_{eff}\right) + c_3\cdot\mathcal{R}_{os}\left(w_{eff}(\bar{r},\Sigma)|w^*_{eff}\right). \tag{3.6}$$

The composition of the out-of-sample CE loss is the same as for the unconditional CE loss given by (3.5). However, because the term on the right hand side in (3.6) is larger than its counterpart for the expected CE loss of the unconditional case, the out-of-sample expected CE loss is clearly larger than the unconditional expected CE loss.

3.2 Expected CE loss of the GMVP

Frahm (2010) suggests using the mean of the GMVP as an estimator for the mean vector of returns:

$$\hat{\mu}_{gmv} = \bar{r}'\hat{w}_{gmv}\cdot\iota = \frac{\bar{r}'\hat{\Sigma}^{-1}\iota}{\iota'\hat{\Sigma}^{-1}\iota}\cdot\iota \tag{3.7}$$

Because $A\cdot\iota = 0$, the efficient frontier reduces to the GMVP when all means are equal. The CE loss in this case is equal to the CE loss using the estimated GMVP. Therefore in terms of the loss function, the estimator for the means given in (3.7) is equivalent to any estimator which is proportional to ι. The estimated portfolio weights of the GMVP using the sample covariance matrix are essentially unbiased (Okhrin and Schmid (2006)). Therefore, for a given γ, the expected CE difference between the theoretical GMVP and the empirical GMVP is:

$$\mathrm{E}\left[CE(w_{gmv}) - CE\left(w_{gmv}(S)\right)\right] = \frac{\gamma}{2}\left[\mathrm{V}\left[w'_{gmv}r_t\right] - \mathrm{V}\left[w_{gmv}(S)'r_t\right]\right].$$

Thus the within-sample CE loss and the out-of-sample CE loss of the empirical GMVP for a given γ are

$$\mathcal{R}(w_{gmv}(S)|w^*_{eff}) = \mathcal{R}(w_{eff}(c\cdot\iota, S)|w^*_{eff}) = \frac{\gamma}{2}\cdot\frac{N-1}{T-N-1}\cdot\sigma^2_{gmv} + \frac{1}{2\gamma}\Delta_{SSR} \qquad (3.8)$$

and

$$\mathcal{R}_{os}\left(w_{gmv}(S)|w^*_{eff}\right) = \mathcal{R}_{os}(w_{eff}(c\cdot\iota, S)|w^*_{eff}) = \frac{\gamma}{2}\cdot\frac{N-1+\Delta_{SSR}}{T-N-1}\cdot\sigma^2_{gmv} + \frac{1}{2\gamma}\Delta_{SSR}, \quad (3.9)$$

respectively. Also, for this special case, the expected CE losses are nonlinear functions of the risk preference parameter with a unique minimum. Therefore, investors with different risk attitudes may face the same expected out-of-sample CE loss.

4 Shrinkage Estimation of the Efficient Portfolio

Based on a Bayesian reasoning, Jorion (1986) proposes a shrinkage estimator for the mean of the form:

$$\hat{\mu}_{shrink} = \eta\cdot\hat{\mu} + (1-\eta)\cdot\hat{\mu}_{gmv}\cdot\iota,$$

where η denotes the shrinkage parameter and $\hat{\mu}_{gmv}$ is the shrinkage target which is equal to the estimated expected return of GMVP as defined in (2.4). The optimal shrinkage parameter of Jorion's Bayes-Stein is given by:

$$\hat{\eta}_{BS} = 1 - \frac{N+2}{(N+2) + T(\hat{\mu} - \hat{\mu}_{gmv} \cdot \iota)'\hat{\Sigma}^{-1}(\hat{\mu} - \hat{\mu}_{gmv} \cdot \iota)} = \frac{\hat{\Delta}_{SSR}}{\hat{\Delta}_{SSR} + \frac{N+2}{T}}.$$

As shown above, the computation of the efficient portfolio using the shrinkage target $\hat{\mu}_{gmv}$ is equivalent to shrinking the mean to any target of form $c \cdot \iota$, where c is an arbitrary constant. For $c = 0$, this shrinkage approach is equivalent to biasing the estimated mean towards zero. Thus estimation risk in the GMVP weight has no impact on the final result. Furthermore, shrinking the mean is also equivalent to directly applying shrinkage estimation to the efficient portfolio weight with the GMVP as the shrinkage target, i.e.

$$w_{shrink}(\eta, \hat{\mu}, \hat{\Sigma}) = \eta \cdot w_{eff}(\hat{\mu}, \hat{\Sigma}) + (1 - \eta) \cdot w_{gmv}(\hat{\Sigma}) = w_{gmv}(\hat{\Sigma}) + \eta \hat{w}_z. \tag{4.1}$$

Equation (4.1) also reveals that shrinking mean returns to the mean of the GMVP is nothing but reducing the investors arbitrage opportunities by lowering the contribution of the (estimated) zero-investment portfolio.

Kan and Zhou (2007) argue that the shrinkage portfolio suggested by Jorion (1986) "can be suboptimal, because it is not constructed for holding optimal position", and propose the optimal shrinkage estimator of the mean for a tangency portfolio. In the following, we derive the optimal shrinkage estimator tailored for the efficient portfolio given in (2.2). Using $\bar{r}$ and S as estimators for the shrinkage weight (4.1) in the risk function given in Proposition 2.2 leads to:

$$\begin{aligned}
& \mathcal{R}(w_{shrink}(\eta, \bar{r}, S)|w^*_{eff}) = \mathcal{R}(w_{eff}(\eta \cdot \bar{r}, S)|w^*_{eff}) \qquad (4.2)\\
= \; & \frac{\gamma}{2} trace\left(\Sigma \cdot \text{Cov}[w_{shrink}(\eta, \bar{r}, S)] + \Sigma \cdot Bias\left(w_{shrink}(\eta, \bar{r}, S)\right)^2\right)\\
= \; & \frac{\gamma}{2}\frac{N-1}{T-N-1}\sigma^2_{gmv} + \frac{\eta^2}{2\gamma}\left[(c_1 + c_2(N-1))\Delta_{SSR} + c_3\frac{N-1}{T}\right] + \frac{\Delta_{SSR}}{2\gamma}\left(1 - \frac{(T-1)\eta}{T-N-1}\right)^2.
\end{aligned}$$

The optimal shrinkage factor η^* can be obtained by minimizing (4.2) and solving the first order condition:

$$\eta^* = \frac{\Delta_{SSR}}{c_3\left(\Delta_{SSR} + \frac{N-1}{T}\right)} \cdot \frac{T-1}{T-N-1} = \frac{(T-N)(T-N-3)}{(T-1)(T-2)} \cdot \frac{\Delta_{SSR}}{\left(\Delta_{SSR} + \frac{N-1}{T}\right)} < 1.$$

With this optimal shrinkage parameter η^*, the expected CE loss is:

$$\mathcal{R}(w_{shrink}(\eta^*, \bar{r}, S)|w^*_{eff}) = \frac{\gamma}{2}\left(\frac{N-1}{T-N-1}\right)\sigma^2_{gmv} + \frac{1}{2\gamma}\Delta_{SSR}\left(1 - \frac{T-1}{T-N-1}\eta^*\right). \quad (4.3)$$

Comparing the expected CE loss of the optimal shrinkage portfolio to the expected loss of the empirical GMVP (see Equation (3.8)), we see that using the shrinkage approach leaves the baseline risk component (first term of (3.8)) unchanged but reduces the theoretical loss of the GMVP. Therefore, the optimal shrinkage portfolio outperforms the GMVP only marginally when the sample size is small or the theoretical loss of the GMVP is small.

The comparison of the risks of the optimal shrinkage portfolio with the risk of the sample efficient portfolio given by (3.5) shows that the expected loss of the optimal shrinkage portfolio in (4.3) is similar to the partial loss of the sample efficient portfolio, which is caused by the estimation error of the covariance matrix (see Equation (3.3)). However, the optimal shrinkage portfolio does not contain the partial loss, which is due to the estimation error in the sample mean and the interaction term, $c_3 \cdot \frac{N-1}{T}$. Thus, in finite samples, the optimal shrinkage portfolio can yield a much higher performance than the sample efficient portfolio. Interestingly, although the CE loss depends on the risk aversion parameter, the true variance of the GMVP and the difference in the squared Sharpe ratios, Δ_{SSR}, the optimal shrinkage parameter is only a function of Δ_{SSR} for a given sample size and number of assets. Therefore optimal shrinkage estimation is valid for any type of investor.

Obviously, the optimal shrinkage parameter η^* is infeasible because it depends on the unknown Δ_{SSR}. This term has to be estimated, which introduces an additional source of estimation error. Contrary to the case of optimal shrinkage tangency portfolio proposed by Kan and Zhou (2007), where both Δ_{SSR} and the mean of the GMVP have to be estimated, η^* depends only on one unknown parameter. This is because, as argued above, shrinking the portfolio weights to the weights of the GMVP is equivalent to shrinking the mean to zero. Therefore, estimates based on the optimal shrinkage parameter η^* are likely to be more reliable than the estimates of shrinkage tangency portfolio along the lines of Kan and Zhou (2007).

Based on the normality assumption and on $\bar{r}$ and S as plug-in estimates, the distribution for the $\hat{\Delta}_{SSR}$ is given by:[5]

$$\frac{T-N+1}{N-1}\hat{\Delta}_{SSR} \sim F_{(N-1,T-N+1)}(T\Delta_{SSR}),$$

where $F_{(N-1,T-N+1)}(T\Delta_{SSR})$ is non-central F distribution with $N-1$ and $T-N+1$ degrees of freedom and noncentrality parameter $T\Delta_{SSR}$. Therefore, the unbiased estimator of Δ_{SSR} is given by:

$$\hat{\Delta}^u_{SSR} = \frac{T-N-1}{T}\hat{\Delta}_{SSR} - \frac{N-1}{T}.$$

When N is large, the unbiased estimator may be negative which would be equivalent to assuming that all assets have negative mean returns. This will cause large short positions in the constructed portfolio. To ensure the positiveness of $\hat{\Delta}_{SSR}$, we can also bound the unbiased estimator by zero:

$$\hat{\Delta}^{mod}_{SSR} = max\left\{\frac{T-N-1}{T}\hat{\Delta}_{SSR} - \frac{N-1}{T}, 0\right\},$$

where $\hat{\Delta}_{SSR}$ is the plug-in estimator of Δ_{SSR} based on the sample mean and the sample covariance matrix.

[5]See Kan and Zhou (2007) for a proof.

5 Calibration to Real Data

The results presented so far are purely analytical. In the following, we therefore provide quantitative evidence on the expected CE loss for realistic magnitudes on the population parameters of the return process for different levels of risk aversion, sample size and portfolio size. More specifically, we consider three different samples containing 1) 5 industry portfolios; 2) 10 industry portfolios; 3) 30 industry portfolios based on monthly returns for the sample period 07/1926 - 09/2009 published on Kenneth French's Web site.[6] For our simulations, we choose the sample estimates from these three data sets for the true mean and true covariance matrix .

5.1 Properties of the Theoretical CE

Table 2 below gives the annualized theoretical CE of the efficient portfolio for different levels of γ. The table also reports the values for the difference in the squared Sharpe ratio, Δ_{SSR}, and the variance of the GMVP, σ^2_{gmv}, based on monthly data for the three different portfolio sizes. While Δ_{SSR}, which reflects the potential financial gains over the minimum variance portfolio, has a positive effect on the CE, the second measure reflects the baseline risk inherent in the portfolio choice problem.

Table 2: Theoretical CE for Different Portfolios and Degrees of Risk Aversion

	Annualized CE (%)								Δ_{SSR}	σ^2_{gmv}
γ	0.04	0.5	1	2	4	6	8	10	(%)	(%)
5-PF	43.51	14.06	12,07	9.97	6.72	3.67	0.68	-2.30	0.2085	0.2452
10-PF	106.01	18.02	13.79	11.05	8.41	6.40	4.56	2.78	0.6348	0.1405
30-PF	427.21	43.44	26.43	17.40	11.85	9.08	7.00	5.20	2.7786	0.1152

Annualized theoretical CE of efficient portfolio, Δ_{SSR} and variance of the GMVP, σ^2_{gmv}, for three different samples. All values are scaled by 10^2.

[6] *http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html*

A comparison of the values for Δ_{SSR} and σ^2_{gmv} reveals that, as the number of asset increases, Δ_{SSR} increases but σ^2_{gmv} decreases only moderately. The two quantities are similar in size in 5 industry portfolio case, but Δ_{SSR} is almost 25 times higher than σ^2_{gmv} in the 30 portfolio case. This explains why the theoretical CE substantially increases when more assets are considered. Since all three portfolio data sets are constructed from the same stocks, the heterogeneity between assets is large in the large dimension cases. However, grouping the assets to larger portfolios lowers the dissimilarity of the asset candidates and consequently restricts the space of portfolio optimization reflected by the values for Δ_{SSR} and σ^2_{gmv}.

Moreover, for all three data sets, the corresponding values of Δ_{SSR} are relatively small with a maximum of 2.78% for the 30 asset case. As shown previously, the difference between the expected CE loss and the expected CE loss for the out-of-sample case mainly depends on the magnitude of Δ_{SSR}. Since Δ_{SSR} is changing only moderately for our data constellations, we refrain from reporting the simulation results for the out of sample case.

The risk aversion level of $\gamma = 0.04$ (first column) is equivalent to a risk tolerance of 50 considered by Chopra and Ziemba (1993). In this case, the investor is nearly risk-neutral and can achieve with the MV portfolio a very high annualized CE. As the γ increases, however, the investor cares more about the risk and the theoretical CE decreases. The CE can even become negative if the investor is too risk averse and, in this case, the risky investment becomes unattractive.

As given by (3.3) and (3.8), Δ_{SSR} and σ^2_{gmv} have different effects on the risk of the estimated efficient portfolio and the GMVP for different levels of γ. For instance, if we use the estimated GMVP, the impact of estimation error of the covariance matrix cannot be neglected in the case of 5 industry portfolios. It is, however, less relevant in the case of 30 industry portfolios where the CE loss of the estimated GMVP is mainly caused by the theoretical difference between the GMVP and the efficient portfolio. In addition, we can also show that as γ increases, the impact of Δ_{SSR} decreases, whereas the impact of σ^2_{gmv} increases.

Theoretical CE and the Impact of Near-Multicollinearity

Theoretical CE and Δ_{SSR} can be heavily affected by the population properties of the asset pool under consideration. In the following, we show how the correlation level affects Δ_{SSR}, σ^2_{gmv} and the CE using the correlation matrix from the three portfolios as the benchmark. More specifically, we keep the means and the variances of the portfolios unchanged but replace the correlation matrix by

$$C_k = k \cdot C_0 + (1-k) \cdot I,$$

where C_0 is the original correlation matrix estimated from the data and k is the parameter determining the strength of the correlation structure. Here, we assume that all correlations change proportionally. It is known that all eigenvectors of C_k are the same as the eigenvector of C_0, but the eigenvalues of C_k are $\lambda_{k,j} = k\lambda_j + (1-k)$, $j = 1, \ldots, N$, where λ_j is the j-th eigenvalue of the original correlation matrix C_0. To ensure the positive definiteness of C_k, all $\lambda_{k,j}$ $j = 1, \ldots, N$ must be positive. Thus, we select the value of k from the interval $[0, (1-\lambda_{0,min})^{-1})$, where λ_{min} is the smallest eigenvalue of C_0.

From Figure 1 we see that, with increasing correlation level, Δ_{SSR} increases substantially and, therefore, the theoretical CE of the efficient portfolio increases. Based on Equation (3.8) the theoretical loss of the GMVP, $\frac{1}{2\gamma}\Delta_{SSR}$, also increases as the correlation level increases. However, since the σ^2_{gmv} is almost unchanged, we can conclude that the expected CE loss of the GMVP caused by estimation risk, which is proportional to σ^2_{gmv}, does not change too much as the correlation level increases (see Equation (3.8) for comparison). Hence, the aggregate loss of the GMVP can be dominated by its theoretical part in the case where asset returns are highly correlated. In addition, in the case of large dimensional portfolio choice problem, Δ_{SSR} is large for high correlations. Thus, the GMVP also performs poorly in this case and the difference between the GMVP and the equally weighted portfolio is small.

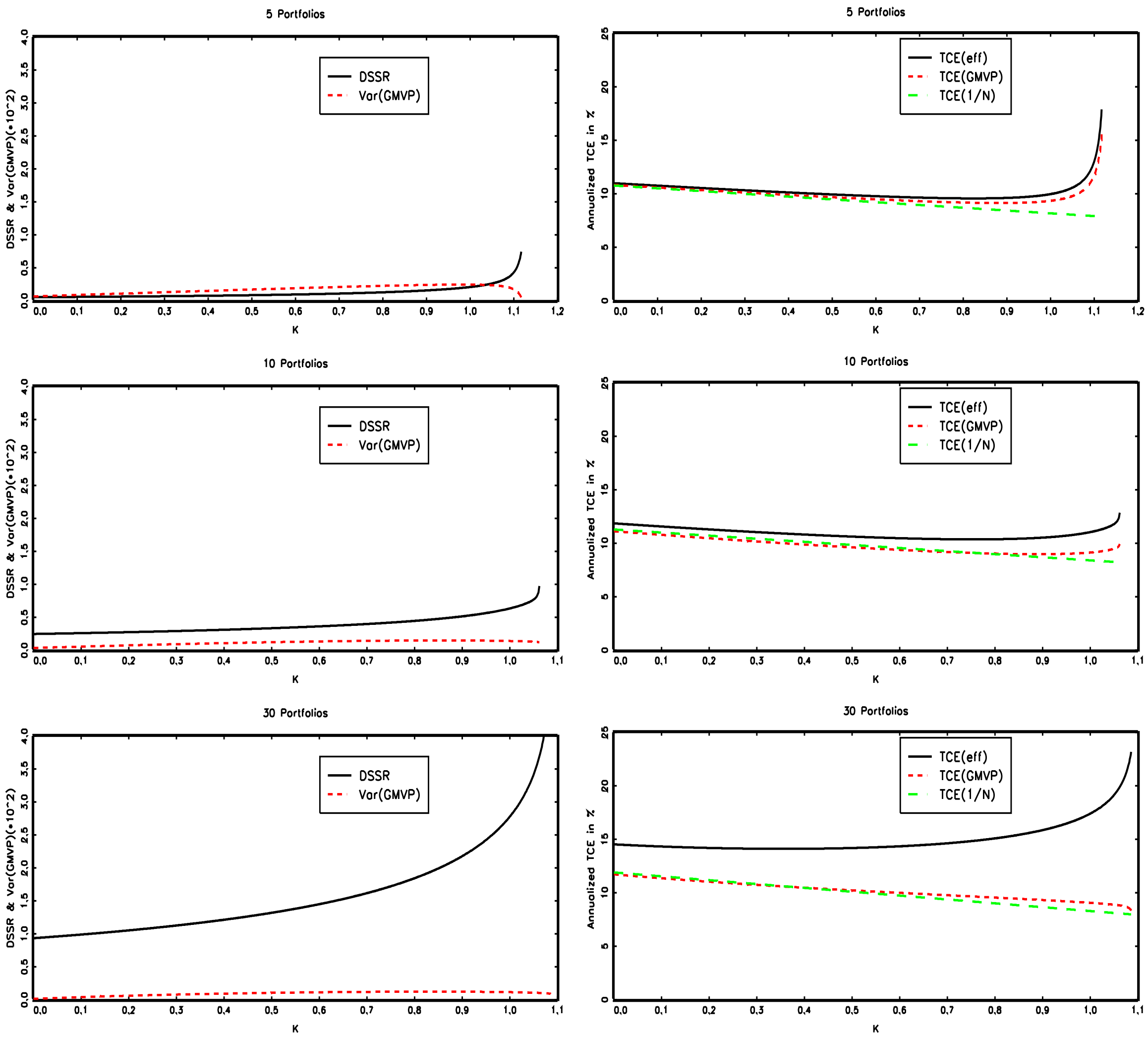

Figure 1: Left Panel: Impact of correlation on Δ_{SSR} (solid line) and σ^2_{gmv} (dashed line) for the three portfolios. Right Panel: Impact of correlation on annualized theoretical CE (in %) on 1) efficient portfolio (TCE(eff), solid); 2), GMVP (TCE(GMVP), small dashes); 3) equally weighted portfolio (TCE(1/N), long dashes). $\gamma = 2$.

Impact of Risk Aversion

Table 3 gives the annualized CE of the efficient theoretical portfolio, the theoretical GMVP and the equally weighted portfolio for our three different data sets. By definition, $CE(w^*_{eff})$ dominates the two other strategies and all CEs are monotonically decreasing functions of the risk aversion level. Their relative performance, however, depends on the specific parameter constellations. Note that, even in this theoretical scenario where estimation risk is still ignored, the performance of the two suboptimal strategies comes close to the efficient portfolio strategy. For instance, in 5 industry portfolio case, the performance of the GMVP is quite close to the performance of the efficient portfolio for γ larger than 2. For the 30 industry portfolio case, however, the theoretical efficient portfolio outperforms the GMVP significantly, even for high levels of risk aversion due to the strong increase of Δ_{SSR}.

In addition, because the GMVP has a relatively small theoretical loss in the 5 portfolio case, it is more attractive than the equally weighted portfolio even for a less risk averse investor. But in the large dimension case, the ranking of the GMVP and the equally weighted portfolio can change when different risk aversion levels are considered. For the less risk averse investor, the equally weighted portfolio could be more attractive.

Table 3: Annualized CE (in %) of efficient portfolio, GMVP and equally weighted portfolio

γ	0.04	0.5	1	2	4	6	8	10
5-PF								
w^*_{eff}	43.51	14.06	12.07	9.97	6.72	3.67	0.68	-2.30
w_{gmv}	12.23	11.56	10.82	9.35	6.41	3.46	0.52	-2.42
$1/N$	11.52	10.73	9.88	8.18	4.77	1.36	-2.05	-5.46
10-PF								
w^*_{eff}	106.01	18.02	13.79	11.05	8.41	6.40	4.56	2.78
w_{gmv}	10.79	10.40	9.98	9.14	7.46	5.77	4.08	2.40
$1/N$	11.68	10.91	10.08	8.40	5.06	1.71	-1.64	-4.99
30-PF								
w^*_{eff}	427.21	43.44	26.43	17.40	11.85	9.08	7.00	5.20
w_{gmv}	10.42	10.10	9.76	9.06	7.68	6.30	4.92	3.54
$1/N$	12.03	11.15	10.20	8.29	4.48	0.67	-3.14	-6.95

Our findings highlight the importance of the level of risk aversion for the evaluation of portfolio strategies. In many horse races of various portfolio strategies presented in the literature, the risk aversion parameter is usually arbitrarily given, e.g. a popular value for the risk aversion parameter is 2 (e.g. Best and Grauer (1991) and DeMiguel, Garlappi, and Uppal (2009)), which is assumed to be the market average risk aversion level. However, when this specific risk aversion level is used to compare the GMVP with the equally weighted portfolio, it is likely that no large differences in the performance of the two strategies will be found and one may be inclined to conclude that complete ignorance of any portfolio optimization strategy by using the equally weighted portfolio is quite meaningful.

5.2 Properties of the Expected CE Loss

As shown before, the theoretical CE loss of the GMVP decreases as the risk aversion level increases. This relationship, however, no longer holds for the expected CE loss based on the plug-in estimated portfolio weights. To illustrate this, in the following we use the 10 and 30 industry portfolios based on sample size $T = 60$ because, in this case, the estimation risk of the GMVP also contributes a considerable fraction to the overall risk of the estimated GMVP. The results are depicted in Figure 2.

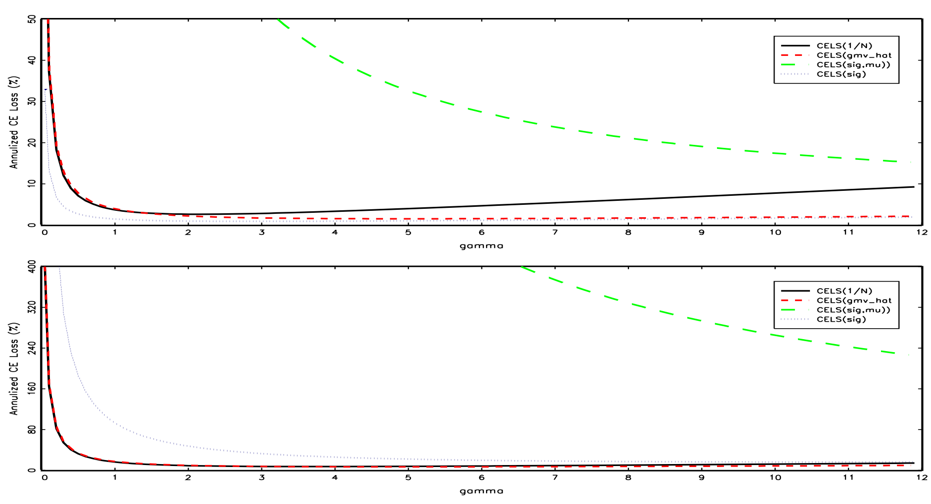

Figure 2: Relation between γ and expected CE loss for different portfolio strategies:
1) CELS(1/N) - solid line: Equally weighted portfolio, $\mathcal{R}((1/N)\iota|w^*_{eff})$;
2) CELS(gmv_hat) - dark dashed line: Empirical GMVP, $\mathcal{R}(w_{gmv}(S)|w^*_{eff})$;
3) CELS(sig,mu) - light dashed line: Empirical efficient PF, $\mathcal{R}(w_{eff}(\bar{r}, S)|w^*_{eff})$;
4) CELS(sig) - dotted line: Empirical efficient PF with μ known, $\mathcal{R}(w_{eff}(\mu, S)|w^*_{eff})$.
Upper panel: 10 portfolio case, lower panel: 30 portfolio case. Sample size: $T = 60$.

Figure 2 shows the (annualized) expected CE losses of different portfolio strategies in the 10-portfolio (upper panel) and the 30-portfolio (lower panel) cases. In the 5 portfolio case, the result is similar to that in the case of 10 portfolios and is not reported here. Accounting for estimation risk changes the ranking of the portfolio strategies completely. The empirical counterpart of the efficient portfolio (plug-in estimator) is inferior to any other portfolio strategy considered over the entire range of γ. The equally weighted portfolio is only a strong competitor for investors with low levels of risk aversion. This explains the results of many horse races where the equally weighted portfolio performs very well. If one assumes, that professional investors are less risk averse than private investors our findings imply that in particular professional investors should be cautious in applying the the empirical counterparts of efficient theoretical portfolio strategies. The comparison of $\mathcal{R}(w_{eff}(\bar{r}, S)|w^*_{eff})$ with $\mathcal{R}(w_{eff}(\mu, S)|w^*_{eff})$ (case 3 and 4 in Figure 2) reveals that the gains of knowing the true mean return vector are considerable. However, in the large dimension (30 portfolio) case, simply estimating Σ and adopting the empirical GMVP outperforms the efficient portfolio strategy based on the true mean. Even in the 10 portfolio case, the empirical GMVP performs as well as this (infeasible) efficient portfolio strategy for $\gamma \geq 2$. Note that, with the exception of the equally weighted portfolio, the expected CE loss is declining in γ, if the value of gamma is within a reasonable range. Because the theoretical CE is always declining in γ, we can conclude that the decrease in the expected CE for the empirical portfolio weights is less pronounced, i.e. the more risk averse investor loses less due to estimation risk than the less risk averse investor.

Relevance of Estimation Risk: Mean vs. Covariance

As shown for the cases 3 and 4 in Figure 2 knowledge of subsets of the parameters of the return distribution can lead to a major reduction in estimation risk. In the following, we provide additional numerical evidence on the relative importance of estimation risk with respect to the mean and the covariance matrix. For this, we compute the percentage of

$\mathcal{R}(w_{eff}(\bar{r}, \Sigma)|w^*_{eff})$ and $\mathcal{R}(w_{eff}(\mu, S)|w^*_{eff})$ in the total expected CE loss of the empirical efficient portfolio.

Table 4 (upper panel) provides results on the expected CE loss for the efficient portfolio if both μ and Σ are estimated by their sample counterparts. In most of the cases, using plug-in portfolio weights leads to substantial expected losses. Only the risk averse investor focusing on small portfolios and using large (in practice rather unrealistic) sample sizes can expect minor losses due to estimation. The expected estimation loss literally explodes for the less risk averse investor with large portfolios (N = 30). In this case, even a large sample size of T=180 (15 years of monthly data) does not really mitigate the problem.

The lower panel of Table 4 reports on the partial risks and the overall risk of the sample efficient portfolio based on the three data sets. With the help of Equation (3.5), we can decompose the overall estimation risk into its three components: i.) the expected CE loss resulting from estimating the mean, ii.) the expected CE loss resulting from estimating the variance and iii) the interaction effect for the latter if, in addition, the mean has to be estimated. Interestingly, the expected CE loss due to estimation of Σ is rather small in percentage terms for all three portfolio sizes and sample sizes. However, its contribution to the overall risk becomes relevant if the mean has to be estimated as well. For T = 60 and N = 30, the scale effect explains almost 84% of the expected CE loss. Our calibration exercise makes clear that horse races for the "best" covariance estimator, which take mean returns as given, assume away a large fraction of estimation uncertainty.

Table 4: Relative expected CE Loss due to Estimation Error in Mean and Covariance (in %)

Annualized Expected Loss of Sample Efficient Portfolios

	5 PF			10 PF			30 PF		
$T \setminus \gamma$	1	2	8	1	2	8	1	2	8
60	52.55	26.44	7.43	159.27	79.87	21.13	2585.39	1293.73	328.62
120	22.88	11.52	3.27	59.13	29.67	7.94	359.98	180.33	46.77
180	14.59	7.35	2.09	35.97	18.05	4.85	173.50	86.95	22.75

Impact of Interaction Effect

	5 PF			10 PF			30 PF		
$T \setminus \gamma$	1	2	8	1	2	8	1	2	8
$s_{(\bar{r})}$					(%)				
60	76.11	75.64	67.32	56.51	56.34	53.25	11.22	11.21	11.03
120	87.43	86.84	76.55	76.10	75.83	70.85	40.28	40.20	38.75
180	91.37	90.74	79.73	83.39	83.08	77.32	55.72	55.59	53.12
$s_{(S)}$					(%)				
60	0.50	1.11	12.00	0.92	1.21	6.64	3.59	3.67	5.19
120	0.50	1.17	12.89	0.88	1.23	7.72	3.41	3.60	7.08
180	0.50	1.19	13.18	0.87	1.24	8.09	3.28	3.50	7.78
$s_{(\bar{r},S)}$					(%)				
60	23.39	23.24	20.69	42.57	42.45	40.11	85.19	85.12	83.78
120	12.07	11.99	10.57	23.02	22.94	21.43	56.31	56.20	54.17
180	8.13	8.07	7.09	15.74	15.68	14.59	41.00	40.91	39.09

i.) $s_{(\bar{r})} = \dfrac{\mathcal{R}(w_{eff}(\bar{r},\Sigma)|w^*_{eff})}{\mathcal{R}(w_{eff}(\bar{r},S)|w^*_{eff})}$ = share of the expected CE loss due to estimation of μ;

ii.) $s_{(S)} = \dfrac{\mathcal{R}(w_{eff}(\mu,S)|w^*_{eff})}{\mathcal{R}(w_{eff}(\bar{r},S)|w^*_{eff})}$ = share of expected CE loss due to estimation of Σ;

iii.) $s_{(\bar{r},S)} = \dfrac{(c_3-1)\mathcal{R}(w_{eff}(\bar{r},\Sigma)|w^*_{eff})}{\mathcal{R}(w_{eff}(\bar{r},S)|w^*_{eff})}$ = interaction effect.

5.3 Shrinkage Portfolio

Although under the normality assumption the true parameters of the return process are estimated efficiently, the evidence provided in this study reveals that the empirical weights of the efficient portfolio imply a high estimation risk, while theoretically less efficient strategies perform comparatively well. In the following, we compare the expected CE loss of different portfolio strategies discussed in the previous sections. Since the modified estimator for Δ_{SSR} introduces an additional source of estimation risk for the shrinkage portfolio and the analytical expected CE loss of this estimated shrinkage portfolio is difficult to obtain, we compute the expected CE loss by means of Monte-Carlo methods using 10000 replications. The expected losses for the different portfolio strategies are depicted in Figure 3.

The performance of the five approaches varies considerably depending on the sample size. Their relative performance is, however, qualitatively the same for the two risk aversion levels $\gamma = 2$ and $\gamma = 8$. Figure 3 confirms the strength of the equally weighted portfolio performance over optimized portfolio strategies for small and moderate sample sizes ($T \leq 240$) and $\gamma = 2$. Only for very large, i.e. in practice unrealistic, sample sizes, the equally weighted portfolio is clearly beaten by the optimized empirical portfolios. In any case, using the empirical GMVP generally dominates the equally weighted portfolio.

If the optimal shrinkage parameter is known, the shrinkage method optimally combines the sample efficient portfolio and the empirical GMVP and significantly reduces the expected CE loss caused by estimation error. This benefit is particularly large if the theoretical risk of the GMVP is large, e.g. in the 30 portfolio case. However, in practice, the estimation error in the shrinkage parameter reduces the benefits of using the shrinkage portfolio. As was shown in the previous section, the shrinkage approach cannot reduce the baseline risk component. Therefore, as the dimension of the portfolio choice problem increases, the baseline risk component dominates, so that shrinkage portfolio also performs poorly and can be even worse than the equally weighted portfolio. Furthermore, if the risk aversion level is high, the optimal shrinkage portfolio, the estimated shrinkage portfolio and the empirical GMVP perform similarly because the true GMVP is theoretically close to the true efficient portfolio and has relative low theoretical CE loss.

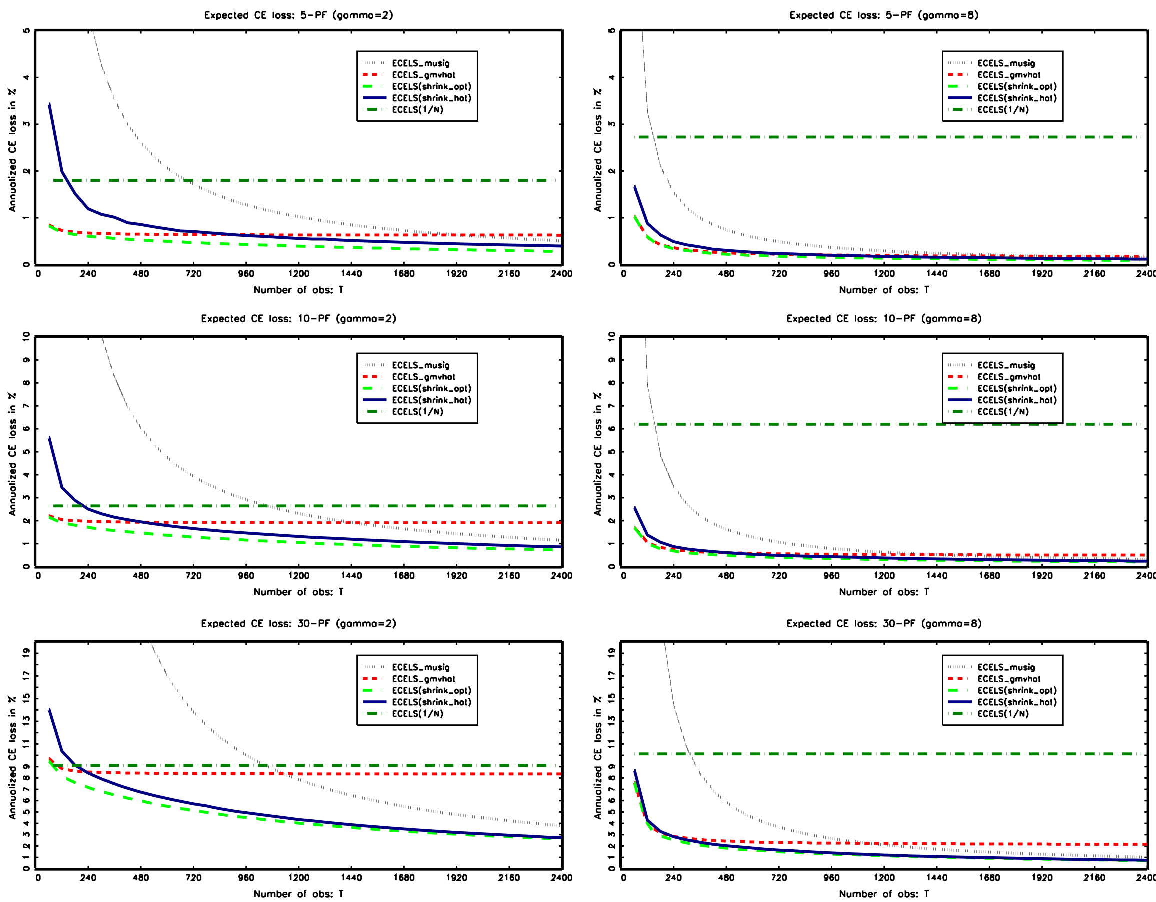

Figure 3: Expected CE loss of estimated efficient portfolio, estimated GMVP, and equally weighted portfolio for two different degrees of risk aversion:
1) ECELS_musig - dotted line: $\mathcal{R}(w_{eff}(\bar{r}, S)|w^*_{eff})$, empirical efficient PF
2) ECELS_gmvhat - short dashed line: $\mathcal{R}(w_{gmv}(S)|w^*_{eff})$, empirical GMVP;
3) ECELS(shrink_opt) - long dashed line: $\mathcal{R}(w_{eff}(\eta^*, \bar{r}, S)|w^*_{eff})$, efficient shrinkage PF;
4) ECELS(shrink_hat) - solid line: $\mathcal{R}(w_{eff}(\hat{\eta}, \bar{r}, S)|w^*_{eff})$, shrinkage PF;
5) ECELS(1/N) - dotted-slashed line: $\mathcal{R}((1/N)\iota|w^*_{eff})$, equally weighted portfolio.

5.3.1 Deviation from Normality

The assumption of i.i.d. normality for the return vector can be easily relaxed if we only consider the estimation risk of the mean given Σ is known. In this case, we only need the assumption of a serially uncorrelated return series to derive the expected CE loss analytically. In the following, we study the robustness of our results for deviations from normality, if the covariance matrix has to be estimated. For this, we simulate i.i.d.

student-t distributed returns with the same mean vectors and covariance matrices as we used previously, i.e.

$$r_t = \sqrt{\frac{\nu - 2}{W_t}} \cdot Y_t + \mu,$$

where $Y \overset{iid}{\sim} N(0, \Sigma)$, $W \overset{iid}{\sim} \chi^2_v$, ν is the degree of freedom of the multivariate t distribution, μ and Σ are the specified mean and covariance of r_t.

In order to study the effect of large kurtosis, we assume 5 degrees of freedom and compute the expected CE losses for the sample efficient portfolio, the empirical GMVP and the estimated shrinkage portfolio. Our estimates of the expected CE losses are based on 10000 simulations. The result is reported in Table 5. If returns are student-t distributed, the performance of the estimated portfolio deteriorates compared to the results based on normal returns, but the difference turns out to be rather small.

Table 5: Expected CE Loss for t-distributed Returns

	CE of $\hat{w}_{eff}$ (%)				CE of $\hat{w}_{gmv}$ (%)				CE of $\hat{w}_{shrink}$ (%)			
	$\gamma = 2$		$\gamma = 8$		$\gamma = 2$		$\gamma = 8$		$\gamma = 2$		$\gamma = 8$	
T	$\mathcal{N}$	t_5	$\mathcal{N}$	t_5	$\mathcal{N}$	t_5	$\mathcal{N}$	t_5	$\mathcal{N}$	t_5	$\mathcal{N}$	t_5
	5-PF											
60	-16,40	-17.73	-6.73	-7.25	9.13	9.03	-0.35	-0.74	7.84	7.75	-0.65	-1.03
180	2.65	2.46	-1.41	-1.65	9.28	9.24	0.25	0.08	8.95	8.91	0.16	-0.02
300	5.73	5.62	-0.54	-0.67	9.31	9.28	0.36	0.25	9.15	9.13	0.32	0.21
	10-PF											
60	-68,23	-73,09	-16,42	-17,59	8,83	8,78	2,85	2,62	6,52	6,43	2,24	2,04
180	-6,88	-7,56	-0,26	-0,47	9,05	9,03	3,73	3,64	8,53	8,44	3,58	3,49
300	1,02	0,80	1,86	1,75	9,09	9,08	3,87	3,82	8,93	8,90	3,83	3,77
60	-1232,35	-1349,00	-310,62	-345,43	7,68	7,63	-0,61	-0,86	4,33	3,98	-1,47	-1,82
180	-67,05	-71,67	-15,12	-16,44	8,80	8,78	3,84	3,74	8,42	8,42	3,74	3,63
300	-22,68	-24,61	-3,58	-4,12	8,92	8,90	4,32	4,27	9,60	9,53	4,49	4,43

6 Conclusion

This paper takes a closer look at the quality of efficient portfolios compared to other portfolio strategies accounting for the budget constraint in a world where the parameters of the return process are unknown and have to be estimated by the investor. The relative performance of the different empirical portfolio strategies depends on the magical quadrangle between i.) the theoretical properties of the return process (e.g. the eigenvalues of the covariance matrix), ii.) the estimation properties of their sample counterparts, notably sample size, iii.) the number of assets and iv.) the investor's risk preferences. While in theory there is a well-defined dominant portfolio strategy, the ranking of portfolio strategies becomes unclear when the portfolio weights are based on estimated parameters. In this case, the ranking depends on the particular parametric constellation within the magical quadrangle. Therefore, in the light of estimation risk, a theoretically suboptimal portfolio can turn into a reasonable choice once sample size, size of the portfolio and risk preferences are taken into account.

Using the concept of implied means, we are able to represent the weights of any portfolio with the theoretically efficient portfolio weights. This allows us to decompose the overall estimation risk into their single components. Unlike previous empirical studies comparing the estimation risks in means and covariances separately, our study clearly shows why and when a precise estimate of covariances becomes necessary. Although the estimation of mean returns is crucial for almost all financial decisions, the impact of estimation error in covariances can never be neglected in the presence of estimation risk in the mean. Thus, to evaluate different covariance estimators, comparing their pure statistical metrics to the true covariance is questionable. An estimate of covariance matrix which has lower statistical risk is not necessarily superior if a financial decision rule is considered.

Based on the property that the weights of the efficient portfolio can be represented as the sum of the weights of the global minimum variance portfolio and the zero investment portfolio, we can show that Bayes-Stein shrinkage estimation of mean returns, shrinkage of the portfolio weights towards the weights of the global minimum variance portfolio as well as plug-in estimation of the efficient portfolio assuming a higher degree of risk aversion are equivalent strategies to reduce estimation risk. In a calibration study, we show that the expected CE loss of the efficient portfolio due to estimation is non-negligible for realistic empirical scenarios. Moreover, we show that shrinkage leads to superior choices of the portfolio weights compared to the empirical efficient portfolio but also compared to the simple $1/N$ strategy.

Admittedly, the results presented in this paper are based on the most simple portfolio set-up allowing us to derive finite sample properties for the estimated portfolio weights and the CE based on estimated portfolio weights. Extending our findings to the case of a dynamic price process, where the investor forms expectations on the return process given past filtration, would be desirable. Moreover, as sample size is a major determinant of estimation risk particularly for large portfolios, optimal empirical portfolio strategies in the presence of structural breaks should be derived combing optimally pre- and post-break information.

References

Best, M. J., and R. R. Grauer (1991): "On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results," *Review of Financial Studies*, 4(2), 315–342.

Black, F., and R. Litterman (1992): "Global Portfolio Optimization," *Financial Analysts Journal*, 48, 28–43.

Britten-Jones, M. (1999): "The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights," *The Journal of Finance*, 54(2), 655–671.

Cho, D. D. (2010): "Estimation risk in covariance," Journal of Asset Management, (forthcoming).

CHOPRA, V. K., AND W. T. ZIEMBA (1993): "The Effect of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice," *Journal of Portfolio Management*, 19(2), 6–11.

DEMIGUEL, V., L. GARLAPPI, AND R. UPPAL (2009): "Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?," *Review of Financial Studies*, 22(5), 1915–1953.

FRAHM, G. (2010): "An Analytical Investigation of Estimators for Expected Asset Returns from the Perspective of Optimal Asset Allocation," University of Cologne, Seminar of Economic and Social Statistics, Discussion Paper 1/10.

FRAHM, G., AND C. MEMMEL (2010): "Dominating estimators for minimum-variance portfolios," *Journal of Econometrics*, 159(2), 289–302.

JAGANNATHAN, R., AND T. MA (2003): "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," *Journal of Finance*, 58(4), 1651–1683.

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.

KEMPF, A., AND C. MEMMEL (2006): "Estimating the Global Minimum Variance Portfolio," *Schmalenbach Business Review*, 58, 332348.

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

OKHRIN, Y., AND W. SCHMID (2006): "Distributional properties of portfolio weights," *Journal of Econometrics*, 134, 235256.

A Appendix

Lemma A.1 *Let I denote the identity matrix. Define matrix $B = \Sigma \cdot A = I - \frac{\iota\iota'\Sigma^{-1}}{\iota'\Sigma^{-1}\iota}$. We have:*

1. *A is semi-positive definite and $A \cdot \Sigma \cdot A = A \cdot B = A$,*

2. *$A \cdot x = 0$ if and only if $x = \iota$,*

3. *$tr(B) = N - 1$.*

Proof A.1 *See e.g .Okhrin and Schmid (2006).*

Proof A.2 (*Proposition 2.1*) *Substituting the solution (2.2) in the CE leads to:*

$$
\begin{aligned}
&CE(w_{eff}) \\
=&\mu'\left(w_{gmv} + \frac{1}{\gamma}\cdot A\cdot\mu\right) - \frac{\gamma}{2}\left(w_{gmv} + \frac{1}{\gamma}\cdot A\cdot\mu\right)'\Sigma\left(w_{gmv} + \frac{1}{\gamma}\cdot A\cdot\mu\right) \\
=&\mu' w_{gmv} - \frac{\gamma}{2}w'_{gmv}\Sigma w_{gmv} + \frac{1}{\gamma}\mu'\cdot A\cdot\mu \\
&-\frac{\gamma}{2}\left(\frac{1}{\gamma}\cdot A\cdot\mu\right)'\Sigma\left(\frac{1}{\gamma}\cdot A\cdot\mu\right) - \gamma w'_{gmv}\Sigma\left(\frac{1}{\gamma}\cdot A\cdot\mu\right)
\end{aligned}
\tag{A.1}
$$

Since $A \cdot \Sigma \cdot A = A$, and $\iota' \cdot A = 0$, we have

$$\frac{\gamma}{2}\left(\frac{1}{\gamma}\cdot A\cdot\mu\right)'\Sigma\left(\frac{1}{\gamma}\cdot A\cdot\mu\right) = \frac{1}{2\gamma}\mu'\cdot A\cdot\mu$$

and

$$w'_{gmv}\Sigma\left(\frac{1}{\gamma}\cdot A\cdot\mu\right) = \frac{1}{\gamma}\cdot\frac{\iota'\Sigma^{-1}}{\iota'\Sigma^{-1}\iota}\Sigma\cdot A\cdot\mu = 0$$

Substituting the GMVP weight $w_{gmv} = \Sigma^{-1}\iota/\left(\iota'\Sigma^{-1}\iota\right)$ in equation (A.1) leads to:

$$
\begin{aligned}
CE(w_{eff}) =&\mu' w_{gmv} - \frac{\gamma}{2}w'_{gmv}\Sigma w_{gmv} + \frac{1}{\gamma}\mu'\cdot A\cdot\mu - \frac{1}{2\gamma}\mu'\cdot A\cdot\mu \\
=&CE(w_{gmv}) + \frac{1}{2\gamma}\mu'\cdot A\cdot\mu \\
=&\frac{1}{2\gamma}\mu'\cdot A\cdot\mu + \frac{1}{\iota'\Sigma^{-1}\iota}(\iota'\Sigma^{-1}\mu - \frac{\gamma}{2})
\end{aligned}
$$

Proof A.3 (*Proposition 2.2*)

$$\begin{aligned}
&CE(w_{eff}) - CE(\hat{w}) \\
=&[\mu' w_{eff} - \frac{\gamma}{2} w_{eff}' \Sigma w_{eff}] - [\mu' \hat{w} - \frac{\gamma}{2} \hat{w}' \Sigma \hat{w}] \\
=&[\mu' w_{eff} - \frac{\gamma}{2} w_{eff}' \Sigma w_{eff}] - [\mu'(w_{eff} + \hat{w} - w_{eff}) - \frac{\gamma}{2}(w_{eff} + \hat{w} - w_{eff})'\Sigma(w_{eff} + \hat{w} - w_{eff})] \\
=&[\mu' w_{eff} - \frac{\gamma}{2} w_{eff}' \Sigma w_{eff}] - \mu'(\hat{w} - w_{eff}) - \mu' w_{eff} \\
&+ \frac{\gamma}{2}(\hat{w} - w_{eff})'\Sigma(\hat{w} - w_{eff}) + \gamma(\hat{w} - w_{eff})\Sigma w_{eff} + \frac{\gamma}{2} w_{eff}' \Sigma w_{eff} \\
=& - \mu'(\hat{w} - w_{eff}) + \frac{\gamma}{2}(\hat{w} - w_{eff})'\Sigma(\hat{w} - w_{eff}) + \underbrace{\gamma(\hat{w} - w_{eff})'\Sigma w_{eff}}_{=(*)}
\end{aligned}$$

Since $w_{eff} = \frac{1}{\gamma}\Sigma^{-1}\mu + (1 - \frac{1}{\gamma}\iota'\Sigma^{-1}\mu)\frac{\Sigma^{-1}\iota}{\iota'\Sigma^{-1}\iota}$ *and* $\sum_i w_i = \iota' w = 1$, *we have :*

$$\begin{aligned}
(*) &= (\hat{w} - w_{eff})'\mu + \underbrace{(\hat{w} - w_{eff})'\iota}_{=0} \frac{\gamma}{\iota'\Sigma^{-1}\iota} - \underbrace{(\hat{w} - w_{eff})'\iota}_{=0} \frac{\iota'\Sigma^{-1}\mu}{\iota'\Sigma^{-1}\iota} \\
&= (\hat{w} - w_{eff})'\mu
\end{aligned}$$

Therefore:

$$CE(w_{eff}) - CE(\hat{w}) = \frac{\gamma}{2}(w_{eff} - \hat{w})'\Sigma(w_{eff} - \hat{w}).$$

The risk function can be easily obtained by taking the expectation of the CE loss.

Proof A.4 (*Proposition 2.3*) *If there is no error in covariance matrix we have:*

$$w_{eff} - w_{eff}(\hat{\mu}, \Sigma) = \frac{1}{\gamma} \cdot A \cdot (\mu - \hat{\mu})$$

Therefore the loss of CE is:

$$CE(w_{eff}) - CE\left(w_{eff}(\hat{\mu}, \Sigma)\right) = \frac{1}{2\gamma}(\mu - \hat{\mu})' \cdot A' \cdot \Sigma \cdot A(\mu - \hat{\mu}) = \frac{1}{2\gamma}(\mu - \hat{\mu})' \cdot A(\mu - \hat{\mu}).$$

The risk function can be easily obtained by taking the expectation of the CE loss.

Proof A.5 (*Proposition 2.4*) *In this proposition we show that for any given portfolio*

$\hat{w}$, *there is a implied mean vector* $\hat{\mu}_{im}$ *such that:*

$$\hat{w} \overset{!}{=} w_{eff}(\hat{\mu}_{im}, \Sigma) = w_{gmv} + \frac{1}{\gamma} \cdot A \cdot \hat{\mu}_{im} = w_{gmv} + \frac{1}{\gamma}\Sigma^{-1} \cdot B \cdot \hat{\mu}_{im}$$

Since ι *and the column vectors of* V *compose a basis of* $\mathbb{R}^N$, *the vector* $\hat{\mu}_{im}$ *can be written as linear combination of this basis, i.e.* $\hat{\mu}_{im} = c \cdot \iota + V \cdot c_0$ *with* $c_0 \in \mathbb{R}^{(N-1)}$. *Then:*

$$\hat{w} \overset{!}{=} w_{eff}(\hat{\mu}_{im}, \Sigma) = w_{gmv} + \frac{1}{\gamma}\Sigma^{-1} \cdot B \cdot \hat{\mu}_{im} = w_{gmv} + \frac{1}{\gamma}\Sigma^{-1} B \cdot V \cdot c_0 \qquad \text{(A.2)}$$

because $A \cdot \iota = 0$. *The solution of equation (A.2) is:*

$$\begin{aligned} c_0 &= \gamma \cdot (V'BV)^{-1} \cdot V' \cdot \Sigma \cdot (\hat{w} - w_{gmv}) \\ &= \gamma \cdot (V'BV)^{-1} \cdot V' \cdot \Sigma \cdot \hat{w} - \gamma \cdot (V'BV)^{-1} \cdot V' \cdot \Sigma \cdot \frac{\Sigma^{-1}\iota}{\iota'\Sigma\iota} \\ &= \gamma \cdot (V'BV)^{-1} \cdot V' \cdot \Sigma \cdot \hat{w} \end{aligned}$$

because $V'\iota = 0$. *Thus,* $\hat{\mu}_{im} = c \cdot \iota + V \cdot c_0 = c \cdot \iota + \gamma \cdot V \cdot (V'BV)^{-1} \cdot V' \cdot \Sigma \cdot \hat{w}$

Proof A.6 (*Proposition 2.5*) *Based on proposition 2.3 and proposition 2.4, the CE loss of an suboptimal portfolio* $\hat{w}$ *can be reformulated as:*

$$\begin{aligned} & CE(w_{eff}) - CE(\hat{w}) \\ = \quad & \frac{1}{2\gamma}(\mu - \hat{\mu}^*_{im})' \cdot A \cdot (\mu - \hat{\mu}^*_{im}) \\ = \quad & \frac{1}{2\gamma}(\mu - \hat{\mu}^*_{im})' \cdot A \cdot \Sigma \cdot A \cdot (\mu - \hat{\mu}^*_{im}) \\ = \quad & \frac{1}{2\gamma}(\mu - \hat{\mu}^*_{im})' \cdot \Sigma^{-\frac{1}{2}} \cdot (I - \frac{\Sigma^{-\frac{1}{2}}\iota\iota'\Sigma^{-\frac{1}{2}}}{\iota'\Sigma^{-1}\iota})' \cdot (I - \frac{\Sigma^{-\frac{1}{2}}\iota\iota'\Sigma^{-\frac{1}{2}}}{\iota'\Sigma^{-1}\iota}) \cdot \Sigma^{-\frac{1}{2}} \cdot (\mu - \hat{\mu}^*_{im}) \end{aligned}$$

where $\hat{\mu}^*_{im}$ *is the mean implied by the portfolio weight* $\hat{w}$.

Let C *denote the correlation matrix and* D *denote the diagonal matrix containing the* *standard deviations. Then* $\Sigma = D \cdot C \cdot D$ *and we have:*

$$
\begin{aligned}
& CE(w_{eff}) - CE(\hat{w}) \\
\leq & \parallel (I - \frac{\Sigma^{-\frac{1}{2}}\iota\iota'\Sigma^{-\frac{1}{2}}}{\iota'\Sigma^{-1}\iota}) \parallel_2^2 \parallel C^{-\frac{1}{2}} \parallel_2^2 \parallel D^{-1} \parallel_2^2 \parallel (\mu - \hat{\mu}^*_{im}) \parallel_2^2 \\
= & \parallel (I - \frac{\Sigma^{-\frac{1}{2}}\iota\iota'\Sigma^{-\frac{1}{2}}}{\iota'\Sigma^{-1}\iota}) \parallel_2^2 \cdot trace\left(C^{-1}\right) \cdot trace\left(D^{-2}\right) \cdot \parallel (\mu - \hat{\mu}^*_{im}) \parallel_2^2 \\
= & \frac{1}{2\gamma}(\sum_{i=1}^{N} \lambda_i^{-1}) \cdot (\sum_{i=1}^{N} \sigma_i^{-2}) \parallel (I - \frac{\Sigma^{-\frac{1}{2}}\iota\iota'\Sigma^{-\frac{1}{2}}}{\iota'\Sigma^{-1}\iota}) \parallel_2^2 \parallel (\mu - \hat{\mu}^*_{im}) \parallel_2^2 \\
= & \frac{1}{2\gamma}(\sum_{i=1}^{N} \lambda_i^{-1}) \cdot (\sum_{i=1}^{N} \sigma_i^{-2})(N-1) \parallel (\mu - \hat{\mu}^*_{im}) \parallel_2^2
\end{aligned}
$$

The last equation holds because:

$$
\begin{aligned}
\parallel I - \frac{\Sigma^{-\frac{1}{2}}\iota\iota'\Sigma^{-\frac{1}{2}}}{\iota'\Sigma^{-1}\iota} \parallel_2^2 &= trace\left(\left[I - \frac{\Sigma^{-\frac{1}{2}}\iota\iota'\Sigma^{-\frac{1}{2}}}{\iota'\Sigma^{-1}\iota}\right]' \left[I - \frac{\Sigma^{-\frac{1}{2}}\iota\iota'\Sigma^{-\frac{1}{2}}}{\iota'\Sigma^{-1}\iota}\right]\right) \\
&= trace\left(\Sigma^{\frac{1}{2}}\Sigma^{-\frac{1}{2}} \left[I - \frac{\Sigma^{-\frac{1}{2}}\iota\iota'\Sigma^{-\frac{1}{2}}}{\iota'\Sigma^{-1}\iota}\right]' \left[I - \frac{\Sigma^{-\frac{1}{2}}\iota\iota'\Sigma^{-\frac{1}{2}}}{\iota'\Sigma^{-1}\iota}\right] \Sigma^{\frac{1}{2}}\Sigma^{-\frac{1}{2}}\right) \\
&= trace\left(\Sigma^{\frac{1}{2}} \cdot A \cdot \Sigma \cdot A \cdot \Sigma^{\frac{1}{2}}\right) \\
&= trace\left(\Sigma^{\frac{1}{2}} \cdot A \cdot \Sigma^{\frac{1}{2}}\right) \\
&= trace\left(\Sigma \cdot A\right) \\
&= N - 1
\end{aligned}
$$