DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 1/07

Linear Statistical Inference for Global and Local Minimum Variance Portfolios

by

Gabriel Frahm

September 26th, 2008

DISKUSSIONSBEITRÄGE ZUR STATISTIK UND ÖKONOMETRIE

SEMINAR FÜR WIRTSCHAFTS- UND SOZIALSTATISTIK
UNIVERSITÄT ZU KÖLN

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

This page intentionally left blank.

DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 1/07

Linear Statistical Inference for Global and Local Minimum Variance Portfolios

by

Gabriel Frahm

September 26th, 2008

Abstract

Traditional portfolio optimization has often been criticized for not taking estimation risk into account. Estimation risk is mainly driven by the parameter uncertainty regarding the expected asset returns rather than their variances and covariances. The global minimum variance portfolio has been advocated by many authors as an appropriate alternative to the tangential portfolio. This is because there are no expectations which have to be estimated and thus the impact of estimation errors can be substantially reduced. However, in many practical situations an investor is not willing to choose the global minimum variance portfolio but he wants to minimize the variance of the portfolio return under specific constraints for the portfolio weights. Such a portfolio is called *local minimum variance portfolio*. Small-sample hypothesis tests for global and local minimum variance portfolios are derived and the exact distributions of the estimated portfolio weights are calculated in the present work. The first two moments of the estimator for the expected portfolio returns are also provided and the presented instruments are illustrated by an empirical study.

Keywords: Estimation risk, Linear regression theory, Markowitz portfolio, Minimum variance portfolio, Portfolio optimization, Top down investment.

2000 MSC: 62F03, 91B28.

JEL Classification: C13, G11.

This page intentionally left blank.

LINEAR STATISTICAL INFERENCE FOR GLOBAL AND LOCAL MINIMUM VARIANCE PORTFOLIOS

GABRIEL FRAHM

UNIVERSITY OF COLOGNE
CHAIR FOR STATISTICS & ECONOMETRICS
DEPARTMENT OF ECONOMIC AND SOCIAL STATISTICS
ALBERTUS-MAGNUS-PLATZ, D-50923 COLOGNE, GERMANY

Abstract. Traditional portfolio optimization has often been criticized for not taking estimation risk into account. Estimation risk is mainly driven by the parameter uncertainty regarding the expected asset returns rather than their variances and covariances. The global minimum variance portfolio has been advocated by many authors as an appropriate alternative to the tangential portfolio. This is because there are no expectations which have to be estimated and thus the impact of estimation errors can be substantially reduced. However, in many practical situations an investor is not willing to choose the global minimum variance portfolio but he wants to minimize the variance of the portfolio return under specific constraints for the portfolio weights. Such a portfolio is called *local minimum variance portfolio*. Small-sample hypothesis tests for global and local minimum variance portfolios are derived and the exact distributions of the estimated portfolio weights are calculated in the present work. The first two moments of the estimator for the expected portfolio returns are also provided and the presented instruments are illustrated by an empirical study.

Keywords: Estimation risk, Linear regression theory, Markowitz portfolio, Minimum variance portfolio, Portfolio optimization, Top down investment.

2000 MSC: 62F03, 91B28.

JEL Classification: C13, G11.

1. Motivation

During the past decades traditional portfolio optimization has often been criticized since it does not account for estimation risk (Jorion, 1986, Kalymon, 1971, Klein and Bawa, 1976, Michaud, 1989). At the beginning of modern portfolio theory (Markowitz, 1952) it was usually supposed that the parameters of interest, i.e. the means and (co-)variances of asset returns can be estimated accurately such that estimation errors remain negligible. Although this conjecture might be true for variances and covariances if the sample size is large enough compared to the number of assets, it is not an appropriate simplification for expected asset returns in most practical situations (Chopra and Ziemba, 1993, Kempf and Memmel, 2002, Merton, 1980). Nowadays many portfolio optimization procedures which take the parameter uncertainty into account can be found in the literature (Black and Litterman, 1992, Frost and Savarino, 1986, Herold and Maurer, 2006, Kan and Zhou, 2007, Scherer, 2004).

Consider a d-dimensional random vector $R = (R_1, \ldots, R_d)$ of asset excess returns at the end of some investment horizon. The excess return of an asset corresponds to the asset return minus the risk-free interest rate and in the following I will usually

Email: frahm@statistik.uni-koeln.de, phone: +49 221 470-4267.

Published online first 2008 in *Statistical Papers*. The original publication is available at http://www.springerlink.com.

drop the prefix 'excess' for convenience. It is assumed that the vector of asset returns is multivariate normally distributed, i.e. $R \sim \mathcal{N}_d(\mu, \Sigma)$, where μ $(d \times 1)$ is an unknown vector of expected asset returns and Σ $(d \times d)$ is an unknown positive-definite matrix containing their variances and covariances.

The *tangential portfolio* (TP) is defined as the portfolio of risky assets which maximizes the Sharpe ratio (see Figure 1), i.e.

$$\underset{(d\times 1)}{w_{\mathrm{TP}}} := \arg\max_{v} \, \mu' v / \sqrt{v'\Sigma\, v}$$

such that the budget constraint $1'v = 1$ is satisfied. Here $v = (v_1, \ldots, v_d)$ symbolizes a vector of portfolio weights and 1 is a vector of ones or the one scalar, respectively. In the following '$(x_1, \ldots, x_d)$' indicates a d-tuple which is understood to be a d-dimensional column vector, whereas '$[\; x_1 \;\; \cdots \;\; x_d \;]$' (without the commas) is a d-dimensional row vector, i.e. $(x_1, \ldots, x_d) \equiv [\; x_1 \;\; \cdots \;\; x_d \;]'$.

An (mean-variance) *efficient portfolio* (EP) can be characterized in terms of the typical mean-variance utility function (or, more precisely, certainty equivalent), i.e.

$$\underset{(d\times 1)}{w_{\mathrm{EP}}} := \arg\max_{v} \big(\mu' v - \alpha/2 \cdot v'\Sigma\, v\big)$$

for some risk-aversion parameter $\alpha > 0$. If the EP satisfies the budget constraint, it can be found on the efficient frontier, i.e. the upper part of the hyperbola given in Figure 1. Otherwise it is located on the capital market line.

A rather simple alternative to the TP or some other EP is given by the so-called *global minimum variance portfolio* (GMVP). This is defined as

$$\underset{(d\times 1)}{w} := \arg\min_{v} \, v'\Sigma\, v$$

under the budget constraint $1'v = 1$. The GMVP can be viewed as an EP after setting $\alpha = \infty$. Any portfolio which minimizes the variance of the portfolio return $R'v$ under some *additional* constraints for the portfolio weights will be called *local minimum variance portfolio* (LMVP).

It is well-known that $w_{\mathrm{TP}} = \Sigma^{-1}\mu/(1'\Sigma^{-1}\mu)$ and $w = \Sigma^{-1}1/(1'\Sigma^{-1}1)$ (a closed-form expression for the LMVP under a set of linear equality constraints for the portfolio weights can be found in Section 3.1). The TP strongly depends on the vector μ of expected asset returns and the same holds for the EP if the investor has a relatively low risk aversion (that means if α is small). In contrast, the GMVP as well as any LMVP is not determined by the unknown parameter μ. However, a LMVP in general will be *inefficient* which is shown by Figure 1.

The GMVP has been advocated by many authors (Jagannathan and Ma, 2003, Kempf and Memmel, 2006, Ledoit and Wolf, 2003). On the one hand choosing the GMVP is closely related to the basic idea of Markowitz (1952), i.e. searching for an efficient portfolio by diversification. On the other hand there are no expected asset returns which have to be estimated and so the impact of estimation errors can be substantially reduced. However, one might ask why it should be appropriate to search for a minimum variance portfolio if the investor is interested in maximizing a mean-variance utility function or the Sharpe ratio according to Tobin's two-fund separation theorem (Tobin, 1958). Thus I would like to explain now the main idea of the present work.

The *suggested* TP can differ substantially from the true one in the presence of estimation risk. Put another way, its *realized* (but not the suggested) Sharpe ratio can be very small since the expected asset returns are unknown and then it might be better to search for some minimum variance portfolio. In particular, the constraints for a LMVP can be chosen in such a way that large volatility assets are preferred (recall that the variances and covariances of asset returns can be much better estimated than their expectations). If some branch contains a larger *risk premium* than another (e.g. the IT sector bears more risk than the finance sector),

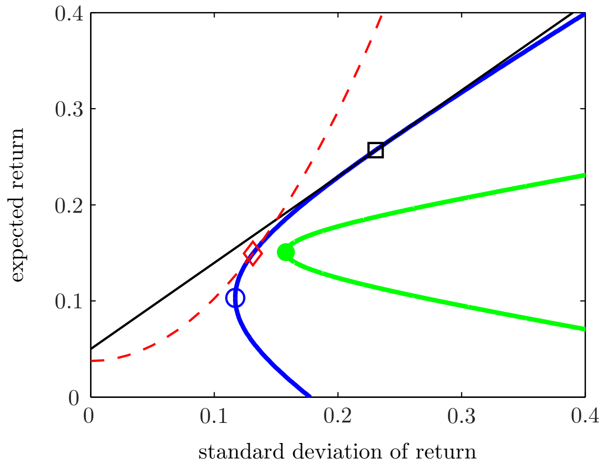

FIGURE 1. Capital market line (straight), utility isoline (dashed), TP (□), EP (◊), GMVP (○), and LMVP (●).

an investor could be simply willing to reap the profit by choosing the corresponding LMVP. Now this is probably closer to the TP or another EP than the GMVP, although the LMVP is inefficient (see Figure 1).

Since there are no expected asset returns which have to be estimated for the LMVP, its realized Sharpe ratio is hopefully larger than the realized Sharpe ratio of the suggested TP. In fact some authors argue that even if portfolio restrictions are binding (which is indicated e.g. by the small hyperbola in Figure 1) they can increase the *out-of-sample* performance (Frost and Savarino, 1988, Jagannathan and Ma, 2003). This is because restricting portfolio weights forces diversification and the investor's decision becomes less vulnerable to estimation risk. Hence, the advertising motto for minimum variance portfolios could be 'A bird in the hand is worth two in the bush'.

Another argument for restricting portfolio weights is that people might have *prior knowledge* apart from empirical data. For instance, investors often believe that some industry sector, region or stock market will 'outperform' another and so they might wish to take the opportunity. Moreover, in many practical situations an investor *must not* choose a mean-variance efficient portfolio. For example, portfolio managers of mutual funds often have to observe certain limits regarding their choice of portfolio weights. This is a typical situation in *top down portfolio management*. That means the set of available assets is divided into some subsets of assets, each subset is divided into some further subsets, etc. These subsets are generally referred to as *asset classes*, according to some industry sector, rating or regional classification. Now, top down portfolio management means that the amount of capital is allocated to the top level partition at first. Given the portfolio weights for that partition, somebody has to choose some optimal portfolio weights for the subsequent asset classes, etc., so that each of the succeeding decisions are limited by the preceding allocations.

As already pointed out by Black and Litterman (1992) as well as Herold and Maurer (2006), combining historical data with 'expert knowledge' (which is usually done in practice) or drawing up some guidelines which must be observed by the decision maker can lead to more reasonable and well-diversified portfolios rather than relying on pure statistical portfolio optimization methods. In this work I will assume that the portfolio weights are generally restricted by a set of linear equality constraints. Thus one might be interested in testing linear hypotheses for the corresponding LMVP rather than the GMVP. I will present standard hypothesis tests for global and local minimum variance portfolios as well as the small-sample distributions of the estimated portfolio weights.

The present work is focused on small-sample rather than large-sample properties but the latter can be easily deduced from the former ones. This is an important issue for I will show that large-sample approximations fail if the sample size is large but the number of observations relative to the number of assets is small. As already mentioned I will concentrate on linear equality constraints though it is clear that in many practical situations inequality constraints play an important role. However, the statistical properties of portfolio weights satisfying inequality constraints cannot be studied by standard econometric methods (Geweke, 1986, Gouriéroux et al., 1982, Wolak, 1987). Investigating the role of linear inequality constraints is left for future research.

In the next section I recall some standard hypothesis tests for the GMVP. The following section deals with hypothesis tests for local minimum variance portfolios. It is shown that, after a suitable transformation of the data, the corresponding tests follow immediately by applying the results of Section 2. In Section 4 the joint distribution of the weights of global and local minimum variance portfolios is derived. The first two moments of an unbiased estimator for the expected portfolio return are also presented. Section 5 contains an empirical study where the following results are applied to stock market data and Section 6 concludes the present work.

2. Hypothesis Tests for the Global Minimum Variance Portfolio

2.1. **Theoretical Foundation.**

Note that $w = \Sigma^{-1}1/(1'\Sigma^{-1}1)$ is a nonlinear function of Σ. However, Kempf and Memmel (2006) noticed that minimizing the variance of the portfolio return can be viewed as a linear regression problem. The return of the GMVP can be written as

$$(1 - w_2 - \ldots - w_d)R_1 + w_2R_2 + \ldots + w_dR_d = \eta + \varepsilon \,, \tag{2.1.1}$$

where $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. By defining $\beta_1 := \eta$, $\beta_j := w_j$, $\Delta R_j := R_1 - R_j$ for $j = 2, \ldots, d$, and $u := \varepsilon$, Eq. 2.1.1 becomes equivalent to

$$R_1 = \beta_1 + \beta_2\Delta R_2 + \ldots + \beta_d\Delta R_d + u \,. \tag{2.1.2}$$

Note that this is a linear regression equation with *stochastic* regressors but the joint normality assumption guarantees that the usual results of econometric theory still hold in this context.

The following proposition is a standard result of linear regression theory. It is crucial for understanding the basic idea of the subsequent derivations and thus it is recalled here for convenience.

Proposition 2.1.1. *Let $Z = (Z_1, \ldots, Z_d)$ be a d-dimensional random vector with positive-definite covariance matrix. Consider the vector*

$$\underset{(d\times 1)}{\beta} = (\beta_1, \ldots, \beta_d) := \arg\min_b \mathrm{E}\big\{(Z_1 - b_1 - b_2Z_2 - \ldots - b_dZ_d)^2\big\},$$

where $b = (b_1, \ldots, b_d)$ and define

$$u := Z_1 - \beta_1 - \beta_2Z_2 - \ldots - \beta_dZ_d \,.$$

The vector β exists and is uniquely defined. More precisely, the subvector $\beta^{\mathrm{s}} := (\beta_2, \ldots, \beta_d)$ is given by

$$\beta^{\mathrm{s}} = \mathrm{Var}(Z^{\mathrm{s}})^{-1}\mathrm{Cov}(Z_1, Z^{\mathrm{s}}) \,,$$

where $Z^{\mathrm{s}} := (Z_2, \ldots, Z_d)$, $\mathrm{Var}(Z^{\mathrm{s}})$ $((d-1)\times(d-1))$ is the covariance matrix of Z^{s}, and $\mathrm{Cov}(Z_1, Z^{\mathrm{s}})$ is the $(d-1)\times 1$ vector of covariances between Z_1 and Z_j $(j = 2, \ldots, d)$. Moreover, the parameter β_1 is given by

$$\beta_1 = \mathrm{E}(Z_1) - \mathrm{E}(Z^{\mathrm{s}})'\beta^{\mathrm{s}}$$

and it holds that $\mathrm{E}(u) = 0$ *as well as* $\mathrm{Cov}(X_j, u) = 0$ *for* $j = 2, \ldots, d$ *.*

The parameters $\beta_1, \ldots, \beta_d$ in Eq. 2.1.2 are chosen in such a way that $\mathrm{E}(u) = 0$ holds and $\mathrm{Var}(u) = E(u^2)$ is minimal, i.e. $\mathrm{Cov}(\Delta R_j, u) = 0$ $(j = 2, \ldots, d)$. So it has been shown that Eq. 2.1.2 indeed is a proper linear regression equation satisfying the standard assumptions of linear regression theory, especially the *strict exogeneity assumption* (Hayashi, 2000, p. 7). For that reason it is possible to develop several exact hypothesis tests for the GMVP by standard methods of econometrics (cf. Kempf and Memmel, 2006).

The next corollary states that the converse of Proposition 2.1.1 is true.

Corollary 2.1.2. *Let* $Z = (Z_1, \ldots, Z_d)$ *be a* d*-dimensional random vector with positive-definite covariance matrix. Search for some numbers* $b_1, \ldots, b_d$ *such that* $\mathrm{E}(u^*) = 0$ *and* $\mathrm{Cov}(Z_j, u^*) = 0$ *for* $j = 2, \ldots, d$ *, where*

$$u^* := Z_1 - b_1 - b_2 Z_2 - \ldots - b_d Z_d \,.$$

The vector $b = (b_1, \ldots, b_d)$ *exists and is uniquely defined by* $b = \beta$ *where* β *is given by Proposition 2.1.1.*

The proof of that corollary follows immediately from the proof of Proposition 2.1.1 (see the appendix) and noting that the linear equation

$$0 = \mathrm{Cov}(Z^{\mathrm{s}}, u^*) = \mathrm{Cov}(Z_1, Z^{\mathrm{s}}) - \mathrm{Var}(Z^{\mathrm{s}}) b^{\mathrm{s}}$$

has a unique solution (due to the positive definiteness of $\mathrm{Var}(Z^{\mathrm{s}})$). Corollary 2.1.2 implies that the strict exogeneity assumption is satisfied only if the error u has minimum variance. Later on it is shown that for that reason the standard test statistics for the GMVP in general must not be applied for testing a LMVP.

2.2. **Statistical Inference.** Of course, in practice the weights of the GMVP are unknown, i.e. they have to be estimated from historical data. Let

$$\underset{(n\times d)}{\mathbf{R}} := \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1d} \\ R_{21} & R_{22} & \cdots & R_{2d} \\ \vdots & \vdots & & \vdots \\ R_{n1} & R_{n2} & \cdots & R_{nd} \end{bmatrix}$$

be a sample of $n > d$ independent copies of R. Now define

$$\underset{(n\times d)}{\mathbf{X}} := \begin{bmatrix} 1 & X_{12} & \cdots & X_{1d} \\ 1 & X_{22} & \cdots & X_{2d} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n2} & \cdots & X_{nd} \end{bmatrix}, \tag{2.2.1}$$

where $X_{ij} := R_{i1} - R_{ij}$ $(i = 1, \ldots, n,\ j = 2, \ldots, d)$ and $\mathbf{Y} := (Y_1, \ldots, Y_n)$ $(n \times 1)$ with $Y_i := R_{i1}$ $(i = 1, \ldots, n)$. Similarly, I will also write $X := (1, X_2, \ldots, X_d)$ $(d \times 1)$, $X^{\mathrm{s}} := (X_2, \ldots, X_d)$ $((d-1) \times 1)$, and $Y \equiv R_1$ (1×1).

According to the standard notation of linear regression theory the linear model represented by Eq. 2.1.2 is given by

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{u} \,, \tag{2.2.2}$$

where $\beta = (\beta_1, \ldots, \beta_d)$ $(d \times 1)$ contains the weights $\beta_2, \ldots, \beta_d$ of the GMVP – except for the first one – as well as the expected return β_1 of the GMVP. Here $\mathbf{u} := (u_1, \ldots, u_n)$ is an $n \times 1$ vector of unobservable errors. Hence, the *ordinary least squares* (OLS) estimator for β can be calculated by

$$\hat{\beta}_{\mathrm{OLS}} = \left(\hat{\eta}, \hat{w}_2, \ldots, \hat{w}_d\right) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \,. \tag{2.2.3}$$

In fact the weights of the GMVP can be estimated by

$$\hat{w}^{\mathrm{s}} := \left(\hat{w}_2, \ldots, \hat{w}_d\right) = \widehat{\Omega}^{-1}\hat{\omega} \,,$$

where $\widehat{\Omega}$ is the sample covariance matrix of X^{s} and $\hat{\omega}$ is the $(d-1)\times 1$ vector of the sample covariances between Y and X_j $(j = 2,\ldots,d)$. The random vector

$$\hat{w} := \big(1 - 1'\hat{w}^{\mathrm{s}}, \hat{w}^{\mathrm{s}}\big)$$

is the best unbiased estimator for the GMVP in the context of normally distributed asset returns (Kempf and Memmel, 2006). Note that if the normal distribution assumption for the asset returns is dropped, in general it cannot be guaranteed that the standard assumptions of linear regression theory are satisfied and thus $\hat{w}$ might become inefficient.

Kempf and Memmel (2006) showed that $\hat{w} = \widehat{\Sigma}^{-1}1/(1'\widehat{\Sigma}^{-1}1)$, i.e. $\hat{w}$ corresponds to the traditional GMVP estimator, where the $d\times d$ matrix

$$\widehat{\Sigma} := \mathbf{R}'\mathbf{R}/n - \overline{\mathbf{r}}\,\overline{\mathbf{r}}\,'$$

represents the sample covariance matrix and $\overline{\mathbf{r}} := \mathbf{R}'1/n$ $(d\times 1)$ is the sample mean vector of R. Further, also the OLS estimator for the expected GMVP return corresponds to its traditional estimator, i.e. $\hat{\eta} = \overline{\mathbf{r}}\,'\hat{w}$.

The relation between the OLS estimator $\hat{\beta}_{\mathrm{OLS}}$ and the residual vector $\hat{\mathbf{u}}$ $(n\times 1)$ can be represented by

$$R_1 = \hat{\eta} + \hat{w}_2\Delta R_2 + \ldots + \hat{w}_d\Delta R_d + \hat{u}$$

or – according to the usual notation of linear regression theory – as

$$\mathbf{Y} = \mathbf{X}\hat{\beta}_{\mathrm{OLS}} + \hat{\mathbf{u}}\,. \tag{2.2.4}$$

Let $\hat{\sigma}^2_{\mathrm{OLS}} := \hat{\mathbf{u}}'\hat{\mathbf{u}}/(n-d)$ be the unbiased OLS estimator for σ^2. It holds that

$$\hat{\sigma}^2 := \hat{w}'\widehat{\Sigma}\,\hat{w} = 1/\big(1'\widehat{\Sigma}^{-1}1\big) = \frac{n-d}{n}\cdot\hat{\sigma}^2_{\mathrm{OLS}}\,,$$

where $\hat{\sigma}^2$ is the traditional estimator for the variance of the GMVP return.

Now consider the fundamental least squares problem

$$(\mathbf{Y} - \mathbf{X}b)'(\mathbf{Y} - \mathbf{X}b) \to \min_b! \tag{2.2.5}$$

under the additional constraint $Hb = h$, where H $(q\times d)$ is a matrix with $\operatorname{rk} H = q \leq d$ and h $(q\times 1)$ some arbitrary vector. The solution of this minimization problem is given by the *restricted least squares* (RLS) estimator

$$\hat{\beta}_{\mathrm{RLS}} := \arg\min_b\,(\mathbf{Y} - \mathbf{X}b)'(\mathbf{Y} - \mathbf{X}b)\,, \qquad \text{s.t. } Hb = h\,. \tag{2.2.6}$$

In the following I will write $\hat{\beta}_{\mathrm{RLS}} = (\hat{\eta}^*, \hat{w}_2^*, \ldots, \hat{w}_d^*)$ and correspondingly

$$R_1 = \hat{\eta}^* + \hat{w}_2^*\Delta R_2 + \ldots + \hat{w}_d^*\Delta R_d + \hat{u}^* \tag{2.2.7}$$

or more compactly

$$\mathbf{Y} = \mathbf{X}\hat{\beta}_{\mathrm{RLS}} + \hat{\mathbf{u}}^* \tag{2.2.8}$$

to indicate that $\hat{\mathbf{u}}^*$ $(n\times 1)$ is the residual vector with respect to the RLS estimator and not to the OLS estimator. The RLS estimator can be calculated explicitly by applying the Lagrange method (Greene, 2003, p. 100). However, in Section 3.2 I will present an alternative method which is more useful in the context of portfolio optimization.

Here only inhomogeneous regressions are taken into consideration and so both $\hat{\mathbf{u}}$ and $\hat{\mathbf{u}}^*$ have zero means. That is to say (2.2.5) indeed leads to the local minimum *variance* portfolio satisfying the given restriction $Hb = h$. However, in contrast to the unrestricted case, each column of $\mathbf{X}$ is correlated with $\hat{\mathbf{u}}^*$ in general. More precisely, $\mathbf{X}'\hat{\mathbf{u}}^* \neq 0$ if the linear restrictions are binding. This is an empirical consequence of Corollary 2.1.2. In the following I will write

$$\hat{w}^* := \big(1 - 1'\hat{w}^{*\mathrm{s}}, \hat{w}^{*\mathrm{s}}\big)\,, \tag{2.2.9}$$

where $\hat{w}^{*\mathrm{s}} := (\hat{w}_2^*, \ldots, \hat{w}_d^*)$.

An exact or, say, small-sample hypothesis test against $H_0 : H\beta = h$ is given by the next theorem. For an alternative representation of that F-test and some applications to financial data see Kempf and Memmel (2006).

Theorem 2.2.1. *Let $\hat{w}$ be the traditional estimator for the GMVP $w = (w_1, \ldots, w_d)$ and $\hat{w}^*$ the RLS estimator given by Eq. 2.2.9. Further, let η be the expected return of the GMVP. If $H\beta = h$ with $\beta = (\eta, w_2, \ldots, w_d)$ it holds that*

$$\frac{n-d}{q} \cdot \frac{(\hat{w} - \hat{w}^*)'\widehat{\Sigma}\,(\hat{w} - \hat{w}^*)}{\hat{\sigma}^2} \sim F_{q,n-d}\,, \tag{2.2.10}$$

where $\hat{\sigma}^2$ denotes the traditional estimator for the variance of the GMVP return.

A similar F-test for the TP (or any other efficient portfolio which is proportional to the TP) has been obtained by Britten-Jones (1999). The result given in Theorem 2.2.1 does not follow from this F-test since Britten-Jones requires the existence of a risk-free asset and the considered portfolios always lie on the capital market line but not on the efficient frontier.

Another important hypothesis is given by $H_0 : \sigma^2 \geq \sigma_0^2$ (for some $\sigma_0^2 > 0$) which can be tested by the next theorem (cf. Kempf and Memmel, 2006).

Theorem 2.2.2. *Consider the traditional estimator $\hat{\sigma}^2$ for the variance σ^2 of the GMVP return. It holds that*

$$\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-d}\,.$$

This is a standard result from linear regression theory (Greene, 2003, p. 50) after noting that $\hat{\sigma}^2 = \hat{\mathbf{u}}'\hat{\mathbf{u}}/n$ and so the proof can be skipped. The parameter uncertainty concerning the variance σ^2 of the GMVP return can be quantified by $\sigma^2 \,|\, \hat{\sigma}^2 \sim \hat{\sigma}^2 n/\chi^2_{n-d}$ either from a *fiducial* (Rao, 1965, Section 5b.5) or *Bayesian* perspective (by using Jeffreys' prior distribution for σ^2), where the estimate $\hat{\sigma}^2$ is considered as fixed. Since $\mathrm{E}(n/\chi^2_{n-d}) = n/(n-d-2)$, it follows that

$$\mathrm{E}\big(\sigma^2 \,|\, \hat{\sigma}^2\big) \approx \frac{\hat{\sigma}^2}{1 - 1/Q}\,,$$

with $Q := n/d > 1$, i.e. the estimation risk essentially depends on the sample size relative to the number of assets. Hence, the capital market is said to be *high-dimensional* if Q – which can be interpreted as the *effective* size of a multivariate sample – is small. In that case small-sample inference must be applied even if the number of observations is large.

Usually an investor not only wants to know whether the variance of the GMVP is bounded by some number σ_0^2 but also to test against $H_0 : \eta \leq \eta_0$, where η represents the true expected return of the GMVP. This can be done by applying the next theorem.

Theorem 2.2.3. *Consider the traditional estimators $\hat{\eta}$ for the expected GMVP return η and $\hat{\sigma}^2$ for the variance of the GMVP return. It holds that*

$$\frac{\hat{\eta} - \eta}{\sqrt{\big\{\hat{\sigma}^2(1 + \bar{\mathbf{r}}'\widehat{\Sigma}^{-1}\bar{\mathbf{r}}) - \hat{\eta}^2\big\}/(n-d)}} \sim t\,(n-d)\,,$$

where $t(n-d)$ denotes Student's t-distribution with $n-d$ degrees of freedom.

The latter theorem completes the repertoire of standard hypothesis tests for the GMVP. In the next section it is shown that the same repertoire can be used also for local minimum variance portfolios after a suitable transformation of the data.

3. Hypothesis Tests for Local Minimum Variance Portfolios

3.1. **Theoretical Foundation.** Consider the LMVP

$$\underset{(d\times 1)}{w^*} = \big(w_1^*,\ldots,w_d^*\big) := \arg\min_v \operatorname{Var}(R'v)\,, \qquad \text{s.t. } Fv = f\,, \tag{3.1.1}$$

where the budget constraint $1'v = 1$ is also satisfied. Here f is a $q\times 1$ vector and F is a $q\times d$ matrix ($q < d$) such that the stacked $(q+1)\times d$ matrix $(1', F)$ has rank $q+1$. Both f and F are assumed to be non-random. Using the definitions from above this can be formulated as a least squares problem, i.e.

$$\underset{(d\times 1)}{\beta^*} := \arg\min_b \mathrm{E}\big\{(Y - X'b)^2\big\} \tag{3.1.2}$$

under a set of linear restrictions affecting only the parameters $b_2,\ldots,b_d$ (i.e. the portfolio weights without the first one). However, due to Corollary 2.1.2 this would not lead to a proper linear regression equation, say

$$R_1 = \beta_1^* + \beta_2^*\Delta R_2 + \ldots + \beta_d^*\Delta R_d + u^*\,, \tag{3.1.3}$$

since u^* generally depends on the regressors $\Delta R_2,\ldots,\Delta R_d$. So the standard test statistics which have been provided in Section 2.2 cannot be applied. However, in the following it will be shown how to reformulate (3.1.2) such that the standard hypothesis tests become applicable.

Consider a matrix $\mathcal{T}$ $(d\times(d-q))$ such that

$$\begin{bmatrix} 1' \\ F \end{bmatrix} \mathcal{T} = \begin{bmatrix} 1' \\ f1' \end{bmatrix}.$$

Then the condition $F\mathcal{T}v = f$ is satisfied for any vector $v \in \mathbb{R}^{d-q}$ with $1'v = 1$. Moreover, it is guaranteed that $1'\mathcal{T}v = 1$, i.e. the budget constraint holds also for $\mathcal{T}v \in \mathbb{R}^d$. Now the LMVP can be simply found by searching for the GMVP with respect to the *transformed* asset return vector

$$R^* = (R_1^*,\ldots,R_{d-q}^*) := \mathcal{T}'R\,.$$

Hence, the least squares problem given by (3.1.2) can be reformulated as

$$\underset{((d-q)\times 1)}{\alpha} := \arg\min_a \mathrm{E}\big\{(Y^* - X^{*\prime}a)^2\big\}\,. \tag{3.1.4}$$

Here $Y^* := R_1^*$ and $X^* := (1, X_2^*,\ldots,X_{d-q}^*)$ with $X_j^* := R_1^* - R_j^*$ for $j = 2,\ldots,d-q$. The corresponding modified linear model

$$R_1^* = \alpha_1 + \alpha_2\Delta R_2^* + \ldots + \alpha_{d-q}\Delta R_{d-q}^* + u^* \tag{3.1.5}$$

is quite similar to the linear regression equation 3.1.3. However, the vector α can be chosen *without any restriction* from $\mathbb{R}^{d-q}$ so that $\operatorname{Var}(u^*)$ becomes minimal and it is always guaranteed that the condition $Fw^* = f$ is satisfied after the re-parameterization

$$w^* := \mathcal{T}(1 - 1'\alpha^{\mathrm{s}}, \alpha^{\mathrm{s}})\,,$$

where $\alpha^{\mathrm{s}} := (\alpha_2,\ldots,\alpha_{d-q})$. Eq. 3.1.5 in fact represents a proper linear regression equation, i.e. $\mathrm{E}(u^*) = 0$ and $\operatorname{Cov}(X_j^*, u^*) = 0$ for $j = 2,\ldots,d-q$.

The LMVP is given by

$$w^* = \frac{\mathcal{T}(\mathcal{T}'\Sigma\,\mathcal{T})^{-1}1}{1'(\mathcal{T}'\Sigma\,\mathcal{T})^{-1}1}$$

and the quantity $\mathcal{T}$ can be derived as follows. Assume that the $(q+1)\times d$ matrix

$$\overline{F} := \begin{bmatrix} 1' \\ F \end{bmatrix} = \begin{bmatrix} \overline{F}_1 & \overline{F}_2 \end{bmatrix}$$

is structured in such a way that $\overline{F}_1$ is a nonsingular $(q+1)\times(q+1)$ matrix and $\overline{F}_2$ is a $(q+1)\times(d-q-1)$ matrix. A structure like this can be always found by a permutation of the columns of F since this has full row rank. Similarly, consider the partition

$$\mathcal{T} = \begin{bmatrix} \mathcal{T}_1 \\ \mathcal{T}_2 \end{bmatrix},$$

where $\mathcal{T}_1$ is a $(q+1) \times (d-q)$ and $\mathcal{T}_2$ is a $(d-q-1) \times (d-q)$ matrix.

Recall that $\mathcal{T}$ has to be such that $\overline{F}\mathcal{T} = (1', f1')$. In the following let

$$\mathcal{T}_2 = \begin{bmatrix} 0 & I_{d-q-1} \end{bmatrix} \tag{3.1.6}$$

so that

$$\overline{F}\mathcal{T} = \overline{F}_1\mathcal{T}_1 + \begin{bmatrix} 0 & \overline{F}_2 \end{bmatrix} = \begin{bmatrix} 1' \\ f1' \end{bmatrix}.$$

That means

$$\mathcal{T}_1 = \overline{F}_1^{-1}\left(\begin{bmatrix} 1' \\ f1' \end{bmatrix} - \begin{bmatrix} 0 & \overline{F}_2 \end{bmatrix} \right). \tag{3.1.7}$$

Note that for the special case $\overline{F} = 1'$, i.e. if there is no additional restriction at all, it holds that $\mathcal{T} = I_d$.

3.2. **Statistical Inference.** In Section 2.2 the minimization problem given by Eq. 2.2.6 has been considered, which involves the expected return estimate $\hat{\beta}_{\mathrm{RLS},1} = \hat{\eta}^* = \bar{\mathbf{r}}'\hat{w}^*$. Note that the $q \times d$ matrix H refers to the expected GMVP return β_1 and the GMVP weights *without* the first one. However, in practical situations linear constraints possibly involve the first portfolio weight by considering the vector $w^+ := (\eta, w_1, \ldots, w_d)$. That means the null hypothesis is given by $H_0\colon Gw^+ = g$ where G is a $q \times (d+1)$ matrix with $\operatorname{rk} G = q$ and g is an arbitrary $q \times 1$ vector. In fact, in that case the LMVP w^* defined by Eq. 3.1.1 has to be found under the budget constraint $1'v = 1$ and

$$G \begin{bmatrix} \bar{\mathbf{r}}' \\ I_d \end{bmatrix} v = g.$$

That means (2.2.6) can be solved in the same manner as (3.1.1) if the sample mean vector $\bar{\mathbf{r}}$ is included in the linear constraint $Fv = f$. Thus any *Markowitz portfolio*

$$\underset{(d\times 1)}{w_{\mathrm{M}}} = \arg\min_v v'\Sigma\, v, \qquad \text{s.t. } \mu' v = \eta_0$$

can be represented as a GMVP after a suitable transformation of the data. However, since in that case the linear constraint is stochastic, the presented methods of statistical inference cannot be applied.

Due to the preceding theoretical arguments the parameter vector α can be readily estimated by the OLS estimator

$$\hat{\alpha}_{\mathrm{OLS}} = \left(\hat{\alpha}_{\mathrm{OLS},1}, \ldots, \hat{\alpha}_{\mathrm{OLS},d-q}\right) := (\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}\mathbf{X}^{*\prime}\mathbf{Y}^*, \tag{3.2.1}$$

where

$$\underset{(n\times(d-q))}{\mathbf{X}^*} := \begin{bmatrix} 1 & X_{12}^* & \cdots & X_{1,d-q}^* \\ 1 & X_{22}^* & \cdots & X_{2,d-q}^* \\ \vdots & \vdots & & \vdots \\ 1 & X_{n2}^* & \cdots & X_{n,d-q}^* \end{bmatrix} \tag{3.2.2}$$

and $\mathbf{Y}^* := (Y_1^*, \ldots, Y_n^*)$ $(n \times 1)$.

The relationship between the residual vector $\hat{\mathbf{u}}^*$ $(n \times 1)$ and the OLS estimator $\hat{\alpha}_{\mathrm{OLS}}$ can be represented by

$$\mathbf{Y}^* = \mathbf{X}^*\hat{\alpha}_{\mathrm{OLS}} + \hat{\mathbf{u}}^*. \tag{3.2.3}$$

After defining $\hat{\alpha}_{\mathrm{OLS}}^{\mathrm{s}} := (\hat{\alpha}_{\mathrm{OLS},2}, \ldots, \hat{\alpha}_{\mathrm{OLS},d-q})$, the OLS estimator for w^* corresponds to

$$\hat{w}^* := \mathcal{T}\left(1 - 1'\hat{\alpha}_{\mathrm{OLS}}^{\mathrm{s}}, \hat{\alpha}_{\mathrm{OLS}}^{\mathrm{s}}\right) \tag{3.2.4}$$

and $\hat{\alpha}_{\mathrm{OLS},1} = \hat{\eta}^*$ is the estimator for the expected return of the LMVP. Hence, $\hat{w}^*$ turns out to be the best unbiased estimator for the corresponding LMVP.

Any null hypothesis concerning the *local* minimum variance portfolio can be implemented in the same way as described at the beginning of this section. Let $w^{*+} := (\eta^*, w_1^*, \ldots, w_d^*)$ be the parameter vector of the LMVP and consider the null hypothesis $H_0^* : Cw^{*+} = c$, where C is some $p \times (d+1)$ matrix with $\operatorname{rk} C = p \leq d - q$ and c is an arbitrary $p \times 1$ vector. This is similar to the null hypothesis $H_0 : Gw^+ = g$. However, for H_0^* there are only $d - q$ degrees of freedom left since the LMVP has been already characterized by q linear restrictions. Of course it has also to be guaranteed that H_0^* does not imply the linear restrictions of the LMVP or the budget constraint. More precisely, consider the linear system of equations

$$(3.2.5) \qquad \begin{bmatrix} \underset{(1\times 1)}{0} & \underset{(1\times d)}{1'} \\ \underset{(q\times 1)}{0} & \underset{(q\times d)}{F} \\ \underset{(p\times 1)}{C_1} & \underset{(p\times d)}{C_2} \end{bmatrix} \begin{bmatrix} \eta^* \\ w_1^* \\ \vdots \\ w_d^* \end{bmatrix} = \begin{bmatrix} \underset{(1\times 1)}{1} \\ \underset{(q\times 1)}{f} \\ \underset{(p\times 1)}{c} \end{bmatrix}$$

with $p + q \leq d$. Now it has to be guaranteed that the $(p + q + 1) \times (d + 1)$ matrix on the left hand side possesses full row rank.

The restricted minimum variance portfolio according to H_0^* is denoted by $\hat{w}^{**}$ and can be calculated as described for the null hypothesis H_0 without using the Lagrange method. Moreover, the standard hypothesis tests derived in Section 2.2 can be applied to local minimum variance portfolios just by transforming the asset returns $R_1, \ldots, R_d$ into the portfolio returns $R_1^*, \ldots, R_{d-q}^*$. Then it holds that

$$(3.2.6) \qquad \frac{n-d+q}{p} \cdot \frac{(\hat{w}^* - \hat{w}^{**})'\widehat{\Sigma}\,(\hat{w}^* - \hat{w}^{**})}{\hat{\sigma}^{*2}} \sim F_{p,n-d+q}\,,$$

provided H_0^* is not binding, as well as

$$\frac{n\hat{\sigma}^{*2}}{\sigma^{*2}} \sim \chi^2_{n-d+q}$$

and

$$\frac{\hat{\eta}^* - \eta^*}{\sqrt{\big\{\hat{\sigma}^{*2}(1 + \bar{\mathbf{r}}'\mathcal{T}(\mathcal{T}'\widehat{\Sigma}\,\mathcal{T})^{-1}\mathcal{T}'\,\bar{\mathbf{r}}) - \hat{\eta}^{*2}\big\}/(n-d+q)}} \sim t\,(n-d+q)\,.$$

That means

(1) the F-distribution given in Theorem 2.2.1,
(2) the χ^2-distribution from Theorem 2.2.2, and
(3) the t-distribution presented in Theorem 2.2.3

simply capture q additional degrees of freedom, where q is the number of linear equalities characterizing the LMVP. Hence, imposing linear restrictions is a simple dimension reduction technique which reduces the parameter uncertainty of portfolio optimization. A similar effect can be also observed for linear inequality constraints like setting upper bounds for the portfolio weights or using short-selling constraints. This is confirmed by several simulation and out-of-sample studies (Eichhorn et al., 1998, Frost and Savarino, 1988, Grauer and Shen, 2000, Jagannathan and Ma, 2003).

It is worth to point out that the GMVP as well as any LMVP can exhibit large positive or negative weights which are not caused by estimation errors. Asset returns in general are dominated by a large principal component representing the *market* or *systematic* risk. There often exist some assets – typically belonging to the finance sector – which strongly depend on the market risk and have a relatively small amount of idiosyncratic risk. In that case extreme negative portfolio weights occur as a matter of principle (Green and Hollifield, 1992). Thus, placing short-selling constraints on the portfolio weights can increase the out-of-sample variance of the portfolio return. Of course, this holds also if linear equality constraints are considered. Nevertheless, Jagannathan and Ma (2003) argue that the negative effect of restricting portfolio weights is usually outweighed by the positive effect of reducing estimation risk. This question will be treated analytically in a different paper.

4. Distribution of the Estimated Portfolio Weights

In the following section I will concentrate on the small-sample distribution of the estimated weights of global and local minimum variance portfolios. This is only loosely connected to hypothesis testing but the small-sample distribution of the estimated portfolio weights might be of interest in its own right.

4.1. Preliminary Definitions.

For the sake of simplicity from now on I will ignore the standard notation of linear regression theory. Recall that $\hat{w}$ denotes the estimator for the GMVP whereas $\hat{w}^*$ is the estimator for some LMVP. Correspondingly, w symbolizes the true GMVP and w^* is the true LMVP. The expected return of the GMVP is denoted by η whereas the expected return of the LMVP is given by η^*. Moreover, σ^2 is the variance of the GMVP return whereas σ^{*2} symbolizes the variance of the LMVP return. The corresponding traditional estimators for these quantities are given by $\hat{\eta}$, $\hat{\eta}^*$, $\hat{\sigma}^2$, and $\hat{\sigma}^{*2}$.

In the following $t_k(a, B, \nu)$ (where $t(\cdot) \equiv t_1(\cdot)$) stands for the k-variate t-distribution with $\nu > 0$ degrees of freedom, location vector a $(k \times 1)$, and positive-semidefinite dispersion matrix B $(k \times k)$, i.e.

$$a + \frac{\zeta}{\sqrt{\chi^2_\nu/\nu}} \sim t_k(a, B, \nu)\,,$$

where $\zeta \sim \mathcal{N}_k(0, B)$ is stochastically independent of χ^2_ν. Here $\zeta \sim B^{1/2}\xi$ with $\xi \sim \mathcal{N}_k(0, I_k)$ and $B^{1/2}$ is some matrix such that $B^{1/2}B^{1/2\,\prime} = B$.

By defining the $(d-1) \times d$ matrix $\Delta := \begin{bmatrix} 1 & -I_{d-1} \end{bmatrix}$ it follows that $\Delta R = X^{\mathrm{s}}$ and thus $\Omega := \Delta\Sigma\Delta'$ denotes the covariance matrix of X^{s}. Analogously, in the context of local minimum variance portfolios the notation $R^* = \mathcal{T}'R$ and $\Delta R^* = X^{*\mathrm{s}}$ will be used. Further, $\Omega^* := \Delta\Sigma^*\Delta'$ is the covariance matrix of $X^{*\mathrm{s}}$, where $\Sigma^* := \mathcal{T}'\Sigma\mathcal{T}$ denotes the covariance matrix of R^*.

4.2. Global Minimum Variance Portfolio.

The next theorem provides the small-sample distribution of the traditional estimator for the GMVP. Another variant of this theorem can be found in Okhrin and Schmid (2006) and so the proof is skipped.

Theorem 4.2.1. *Let $w = (w_1, \ldots, w_d)$ be the GMVP of d assets and $\hat{w} = (\hat{w}_1, \ldots, \hat{w}_d)$ the corresponding traditional estimator given a sample of asset returns with size $n \geq d$. It holds that*

$$(\hat{w}_2, \ldots, \hat{w}_d) \sim t_{d-1}\Big((w_2, \ldots, w_d), \frac{\sigma^2}{n-d+1} \cdot \Omega^{-1}, n-d+1\Big),$$

where Ω is the covariance matrix of ΔR and $\sigma^2 = w'\Sigma w$ is the variance of the GMVP return.

An unbiased estimator for the covariance matrix of $\hat{w}^{\mathrm{s}} = (\hat{w}_2, \ldots, \hat{w}_d)$ is provided by the next corollary.

Corollary 4.2.2. *Consider a sample of asset returns with size $n \geq d+2$ and let $\hat{w} = (\hat{w}_1, \ldots, \hat{w}_d)$ be the traditional estimator for the GMVP. Then the matrix*

$$\widehat{\mathrm{Var}}\big\{(\hat{w}_2, \ldots, \hat{w}_d)\big\} := \frac{\hat{\sigma}^2}{n-d} \cdot \widehat{\Omega}^{-1}$$

is an unbiased estimator for the covariance matrix of $\hat{w}^{\mathrm{s}} = (\hat{w}_2, \ldots, \hat{w}_d)$, *where* $\widehat{\Omega}$ *is the sample covariance matrix of* ΔR *and* $\hat{\sigma}^2$ *is the traditional estimator for the variance of the GMVP return.*

Note that $\hat{w}_1 = 1 - 1'\hat{w}^{\mathrm{s}}$ and from Theorem 4.2.1 it follows that the GMVP estimator $\hat{w}$ is t-distributed with mean w, dispersion matrix $\sigma^2\Delta'\Omega^{-1}\Delta/(n-d+1)$, and $n-d+1$ degrees of freedom. From Proposition 1 of Okhrin and Schmid (2006) it follows that $\sigma^2\Delta'\Omega^{-1}\Delta = \sigma^2\Sigma^{-1} - ww'$ and thus

$$\hat{w} \sim t_d\Big(w, \big(\sigma^2\Sigma^{-1} - ww'\big)/(n-d+1), n-d+1\Big).$$

Moreover, Corollary 4.2.2 implies that

$$\widehat{\mathrm{Var}}(\hat{w}) := \big(\hat{\sigma}^2\widehat{\Sigma}^{-1} - \hat{w}\hat{w}'\big)/(n-d) \tag{4.2.1}$$

is an unbiased estimator for the covariance matrix of $\hat{w}$.

A stochastic representation for $\hat{\eta}$, i.e. the traditional estimator for the expected return of the GMVP could be found after some calculation. However, this is cumbersome and not useful for econometric purposes. In contrast, the first two moments of the distribution of $\hat{\eta}$ can be easily derived. First of all recall that $\bar{\mathbf{r}}$ and $\widehat{\Sigma}$ are stochastically independent. Thus

$$\mathrm{E}(\hat{\eta}) = \mathrm{E}\big\{\mathrm{E}\big(\bar{\mathbf{r}}'\hat{w} \,|\, \widehat{\Sigma}\big)\big\} = \mathrm{E}(\mu'\hat{w}) = \mu'w = \eta\,.$$

Further, it holds that

$$\begin{aligned}\mathrm{Var}(\hat{\eta}) &= \mathrm{E}\big\{\mathrm{Var}\big(\bar{\mathbf{r}}'\hat{w} \,|\, \widehat{\Sigma}\big)\big\} + \mathrm{Var}\big\{\mathrm{E}\big(\bar{\mathbf{r}}'\hat{w} \,|\, \widehat{\Sigma}\big)\big\} \\ &= \mathrm{E}(\hat{w}'\Sigma\,\hat{w}/n) + \mu'\mathrm{Var}(\hat{w})\mu\,,\end{aligned}$$

and after some calculation it follows from Theorem 4.2.1 that

$$\mathrm{E}(\hat{w}'\Sigma\,\hat{w}) = \frac{n-2}{n-d-1}\cdot\sigma^2\,.$$

That means if $n \geq d+2$,

$$\mathrm{Var}(\hat{\eta}) = \mu'\mathrm{Var}(\hat{w})\mu + \frac{n-2}{n-d-1}\cdot\frac{\sigma^2}{n}\,,$$

where

$$\mathrm{Var}(\hat{w}) = (\sigma^2\Sigma^{-1} - ww')/(n-d-1)\,.$$

Note that σ^2/n is the variance of $\bar{\mathbf{r}}'w$, i.e. the variance of the expected GMVP return if w would be known but the expected asset returns $\mu_1, \ldots, \mu_d$ unknown. That means the estimation risk concerning the expected GMVP return can be decomposed into two parts, viz

(1) one part carrying the estimation risk of the portfolio weights and
(2) another part for the estimation risk concerning the expected returns.

More precisely, the variance of $\hat{\eta}$ is an affine-linear transformation of σ^2/n, where $(n-2)/(n-d-1) \geq 1$ and $\mu'\mathrm{Var}(\hat{w})\mu \geq 0$.

4.3. **Local Minimum Variance Portfolios.** From the previous discussion it is clear that any LMVP can be found in the same manner as the GMVP after transforming the asset return vector R into the portfolio return vector R^*. Recall that the LMVP estimator $\hat{w}^*$ can be written as $\hat{w}^* = \mathcal{T}(1 - 1'\hat{\alpha}^{\mathrm{s}}_{\mathrm{OLS}}, \hat{\alpha}^{\mathrm{s}}_{\mathrm{OLS}})$ (see Section 3.2), where

$$\hat{\alpha}^{\mathrm{s}}_{\mathrm{OLS}} \sim t_{d-q-1}\Big(\alpha^{\mathrm{s}}, \frac{\sigma^{*2}}{n-d+q+1}\cdot\Omega^{*-1}, n-d+q+1\Big).$$

Thus it holds that

$$\hat{w}^* \sim t_d\Big(w^*, \big(\sigma^{*2}\mathcal{T}\Sigma^{*-1}\mathcal{T}' - w^*w^{*\prime}\big)/(n-d+q+1), n-d+q+1\Big).$$

TABLE 1. Industry sectors and numbers of assets.

Industry sector	Assets
Consumer Discretionary	54
Energy	14
Consumer Staples	31
Financial	38
Health Care	22
Industrial	40
Information Technology	20
Materials	24
Telecommunications	5
Utilities	26
Σ	274

Similarly, the remaining assertions follow from the theorems and corollaries already derived for the GMVP, simply by substituting d by $d-q$, η (or $\hat{\eta}$) by η^* (or $\hat{\eta}^*$), and σ^2 (or $\hat{\sigma}^2$) by σ^{*2} (or $\hat{\sigma}^{*2}$). For example, according to Eq. 4.2.1 it follows that

$$\widehat{\mathrm{Var}}(\hat{w}^*) := \left(\hat{\sigma}^{*2}\mathcal{T}\widehat{\Sigma}^{*-1}\mathcal{T}' - \hat{w}^*\hat{w}^{*\prime}\right)/(n-d+q)$$

is an unbiased estimator for the covariance matrix of $\hat{w}^*$. Moreover, $\mathrm{E}(\hat{\eta}^*) = \eta^*$ and

$$\mathrm{Var}(\hat{\eta}^*) = \mu'\mathrm{Var}(\hat{w}^*)\mu + \frac{n-2}{n-d+q-1}\cdot\frac{\sigma^{*2}}{n},$$

where

$$\mathrm{Var}(\hat{w}^*) = \left(\sigma^{*2}\mathcal{T}\Sigma^{*-1}\mathcal{T}' - w^*w^{*\prime}\right)/(n-d+q-1).$$

5. Empirical Study

The following empirical study is based on daily asset prices between 1980-01-01 and 2003-11-26 of the 500 stocks listed by the S&P 500 stock index on 26$^{\text{th}}$ November 2003. The data have been kindly provided by Thomson Financial Datastream and the considered asset prices are adjusted for dividends, splits, etc. However, only for 285 stocks the asset prices are available over the whole sample period. The residual 215 time series exhibit missing values caused by IPO's or M&A's during the sample period and are not considered in this study. Moreover, 274 firms could be found to belong to one of 10 industry sectors according to S&P's *Global Industry Classification Standard* (GICS). The other 11 stocks have been also removed from the study.

The risk-free interest rate is calculated by the secondary market 3-month US treasury bill rate (p.a.). The investment period is supposed to be 21 days (i.e. one trading month) and so the corresponding yields have been divided by 12. For example, the treasury bill rate on 1$^{\text{st}}$ January, 1980, corresponds to 12.04% and so the risk-free interest rate between 1980-01-01 and 1980-01-22 is set to 1%. The interest rates are used to calculate the excess returns of each asset.

The sample contains $n = 296$ monthly excess returns for each of the $d = 274$ firms. The estimated expected return of the GMVP corresponds to $\hat{\eta} = 0.18\%$, whereas $\hat{\sigma} = 0.8\%$ is its estimated standard deviation. The latter is obtained by the biased traditional estimator $\hat{\sigma}^2$. After adjusting for the bias the estimated standard deviation corresponds to

$$\sqrt{\frac{n}{n-d}}\cdot\hat{\sigma} = \sqrt{\frac{1}{1-1/Q}}\cdot\hat{\sigma} = 2.94\%$$

TABLE 2. Estimated weights of the GMVP and corresponding standard errors in parentheses.

Sector	Weight	Sector	Weight
Consumer Discretionary	−14.67% (10.24%)	Industrial	27.07% (13.10%)
Energy	14.33% (4.29%)	Information Technology	−4.70% (4.09%)
Consumer Staples	35.50% (9.79%)	Materials	−5.81% (9.64%)
Financial	−17.14% (7.89%)	Telecommunications	18.18% (5.34%)
Health Care	6.02% (7.19%)	Utilities	41.22% (5.88%)

with effective sample size $Q = n/d = 1.08$. Hence, the considered capital market is high-dimensional and the small-sample bias is tremendously large although there are 296 observations.

For the purpose of dimension reduction a pre-allocation is done by aggregating the stocks within each industry sector. More precisely, the asset returns of the firms belonging to the industry sector 'Consumer Discretionary' (see Table 1) are equally weighted by $1/54$, the asset returns belonging to 'Energy' by $1/14$ and so on. Hence, after the pre-allocation there remain 10 portfolios which can be interpreted as sector indices. The estimate for the expected return of the corresponding GMVP (see Table 2) amounts to $\hat{\eta} = 0.33\%$, whereas the estimated standard deviation is $\hat{\sigma} = 3.62\%$. Now there are only $d = 10$ assets (which are the sector indices), $Q = 29.6$ and so the curse of dimension is lifted. Hence, the estimate for σ based on the unbiased estimate for σ^2 corresponds to 3.68%, which is quite similar to $\hat{\sigma}$.

By applying Theorem 2.2.1 one can test for example against the null hypothesis $H_0 : w = 1/d = 0.1 \cdot \mathbf{1}$, i.e. that the GMVP corresponds to the *trivial portfolio*. Thus $q = d - 1 = 9$, $n - d = 286$, $\hat{w}^* = 0.1 \cdot \mathbf{1}$, and the F-statistic corresponds to

$$\frac{n-d}{q} \cdot \frac{(\hat{w} - \hat{w}^*)'\widehat{\Sigma}\,(\hat{w} - \hat{w}^*)}{\hat{\sigma}^2} = 15.7536 > 1.9127 = F^{-1}_{F,9,286}(1-\alpha)$$

with $\alpha = 0.05$. Hence, H_0 can be rejected which means that for the purpose of risk minimization it is not sufficient to choose the trivial portfolio.

The next null hypothesis is given by $H_0 : \sigma^2 \geq \sigma_0^2 = (0.2)^2/12 = 0.33\%$. Due to Theorem 2.2.2 the test statistic is given by

$$\frac{n\hat{\sigma}^2}{\sigma_0^2} = 116.2874 < 247.8302 = F^{-1}_{\chi^2,286}(\alpha)\,.$$

That means the GMVP has a sufficiently low risk of return (i.e. $12\sigma^2 < (0.2)^2$). Another null hypothesis is given by $H_0 : \eta \leq \eta_0 = 0.02/12 = 0.17\%$. For the t-test based on Theorem 2.2.3 one has to calculate the t-statistic

$$\frac{\hat{\eta} - \eta_0}{\sqrt{\big\{\hat{\sigma}^2(1 + \bar{\mathbf{r}}'\widehat{\Sigma}^{-1}\bar{\mathbf{r}}) - \hat{\eta}^2\big\}/(n-d)}} = 0.7352 \not> 1.6502 = F^{-1}_{t,286}(1-\alpha)\,,$$

and so the null hypothesis *cannot* be rejected. Although $\hat{\eta}$ is twice the size of η_0, the estimate for the expected return of the GMVP is not *significantly* larger than $0.02/12$ or, equivalently, $12\eta > 0.02$. This is a typical problem of performance measurement (Frahm, 2007).

Now suppose that an investor wants to put 80% into the sectors 'Energy' and 'Information Technology' and he is searching for the corresponding LMVP. The matrix F given by (3.1.1) corresponds to the row vector $[\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ ]$,

TABLE 3. Estimated weights of the LMVP and corresponding standard errors in parentheses.

Sector	Weight	Sector	Weight
Consumer Discretionary	−19.48% (12.66%)	Industrial	−18.37% (15.57%)
Energy	51.30% (3.81%)	Information Technology	28.70% (3.81%)
Consumer Staples	77.92% (11.36%)	Materials	−21.99% (11.82%)
Financial	−20.03% (9.76%)	Telecommunications	5.84% (6.50%)
Health Care	−16.77% (8.61%)	Utilities	32.88% (7.23%)

$f = 0.8$, the matrix

$$\overline{F}_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

possesses full rank and

$$\overline{F}_2 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

The transformation matrix $\mathcal{T}$ can be simply calculated by (3.1.6) and (3.1.7) and the estimated weights of the LMVP are given in Table 3. Further, the estimate for the expected LMVP return corresponds to $\hat{\eta}^* = 0.45\%$ and $\hat{\sigma}^* = 4.49\%$ for the standard deviation. Both the risk and expected return are apparently higher for the LMVP than for the GMVP. This effect has been already motivated in Section 1 and indicated by Figure 1.

Similar to the F-test conducted above, the null hypothesis is now that the industry sectors are equally weighted (except for 'Energy' and 'Information Technology'). Here $p = 7$, $n - d + 1 = 287$, and it can be found that $\hat{w}_2^{**} = 64.82\%$, $\hat{w}_7^{**} = 15.18\%$, $\hat{w}_1^{**}, \hat{w}_3^{**}, \ldots, \hat{w}_6^{**}, \hat{w}_8^{**}, \ldots, \hat{w}_{10}^{**} = 2.50\%$. The F-statistic given by (3.2.6) amounts to

$$\frac{n - d + q}{p} \cdot \frac{(\hat{w}^* - \hat{w}^{**})' \widehat{\Sigma} \, (\hat{w}^* - \hat{w}^{**})}{\hat{\sigma}^{*2}} = 18.4532 > 2.0416 = F^{-1}_{F,7,287}(1 - \alpha)\,.$$

That means the LMVP is not a trivial one.

Further, the χ^2-test against the null hypothesis $H_0 \colon \sigma^{*2} \geq \sigma_0^{*2} = 0.33\%$ leads to

$$\frac{n\hat{\sigma}^{*2}}{\sigma_0^{*2}} = 178.8119 < 248.7615 = F^{-1}_{\chi^2,287}(\alpha)$$

and so also the LMVP risk of return is sufficiently low. However, for the t-test against $H_0 \colon \eta^* \leq \eta_0^* = 0.17\%$ it holds that

$$\frac{\hat{\eta}^* - \eta_0^*}{\sqrt{\left\{\hat{\sigma}^{*2}(1 + \overline{\mathbf{r}}\,'\mathcal{T}(\mathcal{T}'\widehat{\Sigma}\,\mathcal{T})^{-1}\mathcal{T}'\,\overline{\mathbf{r}}) - \hat{\eta}^{*2}\right\}/(n - d + q)}} = 1.0689\,,$$

whereas $F^{-1}_{t,287}(1 - \alpha) = 1.6502$. Once again it is not possible to proof that the expected excess return of the LMVP is significantly large whilst the t-value obtained for the LMVP (1.0689) exceeds the t-value of the GMVP (0.7352).

6. Conclusion

Traditional portfolio optimization does not take estimation risk into account. Many empirical and numerical studies show that estimation risk is a substantial drawback of pure statistical portfolio optimization techniques. This is an important problem in practice, particularly when the sample size compared to the number of assets is small. In the present work it has been shown that estimation risk can be simply reduced by imposing linear constraints on the portfolio weights. Small-sample hypothesis tests for global and local minimum variance portfolios have been derived by linear regression theory. Further, the joint distribution of the weights as well as the first two moments of the estimator for the expected return of the global or some local minimum variance portfolio have been calculated. The presented results hold in small samples, which is an important fact since large-sample approximations fail if the sample size is large but the number of observations relative to the number of assets is small. Hence, the estimation risk of global and local minimum variance portfolios can be readily controlled by applying the given instruments even in the context of high-dimensional data.

Appendix

Proof of Proposition 2.1.1. Since

$$\begin{aligned} \mathrm{E}\big\{(Z_1 - b_1 - Z^{\mathrm{s}\prime} b^{\mathrm{s}})^2\big\} &= \mathrm{Var}(Z_1 - b_1 - Z^{\mathrm{s}\prime} b^{\mathrm{s}}) + \big\{\mathrm{E}(Z_1 - b_1 - Z^{\mathrm{s}\prime} b^{\mathrm{s}})\big\}^2 \\ &= \mathrm{Var}(Z_1 - Z^{\mathrm{s}\prime} b^{\mathrm{s}}) + \big\{\mathrm{E}(Z_1) - b_1 - \mathrm{E}(Z^{\mathrm{s}})' b^{\mathrm{s}}\big\}^2, \end{aligned}$$

where $b^{\mathrm{s}} := (b_2, \ldots, b_d)$, it is clear that $\beta_1 = \mathrm{E}(Z_1) - \mathrm{E}(Z^{\mathrm{s}})'\beta^{\mathrm{s}}$ and thus $\mathrm{E}(u) = 0$. That means the minimization problem can be solved equivalently by minimizing

$$\mathrm{E}\big\{(Z_1^* - b_2 Z_2^* - \ldots - b_d Z_d^*)^2\big\}, \tag{6.0.1}$$

where $Z_j^* := Z_j - \mathrm{E}(Z_j)$ for $j = 1, \ldots, d$. Now define $Z^{*\mathrm{s}} := (Z_2^*, \ldots, Z_d^*)$ so that (6.0.1) corresponds to

$$\mathrm{E}\big\{(Z_1^* - Z^{*\mathrm{s}\prime} b^{\mathrm{s}})^2\big\} = \mathrm{Var}(Z_1) - 2\mathrm{Cov}(Z_1, Z^{\mathrm{s}})' b^{\mathrm{s}} + b^{\mathrm{s}\prime}\mathrm{Var}(Z^{\mathrm{s}}) b^{\mathrm{s}}.$$

Due to the positive definiteness of $\mathrm{Var}(Z)$ also $\mathrm{Var}(Z^{\mathrm{s}})$ is positive-definite. Hence, this is a simple quadratic minimization problem and its unique solution is given by

$$\beta^{\mathrm{s}} = \mathrm{Var}(Z^{\mathrm{s}})^{-1}\mathrm{Cov}(Z_1, Z^{\mathrm{s}}).$$

Now calculate the $(d-1) \times 1$ vector of covariances between u and Z_j $(j = 2, \ldots, d)$, i.e.

$$\begin{aligned} \mathrm{Cov}(Z^{\mathrm{s}}, u) &= \mathrm{Cov}(Z^{\mathrm{s}}, Z_1 - \beta_1 - Z^{\mathrm{s}\prime}\beta^{\mathrm{s}}) \\ &= \mathrm{Cov}(Z_1, Z^{\mathrm{s}}) - \mathrm{Var}(Z^{\mathrm{s}})\beta^{\mathrm{s}} = 0. \end{aligned}$$

□

Proof of Theorem 2.2.1. From linear regression theory (Greene, 2003, p. 102) it is known that

$$\frac{n-d}{q} \cdot \frac{\hat{\mathbf{u}}^{*\prime}\hat{\mathbf{u}}^* - \hat{\mathbf{u}}'\hat{\mathbf{u}}}{\hat{\mathbf{u}}'\hat{\mathbf{u}}} \sim F_{q,n-d}.$$

Since Eq. 2.2.7 constitutes an inhomogeneous regression it holds that $\hat{\eta}^* = \bar{\mathbf{r}}'\hat{w}^*$ and hence $\hat{\mathbf{u}}^* = \big(\mathbf{R} - 1\bar{\mathbf{r}}'\big)\hat{w}^*$. That means

$$\hat{\mathbf{u}}^{*\prime}\hat{\mathbf{u}}^*/n = \hat{w}^{*\prime}\big(\mathbf{R} - 1\bar{\mathbf{r}}'\big)'\big(\mathbf{R} - 1\bar{\mathbf{r}}'\big)\hat{w}^*/n = \hat{\sigma}^{*2},$$

where $\hat{\sigma}^{*2} := \hat{w}^{*\prime}\widehat{\Sigma}\,\hat{w}^*$. Since $\hat{\sigma}^2 = \hat{w}'\widehat{\Sigma}\,\hat{w}$ and $\hat{w} = \widehat{\Sigma}^{-1}1/(1'\widehat{\Sigma}^{-1}1)$, it follows that

$$\hat{\sigma}^{*2} = \hat{\sigma}^2 + (\hat{w} - \hat{w}^*)'\widehat{\Sigma}\,(\hat{w} - \hat{w}^*).$$

Note also that $\hat{\sigma}^2 = \hat{\mathbf{u}}'\hat{\mathbf{u}}/n$ and thus

$$\frac{\hat{\mathbf{u}}^{*\prime}\hat{\mathbf{u}}^* - \hat{\mathbf{u}}'\hat{\mathbf{u}}}{\hat{\mathbf{u}}'\hat{\mathbf{u}}} = \frac{(\hat{w} - \hat{w}^*)'\widehat{\Sigma}\,(\hat{w} - \hat{w}^*)}{\hat{\sigma}^2},$$

which leads to the desired F-statistic. □

Proof of Theorem 2.2.3. From linear regression theory (Greene, 2003, p. 51) it follows that

$$\frac{\hat{\eta}-\eta}{\sqrt{n\hat{\sigma}^2\left[(\mathbf{X}'\mathbf{X})^{-1}\right]_{11}/(n-d)}} \sim t\,(n-d)\,,$$

where $\left[(\mathbf{X}'\mathbf{X})^{-1}\right]_{11}$ denotes the upper left component of $(\mathbf{X}'\mathbf{X})^{-1}$, i.e.

$$\left[(\mathbf{X}'\mathbf{X})^{-1}\right]_{11} = \left\{n - n\overline{\mathbf{x}}'(\mathbf{X}^{s\prime}\mathbf{X}^{s})^{-1}\overline{\mathbf{x}}n\right\}^{-1} = \frac{\{1-n\overline{\mathbf{x}}'(\mathbf{X}^{s\prime}\mathbf{X}^{s})^{-1}\overline{\mathbf{x}}\}^{-1}}{n}\,,$$

where $\mathbf{X}^{s}$ $(n\times(d-1))$ symbolizes the regressor matrix $\mathbf{X}$ without the column of ones. Note that $\mathbf{X}^{s\prime}\mathbf{X}^{s} = n\big(\widehat{\Omega}+\overline{\mathbf{x}}\,\overline{\mathbf{x}}'\big)$ and due to the binomial inverse theorem (Press, 2005, p. 23) it holds that

$$n\,(\mathbf{X}^{s\prime}\mathbf{X}^{s})^{-1} = \big(\widehat{\Omega}+\overline{\mathbf{x}}\,\overline{\mathbf{x}}'\big)^{-1} = \widehat{\Omega}^{-1} - \frac{\widehat{\Omega}^{-1}\overline{\mathbf{x}}\,\overline{\mathbf{x}}'\widehat{\Omega}^{-1}}{1+\overline{\mathbf{x}}'\widehat{\Omega}^{-1}\overline{\mathbf{x}}}\,.$$

That is

$$1-n\overline{\mathbf{x}}'(\mathbf{X}^{s\prime}\mathbf{X}^{s})^{-1}\overline{\mathbf{x}} = 1-\overline{\mathbf{x}}'\widehat{\Omega}^{-1}\overline{\mathbf{x}} + \frac{(\overline{\mathbf{x}}'\widehat{\Omega}^{-1}\overline{\mathbf{x}})^2}{1+\overline{\mathbf{x}}'\widehat{\Omega}^{-1}\overline{\mathbf{x}}} = \frac{1}{1+\overline{\mathbf{x}}'\widehat{\Omega}^{-1}\overline{\mathbf{x}}}$$

and thus

$$\left[(\mathbf{X}'\mathbf{X})^{-1}\right]_{11} = \frac{1+\overline{\mathbf{r}}'\Delta'\widehat{\Omega}^{-1}\Delta\overline{\mathbf{r}}}{n}\,.$$

Since $\hat{\sigma}^2\Delta'\widehat{\Omega}^{-1}\Delta = \hat{\sigma}^2\widehat{\Sigma}^{-1} - \hat{w}\hat{w}'$ and $\hat{\eta} = \overline{\mathbf{r}}'\hat{w}$, it follows that

$$n\hat{\sigma}^2\left[(\mathbf{X}'\mathbf{X})^{-1}\right]_{11} = \hat{\sigma}^2 + \overline{\mathbf{r}}'(\hat{\sigma}^2\widehat{\Sigma}^{-1}-\hat{w}\hat{w}')\,\overline{\mathbf{r}} = \hat{\sigma}^2(1+\overline{\mathbf{r}}'\widehat{\Sigma}^{-1}\overline{\mathbf{r}}) - \hat{\eta}^2\,.$$

□

Proof of Corollary 4.2.2. Theorem 4.2.1 implies that the covariance matrix of $(\hat{w}_2,\ldots,\hat{w}_d)$ is given by

$$\mathrm{Var}\{(\hat{w}_2,\ldots,\hat{w}_d)\} = \frac{\sigma^2}{n-d-1}\cdot\Omega^{-1}.$$

From Wishart theory it follows that $\widehat{\Omega}^{-1} \sim W_{d-1}^{-1}\big((\Omega/n)^{-1}, n+d-1\big)$ (Press, 2005, p. 117). Hence, it holds that

$$\mathrm{E}\big(\widehat{\Omega}^{-1}\big) = \frac{(\Omega/n)^{-1}}{(n+d-1)-2\,(d-1)-2} = \frac{n}{n-d-1}\cdot\Omega^{-1}$$

(Press, 2005, p. 119). Moreover, from linear regression theory (Greene, 2003, p. 56) it is known that $\hat{\sigma}^2_{\mathrm{OLS}} = \hat{\mathbf{u}}'\hat{\mathbf{u}}/(n-d)$ is a conditionally unbiased estimator for σ^2. That means

$$\begin{aligned}\mathrm{E}\left(\frac{\hat{\sigma}^2_{\mathrm{OLS}}}{n}\cdot\widehat{\Omega}^{-1}\right) &= \mathrm{E}\left\{\mathrm{E}\left(\frac{\hat{\sigma}^2_{\mathrm{OLS}}}{n}\cdot\widehat{\Omega}^{-1}\,|\,\widehat{\Omega}^{-1}\right)\right\} = \mathrm{E}\left(\frac{\sigma^2}{n}\cdot\widehat{\Omega}^{-1}\right)\\ &= \frac{\sigma^2}{n-d-1}\cdot\Omega^{-1} = \mathrm{Var}\{(\hat{w}_2,\ldots,\hat{w}_d)\}\end{aligned}$$

and note that $\hat{\sigma}^2_{\mathrm{OLS}} = n/(n-d)\cdot\hat{\sigma}^2$. □

Acknowledgements

I would like to thank Dr. Christoph Memmel for many fruitful discussions and Professor Karl Mosler as well as an anonymous referee for their valuable comments on the manuscript. Many thanks belong also to Thomson Financial Datastream for their kind support.

References

F. Black and R. Litterman (1992), 'Global portfolio optimization', *Financial Analysts Journal* **48**, pp. 28–43.

M. Britten-Jones (1999), 'The sampling error in estimates of mean-variance efficient portfolio weights', *Journal of Finance* **54**, pp. 655–671.

V.K. Chopra and W.T. Ziemba (1993), 'The effect of errors in means, variances, and covariances on optimal portfolio choice', *Journal of Portfolio Management* **19**, pp. 6–11.

D. Eichhorn, F. Gupta, and E. Stubbs (1998), 'Using constraints to improve the robustness of asset allocation', *Journal of Portfolio Management* **24**, pp. 41–48.

G. Frahm (2007), 'Testing for the best alternative with an application to performance measurement', Discussion paper, University of Cologne, Department of Economic and Social Statistics, Germany.

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

P.A. Frost and J.E. Savarino (1988), 'For better performance: constrain portfolio weights', *Journal of Portfolio Management* **14**, pp. 29–34.

J. Geweke (1986), 'Exact inference in the inequality constraint normal linear regression model', *Journal of Applied Econometrics* **1**, pp. 127–141.

C. Gouriéroux, A. Holly, and A. Monfort (1982), 'Likelihood ratio test, Wald test, and Kuhn-Tucker test in linear models with inequality constraints on the regression parameters', *Econometrica* **50**, pp. 63–80.

R. Grauer and F. Shen (2000), 'Do constraints improve portfolio performance?', *Journal of Banking and Finance* **24**, pp. 1253–1274.

R. Green and B. Hollifield (1992), 'When will mean-variance efficient portfolios be well diversified?', *Journal of Finance* **47**, pp. 1785–1809.

W.H. Greene (2003), *Econometric Analysis*, Prentice Hall.

F. Hayashi (2000), *Econometrics*, Princeton University Press.

U. Herold and R. Maurer (2006), 'Portfolio choice and estimation risk – a comparison of Bayesian to heuristic approaches', *ASTIN Bulletin* **36**, pp. 135–160.

R. Jagannathan and T. Ma (2003), 'Risk reduction in large portfolios: Why imposing the wrong constraints helps', *Journal of Finance* **58**, pp. 1651–1683.

P. Jorion (1986), 'Bayes-Stein estimation for portfolio analysis', *Journal of Financial and Quantitative Analysis* **21**, pp. 279–292.

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

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

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

A. Kempf and C. Memmel (2006), 'Estimating the global minimum variance portfolio', *Schmalenbach Business Review* **58**, pp. 332–348.

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

O. Ledoit and M. Wolf (2003), 'Improved estimation of the covariance matrix of stock returns with an application to portfolio selection', *Journal of Empirical Finance* **10**, pp. 603–621.

H.M. Markowitz (1952), 'Portfolio selection', *Journal of Finance* **7**, pp. 77–91.

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

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

Y. Okhrin and W. Schmid (2006), 'Distributional properties of portfolio weights', *Journal of Econometrics* **134**, pp. 235–256.

S.J. Press (2005), *Applied Multivariate Analysis*, Dover Publications, second edition.

C.R. Rao (1965), *Linear Statistical Inference and Its Applications*, John Wiley.

B. Scherer (2004), 'Resampled efficiency and portfolio choice', *Financial Markets and Portfolio Management* **18**, pp. 382–398.

J. Tobin (1958), 'Liquidity preferences as behavior towards risk', *Review of Economic Studies* **25**, pp. 325–333.

F.A. Wolak (1987), 'An exact test for multiple inequality and equality constraints in the linear regression model', *Journal of the American Statistical Association* **399**, pp. 782–793.