CFR Working Paper No. 05-02

On the Estimation of the Global Minimum Variance Portfolio

Alexander Kempf and Christoph Memmel
On the Estimation of the Global Minimum Variance Portfolio

Alexander Kempf and Christoph Memmel∗

Abstract

Expected returns can hardly be estimated from time series data. Therefore, many recent papers suggest investing in the global minimum variance portfolio. The weights of this portfolio depend only on the return variances and covariances, but not on the expected returns. The weights of the global minimum variance portfolio are usually estimated by replacing the true return covariance matrix by its time series estimator. However, little is known about the distributions of the estimated weights and return parameters of this portfolio. Our contribution is to determine these distributions. The knowledge of these distributions allows us to calculate the extent of the estimation risk an investor faces and to answer important questions in asset management.

Keywords: Global Minimum Variance Portfolio, Weight Estimation, Estimation Risk

JEL classification: C22, G11

∗University of Cologne, Department of Finance, Albertus-Magnus-Platz, 50923 Cologne, Germany, Tel.: +49-221-4702741, Fax: +49-221-4703992, Email: kempf@wiso.uni-koeln.de.
On the Estimation of the Global Minimum Variance Portfolio

Expected stock returns are hard to estimate [see, e.g., Merton (1980)]. Typically, the estimated values differ largely from the true ones. These estimation errors lead to a suboptimal portfolio composition and thus to a poor portfolio performance [see, e.g., Jorion (1991) and Chopra and Ziemba (1993)]. Therefore, several recent papers suggest avoiding the estimation of expected returns [see, e.g., Ledoit and Wolf (2003) and Jagannathan and Ma (2003)]. Instead, they assume that all stocks have equal expected returns. Under this assumption, all stock portfolios differ only with respect to their risk, but not with respect to expected returns. Therefore, the only efficient stock portfolios is the one with the smallest risk, i.e. the global minimum variance portfolio. All investors which optimize the tradeoff between expected return and risk of their portfolio should then combine the global minimum variance portfolio with the risk free asset.

The composition of the global minimum variance portfolio depends only on the covariance matrix of stock returns. Since the covariance matrix can be estimated much more precisely than the expected returns (see Section 1), the estimation risk of the investor is expected to be reduced by focusing on the global minimum variance portfolio.¹ However, little is known about the distribution of the estimated portfolio weights and the extent of the estimation risk. Dickinson (1974) calculates the unconditional distribution of the portfolio weights in the special case of two uncorrelated assets. Ohkrin and Schmid (2004) generalize this result by allowing \( N \) assets with arbitrary correlations. However, the conditional distribution is yet unknown, but it is necessary for calculating test statistics and confidence intervals in small samples. The main contribution of our paper is to derive the conditional distributions of the estimated weights of the global minimum variance portfolio, its estimated expected return and its estimated return variance. Knowing the conditional distributions allows us to answer important questions in asset management, for example: (i) What determines the extent of estimation risk? (ii) Can an investor reduce the portfolio risk significantly by including additional assets in his portfolio?

The paper is organized as follows. In Section 1 we show that the covariance matrix can be estimated much more precisely than expected returns. This provides...
the motivation for focusing on the global minimum variance portfolio. In Section 2 we briefly review the traditional approach of estimating the weights of the global minimum variance portfolio. In Section 3 we present an alternative OLS estimation approach, which leads to identical weight estimates. Using this alternative estimation approach we derive in Section 4 the conditional distributions of the estimated portfolio weights and the estimated return parameters of the global minimum variance portfolio. In Section 5 we show that our OLS approach can also be applied when we give up the assumption of normally distributed asset returns. In Section 6 we apply our results to calculate the estimation risk associated with the estimation of the global minimum variance portfolio. In Section 7 we give some examples of how to apply our results in international asset management. Section 8 concludes.

1 Precision of Parameter Estimates

Assume that there are $N$ stocks in the capital market. We denote the return of stock $i$ from time $t-1$ to $t$ by $r_{t,i}$. The vector $\mu$ contains the expected returns of the $N$ stocks. The $N \times N$ matrix $\Sigma$ contains the return variances and covariances $\sigma_{ij}$. We assume that the returns are multivariate normally distributed. In addition, the returns are identically and independently distributed. Thus, we assume the best possible situation for an investor who wants to estimate the returns distribution parameters. The investor can increase the precision of the estimate by using longer time series. If the length of the time series goes to infinity, both, the expected returns and the covariance matrix, can be estimated exactly. There is no estimation risk. However, real time series are not that long and the distribution parameters cannot be estimated exactly. Estimation risk occurs - even in the best possible situation an investor can face.

Assume that there are $\tau \geq 1$ years of data available to estimate the expected returns $\mu$ and the covariance matrix $\Sigma$. There are $n \geq 1$ subperiods of equal length per year. Thus, the number of observations is $T = \tau n$. The precision of the estimates is given by the variance of the estimators. The variances of the estimators for the expected
return $\mu_i$ and the standard deviation $\sigma_i$ are:

$$\text{var}(\hat{\mu}_i) = \frac{\sigma_i^2}{\tau}. \quad (1)$$

$$\text{var}(\hat{\sigma}_i) \approx \frac{1}{2} \frac{\sigma_i^2}{\tau n} \quad (2)$$

The precision of both estimates is the larger, the more years of data ($\tau$) are available. For $\tau \to \infty$ both variances go to zero, the estimation risk disappears. For a finite number of years ($\tau < \infty$), the covariance matrix can be estimated more precisely than the estimated returns. The precision ratio is given as:

$$\frac{\text{var}(\hat{\mu}_i)}{\text{var}(\hat{\sigma}_i)} \approx 2 n \quad (3)$$

For stock markets, one typically uses daily ($n = 250$), weekly ($n = 52$) or monthly ($n = 12$) observations. Therefore, typical precision ratios are within a range of 24 to 500. Thus, the covariance matrix can be estimated with a much higher precision than the vector of expected returns.

To illustrate the size of the estimation risk we provide a numerical example. Assume that the stock has a volatility $\sigma_i$ of 25%. There are weekly return observations available. Table 1 shows the level of the stock risk $\sigma_i$, the estimation risk of the expected returns $\sqrt{\text{var}(\hat{\mu}_i)}$ and the estimation risk of the volatility $\sqrt{\text{var}(\hat{\sigma}_i)}$. 
Table 1: Precision of Parameter Estimates

<table>
<thead>
<tr>
<th></th>
<th>$T = 1$</th>
<th>$T = 5$</th>
<th>$T = 10$</th>
<th>$T = 20$</th>
<th>$T = 50$</th>
<th>$T = 100$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_i$</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
</tr>
<tr>
<td>$\sqrt{\text{var}(\hat{\mu}_i)}$</td>
<td>0.250</td>
<td>0.112</td>
<td>0.079</td>
<td>0.056</td>
<td>0.035</td>
<td>0.025</td>
</tr>
<tr>
<td>$\sqrt{\text{var}(\hat{\sigma}_i)}$</td>
<td>0.025</td>
<td>0.011</td>
<td>0.008</td>
<td>0.005</td>
<td>0.003</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Table 1 shows that the estimation risk with respect to the volatility is comparably small. The opposite is true for the estimation risk with respect to expected returns. For short time periods, the estimation risk with respect to expected returns is huge, and even for long time series, it does not become negligible. Therefore, an investor might be well advised to abstain from estimating expected returns and to concentrate on the global minimum variance portfolio - even when stock returns have all the desirable features like normality and IID.

2 Traditional Approach

The global minimum variance portfolio ($MV$) is the stock portfolio with the lowest return variance for a given covariance matrix $\Sigma$. It is the solution to the following minimization problem:

$$
\min_{w=(w_1,\ldots,w_N)'} w'\Sigma w \quad \text{s.t.} \quad w'\mathbf{1} = 1
$$

(4)

$\mathbf{1}$ is a column vector of appropriate dimension whose entries are ones and $w = (w_1,\ldots,w_N)'$ is a vector of portfolio weights. The weights $w_{MV} = (w_{MV,1},\ldots,w_{MV,N})'$ of the global minimum variance portfolio are given as

$$
w_{MV} = \Sigma^{-1} \frac{\mathbf{1}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}.
$$

(5)
The expected return $\mu_{MV}$ and the return variance $\sigma^2_{MV}$ of the global minimum variance portfolio are given as

$$\mu_{MV} = \mu' w_{MV} = \frac{\mu' \Sigma^{-1} 1}{1' \Sigma^{-1} 1} \quad (6)$$

and

$$\sigma^2_{MV} = w'_{MV} \Sigma w_{MV} = \frac{1}{1' \Sigma^{-1} 1}. \quad (7)$$

The lower variance bound (7) can only be attained if the covariance matrix $\Sigma$ of the stock returns is known. As pointed out before, the covariance matrix $\Sigma$ is not known but has to be estimated in real markets. Typically, historical return observations are used for this estimation.

The traditional estimation approach is to replace the expected returns $\mu$ and the covariance matrix $\Sigma$ by their maximum likelihood estimators $\hat{\mu}$ and $\hat{\Sigma}$ in the Equations (5) - (7). The estimated portfolio weights $\hat{w}_{MV}$ and return parameters $\hat{\mu}_{MV}$ and $\hat{\sigma}^2_{MV}$ of the global minimum variance portfolio are non-linear functions of the stock return parameter estimates $\hat{\mu}$ and $\hat{\Sigma}$. Therefore, the distributions of $\hat{w}_{MV}$, $\hat{\mu}_{MV}$ and $\hat{\sigma}^2_{MV}$ are hard to determine, even if the distributions of the parameter estimates $\hat{\mu}$ and $\hat{\Sigma}$ are known. The calculation of these distributions is the main contribution of our paper.

3 OLS Approach

We use a regression based approach to determine the weights $w_{MV}$, the expected return $\mu_{MV}$ and the return variance $\sigma^2_{MV}$ of the global minimum variance portfolio. We rewrite the weights of the global minimum variance portfolio as regression coefficients. Without loss of generality we choose the return of stock $N$ to be the dependent variable:

$$r_{t,N} = \alpha + \beta_1 (r_{t,N} - r_{t,1}) + \cdots + \beta_{N-1} (r_{t,N} - r_{t,N-1}) + \varepsilon_t \quad t = 1, \ldots, T > N \quad (8)$$

$\varepsilon_t$ is a noise term that satisfies the standard assumptions of the classical linear regression model regarding errors. The returns are again normal and IID. The
three statements in Proposition 1 describe the relation between the linear regression and the global minimum variance portfolio.

**Proposition 1**

1. The regression coefficients $\beta_1, \ldots, \beta_{N-1}$ in Equation (8) correspond to the portfolio weights $w_{MV,1}, \ldots, w_{MV,N-1}$ of the global minimum variance portfolio:

   $$\beta_i = w_{MV,i}$$

(9)

2. The coefficient $\alpha$ in Equation (8) corresponds to the expected return $\mu_{MV}$ of the global minimum variance portfolio:

   $$\alpha = \mu_{MV}$$

(10)

3. The variance $\sigma^2_{\varepsilon}$ of the noise term $\varepsilon_t$ in Equation (8) corresponds to the variance $\sigma^2_{MV}$ of the global minimum variance portfolio:

   $$\sigma^2_{\varepsilon} = \sigma^2_{MV}$$

(11)

To prove this proposition we define $\beta^{ex}$, $w^{ex}_{MV}$ and $r_t^{ex}$ as column vectors of dimension $N - 1$. These vectors contain the entries $\beta_i$, $w_{MV,i}$ and $r_{t,i}$ with $i = 1, \ldots, N - 1$. The $(N - 1) \times (N - 1)$ matrix $\Omega$ is the covariance matrix of the regressors of Equation (8):

$$\Omega := \text{var} (r_{t,N} \mathbf{1} - r_t^{ex})$$

(12)

The regression coefficients $\beta^{ex}$ are the standardized covariances of the regressors and the dependent variable:

$$\beta^{ex} = \Omega^{-1} \text{cov} (r_{t,N} \mathbf{1} - r_t^{ex}, r_{t,N})$$

(13)

We have to show that the weights $w^{ex}_{MV}$ of the global minimum variance portfolio correspond to the regression coefficients $\beta^{ex}$. The weight $w_{MV,N}$ can then be computed as $1 - (w^{ex}_{MV})' \mathbf{1}$. To prove $\beta^{ex} = w^{ex}_{MV}$ we consider an arbitrary portfolio $P$. Its return is determined by the weight vector $w^{ex}_P = (w_{P,1}, \ldots, w_{P,N-1})'$ and the stock
returns $r_t^{ex}$ and $r_{t,N}$:

$$r_{t,P} = (w_P^{ex})'r_t^{ex} + (1 - (w_P^{ex})')r_{t,N} = r_{t,N} - (w_P^{ex})'(r_{t,N} - r_t^{ex})$$

(14)

The return variance of this arbitrary portfolio $P$

$$\sigma_P^2 = \sigma_N^2 + (w_P^{ex})'\Omega w_P^{ex} - 2(w_P^{ex})'\text{cov} (r_{t,N} - r_t^{ex}, r_{t,N})$$

(15)

is a function of the weights $w_P^{ex}$. To find the weights of the global minimum variance portfolio we minimize (15) with respect to the portfolio weights $w_P^{ex}$. This minimization leads to

$$w_{MV}^{ex} = \Omega^{-1}\text{cov} (r_{t,N} - r_t^{ex}, r_{t,N}).$$

(16)

The weights (16) correspond to the regression coefficients (13). This proves the first statement of Proposition 1. To prove our Statements 2 and 3 we rearrange (8) and use $\beta_i = w_{MV,i}$:

$$\alpha + \varepsilon_t = w_{MV,1}r_{t,1} + \ldots + w_{MV,N-1}r_{t,N-1} + \left(1 - \sum_{i=1}^{N-1} w_{MV,i}\right) r_{t,N}$$

(17)

The right hand side of Equation (17) is the return of the global minimum variance portfolio. Applying the expectation and the variance operator to (17) proves our Statements 2 and 3.

Proposition 1 shows that the traditional approach and the OLS approach lead to identical portfolio weights. However, the result was based on the assumption of known parameters. Next we show that the identity result holds even if we have to estimate the parameters. We define the OLS estimates of the coefficients in Equation (8) as $\hat{\alpha}, \hat{\beta}_1, \ldots, \hat{\beta}_{N-1}$. $\hat{\sigma}_\varepsilon^2 = \frac{1}{T-N} \sum_{t=1}^{T} \varepsilon_t^2$ is the OLS estimate of the variance of $\varepsilon_t$.

**Proposition 2**

1. The traditional weight estimate $\hat{w}_{MV,i}$ equals the OLS estimate:

$$\hat{w}_{MV,i} = \hat{\beta}_i \quad \forall \ i = 1, \ldots, N-1$$

(18)
\[ \hat{w}_{MV,N} = 1 - \sum_{i=1}^{N-1} \hat{\beta}_i \]  

(19)

2. The traditional estimate of the expected return of the global minimum variance portfolio \( \hat{\mu}_{MV} \) equals the OLS estimate:

\[ \hat{\mu}_{MV} = \hat{\alpha} \]  

(20)

3. The traditional estimate of the return variance of the global minimum variance portfolio \( \hat{\sigma}_{MV}^2 \) is a multiple of the OLS estimate of the variance \( \hat{\sigma}_\varepsilon^2 \):

\[ \hat{\sigma}_{MV}^2 = \frac{T - N}{T} \hat{\sigma}_\varepsilon^2 \]  

(21)

First we prove Statement 1. The traditional approach is the solution to the minimization problem

\[ \min_{w_1, \ldots, w_N} \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \hat{\sigma}_{ij}. \]  

(22)

In the OLS approach the regression coefficients are estimated by solving the following minimization problem

\[ \min_{\alpha, \beta_1, \ldots, \beta_{N-1}} \sum_{t=1}^{T} \varepsilon_t^2. \]  

(23)

(23) can be rewritten as

\[ \min_{\alpha, \beta_1, \ldots, \beta_{N-1}} \sum_{t=1}^{T} \left[ -\alpha + \beta_1 r_{t,1} + \ldots + \beta_{N-1} r_{t,N-1} + \left( 1 - \sum_{i=1}^{N-1} \beta_i \right) r_{t,N} \right]^2. \]  

(24)

Since the coefficients \( \beta_i \) correspond to the portfolio weights \( w_i \) (Proposition 1) and since the \( N \) portfolio weights add up to one, we can rearrange Equation (24) as follows:

\[ \min_{\alpha, w_1, \ldots, w_N} \sum_{t=1}^{T} \left[ -\alpha + w_1 r_{t,1} + \ldots + w_N r_{t,N} \right]^2 \quad \text{s.t.} \quad \sum_{i=1}^{N} w_i = 1 \]  

(25)

Differentiating (25) with respect to \( \alpha \) leads to the necessary condition for a minimum:

\[ \alpha = w_1 \hat{\mu}_1 + \ldots + w_N \hat{\mu}_N \]  

(26)
Here $\hat{\mu}_i = \frac{1}{T} \sum_{t=1}^{T} r_{t,i}$ is the estimated mean return of asset $i$. Using (26) we rewrite (25) as

$$\min_{w_1, \ldots, w_N} \sum_{t=1}^{T} [w_1(r_{t,1} - \hat{\mu}_1) + \ldots + w_N(r_{t,N} - \hat{\mu}_N)]^2$$

subject to the condition that the $N$ portfolio weights add up to one. Rearranging the sum in (27) yields another representation of the OLS approach (23):

$$\min_{w_1, \ldots, w_N} T \sum_{i=1}^{N} \sum_{j=1}^{N} \left[ w_i w_j \frac{1}{T} \sum_{t=1}^{T} (r_{t,i} - \hat{\mu}_i)(r_{t,j} - \hat{\mu}_j) \right] = \min_{w_1, \ldots, w_N} T \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \hat{\sigma}_{ij}$$

Thus, the sum of the squared residuals in (23) is equivalent to (28). Since (28) and (22) differ only by the positive factor $T$, both optimization problems produce the same portfolio weights. This proves the first statement of Proposition 2.

Statement 2 can be derived from the necessary condition (26). Replacing $w_i$ by $\hat{w}_{MV,i}$ makes $\hat{\alpha}$ the estimated expected return of the global minimum variance portfolio, which leads to $\hat{\alpha} = \hat{\mu}' \hat{w}_{MV}$. The expression $\hat{\mu}' \hat{w}_{MV}$ equals the traditional estimator $\hat{\mu}_{MV}$.

Statement 3 can be derived accordingly. The sum of the squared residuals (23) equals $T \hat{\sigma}_{MV}^2$. This can be easily seen by rewriting (28) as $T \min_w w' \hat{\Sigma} w$. Its solution $T \hat{w}_{MV}' \hat{\Sigma} \hat{w}_{MV}$ equals $T$ times the estimated variance of the global minimum variance portfolio.

Proposition 2 states that the OLS estimation technique and the traditional approach yield identical estimates of the portfolio weights of the global minimum variance portfolio. Therefore, the estimates of $\hat{\mu}_{MV}$ are identical. The variance estimates differ only by the scalar $(T - N)/T$.

The equivalence of the two estimation approaches allows us to transfer all the distributional results of the OLS approach to the traditional approach. Therefore, we have a powerful yet simple way of deriving the conditional distributions of the estimated weights and return parameters. This is done in Section 4.
4 Conditional Distribution

We estimate the weights of the global minimum variance portfolio using the linear regression (8). We define the $T \times N$ matrix $Z$ which contains the regressors $z_t = (r_{t,N} - r_{t,1}, \ldots, r_{t,N} - r_{t,N-1})'$ of the linear regression (8):

$$Z := \begin{pmatrix}
1 & z'_1 \\
\vdots & \vdots \\
1 & z'_T
\end{pmatrix} = (1 z)$$

The vector $\bar{z} = \frac{1}{T} \sum_{t=1}^{T} z_t$ consists of the arithmetic averages of the regressors.

Proposition 3 gives the conditional distributions of the estimated portfolio weights and return parameters. The information set we condition on consists of the $T \times (N - 1)$ matrix $z$ of return differences.

Proposition 3

1. The OLS estimates of the portfolio weights, $\hat{\beta}^{ex}$, are jointly normally distributed:

$$\hat{\beta}^{ex} | z \sim N \left( w_{MV}^{ex}; \sigma_{MV}^2 (z'z - T\bar{z}\bar{z}')^{-1} \right)$$

2. The OLS estimate of the expected return, $\hat{\alpha}$, is normally distributed:

$$\hat{\alpha} | z \sim N \left( \mu_{MV}; \sigma_{MV}^2 \left( 1/T + \bar{z}'(z'z - T\bar{z}\bar{z}')^{-1}\bar{z} \right) \right)$$

3. Let $\hat{\sigma}_\varepsilon^2$ be the OLS estimate of the variance of the error term $\varepsilon_t$. The following expression is $\chi^2$-distributed:

$$(T - N) \frac{\hat{\sigma}_\varepsilon^2}{\sigma_{MV}^2} \sim \chi^2(T - N)$$

Proposition 3 is based on Proposition 1. The OLS estimator $\hat{B} = (\hat{\alpha}, \hat{\beta}_1, \ldots, \hat{\beta}_{N-1})' = (Z'Z)^{-1}Z'r_N$ with $r_N = (r_{1,N}, \ldots, r_{T,N})'$ is normally distributed:

$$\hat{B} | z \sim N \left( B; \sigma_\varepsilon^2(Z'Z)^{-1} \right).$$
\( B = (\alpha, \beta_1, \ldots, \beta_{N-1})' \) is the parameter vector. From (33) we see directly that the expectations of the conditional estimators \( \hat{\beta}^{\text{ex}} \) and \( \hat{\alpha} \) are \( \beta^{\text{ex}} \) and \( \alpha \). According to Proposition 1, the variance \( \sigma^2_\varepsilon \) is equal to the variance of the global minimum variance portfolio \( \sigma^2_{\text{MV}} \). Using (29) we partition the matrix \( Z'Z \):

\[
Z'Z = \begin{pmatrix} T & T\bar{z}' \\ T\bar{z} & z'z \end{pmatrix}
\]  

(34)

The inversion of the matrix \( Z'Z \) yields:\(^7\)

\[
(Z'Z)^{-1} = \begin{pmatrix} 1/T + \bar{z}'(z'z - T\bar{z}\bar{z}')^{-1}\bar{z} & \bar{z}'(z'z - T\bar{z}\bar{z}')^{-1} \\ (z'z - T\bar{z}\bar{z}')^{-1}\bar{z} & (z'z - T\bar{z}\bar{z}')^{-1} \end{pmatrix}
\]  

(35)

\( \sigma^2_{\text{MV}} \) times the upper left element of the right hand side of (35) is the conditional variance of \( \hat{\alpha} \). \( \sigma^2_{\text{MV}} \) times the lower right element is the conditional covariance matrix of \( \hat{\beta}^{\text{ex}} \).

Proposition 3 states the core results of this paper. It allows us to calculate the estimation risk involved in estimating the global minimum variance portfolio (Section 6) and to carry out statistical tests concerning the estimated weights and return parameters (Section 7).

5 Non-Normal Returns

Throughout the paper we assumed that stock returns are normally distributed and IID. However, our OLS approach can also be used for non-normal returns. Instead of restricting ourselves to the multivariate normal distribution, we now consider the broader class of elliptical distributions. Among others, the class of elliptical distributions comprises the normal distribution and the Student-t-distribution. We choose this class of distributions for two reasons. Firstly, elliptical distributions support mean variance analysis since they fulfill the two requirements: (i) Elliptical distributions can be entirely characterized by their mean and variance and (ii) linear combinations of elliptically distributed random variables are again elliptically distributed. Secondly, elliptical distributions can describe empirical features of stock returns, especially the heavy tails of stock return distributions.
If we no longer assume normally distributed returns, but elliptically distributed returns, the noise term $\varepsilon_t$ in Equation (8) will remain uncorrelated of the regressors $z_t$. However, the error term will not necessarily be independent of the regressors. For instance, the correlation of the squared noise term $\varepsilon_t^2$ and the squared regressors $z_{t,i}^2$ may be different from zero. This means that the standard assumptions of the linear regression model are no longer fulfilled because the noise terms $\varepsilon_t$ are heteroskedastic. The variance of $\varepsilon_t$ varies in a systematic way. Nevertheless, we can apply the OLS methodology. Propositions 1 and 2 remain unaltered because their proofs do not depend on the normality assumption, but Proposition 3 has to be modified: The OLS estimates remain unbiased and consistent but the estimates $\hat{\alpha}$ and $\hat{\beta}$ are only asymptotically normally distributed. In addition, the estimated covariance matrix needs to be modified. To get correct standard errors in the regression, one has to adjust the covariance matrix e.g. by using the White (1980)-correction.

6 Estimation Risk

We now apply our results to calculate the extent of the estimation risk. We again assume the best situation an investor can face: the returns are normal and IID. The estimation risk is the additional out-of-sample return variance due to errors in the estimated portfolio weights. In our Propositions 4 and 5 we calculate the conditional and unconditional estimation risk, respectively. In Proposition 6 we prove that the traditional weight estimator $\hat{w}_{MV}$ leads to the lowest estimation risk of all unbiased estimators.

We consider an investor who uses $T$ return observations $r_1, \ldots, r_T$ to estimate $\hat{w}_{MV}$. Using the estimates $\hat{w}_{MV}$, the investor invests his funds for the period to follow. This strategy yields the out-of-sample return $\hat{r}_{T+1,MV} = \hat{w}'_{MV}r_{T+1}$. Its risk is $\text{var}(\hat{r}_{T+1,MV}|r_1, \ldots, r_T)$ which depends on the realizations of the stock returns from $t = 1$ to $t = T$.

Proposition 4

If the portfolio weights are estimated traditionally, then the conditional out-of-
sample return variance is given by

$$\text{var}(\hat{r}_{T+1, MV} | r_1, \ldots, r_T) = \sigma_{MV}^2 + \tilde{R}(\hat{w}_{MV})$$  \hspace{1cm} (36)$$

with

$$\tilde{R}(\hat{w}_{MV}) = (\hat{w}_{MV} - w_{MV})' \Sigma (\hat{w}_{MV} - w_{MV}).$$ \hspace{1cm} (37)$$

Proposition 4 (proved in Appendix 1) shows that the risk depends on two components. The first component, $\sigma_{MV}^2$, is the innovation risk, i.e. the risk due to the randomness of stock returns. The second component, $\tilde{R}(\hat{w}_{MV})$, is the estimation risk. If the investors knew all return distribution parameters, they would choose (5) as their weights when selecting the global minimum variance portfolio. In such a case there is no estimation risk and (36) reduces to (7). However, since the investor does not know the distribution parameters and has to estimate them instead, his estimated portfolio weights, $\hat{w}_{MV}$, differ from the true ones, $w_{MV}$. This difference leads to the conditional estimation risk $\tilde{R}(\hat{w}_{MV})$. Note that the $\tilde{R}(\hat{w}_{MV})$ is a random variable which takes on only positive values. The more the estimated weights differ from the true ones, the larger $\tilde{R}(\hat{w}_{MV})$ is. The unconditional estimation risk is obtained by applying the expectation operator to var($\hat{r}_{T+1, MV} | r_1, \ldots, r_T$).

**Proposition 5**

*If the portfolio weights are estimated traditionally, then the unconditional out-of-sample return variance is given by*

$$E(\text{var}(\hat{r}_{T+1, MV} | r_1, \ldots, r_T)) = \sigma_{MV}^2 + \bar{R}(\hat{w}_{MV})$$  \hspace{1cm} (38)$$

with

$$\bar{R}(\hat{w}_{MV}) = \sigma_{MV}^2 \frac{N - 1}{T - N - 1}. \hspace{1cm} (39)$$

According to this proposition (proved in Appendix 2) the larger the innovation risk $\sigma_{MV}^2$, the larger the investment universe $N$ and the shorter the estimation period $T$ are, the higher is the unconditional estimation risk $\bar{R}(w_{MV})$.

Proposition 6 proves that the estimation risk cannot be reduced by choosing another unbiased weight estimator. The traditional weight estimator is the best unbiased estimator.
Proposition 6

The traditional weight estimator $\hat{w}_{MV}$ as given in Equation (18) has the lowest unconditional estimation risk $\bar{R}(\cdot)$ of all unbiased weight estimators $\tilde{w}_{MV}$:

$$\bar{R}(\hat{w}_{MV}) \leq \bar{R}(\tilde{w}_{MV}).$$

(40)

This proposition follows from the properties of OLS estimators. In the case of normally distributed error terms, the OLS estimator is the best unbiased weight estimator. According to Proposition 2 this statement is true for the traditional estimator, too. In Appendix 3 we show that this property implies the lowest estimation risk possible.

7 Statistical Inference

In this section we use our results to address problems in international asset allocation. We conduct an empirical study based on international stock data. Our data set consists of monthly MSCI total return indices of the G7 countries Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States. These countries cover the major currency regions (Dollar, Euro, Pound, Yen). All indices are calculated in Euro, i.e. we take the view of an German investor. The data set covers the period from January 1984 to December 2003. We choose the return of the German index as the dependent variable $r_{t,N}$ in the regression (8). We run the regression and obtain estimates of the portfolio weights of the global minimum variance portfolio. In Table 2 we report the weight estimates $\hat{w}_{MV,i}$, their standard errors and the $t$–statistics. As the stock returns show excess kurtosis we applied the White (1980)-correction to calculate standard errors.

Table 2 shows that the UK market has the highest weight in the international global minimum variance portfolio, followed by Japan and USA. Only the weights for the indices of Japan, the UK and USA are significantly different from zero at the 10% level. This suggests that a German investor who only holds German stocks should add American, Japanese and British stocks to his domestic holdings.

To test whether a German investor can exclude several countries from his port-
Table 2: Weight Estimates of the Global Minimum Variance Portfolio

<table>
<thead>
<tr>
<th>Country (i)</th>
<th>Weight $\hat{w}_{MV,i}$</th>
<th>Standard Error</th>
<th>$t-$Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada (Can)</td>
<td>0.0146</td>
<td>0.0998</td>
<td>0.1467</td>
</tr>
<tr>
<td>France (Fra)</td>
<td>0.0756</td>
<td>0.0772</td>
<td>0.9799</td>
</tr>
<tr>
<td>Germany (Ger)</td>
<td>0.1418</td>
<td>0.0881</td>
<td>1.6088</td>
</tr>
<tr>
<td>Italy (Ita)</td>
<td>0.0427</td>
<td>0.0512</td>
<td>0.8336</td>
</tr>
<tr>
<td>Japan (Jap)</td>
<td>0.1909</td>
<td>0.0570</td>
<td>3.3512</td>
</tr>
<tr>
<td>United Kingdom (UK)</td>
<td>0.3536</td>
<td>0.0946</td>
<td>3.7362</td>
</tr>
<tr>
<td>United States (USA)</td>
<td>0.1807</td>
<td>0.1067</td>
<td>1.6947</td>
</tr>
</tbody>
</table>

folio without increasing the risk of his portfolio, we apply the $F-$test as shown in Appendix 4. The $F-$test allows to test several linear restrictions concerning the portfolio weights simultaneously. First we want to know whether a German investor can reduce his portfolio risk by diversifying internationally. We test the hypothesis:

$H_{0,1}$: International diversification does not pay for German investors, i.e. $w_{MV,Can} = w_{MV,Fra} = w_{MV,USA} = w_{MV,Ita} = w_{MV,UK} = w_{MV,Jap} = 0$. The null hypothesis is rejected at the 1%-level ($F(6, 233)-$statistic $= 22.45$). Thus, it pays for a German investor to diversify internationally. Whether adopting a naive diversification strategy or diversifying optimally makes a difference is analyzed next.

$H_{0,2}$: Naive diversification ($w_{MV,i} = 1/7 \forall i$) offers the same risk diversification effect as optimal diversification. $H_{0,2}$ is rejected at the 10% level ($F(6, 233)-$statistic $= 2.06$). We conclude that a German investor is better off choosing the weights according to (5) than by investing equally in all countries. Finally, we want to know whether investing in only one country per currency region reduces the diversification effect significantly. The
countries invested in are Germany (Euro), Japan (Yen), the UK (Pound) and the United States (Dollar).

\[ H_{0,3}: \text{Investing in one country per currency region } (w_{MV,Can} = w_{MV,Fra} = w_{MV,Ita} = 0) \text{ offers the same risk diversification as investing in all countries.} \]

We cannot reject \( H_{0,3} \) (\( F(3, 233) - \text{statistic} = 0.58 \)). The results suggest that covering the major currency regions by choosing only one country for each currency region provides sufficient diversification.

The three hypotheses tested above serve as examples of how to use the results of this paper. Obviously, one can easily find other hypotheses to test with our approach.

8 Conclusion

In this paper we show that the weights of the global minimum variance portfolio are equal to regression coefficients. This allows us to transfer the entire OLS methodology to the estimation of the weights and return parameters of the global minimum variance portfolio. From the OLS methodology we derive the conditional distributions of the estimated portfolio weights and estimated return parameters.

We discuss two applications of our distributional results. The first application is to assess the extent of the estimation risk involved in estimating the global minimum variance portfolio. Our second application is to test important hypotheses in international asset management. These two applications serve as an illustration of the usefulness of our approach.
Appendix 1

Using \( \hat{w}_{MV} = w_{MV} + (\hat{w}_{MV} - w_{MV}) \) we rewrite the conditional out-of-sample return variance as

\[
\text{var}(\hat{r}_{T+1,MV} | r_1, \ldots, r_T) = \hat{w}_{MV}' \Sigma \hat{w}_{MV}
\]

\[
= \sigma^2_{MV} + (\hat{w}_{MV} - w_{MV})' \Sigma (\hat{w}_{MV} - w_{MV})
\]

\[
+ 2w_{MV}' \Sigma (\hat{w}_{MV} - w_{MV}).
\]  

(41)

The last term in (41) can be rewritten as

\[
2(w_{MV}' \Sigma \hat{w}_{MV} - w_{MV}' \Sigma w_{MV}).
\]  

(42)

The first term is the return covariance of a portfolio with the portfolio weights \( \hat{w}_{MV} \) and the global minimum variance portfolio \( w_{MV} \). The second term is the return variance of the global minimum variance portfolio. Huang and Litzenberger (1988), p. 68, prove that the return covariance of an arbitrary stock portfolio and the global minimum variance portfolio is equal to the return variance of the global minimum variance portfolio. Therefore, the last term in (41) drops out. This completes the proof of Proposition 4.
Appendix 2

In this appendix we prove Proposition 5. In Lemma 1 we show how to express the unconditional estimation risk $\bar{R}(\cdot)$ of any unbiased weight estimator $\hat{w}_{MV}$ as a function of the estimator’s unconditional variance $\text{var}(\hat{w}_{MV}^{ex})$. In Lemma 2 we compute the unconditional variance of a specific unbiased weight estimator, the traditional weight estimator. Combining these two lemmata, we obtain the expression for the estimation risk $\bar{R}(\hat{w}_{MV})$ as stated in Proposition 5.

**Lemma 1**

Let $\hat{w}_{MV}$ be any unbiased weight estimate. Then the unconditional out-of-sample return variance is

$$E(\text{var}(\hat{r}_{T+1,MV}|r_1, \ldots, r_T)) = \sigma_{MV}^2 + \bar{R}(\hat{w}_{MV})$$

(43)

with

$$\bar{R}(\hat{w}_{MV}) = \text{tr}[\text{var}(\hat{w}_{MV}^{ex})]\Omega].$$

(44)

Proof of Lemma 1: Using (14) we can rewrite the out-of-sample return as

$$\hat{r}_{T+1,MV} = r_{T+1,N} - (\hat{w}_{MV}^{ex})'(r_{T+1,N1} - r_{T+1}^{ex}).$$

(45)

The unconditional out-of-sample variance is

$$E(\text{var}(\hat{r}_{T+1,MV}|r_1, \ldots, r_T)) = \sigma_{MV}^2 + E((\hat{w}_{MV}^{ex})'(\hat{w}_{MV}^{ex}) - 2E(\hat{w}_{MV}^{ex})'\text{cov}(r_{T+1,N1} - r_{T+1}^{ex}, r_{T+1,N}).$$

(46)

Setting $E(\hat{w}_{MV}^{ex}) = w_{MV}^{ex} + E(\hat{w}_{MV}^{ex} - w_{MV}^{ex})$ we rewrite the expression $E((\hat{w}_{MV}^{ex})'\Omega\hat{w}_{MV}^{ex})$ as

$$E((\hat{w}_{MV}^{ex})'\Omega\hat{w}_{MV}^{ex}) = (w_{MV}^{ex})'\Omega w_{MV}^{ex} + E((\hat{w}_{MV}^{ex} - w_{MV}^{ex})'(\hat{w}_{MV}^{ex} - w_{MV}^{ex})) + 2E(\hat{w}_{MV}^{ex} - w_{MV}^{ex})'\Omega w_{MV}.$$  

(47)

Inserting (47) in (46) and using

$$\sigma_{MV}^2 = \sigma_{N}^2 + (w_{MV}^{ex})'\Omega w_{MV}^{ex} - 2(w_{MV}^{ex})'\text{cov}(r_{T+1,N1} - r_{T+1}^{ex}, r_{T+1,N})$$

(48)
we get
\begin{equation}
E(\text{var}(\hat{\beta}_{T+1,MV} | r_1, \ldots, r_T)) = \sigma_{MV}^2 + E((\hat{\beta}_{MV}^{ex} - w_{MV}^{ex})' \Omega (\hat{\beta}_{MV}^{ex} - w_{MV}^{ex})) . \tag{49}
\end{equation}

Finally we deal with the expression \( E((\hat{\beta}_{MV}^{ex} - w_{MV}^{ex})' \Omega (\hat{\beta}_{MV}^{ex} - w_{MV}^{ex})) \).

\begin{align*}
E((\hat{\beta}_{MV}^{ex} - w_{MV}^{ex})' \Omega (\hat{\beta}_{MV}^{ex} - w_{MV}^{ex})) &= E(\text{tr} ((\hat{\beta}_{MV}^{ex} - w_{MV}^{ex})' \Omega (\hat{\beta}_{MV}^{ex} - w_{MV}^{ex}))) \\
&= E(\text{tr} ((\hat{\beta}_{MV}^{ex} - w_{MV}^{ex}) (\hat{\beta}_{MV}^{ex} - w_{MV}^{ex})' \Omega)) \\
&= \text{tr} (E((\hat{\beta}_{MV}^{ex} - w_{MV}^{ex}) (\hat{\beta}_{MV}^{ex} - w_{MV}^{ex})')) \Omega) \\
&= \text{tr} (\text{var}(\hat{\beta}_{MV}^{ex}) \Omega) \tag{50}
\end{align*}

Lemma 1 results directly from (49) in combination with (50).

The estimation risk given by (44) depends on the estimator’s variance \( \text{var}(\hat{\beta}_{MV}^{ex}) \).

For the traditional estimator we can state this variance explicitly. This is done in Lemma 2.

**Lemma 2**

The unconditional variance of the traditional weight estimator \( \hat{\beta}_{MV}^{ex} \) is

\begin{equation}
\text{var}(\hat{\beta}_{MV}^{ex}) = \sigma_{MV}^2 \frac{1}{T - N - 1} \Omega^{-1}. \tag{51}
\end{equation}

Proof of Lemma 2: From the first statement of Proposition 2 in connection with the first statement of Proposition 3 we get the conditional variance:

\begin{equation}
\text{var}(\hat{\beta}_{MV}^{ex} | z) = \sigma_{MV}^2 (z'z - T \bar{z} \bar{z}')^{-1} \tag{52}
\end{equation}

The variance decomposition theorem provides the relation between the unconditional and conditional variance:

\begin{equation}
\text{var}(\hat{\beta}_{MV}^{ex}) = E(\text{var}(\hat{\beta}_{MV}^{ex} | z)) + \text{var}(E(\hat{\beta}_{MV}^{ex} | z)) \tag{53}
\end{equation}

As the estimator \( \hat{\beta}_{MV}^{ex} \) is unbiased, the second term on the right hand side of (53) is zero. Therefore, it remains to determine the expectation of \( (z'z - T \bar{z} \bar{z}')^{-1} \). The matrix \( (z'z - T \bar{z} \bar{z}') \) is Wishart distributed, which follows from the assumption of
normally distributed returns:

\[
z'z - T\bar{z}\bar{z}' = \sum_{t=1}^{T} (z_t - \bar{z})(z_t - \bar{z})' \sim W(\Omega, T - 1, N - 1) \tag{54}
\]

The expectation of a random matrix whose inverse is Wishart distributed is shown in Press (1972), p. 112:

\[
E \left( (z'z - T\bar{z}\bar{z}')^{-1} \right) = \frac{1}{T - N - 1} \Omega^{-1} \tag{55}
\]

Lemma 2 follows immediately from (55).

Inserting (51) into (44) yields (39). This completes the Proof of Proposition 5.
Appendix 3

Based on (43) of Lemma 1 we can state the difference in the unconditional estimation risk between using an arbitrary unbiased weight estimator $\hat{w}_{MV}$ and using the traditional estimator $\tilde{w}_{MV}$, respectively:

$$\bar{R}(\hat{w}_{MV}) - \bar{R}(\tilde{w}_{MV}) = \text{tr}[\text{var}(\hat{w}_{MV}^{ex})\Omega] - \text{tr}[\text{var}(\tilde{w}_{MV}^{ex})\Omega]$$  \hspace{1cm} (56)

$$= \text{tr}[\Delta\Omega]$$  \hspace{1cm} (57)

with

$$\Delta = \text{var}(\hat{w}_{MV}^{ex}) - \text{var}(\tilde{w}_{MV}^{ex})$$  \hspace{1cm} (58)

As $\hat{w}_{MV}^{ex}$ is the best unbiased estimator, the difference matrix $\Delta$ is at least positive semi-definite. Since the trace of the matrix product of two semi-definite matrices is never negative, the expression $\text{tr}[\Delta\Omega]$ in (57) is not negative, either.\textsuperscript{15} Therefore, there is no unbiased weight estimator with lower unconditional estimation risk than that of the traditional estimator.\textsuperscript{16}
Appendix 4

In this appendix we explicitly give the test statistics used in Section 7. In the case of non-normally distributed returns these statistics are only asymptotically exact and the estimated covariance matrix has to be adjusted as pointed out in Section 5. Let \( q = (q_1, \ldots, q_{N-1})' \) be an arbitrary non-stochastic vector. Then the following statistic is \( t \)-distributed:

\[
\frac{q' \hat{w}_{MV}^{\varepsilon z} - q' \hat{w}_{MV}^{\varepsilon z}}{\sqrt{\hat{\sigma}^2 q'(z'z - \bar{z}\bar{z})^{-1} q}} \sim t(T - N) \tag{59}
\]

Since the estimated weight of asset \( N \) is a linear combination of the other weights, i.e. \( \hat{w}_{MV,N} = 1 - 1' \hat{w}_{MV}^{\varepsilon z} \), we can derive the distribution of \( \hat{w}_{MV,N} \) from (59) by setting \( q = 1 \):

\[
\frac{\hat{w}_{MV,N} - w_{MV,N}}{\sqrt{\hat{\sigma}^2 1'(z'z - \bar{z}\bar{z})^{-1} 1}} \sim t(T - N) \tag{60}
\]

In the third column of Table 2 we report the \( t \)-statistic as computed by (59) for the weights \( i = Can, Fra, Ger, Ita, Jap, Uk \) and by (60) for the weight \( i = US \).

Let \( SSR \) and \( SSR_R \) be the sum of the squared residuals in the unrestricted and restricted regression. Let \( m \leq N - 1 \) be the number of linear independent restrictions. Then the following statistic is \( F \)-distributed:

\[
F = \frac{T - N}{m} \left( \frac{SSR_R}{SSR} - 1 \right) \sim F(m, T - N) \tag{61}
\]

This statistic is calculated for the hypotheses \( H_{0,1} \) to \( H_{0,3} \).
References


Notes

1In this paper we deal only with estimation risk resulting from unknown return distribution parameters. In the more general definition of Bawa, Brown and Klein (1979), estimation risk also includes situations in which not only the parameters, but also the functional form of the distribution is unknown.

2For example, using $\tau = 10$ years of monthly ($n = 12$) data provides us with $T = 120$ observations.

3See Merton (1980). The distribution of the estimated return standard deviation $\hat{\sigma}_t$ is only known asymptotically.

4Note that the error term $\varepsilon_t$ is by construction uncorrelated with all the return differences $r_{t,N} - r_{t,i}$. The absence of correlation allows us to apply the OLS estimation technique.

5We will show in Section 5 that one can apply the OLS approach even when the returns are not normal and the error term does not satisfy the standard assumption of linear regression.

6The superscript $ex$ indicates that the vector has no entry for asset $N$.

7See Greene (2000), p. 34.


9In case, there is not only heteroscedasticity, but also autocorrelation in the data, one hast to use the correction of Newey and West (1987) instead.

10Without the assumption of normality, Proposition 4 remains unchanged. Proposition 6 is weakened: the traditional estimator is only the best linear unbiased estimator. Regarding Proposition 5: Without the normality assumption, we cannot determine the unconditional estimation risk because this determination requires the expectation of a quadratic form which is only known for normally distributed random variables.
This result proves the claim of Jagannathan and Ma (2003).

See Appendix 4 for the exact formula of the test statistic.

Our test is a simplified version of a spanning test. The spanning tests suggested in the literature (see, e.g., Kan and Zhou (2001)) test whether the inclusion of an additional asset changes the minimum variance frontier. Our test focuses not on the whole frontier, but solely on one portfolio of the frontier, the global minimum variance portfolio. If we find a significant change in the global minimum variance portfolio we know that the minimum variance frontier has changed as well. Thus, our test is a sufficient test for spanning. Since the global minimum variance portfolio does not depend on expected returns, our test has a higher power than traditional spanning tests.

Jorion (1985) develops an alternative test to address this question. He uses a maximum likelihood test to check whether a given portfolio is significantly different from the global minimum variance portfolio. While the distribution of the Jorion (1985) test is known only asymptotically, the distribution of our test is known even in small samples.


If we give up the assumption of normality, the traditional estimator is the best linear unbiased estimator. For the Gauss-Markov-Theorem see, e.g., Hayashi (2000), p. 27-29.
CFR Working Paper Series

CFR Working Papers are available for download from [www.cfr-cologne.de](http://www.cfr-cologne.de).

Hardcopies can be ordered from: Centre for Financial Research (CFR), Albertus Magnus Platz, 50923 Koeln, Germany.

### 2004

<table>
<thead>
<tr>
<th>No.</th>
<th>Author(s)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>04-01</td>
<td>I. Chowdhury, M. Hoffmann, A. Schabert</td>
<td>Inflation Dynamics and the Cost Channel of Monetary Transmission</td>
</tr>
<tr>
<td>04-02</td>
<td>A. Kempf, S. Ruenzi</td>
<td>Tournaments in Mutual Fund Families</td>
</tr>
<tr>
<td>04-03</td>
<td>V. Agarwal, W.H. Fung, N.Y. Naik</td>
<td>Risks in Hedge Fund Strategies: Case of Convertible Arbitrage</td>
</tr>
<tr>
<td>04-04</td>
<td>V. Agarwal, N.D. Daniel, N.Y. Naik</td>
<td>Flows, Performance, and Managerial Incentives in Hedge Funds</td>
</tr>
<tr>
<td>04-07</td>
<td>J.J. Merrick, Jr., N.Y. Naik, P.K. Yadav</td>
<td>Strategic Trading Behavior and Price Distortion in a Manipulated Market: Anatomy of a Squeeze</td>
</tr>
<tr>
<td>04-08</td>
<td>N.F. Carline, S.C. Linn, P.K. Yadav</td>
<td>Can the Stock Market Systematically make Use of Firm- and Deal-Specific Factors when Initially Capitalizing the Real Gains from Mergers and Acquisitions</td>
</tr>
<tr>
<td>04-09</td>
<td>A. Kempf, K. Kreuzberg</td>
<td>Portfolio Disclosure, Portfolio Selection and Mutual Fund Performance Evaluation</td>
</tr>
<tr>
<td>04-10</td>
<td>N. Hautsch, D. Hess</td>
<td>Bayesian Learning in Financial Markets – Testing for the Relevance of Information Precision in Price Discovery</td>
</tr>
</tbody>
</table>

### 2005

<table>
<thead>
<tr>
<th>No.</th>
<th>Author(s)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>05-01</td>
<td>S. Frey, J. Grammig</td>
<td>Liquidity supply and adverse selection in a pure limit order book market</td>
</tr>
<tr>
<td>05-02</td>
<td>A. Kempf, C. Memmel</td>
<td>On the Estimation of the Global Minimum Variance Portfolio</td>
</tr>
</tbody>
</table>