Preliminary and Incomplete
Please do not quote
Comments will be appreciated.

Assessing the Risk in Mean-Variance Efficient Portfolios

Gopal K. Basak
Department of Mathematics
University of Bristol
Bristol, BS8 1TW, UK
Gopal.K.Basak@bristol.ac.uk, +44 (0)117 928 9164

Ravi Jagannathan
Department of Finance
Kellogg Graduate School of Business
Northwestern University, Evanston, IL 60208
rjaganna@nwu.edu, 847-491-8338
and
NBER

and Tongshu Ma
Department of Finance
David Eccles School of Business
University of Utah, Salt Lake City, UT84112
fintsm@business.utah.edu, 801-581-7506

November 3, 2003

Abstract

We theoretically show that on average the in-sample estimate of the out-of-sample variance of minimum risk portfolios will be optimistic. Imposing portfolio weight constraints reduce the in-sample optimism. We empirically demonstrate that scaling up the in-sample estimate by a factor that is based on the degrees of freedom of the distribution of the sample covariance matrix will in general be inadequate when the number of assets involved is relatively large. In our sample the variance of the minimum risk portfolios computed using the covariance matrix of returns under the predictive distribution using standard diffuse priors also exhibits in-sample optimism. We develop a Jackknife type estimator for the out-of-sample variance of minimum risk portfolios that is valid when returns are i.i.d or have the exchangeability property that performs rather well in our sample.

1 Introduction

Mean-variance optimization has received wide attention in both academic and practitioner literature. Two inputs are needed for that purpose: the vector of expected returns and the covariance matrix of returns. Typically expected return calculations are based on extensive analysis of various types of information about the firm and macro economic conditions whereas covariance matrix estimators are based on historical return data.

Constructing portfolios that have minimum risk (as measured by variance, or variance of the tracking error with reference to some benchmark) for a given level of expected return is not enough. It is also important that the investor knows what the risk of the optimal portfolio is and what the expected return would be. In this paper we focus on this issue.

It is well recognized in the literature that the in-sample value of the variance of an optimal portfolio constructed using historical return data is a rather optimistic estimate of what the true variance of the portfolio is (Jobson and Korkie (1981)). This phenomenon is termed the in-sample optimism in the literature. However the reasons for the in-sample optimism are not well-understood and there has been no suggestion on how to correct this problem.

It might be argued that by suitably scaling up the in-sample variance by a factor that is related to the degree-of-freedom of the distribution of the estimated covariance matrix, we may be able to correct for in-sample optimism. We show that this is not the case. For that purpose we derive lower bounds on the average out of sample variance of an efficient portfolio and upper bounds on the corresponding average in-sample variance. We show that these bounds provide strict inequalities. When returns over time are drawn from an i.i.d multivariate Normal distribution, we show that the lower bound on the out of sample variance is a scale multiple of the in-sample variance that is related to the degrees of freedom. We demonstrate using simulation methods that the average out of sample variance is substantially larger than the its lower bound and hence degrees of freedom based corrections of the in-sample variance is unlikely to provide sufficiently accurate estimates of the out of sample variance. The example in Table IX of Jagannathan and Ma (2003) illustrates the inadequacy of the standard degrees of freedom correction. They report the properties of the global minimum variance portfolio of the 25 Fama and French (1993) size and book/market sorted portfolios, formed using the sample covariance matrix computed with data on monthly returns for 60 months. The average out-of-sample standard deviation is 15.93% whereas the corresponding out of sample number is 7.44%, i.e., the ratio of the two variances is 4.58 ($= \frac{15.93}{7.44}^2$). In contrast, as we show later, the Classical degrees of freedom based correction is approximately $(\frac{T-1}{T-N+1})$, i.e., 1.63.

Further, degrees of freedom based corrections are difficult to derive when using covariance matrices other than the sample covariance matrix or a factor model based estimates. An example would be the shrinkage estimator of Ledoit. The in-sample optimism is also known to be sensitive to portfolio weight constraints (Jobson and Korkie (1981), Frost and Savarino (1988)). If the constraint is tight, then the in-sample optimism is attenuated. Hence when there are portfolio weight constraints, the degree-of-freedom correction needs to consider the tightness of the constraints as well. Degrees of freedom corrections are not available when efficient portfolios are constructed subject to portfolio weight constraints.

Another possible approach to address the in-sample optimism issue would be to follow the Bayesian approach. In that approach, the portfolio manager takes into account of the fact that the means and covariances used in the optimization procedure are in fact random. Therefore, the optimal portfolio will reflect the manager's beliefs about this randomness, and the in-sample optimism will be avoided. However, with standard diffuse priors commonly used in the literature, we find that the variance of the efficient portfolio computed using the predictive distribution can be substantially below its true variance. Under the Bayesian approach, with diffuse prior, the predictive covariance matrix is the sample covariance matrix multiplied by the factor,

$\frac{T+1}{T-N-2}$. For T = 60 and N = 25, the value of this ratio is 1.9. In contrast, as pointed out earlier, the ratio of the out-of-sample to in-sample variances for the minimum variance portfolio of the 25 book to market and size sorted portfolios in Ma and Jagannathan (2003) is 4.58. This should not come as a surprise given that other authors have also observed the inadequacies of standard diffuse priors and the need for modifying them in other related contexts (see Jacquier, Kane and Marcus (2002) for an example).

We therefore propose a Jackknife type method for estimating the risk in optimal portfolios. We demonstrate that the method provides reasonably accurate estimates of the variance of efficient portfolios.

The rest of the paper is organized as follows. In section 2, we prove that the in-sample estimate of the variance of minimum risk portfolios would on average be strictly less than their corresponding population (out-of-sample) variance, i.e., there will be in-sample optimism. In section 3, we examine the in-sample optimism when using Baye's estimator of the covariance matrix. We develop a Jackknife type estimator for assessing the risk in efficient portfolios in section 4 and empirically evaluate its performance in section 5. We summarize and conclude in section 6.

2 Relation between in-sample and out-of-sample variances of mean-variance efficient portfolios

In this section we show that the out-of-sample (population) variance of global minimum variance and minimum tracking error variance portfolios constructed using sample moments will on average be strictly larger than their corresponding in-sample estimates.

We use the following notation: R_t, is the $N \times 1$ vector of date t returns (or returns in excess of some benchmark return) on N primitive assets; μ is the vector of expected returns; and Σ & $\bar{R}$ are unbiased estimates of Σ and μ based on T observations on returns, $\{R_1, R_2, \ldots, R_T\}$. Let $\mathcal{F}_T$ denote the

σ-field or the information set generated by the returns $\{R_1, R_2, \ldots, R_T\}$. We assume that returns have a multidimensional positive density and are conditionally homoscedastic, i.e., $\Sigma = E\left((R_{T+1} - \mu)(R_{T+1} - \mu)' \mid \mathcal{F}_T\right) = E\left((R_{T+1} - \mu)(R_{T+1} - \mu)'\right)$. However, $\mu = E\left(R_{T+1} \mid \mathcal{F}_T\right)$ need not equal $E\left(R_{T+1}\right)$.

Let w_p denote the N vector of portfolio weights corresponding to the mean-variance efficient portfolio with an expected return of ν, i.e., $w = w_p$ solves the following problem.

$$\begin{aligned} &\min_w \{w'\Sigma w\} && \\ &\text{subject to} && w'\mu = \nu \\ &\sum_{i=1}^{N} w_i = 1, && \underline{w_i} \leq w_i \leq \bar{w}_i, \\ &\text{for some constants} && \underline{w_i}, \bar{w}_i, i = 1, \ldots, N \end{aligned} \tag{1}$$

The Lagrangian function associated with the above minimization problem is given by:

$$f_p(w) = \frac{1}{2} w'\Sigma w - \delta_1(w'\mu - \nu) - \delta_2(w'1_N - 1) - (w - \underline{w})'\theta_1 + (w - \bar{w})'\theta_2 \tag{2}$$

where δ_1, δ_2 are non-negative constants and θ_1 and θ_2 are non-negative vectors. Equation (2) describes a parabolic surface as a function of ω. Since Σ is a positive definite matrix, the surface is strictly parabolic, i.e., the Lagrangian function f_p is strictly convex (upwards) with a unique minimum. Let $w = w_p$ along with $(\delta_1 = \delta_1^p,\ \delta_2 = \delta_2^p,\ \theta_1 = \theta_1^p,\ \theta_2 = \theta_2^p))$ solve the minimization problem given in (1). Then $\underline{w} \leq w_p \leq \bar{w}$; $(w_{p,i} - \underline{w}_i)\theta_{1,i}^p = 0, i = 1 \ldots N$; $(\bar{w}_i - w_{p,i})\theta_{2,i}^p = 0, i = 1 \ldots N$; $({w_p}'\mu - \nu) = 0$; and $({w_p}'1_N - 1) = 0$.

Now consider the sample analogue of the minimization problem given in equation (1) obtained by replacing Σ, and μ in $f_p(\omega)$ by their estimated values, S and $\bar{R}$. The Lagrangian function, $f_s(w)$, associated with this minimization problem is given by:

$$f_s(w) = \frac{1}{2} w'S\omega - \delta_1(w'\bar{R} - \nu) - \delta_2(w'1_N - 1) - (w - \underline{w})'\theta_1 + (w - \bar{w})'\theta_2. \tag{3}$$

Let w_s, together with the auxiliary parameters $(\delta_1^s,\ \delta_2^s,\ \theta_1^s,\ \theta_2^s)$, minimize the sample Lagrangian function, $f_s(w)$, given above.

Note that $w_s' S w_s$ is the in-sample variance of the efficient portfolio with in-sample expected return of ν. The population (out of sample) variance of this portfolio is given by $w_s' \Sigma w_s$. The corresponding out of sample mean is $w_s' \mu$. We want to compare in sample variance, $w_s' S w_s$, with the out of sample variance, $w_s' \Sigma w_s$.

When $w_p \neq w_s$ we obtain the following inequalities.

$$\begin{aligned} \frac{1}{2} w_p' \Sigma w_p &= f_p(w_p) \\ &< f_p(w_s) \\ &= \frac{1}{2} w_s' \Sigma w_s - \delta_1^s (w_s' \mu - \nu) - (w_s - \underline{w})' \theta_1^s + (w_s - \bar{w})' \theta_2^s \\ &= \frac{1}{2} w_s' \Sigma w_s - \delta_1^s (w_s' \mu - \nu). \end{aligned} \tag{4}$$

The first (equality) follows from the definition of the Lagrangian function given in equation (2) and the observation that $(w_p - \underline{w})' \theta_1^p = 0$; $(\bar{w}_i - w_{p,i})' \theta_2^p = 0$; $({w_p}' \mu - \nu) = 0$; and $({w_p}' 1_N - 1) = 0$. The second (inequality) follows from the uniqueness of the minimization problem and the fact that $w_p \neq w_s$. The third (equality) follows from the definition of the function $f_p(w_s)$ and the observation that and ${w_s}' 1_N = 1$. The last (equality) follows from the fact $(w_s, \delta_1^s, d_2^s, \theta_1^s, \theta_2^s)$ minimizes the Lagrangian function $f_s(w)$ given by equation (3); and therefore both $(w_s - \underline{w})' \theta_1^s$ and $(\bar{w} - w_s)' \theta_2^s$ are zero by the complementary slack conditions.

By a similar logic, it follows that:

$$\begin{aligned} \frac{1}{2} w_s' S w_s &= f_s(w_s) \\ &< f_s(w_p) \\ &= \frac{1}{2} w_p' S w_p - \delta_1^p (w_p' \bar{R} - \nu) - (w_p - \underline{w})' \theta_1^p + (w_p - \bar{w})' \theta_2^p \\ &= \frac{1}{2} w_p' S w_p - \delta_1^p (w_p' \bar{R} - \nu). \end{aligned} \tag{5}$$

Since $Prob\{w_s = w_p\} < 1$, $E(S) = \Sigma$, and $E(\bar{R}) = \mu$, taking the expectation on both sides of (5) gives:

$$\begin{aligned} E(f_s(w_s)) < E(f_s(w_p)) &= \frac{1}{2}w_p' E(S)w_p - \delta_1^p(w_p' E(\bar{R}) - \nu)) \\ &= \frac{1}{2}w_p' \Sigma w_p - \delta_1^p(w_p' \mu) - \nu) \\ &= \frac{1}{2}w_p' \Sigma w_p \\ &= f_p(w_p). \end{aligned} \tag{6}$$

Similarly, by taking the expectation of both sides of 4 we get:

$$E(f_p(w_p)) < E(f_p(w_s)) = \frac{1}{2}E(w_s' \Sigma w_s) - E(\delta_1^s(w_s' \mu - \nu)).$$

Note that the in expected values of the in-sample and the out of sample variances of the efficient portfolio are, $E(w_s' S w_s)$ and $E(w_p' \Sigma w_p)$ respectively. The former is also equal to $2E(f_s(w_s))$ (see (5). We are now in a position to relate the two.

From (5) we get, $\frac{1}{2}E(w_s' S w_s) = E(f_s(w_s)) < E(f_s(w_p))$; from (6) we get, $E(f_s(w_p)) = \frac{1}{2}w_p' \Sigma w_p f_p(w_p) = E(f_p(w_p))$; and from (4) we get $E(f_p(w_p)) < E(f_p(w_s))$, which in turn equals $\frac{1}{2}E(w_s' \Sigma w_s) - E(\delta_1^s(w_s' \mu - \nu))$. By combining these inequalities and equalities and multiplying by 2 we get:

$$E(w_s' S w_s) < w_p' \Sigma w_p < E(w_s' \Sigma w_s) - 2E(\delta_1^s(w_s' \mu - \nu)). \tag{7}$$

Now consider the global minimum variance (minimum tracking error variance) portfolio. In that case, the constraint that the expected return on the portfolio should equal some target value in the minimization problems given in (4) and 5) do not apply and the right most term in (7) drops out. Hence (7) becomes:

$$E(w_s' S w_s) < w_p' \Sigma w_p < E(w_s' \Sigma w_s). \tag{8}$$

The third term given above, $E(w_s' \Sigma w_s)$ is the expected value of the conditional variance of the return on the global minimum variance portfolio constructed using the estimated covariance matrix, i.e., $E\left(Var(w_s R_{T+1} \mid \mathcal{F}_T)\right)$.

We will refer to $Var(w_s R_{T+1} \mid \mathcal{F}_T)$ as the out of sample variance and $w_s' S w_s$ as the in-sample variance. Hence the in-sample variance of the global minimum variance portfolio constructed using sample moments will on average be strictly smaller than its out-of-sample variance. The unconditional variance of the portfolio will be larger than $E\left(Var(w_s R_{T+1} \mid \mathcal{F}_T)\right)$ by the variance of $w_s'\mu$, i.e., the variance of the conditionally expected return on the portfolio. In what follows we demonstrate that the difference between the out-of-sample and in-sample variances can be large.

2.1 In-sample optimism when returns have an i.i.d multivariate Normal distribution

Consider the global minimum variance portfolio. Define the in-sample optimism measure, iso as,

$$iso = \frac{E(\text{out of sample variance})}{E(\text{in-sample variance})} = \frac{E(w_s'\Sigma w_s)}{E(w_s' S w_s)}.$$

To obtain a lower bound, $\underline{iso}$, for iso divide all the terms in (8) by $E(w_s' S w_s)$ to get:

$$1 < \frac{w_p'\Sigma w_p}{E(w_s' S w_s)} < \frac{E(w_s'\Sigma w_s)}{E(w_s' S w_s)}, \text{ i.e.,}$$
$$1 < \underline{iso} = \frac{w_p'\Sigma w_p}{E(w_s' S w_s)} < \frac{E(w_s'\Sigma w_s)}{E(w_s' S w_s)} \equiv iso. \tag{9}$$

In what follows we derive an exact expression for $\underline{iso}$ when the vector of returns are drawn from an i.i.d joint Normal distribution, and the sample covariance matrix of returns is based on T observations provided T us sufficiently large.

For convenience we will assume that the lower bound, $\underline{w}$ on the vector of portfolio weights in the minimization problem given in equation (1) and its sample counterpart is the zero vector and the the upper bound, $\bar{w} = \mathbf{1}$, i.e., there are no upper bounds on portfolio weights since portfolio weights must sum to unity. In that case, the vector of optimal portfolio weights corresponding to the global minimum variance portfolio, w_p, satisfy the first order conditions for the optimization problem in (1) given by:

$$\Sigma w_p = \theta_1^p + \delta_2^p \mathbf{1}$$

where θ_1^p and δ_2^p are the optimal values for the nonnegative vector, θ_1, and nonnegative scalar, δ_2, defined in equation (2).

Multiplying both sides of the above equation by Σ^{-1} we get:

$$w_p = \Sigma^{-1}\theta_1^p + \Sigma^{-1}\delta_2^p \mathbf{1}$$

In order to derive an expression for <u>*iso*</u> it is convenient to partition the elements of w_p into two sets, the first consisting of those rows whose elements are zeros and the second consisting of rows with strictly positive elements. Without loss of generality we will assume that the first k rows of w_p are zeros and the next $N-k$ rows are strictly positive, i.e., partition w_p' as (w_{p1}', w_{p2}').

When there are no portfolio weight constraints, the vector of portfolio weights that attain the global minimum variance is given by:

$$w_p^{nc} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}$$

In order to relate the two partitions w_{p1} and w_{p2} to the covariance matrix of returns, Σ, we assume that $w_{p,i}^{nc} < 0 \Rightarrow w_{p,i} = 0$; $w_{p,i}^{nc} > 0 \Rightarrow w_{p,i} > 0$; $w_{p,i}^{nc} = 0 \Leftrightarrow (\Sigma^{-1}\mathbf{1})_i = 0 \Rightarrow$ no restrictions on $w_{p,i}$, and every element of $\Sigma^{-1}\mathbf{1}$ is nonzero.

Partition Σ and Σ^{-1}, w_p and θ_1 as

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \qquad \Sigma^{-1} = \begin{pmatrix} \Sigma^{11} & \Sigma^{12} \\ \Sigma^{21} & \Sigma^{22} \end{pmatrix},$$

and

$$w_p = \begin{pmatrix} w_{p1} \\ w_{p2} \end{pmatrix}, \qquad \theta_1^p = \begin{pmatrix} \theta_{1,1}p \\ \theta_{1,2}^p \end{pmatrix},$$

with Σ_{11} and Σ^{11} being a $k \times k$ matrices and w_{p1} and $\theta_{1,1}^p$ are k-dimensional vectors. In this way, $\Sigma^{11}\mathbf{1}_k + \Sigma^{12}\mathbf{1}_{N-k}$ is a vector with all negative elements and $\Sigma^{21}\mathbf{1}_k + \Sigma^{22}\mathbf{1}_{N-k}$ is a vector with all positive elements. Then $\theta^p_{1,2} = 0$ and

$$w_{p2} = \frac{1}{\mathbf{1}'_{N-k}\Sigma_{22}^{-1}\mathbf{1}_{N-k}}\Sigma_{22}^{-1}\mathbf{1}_{N-k}.$$

When S is the sample covariance matrix,

$$Prob(|S^{-1}\mathbf{1} - \Sigma^{-1}\mathbf{1}|) \to 0$$

as $T \to \infty$. In fact, it can be shown to be an almost sure convergence (i.e., $Prob(\lim_{T\to\infty} |S^{-1}\mathbf{1} - \Sigma^{-1}\mathbf{1}| = 0) = 1$. Thus, for large values of T, for almost all samples $S^{-1}\mathbf{1} < 0$ for first k elements as they are also negative for $\Sigma^{-1}\mathbf{1}$. Therefore, for large values of T, the global minimum variance portfolio constructed using the estimated covariance matrix S would be $w_{s1} = 0$ and

$$w_{s2} = \frac{1}{\mathbf{1}'_{N-k}S_{22}^{-1}\mathbf{1}_{N-k}}S_{22}^{-1}\mathbf{1}_{N-k}.$$

Thus for large enough T the three variances, the in-sample variance of the sample global minimum variance portfolio, the variance of the population global minimum variance portfolio, and the out-of-sample variance of the sample global minimum variance portfolio, are as given below.

The in-sample estimate of the sample global minimum variance is

$$w_s'Sw_s = \frac{1}{\mathbf{1}'_{N-k}S_{22}^{-1}\mathbf{1}_{N-k}},$$

the population global minimum variance is

$$w_p'\Sigma w_p = \frac{1}{\mathbf{1}'_{N-k}\Sigma_{22}^{-1}\mathbf{1}_{N-k}},$$

and the out-of-sample variance for the sample minimum-variance portfolio is

$$w_s'\Sigma w_s = \frac{\mathbf{1}'_{N-k}S_{22}^{-1}\Sigma_{22}S_{22}^{-1}\mathbf{1}_{N-k}}{(\mathbf{1}'_{N-k}S_{22}^{-1}\mathbf{1}_{N-k})^2}.$$

We know that $(T-1)S_{22}$ is Wishart distribution with degree of freedom $(T-(N-k)+1)$, and parameter matrix Σ_{22}. That is,

$$(T-1)S_{22} \sim W_N(T-N+k+1,\ \Sigma_{22}).$$

where we follow the notation of Muirhead (1982). From Theorem 3.2.11 of Muirhead (1982), we know that

$$(\mathbf{1}'_{N-k}[(T-1)S_{22}]^{-1}\mathbf{1}_{N-k})^{-1} \sim W_1(T-N+k+1,\ (\mathbf{1}'_{N-k}\Sigma_{22}^{-1}\mathbf{1}_{N-k})^{-1}).$$

But for any scalar random variable x, that has a Wishart$(T-N+k+1,\ \sigma^2)$ distribution $\frac{x}{\sigma^2}$ is $\chi^2_{T-N+k+1}$ (Muirhead (1982), p. 87.) Since the expectation of the $\chi^2_{T-N+k+1}$ is $(T-N+k+1)$, we get,

$$E\left(\frac{(\mathbf{1}'_{N-k}[(T-1)S_{22}]^{-1}\mathbf{1}_{N-k})^{-1}}{(\mathbf{1}'_{N-k}\Sigma_{22}^{-1}\mathbf{1}_{N-k})^{-1}}\right) = T-N+k+1,$$

i.e.,

$$E\left(\frac{1}{\mathbf{1}'_{N-k}S_{22}^{-1}\mathbf{1}_{N-k}}\right) = \frac{T-N+k+1}{T-1}\frac{1}{\mathbf{1}'_{N-k}\Sigma_{22}^{-1}\mathbf{1}_{N-k}}.$$

This gives the following expression for $\underline{iso}$, the lower bound on the in-sample optimism measure, iso:

$$\underline{iso} = \frac{\left(\frac{1}{\mathbf{1}'_{N-k}\Sigma_{22}^{-1}\mathbf{1}_{N-k}}\right)}{E\left(\frac{1}{\mathbf{1}'_{N-k}\Sigma_{22}^{-1}\mathbf{1}_{N-k}}\right)} = \frac{T-1}{T-N+k+1} \tag{10}$$

To see the effect of nonnegativity constraints, suppose N = 30 and T = 60. Then the ratio of the to the variance of the population global minimum variance portfolio to the average in-sample variance will be (60-1)/(60-30+1) = 1.90. When there are nonnegativity constraints and k = 20, the ratio will be (60-1)/(60-30+20+1) = 1.15. Hence constraints bring the average in-sample variance of the sample global minimum variance portfolio and the variance of the global minimum variance portfolio closer.

The in-sample optimism measure, iso is strictly larger than the lower bound, $\underline{iso}$. Hence scaling up the in-sample variance by a factor that is related to the degrees of freedom associated with the Wishart distribution may not be sufficient to get a reasonable estimate of the out of sample variance.

Table 1 gives the in-sample standard deviation of average variance of the sample global minimum variance portfolio, the standard deviation of the population global minimum variance portfolio and the standard deviation of average variance of the sample global minimum variance portfolio. The simulations correspond to a time series length of 750 return observations (can be thought of as 3 years of daily return data). The parameters for the simulation are done following Jagannathan and Ma (2003). When the number of stocks N is 60, the in-sample standard deviation is 0.419; the standard deviation of the population global minimum variance portfolio is 0.436; and the out-of-sample standard deviation is 0.453. In this case the $\underline{iso}$ is $\frac{0.436}{0.419}^2 = 1.08$ and iso is $\frac{0.453}{0.419}^2 = 1.17$, i.e., the two are rather close. However, when $N = 360$ the corresponding numbers are 1.93 and 3.65.

The in-sample standard deviation when multiplied by the degrees of freedom related scale factor (column 5 in Table 1A) is 0.437 which is 0.96 times the out-of-sample standard deviation. The degrees of freedom correction works rather well for relatively smaller values of N. When N = 360 however, the corresponding numbers are 0.189 and 0.260, i.e., scaling up the in-sample variance by the degrees of freedom related factor still gives a rather optimistic estimate of the out-of-sample variance. In this case $\underline{iso} << iso$. This suggests that a degrees of freedom correction may not be adequate to estimate the out-of-sample variance of the global minimum variance portfolio sufficiently accurately.

Note that when we impose portfolio weight constraints the minimization is done over a smaller set of feasible parameter values. In the extreme case there may only be one set of values for the parameters that would satisfy the constraints. Hence we may expect the in-sample and out-of-sample variances of the sample global minimum variance portfolio as well as the variance of the population global minimum variance portfolio to be closer to each other when we impose portfolio weight constraints. The results in Table 1 (1A, 1B, and 1C) confirm that this intuition.

While imposing portfolio weight constraints improve our ability to assess the out-of-sample variance more accurately, the unconstrained variance minimization problem is still of interest since imposing portfolio weight constraints can increase the minimum risk by a substantial amount. For example, the ratio of out-of-sample to in-sample-variance for the sample covariance matrix, comes down from $\frac{0.260}{0.136}^2 = 3.65$ in Panel A to $\frac{0.413}{0.395}^2 = 1.09$ in Panel C, a substantial reduction in the in sample optimism. However, the out-of-sample standard deviation of the sample global minimum variance portfolio is 0.637 for N = 360. The corresponding number is 0.859 when both nonnegativity and upper bound constraints are imposed, i.e., constraints increase the average variance of the sample minimum variance portfolio by $(0.413/0.260)^2 - 1 = 152\%$. Hence there is a need for examining alternative estimators for the out-of-sample variance of minimum risk portfolios. With this objective we examine the Bayesian estimator of the covariance matrix and the Jackknife type estimator for the out-of-sample variance in the next two sections that follow.

3 In-sample optimism with Bayes' estimator of the covariance matrix

We assume, as in the earlier section, that returns are drawn from an i.i.d multivariate Normal distribution, and the portfolio manager has observations on a time series of returns of length T. For convenience we will consider the global minimum variance (or minimum tracking error variance) portfolio, and ignore portfolio weight constraints by assuming that they are not binding.

Let μ, denote the vector of mean returns and Σ denote their covariance matrix. The portfolio manager does not know the value of μ and Σ, but we as the designer of the simulation exercise do. The sample mean is $\hat{\mu}$. For convenience we use S to denote the MLE of the covariance matrix. The portfolio manager has a diffuse prior about μ and Σ, given by:

$$p(\mu, \Sigma) \propto |\Sigma|^{-(N+1)/2}.$$

We assume that the Bayesian portfolio manager's objective is to choose portfolio weights w to minimize the variance of the portfolio's return in the next period, i.e.,

$$\mathrm{var}(w' R_{T+1}),$$

where R_{T+1} is the next period's return, and $\mathrm{var}(\cdot)$ is the variance under the predictive distribution.

It is well-known that the predictive distribution for R_{T+1} is multivariate Student t (Zellner (1971)p.235-236) given by:

$$\begin{aligned} & p(R_{T+1}|R_1, ..., R_T) && (11)\\ \propto\ & \left[1 + \frac{T}{T+1}(R_{T+1} - \hat{\mu})(TS)^{-1}(R_{T+1} - \hat{\mu})'\right]^{-T/2} && (12)\\ \propto\ & \left[(T-N) + (R_{T+1} - \hat{\mu})(\frac{T+1}{T-N}S)^{-1}(R_{T+1} - \hat{\mu})'\right]^{-(N+(T-N))/2} && (13)\end{aligned}$$

with parameters $(\hat{\mu},\ (\frac{T+1}{T-N}S)^{-1},\ T-N,\ N)$, using Zellner's notation (Zellner (1971), p.383). The $\propto$ sign means positively proportional. So the posterior covariance matrix of R_{T+1} is given by:

$$\frac{T-N}{T-N-2}\ \frac{T+1}{T-N}\ S = \frac{T+1}{T-N-2}\ S.$$

Since the posterior covariance matrix is proportional to S, the global minimum portfolio weights, $w_s B$ would be the same as in the Classical case where S is used as the covariance matrix, w_s, i.e.,:

$$w_{sB} = \frac{1}{\mathbf{1}' S^{-1} \mathbf{1}} S^{-1}\mathbf{1} = w_s$$

The variance of the portfolio using the predictive distribution is given by:

$$\begin{aligned} & w_s' \mathrm{var}(R_{T+1}) w_s && (14)\\ =\ & \frac{1}{(\mathbf{1}' S^{-1}\mathbf{1})^2}\mathbf{1}' S^{-1}\frac{T+1}{T-N-2}\ S S^{-1}\mathbf{1} && (15)\\ =\ & \frac{T+1}{T-N-2}\ \frac{1}{\mathbf{1}' S^{-1}\mathbf{1}}. && (16)\end{aligned}$$

Notice that in the above discussion, we use S to denote the MLE of the covariance matrix, which is $\frac{T-1}{T}$ times the sample covariance matrix. Adjusting for this scaling factor, we find that the variance of the minimum variance portfolio under the predictive distribution is:

$$\frac{(T-1)(T+1)}{T(T-N-2)} \frac{1}{\mathbf{1}' S^{-1} \mathbf{1}},$$

where S is the sample covariance matrix (from now on S denotes the sample covariance matrix). Clearly, this is larger than the in-sample global minimum variance, $\frac{1}{\mathbf{1}' S^{-1} \mathbf{1}}$ by the factor, $\frac{(T-1)(T+1)}{T(T-N-2)}$

In the earlier section we showed that, under the objective probability measure (when portfolio weight constraints are not binding),

$$E\left(\frac{1}{\mathbf{1}' S^{-1} \mathbf{1}}\right) = \frac{T-N+1}{T-1} \frac{1}{\mathbf{1}' \Sigma^{-1} \mathbf{1}}.$$

So the expected value (averaged in the objective probability measure) of the Bayesian estimate of the achieved global minimum variance is

$$\frac{(T-1)(T+1)}{T(T-N-2)} \frac{T-N+1}{T-1} \frac{1}{\mathbf{1}' \Sigma^{-1} \mathbf{1}} = \frac{T+1}{T} \frac{T-N+1}{T-N-2} \frac{1}{\mathbf{1}' \Sigma^{-1} \mathbf{1}}, \quad (17)$$

which is slightly larger than the population global minimum variance of

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

We showed in the last section that the population global minimum variance is less than the expected out-of-sample variance (under the objective probability measure), i.e.,:

$$\frac{1}{\mathbf{1}' \Sigma^{-1} \mathbf{1}} < E\left(\frac{\mathbf{1}' S^{-1} \Sigma S^{-1} \mathbf{1}}{(\mathbf{1}' S^{-1} \mathbf{1})^2}\right), \quad (18)$$

Hence, on average, the variance of the global minimum variance portfolio under the predictive distribution will be smaller than its out-of-sample variance, provided, the difference between the right side and the left side of the of the above is larger than the scale factor $\frac{T+1}{T} \frac{T-N+1}{T-N-2}$,.

In fact, we may expect the difference between the two sides of (18) to increase when N gets larger for a fixed T. This is because in that case, the portfolio weighs constructed using sample estimates will become more unreliable, and the Jensen's inequality effect, that is, the difference between $E(w_s \Sigma w_s)$ and $w_p' \Sigma w_p$, will get large. So scaling $\frac{1}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}$ by the factor $\frac{T+1}{T}\ \frac{T-N+1}{T-N-2}$, is unlikely to be insufficient to compensate for the increased Jensen's inequality effect.

Column 6 of Table 1 gives the variance of the sample minimum variance portfolio under the predictive distribution. As can be seen, the average standard deviation under the predictive distribution is very close to the standard deviation of the population global minimum variance portfolio. However, as the number of stocks become larger the latter substantially understates the population variance of the sample global minimum variance portfolio.

4 Jackknife type estimator of the out-of-sample variance

As before we assume that the portfolio manager has time series data for T periods in the immediate past, and estimates the covariance matrix using these T observations. Let S denote this estimate and w_T denote the portfolio weights formed using this estimated covariance matrix. The return on this portfolio during period $T+1$ is $w_T' R_{T+1}$. Consider the following estimator for $Var\left(w_T' R_{T+1}\right)$.

Suppose we estimate the covariance matrix by dropping the i'th observation. For each i in the interval $[1, T-1]$, estimate the covariance matrix $S_{(T-1,\,i)}$. using the data $\{R_1, \ldots, R_{i-1}, R_{i+1}, \ldots, R_T\}$. Construct the global minimum variance portfolio using $S_{(T-1,i)}$. Clearly, the weights of the global minimum variance portfolio will be a function of $S_{(T-1,i)}$. Call it $w_{(T-1),i}$, which is a N vector. Note that $w_{(T-1),i}' R_i$ has the same distribution as that of $w_T'\, R_{T+1}$ for large enough T, if returns are drawn from an i.i.d distribution.

A natural estimator of the out-of-sample variance, $Var\,(w'_T R_{T+1})$ would be the he sample variance of $[w'_{(T-1,\,1)}R_1,\ w'_{(T-1,\,2)}R_2,\ ...,\ w'_{(T-1,\,T-1)}R_T]$.

We can relax the i.i.d assumption with the assumption that returns have the *exchangeability* property, i.e., for any fixed positive integer k, joint density of $\{R_1, \ldots, R_k\}$, say, $f(r_1, \ldots, r_k) = f(r_{\sigma(1)}, \ldots, r_{\sigma(k)})$ for all permutation $\sigma : \{1, \ldots, k\} \to \{1, \ldots, k\}$). Hence for every i, $S_{(T-1,\,i)}$ will have the same distribution as S_T for moderately large T, and so each $w_{(T-1,\,i)}$ will have approximately the same distribution as w_T. Also, each R_i has the same distribution as R_{T+1}, being stationary, and furthermore, joint distribution of $w_{(T-1,\,i)}$ and R_i will be almost the same as that of w_T and R_{T+1}. Therefore, each $w'_{(T_1,\,i)}R_i$ will have approximately the same distribution as $w'_T R_{T+1}$. So we can use the sample variance of the former to estimate the variance of the latter.

While the exchangeability assumption appears rather restrictive, the empirical evidence we present in subsequent show that the estimator derived based on that assumption performs surprisingly well.

Ma and Jagannathan (2003) show that using daily return data instead of monthly return data improves the performance of minimum risk portfolios. The empirical findings reported Liu (2003)suggest that there may not be much advantage to using higher frequency return data (higher than daily) if the holding period is six months or longer when a year or more of historical daily return data is available. In view of this, in our empirical analysis we form minimum risk portfolios based on covariance matrices estimated using past three years of daily return data and hold the portfolio for six months. At the end of six months we recompute the portfolio weights based on the most recent six months history of daily returns. In what follows we modify the estimator of the out of sample variance we discussed earlier in this section to take this feature into account and allow for the holding period to be longer than the interval over which returns are measured. We will denote the Jackknife-type estimator as the *Jackknife* estimator for convenience.

4.1 Jackknife-type estimator

We now assume that returns are measured more frequently, say once a day for convenience, but the portfolio manager revises the holdings once every few days, say once a month for expositional convenience (could be six months as in our empirical study that follows), with there being l days in a month. Let T denote the number of days of daily return data used to compute the covariance matrix, S_T, and the portfolio weights, w_T; and $m = T/l$ denote the number of months of observation. For convenience we assume that T is an integral multiple of l. The return on this portfolio for the l days following T are: $(w'_T R_{T+1} \dots w'_T R_{T+l})$. Let the sample variance of the sequence, $(w'_T\, R_{T+1}.\dots.w'_T\, R_{T+l})$, be denoted by $Var(w'_T R_{T+l})$, for convenience.

We assume that monthly returns satisfy the exchangeability property. Then, for each month $1 \leq i \leq m$, we estimate the covariance matrix $S_{(T-l,\,i)}$. by deleting the return data for the l days in month i and construct the global minimum variance portfolio using $S_{(T-l,i)}$,. Let the corresponding portfolio weights be denoted, $w_{(T-l,i)}$, which is a N vector. Compute the sample variance of the sequence, of l returns on the efficient portfolio, $\left(w'_{T-l,i}\, R_{(i-1)l+1}.\dots.w'_{T-l,i}\, R_{il}\right)$, for each i, $i = 1 \dots m$. Let $Var(w'_T R_{T+l}, i)$ denote this sample variance. Then the average of the m such variances we compute provides and estimate of the expected value of the out of sample variance, $Var(w'_T R_{T+l})$, that we are interested in.

The intuition is that each $S_{(T-l,\,i)}$ will have the same distribution as S_T, for moderately large T, and so each $w_{(T-1,\,i)}$ will have approximately the same distribution as w_T. Hence, each sequence, $\left(w'_{T-l,i}\, R_{(i-1)l+1} \dots w'_{T-l,i}\, R_{il}\right)$ will have approximately the same distribution as the sequence, $(w'_T\, R_{T+1}.\dots.w'_T\, R_{T+l})$. Hence the Jackknife type estimator given by,

$$Jackknife = \frac{1}{m}\sum_{i=1}^{m} Var(w'_T R_{T+l}, i) \tag{19}$$

provides a consistent estimate of the out of sample variance for moderately large values of T. We report the standard deviation of *Jackknife* since standard deviations are more often used in practice.

5 Empirical Evaluation of the Estimators

The Jackknife estimator we derived assumes that returns are i.i.d (or satisfy the exchangeability condition). However, for evaluating the estimator using data, we need a benchmark estimator that is valid under more general conditions, but may require a much longer time series of data. For that purpose we use the *rolling window benchmark* described in the next subsection. Our discussion of the rolling window benchmark is based on Hall and Jing (1996) and Hall, Jing and Lahiri (1998).

5.1 Rolling Window Estimator

We assume that the portfolio manager has a long time series of data, of length $\mathbf{T}$, available for evaluating different estimators, where $\mathbf{T}$ much larger than T, the number of observations used to form the covariance matrix, S used in constructing efficient portfolios. We consider the general case where the manager uses T daily return observations to form portfolios and holds them for T_1 days. When the holding period is one month, $T_1 = l$, the number of days in a month. When the holding period is one day, $T_1 = 1$. We use the rolling window method described below as the benchmark, that provides a consistent estimator of the out of sample variance, i.e., the expected value of the sample variance of the sequence $w'_T\, R_{T+1} \ldots w'_T\, R_{T+T_1}$, under the weaker assumption that the return generating process is stationary and ergodic. For convenience, we denote the variance of the sequence, $w'_T\, R_{T+1} \ldots w'_T\, R_{T+T_1}$, by $Var(w'_T R_{T+T_1})$

Proposition. Let $\{R_t\}_{1 \leq t \leq \mathbf{T}}$ be the $\mathbf{T}$ number of (daily) returns of the N stocks. Assume that $\{R_t\}$ is a stationary and ergodic process. Let $T,\ T_1$ be positive integers denoting the number of observations used to form efficient portfolios and the length of the holding period respectively. Define $T_2 = \mathbf{T} - T - T_1 + 1$. Let minimum variance portfolio weights be estimated from $\{R_1, R_2, \ldots, R_T\}$ and call it $\hat{w}_1$. Similarly, for $1 \leq i \leq T_2$, $\hat{w}_i$ is estimated from $\{R_i, R_{i+1}, \ldots, R_{T+i-1}\}$. Let let $\hat{\Sigma}_1$ denote the sample covariance matrix of the sequence $\{R_{T+1}, R_{T+2}, \ldots, R_{T+T_1}\}$. For $1 \leq i \leq T_2$, let $\hat{\Sigma}_i$ denote the sample covariance matrix of $\{R_{T+i}, R_{T+i+1}, \ldots, R_{T+T_1+i-1}\}$. Then

$$\frac{\hat{w}_1' \hat{\Sigma}_1 \hat{w}_1 + \cdots + \hat{w}_{T_2}' \hat{\Sigma}_{T_2} \hat{w}_{T_2}}{T_2}$$

converges to the expected (population) out-of-sample variance $E\left(Var(w_T' R_{T+T_1})\right)$, as $T_2 = \mathbf{T} - T - T_1 + 1 \rightarrow \infty$.

Remark. Observe that for $1 \leq i \leq T_2$, $\hat{w}_i' \hat{\Sigma}_i \hat{w}_i$ is nothing but the sample variance of the portfolio returns $\{\hat{w}_i' R_{T+i}, \ldots, \hat{w}_i' R_{T+T_1+i-1}\}$.

Also, it is important to note that moderately large T is essential to make sample covariance matrix positive definite.

Proof. Under the assumption of stationarity and ergodicity we observe that

$$\left|\frac{\hat{w}_1' \hat{\Sigma}_1 \hat{w}_1 + \cdots \hat{w}_{T_2}' \hat{\Sigma}_{T_2} \hat{w}_{T_2}}{T_2} - E(\hat{w}_1' \hat{\Sigma}_1 \hat{w}_1)\right| \tag{20}$$

converges to 0 as $T_2 \rightarrow \infty$.

Remark Suppose $\frac{\mathbf{T}}{T+T_1}$ is a sufficiently large. For expositional convenience we assume that $\frac{\mathbf{T}}{T+T_1} = m$, an integer. Consider using the first T days of observations to estimate the covariance matrix and construct the minimum variance portfolio. Hold the portfolio for T_1 days and compute the sample variance of the T_1 daily returns on the minimum variance portfolio so constructed. At the end of $T + T_1$ days use the immediately preceding T observations to re-estimate the covariance matrix, form the minimum variance portfolio, and hold it from day $T + T_1 + 1$ to day $T + 2T_1$. Compute the sample variance of the T_1 returns on the minimum variance portfolio from day $T + T_1 + 1$ to day $T + 2T_1$. Repeating this process gives m out-of-sample variances. Then the average of the m out-of-sample variances provides a consistent estimate of $Var(w'_T R_{T+T_1})$. In out empirical exercise we will use this version of the rolling window estimator as the benchmark.

5.2 Data and Methodology

In this section we empirically evaluate the performance of the Jackknife estimate of the out-of-sample variance of sample global minimum variance and minimum tracking error variance portfolios formed using a number of covariance matrix estimators. We use daily return data for stocks traded on NYSE, AMEX and NASDAQ for the period May 1964 to April 1999. We start with the last day of April 1967. We choose 200 stocks with the largest market capitalization from all common domestic stocks traded on the NYSE and the AMEX, and with monthly return data for all the immediately preceding 36 months. When a daily return is missing, the equally weighted market return of that day is used instead. We then estimate the covariance matrix of returns on the 200 stocks.

When variance minimization is the objective, we form three global minimum variance portfolios using each covariance matrix estimator. The first portfolio is constructed without imposing any restrictions on portfolio weights, the second is subject to the constraint that portfolio weights should be non-negative, and the third, in addition, faces the restriction that no more than two percent (i.e., 10 times of the equal weight) of the investment can be in any one stock. Each of these portfolios are held for 6 months. Their daily returns are recorded, and at the end of six months, the same process is repeated. This gives time series of post-formation daily returns for each of the 64 non-overlapping six month intervals during the period May 1967 to April 1999 for each covariance matrix estimator. We use the average of the 64 daily return variances as the benchmark i.e., the expected out of sample variance. For convenience of interpretation, we report the square root of the average variance (as the standard deviation) instead of the variance.

For tracking error minimization, following Chan, Karceski, and Lakonishok (1999) we assume the investor is trying to track the return of the S&P 500 index. As in the case of portfolio variance minimization, we construct three tracking error minimizing portfolios for each covariance estimator. Notice that constructing the minimum tracking error variance portfolio is the same as constructing the minimum variance portfolio using returns in excess of the benchmark, subject to the restriction that the portfolio weights sum to one.

5.3 Covariance Matrix Estimators

The first estimator is the sample covariance matrix:

$$S_N = \frac{1}{T-1} \sum_{t=1}^{T} (h_t - \bar{h})(h_t - \bar{h})',$$

where T is the sample size, h_t is a $N \times 1$ vector of stock returns in period t, and $\bar{h}$ is the average of these return vectors.

The second estimator assumes that returns are generated according to Sharpe's (1963) one-factor model given by:

$$r_{it} = \alpha_i + \beta_i r_{mt} + \epsilon_{it},$$

where r_{mt} is the period t return on the value-weighted portfolio of stocks traded on the NYSE, AMEX, and Nasdaq. Then the covariance estimator is

$$S_1 = s_m^2 BB' + D. \tag{21}$$

Here B is the $N \times 1$ vector of β's, s_m^2 is the sample variance of r_{mt}, and D has the sample variances of the residuals along the diagonal and zeros elsewhere.

.

The third estimator is the Fama and French (1993) three-factor model.

We also report the results for Bayesian estimators of the covariance matrices with and without exact factor structure. The Bayesian estimator of the covariance matrix when returns have a factor structure is described in the appendix.

5.4 Simulations

In order to examine how the different estimators would perform when returns are i.i.d multivariate Normal, we generate return data using simulations. For that purpose, every time we read in the returns on the 200 stocks over the 3 year estimation period, we also read in corresponding factor and benchmark returns. We then estimate the betas for all the 200 stocks with respect to the factors, calculated the residual variances, and the covariance matrix of the factors; generate 200 residual returns and 3 factor returns for each of the days in the 3 years assuming i.i.d multivariate Normal distribution; compute the returns on the 200 stocks using the estimated factor betas; form minimum risk portfolios for each covariance matrix estimator; and generate six months of out of sample daily returns. Hence we apply the same procedure to both real data and data generated using simulation.

5.5 Empirical Results

Table 2 panels A and C give the result for the global minimum variance minimum tracking error variance portfolios respectively. When returns have an exact three factor structure and are i.i.d multivariate Normal (Panel A) the in-sample-variance accurately reflect the out of sample variance provided a three factor model is used to estimate covariance matrices, as is to be expected. However, when the sample covariance matrix or the one factor model based covariance matrix estimator is used to construct the minimum variance portfolios, the in-sample variance is substantially smaller than its out of sample counter part. Neither the use of the degrees of freedom based correction nor the use of the Bayesian covariance matrix lead to a sufficient reduction in the in-sample optimism. In all cases the Jackknife type estimator of the out of sample variance is reasonably accurate. The out of sample standard deviation of the global minimum variance portfolio constructed using sample information is the least for the three factor model. Imposing the one factor structure in this case worsens the performance when compared to using the sample covariance matrix.

When portfolio weights are constrained to be nonnegative, the out of sample variances of the minimum variance portfolios associated with the three covariance matrix estimators are about the same as observed by Jagannathan and Ma (2003). However, the in-sample optimism comes down substantially with portfolio weight constraints. While the use of constraints impose a penalty – the out of sample variance goes up – the investor is able to more accurately assess the out of sample variance.

Panel C of Table 2 gives the results when daily return data on stocks traded on NYSE, AMEX and NASDAQ are used. The sample covariance matrix outperforms the three (and one) factor models – the out of sample variance as well as the in-sample optimism are the least for the sample covariance matrix. This indicates that an exact three factor structure may have too large a specification error to describe the daily return data. As is to be expected, with portfolio weight constraints the in-sample optimism comes down substantially. As is the case for simulated data, neither the degrees of freedom correction nor the use of the Bayesian estimator of the covariance matrix help much in reducing the in-sample optimism. While the Jackknife estimator performs better. However, unlike with simulated data, the Jackknife too exhibits some in-sample optimism, suggesting the non-i.i.d nature of the dominant factor in daily returns.

Panel B gives the results for tracking error variance minimization with simulated data. The patterns are similar to the variance minimization case. The in-sample optimism is most for the sample covariance matrix. Jackknife estimates are accurate. The three factor model performs the best and there is no in-sample optimism in that case. However, imposing portfolio weight constraints has little effect.

Panel D gives the results for tracking error minimization with daily return data. The patterns are quite different from what we observed when simulated data was used. The sample covariance matrix provides the lowest out of sample tracking error variance. However, the it has the most in-sample optimism. The out of sample standard deviation of the tracking error for the three factor model is 1.27 times that for the sample covariance matrix. Hence there is substantial benefit to using the sample covariance matrix if tracking error minimization is the objective. This is true with or without portfolio weight constraints. The in-sample optimism does not go away with the use of the three factor model. With the use of factor models, however, the Jackknife type estimator provides a rather accurate estimate of the out of sample variance – it understates the out of sample variance to some extent when the sample covariance matrix is used.

These results suggest that while the Jackknife type estimator (based on the assumption that returns are i.i.d) performs rather surprisingly well, there is scope for improvement through relaxing the i.i.d assumption.

Table 3, Panel A reports the results for the global minimum variance portfolios constructed using monthly return data on 25 size and book/market sorted portfolio made available by French. As can be seen, the Jackknife estimator provides a better assessment of the out of sample variance in all cases. Panel B of Table 3 provides the results for the tangent portfolios. In this case it is not possible to estimate the out of sample mean or or the variance of the tangent portfolio with much accuracy. However, when portfolio weight constraints are in place the Jackknife as well as the in-sample estimates provide a rather accurate assessment of the out of sample variance of the tangent portfolio. In almost all cases the global minimum variance portfolio has a higher out of sample Sharpe Ratio than the Tangent portfolio. This suggests the need for estimators of expected returns that are more precise than the sample mean.

6 Summary

It is a well documented fact that the in-sample value of the variance of a mean variance efficient portfolio typically understates its population variance, i.e., there is *in-sample optimism*. In this paper we theoretically show that this is what one should expect to find.

We obtain a lower bound for in-sample optimism for global minimum variance and minimum tracking error variance portfolios constructed using sample moments when there are portfolio weight constraints. Using simulation methods we show that the in-sample optimism can be large. We find that there is substantial in-sample optimism even with Bayesian covariance matrix estimators under standard diffuse priors when the number of assets are relatively large.

We develop a Jackknife type estimator for the out-of-sample variance of minimum risk portfolios that is valid when returns are identically and independently distributed through time or when they have the exchangeability property. Our empirical analysis indicates that this estimator provides a more accurate estimate of the variance of minimum risk portfolios even though the data does not support the i.i.d assumption.

Appendix

Bayes' estimator of the covariance matrix that has a factor structure

In this appendix we derive the Bayes' estimator of the covariance matrix when it has an exact k factor structure. Let R_t, denote the $N \times 1$ vector of date t returns as before, and f_t, denote the $k \times 1$ vector of date t factors. We assume that returns and factors together are drawn from an i.i.d multivariate Normal distribution and the manager has observations on a time series of T returns and factors.

Let the true mean return vector be μ_r, the true covariance matrix of returns be Σ_r. We use subscript f to denote the corresponding parameters of the vector of factors.

We assume that returns have the following factor structure:

$$R_t = \alpha + \beta f_t + U_t, \tag{A1}$$

where α is a $N \times 1$, vector, β is a $N \times k$, matrix, and U_t is $N \times 1$ vector of residuals that have a multivariate normal distribution with mean zero and a diagonal covariance matrix Ω.

Let R denote the $T \times N$ matrix of returns. Then equation (A1)can be written in matrix form as follows:

$$R = FB + U,$$

where

$$F = \begin{pmatrix} 1 & f_{11} & ... & f_{1k} \\ ... & ... & ... & ... \\ 1 & f_{T1} & ... & f_{Tk} \end{pmatrix}$$

$$B = \begin{pmatrix} \alpha' \\ \beta' \end{pmatrix}$$

and U is defined similarly.

As before, we assume that the Bayesian portfolio manager does not observe $\mu_f, \Sigma_f, \Omega, \alpha$, and β but believes that they are drawn from a prior distribution. From the posterior distribution for $\mu_f, \Sigma_f, \Omega, \alpha$, and β, the portfolio manager computes Σ_r, the covariance matrix of R_t under the predictive distribution. Our derivation below follows that of Wang (2000). But our prior specifications are somewhat different, as will be pointed out below.

Let the prior be the usual diffuse prior:

$$p(\mu_f, \Sigma_f, \alpha, \beta, \Omega) \propto |\Omega|^{-1} |\Sigma_f|^{-(k+1)/2}. \tag{A2}$$

Notice that since Ω is diagonal, $|\Omega|$ is the product of its diagonal elements. Therefore the prior for $|\Omega|$ is a product of the independent priors for its diagonal elements, with the ith one given by:

$$p(\Omega_{ii}) \propto \frac{1}{\Omega_{ii}},$$

which is the commonly used diffuse prior. This prior is equivalent to the following prior for the square roots of the diagonal elements:

$$p(\Omega_{ii}^{1/2}) \propto \frac{1}{\Omega_{ii}^{1/2}}.$$

In what follows we will adopt this latter parameterization for convenience.

Our prior specification (A2) differs from that of Wang's (2000) in two aspects: First, we assume Ω is diagonal while Wang does not. Second, we start with a diffuse prior for α, while Wang assumes that the investor believes the asset pricing model to certain degree, i.e., the prior for α is Normal centered at the zero vector.

Given our prior, the posterior distributions are standard: $p(\Sigma_f|F)$ is inverted Wishart, $p(\mu_f|F, \Sigma_f)$ is Normal, $p(\Omega_{ii}^{1/2}|F, R)$ is inverted Gamma, and $p(B_i|F, R, \Omega_{ii})$ is Normal, as given below:

$$p(\Sigma_f|F) \propto |\Sigma_f|^{-(T+k)/2} \exp\left\{-\frac{T}{2}\text{tr}(\hat{\Sigma}_f\Sigma_f^{-1})\right\} \tag{A3}$$

$$p(\mu_f|F, \Sigma_f) \propto |\Sigma_f|^{-1/2} \exp\left\{-\frac{T}{2}\text{tr}(\mu_f - \hat{\mu}_f)(\mu_f - \hat{\mu}_f)'\Sigma_f^{-1}\right\} \tag{A4}$$

$$p(\Omega_{ii}^{1/2}|F, R) \propto \Omega_{ii}^{-(\nu+1)/2} \exp\left(-\frac{\nu S_{ii}}{2\Omega_{ii}}\right), \quad i = 1, 2, ..., N. \tag{A5}$$

$$p(B_i|F, R, \Omega_{ii}) \propto \Omega_{ii}^{-1/2} \exp\left(-\frac{(B_i - \hat{B}_i)'F'F(B_i - \hat{B}_i)}{2\Omega_{ii}}\right), \forall i. \tag{A6}$$

where $\hat{\Sigma}_f$ and $\hat{\mu}_f$ refer to the maximum likelihood estimates of the respective quantities, B_i is the ith column of B, $\hat{B}_i$ is the least squares estimate of B_i, S_{ii} is the least squares estimate of the residual variance in the factor model regression for the return on stock i, (i.e., sum of squared residuals divided by $T - k - 1$) and $\nu = T - k - 1$.

These distribution functions imply that the posterior distribution of (B, Ω) is independent of that of (μ_f, Σ_f), as pointed out by Wang (2000), with the following posterior moments:

$$E(\Sigma_f|F) = \frac{T}{T-k-2}\hat{\Sigma}_f \equiv \tilde{\Sigma}_f \tag{A7}$$

$$E(\mu_f|F) = \hat{\mu}_f \tag{A8}$$

$$Var(\mu_f|F) = \frac{1}{T-k-2}\hat{\Sigma}_f \tag{A9}$$

$$E(\Omega_{ii}|F, R) = \frac{\nu}{\nu-2}S_{ii} \equiv \tilde{\Omega}_{ii} \tag{A10}$$

$$E(B|F, R, \Omega) = \hat{B} \tag{A11}$$

$$Var(vec(B)) = \tilde{\Omega} \otimes (F'F)^{-1} \tag{A12}$$

Since the posterior distributions of (B, Ω) and (μ_f, Σ_f) are independent, the posterior mean of the return covariance matrix is

$$E(\Sigma_r|F, R) = E(\beta\Sigma_f\beta'|F, R) + \tilde{\Omega}.$$

From the law of iterated expectation, we have

$$E(\beta\Sigma_f\beta'|F, R) = \hat{\beta}\tilde{\Sigma}_f\hat{\beta}' + \text{tr}(G\tilde{\Sigma}_f)\tilde{\Omega}$$

where G is the $k \times k$ submatrix in the lower-right corner of $(F'F)^{-1}$.

From the inverse of partitioned matrices, we know that $G = \frac{1}{T}(\hat{\Sigma}_f)^{-1}$, hence the last term in the above equation is given by $\frac{k}{T-k-2}\tilde{\Omega}$.

Hence the posterior mean of the covariance matrix of the returns is

$$E(\Sigma_r|F,R) = \hat{\beta}\tilde{\Sigma}_f\hat{\beta}' + (1 + \frac{k}{T-k-2})\tilde{\Omega} = \hat{\beta}\tilde{\Sigma}_f\hat{\beta}' + \frac{T-2}{T-k-2}\tilde{\Omega}. \quad \text{(A13)}$$

Given the posterior moments for the model parameters, we can calculate the covariance matrix of R_{T+1} under the predictive distribution. From the law of iterated expectations, we have

$$Var(R_{T+1}|R,F) = E(\Sigma_r|R,F) + Var(\mu_r|R,F). \quad \text{(A14)}$$

The first term on the RHS is given in (A13). The second term is

$$\begin{aligned} & Var(\alpha + b\mu_f|F,R) \\ = \; & Var(E[\alpha + \beta\mu_f|\mu_f,R,F]|R,F) + E[Var(\alpha + \beta\mu_f|\mu_f,R,F)|R,F] \\ = \; & \frac{1}{T-k-2}\hat{\beta}\hat{\Sigma}_f\hat{\beta}' + E[Var(\alpha + \beta\mu_f|\mu_f,R,F)|R,F] \quad \text{(A15)} \end{aligned}$$

Since $\alpha + \beta\mu_f = (I_N \otimes (1\ \mu_f'))vec(B)$, it follows from equation (A12) that

$$\begin{aligned} & Var(\alpha + \beta\mu_f|\mu_f,R,F) \quad \text{(A16)} \\ = \; & (I_N \otimes (1\ \mu_f'))(\tilde{\Omega} \otimes (F'F)^{-1})(I_N \otimes (1\ \mu_f'))' \quad \text{(A17)} \\ = \; & \tilde{\Omega} \otimes [(1\ \mu_f')(F'F)^{-1}(1\ \mu_f')'] \quad \text{(A18)} \\ = \; & \rho\tilde{\Omega} \quad \text{(A19)} \end{aligned}$$

where $\rho = (1\ \mu_f')(F'F)^{-1}(1\ \mu_f')'$. The posterior mean of ρ is,

$$E(\rho|F,R) = \text{tr}\left\{(F'F)^{-1}\begin{pmatrix} 1 & \hat{\mu}_f' \\ \hat{\mu}_f & (T-k-2)^{-1}\hat{\Sigma}_f + \hat{\mu}_f\hat{\mu}_f' \end{pmatrix}\right\}.$$

Notice that,

$$F'F = T\begin{pmatrix} 1 & \hat{\mu}_f' \\ \hat{\mu}_f & \hat{\Sigma}_f \end{pmatrix},$$

and

$$\begin{pmatrix} 1 & \hat{\mu}'_f \\ \hat{\mu}_f & (T-k-2)^{-1}\hat{\Sigma}_f + \hat{\mu}_f\hat{\mu}'_f \end{pmatrix} = \begin{pmatrix} 1 & \hat{\mu}'_f \\ \hat{\mu}_f & \hat{\Sigma}_f \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & -\frac{(T-k-3)}{(T-k-2)}\hat{\Sigma}_f + \hat{\mu}_f\hat{\mu}'_f \end{pmatrix}.$$

Hence $E(\rho|F,R)$ is the sum of two terms. The first term is $\frac{k+1}{T}$. To calculate the second term, we need to get the $k \times k$ lower-right submatrix of $(F'F/T)^{-1}$. Using the inverse of partitioned matrix, this submstrix is

$$(\hat{\Sigma}_f - \hat{\mu}_f\hat{\mu}'_f)^{-1} = \hat{\Sigma}_f^{-1} + \frac{1}{1-\hat{\mu}'_f\hat{\Sigma}_f^{-1}\hat{\mu}_f}\hat{\Sigma}_f^{-1}\hat{\mu}_f\hat{\mu}'_f\hat{\Sigma}_f^{-1}.$$

It can be verified that the second term in the *trace* is therefore,

$$-\frac{k(T-k-3)}{T(T-k-2)} + \frac{1}{T(T-k-2)}\frac{\hat{\mu}'_f\hat{\Sigma}_f^{-1}\hat{\mu}_f}{1-\hat{\mu}'_f\hat{\Sigma}_f^{-1}\hat{\mu}_f} \tag{A20}$$

Combining (A13), (A14),(A15), (A19), and (A20), and adding the term $\frac{k+1}{T}$ from above, we find that the posterior variance of R_{T+1} is given by:

$$\begin{aligned} & Var(R_{T+1}|R,F) \\ = \; & \frac{T+1}{T-k-2}\hat{\beta}\hat{\Sigma}_f\hat{\beta}' + \frac{(T+1)(T-2)}{T(T-k-2)}\tilde{\Omega} \\ & + \frac{1}{T(T-k-2)}\frac{\hat{\mu}'_f\hat{\Sigma}_f^{-1}\hat{\mu}_f}{1-\hat{\mu}'_f\hat{\Sigma}_f^{-1}\hat{\mu}_f}\tilde{\Omega}. \end{aligned} \tag{A21}$$

As observed by Wang (2000), the term $\hat{\mu}'_f\hat{\Sigma}_f^{-1}\hat{\mu}_f$ is the square of the highest Sharpe ratio of the frontier spanned by the sample mean $\hat{\mu}_f$ and variance $\hat{\Sigma}_f$. For the U.S. data, it is definitely less than 0.5. So the last term in (A21) is less than $\frac{1}{T(T-k-2)}\tilde{\Omega}$, which is less than $0.0003\tilde{\Omega}$ if $T \geq 60$ and $k \leq 3$. Hence this term can be ignored for practical purposes.

The portfolio manager's problem is to choose portfolio weights, w, to

$$\min_{w} var(w'R_{T+1}).$$

given the predictive covariance matrix of R_{T+1} in (A21).

We can compare the Bayes estimator of the covariance matrix given in equation (A21) with the sample estimate of the covariance matrix under factor model structure. The sample estimate is given by:

$$S_r = \hat{\beta} S_f \hat{\beta}' + S. \tag{A22}$$

The term S_f, the sample covariance matrix of the factors, differs from the first term of $Var(R_{T+1})$ by a the scale factor, $\frac{(T+1)(T-1)}{T(T-k-2)}$. When $k \leq 3$ and $T \geq 60$, this scale factor is less than 1.091. The term, S, the least squares estimate of the (diagonal) residual covariance matrix (i.e., sum of squared-residuals divided by $T-k-1$), differs from the second term of (A21) by the scale factor, $\frac{(T+1)(T-2)}{T(T-k-2)} \frac{T-k-1}{T-k-3}$, which is less than 1.11 if $k \leq 3$ and $T \geq 60$. Therefore, it appears that the difference may be small whether we use the Bayes' estimator or the sample covariance matrix estimator for factor models.

Also, since the first and second terms of the RHS of the above two equations differ by different factors, the portfolio weights constructed from the sample covariance matrix will differ from the one constructed using the Bayes' covariance matrix estimator. However, since the factors are very similar, we expect the difference in the portfolio weights to be small.

Under the Bayes' estimator, the unconstrained global minimum variance portfolio will have a variance of $(\mathbf{1}'[Var(R_{T+1}|F,R)]^{-1}\mathbf{1})^{-1}$.

REFERENCES

Chan, Louis K. C., Jason Karceski, and Josef Lakonishok, 1999, On portfolio optimization: Forecasting covariances and choosing the risk model, *Review of Financial Studies* 12, 937-974.

Efron, Bradley, 1982, *The Jackknife, the Bootstrap, and Other Resampling Plans,* Society for Industrial and Applied Mathematics, Philadelphia, PA.

Elton, Edwin J., and Martin J. Gruber, 1973, Estimating the dependence structure of share prices-Implications for portfolio selection, *Journal of Finance* 28, 1203-1232.

Fama, Eugene, and Kenneth French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Flemming, Jeff, Cris Kirby, and Barbara Ostdiek, 2002, The economic value of volatility timing using "realized" volatility, *Journal of Financial Economics* (forthcoming).

Foster, D. P., and D. B. Nelson, 1996, Continuous record asymptotics for rolling sample variance estimators, *Econometrica* 64, 139-174.

Frost, Peter A., and James E. Savarino, 1988, For better performance: Constrain portfolio weights, *Journal of Portfolio Management* 15, 29-34.

Hall, P. and Jing, B. Y., 1996, On sample reuse methods for dependent data, Journal of Royal Statistical Society Series B, 58, 727-737.

Hall, P., Jing, B. Y. and Lahiri, S. N., 1998, On the sampling window method for long-range dependent data, Statistica Sinica, 8, 4, 1189-1204.

Jacquier, Eric, Alex Kane, and Alan Mrcus, 2002, Optimal forecasts of long term returns and asset allocation: Geometric, arithmetic, or other means? *Working Paper, Boston College.*

Jagannathan, Ravi, and Tongshu Ma, 2002, Risk reduction in large portfolios: Why imposing the wrong constraint helps, *Journal of Finance* (forthcoming).

Jobson, J. D., and Bob Korkie, 1981, Putting Markowitz theory to work, *Journal of Portfolio Management,* 70-74.

Jorion, Philippe, 1986, Bayes-Stein estimation for portfolio analysis, *Journal of Financial and Quantitative Analysis* 21, 279-292.

Künsch, Hans R. 1989, The Jackknife and the Bootstrap for General Stationary Distributions, *Annals of Statistics* 17, 1217-1241.

Lahiri, S. N. 2002, On the Jackknife-After-Bootstrap Method for Dependent Data and Consistency Properties, *Econometric Theory* 18, 79-98.

Ledoit, Olivier, 1996, A well-conditioned estimator for large dimensional covariance matrices, Working paper, UCLA.

Ledoit, Olivier, 1999, Improved estimation of the covariance matrix of stock returns with an application to portfolio selection, Working paper, UCLA.

Liu, Qianqiu, 2003, On portfolio optimization: How do we benefit from high-frequency data? Working paper, Kellogg Graduate School of Management, Northwestern University.

Merton, Robert C., 1980, On Estimating the expected return on the market: An exploratory investigation *Journal of Financial Economics* 8, 323-361.

Michaud, Richard O., 1998, *Efficient Asset Management,* Harvard Business School Press, Boston, MA.

Muirhead, Robb J., 1982, *Aspects of Multivariate Statistical Theory,* John Wiley and Sons, New York.

Sharpe, William, 1963, A simplified model for portfolio analysis, *Management Science* 9, 277-293.

Schwert, G. Williams, and Paul L. Seguin, 1990, Heteroskedasticity in Stock Returns, *Journal of Finance* 45, 1129-1155.

Wang, Zhenyu, 2000, A shrinkage theory of asset allocation using asset-pricing models, Columbia University working paper.

Zellner, Arnold, 1971, *An Introduction to Bayesian Inference in Econometrics,* John Wiley and Sons, New York.

Table 1
In-sample Optimism: Simulation Evidence

Following Jagannathanand Ma (2003) we assume that returns are generated from a two factor model with the factors drawn from independent standard Normal distributions; betas of firms' stocks with respect to the first factor have a Normal distribution with mean 1 and standard deviation 0.4 in the cross section; betas of stocks with respect to the second factor have a Normal distribution with mean 0 and standard deviation 0.2; and residual variances have a LogNormal cross sectional distribution with mean 0.8 and standard deviation 0.7. For each simulatin we randomly generate a time series of 750 observations for the factors and the residual returns and from them construct the time series of stock returns. We construct the global minimum variance portfolio using the sample covariance matrix and compute its in-sample and out-of-sample variances. We repeat the simulations 10 times and report the average numbers. The upper bound when imposed correspond to 10 times equal weighting. We also report the variance estimate obtained by scaling the in-sample estimate by a degrees of freedom (df) factor, the estimate obtained using the Bayesian covariance matrix estimator, and the Jackknife estimate.

Table 1A

(1)	(2)	(3)	(4)	(5)	(6)	(7)		
No. of stocks	Stdev of Population GMV Porfolio	In-sample stdev of sample GMV	Population stev of sample GMV	Column (4) times DF factor	In-sample stdev of sample GMV using Bayesian Covariance Matrix	Jackknife stdev estimate	Lower Bound on Iso	ISO
Panel A. Unconstrained portfolio								
60	0.436	0.419	0.453	0.437	0.437	0.454	1.08	1.17
180	0.267	0.229	0.305	0.263	0.263	0.302	1.36	1.77
360	0.189	0.136	0.260	0.189	0.190	0.262	1.93	3.65
Panel B. Nonnegativity constrained portfolio								
60	0.554	0.548	0.559	0.563		0.560	1.02	1.04
180	0.419	0.405	0.426	0.440		0.422	1.07	1.11
360	0.340	0.332	0.350	0.393		0.353	1.05	1.11
Panel C. Nonnegativity and upper bound constrained portfolio								
60	0.565	0.559	0.570			0.571	1.02	1.04
180	0.455	0.442	0.461			0.455	1.06	1.09
360	0.405	0.395	0.413			0.412	1.05	1.09

Table 1B

Normalized by Population Standard Deviation of the Sample GMV Portfolio

(1) No. of stocks	(2) Stdev of Population GMV Porfolio	(3) In-sample stdev of sample GMV	(4) Population stev of sample GMV	(5) Column (4) times DF factor	(6) In-sample stdev of sample GMV using Bayesian Covariance Matrix	(7) Jackknife stdev estimate
Panel A. Unconstrained portfolio						
60	0.96	0.92	1.00	0.96	0.96	1.00
180	0.88	0.75	1.00	0.86	0.86	0.99
360	0.73	0.52	1.00	0.73	0.73	1.01
Panel B. Nonnegativity constrained portfolio						
60	0.99	0.98	1.00	1.01		1.00
180	0.98	0.95	1.00	1.03		0.99
360	0.97	0.95	1.00	1.12		1.01
Panel C. Nonnegativity and upper bound constrained portfolio						
60	0.99	0.98	1.00			1.00
180	0.99	0.96	1.00			0.99
360	0.98	0.96	1.00			1.00

Table 2
Empirical
Evidence

Table 2A

Estimates made using simulated data described in text. 36 months of daily data are used to form covariance matrices for computing minimum variance portfolios. The portfolios are rebalanced every 6 months. Each month 200 stocks with the largest market capitalization used for calibrating simulation parameters. Results are based on 8057 days, i.e., 384 months of out of sample returns. Means and standard deviations are annualized and expressed in percentage points. The degrees of freedom adjustment factors are computed using the formula in section 2.1. The Baysian covariance matrix estimators are described in section 3 and in the appendix

	In -sample Std. Dev			Jackknife Std Dev		Out-of-sample	Normalized Std. Dev		Out Std Dev
		DF Adjusted	Bayesian	Mean	Std Dev	Std. Dev	Out/In	Out/Jack	Normalized by Smpl Cov
Smpl Cov	4.60	5.36	5.37	6.27	0.11	6.29	1.37	1.00	1.00
Smpl Cov C	7.71	8.35		8.04	0.21	8.01	1.04	1.00	1.27
Smpl Cov D	7.96			8.25	0.23	8.28	1.04	1.00	1.31
1 Factor	4.70		4.72	6.51	0.19	6.51	1.38	1.00	1.03
1 Factor C	7.35			8.05	0.21	8.00	1.09	0.99	1.27
1 Factor D	7.61			8.25	0.23	8.27	1.09	1.00	1.31
FF 3-factor	5.18		5.20	5.52	0.10	5.56	1.07	1.01	0.88
FF 3-factor C	7.74			8.00	0.21	7.96	1.03	1.00	1.26
FF 3-factor D	7.99			8.22	0.23	8.24	1.03	1.00	1.31
Equally-weighted	13.79			13.50	0.42	13.67	0.99	1.01	2.17

Table 2B

Estimates made using simulated data described in text. 36 months of daily data are used to form covariance matrices for computing minimum tracking error variance portfolios. The portfolios are rebalanced every 6 months. Each month 200 stocks with the largest market capitalization used for calibrating simulation parameters. Results are based on 8057 days, i.e., 384 months of out of sample returns. Means and standard deviations are annualized and expressed in percentage points. The degrees of freedom adjustment factors are computed using the formula in section 2.1. The Baysian covariance matrix estimators are described in section 3 and in the appendix.

	In -sample Std. Dev			Jackknife Std Dev		Out-of-sample	Normalized Std. Dev		Out Std Dev
		DF Adjusted	Bayesian	Mean	Std Dev	Std. Dev	Out/In	Out/Jack	Normalized by Smpl Cov
Smpl Cov	1.42	1.66	1.66	1.95	0.02	1.91	1.34	0.98	1.00
Smpl Cov C	1.44	1.65		1.89	0.02	1.87	1.30	0.99	0.98
Smpl Cov D	1.44			1.89	0.02	1.87	1.30	0.99	0.98
1 Factor	1.62		1.62	1.79	0.02	1.76	1.08	0.98	0.92
1 Factor C	1.62			1.79	0.02	1.76	1.09	0.98	0.92
1 Factor D	1.62			1.79	0.02	1.76	1.09	0.98	0.92
FF 3-factor	1.65		1.65	1.67	0.02	1.65	1.00	0.98	0.86
FF 3-factor C	1.65			1.67	0.02	1.65	1.00	0.98	0.86
FF 3-factor D	1.65			1.67	0.02	1.65	1.00	0.98	0.86
Equally-weighted	2.16			2.40	0.26	2.11	0.98	0.88	1.11

Table 2C

Estimates made using daily data for the period 1964/5 to 1999/4. 36 months of daily data are used to form covariance matrices for computing minimum variance portfolios. The portfolios are rebalanced every 6 months. Each month 200 stocks with the largest market capitalization used for calibrating simulation parameters. Results are based on 8057 days, i.e., 384 months of out of sample returns. Means and standard deviations are annualized and expressed in percentage points. The degrees of freedom adjustment factors are computed using the formula in section 2.1. The Baysian covariance matrix estimators are described in section 3 and in the appendix.

	In -sample Std. Dev			Jackknife Std Dev		Out-of-sample	Normalized Std. Dev		Out Std Dev
		DF Adjusted	Bayesian	Mean	Std Dev	Std. Dev	Out/In	Out/Jack	Normalized by Smpl Cov
Smpl Cov	5.48	6.39	6.40	8.27	0.29	8.98	1.64	1.09	1.00
Smpl Cov C	7.97	8.62		8.75	0.30	9.30	1.17	1.06	1.04
Smpl Cov D	8.24			8.94	0.32	9.40	1.14	1.05	1.05
1 Factor	4.45		4.46	9.12	0.28	9.60	2.16	1.05	1.07
1 Factor C	7.13			8.94	0.29	9.43	1.32	1.05	1.05
1 Factor D	7.40			9.08	0.31	9.48	1.28	1.04	1.06
FF 3-factor	4.91		4.93	8.39	0.26	9.11	1.86	1.09	1.01
FF 3-factor C	7.44			8.81	0.29	9.36	1.26	1.06	1.04
FF 3-factor D	7.70			8.98	0.31	9.43	1.22	1.05	1.05
Equally-weighted	13.61			13.32	0.48	13.57	1.00	1.02	1.51

Table 2 D

Estimates made using daily data for the period 1964/5 to 1999/4. 36 months of daily data are used to form covariance matrices for computing minimum tracking error variance portfolios. The return on the S&P500 is used as the benchmark to be tracked. The portfolios are rebalanced every 6 months. Each month 200 stocks with the largest market capitalization used for calibrating simulation parameters. Results are based on 8057 days, i.e., 384 months of out of sample returns. Means and standard deviations are annualized and expressed in percentage points. The degrees of freedom adjustment factors are computed using the formula in section 2.1. The Baysian covariance matrix estimators are described in section 3 and in the appendix.

	In -sample Std. Dev			Jackknife Std Dev		Out-of-sample	Normalized Std. Dev		Out Std Dev
		DF Adjusted	Bayesian	Mean	Std Dev	Std. Dev	Out/In	Out/Jack	Normalized by Smpl Cov
Smpl Cov	0.93	1.08	1.08	1.33	0.04	1.49	1.61	1.13	1.00
Smpl Cov C	0.94	1.08		1.29	0.03	1.46	1.55	1.13	0.98
Smpl Cov D	0.96			1.32	0.03	1.48	1.53	1.12	0.99
1 Factor	1.59		1.59	2.22	0.05	2.16	1.36	0.97	1.45
1 Factor C	1.59			2.22	0.05	2.16	1.36	0.97	1.45
1 Factor D	1.59			2.22	0.05	2.16	1.36	0.97	1.45
FF 3-factor	1.62		1.63	1.92	0.04	1.89	1.17	0.99	1.27
FF 3-factor C	1.62			1.92	0.04	1.89	1.17	0.99	1.27
FF 3-factor D	1.62			1.92	0.04	1.89	1.17	0.99	1.27
Equally-weighted	2.78			2.99	0.25	2.53	0.91	0.85	1.69
Value-weighted	1.62			1.63	0.07	1.48	0.91	0.91	0.99

Table 3

Global Minimum Variance and Tangency Portfolios of 25 Size and Book/Market Sorted Portfolios

Monthly return data for the period 1963/5 to 2002/4 on the 25 size and book/market sorted portfolios provided by Ken French are used to evaluate the performance of the different estimators. Covariance matrices and expected returns are estimated using 60 months of historical data. Global minimum variance portfolios and Tangent portfolios are constructed based on these estimates and held for 12 months following the estimation period. The procedure is repeated at the end of each holding period. Mean and standard deviations are

	Out of Sample			**In sample**					**Mean Jackknife Estimates**		
Table 3, Panel A: Global variance minimization	Mean	Std. Dev	SR	Mean	Std. Dev	SR	DF Adjusted	Bayesian	Mean	Std. Dev	SR
Monthly smpl cov	11.74	14.85	0.79	9.73	7.60	1.34	9.73	10.25	9.68	13.64	0.71
Monthly smpl cov C	6.85	13.78	0.50	6.41	12.97	0.55	15.09		6.54	13.97	0.47
1-factor model	8.24	16.48	0.50	7.90	8.37	0.92		8.55	7.93	15.86	0.50
1-factor model C	7.29	14.23	0.51	6.74	12.81	0.59			6.70	14.21	0.47
FF-3 factor model	11.07	12.89	0.86	9.00	8.63	1.15		9.00	9.50	12.35	0.77
FF-3 factor model C	6.92	13.94	0.50	6.51	13.01	0.56			6.60	13.97	0.47
Ledoit	9.80	12.45	0.79	8.32	9.36	0.96			8.52	12.23	0.70
Ledoit C	7.04	13.82	0.51	6.55	12.98	0.56			6.66	14.03	0.47

Table 3, Panel B: Tangency portfolios	Out of Sample			In sample					Mean Jackknife Estimates		
	Mean	Std. Dev	SR	Mean	Std. Dev	SR	DF Adjusted	Bayesian	Mean	Std. Dev	SR
Monthly smpl cov	24.03	25.37	0.95	42.31	13.85	2.62			-11.40	236.49	-0.05
Monthly smpl cov C	6.90	16.17	0.43	13.41	16.33	0.88			9.06	16.63	0.54
1-factor model	12.34	39.20	0.31	50.87	19.81	1.82			1054.87	2256.92	0.47
1-factor model C	6.32	16.16	0.39	13.44	15.86	0.91			9.61	16.78	0.57
FF-3 factor model	37.49	54.46	0.69	86.11	32.78	2.51			-38.16	266.97	-0.14
FF-3 factor model C	6.92	16.18	0.43	13.39	16.33	0.88			9.06	16.63	0.59
FF-3 factor model (new)	20.20	37.86	0.53	44.15	44.35	0.96			-82.93	275.48	-0.30
FF-3 factor model (new) C	6.97	16.21	0.43	12.85	16.64	0.83			10.86	16.91	0.64
Ledoit	24.40	46.40	0.53	92.01	43.13	2.11			114.18	825.57	0.14
Ledoit C	6.86	16.16	0.42	13.42	16.25	0.88			9.16	16.64	0.55
Equally Weighted Portfolio	7.49	16.65	0.45	8.08	17.33	0.53			8.08	17.33	0.47

Table 1
In-sample Optimism: Simulation Evidence

Following Jagannathanand Ma (2003) we assume that returns are generated from a two factor model with the factors drawn from independent standard Normal distributions; betas of firms' stocks with respect to the first factor have a Normal distribution with mean 1 and standard deviation 0.4 in the cross section; betas of stocks with respect to the second factor have a Normal distribution with mean 0 and standard deviation 0.2; and residual variances have a LogNormal cross sectional distribution with mean 0.8 and standard deviation 0.7. For each simulatin we randomly generate a time series of 750 observations for the factors and the residual returns and from them construct the time series of stock returns. We construct the global minimum variance portfolio using the sample covariance matrix and compute its in-sample and out-of-sample variances. We repeat the simulations 10 times and report the average numbers. The upper bound when imposed correspond to 10 times equal weighting. We also report the variance estimate obtained by scaling the in-sample estimate by a degrees of freedom (df) factor, the estimate obtained using the Bayesian covariance matrix estimator, and the Jackknife estimate.

Table 1A

(1)	(2)	(3)	(4)	(5)	(6)	(7)		
No. of stocks	Stdev of Population GMV Porfolio	In-sample stdev of sample GMV	Population stev of sample GMV	Column (4) times DF factor	In-sample stdev of sample GMV using Bayesian Covariance Matrix	Jackknife stdev estimate	Lower Bound on Iso	ISO
Panel A. Unconstrained portfolio								
60	0.436	0.419	0.453	0.437	0.437	0.454	1.08	1.17
180	0.267	0.229	0.305	0.263	0.263	0.302	1.36	1.77
360	0.189	0.136	0.260	0.189	0.190	0.262	1.93	3.65
Panel B. Nonnegativity constrained portfolio								
60	0.554	0.548	0.559	0.563		0.560	1.02	1.04
180	0.419	0.405	0.426	0.440		0.422	1.07	1.11
360	0.340	0.332	0.350	0.393		0.353	1.05	1.11
Panel C. Nonnegativity and upper bound constrained portfolio								
60	0.565	0.559	0.570			0.571	1.02	1.04
180	0.455	0.442	0.461			0.455	1.06	1.09
360	0.405	0.395	0.413			0.412	1.05	1.09

Table 1B

Normalized by Population Standard Deviation of the Sample GMV Portfolio

(1)	(2)	(3)	(4)	(5)	(6)	(7)
No. of stocks	Stdev of Population GMV Porfolio	In-sample stdev of sample GMV	Population stev of sample GMV	Column (4) times DF factor	In-sample stdev of sample GMV using Bayesian Covariance Matrix	Jackknife stdev estimate
Panel A. Unconstrained portfolio						
60	0.96	0.92	1.00	0.96	0.96	1.00
180	0.88	0.75	1.00	0.86	0.86	0.99
360	0.73	0.52	1.00	0.73	0.73	1.01
Panel B. Nonnegativity constrained portfolio						
60	0.99	0.98	1.00	1.01		1.00
180	0.98	0.95	1.00	1.03		0.99
360	0.97	0.95	1.00	1.12		1.01
Panel C. Nonnegativity and upper bound constrained portfolio						
60	0.99	0.98	1.00			1.00
180	0.99	0.96	1.00			0.99
360	0.98	0.96	1.00			1.00

Table 2
Empirical
Evidence

Table 2A

Estimates made using simulated data described in text. 36 months of daily data are used to form covariance matrices for computing minimum variance portfolios. The portfolios are rebalanced every 6 months. Each month 200 stocks with the largest market capitalization used for calibrating simulation parameters. Results are based on 8057 days, i.e., 384 months of out of sample returns. Means and standard deviations are annualized and expressed in percentage points. The degrees of freedom adjustment factors are computed using the formula in section 2.1. The Bayesian covariance matrix estimators are described in section 3 and in the appendix.

	In -sample Std. Dev			Jackknife Std Dev		Out-of-sample	Normalized Std. Dev		Out Std Dev
		DF Adjusted	Bayesian	Mean	Std Dev	Std. Dev	Out/In	Out/Jack	Normalized by Smpl Cov
Smpl Cov	4.60	5.36	5.37	6.27	0.11	6.29	1.37	1.00	1.00
Smpl Cov C	7.71	8.35		8.04	0.21	8.01	1.04	1.00	1.27
Smpl Cov D	7.96			8.25	0.23	8.28	1.04	1.00	1.31
1 Factor	4.70		4.72	6.51	0.19	6.51	1.38	1.00	1.03
1 Factor C	7.35			8.05	0.21	8.00	1.09	0.99	1.27
1 Factor D	7.61			8.25	0.23	8.27	1.09	1.00	1.31
FF 3-factor	5.18		5.20	5.52	0.10	5.56	1.07	1.01	0.88
FF 3-factor C	7.74			8.00	0.21	7.96	1.03	1.00	1.26
FF 3-factor D	7.99			8.22	0.23	8.24	1.03	1.00	1.31
Equally-weighted	13.79			13.50	0.42	13.67	0.99	1.01	2.17

Table 2B

Estimates made using simulated data described in text. 36 months of daily data are used to form covariance matrices for computing minimum tracking error variance portfolios. The portfolios are rebalanced every 6 months. Each month 200 stocks with the largest market capitalization used for calibrating simulation parameters. Results are based on 8057 days, i.e., 384 months of out of sample returns. Means and standard deviations are annualized and expressed in percentage points. The degrees of freedom adjustment factors are computed using the formula in section 2.1. The Bayesian covariance matrix estimators are described in section 3 and in the appendix.

	In -sample Std. Dev			Jackknife Std Dev		Out-of-sample	Normalized Std. Dev		Out Std Dev
		DF Adjusted	Bayesian	Mean	Std Dev	Std. Dev	Out/In	Out/Jack	Normalized by Smpl Cov
Smpl Cov	1.42	1.66	1.66	1.95	0.02	1.91	1.34	0.98	1.00
Smpl Cov C	1.44	1.65		1.89	0.02	1.87	1.30	0.99	0.98
Smpl Cov D	1.44			1.89	0.02	1.87	1.30	0.99	0.98
1 Factor	1.62		1.62	1.79	0.02	1.76	1.08	0.98	0.92
1 Factor C	1.62			1.79	0.02	1.76	1.09	0.98	0.92
1 Factor D	1.62			1.79	0.02	1.76	1.09	0.98	0.92
FF 3-factor	1.65		1.65	1.67	0.02	1.65	1.00	0.98	0.86
FF 3-factor C	1.65			1.67	0.02	1.65	1.00	0.98	0.86
FF 3-factor D	1.65			1.67	0.02	1.65	1.00	0.98	0.86
Equally-weighted	2.16			2.40	0.26	2.11	0.98	0.88	1.11

Table 2C

Estimates made using daily data for the period 1964/5 to 1999/4. 36 months of daily data are used to form covariance matrices for computing minimum variance portfolios. The portfolios are rebalanced every 6 months. Each month 200 stocks with the largest market capitalization used for calibrating simulation parameters. Results are based on 8057 days, i.e., 384 months of out of sample returns. Means and standard deviations are annualized and expressed in percentage points. The degrees of freedom adjustment factors are computed using the formula in section 2.1. The Bayesian covariance matrix estimators are described in section 3 and in the appendix.

	In -sample Std. Dev			Jackknife Std Dev		Out-of-sample	Normalized Std. Dev		Out Std Dev
		DF Adjusted	Bayesian	Mean	Std Dev	Std. Dev	Out/In	Out/Jack	Normalized by Smpl Cov
Smpl Cov	5.48	6.39	6.40	8.27	0.29	8.98	1.64	1.09	1.00
Smpl Cov C	7.97	8.62		8.75	0.30	9.30	1.17	1.06	1.04
Smpl Cov D	8.24			8.94	0.32	9.40	1.14	1.05	1.05
1 Factor	4.45		4.46	9.12	0.28	9.60	2.16	1.05	1.07
1 Factor C	7.13			8.94	0.29	9.43	1.32	1.05	1.05
1 Factor D	7.40			9.08	0.31	9.48	1.28	1.04	1.06
FF 3-factor	4.91		4.93	8.39	0.26	9.11	1.86	1.09	1.01
FF 3-factor C	7.44			8.81	0.29	9.36	1.26	1.06	1.04
FF 3-factor D	7.70			8.98	0.31	9.43	1.22	1.05	1.05
Equally-weighted	13.61			13.32	0.48	13.57	1.00	1.02	1.51

Table 2 D

Estimates made using daily data for the period 1964/5 to 1999/4. 36 months of daily data are used to form covariance matrices for computing minimum tracking error variance portfolios. The return on the S&P500 is used as the benchmark to be tracked. The portfolios are rebalanced every 6 months. Each month 200 stocks with the largest market capitalization used for calibrating simulation parameters. Results are based on 8057 days, i.e., 384 months of out of sample returns. Means and standard deviations are annualized and expressed in percentage points. The degrees of freedom adjustment factors are computed using the formula in section 2.1. The Bayesian covariance matrix estimators are described in section 3 and in the appendix.

	In -sample Std. Dev			Jackknife Std Dev		Out-of-sample	Normalized Std. Dev		Out Std Dev
		DF Adjusted	Bayesian	Mean	Std Dev	Std. Dev	Out/In	Out/Jack	Normalized by Smpl Cov
Smpl Cov	0.93	1.08	1.08	1.33	0.04	1.49	1.61	1.13	1.00
Smpl Cov C	0.94	1.08		1.29	0.03	1.46	1.55	1.13	0.98
Smpl Cov D	0.96			1.32	0.03	1.48	1.53	1.12	0.99
1 Factor	1.59		1.59	2.22	0.05	2.16	1.36	0.97	1.45
1 Factor C	1.59			2.22	0.05	2.16	1.36	0.97	1.45
1 Factor D	1.59			2.22	0.05	2.16	1.36	0.97	1.45
FF 3-factor	1.62		1.63	1.92	0.04	1.89	1.17	0.99	1.27
FF 3-factor C	1.62			1.92	0.04	1.89	1.17	0.99	1.27
FF 3-factor D	1.62			1.92	0.04	1.89	1.17	0.99	1.27
Equally-weighted	2.78			2.99	0.25	2.53	0.91	0.85	1.69
Value-weighted	1.62			1.63	0.07	1.48	0.91	0.91	0.99

Table 3

Global Minimum Variance and Tangency Portfolios of 25 Size and Book/Market Sorted Portfolios

Monthly return data for the period 1963/5 to 2002/4 on the 25 size and book/market sorted portfolios provided by Ken French are used to evaluate the performance of the different estimators. Covariance matrices and expected returns are estimated using 60 months of historical data. Global minimum variance portfolios and Tangent portfolios are constructed based on these estimates and held for 12 months following the estimation period. The procedure is repeated at the end of each holding period. Mean and standard deviations are annualized percentage points.

	Out of Sample			**In sample**					**Mean Jackknife Estimates**		
Table 3, Panel A: Global variance minimization	Mean	Std. Dev	SR	Mean	Std. Dev	SR	DF Adjusted	Bayesian	Mean	Std. Dev	SR
Monthly smpl cov	11.74	14.85	0.79	9.73	7.60	1.34	9.73	10.25	9.68	13.64	0.71
Monthly smpl cov C	6.85	13.78	0.50	6.41	12.97	0.55	15.09		6.54	13.97	0.47
1-factor model	8.24	16.48	0.50	7.90	8.37	0.92		8.55	7.93	15.86	0.50
1-factor model C	7.29	14.23	0.51	6.74	12.81	0.59			6.70	14.21	0.47
FF-3 factor model	11.07	12.89	0.86	9.00	8.63	1.15		9.00	9.50	12.35	0.77
FF-3 factor model C	6.92	13.94	0.50	6.51	13.01	0.56			6.60	13.97	0.47
Ledoit	9.80	12.45	0.79	8.32	9.36	0.96			8.52	12.23	0.70
Ledoit C	7.04	13.82	0.51	6.55	12.98	0.56			6.66	14.03	0.47

Table 3, Panel B: Tangency portfolios	Out of Sample			In sample					Mean Jackknife Estimates		
	Mean	Std. Dev	SR	Mean	Std. Dev	SR	DF Adjusted	Bayesian	Mean	Std. Dev	SR
Monthly smpl cov	24.03	25.37	0.95	42.31	13.85	2.62			-11.40	236.49	-0.05
Monthly smpl cov C	6.90	16.17	0.43	13.41	16.33	0.88			9.06	16.63	0.54
1-factor model	12.34	39.20	0.31	50.87	19.81	1.82			1054.87	2256.92	0.47
1-factor model C	6.32	16.16	0.39	13.44	15.86	0.91			9.61	16.78	0.57
FF-3 factor model	37.49	54.46	0.69	86.11	32.78	2.51			-38.16	266.97	-0.14
FF-3 factor model C	6.92	16.18	0.43	13.39	16.33	0.88			9.06	16.63	0.59
FF-3 factor model (*)	20.20	37.86	0.53	44.15	44.35	0.96			-82.93	275.48	-0.30
FF-3 factor model (*) C	6.97	16.21	0.43	12.85	16.64	0.83			10.86	16.91	0.64
Ledoit	24.40	46.40	0.53	92.01	43.13	2.11			114.18	825.57	0.14
Ledoit C	6.86	16.16	0.42	13.42	16.25	0.88			9.16	16.64	0.55
Equally Weighted Portfolio	7.49	16.65	0.45	8.08	17.33	0.53			8.08	17.33	0.47

In "*" we estimate expected returns by setting the alphas to zero in the three factor model